

Dissertationes Forestales 77

Using airborne laser scanning data and digital aerial
photographs to estimate growing stock by tree species

Petteri Packalén
Faculty of Forest Sciences
University of Joensuu

Academic dissertation

To be presented, with the permission of the Faculty of Forest Sciences of the University of Joensuu, for public criticism in the auditorium BOR100 of the University of Joensuu, Yliopistonkatu 7, Joensuu, on 6th February 2009, at 12 o'clock noon.

Title of dissertation: Using airborne laser scanning data and digital aerial photographs to estimate growing stock by tree species

Author: Petteri Packalén

Dissertationes Forestales 77

Thesis Supervisor:

Professor Matti Maltamo

Faculty of Forest Sciences, University of Joensuu, Finland

Pre-examiners:

Professor Annika Kangas

Department of Forest Resource Management, University of Helsinki, Finland

Professor Barbara Koch

Department of Remote Sensing & Land information Systems, Institute for Forest Economy, University of Freiburg, Germany

Opponent:

Associate professor Terje Gobakken

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Norway

ISSN 1795-7389

ISBN 978-951-651-242-9 (PDF)

(2009)

Publishers:

Finnish Society of Forest Science

Finnish Forest Research Institute

Faculty of Agriculture and Forestry of the University of Helsinki

Faculty of Forest Sciences of the University of Joensuu

Editorial Office:

Finnish Society of Forest Science

P.O. Box 18, FI-01301 Vantaa, Finland

<http://www.metla.fi/dissertationes>

Packalén, P. 2009. Using airborne laser scanning data and digital aerial photographs to estimate growing stock by tree species. *Dissertationes Forestales 77*. 41 p. Available at: <http://www.metla.fi/dissertationes/df77.htm>

ABSTRACT

Nowadays modern remote sensing techniques are seen as very attractive alternatives for expensive field measurements in forest inventories. The most promising remote sensing technique for forest inventory purposes is currently Airborne Laser Scanning (ALS). Several studies have indicated that essential forest characteristics such as mean height, basal area and stand volume can be predicted very accurately using ALS data. However, most ALS studies have concentrated on predicting total stand characteristics although species-specific characteristics are of primary interest in Finland. The aim of the thesis is to develop and evaluate methods for species-specific stand level forest inventories using remote sensing.

The study was carried out in two test areas, both of which are located in eastern Finland and represent typical managed boreal forests in Finland. All the modelling was done at plot level and the tree species considered were Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* (L.) Karst.) and deciduous trees as a species group. The simultaneous use of ALS data and aerial photographs forms a basis of the work; the idea is that the ALS data bring information about dimensions and that aerial photographs are useful in order to discriminate between tree species.

In the first sub-study regression modelling combined with fuzzy classification and a variant of the nearest neighbour method called k-MSN were compared for determining species-specific volumes. The k-MSN method provided promising results and in the second sub-study it was extended to simultaneously predict by tree species: volume, stem number, basal area, and diameter and height of the basal area median tree; and to calculate total characteristics as sums of the species-specific estimates. The third sub-study investigated the ability of remote sensing information to predict species-specific diameter distributions. The nearest neighbour approach in which field measured trees were used as such to predict diameter distribution outperformed a theoretical diameter distribution approach in which the parameters of the Weibull distribution were predicted using the k-MSN estimates. In the fourth sub-study a new method was presented which uses the unrectified aerial photographs with known external and internal orientation instead of orthorectified images. This overcomes certain issues that are inherent to the previously used approach: the radiometric correction of aerial photographs, and the proper linkage of ALS points and aerial photographs.

The nearest neighbour method employed in this thesis turned out to be an efficient and versatile approach to estimating both total and species-specific forest characteristics and diameter distributions when using ALS data, aerial photographs and a set of sample plots as source data. The accuracies achieved in the estimation of species-specific characteristics were at least as good as those obtained by the current field inventory practise, and in the case of total characteristics the accuracy was even better.

Keywords: ALS, diameter distribution, k-MSN, non-parametric estimation, species-specific stand attributes

PREFACE

Today it is quite obvious that airborne laser scanning (ALS) will play an important role in small area forest inventories in Scandinavia. This was not so apparent when I started the work presented here. From this point of view I have been lucky; the public interest towards the topic of this thesis has grown considerably during the project. On the other hand this increasing interest has had the side-effect that due to the many other projects related approximately to the same topic that I have been involved in, the present dissertation has been postponed. However, I really cannot complain about this. At present there are already many attempts to bring ALS based forest inventory into commercial production.

ALS produces a three-dimensional georeferenced point cloud which approximates the surface of the earth. In forest inventory we are primarily interested in the quantity and dimensions of trees and therefore an essential processing step is to subtract the terrain level from the original height hits in order to scale heights to the above ground scale. By using ALS data both the terrain and canopy levels may be seen simultaneously. This is the main reason why ALS data is so well-suited for forest inventory purposes. Although laser scanning is such a hot topic at present, the usefulness of spectral images should not be forgotten. There are many applications for which ALS data is not suitable. In this thesis, for instance, aerial images were utilised together with ALS data since the use of spectral data improves the separation of tree species. Besides, it must be remembered that there is no spaceborne laser scanning instrument available at the moment and therefore large-area remote sensing cannot rely on laser scanning.

In the present thesis forest growing stock is estimated by tree species and aerial images are also used. Therefore, it differs slightly from most of the studies in which ALS data has been used to estimate forest characteristics. The estimation of growing stock by tree species is generally speaking a far more difficult task than the estimation of total characteristics. The analogy in the estimation of total characteristics is that height and density distributions of ALS data describe canopy dimensions precisely and that this information corresponds well with many forest attributes, such as with total volume. Unfortunately the same is not valid when estimating species-specific characteristics.

The Faculty of Forest Sciences at the University of Joensuu provided excellent facilities and a working environment for me. I also appreciate the financial support I received from the faculty, since I have been working there as a Senior Assistant in Forest Mensuration Science almost ever since I started the work presented here. A position as a faculty member provides the best imaginable environment to carry out research.

First and foremost, I would like to thank my supervisor, friend and co-author in all papers, Professor Matti Maltamo. We have had numerous fertile discussions about science, football and many other things. I also wish to thank the second co-author Mr. Aki Suvanto. We resolved many interesting issues while implementing the methods presented in this thesis. Mr. Pekka Savolainen provided valuable technical support regarding airborne data and their pre-processing. UPM-Kymmene Ltd funded airborne data and provided their sponsorship for collecting the ground truth material. I would also like to express my gratitude to all my colleagues, friends and parents for their support during this project. All of you know how you are related to this work. Thank you.

Joensuu, June 2008

Petteri Puttonen

LIST OF ORIGINAL ARTICLES

The present thesis is based on the following papers, which are referred to in the text by their Roman numerals:

I. Packalén, P. & Maltamo, M. 2006. Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs. *Forest Science* 52(6): 611–622. <http://www.ingentaconnect.com/content/saf/fs/2006/00000052/00000006/art00001>

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

III. Packalén, P. & Maltamo, M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Canadian Journal of Forest Research* 38: 1750–1760. doi:10.1139/X08-037

IV. Packalén, P., Suvanto, A. & Maltamo, M. 2008. A two stage method to estimate species-specific growing stock by combining ALS data and aerial photographs of known orientation parameters. Manuscript.

In papers I, II and IV the present author performed all the calculations and analysed the results. In paper III he was responsible for most of the calculations and analyses except the estimation of the Weibull-distributions. He was also the principal author of all papers.

CONTENTS

ABSTRACT	3
PREFACE.....	4
LIST OF ORIGINAL ARTICLES	5
1 INTRODUCTION.....	7
1.1 Forest inventory and remote sensing	7
1.2 Inventory requirements	8
1.3 Principle of airborne laser scanning.....	9
1.4 Stand level forest inventory by airborne data	11
1.4.1 Inventory unit.....	11
1.4.2 Aerial photographs.....	12
1.4.3 Airborne laser scanning	13
1.4.4 Species-specific estimation by ALS data	14
1.5 Objectives	15
2 MATERIALS.....	16
2.1 Study areas and field data	16
2.2 Remote sensing data	17
3 METHODS	18
3.1 Radiometric calibration of aerial photographs	18
3.2 Predictor variables	18
3.2.1 ALS data	18
3.2.2 Ortho-rectified images	18
3.2.3 Point proportions.....	19
3.3 Fuzzy classification.....	20
3.4 The k-MSN method	20
3.5 Diameter distribution based on k-MSN imputed trees.....	21
3.6 Diameter distribution based on the Weibull function	22
3.7 Estimation of growing stock by tree species.....	23
3.8 Accuracy assessment	24
4 RESULTS	25
4.1 Volume estimates by tree species	25
4.2 k-MSN estimates of species-specific stand attributes.....	26
4.3 Diameter distributions.....	28
4.4 Rigorous data fusion method	29
5 DISCUSSION	30
REFERENCES	35

1 INTRODUCTION

1.1 Forest inventory and remote sensing

Nowadays modern remote sensing techniques are seen as very attractive alternatives for expensive field measurements in forest inventories. The most promising remote sensing technique for forest inventory purposes is currently Airborne Laser Scanning (ALS). However, the first attempts to exploit remote sensing data in forestry were carried out already a long time ago. Early studies and experiments with the use of aerial photographs were conducted both in Europe and in North America in the 1920s (Sarvas 1938; Myles 1945; Loetsch & Haller 1973). Since then both monoscopic (2D) and stereoscopic (3D) techniques for interpreting aerial images have been widely used in forest inventories and mapping. The first preliminary studies of the use of aerial photographs in Finnish forestry were carried out in the 1920s (Fabritius 1922) and by the 1930s aerial photographs were already being used for forest mapping and the estimation of growing stock volumes (Nyyssönen 1959). In Finnish forestry, however, analytic stereoscopic measurements have never been a common practice, but aerial photographs have been used for decades in practical forestry for stand delineation.

The rapid developments in space technology during the cold war in the 1960s led to the launching of the first non-military earth observation satellite, Landsat MSS, in 1972 (Aronoff 2005). Since then passive satellite data has been used for the large-scale mapping of forests (e.g. deSteiguer 1978; Kilkki & Päivinen 1987; Wulder et al. 2003; Reese et al. 2003). In the Finnish National Forest Inventory, for instance, satellite data is used to increase the efficiency of the inventory by producing wall-to-wall estimates of forest characteristics from which forest resource information may be derived for smaller areas (Tomppo 2006). On the other hand, satellite imagery has not been successfully used in stand level forest inventories. Very high resolution (VHR) satellite imageries have become available for civil use in the past decade. The spatial resolution of VHR data is already approaching that of aerial photographs, so that they may be used to some extent as alternatives for airborne images.

Radio detection and ranging (RADAR) is an active remote sensing technique which has been used in many studies to estimate forest variables. RADAR instruments are usually carried by satellite platforms and their spatial resolution is generally comparable to that of passive satellite imagery. Synthetic aperture radar (SAR) is the most common remote sensing technique used for radar imaging (Hoekman 1985; Rauste et al. 1994; Fransson 2001; Magnusson 2006), with approximately the same level of accuracy in estimating forest volumes as can be achieved with passive satellite data, or even slightly poorer. One reason for the low level of accuracy may be that, with some exceptions, current SAR systems were not designed for forestry purposes and therefore do not meet the needs of forest inventories or mapping (Dobson 2000). Recent developments in satellite SAR imaging are enhanced spatial resolution and diverse polarization modes (Morena et al. 2004).

Airborne Laser Scanning system is a type of RADAR that differs considerably in response from the conventional RS data sources. The method provides a means to measure the distance between the aircraft and the earth's surface. The ALS gives a three-dimensional (3D) georeferenced point cloud which approximates the surface of the earth, i.e. the physical dimensions are measured directly. Several studies have indicated that essential forest characteristics such as mean height, basal area and stand volume can be

predicted very accurately using ALS data (Magnussen & Boudewyn 1998; Næsset et al. 2004; Jensen et al. 2006). Regarding total characteristics, such as the total volume of all tree species, the results have been even better than those achieved with traditional stand level field inventory techniques. However, there is a lack of research in which species-specific forest variables are estimated using ALS data alone, or in combination with some other data source. Because the costs of ALS data have decreased considerably due to technological development, ALS based forest inventory applications have emerged as an interesting and realistic alternative for operative forestry purposes. The promising results achieved with ALS in estimating growing stock were a significant source of motivation and the starting point for the present work.

1.2 Inventory requirements

The aim of a forest inventory is to provide unbiased information for decision-making purposes. Without an accurate estimate of the current state of the forest resources it is impossible to make proper and rational decisions. To meet the user's requirements several types of forest inventories are needed. However, regardless of the type of forest inventory, the aim is to estimate means and totals for measures of forest characteristics over a defined area (Kangas et al. 2006). The characteristics of interest vary and may be, for example, the volume of growing stock, the area of forest damage, the biomass, the carbon content, or the quantity of coarse woody debris. The primary interest may also be the change over time instead of the current state (Duncan & Kalton 1987). Forest inventories are carried out on different scales applying various methods using both field surveys and remote sensing. Global, country-wide and regional inventories are performed in order to support political decision making. Currently, global or country-wide inventories provide information that may aid in decision-making when attempting to mitigate the effects of climate change and global warming, for instance. This thesis presents methods for stand level forest inventories, focussing particularly on the requirements in Finland and therefore only such inventories are discussed here.

A stand is a basic treatment unit in silviculture. It is considered to be a uniform forest area with respect to its growing stock, species composition and site conditions. Stands are usually delineated before the field inventory using aerial photographs, and the stand borders are corrected during the fieldwork as required. Typically, colour-infrared (CIR) photographs are used for delineation; this enables the easy interpretation of deciduous trees, in particular. Currently stand characteristics by tree species are assessed in the field using subjectively placed angle-count sample plots, i.e. large trees are more likely to be sampled. The forest stock characteristics to be estimated are mean age, basal area or stem number and the diameter and height of the basal area median tree. All other stand attributes are estimated from these measurements by employing theoretical diameter distributions and tree level height and volume models. The need for using diameter distributions derives from the tradition of using tree level growth and yield models. Diameter distributions are constructed by tree species and usually scaled according to the basal area. Several diameter distribution models have been used, including the probability functions of beta, Weibull and Johnson SB, and various percentile models (Kilikki et al. 1989; Siipilehto 1999; Kangas & Maltamo 2000b).

Haara and Korhonen (2004) evaluated the accuracy of the current inventory system by means of comprehensive reference measurements. The standard error of the growing stock

volume at the stand level was 24.8%, and the standard errors for the species-specific volumes were 29.3%, 43.0% and 65.0% for pine, spruce and deciduous trees, respectively. Haara and Korhonen (2004) also concluded that due to the subjectivity of the method, the accuracy of the stand attributes is highly dependent on the skills of the surveyor. The same type of variation among surveyors was also reported by Kangas et al. (2004).

The current inventory practice sets the minimum level of accuracy that this thesis aims at. It is also obvious that diameter distributions by tree species are required because in Finland tree-level models are applied almost without exception. Basically, regardless of the inventory method, diameter distributions may be constructed by estimating the same variables as are now measured in the field and then utilizing diameter distribution modelling approach. However, it may be reasonable to unlearn conventional diameter distribution modelling when applying remote sensing methods to forest inventories.

1.3 Principle of airborne laser scanning

The terms around ALS are often obscure and misused among forest specialist. The term LIDAR is an acronym for “Light Detection and Ranging”. Thus it emphasizes the capability of carrying out range measurements. LASER is an acronym for “Light Amplification by Stimulated Emission of Radiation”, whereas the term “laser scanning” refers explicitly to a laser system with scanning capability. All these terms are often used interchangeably although they usually refer to a LIDAR installation carried by an airborne vehicle (ALS). It must be noted that there are other types of LIDAR as well, for instance, space-borne instruments and those that are operated on the ground. ALS has gained a great deal of attention during the last 5-10 years especially in the fields of terrain modelling, urban area modelling and forestry. The present overview of ALS technology has been compiled from many published and unpublished sources. Therefore no references are listed. An introduction to the in-depth theory of ALS is given by Wehr & Lohr (1999).

LIDAR is an optical remote sensing technology which measures the properties of scattered light radiating from of a distant target. LIDAR systems emit highly directional pulses of laser light and measure precisely the time required for a reflection to return from the ground below. As the speed of light is known, approximately 0.3 metres per nanosecond, the distance from the target can be calculated. In addition, the scanner viewing angle at the time each pulse reflection was received and the sensor position and orientation must be available so that the 3D position of the target can be defined. The sensor position is provided by a GPS/IMU system (Global Positioning System and Inertial Measurement Unit) and the viewing angle is recorded by the sensor itself. Thus, in principle, the positions of the laser hits are determined *in-situ*, although in practice the final positions are produced during post-processing, when certain sources of bias are removed.

Although nowadays practically all sensors are capable of recording multiple reflections per pulse, a laser pulse cannot penetrate solid surfaces, due to the wavelength used. Therefore, not every pulse generates multiple returns; only those pulses which encounter features which at least partly allow the light to penetrate them, such as the forest canopy, generate multiple reflections. In considerably simplified terms, the first return is generated when a pulse encounters the canopy and the last return when it encounters the bare earth. In practice there are many pulses which provide only one reflection, so that the first return can be considered equal to the last return. Modern sensors can record at least three reflections but often only the first and last one are eventually utilized. In forestry applications the

accurate determination of the ground level is essential because the primary interest is usually in the scale of the pulse heights relative to the ground level. The accurate results of estimating forest characteristics by ALS data are, indeed, a consequence of the possibility to estimate both ground and canopy level from the data. Conceptually, the first step in forest inventory applications is almost always to create a terrain model and to subtract it from the (orthometric) pulse heights in order to scale heights above ground level.

Some small footprint LIDARs are also capable of digitizing a full waveform, i.e. reflections as a function of height. Full waveform digitization is still rather uncommon, but it is expected that there will be a great deal of methodological development around this topic in the near future. When LIDAR data is being collected, sensors also capture the light intensity of each pulse that is returned, and this intensity can be used to generate orthophoto-like images, although these are not widely used because of the substantial fluctuation in intensity values.

Airborne LIDAR instruments are usually carried by conventional aircraft, which are preferred mostly for economic reasons, but helicopters may be used as well. As the aircraft moves forward the LIDAR system transmits laser pulses across the flight direction, resulting in a Z-shaped pattern of points on the ground. The x spacing of the laser hits on the ground depends on the viewing angle, pulse repetition rate (i.e. how many pulses are generated per second) and flight altitude. Correspondingly, the y spacing is a function of the flight speed, pulse repetition rate and flight altitude. The flight altitude and viewing angle also determine the swath width on the ground. By capturing numerous adjacent flight lines, usually with a 20-40% overlap, full area coverage is achieved. A side view of an ALS point cloud for a mature spruce stand is presented in Figure 1.

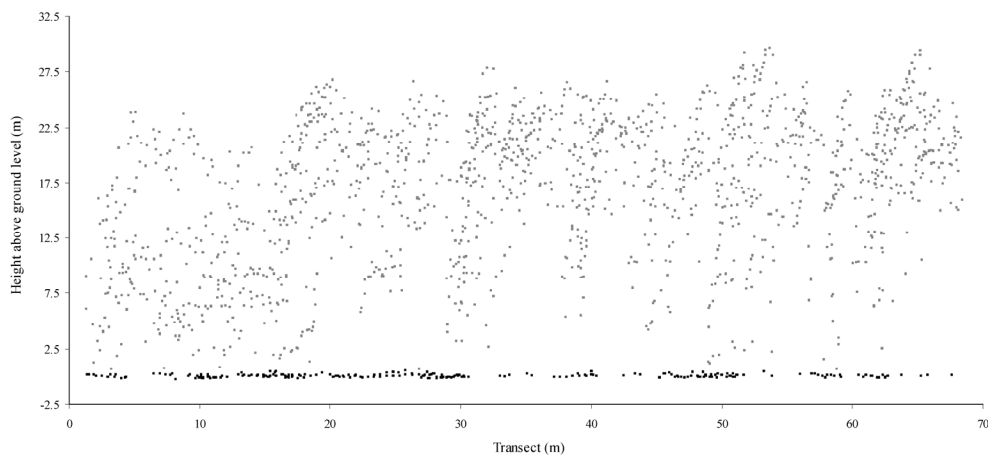


Figure 1. Side view of a 5 metre deep transect of ALS data for a mature spruce stand in the Matalansalo test area acquired with an Optech ALTM 2033 sensor at an altitude of 1500 metres above ground level. Only the first pulses are depicted in the figure. The pulse density is about 0.7 measurements per square metre. The black dots represent pulses classified as ground hits.

Correspondingly, the beam divergence and flight altitude determine the laser beam diameter on the ground, or footprint diameter. The ALS installations discussed here can all be considered small-footprint LIDARs. Large-footprint (e.g. 10-25 m) LIDARs are operated from space and employ slightly different technology, so that they will be ignored here. In small footprint LIDARs the footprint diameter is usually 0.1-1 m. Most commercial ALS systems provide a vertical accuracy of about ± 5 -50 cm depending on the surface characteristics and sub-metre horizontal accuracy. The typical flight altitude is 200-3000 metres above ground level, and the pulse density on the ground is usually something like 0.1-10 measurements per square metre. Most commercial ALS systems currently use a near-infrared wavelength (e.g. 1064 nm) and the pulse repetition rates are in the range 30–200 kHz. What is common to all state-of-the-art technologies, however, is that developments are rapid, and therefore the above parameters and sensor characteristics may no longer be valid even in the near future.

1.4 Stand level forest inventory by airborne data

1.4.1 Inventory unit

In terms of inventory unit, a very similar categorization of inventory approaches can be used with both ALS data and aerial photographs. The unit may be an operative stand, a “microstand”, a plot or a tree. A microstand is a homogeneous patch of an area that is usually smaller than an operative forest stand (Hyvönen et al. 2005). Microstands are always created using image segmentation techniques, and thus are also referred to as (forest) segments. The major difference between microstands and operative stands is that microstands are not designated to be used as a basic treatment unit in silviculture; they are used for inventory purposes only or may be used as stands marked for cutting, for instance.

Automatic tree level inventory methods rely on the delineation of individual tree crowns to segments. Both 2D and 3D techniques have been introduced for this purpose. The idea behind individual tree detection is that their crown dimensions can be correlated with tree characteristics of primary interest such as diameter at breast height and tree height, after which generic models that use diameter and height as explanatory variables may be used to estimate the stem volume. In 2D methods only the crown diameter or area is used to estimate other tree characteristics, while in 3D methods tree height can also be determined. Because of the direct determination of height, 3D methods produce much more accurate estimates for stem characteristics of interest. Tree level inventory approaches produce ideal input data for forest management systems which operate at the single tree level, such as the systems currently used in Finland. One disadvantage, however, is that suppressed trees which are not visible from above cannot be detected and the individual tree interpretation is also generally difficult in a dense forest.

The basis of the methods applied at stand, plot or microstand level is to estimate the required forest attributes for a group of trees directly at that level. Even though all methods entail a natural inventory unit, the applicability of the method is not tied to any particular level. Individual trees, for instance, can be used to compose a plot, or plots can be aggregated to form stands, e.g. using a grid approach (Næsset 2004a). Due to averaging, up-scaling generally increases the accuracy at the aggregated level.

1.4.2 Aerial photographs

Previous studies have presented multiple ways to utilize aerial photographs in forest inventories. These can be categorized to techniques that use visual (or manual) interpretation and those that use computer-aided interpretation. In the era of analogue photographs it was naturally only visual interpretation that was used. However, in the era of digital images visual interpretation is still frequently used, for instance, in the delineation of forest stands. Visual interpretation has played an especially important role in two-phase sampling. The principle is that aerial photographs are used for stratification in the first phase and in the second phase a subset of plots or stands are measured in the field and the results are generalized to the whole area (Poso 1972; Mattila 1985; Spencer & Hall 1988). During the past decades numerous other visual interpretation methods for forest inventory purposes have also been studied at the tree and stand levels using both 2D and 3D techniques (Nyyssönen 1955; Aldrich et al. 1969; Talts 1977; Needham & Smith 1987; Næsset 1991; Gagnon et al. 1993; Anttila 1998). According to Anttila (2005), Nordic research into visual interpretation has mostly been aimed at stand level inventories. In general, visual interpretation is a time-consuming task, which often leads to high costs. Its disadvantage is also its subjectivity, which means that the outcome and its accuracy vary from surveyor to surveyor. The focus has therefore switched towards computer-aided interpretation.

Forest inventory methods which rely on the use of aerial photography may also be divided into 2D and 3D methods. In 2D methods the radiance observed by the camera and the texture it creates onto the image plane are used as predictors. The radiance observed by the camera system is usually referred to as the digital number (DN) or just pixel value, because in practice nowadays only digital images are used and the radiance cannot be converted into reflectance in real world applications. The terms spectral feature and tone are generally used, too. Texture contains information about the spatial distribution of tonal variations within an image. There are several methods which have been used to describe texture in numerical images, the most common of which is probably the principle of grey-tone spatial dependence matrix presented by Haralick et al. (1973). The reason for the use of texture is that it makes sense to utilize all the meaningful information available in the images, and spectral features do not take into account spatial variation at all. The combined use of spectral and textural features has often been found to improve the accuracy by comparison with the use of tone alone (see Anttila 2005 p. 11).

In 3D methods photogrammetry is employed. The fundamental task of photogrammetry is to rigorously establish the geometric relationship (orientations) between the image and the object (Mikhail et al. 2001). Once this relationship is correctly recovered, one can then derive information about the object. In forest inventory applications the aim of photogrammetry is the extraction of tree or stand dimensions. Therefore 3D photogrammetric methods can in a sense be compared with ALS based inventory methods.

A common tree level 2D approach consists of first detecting the tree tops by assuming that the local intensity maxima on the aerial images will represent tree tops, after which the crowns are segmented by means of a region-growing algorithm (Dralle & Rudemo 1996; Lowell 1998; Wulder et al. 2000; Pitkänen 2001; Anttila & Lehtikoinen 2002). Another approach to detecting individual trees is to use predefined instances of sub-images, called templates. This type of technique has been developed and tested by many authors (Pollock 1996, 1999; Larsen 1997; Larsen & Rudemo 1998; Olofsson et al. 2006). Korpela (2000, 2004) extended the template matching approach to a 3D application, which enables tree

height determination as well. Other approaches to the detection of individual tree crowns are valley following (Gougeon 1995), multiple-scale analysis (Brandtberg & Walter 1998) and edge detection (Pinz 1999).

Automatic forest inventory methods which exploit aerial photographs at the stand, microstand or plot level mostly employ spectral features and texture. To mention a few recent studies, approaches of this kind have been employed in Hyypä et al. (2000), Shugart et al. (2000), Holmström et al. (2001), and Hyvönen et al. (2005). However, 3D methods have also been adopted at the stand level, such as the one by Korpela & Anttila (2004).

The drawback of aerial photography is that the radiance observed by the camera is affected by several forms of spectral distortion, such as light fall-off effects and variations in atmosphere and imaging geometry (King 1991; Pellikka 1998; Lillesand et al. 2004). In addition, the camera system, film and post-processing can cause undesirable variation in the pixel values. The most significant source of spectral distortion, at least where forests are concerned, is the bi-directional reflectance effect. This can be perceived in aerial photographs as variation in brightness caused by the fact that the tree crowns in the direction of incoming radiation expose their shady sides to the camera and those in the opposite direction their illuminated sides (Holopainen & Wang 1998). Due to spectral distortion, the same object is not represented with congruent pixel values in different parts of the image or in different images. When spectral features are used in the computer-aided interpretation of forest characteristics it is essential to minimize these anomalies. Several empirical methods in particular and also some semi-empirical ones have been developed and tested for this purpose.

Aerial photography is currently in a transitional phase from analog to digital imaging. Although digital images have almost exclusively been used for some time, usually images have been captured with an analogue frame camera and digitized afterwards by scanning the film. Currently digital aerial cameras are replacing traditional analogue frame cameras and only then is an error prone analog-digital transformation avoided completely. Digital cameras bring also some new characteristics to the aerial imaging, such as capturing 4 multispectral bands simultaneously, separate panchromatic band(s) and an image fusion technique called pan-sharpening. GPS/IMU systems, which provide *in-situ* positioning of the camera and its rotations, are also becoming increasingly common.

The accuracy of forest inventory techniques based on aerial photography has in general been poorer than that achieved with traditional field inventories and only part of the necessary information can be collected in this way. In addition, the author has observed that the promising results achieved with ALS data have decreased considerably the interest in developing inventory methods based solely on aerial photographs.

1.4.3 Airborne laser scanning

The two main approaches for predicting growing stock characteristics using ALS are the laser canopy height distribution approach (also referred to as the “area based statistical approach”), usually used with low-resolution data (Næsset 1997, 2002; Lim et al. 2003; Maltamo et al. 2006b; Jensen et al. 2006; van Aardt et al. 2006), and the individual tree delineation approach (3D), used with high-resolution data (Hyypä & Inkinen 1999; Persson et al. 2002; Popescu et al. 2003; Maltamo et al. 2004; Solberg et al. 2006; Peuhkurinen et al. 2007). Low resolution means in this context that the pulse density at ground level is about 1 per square metre and high resolution about 5-10 per square metre.

The major difference between the laser canopy height distribution and the individual tree delineation approach is that the latter relies on the recognition of individual trees and allometric relationships on the tree level, whereas the former uses height hits directly on the plot, microstand or stand level to estimate growing stock characteristics. A common method in individual tree delineation is to detect trees from an interpolated canopy height model by locating local maxima of the height values. After that trees are segmented around local maxima, for instance by some region growing algorithm. In the canopy height distribution approach regression modelling is the most often used estimation technique, although other methods, such as the non-parametric approach used in this thesis, have also been utilized. Most actual forestry applications have so far been based on the canopy height distribution approach.

ALS data have been used to construct theoretical diameter distribution models in Norway and Finland (Gobakken & Næsset, 2004, 2005; Maltamo et al. 2006a; Bollandsås & Næsset 2007; Maltamo et al. 2007). These studies have employed the parameter prediction method, the Weibull function and percentile-based distributions. All of these studies considered basal area diameter distributions at stand level. In addition, Maltamo et al. (2007) showed that unlike conventional angle count measurements, ALS data do not require the use of basal area distributions: diameter distributions can be formed directly in terms of tree frequencies without a loss of accuracy in volume estimates. Maltamo et al. (2007) also tested the usability of calibration estimation for adjusting the predicted distributions to be compatible with the ALS based estimated stand volume. Mehtätalo et al. (2007) presented an ALS data based parameter recovery method, where values for the parameters of assumed diameter distribution and height-diameter models are determined simultaneously. An advantage of this approach is that if a solution for the formulated system of equations exists, it is always compatible with the predictions of stand characteristics.

Regarding total characteristics, such as the total volume, the accuracies obtained by ALS have been even better than those achieved with traditional stand level field inventory techniques.

1.4.4 Species-specific estimation by ALS data

Most ALS studies have concentrated on predicting total stand characteristics, thus tree species are not taken into account. There are many reasons why tree species are often ignored. One reason may be that either the study has no link to practical requirements or, as in many cases, there is no need for species-specific inventory data. In intensively managed tree plantations, for instance, it is practically always known what the tree species of a stand is. It may also be the case that tree species has already been taken into account by defining the main tree species at the stand level as a separate prestratification stage using aerial images, as is done for example in Norway (Næsset 2004b). A third reason for not taking tree species into account may also be that estimating growing stock is considered to be too difficult a task. This is certainly true in areas where, for instance, 5-10 tree species exist. In Finland, however, there are only two important coniferous tree species, Scots Pine and Norway spruce, and deciduous trees are in the minority, representing less than 20% of the growing stock (Korhonen et al. 2006). This makes it is feasible to estimate growing stock by tree species in Finland.

The majority of species-specific ALS studies have been carried out on the individual tree level, where the unit of classification is obvious, a single tree (e.g. Persson et al. 2003;

Moëffiet et al. 2005; Ørka et al. 2007; Liang et al. 2007). Some studies have employed only ALS data (e.g. Holmgren & Persson 2004; Brandtberg 2007) and some have combined ALS and spectral data (e.g. Koukoulas & Blackburn 2005; Holmgren et al. 2008). Note that in the above mentioned studies tree species classification was carried out but final species-specific end products of forest inventory were not produced. Korpela et al. (2007) seems to be so far the only publication in which an entire individual tree based inventory system that also takes into account tree species was tested and accuracies were reported. In that particular case an experienced photo-interpreter carried out the visual interpretation of tree species by using aerial photographs.

In the case of a plot or stand level inventory individual trees cannot be classified; instead species-specific forest attributes must be estimated directly for a group of trees. It is apparent that at least at the plot or stand level ALS data alone cannot discriminate tree species very well. However, since it contains structural information about the canopy, ALS data differs considerably in response from the conventional spectral images used in remote sensing. Because this structural information contains at least some information about tree species and it differs considerably from the spectral response, it makes sense to combine ALS data with aerial photographs when estimating species-specific forest characteristics. Aerial photographs have broadly been used to discriminate tree species in visual interpretation. However, it is not so obvious that aerial photographs may be used to discriminate tree species accurately in computer-aided image interpretation. Several spectral distortions, changing lighting conditions and tree species which create similar kinds of spectral and textural fingerprints aggravate the interpretation considerably. On the other hand, the virtue of using aerial photographs in a forest inventory is that they usually do not introduce extra costs because they are already needed for stand delineation.

So far there have been very few experiments with ALS data in which the forest growing stock has been estimated by tree species. Nor have there been ALS data based diameter distribution studies where tree species would have been considered.

1.5 Objectives

The overall aim of the thesis is to develop and evaluate methods for species-specific stand level forest inventories using a combination of ALS data and aerial photographs. A specific point of interest is to produce those tree characteristics that are required in Finland, i.e. ultimately diameter distributions by tree species. The specific objectives for papers I-IV are:

- I. To evaluate two different approaches and their accuracy in predicting species-specific volumes and total volume at the plot level. The approaches tested were regression modelling combined with fuzzy classification and k-MSN in a multivariate manner.
- II. To develop and test the ability of the k-MSN method to predict simultaneously numerous species-specific stand attributes and total characteristics as their sums.
- III. To develop methods and test the ability of remote sensing information to predict species-specific diameter distributions. The k-MSN method is extended for the estimation of diameter distributions.
- IV. To further develop the previously presented method (II-III), in particular, the aim was to improve the overall accuracy of tree species-specific estimation and the

applicability of the method in operational use by employing a rigorous data fusion method.

2 MATERIALS

2.1 Study areas and field data

Two study areas were used in this thesis, henceforth referred to as Matalansalo and Juuka. The Matalansalo study area was used in papers I-III, whereas the Juuka area was used in study IV. Both areas represent typical Finnish managed boreal forests and are located in eastern Finland. The Matalansalo area covers approximately 2 000 and Juuka about 12 000 hectares. The study areas are dominated by coniferous tree species, namely Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.). Deciduous trees, mainly the downy birch (*Betula pubescens* Ehrh.) and silver birch (*Betula pendula* Roth.), are usually in the minority in the tree stock. All the deciduous species were lumped together in the present work, as it was already known beforehand that they are virtually impossible to separate by remote sensing methods under Finnish conditions, where the forests are dominated by coniferous species and the vast majority of the deciduous trees are birches. Basically the same materials were used in papers I-III and therefore described only once.

A network of circular sample plots with a radius of 9 metres was established on both areas. The Global Positioning System with differential correction was used to determine the position of the centre of each plot to an accuracy of about 1 meter. Sample plots were placed in the young, middle-aged and mature forests. Seedling and sapling stand development classes were completely left out from this study. The diameter at breast height (dbh), tree and storey class and tree species were measured for each tree with a dbh greater than 5 cm. Height was measured for one sample tree of each species and storey class by plots, and this information was used for calibration of the tree species-specific height models. In Matalansalo a H-D curve by Veltheim (1987) was employed, whereas in Juuka a local H-D model was fitted using a function form by Näslund (1937). The volumes of individual trees were calculated using the species-specific models presented by Laasasenaho (1982). Finally, basal area (G), number of stems (N), volume (V) and diameter (dgm) and height (hgm) of the basal area median tree were calculated for each plot and stand and multiplied out to the hectare level.

In Matalansalo a network of 463 sample plots distributed over 67 forest stands was measured in the summer of 2004. Scots pine is the main tree species on 59% of the plots, Norway spruce on 34% and deciduous trees on 7%. At least two tree species occurred in 92% of the plots. In Juuka a field data of 493 plots were acquired during the consecutive summers of 2005 and 2006. The dominant tree species represented over 90% of the volume in 48% of the plots, 75–90% in 25%, 50–75% in 23% and less than 50% in 3%. Only one tree species occurred in 12% of the plots, two in 34% and three in 54%. Due to ample species mixture both data sets can be considered to represent mixed boreal forests realistically and therefore suitable for studies where species-specific estimation is in focus. In Matalansalo, the plots were placed such a manner that that there were always numerous plots within a stand, i.e. an average of seven plots per stand. This enabled the calculation of the stand level results by taking the mean value for all the plots within one stand. In Juuka, there is generally only one plot for each stand.

2.2 Remote sensing data

In Matalansalo the ALS data was collected at night on August 4, 2004 using an Optech ALTM 2033 laser scanning system operating at an altitude of 1500 m above ground level using a half-angle of 15 degrees. This resulted in a swath width of 800 metres and a nominal sampling density of about 0.7 measurements per square metre. The divergence of the laser beam (1064 nm) was set at 0.3 mrad, which produced a footprint of 45 cm at ground level. Altogether seven laser strips were measured with an overlap of 35 percent, and both first and last pulse data were recorded.

In Juuka the ALS data was collected on July 13, 2005 using an Optech ALTM 3100C laser scanning system. The area was measured from an altitude of 2000 m above ground level using a field of view of 30 degrees and a side overlap of about 20%. This resulted in a swath width of approximately 1050 m and a nominal sampling density of about 0.6 measurements per square meter. Altogether 7 laser strips were measured, one of which was perpendicular to the other strips and used only for calibration purposes. The divergence of the laser beam (1064 nm) was set at 0.3 mrad, which produced a footprint of 60 cm at ground level. The Optech ALTM 3100C laser scanner captures 4 range measurements for each pulse, but here the measurements were reclassified to represent first and last pulse echoes.

A digital terrain model (DTM) was generated from the ALS data in both study areas. First laser points were classified to ground points and other points (the method is explained by Axelsson, 2000) and then a DTM raster with a cell size of 2.5 meters was created by computing the mean of the ground points within each raster cell. Values for raster cells with no data were derived using Delaunay triangulation and bilinear interpolation.

In Matalansalo the aerial photographs were captured on a colour-infrared film to a scale of 1:30 000. The film used recorded the green, red and near-infrared portions of the spectrum. The photographs were taken with a Leica RC30 camera using a UAGA-F 13158 objective with a focal length of 163.18 mm and an anti-vignetting filter (AV525 nm). The films were digitized at a resolution of 14 μm with a Leica DSW600 film scanner and then ortho-rectified and resampled to a pixel size of 0.5 metres. The DTM generated from the ALS data was used in the ortho-rectification. The test area was covered by means of three aerial photographs, all acquired on July 22, 2004. The Landsat 7 ETM scene used for radiometric correction of the aerial photographs in Matalansalo had been captured on June 16, 2002. Since it was known beforehand that there had been no logging in the test area in 2002-2004, the two-year time difference between the acquisition dates of the aerial photographs and the Landsat scene is insignificant.

In Juuka the aerial photographs were taken with a Vexcel UltraCamD digital aerial camera on September 1, 2005. The images were taken at an altitude of 3000 meters above ground level with a sidelap of 65 % and endlap of 80 %. The study area was covered by 260 images. In this thesis only color (red, green, blue) and NIR bands were used, thus pan-sharpening was ignored; Vexcel refers to this processing level as Level-2. Ortho-rectification was also discarded; unrectified images of known internal and external orientation were used instead. The DTM generated from the ALS data was used in the aerial triangulation as the source of height control points in calculating a Z-shift for projection centers and thus adjusting the orientations in height to match the ALS data.

3 METHODS

3.1 Radiometric calibration of aerial photographs

The aerial photographs were corrected radiometrically using the method presented by Tuominen and Pekkarinen (2004) in papers I-III. The idea is to correct anomalies caused by the bi-directional reflectance effect using a set of local adjustment and image material that is less prone to this effect. The bi-directional reflectance-affected intensities of the aerial photographs were adjusted to the local intensities of a reference image, in this case the Landsat 7 ETM, as its BRDF effect is insignificant compared with that of aerial photographs. The local neighbourhoods were defined by means of a moving circle with a radius of 40 metres. The mean value of the neighbourhood on the Landsat ETM image was divided by the mean value of the aerial photograph in the corresponding neighbourhood and further multiplied by the original pixel value in the aerial photograph. The correction was made band-by-band using the approximately equivalent Landsat ETM channel to correct the corresponding aerial photograph channel. The method does not change the shape of the histogram within a moving circle; it only transfers its location.

It was perceived in the preliminary tests that the calibration method used here improves the correlation between the image intensities and species-specific stand attributes. The major benefit of this method is that it does not require any a priori information, hence it is always applicable when a valid reference satellite image is available.

3.2 Predictor variables

3.2.1 ALS data

First the ALS data was converted to above-ground scale (canopy heights) by subtracting the DTM from the orthometric heights. The ground hits were excluded by assuming that any point with a canopy height of less than 2 metres is a ground hit and that the remaining points are canopy hits. After that the first and last pulse height distributions for each sample plot were created from the canopy height hits. Canopy height percentiles from 5 to 100% (h_5, \dots, h_{100}) were computed, and the corresponding proportional canopy densities (p_5, \dots, p_{100}) were also calculated for each of these quantiles. The actual percentile points used in the modelling varied by study. In addition, the proportions of canopy hits vs. ground hits and the mean and the standard deviation of the canopy height hits were determined. All metrics were calculated separately for the first and last pulse data.

3.2.2 Ortho-rectified images

The spectral and textural features calculated from the radiometrically corrected images were used, together with the ALS data, as predictor variables in the modelling of forest attributes in papers I-III. The idea of using aerial photography was to achieve tree species discrimination. The spectral values used in the subsequent modelling were calculated as averages of all pixels within a plot by bands. Apart from average values, median values were used in a corresponding manner in papers II and III.

The textural features were calculated from the grey-tone spatial-dependence matrix on the principle presented by Haralick et al. (1973). These textural features have been widely

used in remote sensing during recent decades (Weszka et al. 1976; Marceau et al. 1990; Muinonen et al. 2001; Coburn & Roberts 2004). There are a large number of parameters to consider when selecting how to calculate these features, and at least the following need to be considered: number of requantification classes, pixel size, window size, direction, lag, channel or channel transformations and the texture measure itself (Franklin et al. 2000; Tuominen & Pekkarinen 2005). There is no unambiguous way to resolve which combination of textural features and parameters should be used. However, it would be impractical to test all the alternatives. Therefore some criteria must be used to reduce the parameter space. Here a preliminary selection of appropriate features and parameters was made based on correlation analysis, discriminant analysis and observations made in previous studies. The final features were calculated using 16 requantification classes and a lag distance of 2.5 metres. Instead of employing the often used moving window technique to calculate textures around the local neighbourhood of each pixel, all the textural features were extracted directly from the circular sample plot.

3.2.3 Point proportions

The proportions of ALS points by tree species were used as predictor variables in paper IV. The first step was to classify each ALS point as a pine, spruce or deciduous hit. This was done by linking spectral values of aerial photographs to ALS points and using these as patterns to classify each point. Due to the problems in linking a known 3D location to an orthorectified image, unrectified images of known internal and external orientations were used instead. Because adjacent images overlap in aerial photography each ALS point may be assigned to many images, i.e. for each ALS point spectral values can be fetched from many aerial photographs. In the remainder of the text p denotes the ALS point which is the first echo and the height of which is greater than the plot level canopy height of the 10th percentile.

Let A denote the whole set of aerial photographs. The set of aerial photographs from which the DN values are fetched for the ALS point p is defined as follows:

$$I_p = \{a \in A | f(a, p)\}, \quad (1)$$

where $f(a, p)$ tests whether the point p lies within an image a . The location of the ALS point in the image plane was calculated by the well known collinearity equations:

$$\begin{aligned} x &= -f \frac{m_{11}(X_p - X_L) + m_{12}(Y_p - Y_L) + m_{13}(Z_p - Z_L)}{m_{31}(X_p - X_L) + m_{32}(Y_p - Y_L) + m_{33}(Z_p - Z_L)} \\ y &= -f \frac{m_{21}(X_p - X_L) + m_{22}(Y_p - Y_L) + m_{23}(Z_p - Z_L)}{m_{31}(X_p - X_L) + m_{32}(Y_p - Y_L) + m_{33}(Z_p - Z_L)}, \end{aligned} \quad (2)$$

where x, y are image coordinates relative to a principal point, f is the focal length, X_p, Y_p, Z_p are ground coordinates of the ALS point, X_L, Y_L, Z_L are ground coordinates of the exposure station, and m_{11}, \dots, m_{33} are coefficients of the rotation matrix defined by the rotation angles ω, φ and κ (Mikhail et al. 2001). Subsequently, internal orientation was used to calculate location in image coordinates and DN values were fetched from the pixel on which the ALS point lies. The mean of the DN values were calculated for each ALS point as follows:

$$p_{mean,b} = \frac{1}{n(I_p)} \sum_{a \in I_p} g(p,b,a), \quad (3)$$

where $g(p,b,a)$ denotes a function that returns the DN value of band $b=\{nir,red,green,blue\}$ from the image a corresponding to the location of p .

A subset of 35 sample plots from Juuka was used as training data in supervised classification. Tree coordinates were not recorded in the field and therefore the training data consisted of only pure one species plots, i.e. all ALS hits within a sample plot represented the same tree species. Since for spruce and deciduous trees there were not enough completely pure one species sample plots, a slight species mixture had to be accepted. Linear Discriminant Analysis (LDA) was used to create a discrimination model using the training data (Venables and Ripley 2002). Mean values $p_{mean,b}$ by bands were used as predictors in this model. Then each ALS point p – not only training data – was classified as a pine, spruce or deciduous hit using this discrimination model. Finally the proportions of ALS points by tree species were calculated for each plot.

3.3 Fuzzy classification

A pattern is a vector of features describing an object. The aim is to create a relationship between a pattern and a class of objects. The relationship between an object and its class may be of a one-to-one kind, producing a hard classification, or of a one-to-many kind, producing a fuzzy classification. A fuzzy classification does not force each object to belong to a single class. Instead, it allows multiple and partial class memberships among the candidate classes (Wang 1990). A membership value between 0.0 and 1.0 is assigned for every class with respect to each object, so that these values generally sum to 1.0 across all the candidate classes. From a probabilistic viewpoint the membership value can be seen as the probability of the object belonging to a particular class (Foody et al. 2003). Note that the term fuzzy classification is used here as a generic term and does not refer exclusively to fuzzy sets. The property of one-to-many classification is that often the signatures (training features) do not represent pure examples of each class. For instance, there are usually many tree species on one sample plot, so that the training data consist of different grades of membership of each class on each plot.

In paper I the fuzzy classification was used to estimate the proportions of plot volume by tree species. Spectral and textural features calculated from the ortho-rectified images were used in classification. Three fuzzy classification approaches were studied: a posteriori probabilities of maximum likelihood (ML) classification, a fuzzy classification based on the underlying logic of fuzzy sets (FZ) and linear mixture modelling (LMM). To overcome the problem of non-categorical signatures, an approach presented by Wang (1990) was employed. This takes into account the fuzziness of signatures and enables the use of partial class memberships.

3.4 The k-MSN method

The k-MSN method is a non-parametric nearest neighbour method which uses canonical correlation analysis to produce a weighting matrix for selecting the k Most Similar Neighbours from the reference data, i.e. observations which are similar to the object of

prediction in terms of the predictor variables. It is possible by means of canonical correlations to find the linear transformations U_k and V_k for the set of dependent variables (Y) and independent variables (X) which maximize the correlation between them:

$$U_k = \alpha_k Y, \text{ and } V_k = \gamma_k X, \quad (4)$$

where α_k represents the canonical coefficients of the dependent variables and γ_k the canonical coefficients of the independent variables. U_k and V_k are ordered in such a manner that the canonical correlation is largest for $k = 1$, second largest for $k = 2$, etc. The predictive power is thus concentrated in the first few canonical components.

The MSN distance metric derived from the canonical correlation analysis is as follows (Moeur & Stage 1995):

$$D_{uj}^2 = (X_u - X_j) \begin{matrix} \Gamma \Lambda^2 \Gamma' \\ \text{p} \times \text{p} \end{matrix} (X_u - X_j)' \quad (5)$$

where X_u is the vector of the independent variables from the target observation, X_j is the vector of the independent variables from the reference observation, Γ is the matrix of canonical coefficients of the predictor variables and Λ is the diagonal matrix of squared canonical correlations. The k-MSN method used in this thesis is the same as the MSN method described by Moeur & Stage (1995), except that the estimates for the target observation are calculated as weighted averages of the k nearest observations in such a manner that the weighting is based on the inverse of the MSN distance. The weight W_{uj} of a reference plot j for the target plot u was calculated as follows:

$$W_{uj} = \frac{\left(\frac{1}{D_{uj}^2} \right)}{\sum_1^k \left(\frac{1}{D_{uj}^2} \right)}, \quad (6)$$

where k is the number of nearest observations and $u \neq j$. One property of the k-MSN method, as also of other nearest neighbour methods, is that they are multivariate, i.e. multiple dependent variables may be estimated simultaneously.

3.5 Diameter distribution based on k-MSN imputed trees

k-MSN imputed diameter distribution consist of trees which occur in the field measured reference plots that are used in the estimation of the target plot. This requires that, in addition to the plot data used in the k-MSN estimation, tree data is available. In this thesis the known attributes of each tree were: species, dbh (mm accuracy) and height (cm accuracy). When constructing diameter distribution for the target plot, the stem frequency represented by each tree on plot u is required. The plot level weight W_{uj} used in the k-MSN estimation can be used for this purpose (see eq. 6). Let T_{uj} denote a set of trees t which are imputed from the reference plot u for the target plot j , and n the number of trees in u :

$$T_{ij} = \{t_{1ij}, t_{2ij}, \dots, t_{n_{ij}}\}. \quad (7)$$

Then the eventual set of trees imputed from the k nearest reference observations for the target plot is as follows:

$$T_j = \{T_{1j}, T_{2j}, \dots, T_{kj}\}. \quad (8)$$

Here diameter distribution can also be considered a set of trees which are simply ordered by diameter and within trees of the same diameter tree height, for instance, varies. Because species-specific distributions were of interest, tree species were treated separately in equations 7 and 8. Let A denote a factor that multiplies stem number to the hectare level:

$$A = \frac{10000m^2}{pa}, \quad (9)$$

where pa is the area of the sample plot. Then the stem number represented by one tree is:

$$N_{i_{ij}} = W_{ij}A. \quad (10)$$

3.6 Diameter distribution based on the Weibull function

The predicted ALS-based species-specific stand variables, i.e. G , dgm and hgm , were used to predict theoretical basal area diameter distributions by tree species by applying the three parameter form of the Weibull function. The parameter prediction models (PPM) of Mykkänen (1986) were used in the case of Scots pine and deciduous tree species and the models of Kilkki et al. (1989) for Norway spruce. The probability density function of the Weibull distribution is:

$$f(x) = \begin{cases} \frac{c}{b} \left(\frac{x-a}{b} \right)^{c-1} \exp\left(- \left(\frac{x-a}{b} \right)^c \right), & a \leq x < \infty \\ 0, & x < a \end{cases} \quad (11)$$

where x is a random variable, i.e. tree diameter in cm, a is the lower limit of the distribution, b is a scale parameter, and c is a shape parameter.

The 1-cm diameter class frequencies were first obtained from the cumulative distribution of relative basal area after which they were scaled to absolute values by multiplying them with the predicted species-specific basal area. Since only trees with a diameter over 5 cm were considered, the basal area proportion of smaller trees were added to diameter classes above 5 cm in a relative way. The number of stems in each diameter class was obtained by dividing the absolute value of the class basal area by the calculated basal area of the mean tree in the given class. The predicted species-specific frequencies of a diameter distribution were further calibrated against the ALS-based number of stems estimate using calibration estimation (see Deville and Särndal 1992; Kangas and Maltamo 2000a).

For the calculation of class-wise stem volumes, tree height and stem volumes were predicted for each class mean tree. Tree heights were predicted using Siipilehto's (1999) H-D curves and stem volumes were calculated using the volume models of Laasasenaho

(1982). Plot level stem volumes were obtained by summing over the class-wise stem volume predictions that were calculated by multiplying volume predictions of class mean trees by the corresponding class-wise stem frequencies.

3.7 Estimation of growing stock by tree species

Paper I compares two approaches for determining species-specific volumes and total volume at the plot level. The first approach consists of two stages, prediction of total volume using ALS data and assignment of this total volume to tree species by fuzzy classification and aerial photographs. Linear mixed-effect modelling was employed to estimate the total volume, because there is a hierarchical structure in the study design. In the second approach, volumes by tree species and the total volume were predicted simultaneously using the k-MSN method based on both ALS data and aerial photographs in one phase. Only species-specific volumes were predicted in the k-MSN approach, the total volume being obtained by summing the species-specific estimates. The number of nearest observations (k) was determined heuristically. In both approaches numerous combinations of predictor variables were tested by manual insertion and deletion and the best combination was selected. In the case of fuzzy classification there were also combinations in which only textural or spectral features were present. Both approaches guarantee that the sum of the species-specific volumes will logically give the total volume.

In paper II the species-specific forest variables volume, stem number, basal area and diameter and height of the basal area median tree were estimated for Scots pine, Norway spruce and deciduous trees and the total characteristics were determined as sums of the species-specific estimates. The median tree characteristics were estimated only by tree species because these are not of interest in Finland when all the tree species are lumped together. The k-MSN method was applied in the estimation in which all the species-specific attributes were estimated simultaneously. Height and density metrics of ALS data and spectral and textural features calculated from the ortho-rectified and radiometrically corrected images were used as predictor variables. It had already been observed in the previous tests that the selection of proper predictor variables for the k-MSN model is a time-consuming task and should be automated somehow. Therefore an algorithm was developed the aim of which is to minimize the weighted average of the relative RMSEs. It should be noted that when there are many dependent variables the cost function used in the minimization is not unambiguous; some control must be maintained over how much weight is given to the different dependent variables. The algorithm developed here takes into account the multiple criteria involved in the problem by letting the user define the weights.

Paper III investigates the ability of remote sensing information to predict species-specific diameter distributions. The first nearest neighbours were established and growing stock was estimated using the k-MSN approach developed in paper II. Then two approaches were compared: a method based on nearest neighbours, and a theoretical diameter distribution model. In the nearest neighbour approach the k-MSN imputed trees were used as such to predict diameter distribution. In the theoretical diameter distribution model the ALS based estimates were used to predict basal area diameter distributions by tree species by applying the Weibull function and further calibration estimation. These methods are explained in detail in sections 3.5 and 3.6.

Paper IV presents a method in which unrectified aerial photographs of known internal and external orientation are utilized. The use of a combination of ALS data and

orthorectified aerial photographs has been studied in papers I-III. However, there are some weaknesses in this approach: aerial photographs need radiometric correction, and the ALS points and aerial photographs are not properly fused due to the radial displacement. When ALS points are linked to unrectified aerial photographs of known orientation parameters, better data fusion is attained. In addition, each ALS point is mapped to several aerial photographs and the average of DN values is utilized; this averaging is considered to be a good substitute for radiometric correction. Bandwise DN averages are then used as patterns to classify each point as a pine, spruce or deciduous hit and the proportions of ALS points by tree species are calculated for each plot. This procedure as a whole is described in section 3.2.3. Finally the same k-MSN approach as earlier was used to estimate growing stock volume by tree species but instead of using aerial photograph based features directly species proportions were used along with ALS predictors in the modelling. A selection of predictor variables was based on an algorithm which has many similarities with the one presented in paper II. An advantage of the algorithm used here is that it enables the modification of numerous variables at the same time, which is useful especially when avoiding local minima. In addition, due to the flexibility of the parametrization, it is easier to control the number of variables included in the solution.

3.8 Accuracy assessment

The Root Mean Square Error (RMSE) is quoted in the literature as an established way of validating the accuracy of a forest inventory. It is a measure of variability that includes the effects of both bias and random error. An absolute RMSE is computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (12)$$

where n is the sample size, e.g. number of plots or stands, y_i is the observed value for sample i and \hat{y}_i is the predicted value for sample i . Correspondingly, the bias in absolute terms is calculated as follows:

$$bias = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}. \quad (13)$$

Both absolute and relative forms of bias and RMSE were used in all papers. The relative figures were computed as percentages by dividing the absolute value by the true mean of the characteristic concerned. In paper II the plots were divided into two groups, for use as modelling data (265 plots) or test data (198 plots), and both modelling and accuracy assessment took place at the plot level. In papers II and III the results were calculated by means of cross-validation, for which an additional condition was imposed, that the similar neighbours were never taken from the same stand as the target plot, to avoid excessively optimistic results. The estimation in papers II and III was carried out at the plot level but the results and reliability characteristics were presented at both the plot and the stand level. The stand level results were computed by taking the mean value for all the plots within one

stand. In paper IV the reliability characteristics were calculated at plot level by means of cross-validation.

In diameter distribution study (III) the accuracy of dominant height and log and pulpwood estimates were considered as well. Dominant height was chosen as test variable because it is not predicted directly but is based on predicted distribution, and it has considerable importance in the prediction of the future development of the tree stock (e.g. Hynynen et al. 2002). Log and pulpwood volumes were chosen as test variables because they can be regarded as the most important end products derived from diameter distribution (Siipilehto 1999). Timber assortments were modelled using the taper curves of Laasasenaho (1982). In addition, the goodness of fit of the predicted diameter distributions was measured with an error index which is fairly similar to the one proposed by Reynolds et al. (1988).

In Paper IV the RMSE was also calculated for the volume of pine, spruce and deciduous trees by taking into account only those plots in which the tree species in question represented more than 10% of the total volume in the plot. This alternative way to calculate RMSE tries to address the issue of numerous “near to zero” observations and it also resembles the approach currently used in Finnish inventories by compartments, i.e. minor tree species are not considered. It was also investigated in paper IV how well the main tree species is preserved in the estimates.

4 RESULTS

4.1 Volume estimates by tree species

The k-MSN method produced considerably more accurate estimates of species-specific volumes (see paper I) than any of the fuzzy classification methods (Table 1), the relative RMSEs for the volumes of pine, spruce and deciduous trees being 45.50%, 61.98% and 92.30%, respectively, where the corresponding figures for the best fuzzy classification-based method, ML, were 62.87%, 81.41% and 148.75%. The other fuzzy classification methods, FZ and LMM, yielded highly imprecise estimates. The relative RMSE of the total volume was 18.88% for the methods based on fuzzy classification and 23.86% for the k-MSN method. Thus the precision of the fuzzy classification approach in which the total volume is predicted beforehand using mixed model regression is slightly better than that of the method in which the total volume is obtained by summing the species-specific volumes,

Table 1. Accuracy and bias of the volume estimates by tree species and the total volume estimate. ML = maximum likelihood classification, FZ = fuzzy classification, LMM = linear mixture modelling and k-MSN = k-Most Similar Neighbour.

Method	RMSE (%)				BIAS (%)			
	<i>Pine</i>	<i>Spruce</i>	<i>Decid.</i>	<i>Total</i>	<i>Pine</i>	<i>Spruce</i>	<i>Decid.</i>	<i>Total</i>
ML	62.87	81.41	148.75	18.88	1.40	10.70	-57.87	-1.48
FZ	83.73	80.35	289.19	18.88	52.00	-2.85	-243.85	-1.48
LMM	79.83	113.83	265.36	18.88	24.65	-21.25	-52.39	-1.48
k-MSN	45.50	61.98	92.30	23.86	1.91	-9.92	3.51	-2.48

although the accuracy of the total volume estimate for the k-MSN method, 23.86% at the plot level, can be considered very good as well.

The biases were also considerably higher in the fuzzy classification approaches than with the k-MSN method (Table 1). ML was the only fuzzy classification method which produced fairly unbiased estimates, except for deciduous trees. The biases of the species-specific estimates in the k-MSN method were minor for pine (1.91%) and deciduous trees (3.51%), only the volume of spruce was moderately overestimated (-9.92%). The level of bias of the total volume estimates was always acceptable.

The final models always contain spectral and textural features from the aerial photographs and laser-based height and density metrics. As expected, it was perceived that there is a clear dependence of the NIR band on the abundance of deciduous trees, and that textural features are useful for discriminating between the coniferous tree species. The manual selection of proper explanatory variables into the k-MSN model turned out to be a laborious task.

4.2 k-MSN estimates of species-specific stand attributes

The optimization of predictor variables in paper II resulted in 25 predictors altogether, 7 of which were aerial photograph based and 18 ALS based. The aerial photograph-based variables included both spectral and textural features. In the final solution the number of nearest observations (k) was set at 5. The accuracy and bias of the estimated stand characteristics are presented in Table 2.

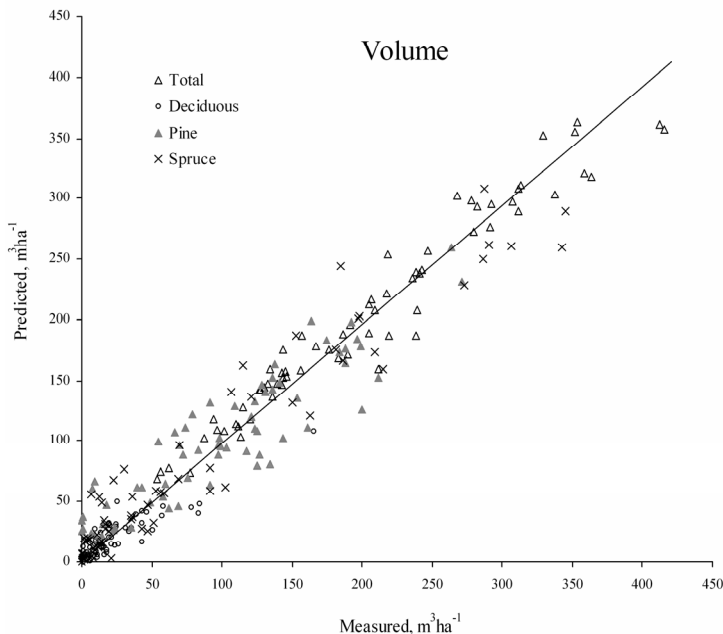


Figure 2. Measured versus predicted volumes by tree species and total volume at the stand level.

The relative RMSEs were highest for the attributes of deciduous trees. This was expected, because their numbers are rather small in the test area and quite often they do not occur at all in the dominant canopy layer. The spruce and pine characteristics were estimated considerably better. Pine was usually predicted slightly better than spruce, but this was partly an outcome of the weights selected for optimization of the predictor variables. The height of the basal area median tree was estimated most accurately, whereas the stem number was the most inaccurate. The stand level species-specific accuracies achieved here can be considered acceptable by comparison with a conventional field inventory by compartments (Haara & Korhonen 2004).

Total characteristics were estimated accurately, especially at the stand level, where the RMSEs for volume and basal area were approximately 10% and for the stem number slightly over 15%. These figures are much better than those attained with current inventory practises in Finland. All the biases were insignificant at the stand level, but at the plot level the bias was significant at the 95% confidence level in the case of deciduous volume, and both dgm and hgm for pine (Table 2). The biases of the total characteristics were always very low.

Table 2. Accuracy and bias of the estimated stand characteristics at the plot and stand levels. ✱ denotes that the bias was significant at the 95% confidence level.

	Plot level		Stand level	
	RMSE (%)	BIAS	RMSE (%)	BIAS
<i>Deciduous</i>				
V, m ³ ha ⁻¹	102.84	2.16✱	62.33	2.46
G, m ² ha ⁻¹	87.76	0.23	52.53	0.29
N, ha ⁻¹	89.69	19.85	47.55	27.92
dgm, cm	45.93	0.20	25.34	-0.17
hgm, m	32.15	0.10	18.40	-0.30
<i>Pine</i>				
V, m ³ ha ⁻¹	51.55	-1.11	28.08	-2.07
G, m ² ha ⁻¹	46.56	-0.22	27.05	-0.34
N, ha ⁻¹	60.53	-13.50	40.81	-13.10
dgm, cm	23.08	0.72✱	16.91	0.11
hgm, m	16.04	0.37✱	8.46	0.09
<i>Spruce</i>				
V, m ³ ha ⁻¹	55.72	-0.61	32.64	-0.41
G, m ² ha ⁻¹	51.19	-0.09	31.30	-0.08
N, ha ⁻¹	63.61	-2.34	38.07	-11.15
dgm, cm	33.06	-0.12	20.15	0.04
hgm, m	29.97	-0.14	17.63	-0.02
<i>Total</i>				
V, m ³ ha ⁻¹	20.51	0.44	10.36	-0.02
G, m ² ha ⁻¹	17.15	-0.08	8.63	-0.14
N, ha ⁻¹	29.95	4.01	15.88	3.66

An example of measured versus predicted growing stock volumes by tree species and total volumes at the stand level is given in Figure 2. The solid line indicates a perfect fit for the prediction line. The trend is predicted fairly well with respect to both the species-specific volumes and total volume. It is worth noting that, due to the small numbers of deciduous trees, the imprecision in terms of the relative RMSE is no longer in evidence when the predictions are plotted against the measured observations, because the absolute scatter is approximately the same or even smaller than that for the other tree species.

4.3 Diameter distributions

The k-MSN distribution produced more precise estimates than the Weibull method for spruce, deciduous trees and total volumes; but in the case of pine, the RMSE was slightly lower when using the Weibull method. Biases were generally much lower in the k-MSN method. With respect to deciduous trees both distribution modelling methods produced some bias, which was partly caused by the biased estimate of the deciduous volume. It is worth noting that the distribution based on k-MSN imputed trees provides exactly the same results as the k-MSN estimation.

Regarding timber assortments the k-MSN distribution provided generally more precise estimates of logwood and pulpwood volumes than the Weibull distribution (Table 3). An exception was the pulpwood volume of deciduous trees, where the Weibull distribution was slightly more precise. The biggest difference was in the case of spruce, where the RMSE of pulpwood volume was 61.92% for the Weibull and 47.71% for the k-MSN distribution. The corresponding figures for logwood were 61.06% and 43.56%. In addition, the Weibull distribution provided highly biased estimates for the pulpwood and logwood volumes of spruce. With respect to estimates of deciduous trees all relative RMSEs are large and mostly biased although in absolute terms these figures are quite small. In the case of pine both methods provided almost equal accuracy.

Dominant height was estimated more precisely by the k-MSN method than the Weibull distribution although the difference was modest (RMSE 5.61% vs. 6.23%). Note that even if dominant height was not predicted directly but was based on predicted distribution the obtained accuracy level was very good. One reason for this is that dominant height does not take into account tree species and ALS data is able to describe the height profile very well. Dominant height was somewhat underestimated (-0.43m) when the Weibull distribution

Table 3. Accuracy of the log- and pulpwood estimates by tree species for both distribution modelling methods.

	<i>Weibull distribution</i>		<i>k-MSN distribution</i>	
	RMSE (%)	BIAS (%)	RMSE (%)	BIAS (%)
Pine Pulpwood	31.86	2.29	30.30	5.38
Spruce Pulpwood	61.92	25.85	47.71	2.05
Deciduous Pulpwood	72.97	2.51	74.30	-9.08
Pine Logwood	40.97	-3.27	39.97	-0.36
Spruce Logwood	61.06	-25.78	43.56	-0.46
Deciduous Logwood	142.93	-37.62	114.54	-17.63

was used. Correspondingly, the k-MSN method slightly overestimated (0.38m) dominant height.

Error indices which show how well predicted diameter distributions correspond to actual ones also indicated the superiority of the k-MSN distribution: k-MSN distributions always fitted better to the observed distribution than the Weibull distribution did. In the cases of pine and deciduous trees the difference was rather small but in the case spruce the k-MSN method was able to describe diameter distribution much better than the Weibull distribution.

4.4 Rigorous data fusion method

A first step in the two staged method developed in paper IV was to classify ALS points as pine, spruce or deciduous hits. Several combinations of variables and their transformations were tested in order to find the ultimate form of the discrimination function. The final LDA model contained logarithms of two band ratios. A reason for using band ratios instead of the original DN values was to further stabilize the discrimination model. The overall classification accuracy of the LDA model in the training data was 89.8%.

The number of variables included in the k-MSN model was kept low in the optimization of predictor variables by setting the parameters accordingly. There are 9 predictors as a whole, 7 of which are ALS based variables and 2 are species proportions of classified ALS hits. It is worth noting that when two species proportions are included in the model, the third one is no longer needed, because this information is already included implicitly in the formulation.

The RMSE and bias of the volume estimates are presented in Table 4. The relative RMSE was considerably lower for pine (33.63%) than for spruce (62.86%) or deciduous trees (83.95%). A reason for the lower relative precision in the case of spruce and deciduous trees lies, of course, in the small mean volume. The total volume was predicted quite accurately (20.28%) by summing species-specific estimates. All the biases were basically negligible. Table 4 also depicts precisions when RMSE is calculated both in absolute and relative terms based on only those plots in which the tree species in question represented more than 10% of the total volume in a plot. The absolute RMSE of pine increased slightly whereas the relative RMSE decreased slightly compared to calculations based on the inclusion of all observations. With respect to spruce and deciduous trees, in absolute terms RMSEs increased substantially but, on the other hand, relative RMSEs

Table 4. Accuracy of the volume estimates by tree species and the total volume estimate at the plot level. Subscript $s_{>10\%}$ denotes that only to those plots in which the tree species in question represented more than 10% of the total volume in a plot were included in the accuracy assessment.

	Pine	Spruce	Decid.	Total
$RMSE, m^3 ha^{-1}$	30.13	26.14	12.95	29.73
$RMSE, \%$	33.63	62.86	83.95	20.28
$BIAS, m^3 ha^{-1}$	0.82	-0.72	-0.82	-0.71
$RMSE_{s_{>10\%}}, m^3 ha^{-1}$	30.78	38.18	20.31	--
$RMSE_{s_{>10\%}}, \%$	30.23	43.05	61.74	--

decreased considerably due to the greater mean volume. It must be taken into account that the results in Table 4 are presented at the plot level. At the stand level the accuracy would be considerably better.

It was also investigated how well the main tree species is preserved in the estimates. The number of pine plots was slightly overestimated whereas the numbers of spruce and deciduous plots were slightly underestimated. The overall classification accuracy of the main tree species was 92.1% at plot level. The most frequent source of misclassification was that spruce plots were classified as pines or vice versa.

5 DISCUSSION

A great deal of interest is currently being shown in remote sensing based forest inventories in Finland. A particular driving force has been the possibility of reducing costs. Because technological development has decreased the costs of ALS data considerably, ALS data based inventory systems have emerged as a realistic alternative for operative forestry purposes. The nearest neighbour method employed in this thesis turned out to be an efficient and versatile approach to estimating total and species-specific forest attributes and diameter distributions when using ALS data, aerial photographs and a set of fixed size sample plots as data sources.

The aim of paper I was to test two approaches to predicting volume by tree species and the total volume using ALS data and aerial photographs. The primary conclusion was that the k-MSN method produced much more precise estimates of species-specific volumes than any fuzzy classification method. It is obvious that the manner in which the FZ and LMM methods were used here is not suitable for estimating species-specific volumes, as both methods produced imprecise and highly biased estimates. ML was the best fuzzy classification method studied, but it still produced biased and imprecise estimates compared with the k-MSN method. Although there was still a great deal of imprecision in the k-MSN estimates, one very good characteristic is that the species-specific estimates sum up to an accurate total volume without any notable bias. The manual testing of different combinations of explanatory variables was fairly time-consuming due to the large number of candidate predictors and their transformations, and the outcome is still more or less subjective. A heuristic variable selection algorithm was therefore implemented in paper II.

The results of paper II were rather encouraging, because the volume estimates were approximately as accurate as in paper I but other species-specific and total attributes were estimated, too. Stand level accuracies were also calculated in paper II, which enabled a comparison with the current field inventory method. The stand level species-specific accuracies were comparable with those achieved in a practical field inventory by compartments (Haara & Korhonen 2004), while the total characteristics were estimated much more accurately than with the current inventory system (Kangas et al. 2004). No obvious bias was detected. One characteristic of the k-MSN modelling approach is that it is of the multivariate type, i.e. a number of dependent variables are estimated simultaneously. Therefore, when selecting parameters for the model based on the goodness of the result, a weight must be given to each sub-component. An explicit weight vector was used here to determine the importance of the dependent variables, and the aim in using a variable selection algorithm was to minimize the weighted average of the RMSEs. This proved to be

a reasonable principle for selecting predictor variables into the k-MSN model, although certainly the algorithm itself could be enhanced. The method used in paper II is straightforward, practical and easy to apply. All the stand characteristics are estimated simultaneously and the dependent variables can be altered as needed. The logical integrity of the results is guaranteed because total characteristics are produced by summing the species-specific estimates.

Diameter distributions by tree species are compulsory in Finland. Therefore the ability of the nearest neighbour method used in this thesis to predict diameter distributions was investigated in paper III. Basically all test criteria indicated that the diameter distribution based on nearest neighbour imputed trees outperforms the theoretical Weibull distribution. Especially the log- and pulpwood volumes were estimated more accurately when using the k-MSN distribution and the same tendency was apparent in the error index, too. Volume estimates were somewhat more accurate using the k-MSN distribution; this is obvious since the k-MSN distribution volumes consist of exactly the same trees that sum up to the volumes of k-MSN estimates. The same applies to other predicted variables too, and therefore calibration estimation, for example, is never needed. The k-MSN distribution is strictly speaking just a set of trees which are ordered by diameter, i.e. a tree list. Each predicted tree has those attributes which are available in the modelling trees. Here the attributes were species, dbh and height; therefore, within trees of the same diameter, tree height varies and the H-D curve is redundant. Generally speaking, if they have been measured, it makes sense to utilize tree diameters rather than a theoretical distribution. However, it must be kept in mind that in the nearest neighbour imputation it is always essential to have reference data that is comprehensive enough. When tree species is considered as well, the required number of sample plots increases substantially because field data must be representative in terms of both size and species compositions.

Certain issues are inherent to the method examined in paper IV. First of all, unrectified aerial photographs with known internal and external orientation are used instead of orthorectified images. The aim of using unrectified images with known orientation is the rigorous linkage of ALS points and aerial photographs. Although there are some restrictions also in this approach when linking ALS points to aerial images, this is a more accurate method for combining spectral characteristics of aerial images and ALS data than the use of orthorectified images. Another peculiar characteristic of the method in paper IV is the avoidance of radiometric correction of aerial photographs. The estimation procedure itself has two separate stages. First, spectral characteristics of aerial photographs are attached to laser points and the points are classified by tree species. At the second stage the proportions of ALS points by tree species and ALS height and density metrics are used to estimate forest characteristics by the k-MSN method. The major difference between the previous studies (I-III) and this one is that here the proportions of ALS points by tree species are used as independent variables, whereas earlier textural and spectral features of aerial photographs were used directly as independent variables in the k-MSN modelling. The accuracies achieved in paper IV are not directly comparable to the previous studies because the test areas and remote sensing material used are different. However, when looking at the species-specific accuracies and measured versus predicted figures it is obvious that the method presented in paper IV works rather well. Although here only plot volumes were considered, the method itself is easily expandable to predict other stand attributes and to generate diameter distributions, as presented earlier.

The essential difference of spectral images and ALS data lies in the type of response. Spectral images depict the intensity of the electromagnetic radiation in the selected

wavelength areas, whereas ALS measures physical dimensions. It is obvious that the latter has a considerably better relation to most forest attributes of interest. ALS based inventory techniques can in a sense be compared with 3D photogrammetric inventory techniques which aim at the extraction of physical dimensions. The difference is that ALS produces physical dimensions directly while photogrammetric techniques rely on some extraction method. As extraction methods are always error prone, it is difficult to imagine an application in which spectral images, such as aerial photographs, produce better information on forest dimensions than ALS. The focus of the present thesis is in tree species-specific estimation. In the beginning of this work it was assumed that ALS data bring information about forest dimensions and that aerial photographs are useful in order to discriminate between tree species. In retrospect this assumption proved to be reasonable.

An inventory system in which aerial photographs are not needed at all for estimating tree species-specific attributes would be advantageous. Then certain difficulties entailed in the computer-aided interpretation of aerial photographs could be avoided. From the operational point of view, when also considering economical aspects, it is especially difficult to utilize aerial images. Inventory areas are often scattered and aerial surveys must be carried out separately in different areas for economical reasons and weather conditions. As a consequence of this, images are taken under different lighting conditions, different points in time, and even with different imaging instruments. It is very challenging to utilize this kind of data due to difficulties originating from the need of radiometric comparability. Experiments should therefore be carried out to see whether aerial photographs could somehow be eliminated altogether. However, in the studies presented here the use of aerial photographs improved the accuracy of the estimated species-specific attributes noticeably, although this comparison is not directly presented in this thesis. Tree level forest inventories based on ALS data may also be a realistic alternative in operational forestry in the near future. Tree level inventories require denser ALS data than the method used here, but technological development will mean that costs will decrease rapidly. There are, however, only a very few reported studies in which entire individual tree based inventory systems have been validated. An approach in which aerial photographs are not needed would be interesting on the individual tree level too.

It is clear that field visits cannot be avoided in stand level forest inventories. First of all, basic information about stands, such as fertility, soil type etc., must be determined in the field; otherwise they can be inherited from the old stand register data – if approximately the same stand delineation still exists. Another issue concerns management proposals. There are basically two options: silvicultural treatments can be determined in the field, or the forest management system can decide on them based on the current state of the forest growing stock and the decision-maker's preferences. Furthermore, the inventory approach studied here is based on the concept of collecting new field data for each region where an inventory is to be carried out. This has the advantage that modelling data is always representative of local conditions and tree species. On the other hand, the inventory area must be large enough for it to be worth collecting new sample plot material each time the method is applied. Another reason for the use of local field data is that otherwise it is difficult to use aerial photographs, because variables based on these are not directly transferable from one dataset to another.

All the deciduous species were lumped together in the present calculations, as it was known beforehand that they are virtually impossible to separate by remote sensing methods under Finnish conditions, where the forests are dominated by coniferous species and the vast majority of the deciduous trees are birches. In addition, single deciduous trees often

occur below the dominant tree layer. The fact that it is very difficult to estimate the characteristics of non-dominant tree layers using remote sensing techniques must be taken into account when putting these accuracies into context. The absolute errors regarding deciduous trees are nevertheless quite small because the numbers of deciduous trees in Finnish forests are rather small. It must also be considered that the current field inventory system produces approximately equally inaccurate estimates for deciduous trees as were obtained here. An obvious advantage of a field-based inventory by compartments is that deciduous tree species can be distinguished. The accuracy of the figures for deciduous tree species other than birch is probably quite poor in the current field inventory system, however, as the minority tree species are grouped together with some other species if their stand level basal area is less than 1m²/ha. Therefore other deciduous tree species are often recorded as birches. It is doubtful whether it will ever be possible to develop a remote sensing based forest inventory method that could discriminate between the different deciduous tree species properly in Finnish conditions. Already the accurate estimation of deciduous trees as species group is a difficult task.

The inventory method studied here is well-suited for wall-to-wall mapping of forest resources. Wall-to-wall coverage may be achieved either by using fixed size grid cells or variable size segments. In the case of segments a crown height model interpolated from the ALS data is a natural choice to carry out the segmentation (see e.g. van Aardt et al. 2006). Grid cells or segments provide a good basis for spatial forest planning in which treatment units are generated without predefined stand boundaries (Heinonen 2007). It must be kept in mind, however, that the accuracy of the estimated forest attributes depends on the inventory unit: at cell or segment level the accuracy is always worse than at stand level. So far research on ALS based forest inventories has focused on young, middle-aged and mature forests. However, in practice it is oversimplified to restrict the inventory to only some predefined set of more or less mature stands – how can one be sure beforehand that it is a mature stand? An unavoidable circumstance in wall-to-wall mapping is that there are patches within stands that belong to other development classes. Stands are delineated from the silvicultural point of view; there is some predefined minimum size of a stand, the shape is restricted, and in practice stand borders generalize natural boundaries. For instance, when applying the inventory there may be some cells or segments within a mature stand that actually belong to sapling stands. On these occasions it must be decided whether sapling stand patches are inventoried along with more mature patches or should they be treated somehow separately. This issue should be considered and studied carefully and more research should be carried out on young sapling stands.

The commonly used measure of accuracy in forest inventory is the relative RMSE. This does not reflect the accuracy properly on such occasions where many near to zero or absolute zero observations occur. This is a typical situation when looking at species-specific estimates and their accuracies obtained in this thesis. The mean volume of deciduous trees, for instance, is low because quite often there are only some deciduous trees in a stand, or deciduous trees are even completely non-existent, i.e. deciduous trees often represent minor tree species. Because of low mean volume the relative accuracy of the volume of deciduous trees is quite poor although in absolute terms the accuracy may be even comparable to pine estimates. This phenomenon is especially apparent when looking at Figure 2 in paper IV. The same tendency applies to spruce as well. On the other hand, in Finnish forest management the main tree species is favoured. This has the consequence that in Finland in managed forests one tree species often dominates and others are in the minority. This again leads to the conclusion that tree species other than the main one which

occur in a stand may be considered less important, and that the current way to assess the accuracy does not take this into account. An alternative measure of accuracy was introduced in paper IV. It can be used as an additional measure of accuracy of species-specific estimates if it holds true that the minority tree species are less important than the main tree species. This measure is identical to RMSE but only those observations where the tree species in question represents more than 10% of the total volume are taken into account. The 10% was selected to be a threshold here but naturally it may be justified to adapt it to reflect better particular requirements. One alternative that also addresses the specific nature of species-specific accuracy is to calculate RMSE separately for the main tree species and for the minor tree species. One should bear in mind why forest inventories are performed and that further applications of the information obtained set certain requirements for its quality. It is exceedingly important to use valid metrics.

REFERENCES

- Aldrich, R.C. & Norick, N.X. 1969. Stratifying Stand Volume on Non-Stereo Aerial Photos. PSW-51, Pacific Southwest Forest and Range Experiment Station, Berkeley, California. 19 p.
- Anttila, P. 1998. Analytyttisellä stereoplotterilla ilmakuvilta tulkittujen puukohtaisten tunnusten tarkkuus. M.Sc. thesis, University of Joensuu, Faculty of Forestry. 36 p. (In Finnish).
- 2005. Assessment of manual and automated methods for updating stand-level forest inventories based on aerial photography. *Dissertationes Forestales* 9. Doctoral thesis, University of Joensuu, Faculty of forestry. 42 p.
- & Lehikoinen, M. 2002. Kuvioittaisten puustotunnusten estimointi ilmakuvilta puoliautomaattisella latvusten segmentoinnilla. *Metsätieteen aikakauskirja* 3/2002: 381–389. (In Finnish).
- Aronoff, S. 2005. *Remote Sensing for GIS Managers*. ESRI Press, Redlands. 524 p.
- Axelsson, P. 2000. DEM Generation from Laser Scanner Data Using Adaptive TIN Models. *Proceedings of the XIXth ISPRS Conference, IAPRS, Vol. XXXIII*, p. 110–117, Amsterdam, The Netherlands.
- 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.
- Brandtberg, T. 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry & Remote Sensing* 61:325–340.
- & Walter, F. 1998. Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. *Machine Vision and Applications* 11: 64–73.
- Coburn, C.A. & Roberts, A.C.B. 2004. A Multiscale Texture Analysis Procedure for Improved Forest Stand Classification. *International Journal of Remote Sensing* 25: 4287–4308.
- deSteiguer, J.E. 1978. Forestry Applications of NASA Remote Sensing Programs. *Journal of Forestry* 76: 208–211.
- Deville, J.-C. & Särndal, C.-E. 1992. Calibration estimators in survey sampling. *Journal of the American Statistical Association* 87: 376–382.
- Dobson, M.C. 2000. Forest information from synthetic aperture radar. *Journal of Forestry* 98: 41–43.
- Dralle, K. & Rudemo, M. 1996. Stem number estimation by kernel smoothing in aerial photos. *Canadian Journal of Forest Research* 26: 1228–1236.
- Duncan, G.J. & Kalton, G. 1987. Issues of Design and Analysis of Surveys across Time. *International Statistical Review* 55: 97–117.
- Fabritius, B. 1922. Flygmaskinen i skogbrukets tjänst. *Forstlig tidskrift* 39. (In Swedish).
- Foody, G.M., Jackson, R.G. & Quine, C.P. 2003. Potential improvements in the characterization of forest canopy gaps caused by windthrow using fine spatial resolution multispectral data: comparing hard and soft classification techniques. *Forest Science* 49: 444–445.

- Franklin S.E., Hall, R.J., Moskal, L.M., Maudie, A.J. & Lavigne, M.B. 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing* 21: 61–79.
- Fransson, J., Smith, G., Askne, J. & Olsson, H. 2001. Stem volume estimation in boreal forests using ERS-1/2 coherence and Spot XS optical data. *International Journal of Remote Sensing* 22: 2777–2791.
- Gagnon, P.A., Agnard, J.P. & Nolette, C. 1993. Evaluating of a soft-copy photogrammetry system for tree-plot measurements. *Canadian Journal of Forest Research* 23: 1781–1785.
- Gobakken, T. & Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research* 19: 529–542.
- & Næsset, E. 2005. Weibull and percentile models for lidar-based estimation of basal area distribution. *Scandinavian Journal of Forest Research* 20: 490–502.
- Gougeon, F. 1995. A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian Journal of Remote Sensing* 21: 274–284.
- Haara, A. & Korhonen, K.T. 2004. Kuvioittaisen arvioinnin luotettavuus. *Metsätieteen aikakauskirja* 4/2004: 489–508. (In Finnish).
- Haralick, R.M., Shanmugam, K. & Dinstein, J. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3: 610–621.
- Heinonen, T. 2007. Developing spatial optimization in forest planning. *Dissertationes Forestales* 34. Doctoral thesis, University of Joensuu, Faculty of forestry. 48 p.
- Hoekman, D.H. 1985. Radar backscattering of forest stands. *International Journal of Remote Sensing* 6: 325–343.
- Holmgren, J. & Persson, Å. 2004. Identifying species of individual trees using airborne laser scanner. *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: 1537–1552.
- Holmström, H., Nilsson, M. & Ståhl, G. 2001. Simultaneous Estimations of Forest Parameters using Aerial Photograph Interpreted Data and the k Nearest Neighbour Method. *Scandinavian Journal of Forest Research* 16: 67–78.
- Holopainen, M. & Wang, G. 1998. The calibration of digitized aerial photographs for forest stratification. *International Journal of Remote Sensing* 19: 677–696.
- Hynynen, J., Ojansuu, R., Hökkä, H., Siipilehto, J., Salminen, H., & Haapala, P. 2002. Models for predicting stand development in MELA System. *Metsäntutkimuslaitoksen Tiedonantoja* 835. 116 p.
- Hyvönen, P., Pekkarinen, A. & Tuominen, S. 2005. Segment-level stand inventory for forest management. *Scandinavian Journal of Forest Research* 20: 75–84.
- Hyypä, J. & Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland* 16: 27–42.
- , Hyypä, H., Inkinen, M., Engdahl, M., Linko, S. & Zhu, Y.-H. 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management* 128: 109–120.
- Jensen, J.L.R., 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.

- Kangas, A. & Maltamo, M. 2000a. Calibrating predicted diameter distribution with additional information. *Forest Science*: 46: 390–396.
- & Maltamo, M. 2000b. Performance of percentile based diameter distribution prediction and weibull method in independent data sets. *Silva Fennica* 34: 381–398.
- , Gove, J.H. & Scott, C.T. 2006. Introduction. In: Kangas, A. & Maltamo, M. (eds.). *Forest inventory - methodology and applications*. Springer, The Netherlands. p. 3–11.
- , Heikkinen, E. & Maltamo, M. 2004. Accuracy of partially visually assessed stand characteristics — a case study of Finnish forest inventory by compartments. *Canadian Journal of Forest Research* 34: 916–930.
- Kilkki, P., Maltamo, M., Mykkänen, R., & Päivinen, R. 1989. Use of the Weibull function in estimating the basal-area diameter distribution. *Silva Fennica* 23: 311–318.
- & Päivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. In: *Remote Sensing-Aided Forest Inventory. Proceedings of Seminars organised by SNS, 10–12 Dec. 1986, Hyytiälä, Finland. Research Notes No 19. Department of Forest Mensuration and Management, University of Helsinki, Helsinki.* pp. 209–215.
- King, D. 1991. Determination and reduction of cover type brightness variations with view angle in airborne multispectral video imagery. *Photogrammetric Engineering & Remote Sensing* 57: 1571–1577.
- Korhonen, K.T., Heikkinen, J., Henttonen, H., Ihalainen, A., Pitkänen, J. & Tuomainen, T. 2006. Suomen metsävarat 2004–2005, *Metsätieteen aikakauskirja*, 1B/2006: 183–221. (In Finnish).
- Korpela, I. 2000. 3-D matching of tree tops using digitized panchromatic aerial photos. Licentiate thesis. University of Helsinki, Faculty of Agriculture and Forestry, Department of Forest Resource Management. 109 p.
- 2004. Individual tree measurements by means of digital aerial photogrammetry. Doctoral thesis, *Silva Fennica Monographs* 3. 93 p.
- & Anttila, P. 2004. Appraisal of the mean height of trees by means of image matching of digitised aerial photographs. *Photogrammetric Journal of Finland* 19: 23–36.
- , Dahlin, B., Schäfer, H., Bruun, E., Haapaniemi, F., Honkasalo, J., Ilvesniemi, S., Kuutti, V., Linkosalmi, M., Mustonen, J., Salo, M., Suomi, O., Virtanen, H. 2007. Single-Tree Forest Inventory Using Lidar and Aerial Images for 3D Treetop Positioning, Species Recognition, Height and Crown Width Estimation. *Proceedings of the ISPRS Workshop Laser Scanning 2007 and Silvilaser 2007, September 12–14, 2007, Finland, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI (Part 3/W52).* pp. 227–233.
- Koukoulas, S. & Blackburn, G.A. 2005. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. *International Journal of Remote Sensing* 26: 431–455.
- Laasasenaho, J. 1982. Taper curve and volume function for pine, spruce and birch. *Communications Instituti Forestalis Fenniae* 108: 1–74.
- Larsen, M. 1997. Crown modeling to find tree top positions in aerial photographs. *Proceedings of the third International Airborne Remote Sensing Conference and Exhibition, Copenhagen, Denmark. ERIM international. Vol 2.* pp. 428–435.
- & Rudemo, M. 1998. Optimizing templates for finding trees in aerial photographs. *Pattern Recognition Letters* 19: 1153–1162.
- Liang, X., Hyyppä, J. & Matikainen, L. 2007. Deciduous-coniferous tree classification using difference between first and last pulse laser signatures. *Proceedings of the ISPRS*

- Workshop Laser Scanning 2007 and Silvilaser 2007, September 12–14, 2007, Finland, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI (Part 3/W52). pp. 253–257.
- Lillesand, T., Kiefer, R. & Chipman, J. 2004. Remote sensing and image interpretation. Fifth edition. New York, Wiley. 784 p.
- Lim, K., Treitz, P., Baldwin, K., Morrison, I. & Green, J. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing* 29: 648–678.
- Loetsch, F. & Haller, K.E. 1973. Forest Inventory. Volume I. BLV Verlagsgesellschaft mbH, München. 463 p.
- Lowell, K. 1998. Evaluation of a moving-window algorithm for detecting individual tree stems from scanned aerial photographs in Victoria's eucalypt forests. *Australian Forestry* 61: 226–234.
- Magnussen, S. & Boudewyn, P. 1998. Derivation of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research* 28: 1016–1031.
- Magnusson, M. 2006. Evaluation of remote sensing techniques for estimation of forest variables at stand level. *Acta universitatis agriculturae Sueciae* 85. Doctoral thesis, Swedish University of Agricultural Sciences, Faculty of Forest Sciences. 38 p.
- Maltamo, M., Erikäinen, K., Packalén, P. & Hyypä, J. 2006a. Estimation of stem volume using laser scanning based canopy height metrics. *Forestry* 79: 217–229.
- , Malinen, J., Packalén, P., Suvanto, A. & Kangas, J. 2006b. Nonparametric estimation of plot volume using laser scanning, aerial photography and stand register data. *Canadian Journal of Forest Research* 36: 426–436.
- , Mustonen, K., Hyypä, J., Pitkänen, J. & Yu, X. 2004. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. *Canadian Journal of Forest Research* 34: 1791–1801.
- , 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: 26–34.
- Marceau, D.J., Howarth, P.J., Dubois, J.M.M. & Gratton, D.J. 1990. Evaluation of the Grey-Level Co-occurrence Matrix (GLCM) Method for Land-Cover Classification Using SPOT Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 28: 513–519.
- Mattila, E. 1985. The combined use of systematic field and photo samples in a large-scale forest inventory in North Finland. *Communicationes Instituti Forestalis Fenniae* 131: 1–97.
- Mehtätalo, L., Maltamo, M., & Packalén, P. 2007. Recovering plot-specific diameter distribution and height-diameter curve using ALS based stand characteristics. *Proceedings of the ISPRS Workshop Laser Scanning 2007 and Silvilaser 2007, September 12–14, 2007, Finland, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI (Part 3/W52)*. pp. 288–293.
- Mikhail, E.M., Bethel, J.S. & McGlone, J.C. 2001. Introduction to modern photogrammetry. John Wiley & Sons. 479 p.
- Moeffiet, T., Mengersen, K., Witte, C., King, R. & Denham, R. 2005. Airborne laser scanning: exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *ISPRS Journal of Photogrammetry & Remote Sensing* 59: 289–309.

- Moeur, M. & Stage, A.R. 1995. Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning. *Forest Science* 41: 337–359.
- Morena, L.C., James, K.V. & Beck, J. 2004. An introduction to the RADARSAT-2 mission. *Canadian Journal of Remote Sensing* 30: 221–234.
- Muinsonen, E., Maltamo, M., Hyppänen, H. & Vainikainen, V. 2001. Forest stand characteristics estimation using a most similar neighbor approach and image spatial structure. *Remote Sensing of Environment* 78: 223–228.
- Mykkänen, R. 1986. Weibull-funktion käyttö puuston läpimittajakauman estimoinnissa. M.Sc. thesis, University of Joensuu, Faculty of Forestry. 80 p. (In Finnish).
- Myles, S. 1945. The Use of Aerial Photographs in Forestry. *Journal of Forestry* 43: 252–257.
- Næsset, E. 1991. Bestemmelse av middeldiameter, høydeklasse, og bruttoverdi i bestand ved hjelp av flybilder. Research paper of Skogforsk. 3/91. 19 p. (In Norwegian).
- 1997. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment* 51: 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: 88–99.
- 2004a. Practical large-scale forest stand inventory using a small airborne scanning laser. *Scandinavian Journal of Forest Research* 19: 164–179.
- 2004b. Accuracy of Forest Inventory Using Airborne Laser-scanning: Evaluating the First Nordic Full-scale Operational Project. *Scandinavian Journal of Forest Research* 19: 554–557.
- , Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å. & Söderman, U. 2004. Laser Scanning of Forest Resources: The Nordic Experience. *Scandinavian Journal of Forest Research* 19: 482–499.
- Näslund, M. 1937. Skogsförsöksanstaltens gallringsförsök i tallskog. *Meddelanden från Statens Skogsförsöksanstalt* 29. 169 p. (In Swedish).
- Needham, T. & Smith, J. 1987. Stem count accuracy and species determination in loblolly pine plantations using 35-mm aerial photography. *Photogrammetric Engineering & Remote Sensing* 53: 1675–1678.
- Nyysönen, A. 1955. On the estimation of the growing stock from aerial photographs. *Communications Instituti Forestalis Fenniae* 46: 1–57.
- 1959. Metsänarvioimistieteellinen tutkimus Suomessa vuosina 1909–1959. *Acta Forestalia Fennica* 70: 1–23. (In Finnish).
- Olofsson, K., Wallerman, J., Holmberg, J. & Olsson, H. 2006. Tree species discrimination using Z/I DMC imagery and template matching of single trees. *Scandinavian Journal of Forest Research* 21: 106–110.
- Ørka, H.O., Næsset, E. & Bollandsås, O.M. 2007. Utilizing airborne laser intensity for tree species classification. Proceedings of the ISPRS Workshop Laser Scanning 2007 and Silvilaser 2007, September 12–14, 2007, Finland, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI (Part 3/W52)*. pp. 300–304.
- Pellikka, P. 1998. Development of correction chain for multispectral airborne video camera data for natural resource assessment. *Fennia* 176: 1–110.
- Persson, Å., Holmgren, J. & Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering & Remote Sensing* 68: 925–932.

- , Holmgren, J. & Söderman, U. 2003. Detection, measurements, and species classification of individual trees for forest inventory and visualization. Proceedings of the Scandlaser scientific workshop on airborne laser scanning of forests, Umeå, Sweden, 2-4 September 2003. pp. 223–234.
- Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J. & Packalén, P. 2007. Pre-harvest measurement of marked stand using airborne laser scanning. *Forest Science* 63: 653–661.
- Pinz, A. 1999. Tree isolation and species classification. In: Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, BC, Feb. 10-12, 1998. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Center. pp. 127–139.
- Pitkänen, J. 2001. Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. *Canadian Journal of Forest Research* 31: 832–844.
- Pollock, R. 1999. Individual tree recognition based on a synthetic tree crown image model. In: Hill, D.A. & Leckie, D.G. (eds.). Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Feb 1998. Victoria, Canada. pp. 25–34.
- Pollock, R.J. 1996. The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown model. PhD-thesis, Computer Science, The University of British Columbia. 158 p.
- Popescu, S.C., Wynne, R.H. & Nelson, R.F. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing* 29: 564–577.
- Poso, S. 1972. A method of combining photo and field samples in forest inventory. *Communications Institutii Forestalis Fenniae* 76(1): 1–133.
- Rauste, Y., Häme, T., Pulliainen, J., Heiska, K. & Hallikainen, M. 1994. Radar-based forest biomass estimation. *International Journal of Remote Sensing* 15: 2797–2808.
- Reese, H., Nilsson, M., Granqvist Pahlén, T., Hagner, O., Joyce, S., Tingelöf, U., Egberth, M. & Olsson, H. 2003. Countrywide Estimates of Forest Variables Using Satellite Data and Field Data from the National Forest Inventory. *Ambio* 32: 542–548.
- 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: 373–399.
- Sarvas, R. 1938. Ilmavalokuvauksen merkityksestä metsätaloudessamme. *Silva Fennica* 48. 49 p. (In Finnish).
- Shugart, H.H., Bourgeau-Chavez, L. & Kasischke, E.S. 2000. Determination of stand properties in boreal and temperate forests using high-resolution imagery. *Forest Science* 46: 478–486.
- Siipilehto, J. 1999. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. *Silva Fennica* 33: 281–301.
- Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single Tree Segmentation Using Airborne Laser Scanner Data in a Structurally Heterogeneous Spruce Forest. *Photogrammetric Engineering & Remote Sensing* 72: 1369–1378
- Spencer, R. & Hall, R. 1988. Canadian large-scale aerial photographic systems (LSP). *Photogrammetric Engineering and Remote Sensing* 54: 475–482.
- Talts, J. 1977. Mätning i storskaliga flygbilder för beståndsdatainsamling. Royal College of Forestry, Department of Forest Mensuration and Management. Research notes 6. 102 p. (In Swedish).

- Tomppo, E. 2006. The Finnish National Forest Inventory. In: Kangas, A. & Maltamo, M. (eds.) *Forest Inventory. Methodology and applications. Managing Forest Ecosystems. Vol. 10.* Springer, Dordrecht. pp. 179–194.
- Tuominen, S. & Pekkarinen, A. 2004. Local radiometric correction of digital aerial photographs for multi source forest inventory. *Remote Sensing of Environment* 89: 72–82.
- & Pekkarinen, A. 2005. Performance of different spectral and textural aerial photograph features in multi-source forest inventory. *Remote Sensing of Environment* 94: 256–268.
- van Aardt, J.A.N., Wynne, R.H. & Oberwald, R.G. 2006. Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *Forest Science* 52: 636–649.
- Veltheim, T. 1987. Pituusmallit männylle, kuuselle ja koivulle. M.Sc. thesis, University of Helsinki, Department of Forest Mensuration and Management. 59 p. (In Finnish).
- Venables, W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S.* Fourth Edition. Springer. 495 p.
- Wang, F. 1990. Fuzzy Supervised Classification of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing* 28: 194–201.
- Wehr, A. & Lohr, U. 1999. Airborne laser scanning – an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing* 54: 68–82.
- Weszka, J., Dyer, C.R. & Rosenfeld, A. 1976. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics* 6: 269–285.
- Wulder, M., Niemann, K.O. & Goodenough, D. 2000. Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery. *Remote Sensing of Environment* 73: 103–114.
- , Dechka, J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudoin, A. & Franklin, S.E. 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. *Forestry Chronicle* 79: 1075–1083.