Forestry Advance Access published April 7, 2010

Comparing different methods for prediction of mean crown height in Norway spruce stands using airborne laser scanner data

M. MALTAMO[1]*, O. M. BOLLANDSÅS[1], J. VAUHKONEN[2], J. BREIDENBACH[1], T. GOBAKKEN[1] AND E. NÆSSET[1]

[1] Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, PO Box 5003, N-1432 Ås, Norway
[2] University of Eastern Finland, Faculty of Science and Forestry, School of Forest Sciences, Yliopistonkatu 7, rakennus Y9, P.O. Box 111, FI-80101 Joensuu, Finland
*Corresponding author. E-mail: matti.maltamo@joensuu.fi

Summary

In this study, different methods were used to predict mean crown height of Norway spruce-dominated stands by means of low pulse density airborne laser scanning (ALS) data. The methods were based on statistical modelling, properties of the laser point clouds or combinations of them. Separate modelling data were used for model calibration and two different validation datasets were used to assess the accuracy of the results. The results obtained were partly contradictory, showing varying performance of different methods using different datasets. However, there were also notable differences between the methods used to obtain crown height by field measurements. The root mean square error figures of crown height predictions were at minimum between 1.0 and 1.5 m. This study showed that statistical modelling based on ALS height metrics was a good approach if the relationship between mean crown height and the ALS information was corresponding in the modelling data and in the application phase. A method based on the alpha shape technique was also an accurate alternative. Methods that rely directly on the laser point cloud to predict mean crown height without any calibration were good alternatives to get relatively accurate results but there are still drawbacks (area of calculation unit) in their applicability.

Introduction

Airborne laser scanning (ALS) data provides excellent possibilities to predict different forest attributes because the height distribution of laser scanner data is strongly related to the vertical structure of the tree canopy (Magnussen and Boudewyn, 1998). Since ALS data comprise three-dimensional (3-D) canopy information through geo-referenced height measurements, characteristics such as tree height or mean height are obvious forest attributes to be obtained (e.g. Næsset, 1997, 2002, 2004a, b; Popescu *et al.*, 2002; Holmgren, 2004; Maltamo *et al.*, 2004; Thomas *et al.*, 2006; Breidenbach *et al.*, 2008b). However, the attributes that have been of main interest in the majority of ALS applications in forests are volume and biomass (e.g. Næsset, 2002, 2004a, b; Holmgren, 2004; Maltamo *et al.*, 2006a; van Aardt *et al.*, 2006; Breidenbach *et al.*, 2008c, Næsset and Gobakken, 2008). Although ALS data are strongly correlated with these attributes, allometric and statistical models are needed to estimate them (e.g. Næsset, 2002; Persson *et al.*, 2002; Takahashi *et al.*, 2005; Maltamo *et al.*, 2006a). Furthermore, ALS data are also correlated with tree diameter since height and diameter distributions of forest stands are correlated. Several studies have modelled diameter distributions from ALS data (Gobakken and Næsset, 2004, 2005; Maltamo *et al.*, 2006a, 2007; Bollandsås and Næsset, 2007; Breidenbach *et al.*, 2008a; Packalén and Maltamo, 2008; Peuhkurinen *et al.*, 2008). All in all, it has been found that ALS data are an excellent information source for forest attribute prediction but the accuracy and methodological approaches have varied a lot.

ALS is a data source that is relatively new to the research and user communities. Thus, there is still much to be explored in relation to the potentials and possibilities of such data. Recent applications have dealt with utilizing ALS data for predicting timber quality information (e.g. Peuhkurinen *et al.*, 2007, 2008; Korhonen *et al.*, 2008; Maltamo *et al.*, 2009b). An interesting tree attribute related to timber quality is crown height, usually understood as the height to the lowest green branch in the continuous tree crown. At tree level, both crown height and height to the lowest dead branch are used to determine the quality classes of sawn logs and the value of the timber (Uusitalo, 1995; Verkasalo *et al.*, 2004). In the case of Scots pine (*Pinus sylvestris*) logs, the length of the dead branch section is an excellent indicator of the quality and value of butt logs (Rikala, 2003; Maltamo *et al.*, 2009b). At stand level, the mean crown height is also a wood quality indicator (Wall *et al.*, 2004).

In addition to indicate timber quality, crown height is used for modelling of tree growth, yield and vitality. Characteristics such as crown ratio and crown length have been used in growth and yield models (e.g. Salminen *et al.*, 2005). These variables have also been used as indicators of tree health and vigour (Smith, 1986). Furthermore, in some canopy fuel models, crown height is one of the input parameters (e.g. Andersen *et al.*, 2005) and crown height is typically used in allometric biomass models for biomass components such as living and dead branches, stem bark and foliage (e.g. Marklund, 1988).

Even if it is an interesting biophysical property in many respects, crown heights are seldom measured in forest inventories neither on tree nor on stand level. This is because the gain of having better information is regarded as being less than the costs of measuring the crown height. In principle, the crown height can be predicted from other variables typically being measured in forest inventories (e.g. Ojansuu and Maltamo, 1995; Petterson, 1997; Uusitalo and Kivinen, 1998). It is, however, very difficult to predict crown height accurately enough using traditional tree and stand variable measurements such as tree diameter, height, stand basal area and age, without stand-level calibration. This is because there is considerably more variation in crown height between stands than within a stand, at least in the case of Scots pine (Mäkinen and Maltamo, 2001). It is also highly influenced by stand density, site fertility, genetic variation and the stand silvicultural history that is usually not known. There are also notable differences in crown height between tree species, e.g. due to the social status of a species in a stand and if the species is shade tolerant or intolerant. In multilayered stands, it can be a complicated task to define the mean crown height.

In many countries and in Scandinavia in particular, ALS is now being used as a primary data source in operational forest inventory. In Norway, where almost all stand-level inventories are performed using ALS (Næsset, 2009), mean crown height has not been used even though it is considered relevant information in terms of selecting among stands for final cutting according to specifications from individual saw mills. However, since ALS provides versatile data for deriving information of the vertical structure of the canopy, it is highly relevant to explore its utilization for derivation of crown height. If the accuracy of such information proves to be on an acceptable level, it should be highly relevant for the forest owners as they can be more specific on which wood qualities they can offer to saw mills.

ALS data have already been used to predict crown height characteristics both at tree and at stand levels (e.g. Næsset and Økland, 2002; Maltamo *et al.*, 2006b; Dean *et al.*, 2009). In studies where tree level crown height is considered, the individual trees have typically been identified from the ALS data first and the crown height have subsequently been derived for the delineated trees (e.g. Holmgren and Persson, 2004; Solberg *et al.*, 2006; Popescu and Zhao, 2008). However, even though ALS is a powerful data source related to canopy, crown height is more difficult to model than tree height. The reason is most likely that laser pulses do not penetrate evenly through the entire canopy.

So far, there have been three main approaches to determine crown height. First of all, crown height has been predicted by using characteristics of the ALS height distribution both at tree and at stand levels (Næsset and Økland, 2002; Andersen *et al.*, 2005; Maltamo *et al.*, 2006b; Popescu and Zhao, 2008; Vauhkonen, 2008). In those studies, height percentiles and other ALS-derived variables have usually been used as independent variables in regression models. Also non-parametric methods have been used (Maltamo *et al.*, 2009b). The accuracy of these model-based approaches has varied from 1.5 m to ~4 m in terms of root mean square error (RMSE). An exception is the study by Maltamo *et al.* (2009b) where nearest neighbour estimation led to even better accuracies than 1.0 m based on cross validation with the training data.

The second approach has been the direct analysis of the height distribution of the laser-derived heights. These methods do not usually need separate training data. Solberg *et al.* (2006) derived crown height on tree level as the height decile with the largest distance downward to its neighbouring decile. Morsdorf *et al.* (2004) used the largest difference between 5% height percentiles as a prediction on tree level. On plot level, Dean *et al.* (2009) used two-parameter truncated Weibull distributions to define crown height. First, they set a truncation point according to the properties of the height distribution. Then they applied a truncated Weibull distribution and used the estimated scale and shape parameters of the truncated Weibull in an original two-parameter Weibull distribution form, i.e. they expected the height distribution of the ALS pulses to approximate the tree canopy.

The third possibility is to utilize properties of the 3-D ALS point cloud to determine crown height at tree level (e.g. Holmgren and Persson, 2004). Different approaches here include polygon models (Pyysalo and Hyyppä, 2002), alpha shape techniques (Holmgren et al., 2008; Vauhkonen, 2008) and voxels (Popescu and Zhao, 2008). The accuracy of direct prediction methods based on properties of the 3-D point cloud has not been as good as in the case of statistical models. Furthermore, the predictions have also tended to be biased as they have been overestimates in most cases. There are also many applications where crown height estimates, which are based on analytical solutions, are used as independent variables in statistical models to support other ALS data variables (e.g. Popescu and Zhao, 2008; Vauhkonen, 2008; Maltamo et al., 2009b).

The aim of this study was to predict mean crown height in Norway spruce (*Picea abies*)-dominated stands. Different methods were compared, including (1) statistical modelling, (2) the use of truncation point derived from the ALS height distribution (cf. Dean *et al.*, 2009), (3) alpha shape techniques (cf. Vauhkonen, 2008) and (4) a method based on height differences between adjacent height deciles (percentiles) (cf. Solberg *et al.*, 2006). In this study, alpha shape predictions and height differences in adjacent height percentiles were applied for the first time at stand level. Furthermore, this was the first time truncation point predictions have been applied at the level of small sample plots. Those predictions were also used as candidate variables in the statistical models. For those methods that required calibration, we used data originating from field plots. The results of all methods were validated on (1) an independent plot-level dataset and (2) a dataset from harvester measurements of final cuttings of stands where crown height was measured on each harvested tree.

Material

Study area

This study was conducted based on data from Aurskog-Høland municipality (59°50′N 11°30′E, 120–390 m above sea level), southeastern Norway. The total area was 960 km^2. The dominant tree species in the area were Norway spruce (*P. abies* (L.) Karst.) and Scots pine (*P sylvestris* L.). The study took advantage of an operational stand-based forest inventory utilizing ALS data that were conducted at the time we were planning this research. Thus, in this study, we utilized aerial photographs, ALS data and field data collected for the operational inventory.

Photo-interpretation and forest stratification

Aerial photographs were acquired 28–29 June 2005 using a digital Vexcel UltraCam D camera. The pixel size on the ground was 25 cm. Stereo photogrammetry based on the aerial photographs was used to delineate the forest stands and to determine tree species, age and site productivity of each stand by manual photo-interpretation. For further details, please refer to Maltamo *et al.* (2009a).

Collection of modelling data

The field data used for modelling were collected within sample plots systematically distributed throughout the study area (systematic sampling design). The sample plot inventory was carried out during fall 2006 and it was also used in the operational inventory. Each plot had a fixed area of 400 m^2. On each plot, diameters at breast height (d.b.h.) ≥4 cm were callipered and tree heights and tree crown heights were measured on sample trees selected with a relascope. This yielded a sample proportional to tree size (tree basal area at d.b.h.). Tree crown height was defined as the lowest point in tree trunk where at least two living branches were found. The number of sample trees per plot ranged between 4 and 13 with an average of 9. The stem diameters were recorded in 2-cm diameter classes. Ground truth value for total plot volume (V), basal area (G) and Lorey's mean height (h_{Lorey}) were computed from the field measurements according to the procedures outlined by Næsset (2004a). These attributes were used when analysing and comparing datasets of this study. Only those plots that were dominated by Norway spruce (>50% volume) and located in mature stands according to the photo-interpretation were chosen for this study. The total number of plots in the operational inventory was 201 but after these restrictions, there were 54 plots left for analysis (Table 1). Even if spruce-dominated stands were selected, all other species in these stands were kept in the analyses. Plot mean crown heights were calculated as basal area weighted mean crown height based on the sample tree measurements. The reason for relying on a weighted mean was that this is more representative of the trees that populate the most valuable timber stock. The plot center coordinates (x, y) were determined by differential Global Navigation Satellite Systems (dGNSS), using dual-frequency receivers observing pseudorange and carrier phase of the Global Positioning System and Global Navigation Satellite System (Næsset, 2001, 2004a).

Collection of sample plot validation data

The sample plot validation data were collected in field during the fall of 2007 and winter 2008. The plots were distributed along five east-west going lines. Initially, a potential plot location was on every 1 km of these lines, but the actual location was guided by additional requirements of that 10 plots should be located in young forest and 30 plots should be located in mature forest. Furthermore, all plots should be fairly close to a road because heavy instruments had to be carried to the plot. If the pre-selected 1-km points on the lines did not satisfy these criteria, a new location was attempted within 300 m to either side. The plots were 1000 m^2 of size unless the plot was so densely populated that the required work exceeded 1 day. These plots were 500 m^2 of size and for this specific study, two of

Table 1: Mean and SD of stand characteristics of the study materials

	Modelling data		Plot-level validation data		Stand-level validation data	
	Mean	SD	Mean	SD	Mean	SD
Mean crown height, m	8.75	2.71	8.04	1.50	7.55	1.32
V, m^3 ha^{-1}	340.32	131.76	311.49	80.15	314.15	75.34
G, m^2 ha^{-1}	32.85	9.31	33.45	5.49	30.75	7.58
h_{Lorey}, m	20.86	3.82	18.56	3.34	18.54	2.24

these plots were used. The selection of plots was restricted to the same species (>50 Norway spruce volume) and age class (mature forest, lower age limit of ~45 years) as the modelling data. Thus, 13 sample plots were used in the validation. On each plot, d.b.h. was callipered starting in the north direction relative to the plot center and going clockwise. Tree heights were measured on sample trees. For the main tree species, 10 sample trees were selected after the following scheme: first, the diameter range was divided into five classes. Then, two sample trees were selected for each diameter class. These trees were the first trees callipered in each class. For each of two secondary tree species, five sample trees were selected with probability proportional to stem basal area. To have five trees per species per plot, the sampling quotient was adjusted according to an initial relascope measurement of the species-wise basal areas. The total number of sample trees per plot ranged from 11 to 20 with an average of 17. The plot center coordinates were determined using dGNSS according to the procedure outlined above. Polar tree coordinates were then registered using a total station. Stand attributes were calculated correspondingly to the modelling data plots (Table 1).

Three different crown height definitions were applied and corresponding measurements taken during the fieldwork. These definitions were (1) the lowest living branch, (2) the lowest branch whorl where half of the branches were living and finally (3) the lowest branch whorl where three-fourth of the branches were living. Because all these definitions were different from the one used in the modelling, an effort was carried out to adjust the crown height of the validation dataset to match the definition of the modelling data. We compared mean plot-level crown heights at the same Lorey's mean height range between the modelling data and the validation data. We found that the two lowest definitions of crown height of the validation data were 1.47 m higher and 1.01 m lower, respectively, than the modelling data. Thus, using either of these crown height definitions from the validation data would lead to biased results. Therefore, we calculated the average of the two lowest measurements of crown height in the plot-level validation data to approximate the definition of the modelling data. The average difference between the crown heights of the modelling data and the validation data after this adjustment was only 0.23 m. Based on these calculations, the average value of the two lowest crown height measurements of the plot-level validation data was selected to be used in the subsequent analysis.

Collection of stand-level validation data

The stand-level validation data were collected during winter and fall in 2007. The data originated from a clearfelling operation of 29 different forest stands over a total area of 40 ha. However, we only used data from stands restricted to the same species (>50% Norway spruce volume) and age class (mature forest) as the modelling data. Thus, 24 harvested stands were used in the validation.

The data were collected using John Deere 1070D harvesters with Timberjack/John Deere H754 harvester heads. The bucking software was the Timbermatic 300, version 2.4.9. The harvesters recorded length and diameter of each tree for every 10 cm section along the stems. Volume of each stem was calculated section wise and summed. The top of the trees did not pass through the harvester head and was thus not measured. We therefore estimated the length of the top based on a function dependent on the last diameter measured by the harvester. The function was calibrated on observed top length data from the same stands as the felling operation took place. The volumes of the tops were calculated under the assumption that their shapes were strictly conical. The operators of the harvesters recorded the crown height from each cut tree manually. This was done by visually determining the crown height point on the log as the log passed through the harvester head. When the crown height point reached the harvester head, the operator registered this into the computer. The definition of crown height was the same as for the modelling data. Each tree was attributed to a delineated stand, i.e. no tree-level coordinates were available. We therefore calculated the mean crown height for each of the harvested stands based on all the trees measured in each specific stand. Since every tree in each stand was measured, the crown height of each tree was weighted with basal area to derive weighted mean stand values (Table 1).

Laser scanner data

Laser data for this study were acquired under leaf-on conditions on 8–10 June 2006. Additional data were acquired on 6 September 2006 to fill in a minor gap in the data acquired in June. A fixed-wing piper PA31-310 aircraft was used in the acquisition. Laser data were collected with an Optech ALTM 3100 laser scanner operated from an altitude of ~1850 m above ground level. The flight speed was 75 ms^{-1} and the pulse repetition frequency was 50 kHz. The scan frequency was 71 Hz, resulting in a point density on the ground of ~0.7 m^{-2}. The maximum scan angle was 15° but pulses emitted at an angle >13° were discarded during the subsequent data processing.

An initial processing of the data was accomplished by the contractor (Blom Geomatics, Oslo, Norway). Planimetric coordinates (x and y) and ellipsoidal height values were computed for all echoes. The entire flight campaign was flown as a block of parallel strips with a swath overlap of 15%. Three flight lines were flown orthogonal to the other flight lines and used in matching and correction for systematic errors between swaths. Ground echoes were found and classified using the progressive Triangular Irregular Network (TIN) densification algorithm (Axelsson (2000) of the TerraScan software; Anonymous (2005)). A TIN model was created from the planimetric coordinates and corresponding heights of the laser echoes classified as ground points. The heights above the ground surface were calculated for all echoes by subtracting the respective TIN heights from the height values of all echoes recorded.

The laser data for the 1000-m^2 validation plots were separated into four equally sized quadrants of 250 m^2 each. The 500-m^2 plots were split into two. Each quadrant (or half) was treated as one observation so that the plot sizes of the training and validation data would better correspond. In reliability calculations, mean crown height predictions for these quadrants were combined to derive pooled estimates across all quadrants of each plot. Correspondingly, in the case of the harvester data, the area of each stand was divided into grid cells of size 400 m^2. However, because cells could be very small and irregular in shape due to the intercepting stand borders (sliver polygons), we required that they should include a minimum of 10 laser echoes, otherwise they were excluded from the data. Low numbers of echoes will tend to increase the variability of the metrics derived from the ALS data (Gobakken and Næsset, 2008), and 10 echoes were thus set as a lower limit. Corresponding to the plot-level reliability calculations, mean crown height predictions for these cells were combined to derive pooled estimates across all grid cells of a stand.

The ALTM 3100 sensor is capable of recording up to four echoes per pulse. In this study, we used the two echo categories classified as 'first of many' and 'single'. Echoes of these two categories were pooled into one dataset, and this aggregated dataset was denoted as 'first' echoes. The height distributions created from the first echoes were used to calculate plot- and cell-level percentiles for 5, 10, 15, . . . , 90, 95 and 100% of the heights (H_5, H_{10}, H_{15}, . . . , H_{90}, H_{95} and H_{100}) (see Næsset, 2002) of the three respective datasets. The height distributions contained only those laser points that were classified as above-ground echoes, using a threshold value of 2 m. H_5, for example, denotes the height at which the accumulation of laser echoes in the vegetation is 5%.

Methods

Modelling alternatives

The different methods to predict mean crown height in this study are listed in Table 2. We applied both direct analysis- and regression analysis-based methods. Most of the analysis was done at the plot level.

Alpha shape technique (A)

Vauhkonen (2008) developed several approaches for extracting information on the 3-D arrangement of a point cloud and using this information for predicting the crown height for individual trees. Most of the approaches of Vauhkonen (2008) require determination of point distances or cross-sectional areas that are, indeed, measured on the scale of individual trees. However, the approach in which connected alpha shape components extracted from the lowest parts of the point cloud are examined ('comp approach'; Vauhkonen, 2008) can be extended to a level of tree groups also. Here, modelling alternative A is a direct application of that approach. Therefore, the approach itself is described only briefly in this context. Rather, the focus is on how the method was applied on the plot level.

The comp approach is an application of computational point geometry, namely the concept of 3-D alpha shapes (Edelsbrunner and Mücke, 1994). An alpha shape is derived from the Delaunay triangulation (see, e.g. Preparata and Shamos, 1985) of a point cloud such that each simplex of the triangulation is compared with the specified alpha value in the computation phase. Those simplices which have an empty circumsphere with a squared radius larger than the defined alpha value are removed. Thus, an alpha shape can be regarded as an alpha-weighted Delaunay triangulation. The alpha value defines the level of detail in the obtained shape and also determines whether the resulting shape is formed from a solid structure or from cavities, holes or even separate components (see Edelsbrunner and Mücke, 1994).

When predicting crown height with the comp approach, an alpha value with one connected component is used as a starting point. The alpha values are traversed in descending order until a new component is split or the minimum height value of the highest component is changed. The first split component is allowed to partly overlap the previous, but otherwise the removal is accepted only if the component is located below the current highest component. If not, the procedure is stopped and the crown height is defined as the height of the lowest vertex in the obtained structure. The computations regarding the alpha shapes were carried out using the functionality of the Open Source library Computational Geometry Algorithms Library (Da and Yvinec, 2007).

Table 2: Different methods to predict mean crown height

Alternative	Prediction approach	Prediction method	Level of the analysis
A	Direct analysis	Alpha shape technique based on the 3-D ALS data	Plot
B	Direct analysis	Truncation point of ALS height distribution	Plot
C	Direct analysis	Truncation point of ALS height distribution	Stand
D	Direct analysis	The point of largest difference in adjacent height percentiles	Plot
E	Direct analysis	The point of the largest difference in adjacent height percentiles	Stand
F	Regression analysis	Modelling based on the variables of the ALS height distribution (ALS height)	Plot
G	Regression analysis	Modelling based on alpha shape technique and ALS height variables	Plot
H	Regression analysis	Modelling based on truncation point and ALS height variables	Plot
I	Regression analysis	Modelling based on the point of the largest difference in adjacent height percentiles and ALS height variables	Plot

In order to take the computation closer to the tree level, vector grids with varying cell sizes were established over the extent of the tested area and the analysis was carried out with the point data of each cell. In other words, each cell was first provided with a crown height estimate, and the plot-level mean crown height value was then aggregated from the per-cell estimates. Cell sizes of 3 × 3, 4 × 4, 5 × 5, 7.5 × 7.5 and 10 × 10 m were tested. The crown height was determined for each cell using the method described above with a restriction that a cell was supposed to have at least five laser echoes with a height value higher than a ground threshold set to 1 m. This limit was set due to the fact that at least four points are needed to generate a single tetrahedron in the comp analysis. The 3 × 3 m grid was ruled out due to this restriction: most cells of this resolution were lacking points for the computation. The mean crown height to the unit considered in the subsequent analysis was calculated as a weighted average, i.e.:

$$\mathrm{CBH}_u = \frac{\sum_{g=1}^{N} A_g \times \mathrm{CBH}_g}{\sum_{g=1}^{N} A_g}, \qquad (1)$$

where CBH_u is the mean crown height prediction for the analysed unit, A_g the area of the cell g within the analysed unit, CBH_g the crown height prediction for the corresponding cell g and N the number of cells g with the estimated crown height >0 and area overlap >50% with the analysed unit. The analysed unit refers either to the plots or to the validation grid cells.

Analysis of truncation point of ALS height distribution (B and C)

In this case, mean crown height was determined in close analogy to what is referred to as a truncation point in Dean *et al.* (2009). They use the truncation point to select the portion of the vegetation height distribution to which they then fitted a truncated Weibull distribution. However, since we needed to deal with considerably smaller units (i.e. sample plots), some adjustments were necessary and, as opposed to Dean *et al.* (2009), the height distribution of the ALS echoes was left in its original form. That is, no transformation to relative values or reversal of values was applied. Vegetation heights of the analysed unit (plot, grid cell and stand) were divided into equal sized vegetation height classes (bins). The number of bins was calculated according to Sturge's formula (Venables and Ripley, 2002) given by ($\log_2 n$ + 1) where n was the amount of echoes within a sample plot. The bin width b was given by range(x)/($\log_2 n$ + 1) with x as the vector of individual echo vegetation heights. The (relative) density of each bin was computed as n_i/n b where n_i was the number of echoes within one of the bins. A vertical bar plot of the relative density over the bin mid positions is referred to as a histogram.

The mode of the binned distribution was determined by ignoring the first two bins that usually revealed highest densities due to large amounts of ground echoes. A series of regression models were fitted to four consecutive relative densities downwards vs the respective bin mid points. The series began with the mode as the highest density and was moved one bin downwards as long as the slope of the regression model did not fall below the limit l (Dean *et al.*, 2009). The truncation point 'trunc' was set to the lower value defining the interval of the highest bin, if the slope fell below the limit or if less than three bins were available for the regression.

The slope limit was heuristically set to a value of 0.05 after examining many histograms. However, if trunc was the lower value of the mode bin (i.e. already the first slope is smaller than l), the procedure was repeated with l set to 0.

Since the method based on truncation points does not directly rely on any area-based statistics, it was possible to calculate a prediction on plot level for modelling data and validation data (B) as well as for grid cells (with same area as sample plots) and for entire stand data without grids using harvester data on stand level (alternative C).

Differences in adjacent height percentiles (D, E)

Crown heights were also predicted by using height differences in adjacent height percentiles (perc). In this case, we detected the largest difference between height percentiles and predicted mean crown height as being the height of the upper one of the percentiles in question (Solberg *et al.*, 2006). However, in our study, we examined height differences in 5% classes, whereas Solberg *et al.* (2006) used 10% classes at tree level. It should also be noted that this method does not need any calibration data. Each value is calculated independently without using information or adjusted parameters from other datasets. Although these predictions were calculated on plot/grid level (D), it was also possible to predict directly from stand-level height percentiles in the case of harvester data (alternative E).

Regression analysis (F–I)

Our final approach was to develop regression models by using stepwise selection of independent variables based on the modelling data (F). We used height percentiles and other ALS-derived variables to predict mean crown height. Transformations of the dependent and independent variables were also calculated to increase the fit of the model. Regression analysis was also applied so that mean crown height predictions derived directly from the properties of ALS data (A, B and D) were used as independent variables in regression models (G, H and I).

Accuracy assessment

The results were validated in terms of RMSE and bias:

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

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

where n is the number of plots or stands, y_i is the observed mean crown height value for plot or stand i and $\hat{y}$ is the predicted mean crown height value for plot or stand i. Additionally, the statistical significance of the bias was tested using Student's t-test.

Results

Predicted plot-level regression models for mean crown height are presented in Table 3. In all models, there were a maximum of two variables only. The highest adjusted R^2 value was obtained for the regression model based on height percentiles (F). Furthermore, the height percentiles used in the models were always below 50%. The relationships between mean crown height and ALS information were found to be linear, i.e. there was no need to use transformations of the independent or dependent variables. Although transformations were selected to be included in the 'best' model in alternative F, the differences in model fit were minor between models with transformed variables and variables on arithmetic scale also in this case. In all alpha shape calculations (A, G), a cell size of 4 × 4 m was used. No height percentiles were able to improve the model for alternative I, which is based on, in fact, a height percentile (D) that varies between different plots. The same was true also for alternative H.

The reliability figures of all methods (A–I) are presented in Table 4. All methods were applied to all three datasets, i.e. the modelling data and two separate and independent validation dataset. The results obtained using the modelling data indicate that the regression model including height percentiles as independent variables (F) was the most accurate alternative. The model error values of the different methods varied between 1.1 and 2.6 m. All methods except the truncation point method (B) were also unbiased when the modelling data were used for validation.

The accuracy seemed to be better in the independent plot-level validation data than using the modelling data. In most cases, the absolute errors in the plot validation data were smaller than 1.5 m. The performance of the method based on truncation point (B) was also unbiased in this dataset. In the case of stand-level harvester data, the main observation of the model validation was that all methods overestimated the crown heights (Table 4). These biases were also statistically significant. Also, the RMSE values from the validation on the harvester data were larger compared with the validation on the modelling data and plot-level validation data. The performance of the alpha shape technique (A) was slightly poorer in this dataset as compared with the plot-level validation dataset although the reliability figures were among the most accurate ones. The performance of the methods based on adjacent height percentiles (D, E, I) was twofold. First, if the predictions were calculated directly at stand level (E), the accuracy was

Table 3: The regression coefficients and reliability figures for the regression models

	Independent variables							
Alternative	Intercept	h_5	$1/(h_5)$	$\ln(h_{40})$	comp	trunc	perc	$\hat{sd}(e)$	Adjusted R^2
F	0.103		−2.259	1.223				0.140	0.821
G	1.087	0.395			0.576			1.387	0.738
H	3.038					0.580		1.964	0.475
I	2.122						0.697	1.629	0.639

The dependent variable is the mean crown height on arithmetic (original) scale except for alternative F where the square root transformation has been applied. $\hat{sd}(e)$ and adjusted R^2 are the standard error and adjusted coefficient of determination of the models, respectively.

Table 4: The performance of different alternatives to predict mean crown height using the modelling dataset and the two independent validation datasets for validation

Prediction alternative	Plot-level modelling data		Plot-level validation data		Stand-level validation data	
	RMSE, m	Bias, m	RMSE, m	Bias, m	RMSE, m	Bias, m
A	1.563	−0.097	1.315	0.273	1.799	−1.154 **
B	2.599	−1.097 *	1.389	−0.029	3.134	−2.834 ***
C	–	–	–	–	2.977	−1.549 **
D	2.003	−0.759	1.553	0.038	3.608	−3.330 ***
E	–	–	–	–	1.791	−0.940 **
F	1.134	0.015	1.071	0.125	1.810	−1.434 ***
G	1.348	−0.004	1.214	0.163	1.683	−1.202 **
H	1.964	−0.000	1.192	0.359	1.887	−1.510 ***
I	1.598	−0.000	1.252	0.367	2.482	−2.170 ***

Statistical significance of t-test: ***P ($T < t$) <0.001, **0.001 < P ($T < t$) < 0.01, *0.01 < P ($T < t$) < 0.05. Performance expressed in terms of RMSE and bias.

among the best ones. Second, grid-based predictions (D and I) were more biased than other methods and also the RMSE figures were higher. As opposed to this, the most accurate alternative among the methods based on the truncation points (B, C and H) was the one based on regression analysis (H). However, the reliability figures (especially bias) were slightly worse compared with the alpha shape technique (A and G) and regression model based on height percentiles (F).

Finally, the performance of some of the most accurate methods (A, E and F) was examined by distributing the validation results from the harvester data according to stand size (Figure 1). The figure shows that in almost every stand, all presented methods led to overestimates. There are no clear trends between methods or according to stand size. Still, it can be seen that the largest absolute differences (>2 m) occur in small stands. In those stands, most of the grid cells are smaller than 400 m^2, which may have effected the predictions at least in the case of regression models (F) and the alpha shape predictions (A).

Discussion and conclusions

In this study, different methods were used to predict mean crown height of Norway spruce-dominated stands by means of ALS data. The methods were based on statistical modelling (F), properties of the laser point clouds (A–E) or combinations of them (G–I). The different methods were compared and separate modelling data and two different validation datasets were used. The results obtained were partly contradictory, showing varying performance of different methods in different datasets. There were also notable differences in the crown height values obtained according to the three different field methods applied.

The properties of the three datasets of this study varied considerably. Both the modelling data and the plot-level validation data consisted of sample tree measurements. There were, however, differences in crown height definitions between these datasets. In the modelling data, crown height was defined as the height to the lowest whorl with at least two living branches, whereas in the plot-level validation data, it was defined in three different ways. To make the measured crown heights comparable between datasets, we used an average of the measurements made according to the two lowest definitions, i.e. the height of lowest living branch and the lowest point where half of the branches were alive in the same whorl. Using the average of the two measurements of the plot-level validation data avoided large biases, which obviously would have appeared by using either of them individually.

In the case of the harvester data, the definition of tree crown height was similar to the one used in modelling data. However, there was over 1-m difference in field average between these datasets (Table 1) and this may have contributed to at least some of the considerable bias inherent in the validation results using the harvester data. Of course, there can be differences in datasets but it should be noted that also the method based on differences between adjacent height percentiles (D and E), which does not at all rely on modelling data, yielded biased results when it was validated with the harvester data. This is an indication of bias in the harvester measurements themselves since the ALS data were all the same in all the three datasets. There are at least two possible reasons for differences in crown heights between these two datasets. Firstly, in the harvester data,

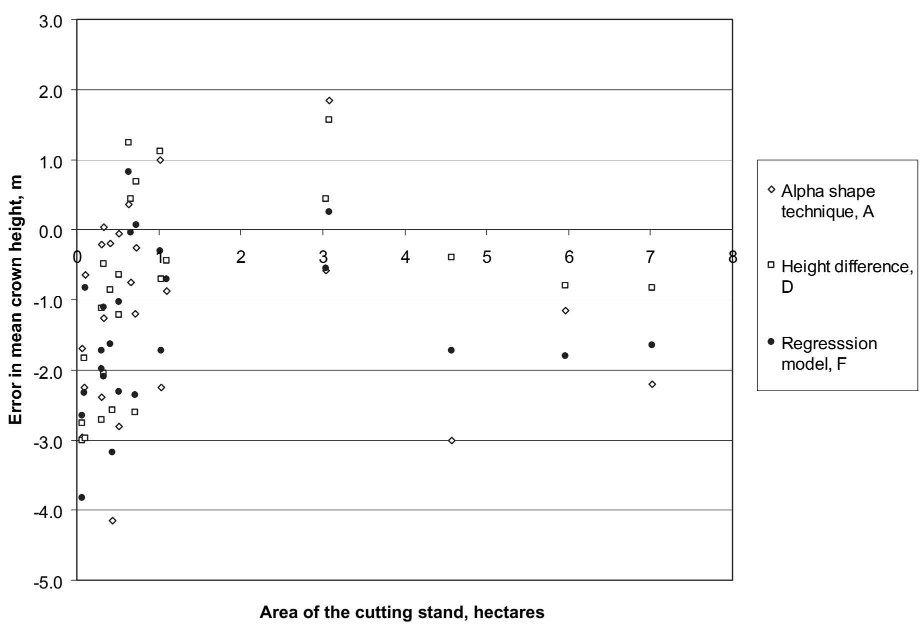

Figure 1. The relationship between stand area and prediction error of mean crown height of three different methods including the alpha shape technique (A), height difference in adjacent height percentiles (E) and regression analysis based on height percentiles (F).

the crown heights of all harvested trees were measured, whereas in the modelling data, crown heights were only measured on trees selected with probability proportional to basal area. Thus, the bias we observed could be due to sampling errors. Secondly, there can also be errors related to the recording of the actual crown height by the operator of the harvester. From the position where the operator observes the stem, it can be hard to see where the crown height starts, especially if there are a lot of branches, intermixing of dead and living branches or if some branches are on the far side of the tree.

The field measurement of tree crown height seems to be delicate process due to the many possible error sources. One aspect not considered here are the differences in crown height between tree species. Since we considered mean crown height in stands dominated by Norway spruce (>50% spruce), it is possible that in certain plots/stands comprising other tree species, differences between species may have had a considerable effect on derived mean values. However, it might be that the influence of tree species on the ALS data characteristics is larger than the effect on the mean field attributes. Finally, Figure 2 indicates the relationship between mean crown height and stand density in terms of basal area in different datasets. The figure shows that the variation in the modelling data was considerably larger than that in the two validation datasets. Additionally, in both plot-level data, the relationships between mean crown height and basal area were positive, i.e. mean crown height increases with increasing basal area. At stand level, the relationship between these attributes was rather stable.

The accuracy of different ALS-based methods varied considerably. The regression model that was based on characteristics of the canopy height distribution (F) was, however, quite accurate in all datasets. The bias was mainly due to differences in crown height definitions between datasets. The addition of mean crown height predictions (alternatives G, H and I) derived directly from the ALS point cloud could not considerably improve the accuracy. Similar regression models for crown height characteristics either at tree or at plot level have been estimated in various previous studies as well (Næsset and Økland, 2002; Andersen *et al.*, 2005; Maltamo *et al.*, 2006b, 2009b; Popescu and Zhao, 2008; Vauhkonen, 2008). The accuracy figures of crown height predictions in the current study were found to be at a similar level as in earlier studies, although for some of the methods we tested, there were also quite low RMSE values between 1.0 and 1.5 m. RMSE values that small have not been reported in previous studies based on regression methods. It seems, however, that in crown height prediction, there is always an error of at least 1 m or so which seems to be inevitable when using ALS information. A likely reason for this 'minimum error level' is the fact that there is no 'foliage carpet' at the crown base that could be a definite lower level of a tree crown and thus be explicitly detected by an airborne laser.

Different techniques based on using alpha shapes have previously been applied successfully at tree level for the prediction of crown height (Holmgren *et al.*, 2008; Vauhkonen, 2008). In the present study, an approach (A) corresponding to that of Vauhkonen (2008) was used, but on a different and larger unit than the individual tree level addressed by Vauhkonen (2008). However, the grid-based approach performed well, especially when it was tested with modelling and plot-level validation datasets. A grid size assumed to be near to the actual tree crown size (4 × 4 m) was found to be best. The performance was slightly poorer using the stand-level validation data, but this was common to all the tested methods. In this study, the mean crown height values were basal area weighted, but still rather unbiased. Thus, with unweighted mean crown height, the method should lead to an overestimation, which is in line

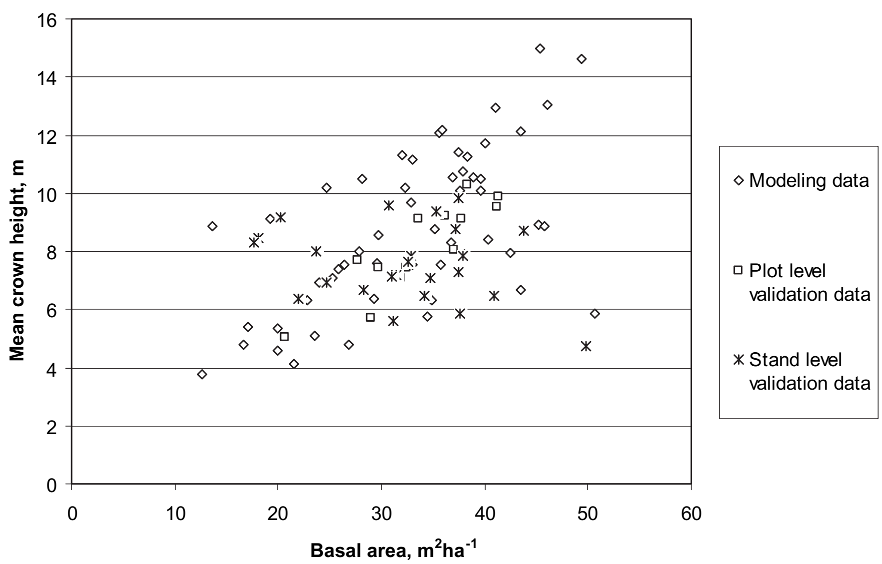

Figure 2. The relationship between stand density in terms of basal area and field measurements of mean crown heights.

with the result by Vauhkonen (2008), who ended up with overestimating the tree-level crown height by around a meter. Finally, it should be noted that the function of this method is based on the assumptions that (a) there are gaps between the canopy and the terrain level and (b) those gaps can be detected from the vertical height profile recorded by the ALS system. Therefore, the existence of high undergrowth would likely affect the predictions by adding bias. Also, although the method should be sensitive to the laser pulse density (Vauhkonen, 2008), only relatively sparse data have been used with this method so far. All in all, it is possible that the application of the alpha shape technique may require data-specific calibration.

The method based on differences between adjacent height percentiles (D and E) applied previously at tree level by Solberg *et al.* (2006) led to rather good accuracy levels in all datasets. However, it was required for this method to perform well that it was applied at full plot or stand level. In the grid-based calculations of the harvester data, the predictions were highly inaccurate. An obvious reason for this is that in the grid calculation, there are always border cells where the area is reduced and this method seems to be extremely sensitive for low number of height observations. However, the accuracy was not improved very much even when all cells smaller than 100 m^2 in size were removed, indicating that there is also a price to pay by removal of cells. Especially, in small stands, most of the grid cells can be smaller than the full size and methods for crown height derivation sensitive to cell size can therefore have limited applicability. For such methods, it seems more relevant to apply the laser point cloud of the entire stand as one unit instead. All in all, this method was found to be very easy to apply and it led at least to very promising initial values of mean crown height.

The performance of methods based on derived truncation point of the ALS height distribution (B and C) led to good accuracy for the plot-level validation data but was otherwise not as accurate as the other methods relying on properties of the ALS data. Secondly, this method yielded rather good predictions for the stand-level validation data when they were applied in regression models (H). It seems that this method potentially also can give good results but the accuracy is varying more than in the case of the rather robust method based on differences between adjacent height percentiles. Unlike in the study of Dean *et al.* (2009), the truncation point was used directly as a surrogate and predictor for crown height for several reasons, i.e. (i) the (truncated) vegetation height distribution does not always seem to follow a Weibull distribution, (ii) especially when using small sample plots (as in this study), too few echoes may be available to support the fit of a distribution function, and (iii) the truncation point and the point of the Weibull density with a small probability (e.g. <0.00001) will be very similar and highly correlated.

Another interesting aspect of this study was that the accuracy of the crown height predictions was not improved when moving from plot level to stand level. Previous studies have indicated that moving from a smaller unit to a larger one (tree–plot–stand) has tended to increase the accuracy of at least some of the derived forest attributes due to the averaging effect (see discussion in Korpela and Tokola, 2006), although the magnitude of this effect seems to vary from one attribute to another (Næsset, 2002, 2004a). The performance of mean crown height may not follow the same pattern since in our study, more accurate stand-level results were obtained only with the method based on height differences in adjacent height percentiles (E). Our plot-level predictions were, however, more accurate than most of the earlier ones obtained at tree level (Solberg *et al.*, 2006, Popescu and Zhao 2008, Vauhkonen 2008). It must be remembered though that we were not able to compare plot- and stand-level predictions by using the same data.

From a practitioner's point of view, we believe that ALS can provide useful information about crown height when selecting stands to be harvested. As an indicator of timber quality, a predicted crown height with an accuracy of, say, 1–2 m can be very informative. In the planning of harvester operations, even biased estimates of crown height can be useful if the objective is to prioritize between similar stands.

Thus, to conclude, this study has indicated that statistical modelling (F) is a good alternative if the relationship between mean crown height and ALS information is corresponding in modelling data and in application phase. The method relying on alpha shape technique (A) is also an accurate alternative but it probably requires at least a small training dataset to verify the result. Direct methods, such as the method based on height differences in adjacent height percentiles (D and E) applied in this study, are good alternatives to get relatively accurate results but there are still drawbacks, such as a minimum reliable calculation area, in their applicability.

Funding

Research Council of Norway (project #173911/I10: 'Assessing wood properties of forest resources'); Forest Trust Fund.

Conflict of Interest Statement

None declared.

Acknowledgements

We wish to thank Viken Skog BA for giving us access to the field data and Blom Geomatics AS for collection and processing of the airborne laser scanner data. We are also grateful to Mr Vegard Lien who organized the collection of the harvester data and took part in the fieldwork on the validation plots in collaboration with Mr Marius Hauglin and Mr John Gunnar Dokk.

References

Andersen, H.E., McGaughey, R.J. and Reutebuch, S.E. 2005 Estimating forest canopy fuel parameters using lidar data. *Remote Sens. Environ.* **94**, 441–449.

Anonymous. 2005 *TerraScan User's Guide*. Terrasolid Ltd, Jyvaskyla, Finland, p. 169. http://www.terrasolid.fi (accessed on 2 October, 2006).

Axelsson, P. 2000 DEM generation from laser scanner data using adaptive TIN models. *Int. Arch. Photogramm. Remote Sens.* **33**, 110–117.

Bollandsås, O.M. and Næsset, E. 2007 Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. *Scan. J. For. Res.* **22**, 33–47.

Breidenbach, J., Gläser, C. and Schmidt, M. 2008a Estimation of diameter distributions by means of airborne laser scanner data. *Can. J. For. Res.* **38**, 1611–1620.

Breidenbach, J., Koch, B., Kaendler, G. and Kleusberg, A. 2008b Quantifying the influence of slope, aspect, crown shape and stem density on the estimation of tree height at plot level using Lidar and InSAR data. *Int. J. Remote Sens.* **29**, 1511–1536.

Breidenbach, J., Kublin, E., McGaughey, R., Andersen, H.-E. and Reutebuch, S. 2008c Mixed-effects models for estimating stand volume by means of small footprint airborne laser scanner data. *Photogramm. J. Finl.* **21**, 4–15.

Da, T.K.F. and Yvinec, M. 2007 *3D Alpha Shapes. CGAL User and Reference Manual*. 3.3 edn. CGAL editorial board (ed.). http://www.cgal.org/Manual/3.3/doc_html/cgal_manual/packages.html#Pkg:AlphaShapes3 (accessed on 17 May, 2009).

Dean, T.J., Cao, Q.V., Roberts, S.D. and Evans, D.L. 2009 Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand. *For. Ecol. Manage.* **257**, 126–133.

Edelsbrunner, H. and Mücke, E.P. 1994 Three-dimensional alpha shapes. *ACM Trans. Graph.* **13**, 43–72.

Gobakken, T. and Næsset, E. 2004 Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scand. J. For. Res.* **19**, 529–542.

Gobakken, T. and Næsset, E. 2005 Weibull and percentile models for LIDAR-based estimation of basal area distribution. *Scand. J. For. Res.* **20**, 490–502.

Gobakken, T. and Næsset, E. 2008 Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* **38**, 1095–1109.

Holmgren, J. 2004 Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand. J. For. Res.* **19**, 543–553.

Holmgren, J. and Persson, Å. 2004 Identifying species of individual trees using airborne laser scanner. *Remote Sens. Environ.* **90**, 415–423.

Holmgren, J., Persson, Å. and Söderman, U. 2008 Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int. J. Remote Sens.* **29**, 1537–1552.

Korhonen, L., Peuhkurinen, J., Malinen, J., Maltamo, M., Suvanto, A. and Packalén, P. *et al.* 2008 The use of airborne laser scanning to estimate sawlog volumes. *Forestry*. **81**, 499–510.

Korpela, I. and Tokola, T. 2006 Potential of aerial image-based monoscopic and multiview single-tree forest inventory – a simulation approach. *For. Sci.* **52**, 136–147.

Magnussen, S. and Boudewyn, P. 1998 Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Can. J. For. Res.* **28**, 1016–1031.

Mäkinen, H. and Maltamo, M. 2001 Puun laatua kuvaavat mallit. [Models describing tree quality]. In *Metsää kuvaavat mallit (In Finnish)*. M. Maltamo and S. Laukkanen (eds). Silva Carelica, University of Joensuu, Faculty of Forestry, Joensuu, Finland, **36**, pp. 111–133.

Maltamo, M., Mustonen, K., Hyyppä, J., Pitkänen, J. and Yu, X. 2004 The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. *Can. J. For. Res.* **34**, 1791–1801.

Maltamo, M., Eerikäinen, K., Packalén, P. and Hyyppä, J. 2006a Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry*. **79**, 217–229.

Maltamo, M., Hyyppä, J. and Malinen, J. 2006b A comparative study of the use of laser scanner data and field measurements in the prediction of crown height in boreal forests. *Scand. J. For. Res.* **21**, 231–238.

Maltamo, M., Suvanto, A. and Packalén, P. 2007 Comparison of basal area and stem frequency diameter distribution modelling using airborne laser scanner data and calibration estimation. *For. Ecol. Manage.* **247**, 26–34.

Maltamo, M., Næsset, E., Bollandsås, O.M., Gobakken, T. and Packalén, P. 2009a Non-parametric prediction of diameter distributions using airborne laser scanner data. *Scand. J. For. Res.* **24**, 541–553.

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P. and Tokola, T. 2009b Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. *Silva Fenn.* **43**, 507–521.

Marklund, L.G. 1988 *Biomassafunktioner för tall, gran och björk i Sverige. [Biomass Functions for Pine, Spruce, and Birch in Sweden] Rapport 45*. Institutionen för skogtaxering, Sveriges Lantbruksuniversitet, Umeå. (In Swedish with English summary).

Morsdorf, F., Meier, E., Kötz, B., Itten, K.I., Dobbertin, M. and Allgöwer, B. 2004 LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sens. Environ.* **92**, 353–362.

Næsset, E. 1997 Determination of mean tree height of forest stands using an airborne laser scanner data. *ISPRS J. Photogramm. Remote Sensing*. **52**, 49–56.

Næsset, E. 2001 Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. *Photogramm. Eng. Remote Sensing*. **67**, 1021–1026.

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

Næsset, E. 2004a Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* **19**, 164–179.

Næsset, E. 2004b Accuracy of forest inventory using airborne laser-scanning: evaluating the first Nordic full-scale operational project. *Scand. J. For. Res.* **19**, 554–557.

Næsset, E. 2009 Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* **113**, 148–159.

Næsset, E. and Gobakken, T. 2008 Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sens. Environ.* **112**, 3079–3090.

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

Ojansuu, R. and Maltamo, M. 1995 Sapwood and heartwood taper in Scots pine stems. *Can. J. For. Res.* **25**, 1928–1943.

Packalén, P. and Maltamo, M. 2008 The estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Can. J. For. Res.* **38**, 1750–1760.

Persson, Å., Holmgren, J. and Söderman, U. 2002 Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sensing*. **68**, 925–932.

Petterson, H. 1997 Functions for predicting crown height of *Pinus sylvestris* and *Picea abies* in Sweden. *Scand. J. For. Res.* **12**, 179–188.

Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J. and Packalén, P. 2007 Preharvest measurement of marked stand using airborne laser scanning. *For. Sci.* **53**, 653–661.

Peuhkurinen, J., Maltamo, M. and Malinen, J. 2008 Estimating species-specific height-diameter distributions and saw log recoveries from ALS data and aerial photographs: a distribution-based approach. *Silva Fenn.* **42**, 625–641.

Popescu, S.C. and Zhao, K. 2008 A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sens. Environ.* **112**, 767–781.

Popescu, S.C., Wynne, R.H. and Nelson, R.F. 2002 Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Comput. Electron. Agric.* **37**, 71–95.

Preparata, F.P. and Shamos, M.I. 1985 *Computational Geometry: An Introduction*. Springer-Verlag, New York.

Pyysalo, U. and Hyyppä, H. 2002 Reconstructing tree crowns from laser scanner data for feature extraction. *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.,* **XXXIV**, Part 3B, 218–221.

Rikala, J. 2003 *Spruce and Pine on Drained Peatlands – Wood Quality and 1 Suitability for the Sawmill Industry*. University of Helsinki, Department of forest resource management, Publications, **35**, 147 pp.

Salminen, H., Lehtonen, M. and Hynynen, J. 2005 Reusing legacy FORTRAN in the MOTTI growth and yield simulator. *Comput. Electron. Agric.* **49**, 103–113.

Smith, D.M. 1986 *The Practice of Silviculture*. 8th edn. John Wiley and Sons, Inc., New York, 527 pp.

Solberg, S., Næsset, E. and Bollandsås, O.M. 2006 Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. *Photogramm. Eng. Remote Sensing*. **72**, 1369–1378.

Takahashi, T., Yamamoto, K., Senda, Y. and Tsuzuku, M. 2005 Predicting individual stem volumes of sugi (*Cryptomeria japonica* D. Don) plantations in mountainous areas using small-footprint airborne LiDAR. *J. For. Res.* **10**, 305–312.

Thomas, V., Treitz, P., McCaughey, J.H. and Morrison, I. 2006 Mapping stand-level forest biophysical variables for a mixed-wood boreal forest using lidar: an examination of scanning density. *Can. J. For. Res.* **36**, 34–47.

Uusitalo, J. 1995 *Pre-harvest Measurement of Pine Stands for Sawing Production Planning*. University of Helsinki, Department of Forest Resource Management, Publications 9.

Uusitalo, J. and Kivinen, V.-P. 1998 Constructing bivariate dbh/dead-branch distribution of pines for use in sawing production planning. *Scand. J. For. Res.* **3**, 509–514.

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

Vauhkonen, J. 2008 Estimating crown base height for Scots pine by means of the 3D geometry of airborne laser scanning data. In *Proceedings of SilviLaser 2008, 8th International Conference on LiDAR Applications in Forest Assessment and Inventory*. R. Hill, J. Rosette and J. Suárez (eds). September 17–19, 2008 Heriot-Watt University, Edinburgh, UK, pp. 616–624.

Venables and Ripley 2002 *Statistics and Computing*. Springer, New York, 112 pp.

Verkasalo, E., Sairanen, P., Kilpeläinen, H. and Maltamo, M. 2004 Modelling the end-use based value of Norway spruce trees and logs by using predictors of stand and tree levels. In *Fourth Workshop. Connection Between Forest Resources and Wood Quality: Modelling Approaches and Simulation Software. September 8–15, 2002. Proceedings*. G. Nepveu (ed). Harrison Hot Springs, British Columbia, Canada, pp. 474–488.

Wall, T., Fröblom, J., Kilpeläinen, H., Lindblad, J., Heikkilä, A. and Song, T. *et al.* 2004 *Harvennusmännyn hankinnan ja sahauksen kehittäminen. [Developing the Procurement and Sawing of Thinned Scots Pine Trees]*. Wood Wisdom – tutkimusohjelman hankekonsortion julkien loppuraportti. Laaja versio. (In Finnish). Finnish Forest Research Institute, Joensuu Research Unit, Joensuu, Finland.

Received 18 September 2009