Remote Sensing of Environment 80 (2002) 88–99

Remote Sensing of Environment

www.elsevier.com/locate/rse

Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data

Erik Næsset*

Department of Forest Sciences, Agricultural University of Norway, P.O. Box 5044, N-1432 Ås, Norway

Received 12 March 2001; received in revised form 19 June 2001; accepted 25 June 2001

Abstract

The mean tree height, dominant height, mean diameter, stem number, basal area, and timber volume of 144 georeferenced field sample plots were estimated from various canopy height and canopy density metrics derived by means of a small-footprint laser scanner over young and mature forest stands using regression analysis. The sample plots were distributed systematically throughout a 1000-ha study area, and the size of each plot was 200 m^2. On the average, the distance between transmitted laser pulses was 0.9 m on the ground. The plots were divided into three strata according to age class and site quality. The stratum-specific regressions explained 82–95%, 74–93%, 39–78%, 50–68%, 69–89%, and 80–93% of the variability in ground-truth mean height, dominant height, mean diameter, stem number, basal area, and volume, respectively. A proposed practical two-stage procedure for prediction of corresponding characteristics of entire forest stands was tested. Sixty-one stands within the study area, with an average size of 1.6 ha each, were divided into 200 m^2 regular grid cells. The six examined characteristics were predicted for each grid cell from the corresponding laser data utilizing the estimated regression equations. Average values for each stand was computed. Most stand level predictions were unbiased ($P > .05$). Standard deviations of the differences between predicted and ground-truth values of mean height, dominant height, mean diameter, stem number, basal area, and volume were 0.61–1.17 m, 0.70–1.33 m, 1.37–1.61 cm, 16.9–22.2% (128–400 ha^{-1}), 8.6–11.7% (2.33–2.54 m^2 ha^{-1}), and 11.4–14.2% (18.3–31.9 m^3 ha^{-1}), respectively.

1. Introduction

During the last 10–15 years, several experiments have been carried out in order to determine various forest stand characteristics, such as tree heights and timber volume, by different airborne laser profiling and scanning systems (e.g., Lefsky, Cohen, et al., 1999; Magnussen & Boudewyn, 1998; Means et al., 1999, 2000; Næsset, 1997a, 1997b; Nelson, Krabill, & Tonelli, 1988). Most studies have focused on mature and old-growth forest, but determination of tree characteristics even in very young stands with tree heights below 6 m has been considered (Næsset & Bjerknes, 2001).

Previous studies have been dealing with thorough and detailed analyses of relationships between different metrics that may be derived from laser data and tree and canopy characteristics. The studies have been based on extensive data from a few intensively sampled study objects with limited spatial extension. Current knowledge about these relationships and how various properties of the applied instruments affect such relationships seem to form a sufficient basis for practical application of laser data in large-scale forest inventories. At least small-footprint systems with footprint diameters of, say, 10–30 cm, are becoming widely available and operated on a commercial basis by a large number of firms (Baltsavias, 1999). In the present study, a procedure for large-scale forest inventory based on data from a small-footprint scanning laser system will be presented and tested.

In Norway, two types of inventories are commonly used, i.e., (1) sample-based inventories where a limited sample of plots with a size of, for example, 200 m^2 each, is used to provide unbiased estimates of growing stock, timber value, etc., at the property, regional or even national level and (2) area-based inventories where the aim is to provide similar estimates of each forest stand of individual properties. The typical size of a stand is 1.5 ha, and the stand-based estimates are used primarily for forest management planning purposes. The procedure presented in this study is intended for use in stand inventories.

* Tel.: +47-64948906; fax: +47-64948890.
E-mail address: erik.naesset@isf.nlh.no (E. Næsset).

0034-4257/01/$ – see front matter
PII: S0034-4257(01)00290-5

Recently, various sampling procedures in which stands of an entire forest could be mapped from laser data using small field samples have been proposed (Means et al., 2000; Næsset, 1997b), and it is advocated that there is a large potential for savings as compared to conventional methods (Means et al., 2000). In practical forest management inventories in Norway, all stands of most forest properties in a municipality are surveyed, and a sample plot inventory that covers the entire municipality is often carried out. Such plot inventories typically comprise an area of 50–500 km^2 and often consist of more than 1000 sample plots. Næsset and Bjerknes (2001) proposed a practical two-stage procedure in which sample plots in very young forest were used to provide individual stand estimates of tree height from canopy heights derived from laser scanner data. This procedure could be applied in old and mature forest as well. In practical applications of laser scanner data, georeferenced sample plots could, in the first stage, be used to develop empirical relationships between various metrics derived from the laser data and tree characteristics measured in field. Such relationships could finally, in the second stage, be used to provide corresponding estimates for each stand from the laser data.

Several key variables in forest planning that at present are measured directly or predicted indirectly by traditional methods utilizing field data collection or stereoscopic aerial photointerpretation may be estimated from laser data. Empirical studies have indicated that mean tree height (Magnussen & Boudewyn, 1998; Magnussen, Eggermont, & LaRiccia, 1999; Means et al., 2000; Næsset, 1997a; Næsset & Bjerknes, 2001; Næsset & Økland, 2001), stand basal area (Means et al., 2000), and stand volume (Means et al., 2000; Næsset, 1997b) may be derived from small-footprint laser data with similar or even better accuracy than that of conventional methods. It has been shown that dominant height may be determined from small-footprint data, at least in very young stands (Næsset & Bjerknes, 2001). Furthermore, application of large-footprint systems has revealed that mean stem diameter and stem number may be estimated from laser data as well (Lefsky, Cohen, et al., 1999). Thus, the objective of the present study was to assess the accuracy of mean tree height, dominant height, mean stem diameter, stem number, stand basal area, and stand volume determined from laser scanner data of a commercially operated small-footprint system in entire young and mature forest stands according to a practical two-stage inventory procedure. The accuracy of the mentioned variables for smaller plots was also assessed.

Previous research has indicated that different age classes will generate laser canopy height distributions with significantly different forms (Lefsky, Cohen, et al., 1999). Furthermore, it has been revealed that relationships between stand characteristics and metrics derived from laser data may vary between tree species (Næsset, 1997a; Nelson, Oderwald, & Gregoire, 1997), and that crown shape which differs between species, affects laser metrics (Nelson, 1997). Commonly used stratification criteria such as age class and site quality, which is correlated with at least the dominant conifer tree species in Norway, may therefore be useful for an efficient stratification of inventories. In the present study, stratification according to age class and site quality was used.

2. Materials and methods

2.1. Study area and stand delineation

A forest area in the municipality of Våler (59°30′N 10°55′E, 70–120 m a.s.l.), southeast Norway, of about 1000 ha was selected for the trial. The main tree species in the area were Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.).

As part of a large-scale practical forest inventory going on in the district, aerial photographs were acquired 13 May 1996. Agfa Aviphot Pan 200 PE1 panchromatic black-and-white film was used, and the approximate scale was 1:15,000. The stand boundaries of all forest stands in the study area were delineated by means of stereoscopic photointerpretation and the boundary coordinates were stored in digital format in a geographical information system (GIS) using the Euref89 geodetic datum. The stands were classified according to criteria such as age class, site index, and tree species. The photointerpretation was accomplished using a Wild B8 stereoplotter equipped with linear encoders. The stereo models were referenced to the terrain by means of elevation points and terrain details of the analog official Economic Map Series. The root mean square error (RMSE) of the planimetric coordinates (x and y) of the Economic Map Series is 2 m. The RMSE of the planimetric coordinates after absolute orientation was expected to be 1.5–2.0 m. Thus, the total planimetric RMSE of the coordinates was approximately 2.5–3.0 m.

Two different ground-truth datasets were acquired in this study, i.e., (1) one consisting of sample plots distributed systematically throughout the entire 1000-ha study area used to establish relationships between stand characteristics and laser data in the first stage of the proposed procedure and (2) a dataset with selected stands for which stand characteristics were predicted from the laser data. The ground-truth stand data set was employed to verify the performance of the practical two-stage procedure.

2.2. Sample plot inventory

A total of 144 circular sample plots classified as young or mature forest according to the preliminary practical photointerpretation were distributed systematically throughout the 1000-ha study area according to a regular grid. The plots were sited independent of the inventoried stands used for verification, but since these stands covered more than 10% of the total study area, some of the sample plots were located within the inventoried stands. Each sample plot was assigned the age class and site quality class of the stand in which it was located according to the preliminary photointerpretation.

The 144 sample plots had an area of 200 m^2 each. They were divided into three strata according to age class and site quality. The first stratum comprised 56 plots classified as young forest of all site qualities, which implies that all mixtures of spruce and pine occurred (Table 1). The second stratum consisted of 36 plots classified as mature forest with poor site quality. Poor site quality was assigned to sites with interpreted H_{40} site index values equal to or less than 11. The H_{40} site index is defined by average age at breast height and the average height of the 100 largest trees per hectare according to diameter at breast height (dominant height). The specific values of the H_{40} index relate to the dominant height at an index age of 40 years (Braastad, 1980; Tveite, 1977). Scots pine was the dominating tree species of the 36 plots. Finally, 52 plots were classified as mature forest with good site quality, i.e., interpreted H_{40} site index values greater than 11. The 52 plots were covered mainly by spruce (Table 1).

Table 1
Summary of sample plot (200 m^2) reference data[a]

Characteristic	Range	Mean
Young forest (n = 56)		
h_L (m)	6.36–22.28	14.09
h_{dom} (m)	6.73–23.42	16.26
d_g (cm)	7.11–22.26	13.28
N (ha^{-1})	550–4600	2056
G (m^2 ha^{-1})	9.68–42.80	25.83
V (m^3 ha^{-1})	41.0–498.2	192.2
Age (year)	26–61	40
Tree species distribution		
Spruce (%)	0–100	52
Pine (%)	0–100	35
Broad-leaved species (%)	1–77	13
Mature forest, poor site quality (n = 36)		
h_L (m)	12.20–22.22	16.45
h_{dom} (m)	12.89–25.20	17.69
d_g (cm)	15.77–35.69	22.84
N (ha^{-1})	150–1100	522
G (m^2 ha^{-1})	7.54–32.93	19.75
V (m^3 ha^{-1})	59.0–280.1	155.0
Age (year)	46–171	126
Tree species distribution		
Spruce (%)	0–100	29
Pine (%)	0–100	64
Broad-leaved species (%)	1–48	7
Mature forest, good site quality (n = 52)		
h_L (m)	12.96–26.16	20.27
h_{dom} (m)	14.80–29.40	22.50
d_g (cm)	16.51–30.52	22.64
N (ha^{-1})	200–2050	722
G (m^2 ha^{-1})	7.50–50.63	27.45
V (m^3 ha^{-1})	54.0–639.8	269.2
Age (year)	41–152	87
Tree species distribution		
Spruce (%)	0–100	70
Pine (%)	0–100	23
Broad-leaved species (%)	1–41	7

[a] h_L = Lorey's mean height, h_{dom} = dominant height, d_g = mean diameter by basal area, N = stem number, G = basal area, V = volume.

Ground-truth data were collected during summer 1999. On each of the 144 plots, all trees with diameter at breast height (d_{bh}) > 4 cm were callipered. On 81 of the plots, tree heights of all trees were measured. On the remaining 63 plots, tree heights were measured on sample trees. The sample trees were selected with equal probability. The number of sample trees per plot ranged from 11 to 23 with an average of 16. The heights were measured with a Vertex hypsometer.

The ground-truth mean height of each plot was computed as the so-called Lorey's mean height (h_L), i.e., mean height weighted by basal area. In young forest, h_L was based on all measured tree heights. In mature forest, h_L was based on measured tree heights of trees with d_{bh} > 10 cm, which conforms to ordinary inventory practice. Dominant height of each plot was computed as the arithmetic mean height of the 100 largest trees per hectare according to diameter (h_{dom}), which is a commonly used definition (Tveite, 1977). Thus, for the 81 plots where all tree heights were measured, h_{dom} was computed as arithmetic mean of the two largest trees according to diameter. For the remaining plots, h_{dom} was computed as arithmetic mean of the two largest trees according to diameter if the heights of both these trees were measured. If height measurements were not available for both the two largest trees according to diameter, h_{dom} was assigned the height of the largest tree according to diameter that was measured. Mean plot diameter was computed as mean diameter by basal area (d_g). In young forest, d_g was computed from diameter of all callipered trees (d_{bh} > 4 cm). In mature forest, d_g was computed from trees with d_{bh} > 10 cm. Stem number was computed as number of trees per hectare (N). In young and mature forest, N was based on counted number of trees with d_{bh} > 4 cm and d_{bh} > 10 cm, respectively. Accordingly, plot basal area (G) was computed as basal area per hectare from diameter measurements of all trees with d_{bh} > 4 cm and d_{bh} > 10 cm, respectively. Volume of each tree was computed by means of volume equations of individual trees, which are based on height and diameter as predictor variables. The height of trees without height measurements was calculated from diameter–height relationships. Total plot volume (V) was computed as the sum of the individual tree volumes for trees with d_{bh} > 4 cm and d_{bh} > 10 cm, respectively. A summary of the ground-truth plot data is displayed in Table 1.

Differential Global Positioning System (GPS) and Global Navigation Satellite System (GLONASS) were used to determine the position of the centre of each sample plot. A Javad Legacy 20-channel dual-frequency receiver observing pseudorange and carrier phase of both GPS and GLONASS was used. Collection of data lasted for about 15–30 min for each plot with a 2-s logging rate and the Euref 89 geodetic datum was used. The antenna height was approximately 3.6 m for all points. Another Javad Legacy receiver was used as base station. The distance between the plots and the base station was approximately 19 km. Coordinates were computed by postprocessing in an adjustment with coordinates and carrier phase ambiguities as unknown parameters using both pseudorange and carrier phase observations (float solution). The Pinnacle version 1.00 (Anon., 1999) postprocessing software package was used. According to the *a priori* positional standard errors reported by Pinnacle, it is likely that the accuracy of the planimetric plot coordinates (*x* and *y*) ranged from < 0.1 to 2.5 m with an average of approximately 0.3 m (Næsset, 2001).

The 144 sample plots were spatially registered according to their centre coordinates and stored in a GIS. The plots were developed in GIS by means of buffers created around each of the plot centres. A radius of 7.98 m was used, which corresponded to circular plots with size 200 m^2.

2.3. Stand inventory

The study comprised a total of 61 stands. They were selected subjectively among the stands delineated by the practical forest inventory in order to represent different combinations of age classes, site quality classes, and tree species mixtures. The preliminary data of each stand from the practical photointerpretation were used for this selection, and the stands were divided into three strata according to the preliminary photointerpretation and the stratification criteria used for the sample plots. A total of 22 stands with an average area of 1.6 ha were classified as young forest, whereas 19 and 20 stands were classified as mature forest with poor and good site qualities, respectively (Table 2).

The ground-truth data of the selected stands were collected during summer and autumn 1998. Circular sample plots were distributed systematically according to a regular grid within each stand. The spacing was between 20 and 60 m and the number of plots per stand ranged from 14 to 30. The size of the individual plots in the young stands and the mature stands were 100 and 200 m^2, respectively. On each plot, all trees with $d_{bh} > 4$ cm and $d_{bh} > 10$ cm were callipered in young and mature stands, respectively. The heights of sample trees selected by horizontal point sampling were measured by a Vertex hypsometer. At least one sample tree was selected on each plot and the number of sample trees per stand ranged from 24 to 84.

The mean height of each stand was computed as the arithmetic mean of the individual sample trees. This mean height corresponds to Lorey's mean height (h_L). The standard error of the stand mean height (variation between trees) ranged from 0.3 to 0.9 m with an average of 0.5 m. Dominant height (h_{dom}) was computed as the arithmetic mean height of those sample trees that corresponded to the 100 largest trees per hectare according to diameter. Mean stand diameter was computed from callipered trees with $d_{bh} > 4$ cm and $d_{bh} > 10$ cm, respectively, as mean diameter by basal area (d_g). Stem number was computed as number of trees per hectare (N) based on counted number of trees with $d_{bh} > 4$ cm and $d_{bh} > 10$ cm, respectively. Accordingly, stand basal area (G) was computed as basal area per hectare of the callipered trees. Diameter–height relationships were developed from the sample trees, and stand volume (V) was calculated from standard volume equations for individual trees and the diameter–height relationships as volume per hectare. Average standard errors of stem number, stand basal area, and stand volume (variation between plots) were 102 ha^{-1}, 2.0 m^2 ha^{-1}, and 19.2 m^3 ha^{-1}, respectively.

Finally, to synchronize the h_L, h_{dom}, d_g, G, and V values to the date the laser data were acquired the individual stand

Table 2
Summary of stand reference data[a]

Characteristic	Range	Mean
Young forest (n = 22)		
Stand area (ha)	0.8–4.2	1.6
h_L (m)	10.30–19.92	13.90
h_{dom} (m)	13.74–21.49	16.62
d_g (cm)	10.01–21.25	13.23
N (ha^{-1})	1044–3078	1844
G (m^2 ha^{-1})	16.36–37.25	23.79
V (m^3 ha^{-1})	101.9–349.4	168.0
Age (year)	44–93	55
Tree species distribution		
Spruce (%)	7–100	48
Pine (%)	0–86	39
Broad-leaved species (%)	0–30	13
Mature forest, poor site quality (n = 19)		
Stand area (ha)	0.8–11.7	1.9
h_L (m)	13.76–19.30	16.37
h_{dom} (m)	15.35–22.05	18.07
d_g (cm)	18.29–24.62	21.17
N (ha^{-1})	283–958	577
G (m^2 ha^{-1})	12.58–30.60	19.84
V (m^3 ha^{-1})	91.8–261.1	154.8
Age (year)	70–153	118
Tree species distribution		
Spruce (%)	4–76	33
Pine (%)	18–92	60
Broad-leaved species (%)	2–22	7
Mature forest, good site quality (n = 20)		
Stand area (ha)	0.7–6.0	1.4
h_L (m)	16.06–23.09	19.77
h_{dom} (m)	18.05–26.77	22.38
d_g (cm)	16.81–24.78	21.24
N (ha^{-1})	526–1520	856
G (m^2 ha^{-1})	17.84–38.82	29.66
V (m^3 ha^{-1})	140.4–414.8	280.5
Age (year)	50–138	93
Tree species distribution		
Spruce (%)	45–94	72
Pine (%)	0–43	18
Broad-leaved species (%)	1–22	10

[a] h_L = Lorey's mean height, h_{dom} = dominant height, d_g = mean diameter by basal area, N = stem number, G = basal area, V = volume.

values were prorated by up to 1.5 years. The prorating of h_L and h_{dom} was done via age- and site quality-specific height growth rates (Braastad, 1975, 1980). The d_g and G values were prorated by means of diameter growth equations (Blingsmo, 1984) and V via volume growth equations (Delbeck, 1965). The prorated values were used as ground-truth. A summary of the ground-truth stand data is displayed in Table 2.

2.4. Laser scanner data

A Piper PA31-310 aircraft carried the ALTM 1210 laser scanning system produced by Optech, Canada. The laser transmits at 1064 nm (near-infrared). The major components of the ALTM 1210 are the infrared laser, the scanner transmitting the laser pulse and receiving the first and last echoes of each pulse, the time interval meter measuring the elapsed time between transmittance and receipt, the GPS airborne and ground receivers, and the inertial reference system reporting the aircraft's roll, pitch, and heading.

The laser scanner data were acquired 8 and 9 June 1999 (Næsset & Bjerknes, 2001). The plane was flown approximately 700 m above the ground and the average speed was 71 m s^{-1}. Forty-three flightlines were flown in a cross pattern. Nineteen parallel flightlines with approximately 50% overlap were flown in one direction and 24 parallel flightlines were flown perpendicular to the first 19 flightlines. Thus, every location in the study area was covered by laser data from four stripes. The pulse repetition frequency was 10 kHz and the scan frequency was 21 Hz. Maximum scan angle was 17°, which corresponded to an average swath width of about 420 m. Pulses transmitted at scan angles that exceeded 14° were excluded from the final dataset. The average footprint diameter was 21 cm and the average distance between footprints was 0.94 m on the sample plots and 0.92 m in the reference stands. First and last returns were recorded.

A complete postprocessing of the first and last pulse data was undertaken by the operating firm, Fotonor, Norway, by means of the proprietary postprocessing software supported by Optech. All ranges measured by the laser at an off-nadir angle, i.e., distances to the ground as well as to the tree canopy, were automatically converted to vertical distances.

After postprocessing, it was experienced that a few long last return ranges that exceeded the distance to the ground by up to 50 m were present in the data. According to the manufacturer, these erroneous ranges were caused by a faulty last return sensor. A second flight was therefore carried out 6 June 2000 to collect last return data only. Flying height corresponded to that of the first flight, but 14 parallel flightlines were flown in only one direction.

The ranging device was calibrated by Optech. The operating firm always calibrates the system after installation in the aircraft. In addition, 30 circular plots were located on plane road segments distributed throughout the study area. Their positions were determined by differential GPS + GLONASS based on accurate dual-frequency carrier phase observations. Based on this calibration, the computed ranges of the first pulse data acquired in 1999 were reduced by 0.13 m. The last pulse ranges collected in 1999 and 2000 were extended by 0.46 m and reduced by 0.11 m, respectively.

The last return data collected in 2000 were used to model the ground surface. In forest areas, a large portion of the last returns of the laser pulses usually represent hits above the ground, such as tree canopy and bushes. The last return data were therefore filtered during postprocessing by the operating firm. In the filtering, local maxima assumed to represent vegetation hits were discarded. A triangulated irregular network (TIN) was generated from the planimetric coordinates (x and y) using the Euref89 geodetic datum and corresponding height values of the individual terrain ground points retained in the last pulse dataset. This TIN model of the ground was denoted as the "digital terrain model" (DTM). The ellipsoidic height accuracy of the DTM was expected to be around 25 cm, at least in relatively plane parts of the area. Studies of terrain surfaces in wooded areas based on more scattered laser datasets than the current and similar filtering procedures indicate an accuracy of 25–30 cm (Kraus & Pfeifer, 1998; Opseth, 1996).

Since the first pulse data and a large portion of the last returns were expected to represent hits above the ground, both first and last return data collected in 1999 were used to derive tree canopy heights. Although some erroneous long ranges were inherent in last pulse data collected in 1999, these errors would certainly not affect the vegetation hits since such hits would represent elevations above the DTM. All first and last return observations (points) were spatially registered to the DTM according to their geographic coordinates. Terrain surface height values were computed for each point by linear interpolation from the DTM. The relative height of each point was computed as the difference between the height of the first or last return and the terrain surface height. In wooded areas, this relative height can be considered as the tree canopy height. In open areas and in gaps between the trees, this height will be close to zero and thus represent ground hits. Observations with a height value less than 2 m were excluded to eliminate ground hits and the effect of stones, shrubs, etc. from the tree canopy datasets (cf. Næsset, 1997a; Nilsson, 1996). The remaining first and last pulse observations were considered to be canopy hits. Various measures of canopy penetration rate have been useful for estimation of, for example, timber volume from laser data (Means et al., 2000; Næsset, 1997b). The total number of transmitted pulses that were recorded as first and last returns, respectively, were therefore of interest. Thus, two datasets were developed from the laser scanner data for the pulses classified as first returns and two datasets for the pulses classified as last returns: (1) geographically registered data on canopy height derived from the first and last returns, respectively, and (2) geographically registered data for all transmitted pulses that were classified as first and last returns, respectively.

2.5. Computations

The first and last pulse datasets were spatially registered to the sample plots and the inventoried reference stands. Pulses that hit outside these objects were excluded from further analysis.

First, multiple regression analysis was used to create stratum-specific relationships between field measurements and laser data. First and last pulse height distributions were created from the laser canopy heights (>2 m, see above) of each 200 m^2 sample plot. A large number of metrics was derived from these distributions. According to previous experiences (Lefsky, Harding, Cohen, Parker, & Shugart, 1999; Magnussen & Boudewyn, 1998; Magnussen et al., 1999; Means et al., 1999, 2000; Næsset, 1997a; Næsset & Bjerknes, 2001; Næsset & Økland, 2001; Nelson, 1997; Nelson et al., 1988, 1997; Ziegler et al., 2000), the following metrics were derived: (1) the quantiles corresponding to the 0, 10, . . ., 90 percentiles of the distributions, (2) the maximum values, (3) the mean values, and (4) the coefficients of variation. Furthermore, several measures of canopy density were derived. Canopy densities were computed as the proportions of first pulse laser hits above the 0, 10, . . ., 90 quantiles of the first pulse height distributions to total number of first pulses. Corresponding densities were computed for the last pulse data as well.

In the multiple regression analysis, multiplicative models were estimated as linear regressions in the logarithmic variables because such models were found to be suitable for estimation of mean height, stem number, and volume by others (Næsset, 1997b; Næsset & Bjerknes, 2001; Næsset & Økland, 2001). The multiplicative model was formulated as (Eq. (1)):

$$Y = \beta_0 h_{0f}^{\beta_1} h_{10f}^{\beta_2} \dots h_{90f}^{\beta_{10}} h_{0l}^{\beta_{11}} h_{10l}^{\beta_{12}} \dots h_{90l}^{\beta_{20}} h_{\mathrm{maxf}}^{\beta_{21}} h_{\mathrm{maxl}}^{\beta_{22}} h_{\mathrm{meanf}}^{\beta_{23}}$$

$$\times h_{\mathrm{meanl}}^{\beta_{24}} h_{\mathrm{cvf}}^{\beta_{25}} h_{\mathrm{cvl}}^{\beta_{26}} d_{0f}^{\beta_{27}} d_{10f}^{\beta_{28}} \dots d_{90f}^{\beta_{36}} d_{0l}^{\beta_{37}} d_{10l}^{\beta_{38}} \dots d_{90l}^{\beta_{46}} \quad (1)$$

whereas the linear form used in the estimation was (Eq. (2)):

$$\begin{aligned} \ln Y = {} & \ln\beta_0 + \beta_1 \ln h_{0f} + \beta_2 \ln h_{10f} + \dots + \beta_{10} \ln h_{90f} \\ & + \beta_{11} \ln h_{0l} + \beta_{12} \ln h_{10l} + \dots + \beta_{20} \ln h_{90l} \\ & + \beta_{21} \ln h_{\mathrm{maxf}} + \beta_{22} h_{\mathrm{maxl}} + \beta_{23} \ln h_{\mathrm{meanf}} \\ & + \beta_{24} \ln h_{\mathrm{meanl}} + \beta_{25} \ln h_{\mathrm{cvf}} + \beta_{26} \ln h_{\mathrm{cvl}} \\ & + \beta_{27} \ln d_{0f} + \beta_{28} \ln d_{10f} + \dots + \beta_{36} \ln d_{90f} \\ & + \beta_{37} \ln d_{0l} + \beta_{38} \ln d_{10l} + \dots + \beta_{46} \ln d_{90l} \quad (2) \end{aligned}$$

where Y= field values of h_L (m), h_{dom} (m), d_g (cm), N (ha^{-1}), G (m^2 ha^{-1}), or V (m^3 ha^{-1}); h_{0f}, h_{10f}, . . ., h_{90f} = the quantiles corresponding to the 0, 10, . . ., 90 percentiles of the first pulse laser canopy heights (m); h_{0l}, h_{10l}, . . ., h_{90l} = the quantiles corresponding to the 0, 10, ..., 90 percentiles of the last pulse laser canopy heights (m); h_{maxf}, h_{maxl} = maximum first and last pulse laser canopy heights (m); h_{meanf}, h_{meanl} = mean of the first and last pulse laser canopy heights (m); h_{cvf}, h_{cvl} = coefficient of variation of the first and last pulse laser canopy heights (%); d_{0f}, d_{10f}, . . ., d_{90f} = canopy densities corresponding to the proportions of first pulse laser hits above the 0, 10, . . ., 90 quantiles to total number of first pulses, d_{0l}, d_{10l}, . . ., d_{90l} = canopy densities corresponding to the proportions of last pulse laser hits above the 0, 10, . . ., 90 quantiles to total number of last pulses.

Stepwise selection was performed to select variables to be included in the final models. No predictor variable was left in the model with a partial F statistic with a significance level greater than .05. The standard least-squares method was used (Anon., 1989). After selection of variables, the selected regression was converted back to original scale by adding half of the variance to the intercept before conversion (Goldberger, 1968).

No independent data were available to assess the accuracy of the estimated regression equations used for prediction on small plots. Cross-validation was therefore used to assess the accuracy. For each stratum (n sample plots), one of the sample plots was removed from the dataset at a time, and the selected models were fitted to the data from the $n-1$ remaining plots. The six studied variables of each stratum were then predicted for the removed plot. This procedure was repeated until predicted values were obtained for all plots.

The objective of the last stage of the proposed procedure was to predict mean height, dominant height, mean diameter, stem number, basal area, and volume of the stands in the trial area. At least, for mean height, it has been argued that for a given area, for example, a given sample plot size, there exist a certain relationship between the height of interest and predictor variables derived from laser canopy height distributions (Magnussen & Boudewyn, 1998). Furthermore, previous research has indicated that by subdividing entire forest stands into grid cells of sizes between 3×3 and 30×30 m, quite accurate stand averages of individual cell estimates of mean height, basal area, and volume may be obtained from laser data (Magnussen & Boudewyn, 1998; Means et al., 2000; Næsset, 1997a, 1997b; Næsset & Bjerknes, 2001; Ziegler et al., 2000). Thus, a regular grid with cell size 14.14×14.14 m corresponding to the sample plot size of 200 m^2 was generated by means of GIS. The grid covered the entire 1000-ha study area. To relate the reference stand boundaries and the four laser datasets to the grid, the grid was laid over the stand boundary map from the preliminary practical photointerpretation and spatially registered to the laser data.

For each 14.14×14.14 m grid cell of every stand, first and last pulse height distributions were created from the laser canopy heights and metrics corresponding to those of the sample plots were derived. The selected stratum-specific regression models were used to predict h_L, h_{dom}, d_g, N, G,

Table 3
Summary of laser scanner data

Dataset	No. of obs.	No. of transmitted pulses[a] (ha^{-1})		No. of canopy hits[a] (ha^{-1})		Mean rate of penetration[a] (%)
		Range	Mean	Range	Mean	
Sample plots						
Young forest	56	9750–14,800	11,428	5500–12,750	9504	17
Mature forest, poor site quality	36	7450–16,100	11,270	2650–12,450	6961	38
Mature forest, good site quality	52	7400–14,300	10,935	4150–13,350	8710	20
Stands						
Young forest	22	10,770–12,874	11,837	6755–10,827	9149	23
Mature forest, poor site quality	19	11,203–13,585	11,857	6042–10,762	7901	33
Mature forest, good site quality	20	10,371–14,096	11,999	7434–13,330	9579	20

[a] Refers to first pulse data.

and *V* of each grid cell. Finally, stand estimates were computed as weighted mean values of the individual cell estimates. The area of each cell was used as weights in order to reduce the influence of "sliver cells" with size <200 m^2 occurring along the stand boundaries (cf. Næsset, 1997b). However, cells with size <150 m^2 were discarded from the computations, since it was assumed that cell size affected the relationship between field measurements and laser data (cf. Magnussen & Boudewyn, 1998). Furthermore, area-based dependent variables, i.e., *N*, *G*, and *V*, were set to 0 in grid cells where no canopy hits were recorded ("no trees") (cf. Næsset, 1997b). In such cells, the predicted values of h_L, h_{dom}, and d_g were discarded from the corresponding mean stand estimates (cf. Næsset, 1997a).

Since the regression equations were estimated from sample plots with size equal to the grid cell size used in each stand, the expected ground-truth values of the stands should be equal to the expected values estimated by the two-stage procedure. The differences between the stand variables estimated from the laser data and the corresponding ground-truth values were evaluated. The statistical significance of the mean differences, i.e., whether systematical differences were present, was assessed by means of two-tailed *t* tests.

Table 4
Differences (*D*) between predicted and ground-truth values of Lorey's mean height (h_L), dominant height (h_{dom}), mean diameter by basal area (d_g), stem number (*N*), basal area (*G*), and volume (*V*), respectively, and standard deviation (S.D.) for the differences in cross-validation of the selected regression equations in Table 6[a]

Variable	Observed mean	*D*		S.D.
		Range	Mean	
Young forest (n = 56 plots)				
h_L (m)	14.09	−2.25–2.90	−0.01 NS	0.87
h_{dom} (m)	16.26	−3.23–3.28	0.02 NS	1.27
d_g (cm)	13.28	−3.88–4.19	0.03 NS	1.66
N (ha^{-1})	2056	−1817–1441	11 NS	649
G (m^2 ha^{-1})	25.83	−8.76–7.83	0.07 NS	3.63
V (m^3 ha^{-1})	192.2	−83.6–63.7	−0.1 NS	29.3
Mature forest, poor site quality (n = 36 plots)				
h_L (m)	16.45	−2.46–3.63	−0.01 NS	1.00
h_{dom} (m)	17.69	−3.41–3.25	−0.01 NS	1.50
d_g (cm)	22.84	−6.71–6.71	0.02 NS	3.20
N (ha^{-1})	522	−325–312	6 NS	135
G (m^2 ha^{-1})	19.75	−13.22–6.99	0.15 NS	4.54
V (m^3 ha^{-1})	155.0	−117.3–75.8	1.1 NS	37.8
Mature forest, good site quality (n = 52 plots)				
h_L (m)	20.27	−2.91–3.94	0.01 NS	1.36
h_{dom} (m)	22.50	−3.90–4.84	0.03 NS	1.54
d_g (cm)	22.64	−9.33–7.16	0.04 NS	2.84
N (ha^{-1})	722	−1026–1867	38 NS	377
G (m^2 ha^{-1})	27.45	−12.25–13.71	0.14 NS	5.64
V (m^3 ha^{-1})	269.2	−153.8–160.6	1.1 NS	62.0

[a] Level of significance: NS = not significant (>.05).

Table 5
Differences (*D*) between predicted and ground-truth values of Lorey's mean height (h_L), dominant height (h_{dom}), mean diameter by basal area (d_g), stem number (*N*), basal area (*G*), and volume (*V*), respectively, and standard deviation (S.D.) for the differences in stand predictions using the selected regression equations in Table 6[a]

Variable	Observed mean	*D*		S.D.
		Range	Mean	
Young forest (n = 22 stands)				
h_L (m)	13.90	−1.50–2.42	0.42*	0.87
h_{dom} (m)	16.62	−2.86–3.04	−0.08 NS	1.33
d_g (cm)	13.23	−3.18–3.68	0.72*	1.60
N (ha^{-1})	1844	−967–689	−90 NS	400
G (m^2 ha^{-1})	23.79	−4.82–6.18	−0.86 NS	2.48
V (m^3 ha^{-1})	168.0	−33.6–52.1	6.2 NS	24.0
Mature forest, poor site quality (n = 19 stands)				
h_L (m)	16.37	−1.31–1.13	−0.09 NS	0.61
h_{dom} (m)	18.07	−1.98–1.25	−0.31 NS	0.70
d_g (cm)	21.17	−1.80–3.60	0.78*	1.61
N (ha^{-1})	577	−212–363	15 NS	128
G (m^2 ha^{-1})	19.84	−2.95–5.47	0.74 NS	2.33
V (m^3 ha^{-1})	154.8	−37.2–35.3	8.2 NS	18.3
Mature forest, good site quality (n = 20 stands)				
h_L (m)	19.77	−2.52–1.85	−0.01 NS	1.17
h_{dom} (m)	22.38	−2.77–2.37	−0.43 NS	1.32
d_g (cm)	21.24	−1.27–4.22	0.98**	1.37
N (ha^{-1})	856	−417–106	−103**	145
G (m^2 ha^{-1})	29.66	−6.22–3.80	−0.67 NS	2.54
V (m^3 ha^{-1})	280.5	−55.1–57.7	0.3 NS	31.9

[a] Level of significance: NS = not significant (>.05). * <.05. ** <.01.

3. Results

First, mean height, dominant height, mean diameter, stem number, basal area, and volume of the 144 sample plots were regressed against the predictor variables derived from the distributions of first and last return laser data classified as canopy hits. Individual models were developed for each of the three strata.

Principal component analysis based on the correlation matrix was used to assess the presence of collinearity in the regression analysis. The square root of the largest eigenvalue divided by the smallest eigenvalue (condition number, κ) was used as a means for suggesting collinearity. A condition number larger than 30 has been proposed to indicate collinearity (Weisberg, 1985). Some of the models initially suggested by the stepwise selection procedure were subject to serious collinearity. The models selected for further analysis were therefore those indicated by the stepwise procedure that fulfilled the requirement of $\kappa < 30$. The maximum condition number of the selected models was 11.7 (Table 6), which indicated no serious collinearity problems of the final models.

All the selected models comprised less than five predictor variables (Table 6). Predictor variables derived from the first as well as the last pulse data were represented. The coefficient of determination (R^2) for mean height (h_L) in young forest was as high as .95. The models for h_L of the two other strata revealed R^2 values of .86 and .82.

In general, higher R^2 values were found in young forest than in the two strata comprising mature forest. For dominant height, stem number, and basal area the R^2 values were in the range between .50 and .93. Coefficient of determination for volume ranged from .80 to .93. The poorest correlation was found for mean diameter with R^2 values between .39 and .78. The highest RMSE was obtained in the models for stem number (.28–.35).

The selected regression equations were transformed back to original scale. Cross-validation of the selected models revealed that the mean differences between predicted and observed values of the 200 m^2 sample plots were far from significance in the statistical sense for all the validated models of the three strata (Table 4). The standard deviations of the differences between predicted and ground-truth values ranged from 0.87 to 1.36 m for mean height (h_L) and from 1.27 to 1.54 m for dominant height (h_{dom}). For mean diameter (d_g), stem number (N), basal area (G), and volume (V), the standard deviations ranged from 1.66 (12.5%) to 3.20 cm (14.0%), 135 to 649 ha^{-1} (25.9–52.2%), 3.63 to

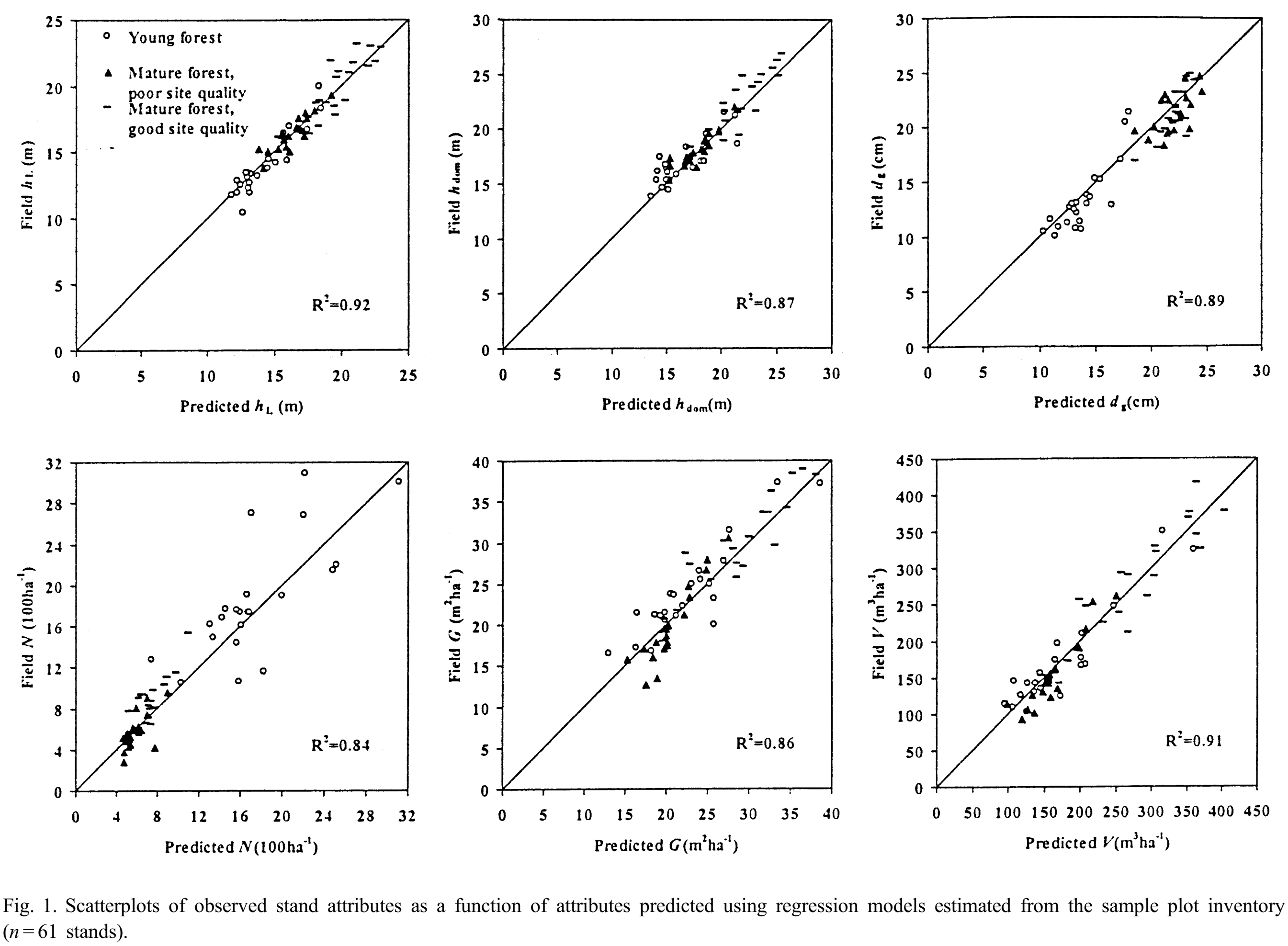

Fig. 1. Scatterplots of observed stand attributes as a function of attributes predicted using regression models estimated from the sample plot inventory ($n = 61$ stands).

5.64 m^2 ha^{-1} (14.1–23.0%), and 29.3 to 62.0 m^3 ha^{-1} (15.2–24.4%), respectively.

The stratum-specific equations were used in the second stage of the proposed procedure to predict mean height, dominant height, mean diameter, stem number, basal area, and volume of the 61 reference stands. The predicted values were compared with field data.

The comparison revealed that five of the 18 computed mean differences (bias) between predicted values and ground-truth were larger than what could be expected due to randomness (Table 5). For h_L, a mean difference of 0.42 m indicating an overestimation by the laser-based procedure was found in young stands. In mature stands, the predicted heights underestimated the field values by 0.01–0.09 m. The predicted dominant heights (h_{dom}) underestimated the ground-truth by 0.08–0.43 m, although these differences were not significant in the statistical sense.

The ground-truth values for mean diameter (d_g) were significantly overestimated by 0.72–0.98 cm, which corresponded to a bias of 3.7–5.4%, whereas a bias of between − 12.0% and 2.6% was found for stem number (N). For G and V, neither of the differences differed significantly from zero at the 5% level. The bias ranged from − 0.86 (− 3.6%) to 0.74 m^2 ha^{-1} (3.7%) for G and from 0.3 (0.1%) to 8.2 m^3 ha^{-1} (5.3%) for V.

The standard deviations between predicted and ground-truth values for h_L and h_{dom} ranged from 0.61 to 1.17 m and from 0.70 to 1.33 m, respectively. The square of the correlation coefficients (R^2) (Pearson product-moment correlations) between predicted and field values of h_L and h_{dom} in the 61 stands were .92 and .87, respectively (Fig. 1).

For d_g, N, and G, the standard deviations for the differences between predicted and ground-truth values ranged between 6.5% and 12.1%, 16.9% and 22.2%, and 8.6% and 11.7%, respectively. The corresponding R^2 values were .89, .84, and .86, respectively.

Standard deviations for V ranged from 18.3 to 31.9 m^3 ha^{-1} or 11.4–14.2%. The R^2 value for the 61 stands was .91.

4. Discussion and conclusions

The results of the current study revealed that the proposed practical two-stage procedure seem to be robust with respect to bias for most of the investigated stand characteristics. Although the sample plots used to estimate the regression equations that related tree and stand characteristics measured in field to predictor variables derived from the laser data were distributed independently of the reference stands throughout a 1000-ha area, the mean difference between laser-derived and ground-truth values were less than 0.5 m for mean height as well as dominant height (Table 5). Except from prediction of stem number in mature forest, the bias for mean diameter, stem number, basal area, and volume was less than 5% in all strata.

In accordance with Magnussen and Boudewyn (1998), it is probably essential for a successful removal of bias that the size of the field sample plots corresponds to the grid cell size used to predict characteristics at stand level and/or that the number of transmitted pulses are almost equal for all plots and all grid cells which the stands are divide into. The distribution of transmitted pulses across the current trial area indicates a variable pulse density in certain parts of the material. Small sample plots will be influenced most seriously by such uneven distributions, since entire stands will represent an average distribution over a larger area. These patterns are quite evident in the material. In, for example, the stratum comprising mature forest with poor site quality, the number of transmitted pulses of the sample plots ranged from 7450 to 16,100 ha^{-1}, whereas the corresponding range for the stands were from 11,203 to 13,585 ha^{-1} (Table 3). It is likely that the uneven distribution of pulses has affected the estimated coefficients of the regression equations and thus introduced both systematic and random errors in the predictions at stand level. On the other hand, in practical cases, the distribution of pulses will never be perfectly regular since the scanner distributing the pulses produces a zig-zag pattern of pulses on the ground and the flight planning often suggests a considerable overlap between adjacent stripes to avoid gaps between them. Steep terrain and complicated topography will also influence the pulse density. Thus, the differences between ground-truth and predicted values in this study are likely to give quite realistic estimates of what might be expected in practical applications.

The standard deviations between predicted and ground-truth values of the plots revealed by cross-validation (Table 4) were up to 160% larger than the corresponding standard deviations obtained for entire stands (Table 5). The relationship between precision on small plots and entire stands illustrates two important aspects. First, small areas are subject to substantial inherent variation around canopy height quantiles leading to highly variable predictions (cf. Magnussen & Boudewyn, 1998). The size of the sample plots and stand grid cells should therefore not be too small. On the other hand, extended plot size will increase the inventory costs. Thus, traditional cost/benefit analysis could be used to balance precision and costs. Second, averaging over several grid cells (stand estimates) will tend to reduce standard errors of mean values and thus improve the precision of stand estimates. In the present study, it is shown empirically that even though the variability at plot level or individual grid cell level is considerable, quite precise estimates may be obtained at the stand level, which is certainly an attractive property of the proposed inventory procedure.

The stratification of plots and stands according to age class and site quality was accomplished in order to separate forest types that were expected to generate distinct distributions of canopy heights from the laser data (cf. Lefsky, Cohen, et al., 1999; Næsset, 1997a; Nelson, 1997; Nelson et al., 1997). It is evident that site quality was a good instrument for stratification according to tree species. The poor sites in the mature forest were dominated by pine and the good sites by spruce (Tables 1 and 2), and for each of the three strata the tree species distributions were quite equal for plots and stands. Furthermore, properties of the distributions of laser canopy heights, such as penetration rate, varied significantly between strata, whereas the penetration rate of plots and stands in corresponding strata were quite equal. For the plots, the mean penetration rates varied between 17% and 38% (Table 3). Stratification according to age class and site quality may therefore be efficient to represent distinct forest types in practical inventories.

The present trial was conducted in an area with small variations in altitude (70–120 m a.s.l.). The flying height above the terrain was therefore quite stabile, and thus the ranges of the laser returns were quite constant. In practical forest inventories in Norway where the altitude may vary from 100 to 1000 m a.s.l. within relatively short distances, it may be difficult to keep the flying height above the terrain constant when an airplane is used as platform. With variations in laser return ranges, the footprint diameter will vary accordingly. Increased footprint size tends to elevate the canopy height inferred from laser data (cf. Aldred & Bonnor, 1985; Nilsson, 1996). It is therefore likely that practical applications of the proposed procedure will benefit from stratification of the forest area according to flying height above the terrain. It is likely that even model-based approaches to derivation of stand characteristics that reduce reliance on auxiliary ground data, such as those proposed by Magnussen and Boudewyn (1998) and Magnussen et al. (1999), would require stratification according to flying height to perform accurately.

In spite of the considerable variation around canopy height quantiles and in laser sampling density of the 200 m^2 plots, the coefficients of determination (R^2) for mean height of .82–.95 and dominant height of .74–.93 (Table 6) seem to correspond to previous findings of small-footprint trials. Means et al. (2000) and Næsset and Økland (2001) reported R^2 values of .93 and .91, respectively, for mean height. Means et al. (2000) used plots with size 2500 m^2, whereas the plots employed by Næsset and Økland (2001) were approximately 50 m^2 each. For dominant height, proportions of explained variation of approximately 60–80% have been encountered for small plots located in very young forest with tree heights less than 11 m (Næsset & Bjerknes, 2001), as well as in entire forest stands (Ziegler et al., 2000). For the other variables considered in this

Table 6
Relationships between logarithmic transformations of ground-based characteristics of the 200 m^2 sample plots (dependent variables) and laser-derived metrics from stepwise multiple regression analysis

Dependent variable[a]	Predictive model[b]	R^2	RMSE	κ
Young forest (n = 56)				
ln h_L	$0.46 + 1.149\ln h_{90l} - 0.28\ln h_{maxl}$	0.95	0.06	5.0
ln h_{dom}	$0.568 + 1.169\ln h_{90l} - 0.286\ln h_{maxl}$	0.93	0.07	5.0
ln d_g	$-0.867 + 0.217\ln h_{10l} + 0.665\ln h_{80l} - 0.805\ln d_{80f}$	0.78	0.12	3.1
ln N	$15.99 - 1.182\ln h_{80l} + 3.08\ln d_{80f}$	0.68	0.28	1.8
ln G	$3.492 + 0.536\ln h_{10f} + 1.388\ln d_{50f}$	0.89	0.14	2.2
ln V	$3.473 + 1.336\ln h_{meanl} + 1.477\ln d_{50f}$	0.93	0.16	2.0
Mature forest, poor site quality (n = 36)				
ln h_L	$0.285 + 1.01\ln h_{90f} - 0.107\ln h_{50l}$	0.86	0.05	2.2
ln h_{dom}	$-0.0187 + 1.002\ln h_{maxf}$	0.74	0.08	1.0
ln d_g	$0.206 + 0.77\ln h_{90l} - 0.312\ln d_{80l}$	0.54	0.12	1.4
ln N	$11.24 + 1.195\ln h_{0f} - 1.662\ln h_{maxf} + 1.156\ln d_{20l}$	0.65	0.30	1.7
ln G	$4.253 + 4.304\ln h_{50f} - 4.022\ln h_{60f} + 0.584\ln d_{90f}$	0.69	0.21	8.5
ln V	$4.951 - 1.278\ln h_{30f} + 5.994\ln h_{50f} - 3.8\ln h_{60f} + 0.766\ln d_{90f}$	0.80	0.20	11.7
Mature forest, good site quality (n = 52)				
ln h_L	$0.35 + 0.529\ln h_{90f} + 0.355\ln h_{maxf}$	0.82	0.07	5.9
ln h_{dom}	$0.525 + 0.23\ln h_{80f} + 0.637\ln h_{maxf} + 0.084\ln d_{10l}$	0.85	0.07	4.3
ln d_g	$0.441 + 0.64\ln h_{90l} - 0.277\ln d_{90l}$	0.39	0.12	1.7
ln N	$10.33 - 0.487\ln h_{0l} - 0.667\ln h_{cvf} + 1.187\ln d_{50f}$	0.50	0.35	1.9
ln G	$3.608 + 2.629\ln h_{80l} - 2.157\ln h_{maxf} + 1.26\ln d_{50f}$	0.75	0.21	3.8
ln V	$3.151 + 3.027\ln h_{80l} - 1.66\ln h_{maxf} + 1.223\ln d_{50f}$	0.80	0.22	3.8

[a] h_L = Lorey's mean height (m), h_{dom} = dominant height (m), d_g = mean diameter by basal area (cm), N = stem number (ha^{-1}), G = basal area (m^2 ha^{-1}), V = volume (m^3 ha^{-1}).

[b] h_{0f}, h_{10f}, h_{30f}, h_{50f}, h_{60f}, h_{80f}, and h_{90f} = the quantiles corresponding to the 0, 10, 30, 50, 60, 80, and 90 percentiles of the first pulse laser canopy heights (m); h_{0l}, h_{10l}, h_{50l}, h_{80l}, and h_{90l} = the quantiles corresponding to the 0, 10, 50, 80, and 90 percentiles of the last pulse laser canopy heights (m); h_{maxf} and h_{maxl} = maxima of first and last pulse laser canopy heights (m); h_{cvf} = coefficient of variation of first pulse laser canopy heights (%); h_{meanl} = arithmetic mean of last pulse laser canopy heights (m); d_{50f}, d_{80f}, and d_{90f} = canopy densities corresponding to the proportions of first pulse laser hits above the 50, 80, and 90 quantiles, respectively, to total number of first pulses; and d_{10l}, d_{20l}, d_{80l}, and d_{90l} = canopy densities corresponding to the proportions of last pulse laser hits above the 10, 20, 80, and 90 quantiles, respectively, to total number of last pulses.

study, only some scattered results where the experimental conditions were somewhat different are available for comparison. Means et al. (2000) reported R^2 values for basal area and volume of .94–.95 and .95–.97, respectively, for 2500 m^2 plots, which is slightly better than .69–.89 and .80–.93 (Table 6) in the present study. On the other hand, results from estimation of volume in stands with an average area of 1.5 ha (Næsset, 1997b) indicated smaller proportions of explained variation (46–89%) than the current trial.

Even though the stand predictions in the second step of the proposed inventory procedure were based on regression equations derived from plots sampled independently of the stands, the stand level precision seem to be fully competitive to that of previous studies based on larger plots and stands. The square of the Pearson product-moment correlation between predicted and field values of .84–.92 (Fig. 1), which are of the same magnitude as those mentioned above, indicate good correspondence between the present and previous trials. As compared to current inventory practice, however, the precision of the applied procedure seems to be superior to that of traditional inventory methods based on field measurements or aerial photointerpretation. The standard deviations between predicted and ground-truth mean height of 0.61–1.17 m (Table 5) are significantly smaller than the 1.1–2.9 m of current methods (Næsset, 1996). As a matter of fact, the average standard error of the ground-"truth" stand mean heights of approximately 0.5 m indicates that the true precision of the laser-based procedure is even higher than 0.61–1.17 m. Prior to the trial, I expected the precision of laser-derived mean height to be around 1–1.5 m. It seems that I underestimated the required sampling effort when the field data collection was planned.

The stand level precision of dominant height, mean diameter, stem number, and basal area was also superior to that of conventional inventory methods. For mean diameter, standard deviations between predicted and ground-truth values at stand level of 8–20% have been reported for methods based on field inventory and photointerpretation (Eid & Nersten, 1996; Næsset, 1995), which is poorer than 6.5–12.1% (1.37–1.61 cm, Table 5) of the present study. For stem number and basal area, a precision of 15–44% (Eid & Nersten, 1996; Næsset, 1996) and 12–22% (Eid, 1996; Eid & Nersten, 1996), respectively, has been found for conventional methods, which is significantly poorer than 16.9–22.2% (128–400 ha^{-1}, Table 5) and 8.6–11.7% (2.33–2.54 m^2 ha^{-1}, Table 5) reported here. Dominant height, however, is not measured directly at stand level by current inventory practice. In Norway, dominant height is computed from mean height, which implies that the precision of traditional methods is poorer than 1.1–2.9 m indicated above and certainly poorer than 0.70–1.33 m (Table 5) revealed in the present study.

Finally, the precision of laser-derived volume of 11.4–14.2% (18.3–31.9 m^3 ha^{-1}, Table 5) seems to be significantly better than that of volume determined by practical field inventories and photointerpretation of 13–33% (Eid & Næsset, 1998). As for reference mean height, the mean ground-"truth" stand volumes seem to be subject to substantial inherent errors (S.E. = 19.2 m^3 ha^{-1}) as compared to the revealed level of precision of laser-derived timber volumes. The true precision may therefore be even higher than indicated here.

Selection of a proper mathematical model for regression functions is essential when estimated regression equations are to be used for prediction purposes. Although previous research has indicated that logarithmic transformations of the variables may be suitable in regression analysis of laser data (Means et al., 2000; Næsset, 1997b; Næsset & Bjerknes, 2001; Næsset & Økland, 2001), there is no guarantee that equations based on multiplicative models and estimated as linear regressions in the logarithmic variables will provide unbiased predictions. In the present study, the regression equations used to predict reference stand characteristics were estimated as linear regressions in the logarithmic variables and then converted back to original scale before prediction. By conversion of the loglinear equations to original scale, a bias will be introduced (e.g., Goldberger, 1968). It is therefore essential that proper corrections are undertaken to avoid serious bias of predictions. Under given circumstances, proper corrections may be rather complicated to accomplish. However, when the differences between the sample plot data and the grid cell data of the stands are small, MSE ≤ 0.5, and $n-k \geq 30$, i.e., the number of observations minus the number of predictor variables, an approximate and simple correction that will introduce an error of less than 1%, is to add half the variance to the regression intercept before conversion (Flewelling & Pienaar, 1981). In the present study, all requirements for using this approximate correction were fulfilled (Tables 1–3, 6). For future practical applications, however, the sampling scheme should be considered carefully since the number of plots assigned to each stratum may restrict the selection of regression models used for stand predictions.

The present trial has indicated that all the six examined stand characteristics can be determined by the proposed two-stage stand inventory procedure with higher precision than conventional methods in practical use today. The precision was fairly high for all variables except stem number. Furthermore, the procedure has proved to yield unbiased estimates in most cases. At least, in steep terrain and areas with large topographic variations, it will be very difficult to keep the flying height constant and thus ensure a stabile laser footprint size. Further research should therefore be devoted to the effect of variable footprint size on the efficiency of the proposed procedure, and to which measures that could be taken to account for such increased variability of the laser measurements. Furthermore, in the present study, a recent and updated photointerpretation was used for stratification of the inventory and identification of the individual stands. There is a large potential for savings if laser data and image data could be collected simultaneously, and stand delineation and characteristics usable for stratification could be derived from existing auxiliary data and automated methods. Combined use of digital image data and laser data in automated segmentation procedures (e.g., Ziegler et al., 2000) should be considered to take full advantage of the structural properties inherent in laser data.

Acknowledgments

This research has been funded by the Research Council of Norway, and is a contribution to research project no. 133303/110. I wish to thank Jo Heringstad and Petter Økseter at the Agricultural University of Norway who carried out some of the fieldwork. Thanks also to Fotonor, Norway, for collection and processing of the airborne laser scanner data.

References

Aldred, A. H., & Bonnor, G. M. (1985). Application of airborne laser to forest surveys. Information Report PI-X-51, Canadian Forest Service, Petawawa National Forestry Institute, Chalk River, 62 pp.

Anon. (1989). SAS/STAT user's guide, version 6, 4th edition, volume 2. SAS Institute, Cary, NC, 846 pp.

Anon. (1999). Pinnacle user's manual. Javad Positioning Systems, San Jose, CA, 123 pp.

Baltsavias, E. P. (1999). Airborne laser scanning: existing systems and firms and other resources. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 164–198.

Blingsmo, K. R. (1984). Diameter increment functions for stands of birch, Scots pine and Norway spruce. Research Paper 7/84, *Norwegian Institute for Forest Research*, Ås, 22 pp.

Braastad, H. (1975). Yield tables and growth models for *Picea abies*. *Reports of the Norwegian Institute for Forest Research*, *31*, 358–537.

Braastad, H. (1980). Growth model computer program for *Pinus sylvestris*. *Reports of the Norwegian Institute for Forest Research*, *35*, 265–359.

Delbeck, K. (1965). Methods for calculation of increment for open forests. *Tidsskrift for Skogbruk*, *73* (1), 5–45.

Eid, T. (1996). Kontroll av skogbruksplandata fra "Understøttet fototakst," Aktuelt fra Skogforsk no. 8-96, Department of Forest Sciences, Agricultural University of Norway, Ås, 21 pp.

Eid, T., & Næsset, E. (1998). Determination of stand volume in practical forest inventories based on field measurements and photo-interpretation: the Norwegian experience. *Scandinavian Journal of Forest Research*, *13*, 246–254.

Eid, T., & Nersten, S. (1996). Problemer omkring registreringer og planlegging for en skogeiendom i Birkenes kommune, Commun. Skogforsk no. 47.12, Dept. of Forest Sciences, Agric. Univ. of Norway, Ås, 56 pp.

Flewelling, J. W., & Pienaar, L. V. (1981). Multiplicative regressions with lognormal errors. *Forest Science*, *27*, 281–289.

Goldberger, A. S. (1968). The interpretation and estimation of Cobb–Douglas functions. *Econometrica*, *35*, 464–472.

Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*, 193–203.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*, 339–361.

Lefsky, M. A., Harding, D., Cohen, W. B., Parker, G., & Shugart, H. H. (1999). Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sensing of Environment*, *67*, 83–98.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, *28*, 1016–1031.

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

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

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., Harmon, M. E., & McKee, W. A. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western Cascades of Oregon. *Remote Sensing of Environment*, *67*, 298–308.

Næsset, E. (1995). Determination of mean diameter by basal area in stands of *Picea abies* and *Pinus sylvestris* in southeastern Norway by means of aerial photographs. *Scandinavian Journal of Forest Research*, *10*, 296–304.

Næsset, E. (1996). Determination of number of stems in coniferous forest stands by means of aerial photo-interpretation. *Scandinavian Journal of Forest Research*, *11*, 76–84.

Næsset, E. (1997a). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*, 49–56.

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

Næsset, E. (2001). Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. *Photogrammetric Engineering and Remote Sensing*, (in press).

Næsset, E., & Bjerknes, K.-O. (2001). Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sensing of Environment*, (in press).

Næsset, E., & Økland, T. (2001). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, (in press).

Nelson, R. (1997). Modeling forest canopy heights: the effects of canopy shape. *Remote Sensing of Environment*, *60*, 327–334.

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*, 247–267.

Nelson, R., Oderwald, R., & Gregoire, G. (1997). Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. *Remote Sensing of Environment*, *60*, 311–326.

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

Opseth, P. E. (1996). Flybåren laserscanning. En test av Optech's airborne laser terrain mapping system (Airborne laser scanning. Testing the Optech airborne laser terrain mapping system). MS Thesis, Department of Surveying, Agricultural University of Norway, Ås, 64 pp.

Tveite, B. (1977). Site index curves for Norway spruce (*Picea abies* (L.) Karst.). *Reports of the Norwegian Institute for Forest Research*, *33*, 1–84.

Weisberg, S. (1985). *Applied linear regression* (2nd ed.). New York: Wiley (324 pp.).

Ziegler, M., Konrad, H., Hofrichter, J., Wimmer, A., Ruppert, G., Schardt, M., & Hyyppä, J. (2000). Assessment of forest attributes and single-tree segmentation by means of laser scanning. *Proceedings of the International Society for Optical Engineering*, *4035*, 73–84.