

# ESTIMATION OF ABOVE- AND BELOW-GROUND BIOMASS IN BOREAL FOREST ECOSYSTEMS

E. Næsset

Agricultural University of Norway, Department of Ecology and Natural Resource Management, P.O. Box 5003, N-1432 Ås, Norway – erik.naesset@ina.nlh.no

**KEY WORDS:** Airborne laser scanning, biomass, forest inventory

## ABSTRACT:

Regression models for above-ground and below-ground biomass were estimated for 143 sample plots in young and mature coniferous forest. Various canopy height and canopy density metrics derived from the canopy height distributions of first as well as last pulse laser scanner data with a sampling density of approximately  $1.1 \text{ m}^{-2}$  were used as independent variables in the regressions. Each of the selected models comprised at least one variable related to canopy height and one related to canopy density. The models for above-ground biomass explained 92% of the variability whereas the models for below-ground biomass explained 86%. The analysis indicated that forest type did not have any significant impact on the estimated models.

## 1. INTRODUCTION

Ever since the very first experiments with airborne laser data in forest inventory around 1980 (e.g. Nelson et al., 1988), above-ground biomass has been one of the major response variables to be derived from the laser data. Airborne lasers have proved to be an efficient technology for precise estimation of above-ground biomass and carbon stocks. Both profiling (Nelson et al., 1988; Nelson et al., 2003b) and scanning (Lim et al., 2003) systems have been used, as well as discrete return small-footprint (Lim et al., 2003; Nelson et al., 2003b) and full waveform large-footprint data (e.g. Means et al., 1999; Lefsky et al., 1999). However, much of the research has been conducted in temperate forests and in ecosystems with dominance of deciduous tree species, and except for a few studies using profiling systems, the biomass studies have been limited to certain very local test sites. As there is an increased interest in carbon accounting in forest ecosystems, there is a need for efficient methods to estimate the above- as well as the below-ground biomass components of the trees.

Previous research has indicated that biophysical variables such as stem volume and basal area are estimated with a higher precision in coniferous forest than in deciduous forest. The objective of the present study was to estimate models for various biomass components of small sample plots in different boreal forest types using small-footprint data from a scanning laser with moderate sampling density, and to assess whether these models differ between forest types.

## 2. MATERIALS AND METHODS

### 2.1 Study area

This study was based on data from a forest inventory in southeast Norway conducted in the municipality of Våler (59°30'N, 10°55'E, 70–120 m a.s.l.). The size of the inventory was approximately 1000 ha. The main tree species were Norway spruce [*Picea abies* (L.) Karst.] and Scots pine (*Pinus*

*sylvestris* L.). Further details about the study area can be found in Næsset (2002b).

### 2.2 Field data

Field data were collected during summer of 1998 (see Næsset, 2002b). In total, 143 circular sample plots were distributed systematically throughout the entire 1000 ha study area according to a regular grid. The plots were pre-stratified according to age and site quality. Three strata were defined; “stratum I” comprised young forest of spruce and pine on poor as well as good sites, “stratum II” comprised mature forest, mainly pine, on poor sites and “stratum III” comprised mature forest, mainly spruce, on good sites (see Table 1). The size of each plot was  $300 \text{ m}^2$  in stratum I and  $400 \text{ m}^2$  in strata II-III.

On each plot, all trees with diameter at breast height ( $d_{bh}$ )  $>4$  and  $>10$  cm were callipered on young and mature plots, respectively, which conforms to ordinary inventory practice in Norway. The heights of sample trees selected with probability proportional to stem basal area at breast height using a relascope were measured by a Vertex hypsometer.

Total above- and below-ground biomass were calculated according to allometric equations for various biomass components of spruce, pine and birch. Below-ground biomass ( $B_b$ ) was calculated by separate equations for roots thinner and thicker than 5 cm, respectively, using  $d_{bh}$  only (Marklund, 1988) or  $d_{bh}$  and tree height ( $h$ ) as independent variables (Mälikönen, 1977). Above-ground biomass ( $B_a$ ) was calculated as the sum of biomass of stump, stem, bark, dead and living branches, and foliage according to  $d_{bh}$  and  $h$  (Mälikönen, 1977; Marklund, 1988). A summary of the ground-truth plot data is displayed in Table 1.

Differential GPS and Global Navigation Satellite System (GLONASS) were used to determine the position of the centre of each sample plot. Dual-frequency receivers observing pseudorange and carrier phase of both GPS and GLONASS were used as rover and base receivers. The estimated 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. Further details are given by Næsset (2002a).

Characteristic	Range		Mean
<i>Young forest – stratum I (n=56)</i>			
$h_L$ (m)	6.5	- 21.2	13.5
$V$ ( $m^3 ha^{-1}$ )	36.3	- 460.1	177.1
$B_a$ ( $Mg ha^{-1}$ )	24.7	- 248.2	107.0
$B_b$ ( $Mg ha^{-1}$ )	4.7	- 43.5	18.7
Tree species distribution			
Spruce (%)	0	- 100	52
Pine (%)	0	- 97	34
Deciduous (%)	0	- 69	14
<i>Mature forest, poor site quality – stratum II (n=36)</i>			
$h_L$ (m)	11.4	- 21.5	16.2
$V$ ( $m^3 ha^{-1}$ )	56.4	- 273.8	150.1
$B_a$ ( $Mg ha^{-1}$ )	37.1	- 147.1	87.6
$B_b$ ( $Mg ha^{-1}$ )	6.8	- 25.2	17.1
Tree species distribution			
Spruce (%)	0	- 100	31
Pine (%)	0	- 100	64
Deciduous (%)	0	- 21	5
<i>Mature forest, good site quality – stratum III (n=51)</i>			
$h_L$ (m)	11.4	- 25.9	20.1
$V$ ( $m^3 ha^{-1}$ )	93.0	- 555.9	267.9
$B_a$ ( $Mg ha^{-1}$ )	51.3	- 302.0	154.1
$B_b$ ( $Mg ha^{-1}$ )	10.8	- 49.6	25.8
Tree species distribution			
Spruce (%)	0	- 100	67
Pine (%)	0	- 100	25
Deciduous (%)	0	- 49	8

<sup>a</sup>  $h_L$ =Lorey's mean height,  $V$ =total stem volume,  $B_a$ =total above-ground biomass,  $B_b$ =total below-ground biomass.

Table 1. Summary of sample plot reference data <sup>a</sup>

### 2.3 Laser scanner data

Laser scanner data were acquired 8 and 9 June 1999 (leaf-on canopy conditions). A Piper PA31-310 aircraft carried the ALTM 1210 laser scanner system. The pulse repetition frequency was 10 kHz. The average flying altitude was approximately 700 m a.g.l., and the average footprint density was  $1.1 m^{-2}$ . According to standard procedures, pulses that were transmitted at scan angles that exceeded  $14^\circ$  were excluded from the final datasets.

Processing of the laser data was accomplished by the contractor (BN Mapping, Norway). Planimetric coordinates ( $x$  and  $y$ ) and ellipsoidal height values were computed for all first and last returns. The last return data were used to model the ground surface. In a filtering operation on the last return data undertaken by the contractor using a proprietary routine, local maxima assumed to represent vegetation hits were discarded. A triangulated irregular network (TIN) was generated from the planimetric coordinates and corresponding height values of the individual terrain ground points retained in the last pulse dataset. Both the original first and last return observations (points) were spatially registered to the TIN according to their coordinates. Terrain surface height values were computed for each point by linear interpolation from the TIN. The relative height of each point was computed as the difference between

the height of the first or last return, respectively, and the terrain surface height.

Observations with a height value less than 2 m above the TIN surface were excluded from both the first and last pulse datasets to eliminate ground hits and the effect of stones, shrubs, etc. from the tree canopy datasets (Nilsson, 1996). Pulses that hit outside the sample plots were excluded from further analysis.

### 2.4 Computations

Height distributions were created separately for first and last pulses classified as canopy hits ( $>2$  m, see above) for each plot inventoried in field, and percentiles for the canopy height for 10%, 20%, ..., 90% were computed. I also computed the mean values and the coefficient of variation of the first and last pulse height distributions. Furthermore, several measures of canopy density were derived. The range between the lowest laser canopy height ( $>2$  m) and the maximum canopy height was divided into 10 fractions of equal length. Canopy densities were then computed for the first and last pulse data separately as the proportions of laser hits above fraction #0 ( $>2$  m), 1, ..., 9 to total number of pulses.

The laser-derived variables were related to the ground-truth values of above- and below-ground biomass using regression analysis.

The following multiplicative regression model was selected based on findings in other studies (Lim et al., 2003):

$$Y = \beta_0 h_{0f}^{\beta_1} h_{10f}^{\beta_2} \dots h_{90f}^{\beta_{10}} h_{0l}^{\beta_{11}} h_{10l}^{\beta_{12}} \dots h_{90l}^{\beta_{20}} h_{\text{meanf}}^{\beta_{21}} h_{\text{meanl}}^{\beta_{22}} \times h_{\text{cvf}}^{\beta_{23}} h_{\text{cvl}}^{\beta_{24}} d_{0f}^{\beta_{25}} d_{1f}^{\beta_{26}} \dots d_{9f}^{\beta_{34}} d_{0l}^{\beta_{35}} d_{1l}^{\beta_{36}} \dots d_{9l}^{\beta_{44}} \quad (1)$$

The linear form used in the estimation was

$$\begin{aligned} \ln Y = & \ln \beta_0 + \beta_1 \ln h_{0f} + \beta_2 \ln h_{10f} + \dots + \beta_{10} \ln h_{90f} \\ & + \beta_{11} \ln h_{0l} + \beta_{12} \ln h_{10l} + \dots + \beta_{20} \ln h_{90l} \\ & + \beta_{21} \ln h_{\text{meanf}} + \beta_{22} \ln h_{\text{meanl}} + \beta_{23} \ln h_{\text{cvf}} + \beta_{24} \ln h_{\text{cvl}} \\ & + \beta_{25} \ln d_{0f} + \beta_{26} \ln d_{1f} + \dots + \beta_{34} \ln d_{9f} \\ & + \beta_{35} \ln d_{0l} + \beta_{36} \ln d_{1l} + \dots + \beta_{44} \ln d_{9l} \end{aligned} \quad (2)$$

where  $Y$ =field values of  $B_a$  or  $B_b$  ( $Mg ha^{-1}$ );  $h_{0f}$ ,  $h_{10f}$ , ...,  $h_{90f}$ =percentiles of the first pulse laser canopy heights for 0%, 10%, ..., 90% (m);  $h_{0l}$ ,  $h_{10l}$ , ...,  $h_{90l}$ =percentiles of the last pulse canopy heights for 0%, 10%, ..., 90% (m);  $h_{\text{meanf}}$ ,  $h_{\text{meanl}}$ =mean of the first and last pulse laser canopy heights (m);  $h_{\text{cvf}}$ ,  $h_{\text{cvl}}$ =coefficient of variation of the first and last pulse laser canopy heights (%);  $d_{0f}$ ,  $d_{1f}$ , ...,  $d_{9f}$ =canopy densities corresponding to the proportions of first pulse laser hits above fraction # 0, 1, ..., 9 to total number of first pulses;  $d_{0l}$ ,  $d_{1l}$ , ...,  $d_{9l}$  = canopy densities corresponding to the proportions of last pulse laser hits above fraction # 0, 1, ..., 9 to total number of last pulses.

Separate regression models were developed for above- and below-ground biomass, but data from all strata were included in each of the two models. Stepwise selection was used to select the variables to be included in these models. The standard OLS method was used (Anon., 1989). No explanatory variable was

left in the models with a partial  $F$  statistic with a significance level greater than 0.05. Multicollinearity issues were addressed by calculating and monitoring the size of the condition number ( $\kappa$ ). Models with  $\kappa > 30$  were not accepted, indicating that there was no serious collinearity inherent in the selected models (Weisberg, 1985).

To evaluate the statistical effects of forest type on the regressions, dummy variables were included in the selected models. Two dummy variables, DUMMY1 and DUMMY2, were applied in each model to identify the effect of each of the three strata. Their values were DUMMY1=0 and DUMMY2=0 if stratum I, DUMMY1=1 and DUMMY2=0 if stratum II, and DUMMY1=0 and DUMMY2=1 if stratum III.

### 3. RESULTS

First, linear regression models with log-transformed variables were estimated without including the dummy variables. The model for above-ground biomass explained 92% of the variability whereas the model for below-ground biomass explained 86% (Table 2). None of the selected models comprised more than three independent variables. Both models contained at least one variable related to canopy height ( $h_{601}$  and  $h_{\text{meanf}}$ ) and one variable related to canopy density ( $d_{1f}$  and  $d_{11}$ ), and variables derived from both first and last pulse data were represented among the selected ones.

It was revealed that including the two dummy variables did not affect the models significantly. The parameter estimates for the two dummy variables were far from statistically significant in both models ( $p=0.3-0.6$ ).

Independent variable	Model			
	$\ln B_a$	$\ln B_a$ with dummy	$\ln B_b$	$\ln B_b$ with dummy
$\beta_0$	1.94***	1.84***	0.55***	0.55**
$\ln h_{601}$			1.12***	1.13***
$\ln h_{\text{meanf}}$	1.32***	1.36***		
$\ln d_{1f}$	0.31*	0.28 NS		
$\ln d_{11}$	0.48***	0.48***	0.59***	0.62***
dummy1		-0.03 NS		0.03 NS
dummy2		-0.04 NS		-0.02 NS
$R^2$	0.92	0.92	0.86	0.86
RMSE	0.14	0.14	0.17	0.17
$\kappa$	5.33	6.16	1.21	3.20

<sup>a</sup> Level of significance: NS = not significant ( $>0.05$ ); \* $< 0.05$ ; \*\* $< 0.01$ ; \*\*\* $< 0.001$ .

<sup>b</sup>  $n=143$  sample plots.

Table 2. Estimated regression models for above-ground ( $B_a$ ) and below-ground ( $B_b$ ) biomass with and without dummy variables (DUMMY1 and DUMMY2) separating the three strata <sup>a,b</sup>

### 4. DISCUSSION

The results of this study indicate that total above-ground biomass can be estimated very precisely in coniferous forest using small-footprint laser scanner data. Using the same laser dataset as in the present study and plots with corresponding plot centers, but smaller plot size (200 m<sup>2</sup>), Næsset (2002a) estimated similar regression models for stem volume. In the

three strata in question, it was found that 93%, 80% and 80% of the variability in stem volume was explained in the three strata in question, respectively. In the present study, 92% of the variability was explained in a single model covering all forest types. This may indicate that total above-ground biomass can be estimated with a similar or even higher precision than stem volume. If so, it is not surprising, since the laser actually measures the canopy – including foliage and branches – which are significant components of the above-ground biomass, whereas the stem is not measured at all.

Based on the experience gained by estimating stem volume from laser data, it was somewhat surprising that the same models seemed to be the most appropriate ones regardless of forest type. However, the main reason why separate models for stem volume often is required for different forest types, is that the relationship between stem form and crown properties (size and shape) differs between forest types (e.g. age classes and species), which means that the laser is not able to capture data relevant to account for the variability between forest types. On the other hand, since the laser measures the canopy, regardless of its shape and density, it actually captures relevant data to derive the biomass components of especially foliage and branches, which means that differences between forest types to a larger degree are inherent in the measured data when it comes to biomass.

There are not many studies of biomass estimation from small-footprint laser data in boreal forest which these results can be compared with. In one of the pioneering studies of airborne laser in forestry, Nelson et al. (1988) estimated regression models for biomass in a pine forest in Georgia, United States. The best model they estimated explained 53% of the variability. However, they used a profiling system and short ground transects (20 m) to establish a field dataset. The results of their study are hardly comparable to later studies based on scanning systems with high sampling density and two or more pulse returns.

In two recent studies, Lim et al. (2003) and Nelson et al. (2003b) estimated above-ground biomass equations from small-footprint scanning and profiling data, respectively. The first of these studies dealt with hardwood forests in Canada and the other one with both hardwoods and conifers in Delaware, United States. The highest proportion of explained variability in the log-linear models estimated by Lim et al. (2003) was 85%, which is somewhat lower than revealed in the present study. The highest proportion of explained variability reported by Nelson et al. (2003b) was 71%, but it is stated that conifer biomass predictions are significantly more accurate than hardwood estimates (Nelson et al., 2003a).

To the very best of my knowledge, the present study is the first one to report models estimated for below-ground biomass, i.e. fine ( $<5$  cm) and coarse ( $>5$  cm) roots. The results indicate that the model fitted the data very well. The ground-truth biomass values used to calibrate the regression models were computed according to standard allometric equations of the same model form as those commonly used for various above-ground components. The  $R^2$ -values reported for the applied equations for roots range between 0.94 and 0.99 (Mälikönen, 1977; Marklund, 1988), so there is no reason to expect the below-ground biomass models estimated in the present study to be less valid than the corresponding above-ground models, unless the allometric equations were based on data from sites with little general validity.

To conclude, the present study has demonstrated that above- and below-ground biomass models can be estimated from laser scanner data in coniferous forest with an accuracy which is comparable or even better than what is expected in hardwood forests. Biomass equations also seem to be less influenced by forest type than, for example, stem volume. Further research should be carried out to confirm the validity of these results in boreal forest, and to assess the precision of predictions using independent testing. The latter task is of great importance for future applications of laser scanning in monitoring of biomass and carbon stocks.

Nelson, R., Valenti, M.A., Short, A. and Keller, C., 2003b. A multiple resource inventory of Delaware using airborne laser data. *BioScience*, 53: 981-992.

Nilsson, M., 1996. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sens. Environ.*, 56: 1-7.

Weisberg, S., 1985. *Applied linear regression (2<sup>nd</sup> ed.)*. Wiley, New York, 324 pp.

## REFERENCES

Anon., 1989. *SAS/STAT User's guide, version 6, 4<sup>th</sup> edition, volume 2*, SAS Institute, Cary, NC, 846 pp.

Lefsky, M. A., Harding, D., Cohen, W. B., Parker, G. and Shugart, H.H., 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sens. Environ.*, 67: 83-98.

Lim, K., Treitz, P., Baldwin, K., Morrison, I. and Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Can. J. Remote Sensing*, 29: 658-678.

Mälkönen, E., 1977. Annual primary production and nutrient cycle in birch stands. *Comm. Inst. For. Fenniae*, 91(5): 1-35.

Marklund, L.G., 1988. *Biomass functions for pine, spruce and birch in Sweden*. Dep. of Forest Surveying, Swedish Univ. of Agric. Sciences, Umeå, Report 45, 73 pp.

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., Harmon, M. E. and McKee, W. A., 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sens. Environ.*, 67: 298-308.

Næsset, E. 2002a. 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., 2002b. Determination of mean tree height of forest stands by means of digital photogrammetry. *Scand. J. For. Res.*, 17: 446-459.

Nelson, R., Krabill, W. and Tonelli, J., 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sens. Environ.*, 24: 247-267.

Nelson, R., Short, A. and Valenti, M., 2003a. *Measuring biomass and carbon in Delaware using an airborne profiling lidar*. In: J. Hyypä, E. Næsset, H. Olsson, T. Granquist Palen and H. Reese (Editors), Scandlaser Scientific Workshop on Airborne laser scanning of forests. Working Paper 112. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics, Umeå Sweden, pp. 134-144.