COMPARISON BETWEEN AN AREA-BASED AND INDIVIDUAL TREE DETECTION METHOD FOR LOW-PULSE DENSITY ALS-BASED FOREST INVENTORY

M. Vastaranta [a,]*, M. Holopainen [a], R. Haapanen [b], X. Yu [c], T. Melkas [d], J. Hyyppä [c], H. Hyyppä [e]

[a] University of Helsinki, Department of Forest Resource Management, Finland -(mikko.vastaranta, markus.holopainen)@helsinki.fi
[b] Haapanen Forest Consulting -reija.haapanen@haapanenforestconsulting.fi
[c] Finnish Geodetic Institute -(xiaowei.yu, juha.hyyppa)@fgi.fi
[d] Metsäteho Ltd -timo.melkas@metsateho.fi
[e]Helsinki University of Technology, Research Institute of Modelling and Measuring for the Built Environment, Finland - hannu.hyyppa@tkk.fi

KEY WORDS: Forest inventory, Laser scanning, k-NN, Individual tree detection, Accuracy

ABSTRACT:

Airborne laser scanning (ALS)-based forest inventories are being increasingly used. The two main approaches in deriving forest information from small-footprint laser scanner data are the area-based and individual tree detection (ITD) methods. In the present study we test the accuracies of an area-based k-nearest neighbour (k-NN)-method and ITD with a practical low-pulse density ($1.8/m^2$) ALS dataset at the plot level. The research material consists of 333 treewise measured circular plots from southern Finland. A test dataset of 97 plots was selected for stand characteristic accuracy observation. The root-mean-squared errors (RMSEs) for basal area, total volume, mean height and mean diameter were with ITD 26.7%, 26.9%, 8.2% and 12.2% and with k-NN 23.1%, 23.7%, 12.4% and 14.8%, respectively. The results obtained in this study demonstrated that both ALS-based inventory methods are practical, even with low-pulse density data. A Combination method could be developed to utilize the strengths of both methods and should be further investigated.

1. INTRODUCTION

Traditionally, forest management planning in Finland has been based on standwise field inventory. Currently, it is being replaced by airborne laser scanning (ALS)-based inventory. ALS is the most accurate remote-sensing (RS) technique for forest inventory, providing relative accuracies ranging between 10% and 20% at the stand level (Hyyppä and Hyyppä 1999, 2001; Hyyppä et al. 2000). The current data acquisition cost is comparable to that of the traditionally used inventory method. ALS devices providing small-footprint diameters (10–30 cm) allow accurate height determination of the forest canopy (e.g. Næsset 1997; Magnussen and Boudewyn 1998; Magnussen et al. 1999; Means et al. 1999). The two main approaches in deriving forest information from small-footprint laser scanner data have been an area-based method (Naesset 1997) and individual tree detection (ITD) (Hyyppä and Inkinen 1999). In the former method, percentiles of the distribution of laser canopy heights are used and the estimation is based on regression or nonparametric estimation methods. For example, Næsset (2002), Lim et al. (2003), Holmgren and Persson (2004) and Maltamo et al. (2006) showed that this approach produces highly reliable estimates of stand-level variables. By increasing the number of laser pulses per square metre, individual trees can be recognized and measured directly from laser data (e.g. Hyyppä and Inkinen 1999; Persson et al. 2002; Leckie et al. 2003; Popescu et al. 2003). However, although tree species can be identified with about 80-90% accuracy for individual trees (Persson et al. 2002, Holmgren and Persson 2004), using very-high-resolution imagery and ALS data, the operational is still missing. Tree species classification is seen as weakness of ALS-based inventories and the practical solution is needed.

Due to its lower pulse density, the area-based method is considered the more cost-efficient approach, although it needs large amounts of expensive fieldwork, compared with the ITD to perform accurately. The ITD approach provides the means for assessing the stand diameter or height distribution which is, in turn, invaluable in forest planning-related simulation and optimization, logging operation planning and wood supply logistics. The ITD method needs only a small ground truth dataset to calibrate ALS-based tree measurements.

In previous studies, area-based- and ITD-inventories have been tested with different datasets: low- and high-pulse. However, from the practical point of view, it is important to test these methods with datasets that are cost-efficient to use in practice. In the present study we test the accuracies of the two main ALS-based inventory methods, ITD and area-based k-nearest neighbour (k-NN) estimation with a practical low-pulse density ALS dataset combined with same-date aerial photographs.

2. METHOD

2.1 Study area and data

The research material comprised of 333 treewise measured circular plots in an app. 2000-ha managed forested area located in the vicinity of Evo, Finland (61.19° N, 25.11° E). Field measurement data from fixed-radius (10 m) field plots were collected from the study area in 2007 and 2008. Sampling of

* Corresponding author.

the field plots was based on prestratification of existing stand inventory data. The plots were located with Trimble's GEOXM 2005 Global Positioning System (GPS) device (Trimble Navigation Ltd., Sunnyvale, CA, USA), and the locations were postprocessed with local base station data, resulting in an average error of app. 0.6 m. The following variables were measured of trees having a diameter-at-breast height (dbh) of over 5 cm: location, tree species, dbh, height, lower limit of living crown and crown width. The volumes were calculated with standard Finnish models (Laasasenaho 1982). The plot-level data were obtained by summing the tree data.

A test dataset (Table 1) of 97 plots was selected for accuracy observation. The remaining 264 plots were used as modelling data for the area-based method. The test data set included 17 plots that were clear-cut after the field measurements. Total volume estimation accuracy from those plots were observed separately and the ground truth for volume was measured with a harvester.

	min	max	mean	std
BA, m^2ha^{-1}	5.5	43.4	26.7	6.7
Vol, m^3ha^{-1}	47.5	530.5	245.9	84.9
Hg, m	10.0	37.9	19.4	4.8
Dg, cm	11.7	43.2	22.7	6.3

Table 1. Range, mean and standard deviation of the stands basal-area (BA), total volume (Vol), mean height (Hg) and mean diameter (Dg) in the test data (n=97).

The ALS data were acquired in midsummer 2006; the flying altitude was 1900 m. The density of the pulses returned within the field plots was $1.8/m^2$ (only, first (F), intermediate or last (L); $1.3/m^2$ if only the first pulses were considered). A digital terrain model (DTM) and consequently, heights above ground level, were computed by the data provider. Same-date aerial photographs were obtained with a digital camera, as well. The photographs were orthorectified, resampled to a pixel size of 0.5 m and mosaicked to a single image covering the entire area. Only the near-infrared (NIR), red (R) and green (G) bands were available.

In all 172 statistical and textural features were extracted from the RS material. The extraction window was generally 20 x 20 m, which was proved suitable in previous studies (e.g. Holopainen and Wang 1998). The features included means and standard deviations of spectral values and ALS height and intensity, Haralick textural features (Haralick et al. 1973; Haralick 1979) derived from spectral values, ALS height and intensity, and 'standard texture' referring to a set of averages and standard deviations of spectral values, ALS height and intensity calculated within a 32 x 32 pixel window. In the case of ALS, these were derived from the first pulse data only, from point data rasterized to 0.5-m pixel size. The Haralick textural features were computed from four directions: 0˚, 45˚, 90˚ and 135˚. Additionally, the height statistics for the first and last pulses were calculated as in Suvanto et al. (2005): mean and maximum height (h_{mean}, h_{max}), standard deviation and coefficient of variation of height (h_{std}, h_{cv}), heights where certain relative amounts of laser points had accumulated (p05-p95), as well as percentages of laser points accumulated at various relative heights (r05-r95). Only pulses exceeding a 2-m height limit were included to remove hits to ground vegetation and bushes. Finally, percentages of points over 2 m in height were added (F_{vege}, L_{vege}). The means and standard deviations of ALS height were included only once in the final dataset, where the total number of features was 172. All features were standardized to a mean of 0 and std of 1.

2.2 Individual tree detection method

The canopy height model (CHM) was computed as the difference between the digital surface model (DSM), representing the top of the crowns, and the DTM. The DSM of the crown was obtained by taking the highest value (z-value) of all laser hits within each pixel. The value for missing pixels was obtained, using Delauny triangulation and linear interpolation. Single tree-based segmentations were performed on the CHM images, using a minimum curvature-based region detector. During the segmentation processes, the tree crown shape and location of individual trees were determined. First, the CHM was filtered by a Gaussian filter and then minimum curvatures were calculated. The CHM image was then scaled, based on the computed minimum curvature, and local maxima were examined in a given neighbourhood and used as markers in the following marker-controlled watershed transformation for tree crown delineations. Each segment was considered to represent a single tree crown. Laser-based tree heights were obtained from the pixel with the highest CHM value within each segment. These heights were calibrated with heights measured in the field. A dbh was predicted for the trees, using the laser-based tree height and model developed by Kalliovirta and Tokola (2005) for boreal forest stands in southern Finland. Tree species were assumed to be known. In ITD, plot-level stand characteristics were computed from the ALS-based tree heights and estimated diameters.

2.3 Area-based k-NN

A reduced set of features was used in the estimation. 11 features were selected among the 172 ALS- and aerial photograph-based features, using genetic algorithms (see Holopainen et al. 2008), and included percentages of points over 2 m in height (from the first and last pulses, separately), heights at which 30% of the laser pulses had accumulated (first and last pulse separately), height where 90% of the first pulses were accumulated, mean height in a 32 x 32 pixel window, angular second moment 45˚ of intensity, local homogeneity 90˚ of height, average NIR, std of NIR of 64 blocks within the 32 x 32 pixel window, and std of G of 1024 blocks within the 32 x 32 pixel window.

The estimation method was k-NN, which has long been used in Finnish RS-aided forest inventory applications (e.g. Kilkki and Päivinen 1987; Muinonen and Tokola 1990; Tomppo 1991). The nearest neighbours were determined by calculating the Euclidean distances between the observations in the n-dimensional feature space. The number of nearest neighbours was set at 5.

2.4 Accuracy observation

The accuracy of the estimated plot-level stand characteristics was observed by calculating the root-mean-squared error (RMSE):

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

$$RMSE\% = 100 * \frac{RMSE}{\bar{y}} , \quad (2)$$

where n = number of plots
y_i = observed value for plot i
$\hat{y}_i$ = predicted value for plot i
$\bar{y}_i$ = observed mean of the variable in question.

The studied stand characteristics were basal area (BA), total volume (Vol), mean height (Hg) and mean diameter (Dg).

3. RESULTS

The accuracies of the ITD- and k-NN inventory methods were studied with low-pulse density ALS data. The plot-level stand characteristic estimation errors are presented in Table 2 for BA, Vol, Hg and Dg. Bias was calibrated using the field data. For basal area and total volume, k-NN gave more accurate estimates. The RMSE in the basal-area was 6.2 m^2ha^{-1}(23.1%) with k-NN and 7.1 m^2ha^{-1} (26.7%) with ITD. The RMSEs were for total volume 58.4 m^3ha^{-1} (23.7%) and 66.2 m^3ha^{-1} (26.9%), respectively. Less accurate estimates for total volume and basal area were expected with ITD than with k-NN. With low-pulse density data, large numbers of trees were missing from the ITD, resulting in unaccurate estimates, especially for sum characteristics.

		BA	Vol	Hg	Dg
RMSE	ITD	7.1	66.2	1.6	2.8
	k-NN	6.2	58.4	2.4	3.4
RMSE-%	ITD	26.7	26.9	8.2	12.2
	k-NN	23.1	23.7	12.4	14.8

Table 2. Accuracy of the stand characteristics estimated by ITD and k-NN (n=97).

The volume estimation errors are presented in Fig. 1 for both methods without calibration. The k-NN estimation for total volume and for basal area tend to overestimate stands with low volume and underestimate large-volume stands. Results obtained with k-NN could be enhanced if this trend were taken into account. In the present study bias from the estimates was calibrated with a constant. Calibration is vital, especially for ITD, if low-pulse density data are used. The ITD estimates for total volume and basal area are clearly underestimated without calibration.

The stand mean characteristics, height and diameter, were estimated slightly more accurately with the ITD. RMSEs were 1.6 m (8.2%) for mean height and 2.8cm (12.2%) for mean diameter. With k-NN accuracies were 2.4m (12.4%) and 3.4 cm (14.8%), respectively. The mean diameter estimation errors are presented in figure 2. Both methods tend to underestimate large-diameter stands. Although the models used to estimate the individual trees diameter in ITD were not accurate at the single-tree level (RMSE 40 mm), the plot-level mean diameters were estimated more accurately than with k-NN.

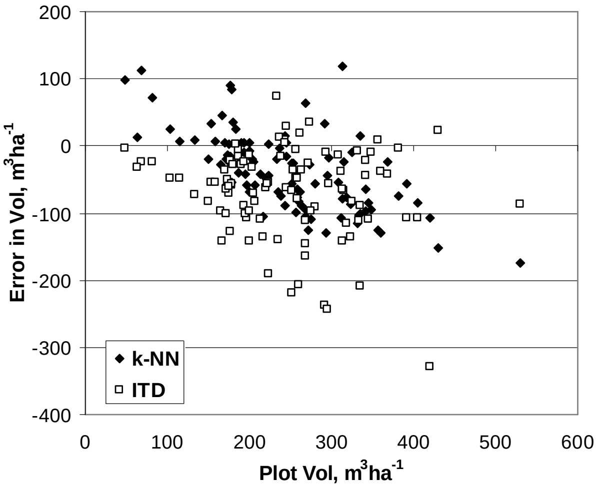

Figure 1. Volume estimation errors (Vol_{Est}-Vol_{Obs}) without bias calibration for both ALS-based inventory methods.

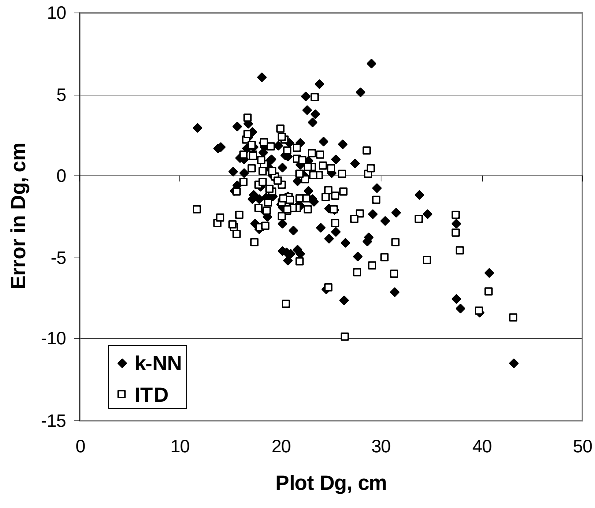

Figure 2. Mean diameter estimation errors (Dg_{Est}-Dg_{Obs}) without bias calibration for both ALS-based inventory methods.

3.1 Volumes measured with harvester

The accuracy of the estimated plot total volumes were studied separately with the 17 plots that were clearcut after the field measurements. In that case, the ground truth for volumes were measured with a harvester. This approach provided an opportunity to examine the accuracy of the field measurements besides the ITD- and k-NN-methods. In clear-cut plots the accuracy of the volume estimation is as accurate with ITD as with k-NN (Table 3). Relative estimation accuracies for total volume enhanced from 26.9% to 20.6% (ITD) and from 23.7% to 20.1% (k-NN) when harvester measurements were used as a ground truth. This can be explained mostly by greater mean volume in clearcut plots (288.4 m^3ha^{-1}) than in the whole data set (245.9 m^3ha^{-1}). Absolute errors enhanced remarkably only in ITD, which is assumed to perform best in mature stands.

		Vol
RMSE	Field	38.1
	k-NN	58.1
	ITD	59.4
RMSE-%	Field	13.2
	k-NN	20.1
	ITD	20.6

Table 3. Accuracy of the plot total volume (n=17).

Field measurements include errors that are seldom taken into account in this kind of studies. RMSE for total volume in the field measurements were 38.1 m^3ha^{-1} (13.2%), which is notable. When traditional field measurements are used as a ground truth, errors in it have an influence on the accuracy of the observed inventory methods.

4. CONCLUSIONS

Low-pulse density ALS data are used for forest inventories in practice; higher pulse density is seen as being too costly. Here we focused on testing two main methods for carrying out ALS-based forest inventory with practical low-pulse density data combined with aerial photographs.

Our results showed that although area-based k-NN inventory gave more accurate results for a total volume and basal area, the ITD method is as accurate in estimating mean stand characteristics, such as mean diameter and height. Our study provides ALS data-based accuracy estimates from a relatively heterogeneous area in southern Finland. In conclusion, we can say that the accuracies were in line with other Finnish studies using low-pulse density data (e.g. Suvanto et al. 2005; Maltamo et al. 2006; Packalén and Maltamo 2007; Packalén et al. 2008). From the practical point of view, our results achieved with ITD with low pulse density are promising and warrant further study.

In this study aerial photographs were used besides low-pulse density ALS data in area-based k-NN inventory. The main advantage of including them in the process comes from the improved accuracy of species-specific stand characteristics (Packalén and Maltamo 2007). Although these characteristics were not estimated in this study, they will be required in any operational ALS-based forest inventory application. Furthermore, aerial photographs are easy acquire and use.

Important differences among these methods include bias of the estimates and the amount of fieldwork needed. With low-pulse density data, the bias in the ITD method was notably greater in total volume and basal area than in the k-NN-method. A solution is needed in ITD to better discriminate individual trees in multistorey stands and in dense stands. In the present study, the ITD method discriminated an average of 65% of the trees in a plot, resulting in bias. Bias was calibrated from both methods with field measurements. The amount of fieldwork needed has a major negative effects on the cost-efficiency of inventory. Area-based methods require more expensive fieldwork than ITD. The quality of area-based inventory is dependent on the quality and amount of fieldwork, since ITD requires fieldwork only to calibrate the estimates. Area-based inventories are usually carried out with hundreds of reference plots, although the method may not require as much reference data if the measurements are directed carefully (e.g. Packalén et al. 2008). The feasibility for reducing the amount of fieldwork needed in both methods should be investigated.

We noted that the models (Kalliovirta and Tokola 2005) used in ITD to estimate dbh were not accurate at the single-tree level. In the present dataset the RMSE for a single tree's dbh was 40 mm. When both crown diameter and tree height were used, the RMSEs were even greater. Still, plot-level mean diameter were estimated more accurately with ITD than k-NN-method.

Accuracy in the field measurements were observed using plots that were clearcut and measured with a harvester. This observation showed that field measurements include errors that should be noticed and may have an influence on the accuracy of the evaluated inventory methods. However, our comparison was done in 17 mature stands, thus the results cannot be generalized to other stand development stages.

The results obtained here demonstrated that both ALS-based inventory methods are practical, even with low-pulse density data. A combination method could be developed to utilize the strengths of both methods and should be further investigated.

5. REFERENCES

Haralick, R. M., Shanmugan, K., and Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), pp. 610-621.

Haralick, R., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), pp. 786-804.

Holmgren, J. and Persson, Å., 2004. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90, pp. 415-423.

Holopainen, M., Haapanen, R., Tuominen, S. & Viitala, R., 2008. Performance of airborne laser scanning- and aerial photograph-based statistical and textural features in forest variable estimation. In Hill, R., Rossette, J. and Suárez, J. 2008. *Silvilaser 2008 proceedings*, pp.105-112.

Holopainen, M. and Wang, G., 1998. The calibration of digitized aerial photographs for forest stratification. *International Journal of Remote Sensing*, 19, pp. 677-696.

Hyyppä, H., and Hyyppä, J., 1999. Comparing the accuracy of laser scanner with other optical remote sensing data sources for stand attribute retrieval. *The Photogrammetric Journal of Finland*, 16, pp. 5-15.

Hyyppä, J. and Inkinen, M., 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland*, 16, pp 27-42.

Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S. and Zhu, Y-H., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, 128, pp. 109-120.

Hyyppä, J. and Hyyppä, H., (Eds.) 2001. High-Scan -Assessing forest stand attributes by integrated use of high-resolution satellite imagery and laserscanner. Contact No. ENV4-CT98-0747 of European Commission. Final Report. September 2001, 81pp.

Kalliovirta, J. and Tokola, T., 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. *Silva Fennica*, 39(2), pp. 227-248.

Kilkki, P., and Päivinen, R., 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. Department of Forest Mensuration and Management, University of Helsinki. Research notes, 19, pp. 210-215.

Laasasenaho, J., 1982. Taper curve and volume function for pine, spruce and birch. *Communicationes Instituti Forestalis Fenniae*, 108, pp. 1-74.

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L. and Shreenan, R. (2003) Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Forest Research*, 29(5), pp. 633–649.

Lim, K., Treitz, P., Baldwin, K., Morrison, I. and Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, 29, pp. 648-678.

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 pp. 1016–1031.

Magnussen, S., Eggermont, P. and LaRiccia, V. N., 1999. Recovering tree heights from airborne laser scanner data. *Forest Science*, 45, pp. 407-422.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A. and Kangas, J., 2006. Non-parametric estimation of stem volume using laser scanning, aerial photography and stand register data. *Canadian Journal of Forest Research*, 36, pp. 426-436.

Means, J., Acker, S., Harding, D., Blair, J., Lefsky, M., Cohen, W., Harmon, M. and McKee, 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, pp. 298-308.

Muinonen, E. and Tokola, T., 1990. An application of remote sensing for communal forest inventory. *Proceedings from SNS/IUFRO workshop: The usability of remote sensing for forest inventory and planning*, 26-28 February 1990, Umeå, Sweden. Remote Sensing Laboratory, Swedish University of Agricultural Sciences, Report 4, pp. 35-42.

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

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

Packalén, P. and Maltamo, M., 2007. The k-MSN method in the prediction of species specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sensing of Environment*, 109, pp. 328-341.

Packalén, P., Pitkänen, J., Maltamo, M., 2008. Comparison of individual tree detection and canopy height distribution approaches: a case study in Finland. In Hill, R., Rossette, J. and Suárez, J. 2008. *Silvilaser 2008 proceedings*, pp. 22-29.

Persson, Å., Holmgren, J. and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrametric engineering & Remote Sensing*, 68, pp. 925-932.

Popescu, S.C., Wynne, R.H. and Nelson, R.F., 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29, pp. 564-577.

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

Tomppo, E., 1991. Satellite image-based national forest inventory of Finland. *International Archives of Photogrammetry and Remote Sensing*, 28, pp. 419-424.

ACKNOWLEDGEMENTS

This study was made possible by financial aid from the Finnish Academy project Improving the Forest Supply Chain by means of Advanced Laser Measurements.