In: Wagner W., Székely, B. (eds.): ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS, Vol. XXXVIII, Part 7B

COMPARISON OF GRID-BASED AND SEGMENT-BASED ESTIMATION OF FOREST ATTRIBUTES USING AIRBORNE LASER SCANNING AND DIGITAL AERIAL IMAGERY

S. Tuominen [a, *], R. Haapanen [b]

[a] Finnish Forest Research Institute, Metsäntutkimuslaitos, PL 18, 01301 Vantaa, Finland - sakari.tuominen@metla.fi
[b] Haapanen Forest Consulting, Kärjenkoskentie 38, 64810 Vanhakylä, Finland - reija.haapanen@haapanenforestconsulting.fi

KEY WORDS: Forestry, LIDAR, Inventory, Photography, Segmentation

ABSTRACT:

Forest management planning in Finland is currently adopting a new-generation forest inventory method, which is based on interpretation of airborne laser scanning data and digital aerial images. The inventory method is based on systematic grid, where the grid elements serve as inventory units, for which the laser and aerial image data are extracted and, for which the forest variables are estimated. As an alternative or a complement to the grid elements, image segments can be used as inventory units. The image segments are particularly useful as the basis for generating the silvicultural treatment and cutting units, since their borderlines should follow the actual stand borders, whereas the grid elements typically cover parts of several forest stands. In this study we carried out an automatic segmentation of two study areas on the basis of laser and aerial image data with a view to delineating ecologically homogeneous micro-stands. Further, we extracted laser and aerial image features both for systematic grid elements and segments. For both units, the set of features used for estimating the forest attributes were selected using a genetic algorithm, which aims at minimizing the estimation error of the forest variables. The estimation accuracy produced by both approaches was assessed by comparing their estimation results. The preliminary results indicate that despite of the theoretical advantages of the image segments, the laser and aerial features extracted from grid elements seem to work better than features extracted from image segments in estimating forest attributes.

1. INTRODUCTION

In Finland, the forest inventory for the forest management planning has traditionally been based on visual inventory by stands. In this method, the forest stands that are delineated on the basis of aerial photographs and their growing stock and site-related characteristics are measured or estimated in field. The method is considered too labour-intensive, and it requires large amount of the fieldwork. Thus, the visual inventory method will be replaced by a new-generation forest inventory method, which is now at the pilot phase and which employs more remote sensing data and less fieldwork.

The new generation forest inventory method will be based on interpretation of airborne laser scanning (ALS) data and digital aerial imagery using field sample plots as a reference data. Statistically the new generation forest inventory method is based on two-phase sampling with stratification, where inventory database is based on systematic grid of sample units (i.e. grid elements as sample units), and the size of grid elements should correspond to the size of field plots. Field measurements are allocated into strata that are derived on the basis of earlier stand inventory data. The typical remote sensing data sources that are used in the new generation forest inventory system are low density ALS-data (typically 1–2 laser pulses/m^2) and color-infrared (CIR) digital aerial imagery with spatial resolution of approx. 0.5 m.

The inventory units in the new generation forest inventory method are typically defined by a systematic grid of sample units. The forest variables are estimated for each grid element that covers a square shaped area. As an alternative to the grid based approach, the use of automatic stand delineation has been studied for defining the inventory units. Automatically delineated stands (i.e. image segments) have an advantage compared to grid elements (e.g. Pekkarinen & Tuominen, 2003; Hyvönen et al., 2005). They can be delineated in such a way that they follow exactly the actual stand borders, whereas the grid elements are spatially "sparse" in relation to the actual borders of stands and other ecological units in forest, so they do not follow the borderlines accurately and they usually cover trees from more than one stand (e.g. Pekkarinen & Tuominen, 2003). On the other hand, grid elements are unambiguously defined by their coordinates and, thus, the same units can be used in consecutive inventories.

In delineating forest stands the primary input variables are the height of the trees and tree species composition (or dominancy). Stand density usually is a secondary parameter for stand delineation. The height of the trees can be derived on the basis the ALS-data but, on the other hand, ALS data with the applied pulse density does not serve well the purpose of the recognition of tree species. Thus, optical aerial imagery is needed for the estimation of the tree species composition.

The objective of this study was to find a suitable combination of laser and aerial data for automatic stand delineation and to test

* Corresponding author.

the laser and aerial image features extracted from grid elements and automatically delineated stand polygons in order to find out which unit serves better the purpose of extracting image features for estimating forest attributes.

2. MATERIAL AND METHODS

2.1 Study areas

The laser scanning and aerial image based estimation was tested in two study areas. Study area 1 was located in the municipality of Lammi in southern Finland and its field data consisted of 282 fixed-radius (9.77 m) circular field sample plots that were measured in 2007. Study area 2 was located in eastern Finland and 546 fixed radius (9 m) sample plots measured in 2009 were available as field reference data here. For the field sampling both study areas were stratified on the basis of earlier stand inventory data and the field sample plots were allocated to these strata in order to cover all types of forest in the study areas.

There was some variation between the forest characteristics of the two study areas. In study area I the total growing stock was more evenly distributed between tree species groups pine, spruce and deciduous trees, whereas the study area 2 was clearly dominated by spruce. Furthermore, the study area 2 had somewhat higher average stand volume, as well as larger distribution of sample plot volumes. The statistics of the study areas based on the sample plots are presented in table 1.

	Study area 1		**Study area 2**	
	Average	**Max**	**Average**	**Max**
Total volume, m3/ha	178.7	575.4	205.5	798.5
Volume of pine, m3/ha	69.8	560.6	52.7	561.8
Volume of spruce, m3/ha	63.7	575.4	109.3	739.2
Volume of deciduous, m3/ha	45.2	312.0	43.5	400.4
Height, m	17.0	30.5	17.3	35.2
Diameter, cm	21.1	50.2	21.9	60.3

Table 1. Forest statistics of the study areas (average and maximum values of sample plots)

2.2 Remote sensing data

In study area 1 the remote sensing data consisted of color-infrared digital aerial imagery (containing near-infrared, red and green bands) and ALS data acquired from a flying altitude of 1900 m with the density of 1.8 returned pulses per square meter. In study area 2 the remote sensing data consisted of color-infrared (containing near-infrared, red and green bands) and natural color (red, green and blue bands) digital aerial imagery and ALS data acquired from a flying altitude of 2000 m with the density of 0.6 returned pulses per square meter. Here the aerial image data was combined to a 4-band composite image containing blue, green, red and near-infrared bands.

The aerial images were ortho-rectified and resampled to a spatial resolution of 0.5 m. The ALS point data was also interpolated to a raster image format (height and intensity images) using second degree polynomial model. The output laser images had similar spatial resolution as the aerial images.

2.3 Automatic image segmentation

Automatic stand delineation was carried out in the study areas by automatic segmentation of aerial images and ALS data interpolated to raster format. The segmentation was carried out in two phases. In the first phase initial segmentation was done using a modified implementation of Narendra & Goldberg (1980) algorithm, which employs local edge gradient. This method typically produces a large number of small polygons, and the objective is to find all potential segment borders at this phase. In the second phase the initial segments were processed using a region merging algorithm that was guided by parameters such as desired minimum size of final segments and the similarity/dissimilarity of the segments to be merged (t-ratio threshold).

In this study, the initial segmentation was based entirely on laser height corresponding mainly to the stand height (Mustonen et al., 2008). The merging of initial segments into larger spatial units (i.e. final segments) was carried out on the basis of laser and aerial image data, taking into account also the tree species structure of the initial segments. Two automatic segmentations with minimum segment sizes of 350 m^2 and 0.1 ha were carried out in both study areas.

2.4 Extraction of laser and aerial image features

Three remote sensing feature data sets were extracted from each of the study areas. In these sets the remote sensing features were allocated to each sample plot from a square window or a segment in which the sample plot was located. The feature set Grid was extracted from a 20 x 20 meter square window centered around each sample plot. The feature set Seg350 was extracted from image segments, whose minimum size was set as 350 m^2. The feature set Seg1000 was extracted from image segments, whose minimum size was set as 0.1 ha.

The following statistical and textural features were extracted from the remotely sensed material for each feature data set:

- Means, standard deviations and Haralick textural features (Haralick et al. 1973, Haralick 1979) of spectral values of aerial photographs, ALS height and intensity (first pulse only)
- Height statistics for the first and last pulses of the points inside the field plot area or the segment area (Suvanto et al., 2005). These included mean, standard deviation (std), maximum, coefficient of variation, heights where certain percentages of points had accumulated and percentages of points accumulated at certain relative heights. Only points over 2 m in height were considered and the percentage of points over 2 m in height was included as a variable.
- A number of std's extracted from a 32 x 32 pixel window using block sizes from 1 to 8 pixels.

All features were standardized to a mean of 0 and std of 1.

2.5 Selection of features for the estimation of forest attributes

The *k*-nearest neighbor (*k*-nn) method was used to estimate forest variables (e.g. Kilkki & Päivinen, 1987; Tokola et al., 1996). . The value of *k* was set to 5, euclidean distances were used to measure closeness in the feature space and the nearest neighbors were weighted with the squared inverse distances.

The accuracy of the estimates produced by the k-nn estimator was tested via leave-one-out cross-validation on the field plots by comparing the estimates of each field plot to the measured value (ground truth) of the plot. The accuracy of the estimates was measured by the relative root mean square error RMSE (Equation 1).

$$RMSE\% = 100 * \frac{RMSE}{\bar{y}} \qquad (1)$$

where:

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

y_i = measured value of variable y on plot i
$\hat{y}_i$ = estimated value of variable y on plot i
$\bar{y}$ = mean of the observed values
n = number of plots

Automatic feature selection was carried out using a simple genetic algorithm presented by Goldberg (1989), and implemented in the GAlib C++ library (Wall 1996). The GA process starts by generating an initial population of strings (chromosomes or genomes), which consist of separate features (genes). The strings evolve during a user-defined number of iterations (generations). The evolution includes the following operations: selecting strings for mating using a user-defined objective criterion (the better the more copies in the mating pool), letting the strings in the mating pool to swap parts (crossing over), causing random noise (mutations) in the offspring (children), and passing the resulting strings into the next generation.

In the present study, the starting population consisted of 300 random feature combinations (genomes). The length of the genomes corresponded to the total number of features in each step, and the genomes contained a 0 or 1 at position *i*, denoting the absence or presence of image feature *i*. The number of generations was 30. The objective variable to be minimized during the process was a weighted combination of relative RMSEs of *k*-nn estimates for mean total volume, mean volumes of Scots pine, Norway spruce and deciduous species, mean diameter and mean height, with total volume having a weight of 50%, and the remaining variables 10% each. Genomes that were selected for mating swapped parts with each other with a probability of 80%, producing children. Occasional mutations (flipping 0 to 1 or vice versa) were added to the children (probability 1%). The strings were then passed to the next generation. The overall best genome of the current iteration was always passed to the next generation, as well. Four successive steps (all including 30 generations) were taken to reduce the number of features to a reasonable minimum. Only features belonging to the best genome in each step were included in the next step. Feature selection was run separately for both areas and each feature extraction unit (field plot, small segments, large segments).

There were 14 (study area 1) or 19 (study area 2) features selected into the final Grid sets, 17 or 19 into the Seg350 sets and 12 or 17 into the Seg1000 sets. Of the selected features, majority (63–79%) were based on the ALS data.

3. RESULTS AND DISCUSSION

In both study areas the features extracted from square grid elements worked better in estimating the forest attributes than the features extracted from image segments. Furthermore, features from image segments derived using minimum size of 350 m^2 performed better in the estimation than features extracted from larger segments (minimum size 0.1 ha).

Study area 2 had generally better estimation accuracy compared to data sets of study area 1. The main reason for this is probably the higher number of sample plots in study area 2, which gives higher number of potential nearest neighbors for each sample plot in the *k*-nn estimation. The estimation accuracy results for the forest attributes used in this study are presented in tables 2 and 3.

	GRID	SEG350	SEG1000
Height	18.5	22.4	25.5
Diameter	25.5	27.7	32.0
Total volume	27.8	34.0	36.6
Volume of pine	74.2	77.1	99.9
Volume of spruce	83.9	87.5	103.3
Volume of deciduous sp.	85.3	88.7	93.9

Table 2. Estimation results for the feature sets (relative RMSE, %) of study area 1

	GRID	SEG350	SEG1000
Height	12.5	13.9	16.5
Diameter	19.8	23.1	25.2
Total volume	29.6	32.9	36.6
Volume of pine	125.2	138.5	137.0
Volume of spruce	59.0	61.5	63.8
Volume of deciduous sp.	99.2	113.4	111.3

Table 3. Estimation results for the feature sets (relative RMSE, %) of study area 2

There were large differences between the study areas in the estimation accuracy of the volumes per tree species groups. Apparently, the differences were caused by the different tree species structure of the two study areas. Typically, the dominant tree species had the highest estimation accuracy, and the less dominant lowest. On the other hand, the volume of deciduous trees had better estimation accuracy compared to the minority coniferous tree species group since the presence of the deciduous trees is more easily recognizable in the aerial images.

Despite the theoretical advantages of the segment-based approach, the features extracted for segments did not perform well in the estimation procedure. There are some possible reasons for this. First, the field data was measured per sample plots and not per segments. Because of this the areas of the field measurement and the extracted remote sensing features correspond to each other best in the feature set Grid. Furthermore, the automatic segmentation often produces segments that are irregularly shaped, i.e. not compact, and in forest stands with large trees the segment borders are typically located in gaps between trees, in which case the variation within the segments may be more significant than the variation between segments. On the other hand, using geographically larger segments in extracting the features typically resulted in lower estimation accuracy compared to other feature sets, which indicates that the larger the units are the more internal variation they have.

Based on the results of this study the most feasible inventory procedure utilizing ALS and aerial image data seems to be the following: 1) estimation based on ALS data and aerial imagery for the systematic grid elements, 2) automatic segmentation utilizing ALS height, ALS intensity and aerial imagery, 3) deriving the estimates for image segments on the basis of the estimates of grid elements and 4) manual combination of image segments for deriving spatial units for forest management purposes.

4. REFERENCES

Haralick, R., 1979. Statistical and structural approaches to texture. *Procedings-IEEE*, 67(5), 786–804.

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

Hyvönen, P., Pekkarinen, A. and Tuominen, S., 2005. Segment-level stand inventory for forest management. *Scandinavian Journal of Forest Research* 20(1), 75–84

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*, 19:210–215.

Mustonen, J., Packalén, P. and Kangas, A., 2008. Automatic segmentation of forest stands using a canopy height model and aerial photography. *Scandinavian Journal of Forest Research* 23(6): 534–545.

Narendra, P. and Goldberg, M., 1980. Image segmentation with directed trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Pami-2: 185–191.

Pekkarinen, A. and Tuominen, S., 2003. Stratification of a forest area for multisource forest inventory by means of aerial photographs and image segmentation. In P. Corona, M. Köhl, & M. Marchetti (Eds.), *Advances in forest inventory for sustainable forest management and biodiversity monitoring. Forestry Sciences*, vol. 76. (pp. 111–123). Kluwer Academic Publishers.

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

Tokola, T., Pitkänen, J., Partinen, S. and Muinonen, E., 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing*, Vol. 17, No. 12, pp. 2333–2351.

ACKNOWLEDGEMENTS

The authors wish to thank M.Sc. Risto Viitala at the HAMK University of Applied Sciences and Lic.Sc. Juho Heikkilä at Forestry Development Centre Tapio for providing the field and remote sensing materials for this study.