Comparison of a grid-based and segment-based estimation of forest attributes using airborne laser scanning and digital aerial imagery

S. Tuominen[1] and R. Haapanen[2]

[1] Finnish Forest Research Institute, Metsäntutkimuslaitos, PL 18, 01301 Vantaa, Finland sakari.tuominen@metla.fi

[2] Haapanen Forest Consulting, Kärjenkoskentie 38, 64810 Vanhakylä, Finland
reija.haapanen@haapanenforestconsulting.fi

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

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), for which the forest variables are estimated. The size of grid elements should correspond to the size of field plots. In the inventoried area, field measurements are allocated to strata that are derived on the basis of earlier stand inventory data. 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 ground resolution of approx. 0.5 m.

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: they can be delineated in such a way that they follow exactly the actual stand borders. The grid elements, in turn, 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 do 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 photograph data for automatic stand delineation and to compare laser and aerial photograph features extracted from grid elements with automatically delineated stand polygons to find out which unit serves better the purpose of extracting image features for estimating forest attributes.

Material

The study area was located in the municipality of Lammi in southern Finland and it covered approx. 1800 ha of forest. The field data consisted of 282 fixed-radius (9.77 m) circular field sample plots that were measured in 2007. For field sampling the study area was stratified into 40 strata on the basis of earlier stand inventory data and the field plots were proportionally allocated to the strata.

The remote sensing data consisted of CIR digital aerial imagery (containing near-infrared, red and green bands) and ALS data. The aerial images were ortho-rectified and resampled to a spatial ground resolution of 0.5 m. The ALS data were acquired from a flying altitude of 1900 m, and the density of the returned pulses within the field plots was 1.8/m^2. The ALS point data were interpolated to a raster format (height and intensity images) using second degree polynomial model. The output laser images had spatial resolution similar to the aerial photographs.

The following features were extracted from ALS data and aerial photographs.

1. Means and standard deviations of spectral values of aerial photographs
2. Haralick textural features of aerial photographs derived from spectral values (Haralick et al. 1973, Haralick 1979)
3. ALS height and intensity
4. Haralick textural features of ALS height and intensity data interpolated to a raster format (Haralick et al. 1973, Haralick 1979)
5. ALS height statistics for the first and last pulses (F, L) (Suvanto et al. 2005):
   - mean and maximum height
   - standard deviation and coefficient of variation of height
   - heights where certain relative amounts of laser points had accumulated (p05-p95), percentages of laser points accumulated at various relative heights (r05-r95)
   - percentages of points over 2 m in height
6. Standard deviation of ALS and aerial photograph pixel values of blocks, into which the window was divided. The block sizes corresponded to 1x1, 2x2, 4x4, and 8x8 pixels.

The total number of laser and aerial photograph features in final dataset was 172. The full set of features was extracted to grid elements using 20 x 20 m window, except in the case of standard deviations of pixel blocks, for which a window of 32 x 32 pixels was used (16 x 16 m). Since the different features had very different scales of variation, all features were standardized to a mean of 0 and standard deviation of 1.

Methods

Automatic stand delineation by image segmentation

The automatic stand delineation was carried out by automatic segmentation of aerial photographs and ALS data interpolated to raster format. The segmentation was carried out in two phases. In the first phase an 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). The initial segmentation was based entirely on laser height (Mustonen et al. 2008). The merging of initial segments into larger spatial units was carried out on the basis of laser and aerial image data. ALS height and intensity and all aerial photograph bands were tested as input for merging the segments.

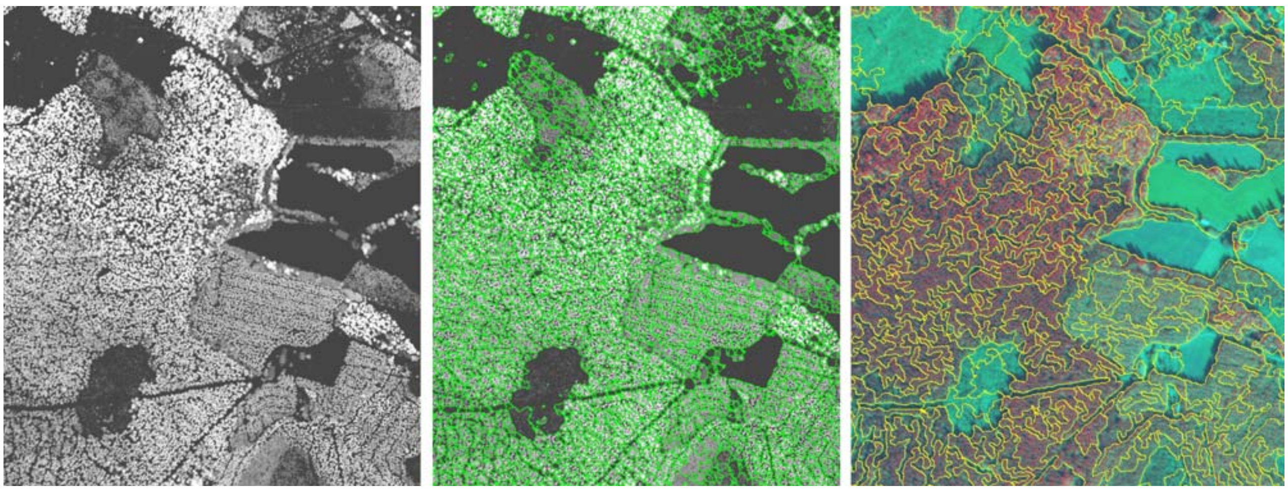

Figure 1. Initial segments (left) on laser height image and final segments on a combination of laser and aerial image channels

Estimation and feature selection

The estimation method used in this study was *k* nearest neighbors estimator (e.g. Kilkki & Päivinen 1987, Tokola et al. 1996). The nearest neighbors were determined by the Euclidean distances between the observations in the feature space defined by the applied image features, and the number of nearest neighbors was set at 5. Leave-one-out cross-validation was used for testing the estimates, and the accuracy of the estimates was measured by their root mean square errors (RMSE). The selection of the image features for the *k*-nn estimation procedure was carried out by a genetic algorithm. The feature selection procedure has the following phases:

1. Create an initial population of feature combinations at random.
2. Evaluate the fitness of each feature combination, by calculating value for objective variable using *k*-nn and cross-validation. The objective variable was a weighted combination of relative RMSEs of total volume, volume of pine, volume of spruce, volume of deciduous trees, diameter and height, with total volume having a weight of 50%, and the remaining variables 10% each.

3. Select desired percentage of the fittest feature combinations.
4. Let the best feature combinations change parts with each other to produce offspring.
5. Add mutations to the offspring.
6. Pass the offspring to the next generation.
7. Go back to step 2 and continue looping as long as desired amount of iterations have been performed.

The selection of image features by genetic algorithm resulted in a total of 11 features that were selected for the k-nn estimation. These 11 features were extracted for both grid elements and image segments. The forest variables that were estimated were the total volume of growing stock (m^3/ha), basal area (m^2/ha), mean height (m), diameter at breast height (DBH; cm), volumes of pine, spruce and deciduous trees (m^3/ha).

Results

The features extracted from grid elements worked better than features extracted from image segments in estimating forest attributes. Image segments derived using different input channels had no clear differences in the estimation. When assessing visually the different image segmentations, the image segments derived using laser height, laser intensity and aerial image NIR or NIR/R ratio produced the best automatic delineation of stands regarding their compactness, shape, correctness of borderlines. The relative RMSEs of the forest variable estimates that were derived using features extracted from different spatial units are presented in Table 1.

Table 1. Relative RMSEs (%) of estimated forest attributes using different units for extracting ALS and aerial image features

Feature extraction unit	volume	basal area	height	diameter	volume of pine	volume of spruce	volume of broadleaved species
GRID	30.8	25.6	18.2	24.9	88.5	86.7	87.5
SEG h	41.8	35.6	23.8	31.0	103.4	106.6	106.3
SEG h+i	42.9	35.5	24.3	32.6	112.8	108.7	110.1
SEG h+i+nir	46.1	38.1	23.4	30.8	109.3	115.1	109.9
SEG nir+r+g	43.4	35.8	25.0	32.4	109.4	110.9	115.8

GRID = features extracted for grid elements
SEG = features extracted for segments, segment region merging by: h = laser height, i = laser intensity, nir = aerial image NIR band, r = aerial image red band, g = aerial image green band

Conclusions

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. Second, the segments were larger than the size of grid elements, which may have caused more variation within the segments. Third, variation within segments may have been more significant than variation between the segments especially in stands with large trees (segment borders are typically located in gaps between trees).

Based on the results of this study the most feasible inventory procedure at present 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.

References

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

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

Kilkki, P., & R. Päivinen, 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.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J. & Vehmas, M. 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. Remote Sensing of Environment, 90: 319–330.

Mustonen, J., Packalén, P. & 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. & 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. & Kangas, J., 2005. Kuviokohtaisten puustotunnusten ennustaminen laserkeilauksella. Metsätieteen aikakauskirja, 4/2005: 413-428.

Tokola, T., Pitkänen, J., Partinen, S. & 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.