Performance of airborne laser scanning- and aerial photograph-based statistical and textural features in forest variable estimation

Markus Holopainen[1], Reija Haapanen[2], Sakari Tuominen[3] & Risto Viitala[4]

[1]University of Helsinki, Department of Forest Resource Management,
markus.holopainen@helsinki.fi
[2]Haapanen Forest Consulting reija.haapanen@haapanenforestconsulting.fi
[3]Finnish Forest Research Institute sakari.tuominen@metla.fi
[4]HAMK University of Applied Sciences risto.viitala@hamk.fi

Abstract

In the present study we tested the performance of different combinations of airborne laser scanning (ALS) and aerial photograph-based features in the estimation of forest variables. The combinations were subsets of a total of 172 features extracted from the remotely sensed material. The subsets were based on expert judgment or a genetic algorithm (GA). The non-parametric *k*-nearest neighbour (*k*-NN) algorithm was applied to derive the estimates. The best performing feature set was obtained after four consecutive steps of GA, each starting with the best features found in the previous step. The best set contained 11 features, 8 of them originating from the ALS data. This set was further weighted with a downhill simplex algorithm, and a relative mean volume RMSE of 27.1% was obtained. The results were slightly worse than in other Finnish ALS studies, most probably due to a larger amount of deciduous trees and greater variation of forests in the study area.

Keywords: k-NN, feature selection, genetic algorithm, species-specific estimates

1. Introduction

The main approaches to deriving forest information from small-footprint airborne laser scanning (ALS) data are plot-level estimation based on features derived from height information (e.g., Næsset 1997, Suvanto *et al*. 2005) and individual tree detection (e.g., Hyyppä and Inkinen 1999, Maltamo *et al*. 2004). The latter method is computationally heavier and requires greater pulse density; thus in large-area inventories the plot-level approach can, at least currently be considered more cost-efficient.

Estimation accuracies can typically be improved with a combination of data sources with complementary properties. Examples are datasets comprising of Landsat-type satellite images with good spectral resolution, and colour-infrared aerial photographs (Haapanen and Tuominen 2008) or even black-and-white photographs, with good spatial resolution (Tuominen and Haakana 2005).

High spatial resolution, a property of e.g., aerial photographs and ALS data, allows the use of two-dimensional (2D) textural features - even three-dimensional (3D) in the latter case. Generally, adding more features in the estimation process improves the output accuracy, but with increasing dimensionality the distinctive capacity of the data may weaken, with increasing noise. Therefore, the dimensionality of large datasets must be reduced. The usefulness of any input variable can be studied by measuring the correlation between the image features and forest attributes. In cases of large feature sets this is extremely tedious. Furthermore, the image features are often highly correlated, and adding additional variables having high correlation with the other variables does not generally improve the estimation accuracy (although it is still possible). Guyon and Elisseef (2003) showed that even a useless variable may be useful when taken with others, and two useless variables can be useful together. Thus, filters that rank features based on correlation coefficients are not sufficient and subset selection algorithms or feature transformation is needed. Principal component analysis is one example of feature transformation, while e.g., stepwise regression (backward or forward selection) or genetic algorithms (GAs) can be used to construct subsets of features. GAs are search algorithms that mimic natural selection and natural genetics (Goldberg 1989). Kudo and Sklansky (2000) compared several feature selection algorithms and concluded that sequential floating search methods worked best for small- and medium-scale problems, whereas for problems with a large number of dimensions (>50), the GAs worked best.

In model construction, it is important to base the feature selection on the researcher's knowledge of the phenomenon and the variables affecting it; thus the use of stepwise selection methods is generally discouraged. However, there are situations in which the superiority of variables A and B over C and D is not clear. In remote sensing (RS), the relationships of recorded radiation or returned laser pulses and forest variables are not too straightforward (the exception being the canopy surface generated from laser height readings) and there are numerous potentially useful statistical/textural variables that can be extracted from the data. Therefore, the use of automated selection methods is justified to a certain extent.

In the present study, we examined the predictive capacity of several feature sets extracted from aerial photographs and low-pulse ALS data. While it is known, that the laser-based features perform far better than the aerial photograph-based features when estimating mean height, mean volume, etc., a combination is better when detecting tree species (Maltamo *et al*. 2006; Packalén and Maltamo 2006, 2007). One of the tested feature sets was based on automatic selection with GAs, others on expert knowledge. The estimation was carried out with the nonparametric *k*-nearest neigbour (*k*-NN) algorithm and we operated at the field plot level. The forest variables estimated included the mean volume of growing stock (m^3/ha), basal area (m^2/ha), height (m), diameter at breast height (DBH; cm), and the volumes of Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* H. Karst.), and deciduous trees (m^3/ha).

2. Material and methods

2.1 Study area and field measurements

The study area is located in Evo, Finland (61.19ºN, 25.11ºE) and it consists of approximately 2000 ha of managed boreal forest. The average stand size in the area was slightly less than 1 ha. Field measurement data from 282 fixed-radius (9.77 m) field plots were collected from the study area in summer 2007. The sampling of the field plots was based on prestratification of existing stand inventory data. There was a 1-year gap between the acquisition of RS data (see section 2.2) and field data measurements; only plots that had remained untreated during the year were measured and the latest growth in height was subtracted. 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 DBH of over 5 cm: location, tree species, crown class, DBH, height, lower limit of living crown and crown width. The volumes were calculated with standard Finnish models. Plot-level data were obtained by summing the tree data. The values of forest attributes of plots located in clear-cut areas or treeless mires were set at zero. The basic characteristics of the field data are presented in Table 1. Of the mean volume, 40% was Scots pine, 35% Norway spruce and 24% deciduous trees, mainly birch (*Betula* L.).

Table 1. Characteristics of the field plots.

	Mean	Min	Max	Std
Basal area-weighted mean height, m	17.0	0	30.5	6.7
Basal area-weighted mean DBH, cm	21.1	0	50.3	9.4
Basal area, m^2/ha	19.9	0	45.5	10.3
Mean volume of growing stock, m^3/ha	179.0	0	575.4	115.4
Mean volume of pine, m^3/ha	69.9	0	560.6	89.8
Mean volume of spruce, m^3/ha	63.5	0	575.4	94.8
Mean volume of deciduous trees, m^3/ha	42.9	0	302.2	51.2
Mean volume of other tree species, m^3/ha	2.7	0	210.1	19.0

2.2 Remote sensing material

The ALS data were acquired in midsummer 2006. The flying altitude was 1900 m. The density of the returned pulses within the field plots was 1.8/m^2 (only, first, intermediate or last; 1.3/m^2 if only or first pulses were considered). A digital elevation model (DEM) 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 mosaiced to a single image covering the entire area. Only near-infrared (NIR), red (R) and green (G) bands were available.

Several statistical and textural features were extracted from the RS material. The extraction window was generally 20 x 20 m, which was proved suitable in earlier 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. The Haralick textural features were computed from 4 directions: 0, 45, 90 and 135°. Additionally, the height statistics for the first and last pulses (F, L) were calculated as in Suvanto *et al.* (2005): mean and maximum height (hmea, hmax), standard deviation and coefficient of variation of height (hstd, hcv), 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 in order to remove hits to ground vegetation and bushes. Finally, percentages of points under 2 m in height were added (Fvege, Lvege; in Suvanto *et al.* 2005 the opposite, percentages of points over 2 m in height, was used). 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.

2.3 Methods

2.3.1 Estimation algorithm

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 nearest plots were weighted with inverse squared distances. The number of nearest neighbours was set at 5. Leave-one-out cross-validation was applied to calculate the results within the field dataset. The accuracy of the estimates was assessed by calculating the root-mean squared error (RMSE) of the studied variables.

2.3.2 Feature selection

Nine feature sets were created for tests:

- A: all aerial photograph features (72)
- B: all laser features (100)
- A + B (172)
- C: aerial photograph spectral features from three bands + their std's (6)
- D: ALS hmea, hmax, hstd, hcv, and vege of first and last pulse separately (10)
- C + D (16)
- E: D + local homogeneity of ALS height of first pulse (four directions) (14)
- F: D + ALS intensity and its std of first pulse (12)
- GA1-GA4: features selected by a genetic algorithm, starting from set A+B (all 172 features). From each step, the best features were fed to next step, e.g., from GA1 to GA2.

Feature sets A, B and A+B were created for benchmarking the results. Sets C, D and C+D were small feature sets containing simple statistics such as averages and variations. Set E was constructed of central laser height features and one height-based Haralick texture, local homogeneity, which performed well in an earlier study by Tuominen and Pekkarinen (2005), when derived from aerial photograph features. At this point in our project, the intensity values were not calibrated in any way, and thus little was expected from them. However, we created a subset containing two ALS intensity-based features, as well (F).

Automatic feature selection was carried out using a simple GA presented by Goldberg (1989), implemented in the GAlib C++ library (Wall 1996). It performed well in an earlier feature selection study by Haapanen and Tuominen (2008). 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 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. Genomes that were selected for mating swapped parts with each other with a probability of 60%, 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 (GA1-GA4). Only features belonging to the best genome in each step were included in the next step. The parameters used were selected via some explorative tests.

Even after careful selection, the features are not equally important in describing the forest attributes and should be weighted. Here we searched for optimal weights for the best subset of features by a downhill simplex method (Nelder and Mead 1965). In the search, the objective was to minimize the RMSE of the mean volume estimates.

3. Results

Estimation errors (RMSE%) obtained using the studied datasets are presented in Table 2 for mean total volume, mean height, basal area, mean DBH and species-specific mean volumes. The results are as follows:

- The ALS-based features (sets B, D, E, F) performed far better than aerial photograph-based features (sets A and C).
- Simply adding aerial photograph features into laser feature sets (A+B, C+D) gave worse results than the laser sets in question (B, D), except in the case of species-specific volumes.
- No expert judgment-based selection was able to surpass the set of all extracted laser features (B), when all variables were considered.
- However, even the first round of the GA produced lower RMSEs for most of the variables compared with full laser feature set B, by combining ALS- and aerial photograph-based features in a successful way.
- When all variables were considered, the most usable results were already obtained in step 3 of the GA process (GA3) with 19 features. Therefore, both GA3 and GA4 were weighted.
- After feature weighting, GA4 produced the lowest RMSEs and the weighted set of 11 features represents our final result. Both ALS and aerial photograph-based features were included. A test was also run without the aerial photograph features, but the accuracies then again lowered.

The 11 features selected into the final set were Fvege, Hvege, Fp30, Lp30, Fp90, mean height in the 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.

Table 2*: K*-NN estimation results obtained with tested datasets. The feature set giving the best results (GA4 weighted) is in bold face.

		RMSE, %						
Dataset	Bands	Mean height	BA	Mean DBH	Mean volume	Pine volume	Spruce volume	Deciduous volume
A	72	30.9	44.4	35.5	57.2	111.2	120.8	104.0
B	100	17.9	28.1	23.6	32.2	88.5	106.9	89.2
A+B	172	20.0	30.5	26.1	34.7	87.4	93.2	83.0
C	6	33.6	44.2	39.1	57.1	106.7	124.3	95.3
D	10	19.1	26.7	25.2	32.7	96.0	97.9	92.7
C+D	16	19.5	30.1	25.2	35.1	89.3	91.5	80.8
E	14	18.7	26.1	24.4	31.6	92.2	97.0	90.2
F	12	19.6	26.8	25.4	32.4	95.0	98.2	94.0
GA1	85	19.6	27.6	25.8	31.6	82.9	90.0	78.5
GA2	41	18.3	26.1	24.8	29.0	85.5	89.1	79.6
GA3	19	17.2	23.9	23.0	27.9	81.8	86.5	81.0
GA3 weighted	19	17.2	23.3	23.2	27.1	82.1	84.2	83.2
GA4	11	17.5	24.2	24.5	28.4	82.1	84.4	79.2
GA4 weighted	**11**	**16.9**	**23.2**	**23.8**	**27.1**	**81.4**	**84.6**	**79.6**

4. Discussion

In the present study, we tested the estimation of the most important forest attributes with a combination of ALS and aerial photograph data, using feature selection and the nonparametric *k*-NN algorithm. Based on our results, the lowest RMSEs, all variables considered, were obtained with a relatively small subset of the original features, comprising of both ALS and aerial photograph-based features. It was found via a GA-based feature selection process. Further weighting of the features was able to slightly lower the RMSEs of most of the variables. Our final RMSE for the mean volume was 27.1% of the mean. In comparison to results obtained using purely aerial photograph-based features, the drop in RMSE% was app. 30 percentage points and in comparison to other ALS-based results, 5-6 percentage points. The poor performance of aerial photograph-based features was in line with earlier studies (e.g. Haapanen and Tuominen 2008). The low spectral and radiometric resolution of these optical area data cannot distinguish forest characteristics: the grey values saturate at relatively low forest volumes (approx. 250 m^3/ha in this study). Even when complemented with textural features, the RMSEs tend to be only slightly lower than with Landsat-type satellite images, which in turn produce field plot level RMSEs of 60% or greater (Haapanen and Tuominen 2008). Naturally, this applies only to large-scale forest inventories based on two-phase sampling, and the situation in applications based on single tree detection is different.

Aerial photograph-based features lowered the estimation accuracies of general forest variables in sets A+B and C+D, which were constructed in a straightforward way. However, the species-specific accuracies were improved, compared with laser-based (or aerial photograph-based features). After the feature selection and weighting, all variables were more accurately estimated with a combination of laser and aerial photograph-based features, than with solely laser-based features. This implies that some aerial photograph-based features can improve the estimation of general forest parameters, as well.

Our results were poorer than the plot-level ALS results in a study area in eastern Finland presented by Suvanto *et al*. (2005), obtained by regression functions, or by Packalén and Maltamo (2007), obtained with a *k*-most similar neighbour (*k*-MSN) method using ALS features and aerial photographs. Our study area had greater variation in forest parameters and a larger proportion of deciduous trees, both being properties that reduce the estimation accuracy (Naesset 2004a; Maltamo *et al*. 2004). In comparison to the results in a study area in southern Finland (Maltamo *et al*. 2004), where the amount of deciduous trees is larger and understories denser than in eastern Finland, the relative mean volume RMSEs were similar (25% vs. our 27%). However, Maltamo *et al*. (2004) were able to reduce the relative mean volume RMSE to 16% by predicting the small trees separately.

To improve the estimation in forest areas with deciduous stands, stratification by cover types (Næsset 2004a) was suggested and later implemented based on aerial photograph-aided prestratification (Næsset 2004b). Aerial photographs were also integrated into the estimation process (Maltamo *et al*. 2006; Packalén and Maltamo 2006, 2007). Our approach resembled those of the latter studies: we fed the aerial photograph features together with the ALS features into a feature selection process. The feature selection criterion was tailored to take the tree species-specific volumes into account.

The large amount of deciduous trees is again seen in the estimation accuracies of the tree-species specific volumes: our accuracies for pine and spruce were lower, but for deciduous trees higher than in the studies by Packalén and Maltamo (2006, 2007). Packalén and Maltamo (2007) obtained RMSEs of 20.5%, 51.2%, 55.7% and 102.8% for mean total volume and mean volumes of pine, spruce and deciduous trees, respectively. Clear-cut areas or small seedling stands were excluded (minimum volume was 54 m^3/ha). When we removed volumes under 50 m^3/ha, our corresponding results were 25.6%, 75.7%, 77.9% and 74.7%, respectively. Weighting had no effect on these results.

Regression, in which each variable is separately modelled, produces more accurate variable-specific results than *k*-NN. However, *k*-NN and its special case *k*-MSN (based on canonical correlations and Mahalanobis distance; Moeur and Stage 1995) have the property of predicting all required variables simultaneously, preserving the concordance between variables. The *k*-MSN method is probably able to perform better than the *k*-NN method.

We did not perform extensive sets of GA runs with varying parameters and repetitions in this study. Since the *k*-NN is sensitive to a large number of features, which is the case at the upper levels of GA runs, our next step will be to test the feature selection separately within both datasets. Better feature sets can probably be found by continuing these efforts. However, our results were promising, and the features selected are a logical mixture of ALS and aerial photograph-based features. Of these features, Fvege and Lvege appeared in all but one of the regression models by Suvanto *et al*. (2005), as well: mean volume, basal area, stem number and mean diameter (it was not needed in the height model). Various height statistic features were selected in both studies. Three aerial photograph features were selected for our final set: mean of NIR values and two standard texture features based on NIR and G. The presence of NIR is logical, since it helps to separate deciduous trees from conifers. The mean of R could also have entered into the final dataset, but the proportion of ALS hits of under 2 m in height (Fvege, Lvege) apparently described the amount of vegetation biomass better than the R band.

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 operating on low pulse density data (Suvanto *et al*. 2004; Maltamo *et al*. 2006; Packalén and Maltamo 2007), but slightly poorer. In our data, the proportion of deciduous trees was considerable, and forests of all development classes were included, as well as both mineral soil and mire sites. This method is suitable for large area forest inventories, since it works with low pulse density and is simple. The feature selection algorithm tested (GA) worked well, outperforming the selections made by the researchers. However, stepwise regression could have performed as well (Haapanen and Tuominen 2008). The ALS data were superior to aerial photograph data (which in turn are slightly better than Landsat-type satellite image data; Haapanen and Tuominen 2008). However, some aerial photograph features were selected to the best performing feature set. More elaborate processing of intensity data (calibration) or higher pulse density of ALS data may eliminate the need for aerial photographs. Bearing in mind the further use of the resulting estimates, the species-specific estimates are a disappointment. If the estimates are to be used as input data in decision-making, or when simulating forest development, far more accurate estimates are needed. However, these figures concern the plot level only, and stand level estimates have typically been better, and similar or even more accurate than those obtained by field inventory of forest stands (Packalén and Maltamo 2007).

Acknowledgements

This study was made possible by financial aid from the Finnish Academy. HAMK University of Applied Sciences provided the ALS and field measurement data.

References

Goldberg, D.E., 1989. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley Publishing Company, Reading, Massachusetts, 412 p.

Guyon, I., and Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* ,3, 1157-1182.

Haapanen, R. and Tuominen, S., 2008. Data combination and feature selection for multi source forest inventory. *PE&RS* 74(7), 869-880.

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

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

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

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.

Kudo, M. and Sklansky, J., 2000. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33, 25-41.

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*, 35-42.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J. and 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.

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, 426-436.

Moeur, M., and Stage, A.R., 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science*, 41, 337–359.

Næsset, E., 1997. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, 51, 246-253.

Næsset, E., 2004a. Practical large-scale forest stand inventory using a small footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19, 164-179.

Næsset, E., 2004b. Accuracy of forest inventory using airborne laser-scanning: Evaluating the first Nordic full-scale operational project. *Scandinavian Journal of Forest Research*, 19, 554-557.

Nelder, J.A., and Mead, R., 1965. A simplex method for function minimization. *Computer Journal*, 7, 391-98.

Packalén, P. and Maltamo, M., 2006. Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. *Forest Science*, 56, 611-622.

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, 328-341.

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

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

Tuominen, S., and Haakana, M., 2005. Landsat TM imagery and high altitude aerial photographs in estimation of forest characteristics. *Silva Fennica*, 39(4), 573-584.

Tuominen, S., and Pekkarinen, A. 2005. Performance of different spectral and textural aerial photograph features in multi-source forest inventory. *Remote Sensing of Environment*, 94, 256-268.

Wall, M., 1996. GAlib: A C++ Library of Genetic Algorithm Components Version 2.4 Documentation, revision B. Massachusetts Institute of Technology. 101 p.