Forestry An International Journal of Forest Research

Institute of Chartered Foresters

Forestry, Vol. 85, No. 1, 2012. doi:10.1093/forestry/cpr051
Advance Access publication date: 11 October 2011

Comparative testing of single-tree detection algorithms under different types of forest

JARI VAUHKONEN[1]*, LIVIU ENE[2], SANDEEP GUPTA[3], JOHANNES HEINZEL[3], JOHAN HOLMGREN[4], JUHO PITKÄNEN[5], SVEIN SOLBERG[6], YUNSHENG WANG[3], HOLGER WEINACKER[3], K. MARIUS HAUGLIN[2], VEGARD LIEN[2], PETTERI PACKALÉN[1], TERJE GOBAKKEN[2], BARBARA KOCH[3], ERIK NÆSSET[2], TIMO TOKOLA[1] AND MATTI MALTAMO[1]

[1]School of Forest Sciences, University of Eastern Finland, PO Box 111, Joensuu, FI 80101, Finland
[2]Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, PO Box 5003, Ås, N 1432, Norway
[3]Department of Remote Sensing and Landscape Information Systems, Albert-Ludwig-University, Tennenbacherstr. 4, Freiburg, D 79106, Germany
[4]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Skogsmarksgränd, Umeå, SE 90183, Sweden
[5]Finnish Forest Research Institute, Joensuu Research Unit, PO Box 68, Joensuu, FI 80101, Finland
[6]Norwegian Forest and Landscape Institute, PO Box 115, Ås, N 1431, Norway
*Corresponding author. E-mail: jari.vauhkonen@uef.fi

Summary

Airborne laser scanning data and corresponding field data were acquired from boreal forests in Norway and Sweden, coniferous and broadleaved forests in Germany and tropical pulpwood plantations in Brazil. Treetop positions were extracted using six different algorithms developed in Finland, Germany, Norway and Sweden, and the accuracy of tree detection and height estimation was assessed. Furthermore, the weaknesses and strengths of the methods under different types of forest were analyzed. The results showed that forest structure strongly affected the performance of all algorithms. Particularly, the success of tree detection was found to be dependent on tree density and clustering. The differences in performance between methods were more pronounced for tree detection than for height estimation. The algorithms showed a slightly better performance in the conditions for which they were developed, while some could be adapted by different parameterization according to training with local data. The results of this study may help guiding the choice of method under different forest types and may be of great value for future refinement of the single-tree detection algorithms.

Introduction

Since around 1995, a large number of scientific studies have indicated a huge potential of airborne laser scanning (ALS) data to provide highly accurate estimates of important biophysical parameters of forests, like tree height, timber volume (e.g. Næsset, 1997; Magnussen and Boudewyn, 1998; Means *et al.*, 2000), and other parameters related to the structure and distribution of the tree layer (e.g. Zimble *et al.*, 2003; Maltamo *et al.*, 2005). The results have been most encouraging for coniferous forests. Overviews are provided by Lim *et al.* (2003), Næsset *et al.* (2004) and Hyyppä *et al.* (2008). Currently, there are two main approaches for using ALS to characterize forest resources: (1) an area-based approach typically providing data at stand level and (2) a single-tree approach where individual trees are the basic unit of the assessment.

In the single-tree approach, an initial task is to detect the trees, for which purpose different methods have been developed by different research groups (e.g. Kaartinen and

Hyyppä, 2008). Typically, the ALS-based height values are analyzed for detecting local height maxima, which are assumed to represent the treetops. To restrict the computational burden in processing dense point data, the local maxima search is usually performed from a canopy height model (CHM), which is interpolated from the point data. The cell values in the CHM represent the height difference between the top of the vegetation and the ground level, i.e. the canopy height. The differences between the algorithms are typically related to either (1) adjusting the CHM smoothing in order to obtain a desired number of local maxima in varying canopy conditions or (2) post-processing the result by further analysis of segment or point data properties (cf. Hyyppä *et al.*, 2008). The developed methods provide direct measurements of the position, height, and canopy shape of the trees, with best results for trees dominating in the canopy layer.

The single-tree approach performs best with dense laser scanning (5–10 laser pulses m^{-2}); however, for older trees with large canopies even as few as 2 pulses m^{-2} may be sufficient (Kaartinen and Hyyppä, 2008). Not all trees are usually detected, this being mainly dependent on scanning density and forest structure. In Scandinavia and Central Europe, rates of correctly detected trees higher than 70 per cent have been reported (Hyyppä *et al.*, 2001; Persson *et al.*, 2002; Koch *et al.*, 2006; Solberg *et al.*, 2006). The success rate in deciduous forests has generally been lower due to more complex crown shape and structure in the deciduous trees (Brandtberg *et al.*, 2003; Koch *et al.*, 2006). In Germany, success rates of ~50–60 per cent have been reported (Koch *et al.*, 2006; Heinzel *et al.*, 2011), which coincide well with similar studies in North America (e.g. Falkowski *et al.*, 2008).

Most tree detection studies have so far been carried out at limited test sites and are restricted to few tree species. Kaartinen and Hyyppä (2008) compared 12 different algorithms on two test sites located in southern Finland. According to them, the extraction method was the main factor influencing on the accuracy, the percentage of detected trees varying from 25 to 90 per cent for the various methods. However, their conclusions were based on results from test sites with fairly simple forest conditions and thus overall suitable for single-tree detection. Forest type and thus varying canopy structure is an important factor affecting the performance of the algorithms.

The purpose of this study was to test and compare the accuracy of single-tree detection algorithms under different types of forest. Compared with previous international comparisons (Kaartinen and Hyyppä, 2008), there are fewer algorithms involved, but the variation in forest types is much greater.

Materials and methods

An overview

The comparison performed in this study covered six tree detection algorithms. ALS data and a small sample of field data were delivered to the operators of these, requesting to extract treetop positions of each test site for validation. Each ALS dataset included at least XY coordinates and height above ground values, which were calculated in the pre-processing stage, i.e. all tested algorithms used standardized height values within each dataset, but the accuracy of the height values cannot be compared between the datasets. The extent of the associated field data varied, with a purpose to adapt the algorithms to the local conditions by training. The accuracy of tree detection and height estimation was assessed for each algorithm and test site.

Test sites and data

The test sites considered in this study were located in Brazil (16° 05′ S, 39° 24′ W), Germany (49° 02′ N, 8° 25′ E), Norway (59° 50′ N, 11° 30′ E) and Sweden (58° 30′ N, 13° 40′ E). The Brazilian test site was an even-aged pulpwood plantation growing Eucalyptus. The German data were pre-stratified to coniferous and deciduous plots. The dominant conifer species was Scots pine (*Pinus sylvestris* L.), while the deciduous plots were composed of oaks (*Quercus rubra* L., *Quercus petraea* (Mattuschka) Liebl.), European beech (*Fagus sylvatica* L.) and silver birch (*Betula pendula* Roth.). The species composition in the Scandinavian data consisted of the Scots pine and Norway spruce (*Picea Abies* (L.) H. Karst.) and to a lesser degree of deciduous trees, mainly birches (*Betula pendula* Roth., *Betula pubescens* Ehrh.). Of these two datasets, the Swedish data included notably larger trees than the Norwegian data. The main properties of the field data are given in Table 1.

The plots were positioned in the field using different methods between the datasets. In Germany, the plot centers were located in a traverse network and the tree stems were positioned relative to the plot center using a compass and a

Table 1: The main attributes of the field data. Number of plots (*N*), plot area (*A*) and averages (standard deviations) of trees per hectare (TPH), basal area (BA) and basal area-weighted mean tree diameter (*D*)

Site	*N*	*A* (ha)	TPH	BA (m^2 ha^{-1})	*D* (cm)
Brazil	19	0.05	807 (46)	22.6 (6.6)	18.5 (2.5)
Germany, deciduous	6	0.09	733 (331)	25.2 (4.9)	35.8 (16.5)
Germany, coniferous	5	0.09	664 (347)	28.8 (3.3)	30.6 (9.5)
Norway	40	0.1 (*N* = 36); 0.05 (*N* = 4)	1093 (485)	24.3 (8.4)	22.4 (4.8)
Sweden	17	0.64 (*N* = 10); 0.1 (*N* = 7)	610 (253)	34.6 (10.6)	32.9 (7.7)

measuring tape. The plot center points and the stems were expected to have accuracies <5 and <15 cm, respectively. In the other areas, satellite positioning and a total station were used to determine the plot and tree positions, respectively. In Brazil, the plot centers were determined using a real-time differential correction signal from the OmniSTAR satellite (http://www.omnistar.com). In Norway, differential post-processing by Pinnacle software package (version 1.00) was applied. Based on the positional standard errors reported by the Pinnacle software, the estimated accuracy of the planimetric plot coordinates ranged from <0.1 to 0.35 m, with an average of 0.12 m. In Sweden, in addition to the GPS measurements, the field data were matched with the ALS data using the spatial pattern of trees from both data sources (Olofsson *et al.*, 2008). The matching algorithm corrected the field plot and tree positions with an estimated translation and rotation for each plot. The translation in either x or y direction was at most 1.5 m and the rotation at most 2 degrees.

The extent of field measurements varied between datasets. Circular plots were used in the Brazilian and Norwegian site and rectangular plots in the German and Swedish sites, with plot sizes given in Table 1. The data consisted of full plots with all trees positioned in the field, except in Brazil, where plot positions were available. Tree heights were typically measured for a proportion of trees only, when the missing heights were modeled using local height curves. Specifically, these were constructed separately for each plot and species using the plot-specific measurements to derive the parameters for the model (Näslund, 1937).

The sample of field data that was given for training the algorithms also varied between datasets. To obtain a reasonable number of observations for reliable validation, the training data were included in the validation dataset, but a comparative analysis was done using independent validation data. The Brazilian site had separate validation plots, but these were measured from the same stands as the training data. In the German site, the training data included three deciduous and two coniferous plots also used for validation and nine plots from a separate area, where only a proportion of the trees had been measured. In Norway, the training data covered 10 plots, and in Sweden, the northern halves of the 80 × 80-m plots (N = 10; Table 1) were used for training. The data were chosen to represent as much variation in the study areas as possible.

The ALS data used in this study were acquired with different instruments and flying heights varying between the datasets, as described in Table 2. The point densities were ~1.5 (Brazil), 7 (Norway) and 30 pulses m^{-2} (Sweden). In Germany, two instruments were used with a pulse density of 16 and 7 pulses m^{-2}, respectively (see Table 2). The results of the German data are presented with respect to the formerly mentioned dataset only, except where these two data sources are compared.

Tree detection methods

The applied methods are listed below, but for detailed description, the reader should consult the primary publications. In the following text, the algorithms will be referred to by the numbers of the subsections below.

#1: Cluster formation using modified k-means approach

The algorithm used local height maxima as seed points (Gupta *et al.*, 2010). The unwanted local maxima were filtered using three-dimensional (3D) Euclidean distance criteria, which were set according to tests using training data. A k-means algorithm was applied to cluster the point data according to the seed points. An empirical height reduction factor was employed on the 3D ALS data and the respective seed points to minimize the bias and improve the grouping of similar objects.

#2: A voxel layer single tree modelling algorithm

The point data were projected into a voxel space, i.e. a regular grid in 3D space, where density images were calculated from sequential height layers (Wang *et al.*, 2008). The images were traced from top to bottom by a hierarchical morphological algorithm, assuming the amount of points to be higher where a tree crown occurs. Compared with Wang *et al.* (2008), the method was extended with an algorithm for merging split tree crowns, based on the horizontal distance of the two tree tops in relation to the crown radii, the vertical height differences of the two tree tops in relation to the crown length of the higher tree, and the difference of the two crown base heights.

#3: Adaptive segmentation based on Poisson forest stand model

The method employed a twofold strategy: (1) controlling the amount of CHM smoothing and (2) obtaining a CHM resolution suitable for representing the smallest tree crowns (Ene *et al.*, 2011). It was assumed that raw estimates of the

Table 2: Main properties of the ALS datasets

	Brazil	Germany (a)	Germany (b)	Norway	Sweden
Acquisition date	16 August 2008	30 August 2007	28 August 2008	6 June 2006	24 April 2007
Instrument	Optech ALTM 3100	TopoSys Harrier 56	Toposys Falcon II	Optech ALTM 3100	TopEye MKII
Density (nominal) (m^{-2})	1.5	16	7	7.4	30
Footprint (m)	0.36	0.23	0.35	0.2	0.13
Mean altitude (m)	1200	450	700	800	130
Field of view (°)	30	22.5	21.6	10	20/14*

* Elliptic scan pattern: cross flight direction/forward–backward.

stem density or tree spacing can be obtained and that the trees are randomly located within plots. The CHMs were first interpolated to various resolutions accommodated to each training dataset. Two runs of the pit-filling algorithm (Ben-Arie *et al.*, 2009) were applied to each CHM, followed by low-pass filtering using a binomial kernel with size proportional to the expected nearest neighbor distance between trees.

#4: Local maxima detection with residual height adjustment

The first echoes were interpolated into a digital surface model (DSM) with a 25-cm spatial resolution using a minimum curvature algorithm (Solberg *et al.*, 2006). The DSM was smoothed by running a Gaussian 3 × 3 filter a given number of times. Tree candidates were taken as local maxima in a 3 × 3 neighborhood. The height deviation of the first echoes from the DSM was calculated providing a residual height distribution, and the DSM and local maxima were adjusted by adding a given residual height percentile typically being the 70 percentile. The window size, the number of Gaussian runs and the residual height percentile adjustment were set specifically for each test site, by running tests with the field-measured trees, i.e. their position and height.

#5: Segmentation based on geometric tree crown models

An image was created by calculating the correlation between the laser-based height values and a geometric tree crown model (Pollock, 1996) placed at the centre of a raster cell (Holmgren and Wallerman, 2006; Holmgren *et al.*, 2010). The smoothed correlation image was used in the tree detection by first marking each raster cell with a non-zero CHM value and a positive correlation value as seed points. The location of each seed was updated to the neighbour cell with the highest correlation, this being repeated until a local maximum of the correlation surface had been reached. The seeds with the final location at the same local maximum defined a tree crown segment. For each segment, the geometric tree crown model was used to decide if the segment should be merged to a neighbour segment. In the training phase, the merging criteria were defined according to the known tree positions.

#6: Adaptive filtering based on CHM height values

In this method, the CHMs were low-pass filtered using Gaussian kernels, increasing the size of the smoothing window as a stepwise function of the heights of the CHM (Pitkänen *et al.*, 2004). The CHMs used with this method were interpolated to a grid of 0.5 m by taking the maximum first return height value within a radius of 0.5 m. The empty cells were filled by taking the average from a 3 × 3 window, and the interpolation was repeated successively until every cell had a height value. The algorithm required the determination of the kernel widths (sigma) and the height classes for which the sigma are applied. These were selected separately for each dataset based on training data.

Evaluation criteria and performance measures

The performance of the tree detection was evaluated by comparing

1. Detection rate, i.e. estimated number of trees in proportion to the number of trees measured in the field (denoted DET%)
2. Root mean squared error (RMSE) and bias of stem number ($RMSE_N$, $BIAS_N$) and the difference of estimated and observed plot-level mean height, in meters ($DIFF_H$)
3. The number of trees not detected, i.e. trees that could not be linked to any of the treetop candidates located by the algorithms (omission errors, OM%) and vice versa, false trees, i.e. the number of treetop candidates that could not be linked to any tree (commission errors, COM%)
4. RMSE and bias of tree height, in meters ($RMSE_H$, $BIAS_H$)

The RMSE and bias were calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}} \quad \text{and} \quad (1)$$

$$\mathrm{Bias} = \frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)}{n}, \quad (2)$$

where n is the number of observations, and x_i and $\hat{x}_i$ are the reference (field inventory) and estimated attributes, respectively, for the tree or plot i. In the case of individual tree height, only measured observations were used in the calculation of RMSE and bias.

The tree positions were not recorded in the field in the Brazilian dataset, so that the evaluation was carried out at the plot-level only. In order to evaluate omission, commission,

Table 3: Stand density, spatial pattern and tree competition indices used in this study

Abbreviation	Description	Equation
TPH	Trees per hectare	$\frac{N_k}{A_k/10000}$
BA	Basal area	$\sum_i \frac{\pi}{4} D_{ik}^2$
SDI	Stand density index	$\mathrm{TPH}_k \times \left(\frac{\bar{D}_k}{10}\right)^{1.605}$
CCF	Crown competition factor	$\frac{\sum_i \pi R_{ik}^2}{A_k}$
CEI	Clark–Evans index	$2\bar{d}_k\sqrt{N_k/A_k}$
OI	Crown overlap index	$\sum_{i,j} \frac{A_{ik \cap jk}}{A_{ik}}, i \neq j$
EI	Elevation angle index	$\sum_{i,j} \arctan\left(\frac{h_{jk} - 0.8 \times h_{ik}}{d_{ik,jk}}\right)$

N = number of trees; A = plot area; D = diameter at breast height; R = crown radius; d = tree-to-tree distance; i, j and k = indices for tree i, neighbor tree j and plot k.

Table 4: Quantitative success of the tree detection

Site	Algorithm	DET%	$RMSE_N$	$BIAS_N$	$DIFF_H$	$RMSE_H$	$BIAS_H$
Brazil	3	97.3	162	−22	−1.0	1.9	−0.5
	4	63.9	302	−291	0.2	2.6	0.7
	5	93.2	267	−55	−1.9	1.8	−0.6
	6	90.0	230	−80	−1.5	2.0	−0.3
Germany	1	69.5	375	−214	3.1	2.5	2.1
	2	83.9	302	−113	3.5	4.6	3.3
	3	49.1	430	−358	3.5	4.9	3.6
	4	100.7	247	5	1.9	2.1	1.2
	5	80.1	276	−139	2	2.5	1.8
	6	65.3	300	−262	2.5	4.1	2.2
Norway	1	49.4	628	−545	2.3	2.7	2.2
	2	51.3	622	−488	2.5	2.8	2.1
	3	52.8	576	−483	2.2	2.7	1.9
	4	56.8	605	−475	1.6	2.8	1.5
	5	68.1	460	−316	1.4	2.2	1.2
	6	45.2	685	−568	2.4	3.2	2.4
Sweden	1	77.6	564	−275	2.5	3.7	1.8
	2	65.1	674	−358	2.6	5.3	4.1
	3	68.9	637	−383	3.5	4.8	3.5
	4	71.6	614	−349	2.1	3.9	2
	5	85.8	415	−175	2.5	3.8	2.5
	6	68.5	668	−388	3.4	4.4	3.2

and errors on tree height, the treetop candidates detected from ALS data were linked to the field-observed trees according to the following procedure:

1 Maximum crown width (CW) was estimated from diameter at breast height and tree height (H) using species-specific models by Pretzsch *et al.* (2002) and Nagel *et al.* (2002).
2 Those treetop candidates that were located within the CW were initially accepted. When more than one candidate fell within the CW, the candidate having the smallest Euclidean 3D distance with the top of the field-observed tree was selected.
3 The height values of the linked pairs were compared to remove gross linking errors. A plotwise regression model $H_{candidate} \sim f(H_{field})$ was constructed, removing those tree-candidate links which had a height difference larger than 2× standard error of the model (cf. Breidenbach *et al.*, 2010).

The effects of forest structure on algorithm performance were quantified focusing on well-known stand density and spatial pattern indices (e.g. Biging and Dobbertin, 1995). The measures of algorithm performance were evaluated against trees per hectare (TPH), basal area (BA), stand density index (SDI), crown competition factor (CCF) and Clark–Evans index (CEI) calculated at plot level (Table 3). Additionally, we used a crown overlap index (OI) calculated from modelled CW (Biging and Dobbertin, 1992) and an additive competition index based on elevation angle sums (elevation angle index, EI) (Miina and Pukkala, 2000) to quantify the effect of tree competition to the algorithm performance at the level of individual trees. All indices are summarized in Table 3.

Table 5: Qualitative success of the tree detection

Site	Algorithm	OM%	COM%	$RMSE_H$	$BIAS_H$
Germany	1	53.4	32.9	2.0	−0.8
	2	54.8	46.1	1.7	−1.1
	3	60.6	19.6	1.7	−0.9
	4	39.0	39.4	2.1	−1.3
	5	42.4	28.2	1.7	−0.8
	6	48.8	21.6	1.5	−0.8
Norway	1	61.9	23.0	1.9	−0.2
	2	64.1	30.0	1.8	−0.4
	3	55.9	16.6	1.7	−0.2
	4	54.8	20.4	1.9	−0.4
	5	49.2	25.4	1.9	−0.2
	6	61.1	13.9	1.7	−0.2
Sweden	1	45.3	29.6	1.3	−0.1
	2	49.7	22.8	1.2	−0.1
	3	41.0	14.4	1.2	0.0
	4	35.7	10.3	1.4	−0.8
	5	30.6	19.1	1.3	0.0
	6	36.2	6.9	1.3	0.0

Results

On average, the number of trees found by the algorithms corresponded to 65 per cent of the number of trees measured in the field. The average tree detection rate varied between the test sites, being 86 per cent for Brazilian, 75 per cent for German, 54 per cent for Norwegian, and 73 per cent for Swedish plots. The average percentages of treetop candidates linked to field trees were 48 per cent (Germany), 42 per cent (Norway) and 60 per cent (Sweden). Thus, the

algorithms generally resulted in more commission errors in the German data.

The results in the independent validation data, i.e. not including the training plots were nearly similar or slightly better than those reported in the previous paragraph, except in the German data where more commission errors were produced. The average tree detection rates were 91 per cent (Germany), 55 per cent (Norway) and 77 per cent (Sweden) and the average percentages of treetop candidates linked to field trees 50 per cent (Germany), 44 per cent (Norway) and 60 per cent (Sweden). In the German data, almost a half of the observations were removed, which decreased the reliability of interpreting the results. Thus, a detailed performance evaluation is reported below for the combined dataset.

Algorithm #5 usually found a number of trees overall most close to field observations, especially in the Scandinavian

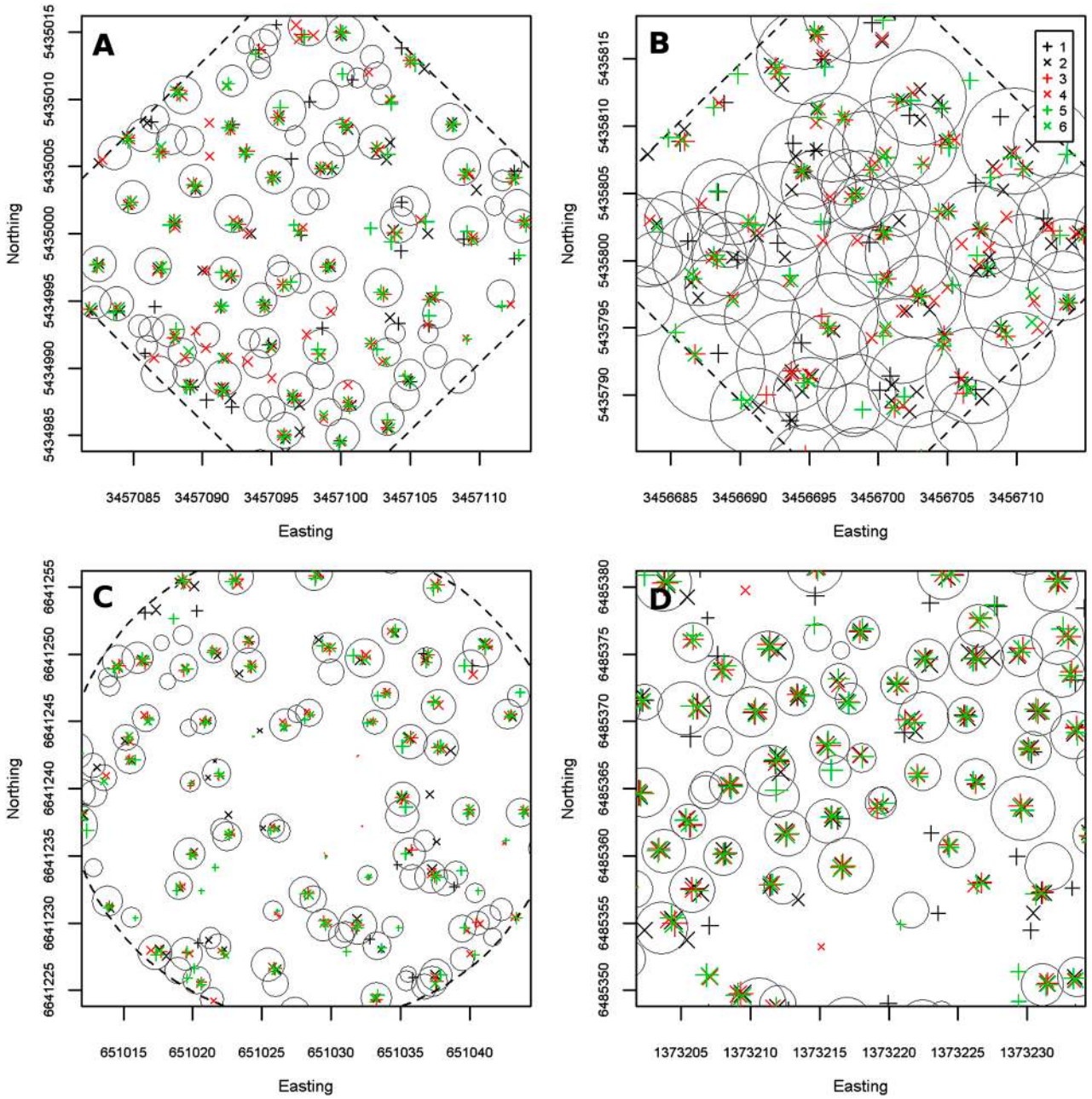

Figure 1. Interpretation results using all tested algorithms illustrated on top of a map of trees, where circles represent the predicted crown diameter based on field data. The plot borders, if visible, are drawn with a dashed line. Example plots: A, Germany (coniferous); B, Germany (deciduous); C, Norway; D, Sweden.

data, where the detection rate was 68 per cent for the Norwegian site and 86 per cent for the Swedish site (Table 4). Algorithms #3 and #2 were more accurate in Brazil and Germany, respectively, by ~4 percentage points compared with algorithm #5. In the German site, the result of algorithm #4 matched the number of field-measured trees (detection rate 100.7 per cent), but its performance varied most between the datasets. Algorithms #3 and #6 usually resulted in lowest number of trees. Algorithms #2 and #1 performed well in the German and Swedish datasets, respectively, but at an average level in others. These two algorithms were not applied to Brazilian data due to low pulse density of that dataset.

Algorithms #5 and #4 resulted in most treetop candidates linked to field trees (Table 5). The omission errors produced by these algorithms were 39 and 42 (Germany), 55 and 49 (Norway) and 36 and 31 per cent, respectively. Compared with these, algorithms #6 and #3 produced usually 6–12 percentage points more omission errors, but generally resulted in 4–23 percentage points lower commission error rates. Considering a summation of these error components, the most accurate algorithms were #6 and #5 in the German data; algorithm #3 followed by #4–6 in the Norwegian data; and algorithms #6, #4 and #5, in this order, in the Swedish data. Algorithms #1 and #2 performed poorer according to these metrics.

Treetop positions extracted by algorithms #1 and #2 most often deviated from other the methods, while these were extracted mainly from the same locations (Figure 1). In the German data, the treetop positions generally had a lower match with the field-measured trees (Figure 1B) and there was a considerable difference in the performance of the algorithms between coniferous and deciduous plots (Table 6). The detection rates of algorithms #1, #3 and #4 were 13–41 lower in the deciduous plots, thus producing a higher amount of omission errors. The detection rates of algorithms #5 and #6 had up to 3 percentage points difference between the coniferous and deciduous trees and algorithm #2 resulted in a 9 percentage points higher detection rate in the deciduous plots. In turn, the commission error rates of algorithms #5, #6 and #2 were 9–13 percentage points higher in the deciduous plots, while the difference in the commission error rates was less than 5 percentage points for the other three algorithms. Algorithm #1 differed from its overall performance (Table 5) in that it found a considerably high number of trees from the coniferous plots and resulted in less commission errors in the deciduous plots.

The data density had no clear effects to the tree detection, as compared in the German site (Table 6). Less trees were usually detected from the lower density dataset (b), but the commission error rates were clearly higher in the denser dataset (a). Algorithm #1 clearly benefitted from the increase in the point density. The tree detection rate of algorithms #2 and #5 also improved but at the cost of increased commission errors. On the other hand, algorithm #3 showed a stable performance between different datasets and types of forest (Table 6).

None of the algorithms differed significantly from their respective overall performance (Table 5), when the detection of trees of different sizes was considered (Figures 2–4). Large dominant trees were most often linked to field trees (Figure 2). Additionally, the same dominant trees were found by all algorithms, while most of the suppressed trees could not be found at all (Figure 4). The differences between algorithms were minor according to tree-level indices (Figure 4). Commission errors were more emphasized on low heights (Figure 3), but there were considerable variations between algorithms with respect to these (Table 5, Figure 3).

On average, the tree detection rate decreased and the omission error rate increased with increasing tree density and clustering in the spatial pattern (Table 7). Correlations of −0.68 and −0.74 were observed between TPH and tree detection rate and CEI and omission error rate, respectively. The relationships were not linear and in fact stronger than according to the correlation (Figure 5A). The number of commission errors reduced with increasing tree density but increased when the trees were more grouped (higher CCF). The latter showed a correlation

Table 6: Differences in algorithm performance between coniferous (C) and deciduous (D) plots in German ALS datasets (a) and (b)

		DET%		OM%		COM%	
Algorithm	Data source	C	D	C	D	C	D
1	(a)	92.6	52.0	33.4	60.1	28.2	23.3
	(b)	85.6	45.5	38.8	62.1	28.5	16.7
2	(a)	78.9	87.6	48.2	53.8	34.3	47.3
	(b)	60.2	68.4	51.8	56.1	20.0	35.8
3	(a)	53.2	46.0	51.8	59.6	9.4	12.1
	(b)	54.2	43.2	51.5	63.6	10.5	15.8
4	(a)	111.4	92.7	32.4	43.9	39.3	39.5
	(b)	116.7	87.4	31.8	42.7	41.5	34.4
5	(a)	78.3	81.6	34.8	42.7	16.7	29.7
	(b)	59.2	61.4	41.8	49.2	1.7	17.3
6	(a)	65.2	65.4	41.1	46.7	9.7	18.5

Algorithm #6 was evaluated only in dataset (a).

of 0.36, but otherwise the commission error rate was less correlated with the stand attributes (Table 7). Most between-algorithm variations came from the commissions (Figure 5F), but all algorithms behaved in relatively similar way with respect to transition in the stand attributes (Figure 5B,D,F).

There were less differences between the algorithms in tree height estimation. Except for the Brazilian dataset, the mean tree height was generally overestimated by 1.5–3.4 m (Table 4) and the difference in estimated and measured height had a correlation of 0.73 with the plot-level basal area (Table 7, Figure 6A). With BA < 20, the plot height was generally underestimated, while the increase in BA turned the values to overestimates.

In estimating height for the detected individual trees, the performance of all algorithms was almost equal, with RMSEs varying ≤ 0.2 m in the Norwegian and the Swedish data and ≤ 0.6 m in the German data (Table 5). Algorithm 4 usually resulted in lowest accuracy especially in the German and the Swedish data (Table 5, Figure 6F), with underestimates of 1.3 and 0.8 m, respectively, while the underestimation with the other algorithms was on average 0.9 m in the German data and ≤0.2 m in the Scandinavian data. Stand density had an effect on the bias (correlation

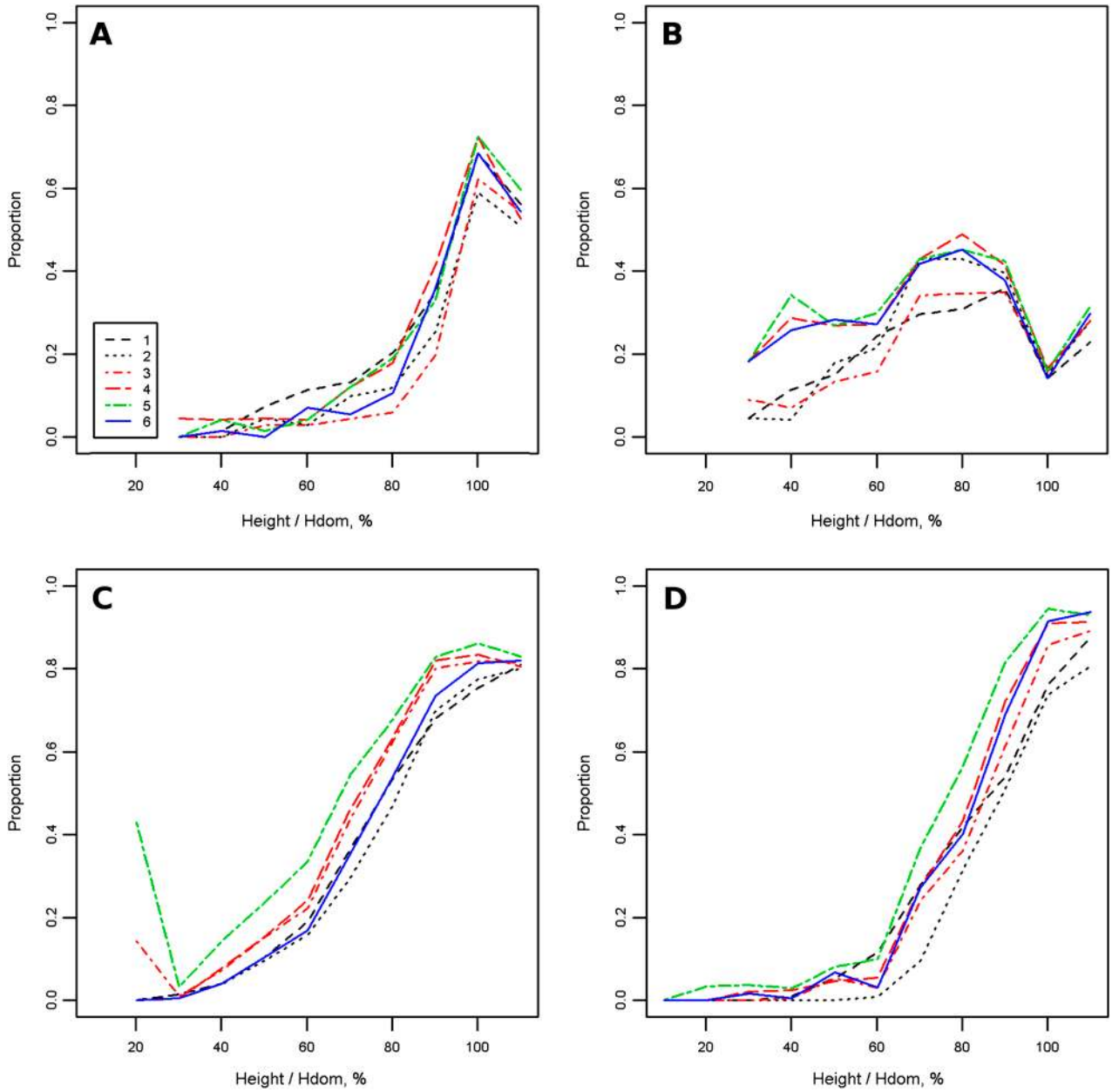

Figure 2. Proportion of treetop candidates linked to field trees as a function of tree height relative to plot-level dominant height (Hdom). Example plots: A, Germany (coniferous plots); B, Germany (deciduous plots); C, Norway; D, Sweden.

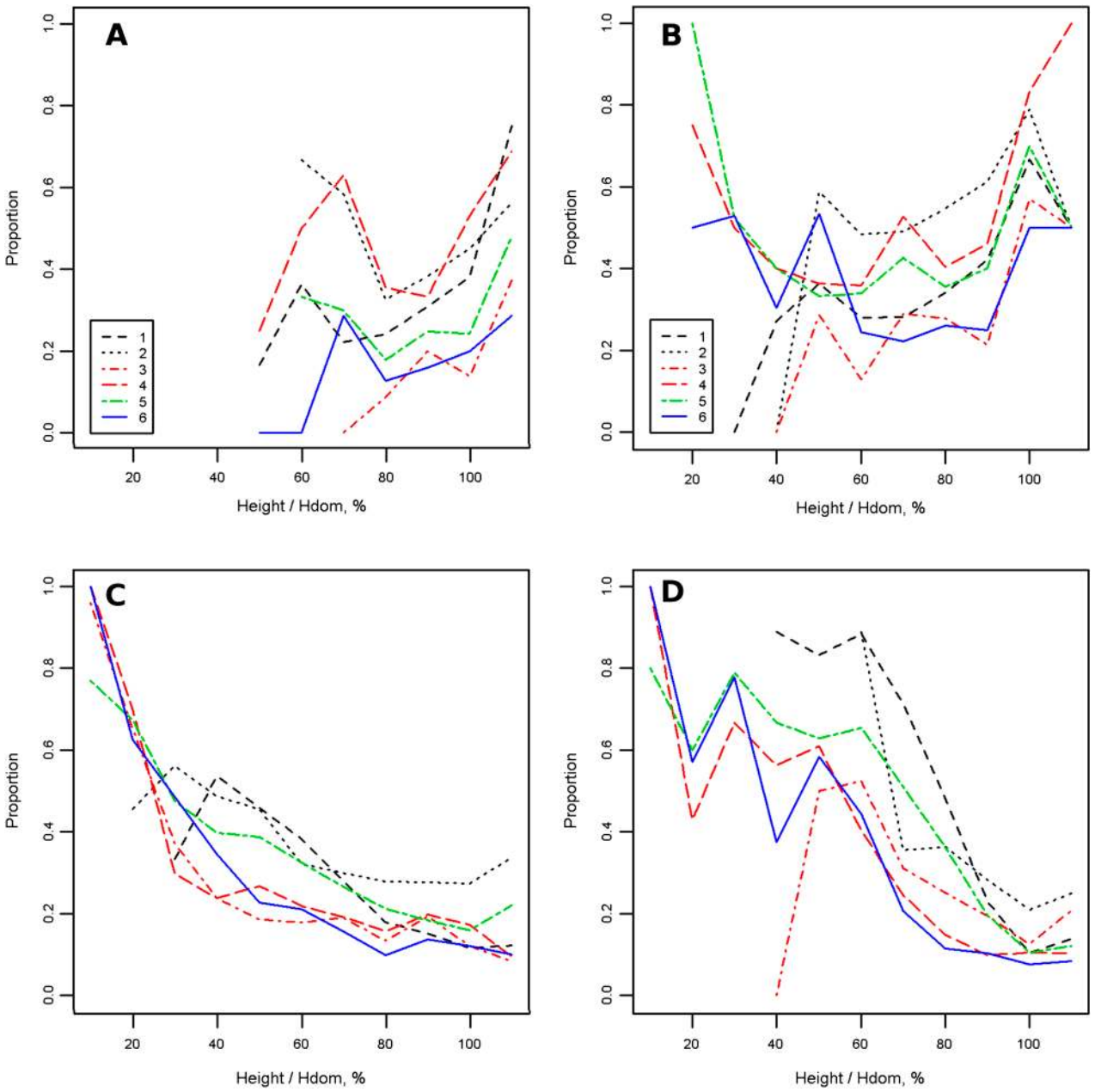

Figure 3. Proportion of commission errors as a function of tree height relative to plot-level dominant height (Hdom). Example plots: A, Germany (coniferous plots); B, Germany (deciduous plots); C, Norway; D, Sweden.

−0.42 with SDI) but generally the forest structure affected less to these figures (Table 7, Figure 6C,E). Figure 6(C,E) shows a different correlation structure between the Norwegian and the Swedish data, such that error levels in the Swedish data seem more independent from the stand indices.

Discussion

The results of this study showed in general less differences between the tested algorithms than the earlier comparison performed by Kaartinen and Hyyppä (2008). The principles of the algorithms tested here are obviously more similar than in the study of Kaartinen and Hyyppä (2008), which included a wider range of different algorithms. The results of the present comparison showed more remarkable differences between the methods in tree detection rates rather than in height estimation accuracies.

There were two basic differences between the algorithms tested in this study. First, algorithms #1 and #2 use point data, whereas the others require a CHM interpolation step. Second, algorithms differ in the ability to employ training data in setting proper parameters for the interpretation.

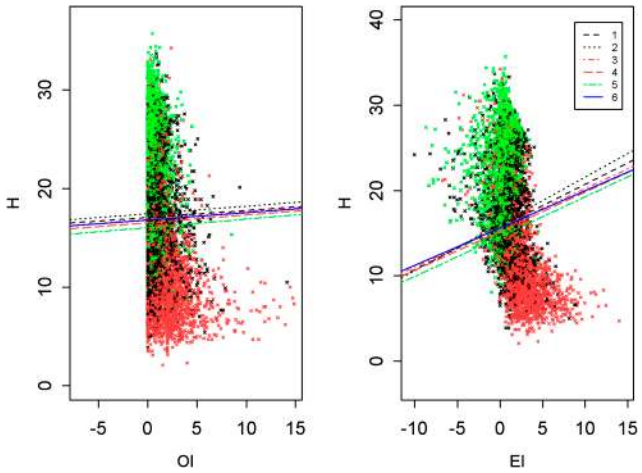

Figure 4. Detected trees according to tree-level indices. Trees detected by all and none algorithms are indicated by green and red, respectively. The lines are drawn perpendicular to a vector of means of the detected and undetected trees of each algorithm. The colour version of this figure is available at *Forestry* online.

Table 7: Pearson's correlation coefficients between the average algorithm performance and stand density and spatial pattern indices

Performance measure	TPH	BA	SDI	CCF	CEI
DET%	−0.68	0.07	−0.13	0.15	0.50
OM%	0.59	−0.14	−0.05	0.17	−0.74
COM%	−0.22	−0.08	−0.27	0.36	−0.34
$DIFF_H$	−0.29	0.73	0.53	0.18	0.22
$RMSE_H$	0.20	−0.02	0.04	0.28	−0.23
$BIAS_H$	−0.17	−0.39	−0.42	−0.33	−0.03

Here local training data were available, and due to the training ability, the performance of some algorithms could be clearly improved. Algorithms #1 and #2 were developed in Central Europe, and they resulted in a better performance in those conditions compared with Scandinavia. On the other hand, Scandinavian algorithms #3–6 performed fairly well even in the non-Scandinavian test sites.

The data used here varied in extent, and the detailed analysis of the algorithm performance was performed in the dataset, in which the algorithm training data and independent validation data were combined. A comparative analysis using the plots which were not included for algorithm training showed, however, only a minor difference in the results, except for the German data. In that dataset, the number of independent validation observations considerably reduced, which may have affected the results. Furthermore, since the training data were selected to describe the overall variation in the area and thus include the extremes in the area, the differences likely reflect the change in the forest conditions between the datasets rather than difference in the algorithm performance.

The datasets of this study differed by their properties such as the point density. A comparison using the German data showed that an increase in the density benefitted especially the point-based algorithms of this study. A similar conclusion was reported by Reitberger *et al.* (2009), who highlighted the benefit of full waveform data especially in segmentation task. The Brazilian dataset, on the other hand, produced a special situation for single-tree detection since the data were sparse compared with other datasets, but the trees appeared in rows with known spacing. Algorithm #3 used initial estimates of stem density in the tree detection and clearly outperformed other algorithms in the Brazilian test site. However, these results were not verified by qualitative criteria as applied to other datasets. As examined visually, a portion of treetop candidates generally originated from large branches rather than actual tree tops.

In some cases, the quantitative and qualitative success criteria were found contradictory. Consider, for example, the plot-level mean tree height: the fewer trees an algorithm detected, the more inaccurate the height estimate since the detection result is focused on the largest trees on the plot. This is still logical, however, as increasing commission often improves the accuracy of plot-level mean height since commission trees are normally not as large as properly detected trees. On the other hand, algorithms #3 and #6, which found the lowest number of trees, were benefitting from the qualitative criteria. These algorithms also behaved most stable when the properties of the datasets changed.

Linking treetops to field trees was clearly more successful in Scandinavian datasets, while it was known that the lateral difference between tree tops and stem foot positions could be up to 7 m in the German data. Thus, the commission error rates between different datasets cannot directly be compared. Furthermore, strict linking criteria were applied, which may partly explain the number of commission errors, but not the high variation between the algorithms. Further work in developing these algorithms should clearly focus on reducing the false height maxima caused by varying canopy structure. One obvious way according to these results would be to remove unacceptably low height maxima in relation to the obtained dominant height.

As opposed to earlier studies, here the algorithm performance was analyzed with respect to forest structure parameters. None of the algorithms showed special sensitivity to these parameters, but the general success of tree detection was found to be strongly related to stand density and spatial pattern of trees. The effect of these parameters to the algorithm performance was in fact surprisingly consistent, considering the differences in the canopy structure of the test sites. Based on our findings, there are clear limits in terms of forest density and clustering where single-tree detection can be successfully applied. Future studies should examine whether corresponding forest structure parameters could be extracted from ALS data (cf. Falkowski *et al.*, 2008), which would help adjusting the algorithm parameterization to correspond the canopy structure conditions or estimating the success of tree detection prior to performing the actual analysis.

Further tests should be carried out to compare algorithms with respect to using the delineated point clouds

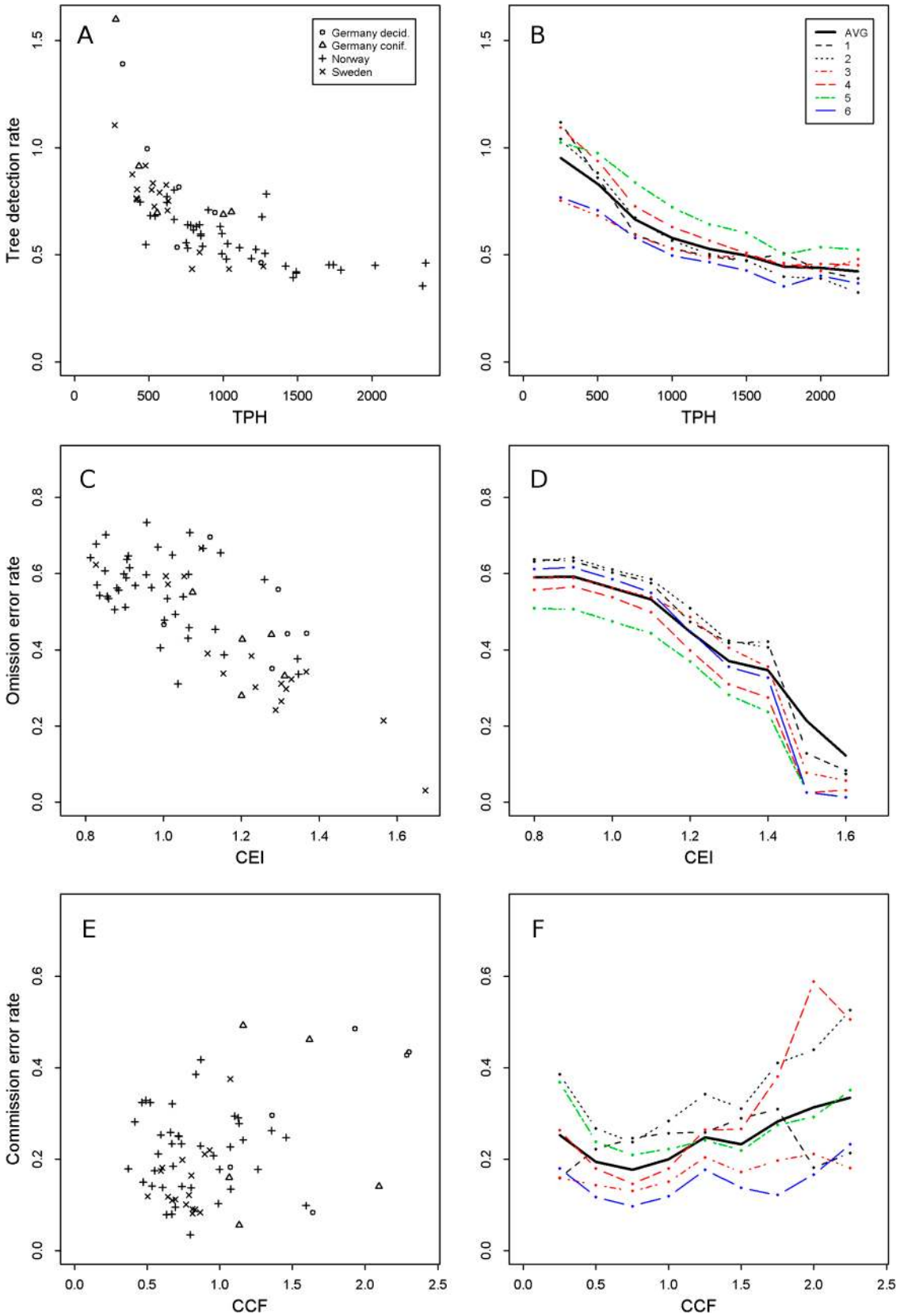

Figure 5. Success of tree detection as a function of stand density and spatial pattern indices. Average algorithm performance (scatter plots, left) and the performance of individual algorithms expressed as a moving average (line plots, right).

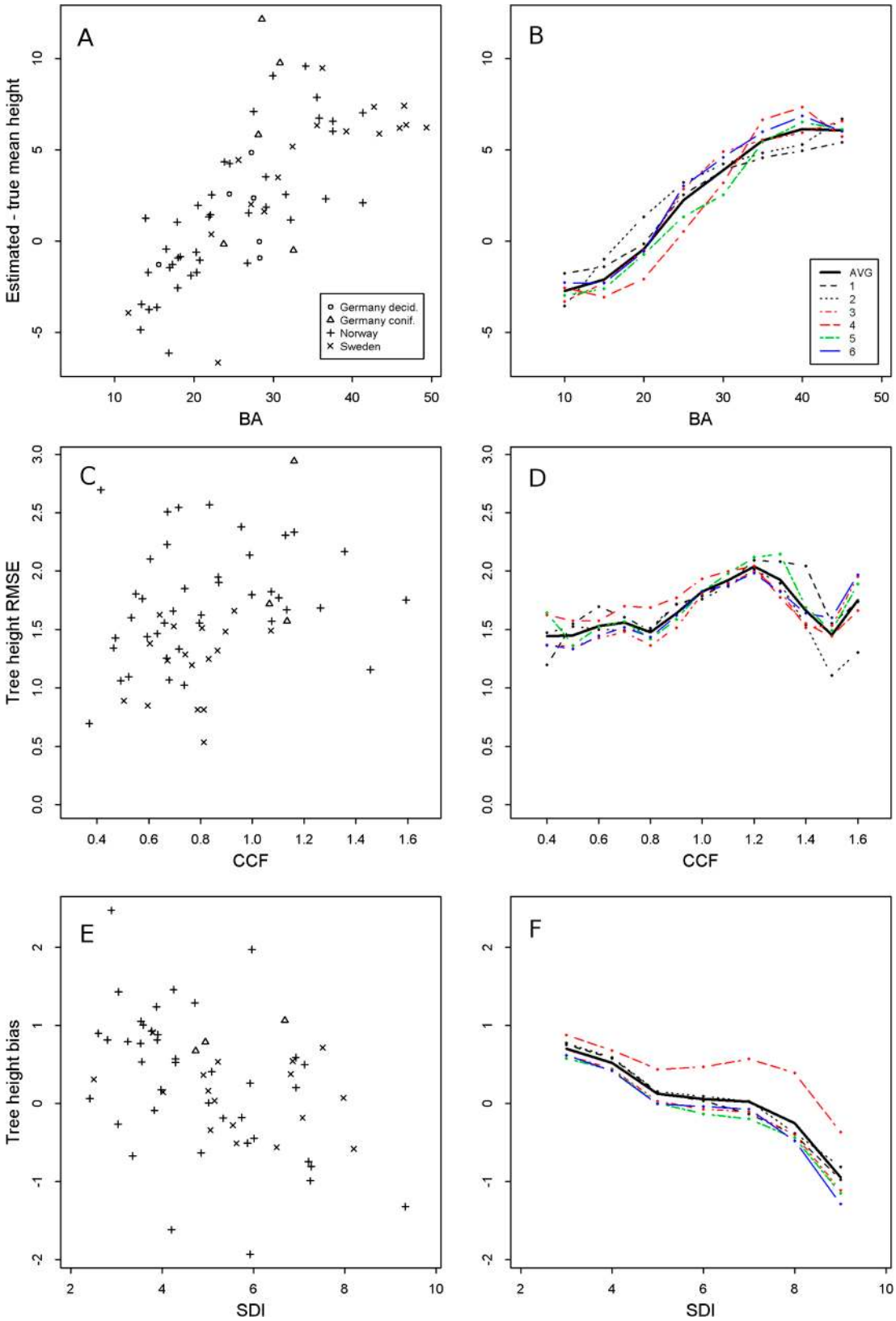

Figure 6. Success of tree height estimation as a function of stand density and spatial pattern indices. Average algorithm performance (scatter plots, left) and the performance of individual algorithms expressed as a moving average (line plots, right).

for predicting attributes such as tree species and stem volume since these are the actual end-products expected from ALS data interpretation rather than only detecting trees. Currently, little is known on tree crown segmentation accuracy, yet further analyses are typically based on the properties of the point cloud (e.g. Holmgren and Persson, 2004; Vauhkonen *et al.*, 2010), which calls for high segmentation accuracy in both horizontal and vertical directions. However, the estimation of species-specific attributes would place challenges for an international comparison due to requirements for a local understanding of tree morphology to differentiate species (Korpela *et al.*, 2010) and a highly extensive field reference data to construct allometric relationships (Kalliovirta and Tokola, 2005), for example.

Conclusions

This study validated and compared single-tree detection algorithms under different types of forest. The differences in performance between methods were found to be more pronounced for tree detection than for height estimation. In general, the algorithms showed very similar performance, which was more influenced by forest structure than by algorithm. Specifically, the success of tree detection was found to be dependent on tree density and clustering.

Funding

EU Wood Wisdom-Net ERA-Net Framework, project 'New Technologies to Optimize the Wood Information Basis for Forest Industries—Developing an Integrated Resource Information System (WW-IRIS)'.

Conflict of interest statement

None declared.

References

Ben-Arie, J.R., Hay, G.J., Powers, R.P., Castilla, G. and St-Onge, B. 2009 Development of a pit filling algorithm for LiDAR canopy height models. *Comput. Geosci.* **35**, 1940–1949.

Biging, G.S. and Dobbertin, M. 1992 A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *For. Sci.* **38**, 695–720.

Biging, G.S. and Dobbertin, M. 1995 Evaluation of competition indices in individual tree growth models. *For. Sci.* **41**, 360–377.

Brandtberg, T., Warner, T.A., Landenberger, R.E. and McGraw, J.B. 2003 Detection and analysis of individual leaf-off tree crowns in small footprint, high density lidar data from the eastern deciduous forest in North America. *Remote Sens. Environ.* **85**, 290–303.

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T. and Solberg, S. 2010 Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sens. Environ.* **114**, 911–924.

Ene, L., Næsset, E. and Gobakken, T. 2011 Single tree detection in heterogeneous boreal forests using airborne laser scanning and area based stem number estimates. *Int. J. Remote Sens.*

Falkowski, M.J., Smith, A.M.S., Gessler, P.E., Hudak, A.T., Vierling, L.A. and Evans, J.S. 2008 The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. *Can. J. Remote Sens.* **34**, (Suppl. 2), S338–S350.

Gupta, S., Koch, B. and Weinacker, H. 2010 Tree species detection using full waveform lidar data in a complex forest. In *Technical Commission VII Symposium—100 Years of ISPRS Vienna, Austria, July 5–7, 2010*. W. Wagner and B. Székely (eds). Vienna University of Technology, Vienna, Austria, pp. 249–254.

Heinzel, J., Weinacker, H. and Koch, B. 2011 Prior knowledge based single tree extraction. *Int. J. Remote Sens.* **32**, 4999–5020.

Holmgren, J. and Persson, Å. 2004 Identifying species of individual trees using airborne laser scanner. *Remote Sens. Environ.* **90**, 415–423.

Holmgren, J. and Wallerman, J. 2006 Estimation of tree size distribution by combining vertical and horizontal distribution of LIDAR measurements with extraction of individual trees. In *Workshop on 3D Remote Sensing in Forestry, 14–15 February 2006*. T. Koukal and W. Schneider (eds). University of Natural Resources and Applied Life Science, Vienna, Austria, pp. 168–173.

Holmgren, J., Barth, A., Larsson, H. and Olsson, H. 2010 Prediction of stem attributes by combining airborne laser scanning and measurements from harvesting machinery. SilviLaser 2010, the 10th International Conference on LiDAR Applications for Assessing Forest Ecosystems, 14–17 September 2010, Freiburg, Germany, Proceedings USB disc, 10 pp.

Hyyppä, J., Kelle, O., Lehikoinen, M. and Inkinen, M. 2001 A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Trans. Geosci. Remote Sens.* **39**, 969–975.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X. and Maltamo, M. 2008 Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *Int. J. Remote Sens.* **29**, 1339–1366.

Kaartinen, H. and Hyyppä, J. 2008 *EuroSDR/ISPRS Commission II project: "Tree Extraction"—final report*. Official publication no. 53. EuroSDR, Frankfurt am Main, Germany, 60 p. http://bono.hostireland.com/~eurosdr/publications/53.pdf (accessed on 26 September, 2011).

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 Fenn.* **39**, 227–248.

Koch, B., Heyder, U. and Weinacker, H. 2006 Detection of individual tree crowns in airborne lidar data. *Photogramm. Eng. Remote Sens.* **72**, 357–363.

Korpela, I., Ørka, H.-O., Maltamo, M., Tokola, T. and Hyyppä, J. 2010 Tree species classification in airborne LiDAR data: influence of stand and tree factors, intensity normalization and sensor type. *Silva Fenn.* **44**, 319–339.

Lim, K., Treitz, P., Wulder, M., St-Onge, B. and Flood, M. 2003 LiDAR remote sensing of forest structure. *Prog. Phys. Geogr.* **27**, 88–106.

Magnussen, S. and Boudewyn, P. 1998 Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Can. J. For. Res.* **28**, 1016–1031.

Maltamo, M., Packalén, P., Yu, X., Eerikäinen, K., Hyyppä, J. and Pitkänen, J. 2005 Identifying and quantifying structural characteristics of heterogenous boreal forests using laser scanner data. *For. Ecol. Manage.* **216**, 41–50.

Means, J.E., Acker, S.A., Brandon, J.F., Renslow, M., Emerson, L. and Hendrix, C.J. 2000 Predicting forest stand characteristics with airborne scanning lidar. *Photogramm. Eng. Remote Sens.* **66**, 1367–1371.

Miina, J. and Pukkala, T. 2000 Using numerical optimization for specifying individual tree competition models. *For. Sci.* **46**, 277–283.

Nagel, J., Albert, M. and Schmidt, M. 2002 Das waldbauliche Prognose- und Entscheidungsmodell BWINPro 6.1 (in German). *Forst Holz.* **57**, 486–493.

Näslund, M. 1937 Skogsförsöksanstaltens gallringsförsök i tallskog (in Swedish). *Medd. Stat. Skogsförsöksanst.* **29**, 169.

Næsset, E. 1997 Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogrammetry Remote Sens.* **52**, 49–56.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M. *et al.* 2004 Laser scanning of forest resources: the Nordic experience. *Scand. J. For. Res.* **19**, 482–499.

Olofsson, K., Lindberg, E. and Holmgren, J. 2004 A method for linking field-surveyed and aerial-detected single trees using cross correlation of position images and the optimization of weighted tree list graphs. Proceedings of SilviLaser 2008, the 8th International Conference on LiDAR Applications in Forest Assessment and Inventory, 17–19 September, 2008, Edinburgh, UK, R.A. Hill, J. Rosette and J. Suárez (eds). 10 pp. http://geography.swan.ac.uk/silvilaser/papers/oral_papers/Data%20Fusion/Olofsson.pdf (accessed on 26 September, 2011).

Persson, Å., Holmgren, J. and Söderman, U. 2002 Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sens.* **68**, 925–932.

Pitkänen, J., Maltamo, M., Hyyppä, J. and Yu, X. 2004 Adaptive methods for individual tree detection on airborne laser based canopy height model. In *Proceedings of ISPRS Working Group VIII/2:* "*Laser-scanners for forest and landscape assessment*". M. Theis, B. Koch, H. Spiecker and H. Weinacker (eds). University of Freiburg, Germany, pp. 187–191.

Pollock, R.J. 1996 *The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model.* Ph.D. Thesis. University of British Columbia, Vancouver, Canada. 172 pp.

Pretzsch, H., Biper, P. and Durský, J. 2002 The single-tree based simulator SILVA: construction, application and evaluation. *For. Ecol. Manage.* **162**, 3–21.

Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U. 2009 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogrammetry Remote Sens.* **64**, 561–574.

Solberg, S., Næsset, E. and Bollandsås, O.M. 2006 Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogramm. Eng. Remote Sens.* **72**, 1369–1378.

Vauhkonen, J., Korpela, I., Maltamo, M. and Tokola, T. 2010 Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sens. Environ.* **114**, 1263–1276.

Wang, Y., Weinacker, H., Koch, B. and Sterenczak, K. 2008 Lidar point cloud based fully automatic 3D single tree modelling in forest and evaluations of the procedure. *Int. Archives Photogrammetry Remote Sens Spatial Inform. Sci.* **XXXVII**, 45–51.

Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C. and Gerard, P.D. 2003 Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sens. Environ.* **87**, 171–182.

Received 15 March 2011