Article

Single-Sensor Solution to Tree Species Classification Using Multispectral Airborne Laser Scanning

Xiaowei Yu [1],*, Juha Hyyppä [1], Paula Litkey [1], Harri Kaartinen [1], Mikko Vastaranta [2] and Markus Holopainen [2]

[1] Finnish Geospatial Research Institute, National Land Survey, Geodeetinrinne 2, FI-02430 Masala, Finland; juha.hyyppa@nls.fi (J.H.); paula.litkey@nls.fi (P.L.); harri.kaartinen@nls.fi (H.K.)
[2] Department of Forest Sciences, University of Helsinki, FI-00014 Helsinki, Finland; mikko.vastaranta@helsinki.fi (M.V.); markus.holopainen@helsinki.fi (M.H.)
* Correspondence: xiaowei.yu@nls.fi; Tel.: +358-50-4126536

Abstract: This paper investigated the potential of multispectral airborne laser scanning (ALS) data for individual tree detection and tree species classification. The aim was to develop a single-sensor solution for forest mapping that is capable of providing species-specific information, required for forest management and planning purposes. Experiments were conducted using 1903 ground measured trees from 22 sample plots and multispectral ALS data, acquired with an Optech Titan scanner over a boreal forest, mainly consisting of Scots pine (*Pinus Sylvestris*), Norway spruce (*Picea Abies*), and birch (*Betula* sp.), in southern Finland. ALS-features used as predictors for tree species were extracted from segmented tree objects and used in random forest classification. Different combinations of features, including point cloud features, and intensity features of single and multiple channels, were tested. Among the field-measured trees, 61.3% were correctly detected. The best overall accuracy (OA) of tree species classification achieved for correctly-detected trees was 85.9% (Kappa = 0.75), using a point cloud and single-channel intensity features combination, which was not significantly different from the ones that were obtained either using all features (OA = 85.6%, Kappa = 0.75), or single-channel intensity features alone (OA = 85.4%, Kappa = 0.75). Point cloud features alone achieved the lowest accuracy, with an OA of 76.0%. Field-measured trees were also divided into four categories. An examination of the classification accuracy for four categories of trees showed that isolated and dominant trees can be detected with a detection rate of 91.9%, and classified with a high overall accuracy of 90.5%. The corresponding detection rate and accuracy were 81.5% and 89.8% for a group of trees, 26.4% and 79.1% for trees next to a larger tree, and 7.2% and 53.9% for trees situated under a larger tree, respectively. The results suggest that Channel 2 (1064 nm) contains more information for separating pine, spruce, and birch, followed by channel 1 (1550 nm) and channel 3 (532 nm) with an overall accuracy of 81.9%, 78.3%, and 69.1%, respectively. Our results indicate that the use of multispectral ALS data has great potential to lead to a single-sensor solution for forest mapping.

Keywords: multispectral laser scanning; ALS; individual tree detection; tree species classification; random forest

1. Introduction

Knowledge of tree species plays an important role in forest management and planning. The optimum output, requested by forest companies from the forest mapping process, is the species-specific size distribution of the trees. The traditional method, based on field inventory work for tree species identification, is labor intensive, time consuming, and limited by spatial extent.

Therefore, remote sensing techniques were introduced, such as the interpretation of large-scale aerial color or infra-red images [1,2]. Although remotely-sensed data have been widely used for forest applications, traditional optical remote sensing techniques suffer from a lack of the ability to capture three-dimensional forest structures, particularly in unevenly-aged, mixed species forests with multiple canopy layers [3]. Recent developments in active remote sensing, particularly laser scanning techniques, have shown potential in forest mapping and other applications because of the capability to capture three-dimensional (3D) information of forests [4–11].

Airborne laser scanning (ALS) is a useful tool for retrieving biophysical variables and for updating forest inventory maps. The successful use of ALS data has been demonstrated for a variety of applications. For example, ALS has been used to estimate tree height [6,7], identify tree species [8–10], and estimate tree volume, biomass [11–13], and growth [14,15]. Tree species information at an individual tree level is particularly useful in growth and yield estimates, and has been primarily studied for forest applications, such as updating forest inventories. Tree species classification using ALS has not been intensively studied, when compared with studies on the successful use of ALS for other forest attribute mapping, because of the lack of spectral information. Brandtberg [9] classified three leaf-off individual deciduous tree species (oaks, red maple, and yellow poplar) in West Virginia, USA, using high density laser data, and reported 64% total accuracy. Holmgren and Persson [8] classified Norway spruce and Scots pine in Remningstorp, Sweden, using ALS-derived point and intensity features, and achieved an accuracy of 95%. Ørka et al. [16] classified three species (spruce, birch and aspen) at the Ostmarka natural forest in southern Norway. Suratno et al. [17] classified ponderosa pine, Douglas-fir, western larch, and lodgepole pine, in a western North American montane forest using low density ALS data, and achieved a classification accuracy of 95% at the dominant species level, and 68% for individual trees.

Intensity was also demonstrated to be useful information for tree species identification. Ørka et al. [18] reported an accuracy of 73% when classifying conifers and deciduous trees, solely based on intensity information. Korpela et al. [19] classified Scots pine, Norway spruce, and birch, by using intensity variables at Hyytiälä in southern Finland, and showed that intensity features can contribute to a classification accuracy of 88% among the three species. With full-waveform (FWF) lasers, the total received power corresponding to the backscattering cross-section can be calculated, which provides information on the objects, from the intensity waveform.

Previous studies have demonstrated that FWF data and the derived metrics can be used to improve the performance of tree species classification. For example, Yao et al. [20] demonstrated the usefulness of waveform features for the classification of deciduous and coniferous trees. Heinzel and Koch [21] analyzed a set of waveform features and identified the most predictive features for classifying up to six tree species. Cao et al. [22] demonstrated that full-waveform data and derived metrics have significant potential for tree species classification in the subtropical forests, and results demonstrated that all tree species were classified with relatively high accuracy (68.6% for six classes, 75.8% for four main species, and 86.2% for conifers and broadleaved trees).

Previous studies have also revealed that combining multispectral information with 3D ALS data can lead to improvement in the accuracy of tree extraction and tree species classification, as we can take advantage of both datasets. For example, Naidoo et al. [23] concluded that the use of ALS and hyperspectral data yielded the highest classification accuracy and prediction success for the eight savanna tree species, with an overall classification accuracy of 87.68%. Zhou et al. [24] demonstrated that the ALS intensity data can contribute to the classification of shaded areas in an urban environment where high resolution digital aerial imagery alone did not produce good results. The fusion of high resolution (satellite or aerial) remote sensing and ALS data can achieve mutual benefits for compensating the lack of 3D structure from imagery and multi-spectral information from ALS data. With respect to the success of these case studies, multi-sensor data fusion seems to be a feasible solution, especially for the mapping of land cover over large areas.

However, there are challenging factors that limit the effective operational use of the fused datasets [25,26]. For example, geometric and radiometric registration between two datasets is demanding, because of the fact that data are normally acquired at different times, using different sensors. It is also costly to make measurements with two sensors, particularly in the boreal forest zone where the measurements can seldom be carried out during a single flight, because ALS measurements can be taken two to four times longer than aerial/hyperspectral measurements during a day, since ALS does not depend on sun light illumination. Furthermore, in contrast to passive imagery, laser scanning always views the targets at the zero degree phase angle in a narrow off-nadir viewing geometry and the transmitted energy is also controllable, thus the interpretation of the laser intensity is less complex than in the case of passive airborne images [27]. The recently developed multispectral laser scanning technique is therefore becoming an attractive option for forest mapping, because it can provide not only a dense point cloud, but also spectral information which can simplify data processing and facilitate the interpretation of data. There are a couple of studies that have demonstrated the potential of multispectral ALS for classifying tree species [28,29]. In Lindberg et al [28], multispectral data were acquired with separate instruments and from different flights—an analogue to Titan multispectral data. The study described the characterization of tree species from ALS data, using three wavelengths: 1064 nm, 1550 nm, and 532 nm, and a point density over 20 point/m^2. However, classification accuracy was not reported. In St-Onge and Budei [29], values for the mean and standard deviation of the intensity in three channels of Titan multispectral ALS, were used in the classification of broadleave vs. needleleaf trees (level 1), and eight genera (level 2) in a suburb of the city of Toronto, Canada. Random forest classification produced a classification error of 4.59% in the case of the level 1 classification (broadleave vs. needleleaf trees), and of 24.29% in the case of the level 2 classification. The point density of the data used was 10.6 first returns/m^2 per channel. Currently, the cost of data acquisition of multispectral ALS is relatively higher than that of aerial images and ALS data, if they are acquired from the same flight. However, it is expected that this cost will decrease in the future, as the technology advances. Therefore, it is worth investigating the potential of multispectral laser scanning for forest inventories, particularly for tree species classification. The objectives of this study are to evaluate the feasibility of multispectral ALS data for tree species classification with intensive field measurements, and to investigate the information content of features derived from both point cloud and intensity. The study was conducted in a boreal forest using 1903 trees in 22 plots.

2. Study Area and Materials

2.1. Test Site

The 5 km × 5 km study area, located in Evo, southern Finland (61.19°N, 25.11°E), belongs to the southern Boreal Forest Zone. It contains approximately 2000 ha of managed boreal forest, having an average stand size of slightly less than 1 ha. The area comprises a broad mixture of forest stands, varying from natural to intensively managed forests. The elevation of the area varies from 125 m to 185 m above sea level. Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) are the dominant tree species in the study area, and contribute 40% and 35% of the total volume, respectively, whereas the share of deciduous trees (mainly birches, *Betula* sp.) constitutes only 24% of the total volume.

2.2. Field Measurements

Field measurements were undertaken in the summer of 2014 and consisted of individual tree measurements for 91 plots in Evo. Sample plots, with dimensions of 32 m × 32 m, were selected, based on the prestratification of ALS data to distribute plots over various stand height and density classes. Sample plot locations were determined using the geographic coordinates of the plot center and its four corners. Plot center positions were measured using a total station (Trimble 5602), which was oriented to the local coordinate system using ground control points measured with VRS-GNSS (Trimble R8) in open areas, close to the plot. Terrestrial laser scanning was also used to assist tree mapping in the field.

After field measurements had been made, the tree map was further verified by comparing it with ALS data. If there was a discrepancy between the two data, the plot was manually corrected to match the ALS data, ensuring a positional accuracy of 0.5 m. Detailed information on the establishment of the sample plots can be found in Yu et al. [30].

From the sample plots, all trees with a diameter at beast height (DBH) exceeding 5 cm, were tallied with steel calipers from two directions perpendicular to each other, and a mean was taken as the value for the DBH. Tree height was measured using an electronic hypsometer. Height measurement accuracy is expected to be approximately 0.5 m. Tree species was also recorded. Among 91 sample plots, 22 plots were fully covered by the airborne laser scanning data and used in this study (Figure 1). The descriptive statistics of 22 sample plots, and the sample trees by species, are summarized in Tables 1 and 2, respectively.
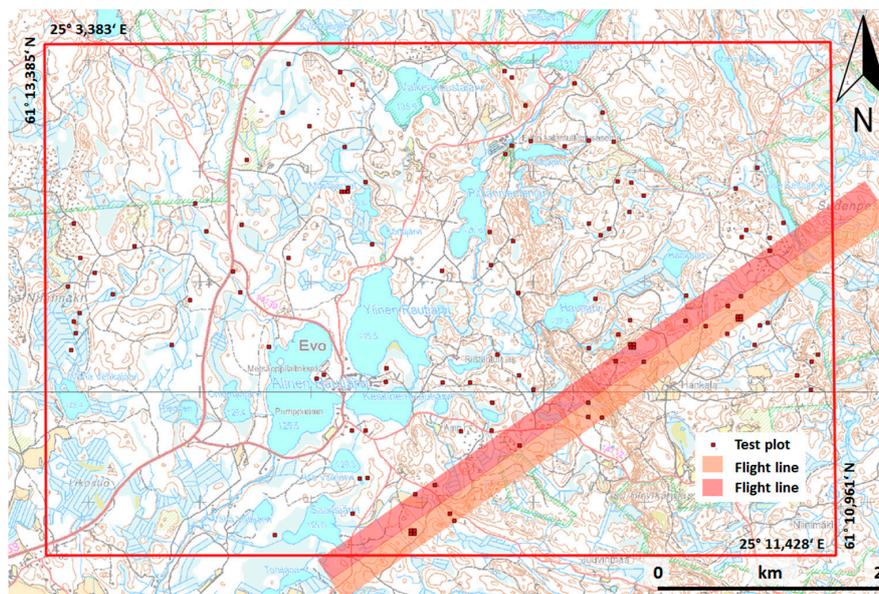

Figure 1. Study area, airborne laser scanning coverage, and sample plots.

Table 1. The descriptive statistics of Lorey's height (H_g), basal area weighted mean diameter (D_g), basal area (G), stem volume (VOL), aboveground biomass (AGB), and trees per hectare (TPH) in the 22 sample plots.

	Minimum	Maximum	Mean	Standard Deviation
H_g (m)	10.02	31.09	21.09	4.41
D_g (cm)	13.92	46.42	25.78	7.50
G (m^2/ha)	6.60	43.17	26.79	7.83
VOL (m^3/ha)	34.46	518.39	270.14	110.04
AGB (Mg/ha)	19.06	230.63	134.49	48.33
TPH (trees/ha)	342	3057	940	554

Table 2. The descriptive statistics of sample trees by tree species.

		Minimum	Maximum	Mean	Standard Deviation	Number of Trees
Pine	Tree height (m)	2.30	28.20	17.29	4.76	839
	DBH (cm)	5.00	39.80	19.37	6.92	
Spruce	Tree height (m)	2.20	35.30	14.32	8.94	630
	DBH (cm)	5.00	57.90	16.27	11.75	
Birch	Tree height (m)	2.00	30.20	16.89	4.80	434
	DBH (cm)	5.10	55.80	14.61	6.42	

2.3. Airborne Laser Scanning Data

Airborne laser scanning data were acquired on the 21st of August 2015, using an Optech Titan multispectral system, operating at a pulse rate of 300 kHz per channel. Optech Titan is the first commercial airborne laser scanner which operates with three channels. The spectral channels are two infrared ones, of 1550 nm (channel 1) and 1064 nm (channel 2), and a green channel of 532 nm (channel 3). The three channels are oriented in different directions, so that the 1064 nm channel is pointing nadir, the 1550 nm channel is positioned 3.5 degrees forward, and the 532 nm channel is positioned 7 degrees forward. As a result, laser pulses are not registered from exactly the same location in each channel. The data in this study were collected from an altitude of 400 m above sea level, resulting in an average pulse density of approximately 3×21 pulses per m^2, and the footprint sizes in diameter were 14 cm in channel 1 and 2, and 28 cm in channel 3 (beam divergence of 0.35 mrad in channel 1 and 2, and 0.7 mrad in channel 3). The system was configured to record up to five echoes per pulse, and intensities were also recorded for each return and channel.

3. Methods

3.1. Preprocessing of Multispectral ALS data

Recorded intensity is the amount of energy reflected back (i.e., backscattered) to the laser sensor, which is a function of several variables, such as target surface characteristics (reflectance, wetness and roughness), environmental effects (atmospheric transmittance, moisture), and acquisition parameters and instruments [31,32]. It is therefore necessary to calibrate intensity values for compensating the impact of these factors and achieving better classification accuracy. In this study, a simplified model was used for the return intensity calibration, in order to correct for range according to the Equation (1) with an exponential factor of 2.5 [27] for forest area, since the environmental factors can be considered stable, and the same acquisition parameters and instruments were maintained during the survey.

$$I_c = I * \left(\frac{R}{Rs} \right)^{2.5} \tag{1}$$

where I_c is the normalized intensity, I is the raw intensity, R is the sensor to target range, and Rs is the reference range or average flying height (in this study Rs = 400 m). The physical explanation for the exponential factor of 2.5 is that the laser beam is affected by the mixture of targets illuminated by the laser beam, such as leaves, dense needle groups (exponential factor close to 2), and the branches and needles (exponential factor close to 3). Correction was separately completed for each channel.

Strip matching between flight lines and between channels was performed by the data provider. Afterwards, the ALS point clouds were processed to separate ground returns from vegetation returns, using the progressive triangulated irregular network (TIN) densification method proposed by Axelsson [33]. Point cloud of channel 2 was used in this process, in order to reduce the amount of data provided that ground returns were dense enough to represent the variation of the terrain. The ALS data from the three channels were then normalised by removing ground elevation from the laser height measurements based on the digital terrain model created from classified ground points. The normalized point cloud was further processed for individual tree detection.

3.2. Individual Tree Detection

Individual trees were detected using a minimum curvature-based algorithm [34], which started with the creation of canopy height model (CHM). The method has two major steps: firstly, the tree tops were found by a local maximum filtering algorithm. Secondly, tree crowns were delineated using a watershed algorithm. CHM was created by taking the maximum value of normalized laser points within a grid cell of size 0.5 m. In the first step, CHM was smoothed by Gaussian filtering and stretched by minimum curvature, and then local maxima were detected from processed CHM.

These local maxima were considered as tree tops and used as seeds in the following step, where crown was delineated by a marker-controlled watershed algorithm, with a background mask of a 2 m height threshold, i.e., if the CHM value was less than 2 m, the pixel was classified as "background". During the segmentation processes, the tree crown shape and location of individual trees were determined, based on the segment outline and the location of maximum hit within the segment. In this study, points of first returns from all three channels were used to create CHM.

Detected trees were then linked with the trees measured in the field by an automatic matching algorithm based on the Hausdorff distance [14]. In the matching procedure, the distance in 3D space between the detected tree and the field-measured tree was used as a matching criterion. If a field-measured tree and a detected tree were the closest to each other, and the distance between them was less than a threshold, the tree was considered as correctly detected. Given the possible difference in tree location measurements from ALS data (at tree top), and in the field (at tree root) and tree height underestimation by laser scanning, a 5 m threshold distance was used to reject a match. The field-measured trees without any link to a tree segment were considered as non-detectable trees, resulting in an omission error, and a tree segment without a link to a reference tree resulted in a commission error.

3.3. Features Derivation from Multispectral ALS Data

In order to classify and characterize the object properly, we can use geometry (from point clouds) and spectral information. For each extracted tree segment, several features were derived from multispectral ALS data and used in tree species classification. They can be grouped into three categories: point cloud features, single-channel intensity (SCI) features and multi-channel intensity (MCI) features. For point cloud features, points falling within each individual tree segment were extracted from all returns, and the normalized heights of these points were used for deriving the tree features. The features were calculated based on points over a height threshold of 2 m above ground from all channels, including maximum height (Hmax), mean height (Hmean), standard deviation of height (Hstd), range of the height (Hrange) represented by the difference between the lowest and highest points, penetration rate as the ratio of points below 2 m to the total number of points, crown area (CA) and volume (CV) estimated by a 2D and 3D convex hull of the points, and crown diameter (CD). In addition, height percentiles (HP_{10} to HP_{90}) from 10% to 90%, with an increment of 10%, were calculated. Furthermore, density-related features were calculated by dividing the height into 10 equal intervals, and calculating the ratio of points within each interval to the total number of points (D_1 to D_{10}). As SCI features, we calculated the minimum, maximum, mean, standard deviation, skewness, kurtosis of the intensity, and percentiles of intensity at 5%, and from 10% to 90%, with a 10% increment for each channel. MCI features included the intensity ratio between each channel and the sum of all channels, and the subtraction of channel 2 and 3, divided by the sum of channel 2 and 3. In total, 145 tree features were generated and used in the analysis. More detailed definitions and explanations are given in Table 3.

Table 3. Tree features derived from normalized point data and spectral information. (superscript i = 1, 2, and 3 for channel 1, 2, and 3)

Feature	Definition
Point cloud features	
Hmax	Maximum of the normalized heights of all points
Hmean	Arithmetic mean of normalized height of all points above 2 m threshold
Hstd	Standard deviation of normalized height of all points above 2 m threshold
Hrange	Range of normalized height of all points above 2 m threshold
P	Penetration as a ratio between number of returns below 2 m and total returns
CA	Crown area as the area of convex hull in 2D
CV	Crown volume as the convex hull in 3D
CD	Crown diameter calculated from crown area considering crown as a circle.

Table 3. *Cont.*

Feature	Definition
HP_{10} to HP_{90}	10% to 90% percentiles of normalized height of all points above 2 m threshold with a 10% increment
D_1 to D_{10}	$D_i = N_i/N_{total}$, where i = 1 to 10, N_i is the number of points within ith layer when tree height was divided into 10 intervals starting from 2 m, N_{total} is the number of all points.
Single-channel Intensity features	
I^i_{min}	Minimum of intensity
I^i_{max}	Maximum of intensity
I^i_m	Mean of intensity
I^i_{std}	Standard deviation of intensity
I^i_{sk}	Skewness of intensity
I^i_{range}	Range of intensity
I^i_{kut}	Kurtosis of intensity
$I^i_{5,10 \text{ to } 90}$	Percentiles of intensity at 5% and from 10% to 90% with 10% increment.
Multi-channel intensity features	
$R^i_F = I^i_F/(I^1_F + I^2_F + I^3_F)$	Ratios of intensity features, F refers to different single-channel intensity features.
$N_F = (I^2_F - I^3_F)/(I^2_F + I^3_F)$	Index of intensity features

3.4. Feature Selection and Tree Species Classification

Introduced by Breiman [35], random forests (RF) is a technique which consists of an ensemble of decision trees, using a majority vote for the final prediction. RF has shown successful performances in many applications, such as in the classification of urban scenes [36] and forest attribute prediction [34,37]. In this study, tree species were estimated based on prediction models by RF using tree features as predictors and tree species as a response for correctly detected trees. Although RF is able to deal with high dimensional data [38], the results of classification can be significantly improved if only the most important features are used [39]. Considering the number of observations and the correlation between the features in this study, it was necessary to reduce the feature dimension to avoid overfitting. The RF built-in measure of feature importance was used to search for a subset of predictors that optimally model responses, subject to constraints which minimize the correlation among the features. In this study, 15 of the most important features were selected for each experimental setting by measuring how influential the predictor was at predicting the response. The parameter settings for RF in each classification were as follows: 200 decision trees were built, with four predictors being randomly selected for the best splitting at the nodes, when decision trees were built.

3.5. Evaluation of Accuracy

The accuracy of tree species classification was evaluated by comparing the classified tree species with the reference tree species recorded in the field for correctly detected trees. The result of the comparison can be represented by an error matrix. Four widely-used measures, i.e., producer's accuracy, user's accuracy, overall accuracy (OA), and Kappa coefficient, were computed for evaluating the performance of the classification. To avoid overfitting of the classification model, independent validation was conducted by equally dividing available data into two sets: one for training the classification model, and the other for testing the performance.

We evaluated different combinations of extracted features for their predictive power as follows: (i) point cloud features as predictors, (ii) SCI features as predictors, (iii) MCI features as predictors, (iv) point cloud and SCI features as predictors, and (v) all features as predictors. The McNemar test was used to determine whether there are statistically significant differences between pairs of classifications, with the different predictor settings mentioned above (e.g., point cloud features vs. SCI features, SCI vs. MCI features vs. all features, and so on).

We also classified four categories of trees, to analyze how crown positioning affected classification accuracy. Thus, the field-measured trees were divided into four categories, based on the distance and height difference of neighbor trees as follows:

- Isolated or dominant trees that are well separated from other trees (distance to neighbor trees is greater than 3 m or tree height is greater than neighbor tree by 2 m) (referred to as C1).
- Group of trees: trees are growing closely to each other (distance less than 3 m) and have a similar height (height difference is less than 2 m) (referred to as C2).
- Tree next to a larger tree: the distance of a tree to a neighbor tree is greater than 1.5 m and the height is smaller than the height of neighbor tree by at least 2 m (referred to as C3).
- Tree under a larger tree: the distance of a tree to a neighbor tree is less than 1.5 m and the height is smaller than the height of neighbor tree by at least 2 m (referred to as C4).

The number of trees in each category was 580 in C1, 552 in C2, 590 in C3, and 181 in C4.

4. Results

4.1. Accuracy of Individual Tree Detection

The accuracy of individual tree detection was evaluated by comparing tree segments with field reference data. Overall, out of 1903 trees, 61.3% of trees were correctly detected. Most of the undetectable specimens were understory trees and trees that were near to a larger tree. At plot level, the detection rate varied between 50% and 98%. In the dense plots, the tree detection rate was lower than that in the sparse plots. The detection rate was also affected when the plot was located near the boundary of the data coverage, where the point distribution was not optimum, i.e., the points in one direction were denser than in a perpendicular direction. When considering the different categories of trees, the detection rate was 91.9% for C1, and 81.5%, 26.4%, and 7.2%, for C2, C3, and C4, respectively. A higher detection rate was expected for trees in C1, because the crown boundary was well defined. For trees in C2, there was a tendency to merge trees into one segment if they were close to each other, whereas in C3, trees were more likely to merge with neighbor trees, leading to a low detection rate. In C4, trees were often not detectable because individual tree detection was based on a CHM where taller trees overtopped the tree underneath. An example of individual tree detection is shown in Figure 2.
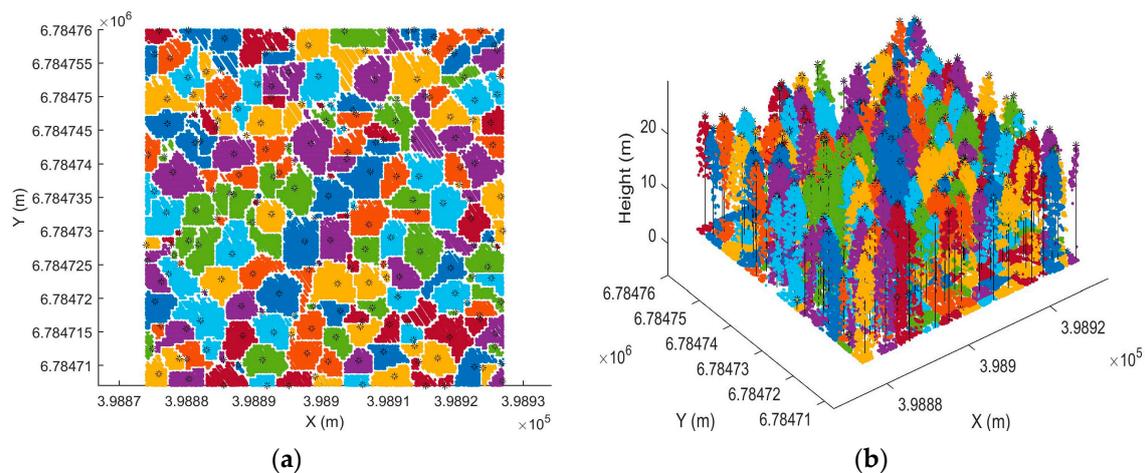

Figure 2. Result of individual tree detection for one plot. (**a**) top view, (**b**) 3D view.

4.2. Classification with Different Combinations of Features

The confusion matrix of classification and the result of accuracy evaluation are presented in Tables 4–8 for the species classification based on the different combination of features, i.e., point cloud features alone (Table 4), SCI features (Table 5), MCI features (Table 6), point cloud and SCI features (Table 7), and all features combined (Table 8). The highest level of accuracy (85.9%) was obtained with a combination of point cloud and SCI features. Point cloud features alone produced the lowest overall accuracy of 76.0%, while single-channel intensity features produced an overall accuracy of 85.4%. A McNemar test indicated no significant difference between classifications based on SCI and all features, at a 5% significant level ($p = 0.69$). Additionally, there was no difference between classifications based on SCI features, and the combination of point cloud and SCI features ($p = 0.58$). This suggested that point cloud features did not provide more information for classification. McNemar tests showed that the difference between classifications based on other pairs of features, were all significant at a 5% significant level (Table 9). Classification accuracy also varied between species. The best accuracies were obtained for pine trees with a 97.5% producer's accuracy using point cloud and SCI features, and for spruce trees with a 78.2% producer's accuracy using SCI features, followed by birch trees with a 71.8% producer's accuracy using all features. SCI features produced slightly better results than MCI features ($p = 0.03$). The results suggested that MCI features do not provide more information than SCI features.

Table 4. Confusion matrix and accuracy evaluation of classification with 15 selected point cloud features and test data.

		Predicted			Producer (%)
		Pine	Spruce	Birch	
Reference	Pine	294	10	22	90.18
	Spruce	24	84	11	70.59
	Birch	54	17	60	45.80
User (%)		79.03	75.68	64.52	OA = 76.04%, Kappa = 0.57

Table 5. Confusion matrix and accuracy evaluation of classification with 15 selected single-channel intensity features and test data.

		Predicted			Producer (%)
		Pine	Spruce	Birch	
Reference	Pine	306	8	12	93.87
	Spruce	15	93	11	78.15
	Birch	24	14	93	70.99
User (%)		88.70	80.87	80.17	OA = 85.42%, Kappa = 0.75

Table 6. Confusion matrix and accuracy evaluation of classification with 15 selected multi-channel intensity features and test data.

		Predicted			Producer (%)
		Pine	Spruce	Birch	
Reference	Pine	299	9	17	92.00
	Spruce	19	85	16	70.83
	Birch	23	22	86	65.65
User (%)		87.68	73.28	72.27	OA = 81.60%, Kappa = 0.68

Table 7. Confusion matrix and accuracy evaluation of classification with 15 selected point cloud and single-channel intensity feature combination and test data.

		Predicted			Producer (%)
		Pine	**Spruce**	**Birch**	
Reference	Pine	317	3	5	97.54
	Spruce	18	85	17	70.83
	Birch	29	9	93	70.99
User (%)		87.09	87.63	80.87	OA = 85.94%, Kappa = 0.75

Table 8. Confusion matrix and accuracy evaluation of classification with 15 selected features among all features and test data.

		Predicted			Producer (%)
		Pine	**Spruce**	**Birch**	
Reference	Pine	306	8	11	94.15
	Spruce	13	93	14	77.50
	Birch	28	9	94	71.76
User (%)		88.18	84.55	78.99	OA = 85.59%, Kappa = 0.75

Table 9. McNemar tests on pairs of the classifications using different combination of features. The number in the table is *p* value. The number with a superscript * indicated that the difference between classifications is significant at a 5% significant level.

Feature	Point Cloud	SCI	MCI	Point Cloud + SCI
SCI	$1.4 \times 10-7$ *			
MCI	$2.4 \times 10-4$ *	0.03 *		
Point cloud + SCI	$1.3 \times 10-11$ *	0.58	0.01 *	
All	$7.2 \times 10-8$ *	0.69	0.02 *	0.66

4.3. Classifications for Four Defined Categories of Trees

We also examined classification accuracy for four categories of trees based on the 15 best features among point cloud and SCI features, because the use of all features does not improve the accuracy. A 10-fold cross-validation strategy was applied in this case because the number of trees in C3 and C4 was low. The obtained accuracies varied widely. For isolated and dominant trees, an overall accuracy of 90.47% was achieved (Table 10). The corresponding accuracy was 89.80% for trees in C2 (Table 11), 79.09% for trees in C3 (Table 12), and 53.85% for trees in C4 (Table 13). As can be seen, very high accuracy was achieved for isolated and dominant trees, and for groups of trees. The accuracy was about 36 percentage points lower for suppressed trees. For dominant trees, both pine and spruce achieved a high accuracy of over 90%, because of their well identified conical shape. For birch, moderate accuracy was obtained. Overall, pines are classified with higher accuracy and less misclassifications, while birches tend to be misclassified as pine for all four categories of trees, resulting in a low user's accuracy for pine. Spruces are more likely to be mixed with both pine and birch.

Table 10. Confusion matrix of classification with point cloud and single-channel intensity features for isolated and dominant trees (C1).

| | | Predicted | | | Producer (%) |
		Pine	Spruce	Birch	
	Pine	299	0	6	98.03
Reference	Spruce	10	128	4	90.14
	Birch	20	11	57	64.77
User (%)		90.88	92.09	85.07	OA = 90.47%, Kappa = 0.83

Table 11. Confusion matrix of classification with point cloud and intensity features for group of trees (C2).

| | | Predicted | | | Producer (%) |
		Pine	Spruce	Birch	
	Pine	268	6	2	98.41
Reference	Spruce	4	34	10	67.57
	Birch	19	5	103	77.59
User (%)		92.10	75.56	89.57	OA = 89.80%, Kappa = 0.80

Table 12. Confusion matrix of classification with point cloud and intensity features for trees next to a larger tree (C3).

| | | Predicted | | | Producer (%) |
		Pine	Spruce	Birch	
	Pine	63	1	2	95.45
Reference	Spruce	6	30	8	68.18
	Birch	10	5	28	65.12
User (%)		79.75	83.33	73.68	OA = 79.09%, Kappa = 0.67

Table 13. Confusion matrix of classification with point cloud and intensity features for trees under a larger tree (C4).

| | | Predicted | | | Producer (%) |
		Pine	Spruce	Birch	
	Pine	4	0	0	100
Reference	Spruce	1	2	2	40.0
	Birch	2	1	1	25.0
User (%)		57.14	66.67	33.33	OA = 53.85%, Kappa = 0.32

4.4. Feature Importance

We also investigated which input features and channels are most relevant for tree species classification based on the measure provided by the RF algorithm for assessing feature importance. If a feature is influential in the prediction, then permuting its values should affect the model error. If a feature is not influential, then permuting its values should have little or no effect on the model error. Table 14 lists the top five features in the classifications based on different combinations of the features. In the classification based on point cloud features, the most important features were penetration and higher level percentiles. Two density-related features at higher and middle layers were also scored as important as higher percentiles. For the case of classification based on the SCI features, the wavelength of 1064nm (Channel 2) seems to contain more information for separating pine, spruce, and birch, followed by wavelengths of 1550nm (channel 1), and 532nm (channel 3). The classification based on the features of the three separate channels also confirmed analysis with an overall accuracy of 81.9%, 78.3%, and 69.1%, for channel 2, 1, and 3, respectively. The difference between the pairs of

classification is significant at a 5% significant level based on McNemar tests. Minimum values and the 90% percentile of intensity are two of the most powerful predictors for all channels in such cases. In MCI-based classification, ratios at higher percentiles for the three channels were among the most important features. Overall, when all features were considered, the minimum intensity of channel 2 and 3, the ratio at 90% percentile for channel 2 and 3, and one point cloud features (P), are among the top five most importance features.

Table 14. The features have the most predictive power in different classification scenarios. A detailed explanation of the features can be found in Table 3. The number in parentheses is the score for the feature. The higher the score, the more important the feature.

Cases	Top 5 features
Point cloud features	P (3.8), D_9 (1.6), Hmax (1.5), D_5 (1.4), HP_{90} (1.3)
SCI features	I^2_{min} (1.9), I^2_{p90} (1.6), I^1_{sk} (1.4), I^1_{p90} (1.5), I^3_{p90} (1.5)
MCI features	R^3_{p90} (1.7), R^2_{p90} (1.4), R^2_{range} (1.4), R^1_{p80} (1.3), N_{p90} (1.3)
Point cloud and SCI features	I^2_{min} (2.0), Hmax (1.5), I^2_{p90} (1.5), I^3_{p90} (1.8), P (1.6)
All features	I^2_{min} (1.8), R^3_{p90} (1.7), P (1.5), I^3_{min} (1.4), R^2_{p90} (1.2)

5. Discussion

In this study, we explored the potential of multispectral ALS data in tree species classification of a boreal forest. Results showed that multispectral ALS data can be used to separate three main tree species, i.e., pine, spruce, and birch, with a high overall accuracy of 85.9% in the best case scenario, which was based on the combined use of point cloud and SCI features. Overall, the results indicated that the intensity of the three channels contains more information for tree species classification than point cloud data. When using the intensity of the three channels, both the producer's and user's accuracies for single tree species were improved, as well as the overall accuracy compared with the results obtained from point cloud data. However, different types of features are more influential on certain tree species. For example, intensity features are more powerful in separating birch from pine and spruce (produce's accuracy improved from 45.8% to 71%, and user's accuracy from 64.5% to 80.2%, when compared with those using point cloud features). With the inclusion of point cloud features, the classification accuracy of intensity features was improved by only 0.5 percentage point, while the corresponding value was 10 percentage points when adding intensity features to point cloud features.

The individual tree-detection rate was not very high in this study. Two factors influenced this. Firstly, individual tree detection was based on CHM, so most of the understory trees were not detectable and 3D information of the dense point cloud was not fully utilized. Secondly, distribution of the point cloud was not optimal, as it was denser in scanning direction than flight direction. The uneven distribution of points affected the results of individual tree detection, as the detection rate tended to decrease when the plot was located near the boundary of data coverage where uneven distribution was more severe. In order to improve individual tree detection, we recommend developing methods which can fully utilize the 3D information provided by point cloud. Multispectral information could also be useful for improving the accuracy of individual tree detection. When point cloud and spectral information are used in tree detection, a simultaneous classification is possible, such that the knowledge relevant to each can aid in the analysis of the other. Ultimately, this could lead to the improvement of accuracy of individual tree detection and classification, as well as computational advantages.

A large variance in feature values can be found, due to the irregular geometry of the canopy surface and varying degrees of penetration. There were more points penetrated, thus reaching the ground, in channels 1 and 3, than in channel 2. One potential factor that contributed to this was the forward viewing geometry for channels 1 and 3. There were also more returns in channels 2 and 3, than in channel 1. For the same point cloud features, the values in channel 3 were higher than those in channels 1 and 2, while channels 1 and 2 produced similar values. This trend was observed for all three species and could be one reason why the point cloud features did not significantly improve the

classification, when used with SCI features. SCI features also overlapped between species. However, the degree of overlap varied among the features and channels. In general, higher percentiles of the intensity distribution and minimum intensity value were more separated than the lower percentiles. For example, the maximum intensity was smaller for pine than spruce in both channel 1 and channel 2, while similar values were observed for pine and spruce in channel 3. There were more overlapping values and variations at lower percentiles of intensity distribution, among tree species in all channels.

MCI features have been used to reduce the radiometric effects on multispectral images and improvements in classification have been reported. In this study, the use of similar ratios and indices did not improve the classification. The reason for this could be that the laser scanner is an active instrument, and recorded intensity mainly depends on the instrument design, measurement range, and reflectance of the targets. If the same instrument has been used for data acquisition and the range effect has been corrected, the major factor affecting recorded intensity is the targets illuminated by the laser. Therefore, the intensity itself is good enough to characterize the objects.

The results in this study are in agreement with previous results, in which tree species were classified using ALS combined with multi/hyper-spectral data, although the studies cannot be compared directly because of the differences in the data used, and the number and type of species identified. For example, Dalponte et al [40] reported a kappa accuracy of 0.89 when classifying three boreal tree species (pine, spruce and broad-leaves), using hyperspectral and ALS data with the manual detection of trees. The higher kappa coefficient obtained in their study could be a result of better delineation of individual trees by manual detection, and a higher spectral resolution. Jones et al [41] achieved an overall accuracy of 73% for classifying 11 species in coastal south-western Canada, using hyperspectral and ALS data. The lower accuracy could be explained by the higher number of species recognised in the study. This indicates that multispectral ALS data contains similar information to the fusion of multispectral images and ALS data. Compared with the previous study, which used a multispectral ALS of similar density for tree species classification, St-onge and Budei [29] reported a classification error of 4.59% in the case of the level 1 classification (broadleave vs. needleleaf trees), and of 24.29% in the case of level 2 classification (eight genera), using intensity features (mean and standard deviation of intensity in three channels). The different number of species could be the reason for the difference in accuracy.

The use of a single source of data apparently has advantages over the use of fused data, with respect to data processing. For example, geometric and radiometric calibrations between different data sources produce big challenges, and require much effort to compensate the changes in illumination conditions and vegetation [26]. Furthermore, previous studies have shown that background signal reduced classification accuracies when using multispectral/hyperspectral images [42–44]. In contrast, multispectral ALS data can easily separate the reflections of vegetation from the reflections of the ground, thus background influences on the results, like soil, could be minimised. Therefore, the accuracy of the classification could be improved with the use of multispectral ALS data. However, this issue needs to be explored further in order to investigate the extent to which the accuracy can be improved with multispectral ALS data.

The intensity values of different returns are affected by the vertical structures of trees. In theory, the intensity of only returns can be radiometrically calibrated with high reliability. The first of many returns is distorted by the signal penetrating to the second and other layers. However, there is still valuable information of all return intensities confirmed by this study. In the future, it should be studied whether it is possible to calibrate the intensities of multiple returns in a better way, by taking into account the attenuated part of the signal and the part that causes other returns.

The major drawback with applied Titan data was the inhomogeneous distribution of the point cloud. In the across track, the point spacing was significantly smaller than that in the along track. Either lower aircraft speed or higher scan frequency should be achieved to provide more homogenous point spacing. Another drawback is that the points from the three channels are not registered from the same location, which means that it is not multispectral data in the conventional sense. As a result,

pixel/point wise classification cannot be performed; instead, object-based analysis has to be carried out, like in this study. The accuracy of the classification may also deteriorate, because the backscatter from different channels could come from different parts of the objects. The impact of such system design on classifications needs further investigation. Regardless of these drawbacks, multispectral ALS data are still a valuable data source for tree species classification, as shown in this study.

Currently, it is more expensive to acquire multispectral ALS data than aerial images and single-channel ALS. However, it is anticipated that the price will drop as technology develops, and the market is growing. Furthermore, ALS data can be acquired during both the day and night, which partly compensates for the cost of the data acquisition. Therefore, multispectral ALS data could be a cost-effective solution for species-specific mapping of forests in the future, and it has the potential to increase the automation of the whole processing chain.

6. Conclusions

In this study, we assessed the potential utility of single-sensor multispectral ALS data for tree species classification in mixed coniferous forests in a boreal zone. The results suggest that additional information, provided by multispectral laser scanning, may be a valuable source of information for tree species classification of pine, spruce, and birch, which are the main tree species found in boreal forest zones. The best overall classification accuracy achieved was 85.9% using point cloud and SCI features, which was not significantly different from the ones in which all features, or solely SCI features were used. Point cloud features alone achieved an accuracy of 76.0%. Channel 2 performed the best when separating pine, spruce, and birch, followed by channel 1 and channel 3, with overall accuracies of 81.9%, 78.3%, and 69.1%, respectively.

This preliminary study has demonstrated the potential of multispectral airborne laser scanning for possible future solutions for automatic single-sensor forest mapping. It is expected that multispectral airborne laser scanning can provide highly valuable data for forest mapping. However, there are many aspects of multispectral ALS that need to be investigated further, for example: how will multispectral ALS data perform in other forest zones where the number of species composition is higher? Is it possible to derive more useful features to improve the classification? From a practical point of view, future studies could explore the possibility to improve the accuracy of forest inventory mapping using species information obtained from this study.

Author Contributions: X. Yu and J. Hyyppä designed the experiments. X. Yu carried out the research, analyzed the data and wrote the first draft of the paper. P. Litkey performed intensity calibration. M. Holopainen and M. Vastaranta were responsible for the field measurements. H. Kaartinen contributed material collection. All co-authors assisted in writing and improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gillis, M.; Leckie, D. Forest inventory update in Canada. *For. Chron.* **1996**, *72*, 138–156. [CrossRef]
2. Waser, L.T.; Ginzler, C.; Kuechler, M.; Baltsavias, E.; Hurni, L. Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sen. Environ.* **2011**, *115*, 76–85. [CrossRef]
3. Lovell, J.L.; Jupp, D.L.; Culvenor, D.S.; Coops, N.C. Using airborne and ground based ranging LiDAR to measure canopy structure in Australian forests. *Can. J. Remote Sens.* **2003**, *29*, 607–622. [CrossRef]
4. Coops, N.C.; Hilker, T.; Wulder, M.A.; St-Onge, B.; Newnham, G.; Siggins, A.; Trofymow, J.T. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* **2007**, *21*, 295–310. [CrossRef]

5.   Wulder, M.A.; White, J.C.; Nelson, R.F.; Næsset, E.; Ørka, H.O.; Coops, N.C.; Hilker, T.; Bater, C.W.; Gobakken, T. LiDAR sampling for large-area forest characterization: A review. *Remote Sens. Environ.* **2012**, *121*, 196–209. [CrossRef]

6.   Næsset, E.; Økland, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sens. Environ.* **2002**, *79*, 105–115. [CrossRef]

7.   Clark, M.L.; Clark, D.B.; Roberts, D.A. Small-footprint LiDAR estimation of subcanopy elevation and tree height in a tropical rain forest landscape. *Remote Sens. Environ.* **2004**, *91*, 68–89. [CrossRef]

8.   Holmgren, J.; Persson, Å. Identifying species of individual trees using airborne laser scanning. *Remote Sens. Environ.* **2004**, *90*, 415–423. [CrossRef]

9.   Brandtberg, T. Classifying individual tree species under leaf-off and leaf-on conditions using airborne LiDAR. *ISPRS J. Photogramm. Remote Sens.* **2007**, *61*, 325–340. [CrossRef]

10.   Lindberg, E.; Eysn, L.; Hollaus, M.; Holmgren, J.; Pfeifer, N. Delineation of tree crowns and tree species classification from full-waveform airborne laser scanning data using 3-D ellipsoidal clustering. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3174–3181. [CrossRef]

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

12.   Hollaus, M.; Wagner, W.; Maier, B.; Schadauer, K. Airborne laser scanning of forest stem volume in a mountainous environment. *Sensors* **2007**, *7*, 1559–1577. [CrossRef]

13.   Ahmed, R.; Siqueira, P.; Hensley, S. A study of forest biomass estimates from LiDAR in the northern temperate forests of New England. *Remote Sens. Environ.* **2013**, *130*, 121–135. [CrossRef]

14.   Yu, X.; Hyyppä, J.; Kukko, A.; Maltamo, M.; Kaartinen, H. Change detection techniques for canopy height growth measurements using airborne laser scanning data. *Photogram. Eng. Remote Sens.* **2006**, *72*, 1339–1348. [CrossRef]

15.   Yu, X.; Hyyppä, J.; Kaartinen, H.; Maltamo, M.; Hyyppä, H. Obtaining plotwise mean height and volume growth in boreal forests using multi-temporal laser surveys and various change detection techniques. *Int. J. Remote Sens.* **2008**, *29*, 1367–1386. [CrossRef]

16.   Ørka, H.O.; Næsset, E.; Bollandsås, O.M. Utilizing airborne laser intensity for tree species classification. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *36*, 300–304.

17.   Suratno, A.; Seielstad, C.; Queen, L. Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 683–693. [CrossRef]

18.   Ørka, H.O.; Naesset, E.; Bollandsas, O.M. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sens. Environ.* **2009**, *113*, 1163–1174. [CrossRef]

19.   Korpela, I.; Ørka, H.O.; Maltamo, M.; Tokola, T. Tree species classification using airborne LiDAR-effects of stand and tree parameters, downsizing of training set, intensity normalization and sensor type. *Silva Fenn.* **2010**, *44*, 319–339. [CrossRef]

20.   Yao, W.; Krzystek, P.; Heurich, M. Tree species classification and estimation of stem volume and DBH based on single. *Remote Sens. Environ.* **2012**, *123*, 368–380. [CrossRef]

21.   Heinzel, J.; Koch, B. Exploring full-waveform LiDAR parameters for tree species classification. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 152–160. [CrossRef]

22.   Cao, L.; Coops, N.C.; Innes, J.L.; Dai, J.; Ruan, H. Tree species classification in subtropical forests using small-footprintfull-waveform LiDAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 39–51. [CrossRef]

23.   Naidoo, L.; Cho, M.A.; Mathieu, R.; Asner, G. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 167–179. [CrossRef]

24.   Zhou, W.; Huang, G.; Troy, A.; Cadenasso, M. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. *Remote Sens. Environ.* **2009**, *113*, 1769–1777. [CrossRef]

25.   Packalén, P.; Suvanto, A.; Maltamo, M. A two stage method to estimate species-specific growing stock. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1451–1460. [CrossRef]

26.   Puttonen, E.; Suomalainen, J.; Hakala, T.; Räikkönen, E.; Kaartinen, H.; Kaasalainen, S.; Litkey, P. Tree species classification from fused active hyperspectral reflectance and LiDAR measurements. *For. Ecol. Manag.* **2010**, *260*, 1843–1852. [CrossRef]

27. Korpela, I.; Orka, H.; Hyyppa, J.; Heikkinen, V.; Tokola, T. Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 369–379. [CrossRef]

28. Lindberg, E.; Briese, C.; Doneus, M.; Hollaus, M.; Schroiff, A.; Pfeifer, N. Multi-wavelength airborne laser scanning for characterization of tree species. In Proceedings of SilviLaser 2015, La Grande Motte, France, 28–30 September 2015; pp. 271–273.

29. St-Onge, B.; Budei, B.C. Individual tree species identification using the multispectral return intensities of the Optech Titan LiDAR system. In Proceedings of SilviLaser 2015, La Grande Motte, France, 28–30 September 2015; pp. 71–73.

30. Yu, X.; Hyyppä, J.; Karjalainen, M.; Nurminen, K.; Karila, K.; Vastaranta, M.; Kankare, V.; Kaartinen, H.; Holopainen, M.; Honkavaara, E.; et al. Comparison of laser and stereo optical, SAR and InSAR point clouds from air- and space-borne sources in the retrieval of forest inventory attributes. *Remote Sens.* **2015**, *7*, 15933–15954. [CrossRef]

31. Bright, B.C.; Hicke, J.A.; Hudak, A.T. Estimating aboveground carbon stocks of a forest affected by mountain pine beetle in Idaho using LiDAR and multispectral imagery. *Remote Sens. Environ.* **2012**, *124*, 270–281. [CrossRef]

32. Ahokas, E.; Hyyppä, J.; Yu, X.; Liang, X.; Matikainen, L.; Karila, K.; Litkey, P.; Kukko, A.; Jaakkola, A.; Kaartinen, H.; et al. Towards automatic single-sensor mapping by multispectral airborne laser scanning. In Proceedings of XXIII ISPRS Congress, Commission III, Prague, Czech Republic, 12–19 July 2016; pp. 155–162. [CrossRef]

33. Axelsson, P. DEM generation from laser scanner data using adaptive TIN models. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2000**, *33*, 110–117.

34. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37. [CrossRef]

35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

36. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne LiDAR and multispectral image data for urban scene classification using Random Forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66. [CrossRef]

37. Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Hall, D.E.; Falkowski, M.J. Nearest neighbour imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sens. Environ.* **2008**, *112*, 2232–2245. [CrossRef]

38. Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef] [PubMed]

39. Millard, K.; Richardson, M. On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping. *Remote Sens.* **2015**, *7*, 8489–8515. [CrossRef]

40. Dalponte, M.; Ørka, H.O.; Ene, L.T.; Gobakken, T.; Næsset, E. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sens. Environ.* **2014**, *140*, 306–317. [CrossRef]

41. Jones, T.G.; Coops, N.C.; Sharma, T. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sens. Environ.* **2010**, *114*, 2841–2852. [CrossRef]

42. Shang, X.; Chisholm, L.A. Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2481–2489. [CrossRef]

43. Adelabu, S.; Mutanga, O.; Adam, E.; Cho, M.A. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *J. Appl. Remote Sen.* **2013**, *7*. [CrossRef]

44. Carleer, A.; Wolff, E. Exploitation of very high resolution satellite data for tree species identification. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 135–140. [CrossRef]