Using multispectral ALS for tree species identification

Arvid Axelsson

Using multispectral ALS for tree species identification

Arvid Axelsson

Keywords: LiDAR, indvidual trees, ITC, spectral

Preface and acknowledgements

This master's thesis is written as the final part of my degree as Jägmästare, a professional master of science in forestry. The thesis is written over the course of two semesters, i.e., 60 credits. An important goal when writing my master's thesis with this scope has been to keep it at a level where the results are possible to publish in a scientific journal.

Abstract

Accurate and large area tree species classification is an important subject with problems that have not yet been completely solved. For both nature conservation and wood production purposes, a detailed description of tree species composition would be useful. The objective of this master's thesis is to explore how tree species differ in spectral and structural properties using multispectral airborne laser scanning data from the Optech Titan X system. Remote sensing data was gathered from Remningstorp, Västra Götaland in Sweden on 21st July 2016. Field data contained 179 solitary trees from nine species. Two new methods for feature extraction are tested and compared to features of height and intensity distributions. The features that were most important for tree species classification were those from the upper part of the crown. Spectral features provided a better basis for tree species classification than structural features. Using single, first or all returns gave only a small difference in cross-validation correctness rate. The best classification model was created using multispectral distribution features of all returns, with an correctness rate of 77.09 %. Spruce and pine had a 100 % overall classification accuracy and were not confused with any other species. Linden was the deciduous species with a large sample that was most frequently confused with many other deciduous species.

1 Introduction

Airborne laser scanning (ALS) has been used to produce nationwide estimations of forest variables, such as height, volume and basal area, with high accuracy (Næsset et al. 2004; Nilsson et al. 2017; Skogsstyrelsen 2016). Estimations are made by using regression analysis of height distribution and density features of the ALS-derived point cloud. As an alternative to the common subjective, manual inventory used when making a forest management plan, an efficient, automated and objective forest inventory can be performed with only a few existing field plots and remote sensing data. Some forest properties however, have not yet been possible to estimate using remote sensing techniques, e.g., nature conservation values and tree species.

In Sweden, half of the close to 2,300 forest-dwelling and red-listed species are dependent on deciduous tree species and many of these rely on a single tree species for survival (Sandström et al. 2015). By using information on stand tree species composition and distributions in the landscape, it is possible to make predictions on where certain threatened species might occur. If species of trees in a stand can be classified, the possible habitats in that area may also be evaluated (Reese et al. 2002).

The most common groups to use when classifying tree species in Sweden are *spruce*, *pine*, and *other*, all deciduous species fall in the *other*-class (Dalponte et al. 2013; Holmgren and Persson 2004). Some have tried to classify different deciduous species, but with moderate success (Brandtberg 2007; Brandtberg et al. 2003). Several different data sources have been used, from passive multispectral array sensors (Reese et al. 2003) to terrestrial laser scanning (Lin and Herold 2016), as well as combinations of data sources (Holmgren, Persson, and Söderman 2008).

Multispectral data have previously only been obtainable through passive optical sensors, such as cameras and electro-optical scanners, where an average brightness within each pixel is recorded (Lillesand, Kiefer, and Chipman 2007). Manual stereo interpretation of aerial photographs have long been the standard when making estimations of vegetation (Åge 1983). The method relies on differences in reflectance and texture visible in photographs. In general, deciduous species will reflect more light in the near infrared spectrum than coniferous, and spruce have different crown structure than pine (Axelson and Nilsson 1993). Automated methods for tree species classification using aerial photography have been developed (Dalponte et al. 2013), but these deal with data from the surface of the canopy which is visible from the camera position. The top layer is the one reflecting most of the sunlight, and a measurement using passive sensors provides more information about the surface than spectral characteristics further down in the canopy. Other limitations are that shadows

and lighting conditions may affect each image differently due to varying solar angle.

The ongoing technological evolution is affecting ALS so that even more information can be produced than before, such as intensity values of returns in multiple wavelengths and full waveform returns. At the time of writing, an operational multispectral ALS system is a new technology, first used in 2015 (Matikainen, Hyyppä, and Litkey 2016), and not yet fully evaluated for forestry purposes. When using such a laser scanner, colors further down in the canopy are revealed (Cottin, Fleming, and Woodhouse 2015), and this information might be very well suited for tree species estimations (Lindberg et al. 2015). Both data from multispectral ALS, using only intensity values, (St-Onge and Budei 2015) and geometric information from ordinary ALS (Lindberg et al. 2014) has already been shown to provide a good basis for tree species classification. Equipment providing an inherent combination of these data may be the future for remote sensing in forestry.

1.1 Goals

The objective of this master's thesis is to explore how different Swedish tree species differ in spectral properties and structure as measured with multispectral ALS. Since the technology is so new, no consensus on how to use data exists and an exploration of data will be needed to know how it could be used in the future. The conventional way of estimating forest variables is by using height distributions (Næsset et al. 2004). This method uses information in one dimension only and does not utilize the three dimensional nature of ALS-data. Using distributions must not necessarily be the best method when working with multispectral data. To achieve the goal, a number of questions were to be answered:

1. What features are the most effective for tree species classification and what are their commonalities?

2. How do spectral and structural features contribute to tree species classification accuracy?

3. Should spectral and structural information from all returns, first returns or single returns be used?

4. What are the differences in classification accuracy between the best feature combination, the best combination of only spectral features and the best combination of only structural features?

2 Materials and Methods

2.1 Data

The study area was Remningstorp, located in Västra Götaland, Sweden (N 58°27′18.35″, E 13°39′08.03″). The scanned area includes the estate of Remningstorp, as well as the nature reserve of Ea hage and forested grazing lands around. Remningstorp estate has an area of more than 1,500 hectares, covered mainly with spruce in stands managed for wood volume production. Ea hage is dominated by different deciduous species, mainly oak (*Quercus robur* L.) and linden (*Tilia cordata* Mill.). The wood pastures were also mainly covered with deciduous trees but they were distributed more sparsely.

2.1.1 Laser scanner data

Multispectral laser data had been gathered on 21st July 2016 using the Optech Titan X ALS-system. Three channels was produced by the system: 1550 nm (channel 1, short wave infrared), 1064 nm (channel 2, near infrared) and 532 nm (channel 3, green) (Teledyne Optech 2015). Due to the scanning being performed at a low altitude, this data had a very high resolution and a precise measure of intensity of the return signal. To make a digital terrain model for height normalization, laser data from the Swedish land survey was used.

ALS data was in the form of discrete returns, with no full waveform data of the return pulse. Each peak intensity was identified as a return, or point, with a coordinate that can be seen as the vector from where the airplane was located at the time of measurement. The set of coordinates can be seen as a point cloud where the returns are plotted three-dimensionally.

2.1.2 Field inventory data

During the autumn of 2016, field measurements were made on 195 individual trees of nine genera (classes): spruce (*Picea abies* (L.) H. Karst), pine (*Pinus sylvestris* L.), birch (*Betula pendula* Roth and *Betula pubescens* Ehrh.), oak (*Quercus robur* L.), ash (*Fraxinus excelsior* L.), linden (*Tilia cordata* Mill.), wild cherry (*Prunus avium* L.), maple (*Acer platanoides* L.), and alder (*Alnus glutinosa* (L.) Gaertner). The only genera where more than one species were present was *Betula*, therefore the term *species* has been used to describe the other species as well as the *Betula* genera. For each tree a number of properties were recorded: species, height and live crown height. GPS-coordinates delineating the tree crown were also recorded, with a minimum of three coordinates. Trees standing close to each other compete and their crown shapes change. To

accommodate for this difference in shape, if a tree stood close to other trees, additional points were recorded between them. This was done to improve the accuracy of the crown boundary and let a deformation of the crown shape, caused by a neighbouring tree, weigh in on the final crown position and radius, as specified in section 2.2.1. The device used was an RTK-GPS that use fixed base stations to improve positioning accuracy. Live crown height was defined by the height from ground to the point where the first living branch met the stem. Any additional information that might be of value, such as diseases or damages, was also noted.

Since only one spruce (*P. abies*) with a free standing crown was found during inventory, a number of circular field plots from another inventory—first made in 2014 with a follow-up inventory in 2016—were used to get more observations in that class. These field plots were visited to measure live crown height while other variable values of the plots were available from the 2016 inventory. Trees in the plots were manually delineated in software (Quick Terrain modeler) from the multispectral point cloud. Points for delineation were chosen in the same way as those collected with GPS for other trees.

Due to the nature of wood pastures and the history of the area around Remningstorp and Ea hage, some tree species was very easily identified by height alone. Oaks in the area were very large while the largest ashes found were dead or almost dead due to ash dieback. The only ashes found that was healthy enough to have a full crown of leafs were much smaller than the oaks. To investigate structural patterns the features had to be size independent.

2.2   Data management and feature design

2.2.1   Preprocessing

Some trees with missing live crown height or height values were removed from the field data. Individual trees that had several stems had a height measurement for each stem. Since multiple stems of one tree together form a single crown, the highest value of height and lowest of live crown height were chosen as values for the individual tree.

To classify species, features of the tree crown were to be used. The exact location of the stem was therefore not as relevant as cutting the whole crown from the point cloud. That was done using the horizontal spread of the crown rather than a circle with a center where the stem meets the ground. To produce a center coordinate for the horizontal spread of the crown, a center of mass for the delineating coordinates was used, i.e., the mean horizontal coordinate for the delineating coordinates. Then the mean distance from the crown center to delineating coordinates was used as radius for a circle when cutting the point cloud.

Return intensity had been recorded by the ALS system as measured. The return intensity, P_r, of a laser pulse is inversely proportional to the distance squared, and was calculated as

$$P_r = \rho \frac{M^2 A}{2\pi R^2} P_t,\qquad(2.1)$$

where ρ is reflectivity of the target, M is atmospheric transmission, A is illuminated area (the footprint, which in this case is assumed to be the same

as target area), 2π is linked to the assumption of the target's bidirectional reflectance distribution function (BRDF, in this example, a lambertian surface), R is the distance from target to scanner, and P_t is transmitted power (Baltsavias 1999). All points of a certain wavelength are assumed to have the same values for atmosperic transmission, footprint, BRDF, and transmitted power. Using the distance, which was calculated using the vector components, $\vec{x}$, $\vec{y}$, and $\vec{z}$, of each return:

$$R^2 = \vec{x}^{\,2} + \vec{y}^{\,2} + \vec{z}^{\,2}, \qquad (2.2)$$

the intensity can be corrected so that it describes the target's reflectivity as a relative value.

The z-coordinates of the point cloud was normalized as height above ground, instead of height above sea level, using the National Land Survey laser data and constructing a triangulated irregular network (TIN) from points classified as ground points. Some points were removed based on height and live crown height values in the field data. All points 10 meters above the tree height were removed to avoid error sources such as birds. In some cases points below the live crown height were also removed when constructing certain features as specified in section 2.2.2.

2.2.2 Features

First and subsequent returns of a pulse will have return intensities dependent on each other. A function of first return intensity can be used to predict the intensity of the second return (Wagner et al. 2008). As the first return contains information on subsequent returns, the main comparisons of feature sets were made using a point cloud of first returns. Nevertheless, to examine the importance of return number within a pulse, all feature sets were calculated separately for point clouds of all, first, and single returns.

Three types of feature sets were calculated: ellipsoid layers, horizontal layers, and distributions, these were produced for each channel. A visualization of ellipsoid and horizontal layer features is shown in fig. 2.1. Each feature set contained two types of features: structural and intensity features. The structural part of the distribution feature set would look similar to the horizontal layer feature set, if viewed from the side, but the separating planes would have different distance depending on point density height-wise.

The channel from which a given feature was derived is denoted as a superscript (C1, C2 or C3) to that feature. Intensity and structure subsets of the features in each feature set were made to determine the subsets' importance for classification.

Ellipsoid layers

This way of describing laser data was designed to catch differences in the density and reflectance of tree crowns from the outside and in. Maple (*A. platanoides*) has a very dense crown with large flat leaves, while ash (*F. excelsior*) grows more sparse with thinner leaves. These differences might be measurable if an ellipsoid is fitted to the crown and ellipsoidal layers are compared to each other.

Features were computed using a point cloud where points below the field measured live crown height were removed. Three main feature types where
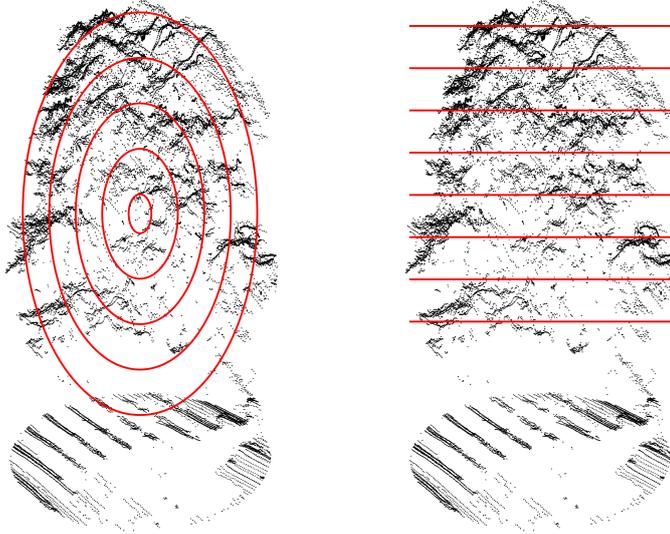
(a) Ellipsoid layers of a tree crown viewed from the side.

(b) Horizontal layers of a tree crown viewed from the side.

Figure 2.1. Two feature set calculation methods visualized. The stripes on the ground are also present in the tree crown above, but less apparent. This pattern is inherent to the scanner type.

computed: J_y^{Cx}, E_y^{Cx}, and AC. The intensity features, J, are the mean return intensity within layers or the mean intensity outside the outmost ellipsoid. Structural features, E, are relative point densities within ellipsoidal layers or the percentage of all points that falls outside of the outmost ellipsoid. AC is also a structural feature, namely the ratio of a and c (see eqs. (2.3) and (2.4)), describing the ellipsoid shape (a value of 1 would be a perfect sphere). Superscript variable x is the channel number (1, 2, or 3) and subscript variable y is the ellipsoidal layer ranging from 0 (outmost) and up to the innermost. y may also be set to e which denotes that this is a value taken from outside of the outmost ellipsoid. E with a subscript e stands for the percentage of all points that fall outside of the outmost ellipsoid, and J with a subscript e stands for the mean intensity outside the outmost ellipsoid.

An example of what a tree crown might look like is presented in figure 2.1a. Using the general shape of the point cloud, an ellipsoid was fitted to the crown. The ellipsoid equation used was

$$\frac{x_s^2}{a^2} + \frac{y_s^2}{b^2} + \frac{z_s^2}{c^2} = 1,$$

where (x_s, y_s, z_s) are coordinates for any point on the surface of the ellipsoid, and a, b, and c are shape-determining constants detailed in eqs. (2.3) and (2.4). To make the ellipsoid horizontally symmetrical (a spheroid), b was set to be

equal to a. The center of the tree was defined as

$$x_c = \bar{x}_p,$$
$$y_c = \bar{y}_p, \text{and}$$
$$z_c = \frac{\max z_p - \min z_p}{2},$$

where x_p, y_p, and z_p are x-, y-, and z-coordinates for the points in the point cloud of the tree. For the outmost ellipsoid, the shape-determining variables were calculated using each point's distance from the center of the tree,

$$a = 2 \times \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2}_{95} \tag{2.3}$$
$$c = \max |z_p - z_c|. \tag{2.4}$$

The value of a being defined as two times the 95th percentile of the distances from the center to a point results in ellipsoids stretching outside of the point cloud. The motive to have a horizontal radius much larger than the data extent is that many fully grown deciduous trees have a flat top of the crown; if only the 95th percentile had been used, too many points would fall outside of the outmost ellipsoid for the analysis to make any sense. There are also parts of the crown that are not inside any ellipsoid. Points in these areas were also analyzed.

All subsequent ellipsoidal layers were created using the same values but with a constant of 0.5 meter subtracted from each radius. The number of ellipsoid layers were determined by dividing the smaller value of a and c by the layer thickness constant and truncating the result. For each layer like those in figure 2.1a, the point density, in points per cubic meter, was divided by the mean density inside of the outmost ellipsoid. In this way the density within the layer was compared to the density within the crown, eliminating size as a factor for tree species determination.

Horizontal layers

The horizontal layer feature set was created in a similar way as the ellipsoid feature set. This feature set was also computed using a point cloud where points below field measured live crown height were removed. The main features types were K_y^{Cx} and L_y^{Cx}, where x is a channel number and y is the layer numbered from top to bottom starting from zero. The intensity features, K, are the mean return intensities within layers, and structural features, L, are relative point density within layers.

With the same constant layer thickness as when computing the ellipsoid feature set, the point cloud was sliced from the top and in each layer mean intensity and relative point density were computed. Relative point density was calculated using a cylinder, defined by the layer thickness and a radius set to the maximum horizontal distance from the center to any point, the density in the layer was then divided by the density in the whole cylinder. Figure 2.1b shows an example of layers in a tree crown.

Distributions

A common way for estimating forest variables from ALS-data is by using features describing the height distribution of the point cloud. One common feature variant is height percentiles (Koenig and Höfle 2016), i.e., the height under which a certain percentage of returns are found (Wonnacott and Wonnacott 1990). Percentiles describe a distribution, it can be applied to height values or intensity values. A point cloud where ground points were included was used to compute the distribution features. The distribution feature subsets consisted of the features presented in table 2.1. Subscripts to percentile features indicates what percentile it is, 1 stands for the 1st percentile and 95 for the 95th. To make the height percentiles independent of the tree height, they were all normalized by dividing them with the 99th height percentile. These features describe height and intensity distribution separately, but not any relation between them.

Feature type	Structural	Intensity
Percentiles	P_1, P_5, P_{10}, P_{25}, P_{50}, P_{75}, P_{90}, P_{95}	Q_1, Q_5, Q_{10}, Q_{25}, Q_{50}, Q_{75}, Q_{90}, Q_{95}, Q_{99}
Average		$\bar{Q}$
Skewness	α	β
Kurtosis	γ	δ
Canopy density	dns	

Table 2.1. Features of the distribution feature set. Percentile is denoted as a subscript.

2.3 Classification model and feature selection

Classification of tree species was made using linear discriminant analysis (LDA) models, which is a well known method for classification using continuous features (Fisher 1936). The method is susceptible to overfitting, meaning that, if used with too many explanatory features and a too small sample size, it will be able to classify all of the training data perfectly but not necessarily classify evaluation data well. To minimize overfitting, by selecting a limited number of features, two methods for feature selection were used: feature ranking (Guyon and Elisseeff 2003) and stepwise feature selection (Weihs et al. 2005). When selecting features, the criterion to stop adding more features was that the feature combination had to have fewer features than there were classes (genera). Stepwise feature selection was also restricted so that additional features were not allowed to make the classification accuracy worse; if no additional feature was found, the algorithm would terminate with less than the maximum number of features.

Single feature performance was evaluated using the F-ratio (also known as Fisher's criterion or the F statistic, commonly used in ANOVA, see Samuels, Witmer, and Schaffner 2012), which describes how much variability there is between the classes compared to within classes. This ranking is optimal for LDA (Guyon and Elisseeff 2003) and features could be ordered by their F-ratio. To compare which individual features that contained the most information on

tree species, the best performing in each feature set and feature subset were extracted and used for an LDA. Results were recorded as a model where included features were ordered by their F-ratio and a correctness rate for the model. Correctness rate is the number of accurately classified trees as a percentage of the total in a leave-one-out cross-validation (Garczarek 2002).

Two features that by themselves provide a good basis for classification must not be the best combination, if they tend to provide the same information (Guyon and Elisseeff 2003). There is for every feature set one or many combinations with the highest correctness rate, but to find one of those, an exhaustive search must be performed. Due to computational limitations this was not feasible, and an heuristic approach was taken. To create a combination of features for classification, a stepwise feature selection algorithm, in which each subset was evaluated using correctness rate, was used (Weihs et al. 2005). This method does not necessarily find the best combination, due to its heuristic nature, but it does find a good combination.

2.4   Species level classification accuracy

Accuracy can be evaluated in different ways. To examine classification accuracy and confusion between species, confusion matrices were used (Lillesand, Kiefer, and Chipman 2007). The confusion matrix contains a summary of classification results for each actual class, as well as user's, producer's and overall accuracy. The user's accuracy is the number of individuals which have been correctly classified as a percentage of all those that have been classified as that particular class. Producer's accuracy is the number of actual instances of a class that have been classified correctly, as a percentage of the total number of actual instances of that class. Overall accuracy is the number of correctly classified trees as a percentage of the total number of trees, irrespective of class. Note that these accuracies are different from the correctness rate described in section 2.3.

To investigate how structural and spectral features contributed to classification accuracy of different species, a number of confusion matrices were made. The most interesting model should be the one that performs best, so one confusion matrix was made using the feature combination that gave the highest correctness rate in cross-validation. An LDA-model was created using those features and all field inventoried trees. The model was then applied to classify the same trees. As a result of using the same data for training and evaluation, some overfitting might have occured. Due to the feature selection made, the effect of overfitting should be small. Additional confusion matrices were created for models built on only spectral and structural subsets of the feature set from which the best performing model was acquired.

3 Results

Complete cross-validation and feature selection results are presented in the tables of appendix A.

3.1   Features for tree species classification

Features used for tree species classification chosen by their F-ratio is presented in table 3.1, ordered from the feature with the best (highest) F-ratio in that feature set to the worst of the eight presented, those with a lower F-ratio are not included.

Feature set	Feature							
	1	2	3	4	5	6	7	8
Ellipsoid	$J_\mathrm{e}^{\mathrm{C1}}$	$J_\mathrm{e}^{\mathrm{C2}}$	$J_\mathrm{e}^{\mathrm{C3}}$	J_1^{C1}	J_2^{C1}	J_0^{C1}	AC	J_1^{C2}
F-ratio	61.85	41.07	35.14	28.10	27.09	23.71	23.67	16.86
Horizontal	K_2^{C1}	K_3^{C1}	K_4^{C1}	K_2^{C2}	K_1^{C1}	K_5^{C1}	K_0^{C1}	K_1^{C3}
F-ratio	25.74	24.44	20.88	20.86	19.86	17.74	17.24	16.53
Distribution	Q_{75}^{C1}	Q_{50}^{C1}	$\bar{Q}^{\mathrm{C1}}$	Q_{90}^{C1}	Q_{99}^{C2}	Q_{95}^{C2}	Q_{90}^{C2}	Q_{95}^{C1}
F-ratio	75.25	72.20	71.23	65.86	61.09	60.58	59.44	57.90

Table 3.1. Features ranked by their F-ratio. Only first returns and complete feature sets were used. Feature number one was the best in the feature set to use if only one were to be used for classification. J, K and Q are intensity variables, and AC is a value describing the shape of the ellipsoid. Superscripts signifies channel. Subscripts of J indicates ellipsoidal layer enumerated from the outmost (0) to the innermost and e pertains to what is outside the outmost ellipsoid. Subscripts of K signifies horizontal layer enumerated from the topmost (0) to the bottommost. Subscripts of Q is the intensity percentile number, and $\bar{Q}$ is the mean intensity. A summary of all subscripts and superscripts are available in table A.1, in appendix A.

The ellipsoid data had four ellipsoid layers, numbered, outside and in, from zero to three, and data from outside of the outmost ellipsoid, with an e as subscript. In horizontal-layer data, seven layers had been constructed, numbered from zero to six with no features from outside the top or bottom layer. As described in section 2.2.2, the distribution statistics had 11 features for the intensity distribution and height distribution each. Percentiles used in both

intensity and height percentile features were 1, 5, 10, 25, 50, 75, 90, and 95. There was also a 99th intensity percentile.

When the eight features with the highest F-ratio were selected, the only structural feature among them were the AC-ratio, i.e., the ellipsoid shape of the crown. Another notable property of these features is that only two are from the 532 nm (C3) data, namely intensity outside of the ellipsoid and intensity in the second layer from the top, the rest are from 1064 nm (C2) and 1550 nm (C1).

In general features containing information from the upper part of the tree contained more information on tree species. The three features from the ellipsoid feature set with the highest F-ratio was those with intensity information from outside the outmost ellipsoid. Horizontal layers ranging in number from zero to five were among the top eight when ordered by F-ratio, but five of eight features were from the three top horizontal layers. From the distribution feature set, only features from the average or median height and up (depending on distribution skewness) are among the eight best features; none of the 1st, 5th, 10th of 25th percentile features were among the eight with the highest F-ratio.

Combinations of features that could be used to produce a model with a high correctness rate are presented in table 3.2. Each row presents the best combination of features from that feature set, that was found using stepwise feature selection.

Feature set	Features							
Ellipsoid	J_0^{C1}	J_e^{C1}	J_2^{C2}	J_e^{C2}	$\bar{E}^{C2}$	J_0^{C3}	J_e^{C3}	AC
Horizontal	K_2^{C1}	K_4^{C1}	L_0^{C1}	K_2^{C2}	K_4^{C3}	K_5^{C3}	L_1^{C3}	L_2^{C3}
Distribution	P_{90}^{C1}	$\bar{Q}^{C1}$	Q_{05}^{C1}	Q_{50}^{C1}	Q_{90}^{C2}	Q_{95}^{C2}	Q_{01}^{C3}	Q_{90}^{C3}

Table 3.2. Features from different feature sets selected by stepwise feature selection. Only first returns and complete feature sets were used. J, K and Q are intensity features. E, AC, L and P are structural features. Superscripts signifies channel. Subscripts of J indicates ellipsoidal layer enumerated from the outmost (0) to the innermost and e pertains to what is outside the outmost ellipsoid. Subscripts of K and L signifies horizontal layer enumerated from the topmost (0) to the bottommost. Subscripts of P and Q are height and intensity percentile number respectively. $\bar{Q}$ is the mean intensity. A summary of all subscripts and superscripts are available in table A.1, in appendix A.

When features were combined to create an accurate classification model using stepwise feature selection, features from all channels were selected and all combinations had both spectral and structural features present. The combination from the ellipsoid feature set was primarily composed of intensity metrics from the outer parts of the tree crown; the only structural feature was AC.

The feature combination from the horizontal layer feature set used four intensity features and three structural features. Only one, an intensity feature, was from channel 2, i.e., 1064 nm. In the combination selected from the distribution feature set, intensity percentiles from the 1st to the 95th were included. Only one height distribution was selected by the algorithm: P_{90}^{C1}.

3.2 Spectral and structural information

Regardless of what feature set was used, multispectral data always gave a higher cross-validation correctness rate than monospectral data, as shown in table 3.3. The model using multispectral data and all features from the distribution data set had the highest correctness rate. The greatest difference between multispectral and monospectral data was found when all variables were available for selection. The smallest difference between monospectral and multispectral data in correctness rate was in the models produced with structural features only.

Feature subset	Ellipsoid	Horizontal	Distribution
Multispectral			
Intensity	65.36	58.66	75.98
Structure	43.02	34.64	41.34
All	68.72	60.34	76.54
Monospectral (1064 nm)			
Intensity	44.13	40.78	67.60
Structure	40.22	30.73	35.75
All	56.42	43.58	66.48

Table 3.3. Correctness rate as a percentage for models built using first returns only. Rows are subsets of either intensity, structure or all features from the feature sets that are the columns. Accuracies using multispectral and monospectral (1064 nm) data are presented. For accuracies in other channels, see appendix A.

3.3 Return number

Table 3.4 shows the correctness rate for different feature sets when using different kinds of returns. In two of three cases, the multispectral data using all returns gave the highest correctness rate in cross-validation. The exception was the ellipsoid feature set where first returns gave a better classification accuracy. When looking at the monospectral data, both ellipsoid and horizontal layer feature sets had their respective highest correctness rate when using all returns. The distribution feature set from multispectral data gave the same accuracy when using first and single returns. The horizontal layer feature set from monospectral data also resulted in the the same correctness rate whether first or single returns were used. Differences in correctness rate depending on return type was less than one percentage point in the multispectral distribution feature sets. The largest difference between return types, in percentage points, was found in the ellipsoid feature set from monospectral data.

Returns	Ellipsoid	Horizontal	Distribution
Multispectral			
Single	63.13	59.78	76.54
First	68.72	60.34	76.54
All	67.60	62.57	77.09
Monospectral (1064 nm)			
Single	44.13	43.02	64.25
First	56.42	43.58	66.48
All	57.54	48.60	70.39

Table 3.4. Correctness rate as a percentage using point clouds of either all returns, only first returns or only single returns. Each feature set was computed separately for these point clouds. Models were created using stepwise selection with both spectral and structural features available.

3.4 Species level classification accuracy

The model with the highest cross-validation correctness rate was created using the distribution feature set and consisted of spectral features only. As a result, only two confusion matrices are presented in tables 3.5 and 3.6.

Classification[†]	Known types[†]									User's (%)
	AL	M	B	AS	S	C	P	O	L	
AL	0	0	0	0	0	0	0	0	0	—
M	0	12	0	0	0	0	0	0	2	86
B	0	0	24	0	0	3	0	0	4	77
AS	1	0	0	19	0	1	0	2	2	76
S	0	0	0	0	15	0	0	0	0	100
C	0	0	0	0	0	1	0	0	0	100
P	0	0	0	0	0	0	29	0	0	100
O	5	0	1	4	0	0	0	35	7	67
L	1	1	1	0	0	0	0	1	8	67
Producer's (%)	0	92	92	83	100	20	100	92	35	**80**

Table 3.5. Confusion matrix for the best performing model. †: AL, alder; M, maple; B, birch; AS, Ash; S, spruce; C, cherry; P, pine; O, oak; L, linden. Features used in the model were computed from all returns. Features used were: Q_{10}^{C1}, Q_{75}^{C1}, $\max Q^{C2}$, Q_{90}^{C2}, Q_{95}^{C2}, Q_{99}^{C2}, Q_{05}^{C3} and Q_{75}^{C3}.

In the confusion matrix of table 3.5, classification accuracy of the best performing model is presented. The stepwise feature selection algorithm chose only intensity features (Q_{10}^{C1}, Q_{75}^{C1}, $\max Q^{C2}$, Q_{90}^{C2}, Q_{95}^{C2}, Q_{99}^{C2}, Q_{05}^{C3} and Q_{75}^{C3}) from the distribution feature set. The resulting model had a cross-validation correctness rate of 77.09 %. As producer's accuracy shows, none of the alders were correctly classified, and cherry and linden classification performed poorly. Spruce and pine on the other hand, got 100 % classification accuracy both in producers and users accuracy. Classification accuracy of maple, birch and ash was fairly high, both in user's and producer's accuracy. Producer's accuracy of oak was 92 % but only 67 % in user's accuracy. The overall accuracy of

classification was 80 %, a little bit higher than the cross-validation correctness rate of 77.09 %.

Classification[†]	Known types[†]									User's (%)
	AL	M	B	AS	S	C	P	O	L	
AL	1	0	1	0	0	0	0	0	1	33
M	1	7	0	2	1	0	0	2	1	50
B	1	1	14	2	1	2	0	2	3	54
AS	0	2	2	13	2	0	1	2	0	59
S	0	0	0	0	9	0	1	1	0	82
C	0	0	0	0	1	1	0	0	1	33
P	0	0	0	0	0	0	25	0	0	100
O	1	1	7	5	1	2	2	28	6	53
L	3	2	2	1	0	0	0	3	11	50
Producer's (%)	14	54	54	57	60	20	86	74	48	**61**

Table 3.6. Confusion matrix for the structural feature subset of the feature set providing the best performing model. †: AL, alder; M, maple; B, birch; AS, Ash; S, spruce; C, cherry; P, pine; O, oak; L, linden. Features used in the model were computed from all returns. Features used were: P_{75}^{C1}, dns^{C1}, α^{C2}, P_1^{C2}, P_{25}^{C2}, α^{C3}, P_{25}^{C3} and P_{95}^{C3}.

The confusion matrix of the model built with only structural features (P_{75}^{C1}, dns^{C1}, α^{C2}, P_1^{C2}, P_{25}^{C2}, α^{C3}, P_{25}^{C3} and P_{95}^{C3}) in table 3.6 shows that such a model does not perform as well overall as the model built with spectral features. The overall accuracy of the model was 61 % while the cross-validation correctness rate was 55.31 %. No species was perfectly discriminated from the others. Only alder had a producer's and user's accuracy higher than in table 3.5. Linden had a higher producer's accuracy, but lower user's accuracy, meaning that more trees have been classified as linden. Cherry had the same producer's accuracy but lower user's accuracy. Maple, birch, ash, spruce, pine and oak had a lower producer's accuracy, and all these except pine had a lower user's accuracy as well.

4 Discussion

4.1 Conclusions

When features were evaluated by their F-ratio, the best features for tree species classification were those that describe the upper part of the tree crown. F-ratio describes how large the between-group variability is compared to within-group variability. When using LDA, this is the optimal way to rank variables, since the classification method works similarly (Guyon and Elisseeff 2003).

The highest F-ratios for the different feature sets were found in intensity features. The only structural feature among all the best features from all data sets was the *AC*-ratio. This indicates that spectral data is highly relevant for tree species identification, which is known to be the case when using infrared aerial photography (Axelson and Nilsson 1993). Of all the features in table 3.1, only two came from the 532 nm data (C3).

In the feature combinations (table 3.2) produced by stepwise feature selection, there were also more spectral features than structural. Nevertheless, each combination contained at least one structural feature. Eight features were from the 532 nm data. This indicates that, while 532 nm data might not be useful as a single source of information, it is useful when combined with other channels as it provides new information to the model.

The exact cause of the relatively poor performance of 532 nm data can not be identified without further studies. There are nonetheless some ways in which channel 3 differs from the other two channels. Due to eye safety regulations, the divergence had to be greater for visible light than for NIR and SWIR channels. As a result, the footprint of the green channel was larger and had less power per area. Some atmospheric effects might also be more prominent for different wavelengths of light. These are two things that might be further examined.

Multispectral information gave a higher correctness rate than the 1064 nm channel separately. A combination of spectral and structural features, using first returns only, provided the highest correctness rate. When using first returns only, subsets of features containing only intensity or structural information provided a lower correctness rate. This was true even when only the near infrared channel was used. Classifying tree species with only structural features when spectral information for the returns is available would not be advisable. Even when using only one channel, the intensity features gives a higher correctness rate than the structural features. The fact that 1064 nm gives a correctness rate of 67.6 % when only intensity distribution features were used is interesting. Many ALS-systems have an adaptive output effect, meaning that if there is a drop in return intensity, the output effect will increase. The value of the output intensity is not always recorded and this limits intensity-based clas-

sification to data from those ALS-systems that do have either constant output effect or records the output effect. In the Optech Titan X system the output effect is held constant, allowing return intensity variations to become apparent.

Whether single, first or all returns were used to compute features did not seem to matter much. The best performing model was built on distribution features from all returns, with a correctness rate of 77.09 %. On the other hand, both first and single returns gave an correctness rate of 76.54 %, which in a practical application would be a negligible difference compared to 77.09 %. This small difference suggests that most of the information is available in the first or only return, something that agrees with previous studies (Wagner et al. 2008). Nevertheless, there seems to be at least some information in the returns following the first, which results in the slightly higher correctness rate. No feature set was created where relations between return numbers were used for classification. Proportions of first and last returns have been used as basis for tree species classification, with a high accuracy for classifying pine and spruce (Holmgren and Persson 2004). Return intensity was corrected by distance but could also be corrected by the intensity of previous returns of the same pulse, there is, however, no established method for doing this. Using proportion of returns or return intensity corrected by previous return intensity is something that might be interesting to examine in multispectral ALS-data in future studies.

Spruce, pine and deciduous classes are easily discriminated from each other, by using features in multispectral intensity data, as shown in the confusion matrix of table 3.5. The species that were confused with each other are the deciduous. Alder and cherry trees were relatively few in this sample, seven and five respectively, which made errors in those classes more difficult to interpret. Linden, with 23 individual trees, seemed to be easily confused with any other species. When comparing tables 3.5 and 3.6, Pine seems to be easily classified when using either spectral or structural features. Oak also have a fairly high producer's accuracy when using structural features. Maple, birch and ash seems to be much easier to distinguish in spectral than in structural features. Spruce seems to be more difficult to classify correctly using only structural features, this might have to do with the fact that spruces were almost exclusively found in closed-canopy stands. A closed canopy might affect structural features so that spruce trees in this sample does not have a typical shape. Spruce crowns may grow into each other and become difficult to separate, meaning that structural features no longer describe an individual tree, but rather a mixture of neighboring trees.

It is clear that multispectral ALS data provides plenty of information on tree species. Features best suited for species classification are from the upper parts of the crown, and intensity features contain more information than structural features. Cross-validation correctness rate was the main metric by which model performance was measured. The model with an correctness rate of 77 % had an overall accuracy in the confusion matrix of 80 %, which may be due to some overfitting when all trees was used both for training and evaluating the model. With such a high overall accuracy, combined with the fact that pine and spruce were perfectly discriminated from deciduous species, a model built with these features might be very useful for both nature value inventory and forestry planning. Further studies are needed to examine the cause of specific

error rates, e.g., why linden have been confused with so many other tree species and how a larger sample of alder and cherry would affect the result.

The model with the highest correctness rate was one that used only intensity features. On the other hand, the best performing model in table 3.3 was the multispectral where features had been selected from both intensity and structure features. The difference between these two models is that the features in the first model were computed using first returns only, and the features in the second were computed using all returns. There are some possible explanations as to why the best performing model used only spectral features, e.g., the returns subsequent to the the first provided more intensity information, the subsequent returns introduced noise in structural information, or that statistical noise affected output of the feature selection algorithm. Since the model where features were computed using all returns provided a higher correctness rate, it is probable that the subsequent returns provided more information on intensity, allowing a better classification to be made. Nevertheless, the higher correctness rate might also be a result of the heuristic approach in the feature selection algorithm, i.e., suboptimum combinations might have been selected.

Feature sets performed differently from each other: distribution features gave the highest correctness rate followed by ellipsoid layer features and lastly horizontal layer features. The ellipsoid and horizontal layer features were constructed to find combinations of spectral and structural features within an ellipsoid or horizontal layer that could be used for tree species identification; spectral features were related to structural features by deriving them from the same ellipsoid or horizontal layer. Distribution features on the other hand, did not relate height and intensity distribution to each other in such a way. The idea that trees such as ash, with a sparse and light green crown, could be separated from maple, with a dense and dark green crown, using ellipsoids or horizontal layers proved to be functional but was surpassed by features describing height and intensity distributions of the tree as a whole.

4.2   Weaknesses and the future

There are some areas where materials and methods could be improved for future studies. Both the field and ALS data are very specific. Only solitary trees were used, the reason being that this kind of ALS data had not been thoroughly examined, and clearly distinguishable crowns made exploration of data easier. A more real-world usage would be in closed-canopy forest stands using either single-tree or area-based methods. Solitary trees differ in shape from trees in a closed canopy, but since the best features for tree species classification came from the upper part of the canopy, the data should be usable in closed stands as well. Some features may be more difficult than other to construct without information about the live crown height. Ellipsoid features might have to be adapted for a real-world usage. One way to adapt them could be by using segmentation algorithms and fit half an ellipsoid over the automatically segmented crown. The ALS-data were gathered from a fairly low flying altitude of 400 meters. This is not practical as each swath will only cover a small area. As a comparison, the New National Height model of Sweden was created using data gathered from an altitude of 1,700 to 2,300 meters (Lantmäteriet 2016), 400 meters is not even a quarter of that altitude.

When choosing and calculating candidate features, a number of completely subjective choices were made. The ellipsoid and horizontal layer thickness was set to 0.5 meter. There was no reason behind this other than that it resulted in a decent number of ellipsoidal layers. The horizontal layer thickness was set to match the the ellipsoidal layers. Distribution features were also chosen subjectively. There were more high percentiles than low, because it was presumed that a high value might be more interesting than a low, as low values in height would be the ground and low intensity values could have been the result of light scattering so that only a small fraction of the output effect returned to the scanner. This assumption seems to be true, as none of the lower percentiles was included in table 3.1 or table 3.2. The only spectral features used were taken from the intensity distribution, either within an ellipsoid, horizontal layer or the whole tree, of each channel separately. No combination of intensities in different channels, analogue to NDVI (Normalized Difference Vegetation Index, for more information see Axelson and Nilsson 1993), was used as a feature. The feature sets used have been treated as separate, there is however no restriction for combining different feature sets. These subjective choices were made because there is an endless amount of different theoretical features that could be extracted from a point cloud. Technically the amount of information is reduced when any data points are averaged, this is done to reduce complexity of data. An objective alternative to manual feature design is machine learning algorithms, such as neural networks, that creates features as the particular algorithm finds fit. The drawback with such a method would be that features could become very synthetic and hard to interpret.

There is a multitude of different feature selection algorithms. The two used were chosen on grounds of simplicity (F-ratio) and optimal performance in relation to the classification method (stepwise feature selection). Some examples are: other ways to rank variables than by F-ratio and using principal components of features. Guyon and Elisseeff (2003) with their paper, titled *An Introduction to Variable and Feature Selection*, provides an overview of many different options.

A Selected variables

Variable	Description
CR	Correctness rate
E	Density in ellispoid layer
J	Intensity in ellipsoid layer
L	Density in horizontal layer
K	Intensity in horizontal layer
P	Height percentile
Q	Intensity percentile
I	Intensity
α	Height skewness
β	Intensity skewness
γ	Height kurtosis
δ	Intensity kurtosis
AC	Ratio of ellipsoid a- and c-values

Sub- and superscript	Description
X_{e}	External, values outside ellipsoid
$\bar{X}$	Arithmetic mean of X
$X_{\min}$	Minimum of X
$X_{\max}$	Maximum of X
X_n	n is a number signifying which in order
$X^{\mathrm{C}n}$	n is 1, 2 or 3 and signifies channel number

Table A.1. Explanation of variables

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
All	65.92	J_e^{C1}	J_e^{C3}	J_e^{C2}	AC	J_0^{C1}	J_2^{C3}	J_1^{C1}	J_2^{C1}
Intensity	62.01	J_e^{C1}	J_e^{C3}	J_e^{C2}	J_0^{C1}	J_2^{C3}	J_1^{C1}	J_2^{C1}	J_0^{C2}
Structure	36.87	AC	E_e^{C2}	E_e^{C1}	E_e^{C3}	E_0^{C3}	E_1^{C1}	E_1^{C1}	$\bar{P}^{C3}$
First returns									
All	65.36	J_e^{C1}	J_e^{C2}	J_e^{C3}	J_1^{C1}	J_2^{C1}	J_0^{C1}	AC	J_1^{C2}
Intensity	60.34	J_e^{C1}	J_e^{C2}	J_e^{C3}	J_1^{C1}	J_2^{C1}	J_0^{C1}	J_1^{C2}	E_2^{C2}
Structure	35.75	AC	E_e^{C1}	E_e^{C2}	E_e^{C3}	E_1^{C1}	E_0^{C3}	$\bar{P}^{C3}$	E_1^{C1}
Single returns									
All	60.89	J_e^{C1}	J_e^{C3}	J_e^{C2}	AC	J_1^{C1}	J_0^{C1}	J_1^{C2}	J_0^{C2}
Intensity	55.31	J_e^{C1}	J_e^{C3}	J_e^{C2}	J_1^{C1}	J_0^{C1}	J_2^{C2}	J_0^{C2}	J_1^{C3}
Structure	30.17	AC	E_e^{C1}	E_e^{C3}	E_1^{C1}	E_1^{C2}	E_e^{C2}	$\bar{P}^{C3}$	E_1^{C1}

Table A.2. Ellipsoid layer features ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
All	52.51	K_2^{C1}	K_3^{C1}	K_4^{C1}	L_2^{C2}	K_5^{C3}	K_1^{C1}	K_0^{C1}	K_1^{C3}
Intensity	52.51	K_2^{C1}	K_3^{C1}	K_4^{C1}	L_2^{C2}	K_5^{C3}	K_1^{C1}	K_0^{C1}	K_1^{C3}
Structure	34.08	L_1^{C3}	L_1^{C2}	L_1^{C1}	L_1^{C1}	L_1^{C1}	L_2^{C1}	L_2^{C2}	L_0^{C3}
First returns									
All	48.6	K_2^{C1}	K_3^{C1}	K_4^{C1}	L_2^{C2}	K_1^{C1}	K_5^{C1}	K_0^{C1}	K_1^{C3}
Intensity	48.6	K_2^{C1}	K_3^{C1}	K_4^{C1}	L_2^{C2}	K_1^{C1}	K_5^{C1}	K_0^{C1}	K_1^{C3}
Structure	30.17	L_1^{C3}	L_1^{C1}	L_1^{C2}	L_1^{C1}	L_1^{C1}	L_2^{C1}	L_0^{C3}	L_2^{C3}
Single returns									
All	54.19	K_2^{C1}	K_1^{C1}	K_0^{C1}	K_3^{C1}	L_2^{C2}	K_1^{C2}	K_4^{C1}	K_1^{C3}
Intensity	54.19	K_2^{C1}	K_1^{C1}	K_0^{C1}	K_3^{C1}	L_2^{C2}	K_1^{C2}	K_4^{C1}	K_1^{C3}
Structure	27.93	L_2^{C1}	L_1^{C1}	L_0^{C3}	L_4^{C1}	L_4^{C3}	L_2^{C2}	L_1^{C1}	L_1^{C1}

Table A.3. Horizontal layer features ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
All	68.72	Q_{90}^{C1}	Q_{75}^{C1}	Q_{95}^{C1}	Q_{99}^{C2}	Q_{95}^{C2}	$\bar{Q}^{C1}$	Q_{90}^{C2}	Q_{90}^{C3}
Intensity	68.72	Q_{90}^{C1}	Q_{75}^{C1}	Q_{95}^{C1}	Q_{99}^{C2}	Q_{95}^{C2}	$\bar{Q}^{C1}$	Q_{90}^{C2}	Q_{90}^{C3}
Structure	40.22	γ^{C3}	P_{10}^{C2}	P_{10}^{C1}	P_{05}^{C2}	P_{25}^{C2}	α^{C3}	P_{25}^{C1}	P_{10}^{C3}
First returns									
All	68.72	Q_{75}^{C1}	Q_{50}^{C1}	$\bar{Q}^{C1}$	Q_{90}^{C1}	Q_{99}^{C2}	Q_{95}^{C2}	Q_{90}^{C2}	Q_{95}^{C1}
Intensity	68.72	Q_{75}^{C1}	Q_{50}^{C1}	$\bar{Q}^{C1}$	Q_{90}^{C1}	Q_{99}^{C2}	Q_{95}^{C2}	Q_{90}^{C2}	Q_{95}^{C1}
Structure	34.08	dns^{C1}	dns^{C3}	P_{10}^{C1}	dns^{C2}	P_{05}^{C1}	P_{10}^{C2}	P_{25}^{C1}	P_{10}^{C3}
Single returns									
All	58.1	Q_{50}^{C1}	Q_{75}^{C1}	Q_{25}^{C1}	$\bar{Q}^{C1}$	Q_{10}^{C1}	Q_{05}^{C1}	Q_{75}^{C2}	$\bar{Q}^{C2}$
Intensity	58.1	Q_{50}^{C1}	Q_{75}^{C1}	Q_{25}^{C1}	$\bar{Q}^{C1}$	Q_{10}^{C1}	Q_{05}^{C1}	Q_{75}^{C2}	$\bar{Q}^{C2}$
Structure	40.22	dns^{C1}	dns^{C2}	dns^{C3}	P_{10}^{C2}	P_{10}^{C1}	P_{05}^{C2}	P_{01}^{C2}	P_{25}^{C1}

Table A.4. Distribution features ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	53.07	J_e	AC	J_0	J_1	J_2	E_e	E_0	E_1
1064 nm	51.96	J_e	AC	J_0	J_1	J_2	E_e	$\bar{P}$	E_0
532 nm	56.98	J_e	AC	J_2	J_1	J_0	E_e	E_0	$\bar{P}$
First returns									
1550 nm	52.51	J_e	J_1	J_2	J_0	AC	E_e	E_0	E_1
1064 nm	51.96	J_e	AC	J_1	J_2	J_0	E_e	E_2	E_0
532 nm	52.51	J_e	AC	J_2	J_0	J_1	E_e	E_0	$\bar{P}$
Single returns									
1550 nm	46.93	J_e	AC	J_1	J_0	E_e	E_1	E_0	$\bar{P}$
1064 nm	35.75	J_e	AC	J_1	J_0	E_1	E_e	$\bar{P}$	E_0
532 nm	47.49	J_e	AC	J_1	J_0	E_e	$\bar{P}$	E_1	E_0

Table A.5. Ellipsoid layer features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	34.08	AC	E_e	E_0	E_1	$\bar{P}$	E_2		
1064 nm	40.22	AC	E_e	$\bar{P}$	E_0	E_1	E_2		
532 nm	37.43	AC	E_e	E_0	$\bar{P}$	E_2	E_1		
First returns									
1550 nm	36.87	AC	E_e	E_0	E_1	$\bar{P}$	E_2		
1064 nm	36.31	AC	E_e	E_2	E_0	$\bar{P}$	E_1		
532 nm	36.31	AC	E_e	E_0	$\bar{P}$	E_2	E_1		
Single returns									
1550 nm	30.73	AC	E_e	E_1	E_0	$\bar{P}$			
1064 nm	31.28	AC	E_1	E_e	$\bar{P}$	E_0			
532 nm	30.73	AC	E_e	$\bar{P}$	E_1	E_0			

Table A.6. Ellipsoid layer structural features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	44.69	J_e	J_0	J_1	J_2				
1064 nm	39.11	J_e	J_0	J_1	J_2				
532 nm	41.9	J_e	J_2	J_1	J_0				
First returns									
1550 nm	45.25	J_e	J_1	J_2	J_0				
1064 nm	41.34	J_e	J_1	J_2	J_0				
532 nm	39.66	J_e	J_2	J_0	J_1				
Single returns									
1550 nm	46.93	J_e	J_1	J_0					
1064 nm	39.66	J_e	J_1	J_0					
532 nm	38.55	J_e	J_1	J_0					

Table A.7. Ellipsoid layer spectral features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	36.87	K_2	K_3	K_4	K_1	K_0	K_5	K_6	L_1
1064 nm	38.55	K_2	K_1	K_3	K_4	K_0	K_5	L_1	L_0
532 nm	47.49	K_5	K_1	K_2	K_4	K_6	K_3	K_0	L_1
First returns									
1550 nm	41.34	K_2	K_3	K_4	K_1	K_5	K_0	K_6	L_1
1064 nm	35.75	K_2	K_1	K_3	K_4	K_0	K_5	K_6	L_1
532 nm	43.58	K_1	K_2	K_5	K_3	K_4	K_6	K_0	L_1
Single returns									
1550 nm	45.81	K_2	K_1	K_0	K_3	K_4	L_2	L_4	L_0
1064 nm	39.66	K_2	K_1	K_3	K_0	K_4	L_0	L_2	L_1
532 nm	36.87	K_1	K_2	K_0	K_3	K_4	L_0	L_4	L_1

Table A.8. Horizontal layer features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	36.31	L_1	L_0	L_2	L_3	L_4	L_6	L_5	
1064 nm	28.49	L_1	L_0	L_2	L_4	L_3	L_5	L_6	
532 nm	26.82	L_1	L_0	L_2	L_4	L_3	L_6	L_5	
First returns									
1550 nm	32.4	L_1	L_0	L_2	L_3	L_4	L_6	L_5	
1064 nm	27.37	L_1	L_0	L_2	L_4	L_6	L_5	L_3	
532 nm	26.26	L_1	L_0	L_2	L_6	L_4	L_5	L_3	
Single returns									
1550 nm	19.55	L_2	L_4	L_0	L_1	L_3			
1064 nm	20.11	L_0	L_2	L_1	L_4	L_3			
532 nm	21.23	L_0	L_4	L_1	L_2	L_3			

Table A.9. Horizontal layer structural features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	42.46	K_2	K_3	K_4	K_1	K_0	K_5	K_6	
1064 nm	34.64	K_2	K_1	K_3	K_4	K_0	K_5	K_6	
532 nm	44.13	K_5	K_1	K_2	K_4	K_6	K_3	K_0	
First returns									
1550 nm	39.66	K_2	K_3	K_4	K_1	K_5	K_0	K_6	
1064 nm	32.96	K_2	K_1	K_3	K_4	K_0	K_5	K_6	
532 nm	40.78	K_1	K_2	K_5	K_3	K_4	K_6	K_0	
Single returns									
1550 nm	48.04	K_2	K_1	K_0	K_3	K_4			
1064 nm	39.11	K_2	K_1	K_3	K_0	K_4			
532 nm	34.64	K_1	K_2	K_0	K_3	K_4			

Table A.10. Horizontal layer spectral features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	62.57	Q_{90}	Q_{75}	Q_{95}	$\bar{Q}$	Q_{50}	Q_{99}	P_{10}	P_{25}
1064 nm	63.69	Q_{99}	Q_{95}	Q_{90}	Q_{75}	$\bar{Q}$	Q_{max}	δ	Q_{50}
532 nm	53.07	Q_{90}	Q_{75}	Q_{95}	$\bar{Q}$	Q_{50}	γ	Q_{99}	α
First returns									
1550 nm	60.34	Q_{75}	Q_{50}	$\bar{Q}$	Q_{90}	Q_{95}	Q_{25}	Q_{10}	Q_{99}
1064 nm	60.89	Q_{99}	Q_{95}	Q_{90}	$\bar{Q}$	Q_{75}	Q_{50}	Q_{max}	Q_{25}
532 nm	56.98	Q_{90}	Q_{95}	Q_{75}	$\bar{Q}$	Q_{50}	Q_{25}	Q_{10}	Q_{05}
Single returns									
1550 nm	46.37	Q_{50}	Q_{75}	Q_{25}	$\bar{Q}$	Q_{10}	Q_{05}	Q_{01}	Q_{90}
1064 nm	49.72	Q_{75}	$\bar{Q}$	Q_{90}	Q_{95}	Q_{50}	Q_{99}	Q_{25}	Q_{10}
532 nm	41.34	Q_{75}	Q_{50}	Q_{25}	Q_{10}	Q_{05}	$\bar{Q}$	Q_{01}	Q_{min}

Table A.11. Distribution features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	33.52	P_{10}	P_{25}	P_{05}	P_{50}	dns	P_{75}	P_{90}	P_{95}
1064 nm	31.28	P_{10}	P_{05}	P_{25}	P_{50}	P_{75}	dns	P_{01}	P_{90}
532 nm	37.99	γ	α	P_{10}	P_{25}	P_{05}	P_{50}	dns	P_{75}
First returns									
1550 nm	35.75	dns	P_{10}	P_{05}	P_{25}	P_{50}	P_{75}	P_{01}	α
1064 nm	30.73	dns	P_{10}	P_{25}	P_{05}	P_{01}	P_{50}	P_{75}	α
532 nm	40.22	dns	P_{10}	γ	P_{25}	α	P_{01}	P_{05}	P_{50}
Single returns									
1550 nm	30.73	dns	P_{10}	P_{25}	P_{05}	α	γ	P_{50}	P_{01}
1064 nm	34.08	dns	P_{10}	P_{05}	P_{01}	P_{25}	P_{50}	P_{90}	P_{75}
532 nm	31.28	dns	P_{25}	P_{05}	P_{10}	P_{01}	P_{90}	P_{75}	P_{50}

Table A.12. Structural distribution features for each channel ordered by F-ratio.

Feature set	CR	Feature							
		1	2	3	4	5	6	7	8
All points									
1550 nm	55.31	Q_{90}	Q_{75}	Q_{95}	$\bar{Q}$	Q_{50}	Q_{99}	Q_{25}	β
1064 nm	63.69	Q_{99}	Q_{95}	Q_{90}	Q_{75}	$\bar{Q}$	Q_{max}	δ	Q_{50}
532 nm	49.16	Q_{90}	Q_{75}	Q_{95}	$\bar{Q}$	Q_{50}	Q_{99}	Q_{25}	Q_{05}
First returns									
1550 nm	60.34	Q_{75}	Q_{50}	$\bar{Q}$	Q_{90}	Q_{95}	Q_{25}	Q_{10}	Q_{99}
1064 nm	60.89	Q_{99}	Q_{95}	Q_{90}	$\bar{Q}$	Q_{75}	Q_{50}	Q_{max}	Q_{25}
532 nm	56.98	Q_{90}	Q_{95}	Q_{75}	$\bar{Q}$	Q_{50}	Q_{25}	Q_{10}	Q_{05}
Single returns									
1550 nm	46.37	Q_{50}	Q_{75}	Q_{25}	$\bar{Q}$	Q_{10}	Q_{05}	Q_{01}	Q_{90}
1064 nm	49.72	Q_{75}	$\bar{Q}$	Q_{90}	Q_{95}	Q_{50}	Q_{99}	Q_{25}	Q_{10}
532 nm	41.34	Q_{75}	Q_{50}	Q_{25}	Q_{10}	Q_{05}	$\bar{Q}$	Q_{01}	Q_{min}

Table A.13. Distribution spectral features for each channel ordered by F-ratio.

Feature set	CR	Selected features
All points		
All	67.6	J_e^{C1} E_1^{C1} J_e^{C2} J_e^{C3} AC
Intensity	63.69	J_0^{C1} J_e^{C1} J_e^{C2} J_e^{C3}
Structure	43.58	E_e^{C1} E_2^{C3} AC
First returns		
All	68.72	J_0^{C1} J_e^{C1} E_2^{C2} J_e^{C2} $\bar{P}^{C2}$ J_0^{C3} J_e^{C3} AC
Intensity	65.36	J_0^{C1} J_e^{C1} J_1^{C2} E_2^{C2} J_e^{C2} J_0^{C3} J_e^{C3}
Structure	43.02	E_e^{C1} E_e^{C2} AC
Single returns		
All	63.13	J_1^{C1} J_e^{C1} E_1^{C1} J_e^{C2} E_1^{C1} E_1^{C2} J_e^{C3}
Intensity	60.89	J_e^{C1} J_e^{C2} J_e^{C3}
Structure	35.75	E_1^{C1} E_e^{C1} $\bar{P}^{C1}$ E_1^{C1} E_1^{C3} E_e^{C3} $\bar{P}^{C3}$ AC

Table A.14. Ellipsoid layer features selected using stepwise feature selection

Feature set	CR	Selected features
All points		
All	62.57	K_2^{C1} K_4^{C1} L_2^{C1} L_2^{C2} K_5^{C2} L_3^{C2} K_4^{C3} K_5^{C3}
Intensity	59.78	K_2^{C1} K_3^{C1} K_4^{C2} K_5^{C3}
Structure	40.78	L_1^{C1} L_1^{C1} L_4^{C1} L_5^{C2} L_0^{C3} L_1^{C3} L_2^{C3}
First returns		
All	60.34	K_2^{C1} K_4^{C1} L_1^{C1} L_2^{C2} K_4^{C3} K_5^{C3} L_1^{C3} L_2^{C3}
Intensity	58.66	K_2^{C1} L_2^{C2} K_4^{C3}
Structure	34.64	L_1^{C1} L_3^{C1} L_1^{C2} L_4^{C3}
Single returns		
All	59.78	K_1^{C1} K_2^{C1} K_4^{C1} K_5^{C3} L_4^{C3}
Intensity	60.34	K_1^{C1} K_2^{C1} K_4^{C1} K_0^{C3} K_1^{C3} K_4^{C3}
Structure	27.37	L_1^{C1} L_1^{C1} L_4^{C1}

Table A.15. Horizontal layer features selected using stepwise feature selection

Feature set	CR	Selected features
All points		
All	77.09	Q_{10}^{C1} Q_{75}^{C1} Q_{max}^{C2} Q_{90}^{C2} Q_{95}^{C2} Q_{99}^{C2} Q_{05}^{C3} Q_{75}^{C3}
Intensity	77.09	Q_{10}^{C1} Q_{75}^{C1} Q_{max}^{C2} Q_{90}^{C2} Q_{95}^{C2} Q_{99}^{C2} Q_{05}^{C3} Q_{75}^{C3}
Structure	55.31	P_{75}^{C1} dns^{C1} α^{C2} P_{01}^{C2} P_{25}^{C2} α^{C3} P_{25}^{C3} P_{95}^{C3}
First returns		
All	76.54	P_{90}^{C1} $\bar{Q}^{C1}$ Q_{05}^{C1} Q_{50}^{C1} Q_{90}^{C2} Q_{95}^{C2} Q_{01}^{C3} Q_{90}^{C3}
Intensity	75.98	$\bar{Q}^{C1}$ Q_{05}^{C1} Q_{50}^{C1} Q_{90}^{C2} Q_{95}^{C2} Q_{01}^{C3} Q_{90}^{C3}
Structure	41.34	P_{10}^{C1} dns^{C1} P_{10}^{C2} P_{25}^{C3}
Single returns		
All	76.54	P_{05}^{C1} $\bar{Q}^{C1}$ Q_{10}^{C1} Q_{90}^{C1} $\bar{Q}^{C2}$ Q_{10}^{C2} Q_{75}^{C2} Q_{75}^{C3}
Intensity	74.3	$\bar{Q}^{C1}$ Q_{10}^{C1} Q_{50}^{C1} $\bar{Q}^{C2}$ Q_{10}^{C2} Q_{75}^{C2} Q_{min}^{C3} Q_{75}^{C3}
Structure	43.58	γ^{C1} P_{10}^{C2} dns^{C2} α^{C3} P_{90}^{C3}

Table A.16. Distribution features selected using stepwise feature selection

26

Feature set	CR	Selected features						
All points								
1550 nm	59.22	J_2	J_e	E_e	AC			
1064 nm	57.54	J_2	J_e	E_1	E_2	E_e	AC	$\bar{P}$
532 nm	59.22	J_e	E_e	AC				
First returns								
1550 nm	56.98	J_2	J_e	E_1	E_e	AC		
1064 nm	56.42	J_2	J_e	E_0	E_e	AC	$\bar{P}$	
532 nm	57.54	J_e	E_0	E_e	AC			
Single returns								
1550 nm	50.84	J_0	J_e	E_e	AC			
1064 nm	44.13	J_e	E_0	AC				
532 nm	50.84	J_e	E_e	AC	$\bar{P}$			

Table A.17. Ellipsoid layer features of each channel selected using stepwise feature selection

Feature set	CR	Selected features		
All points				
1550 nm	41.9	E_e	AC	
1064 nm	40.78	E_2	E_e	AC
532 nm	41.34	E_2	E_e	AC
First returns				
1550 nm	41.9	E_e	AC	
1064 nm	40.22	E_e	AC	
532 nm	42.46	E_e	AC	
Single returns				
1550 nm	34.64	E_e	AC	$\bar{P}$
1064 nm	34.08	E_e	AC	$\bar{P}$
532 nm	33.52	E_e	AC	

Table A.18. Ellipsoid layer structural features of each channel selected using stepwise feature selection

Feature set	CR	Selected features		
All points				
1550 nm	46.93	J_e		
1064 nm	41.9	J_e		
532 nm	43.58	J_2	J_e	
First returns				
1550 nm	49.72	J_0	J_2	J_e
1064 nm	44.13	J_e		
532 nm	43.02	J_e		
Single returns				
1550 nm	46.93	J_0	J_1	J_e
1064 nm	43.02	J_e		
532 nm	41.34	J_e		

Table A.19. Ellipsoid layer spectral features of each channel selected using stepwise feature selection

Feature set	CR	Selected features
All points		
1550 nm	50.28	K_1 K_3 K_5 L_0 L_1 L_6
1064 nm	48.6	K_2 K_3 K_6 L_0 L_2
532 nm	48.04	K_1 K_2 K_3 K_6 L_2 L_6
First returns		
1550 nm	49.16	K_1 K_2 K_6 L_0 L_1 L_3
1064 nm	43.58	K_0 K_2 L_2 L_4
532 nm	46.93	K_1 K_2 K_3 L_1 L_2
Single returns		
1550 nm	49.16	K_0 K_1 K_2 K_3 K_4 L_0
1064 nm	43.02	K_0 K_1 K_2 L_0
532 nm	41.34	K_1 L_1 L_2 L_3 L_4

Table A.20. Horizontal layer features of each channel selected using stepwise feature selection

Feature set	CR	Selected features
All points		
1550 nm	37.43	L_0 L_1 L_3 L_4 L_5
1064 nm	31.84	L_0 L_1 L_2 L_4 L_5
532 nm	28.49	L_0 L_1 L_2 L_4 L_6
First returns		
1550 nm	33.52	L_0 L_1 L_2 L_3 L_4 L_6
1064 nm	30.73	L_1 L_2 L_4
532 nm	32.4	L_0 L_1 L_2 L_4 L_6
Single returns		
1550 nm	27.37	L_0 L_1 L_4
1064 nm	24.02	L_0 L_2
532 nm	27.93	L_1 L_4

Table A.21. Horizontal layer structural features of each channel selected using stepwise feature selection

Feature set	CR	Selected features
All points		
1550 nm	46.37	K_3 K_5
1064 nm	44.13	K_2 K_6
532 nm	46.37	K_0 K_1 K_2 K_4 K_5 K_6
First returns		
1550 nm	44.69	K_2 K_4 K_5
1064 nm	40.78	K_2 K_3 K_6
532 nm	43.58	K_0 K_2 K_3 K_4 K_6
Single returns		
1550 nm	44.69	K_0 K_1 K_4
1064 nm	43.02	K_0 K_1 K_2
532 nm	39.11	K_1 K_2

Table A.22. Horizontal layer spectral features of each channel selected using stepwise feature selection

Feature set	CR	Selected features							
All points									
1550 nm	65.92	P_{10}	P_{25}	Q_{05}	Q_{75}	Q_{90}			
1064 nm	70.39	P_{01}	P_{05}	P_{25}	δ	Q_{05}	Q_{90}	Q_{95}	Q_{99}
532 nm	63.69	γ	P_{10}	Q_{05}	Q_{10}	Q_{25}	Q_{75}	Q_{95}	dns
First returns									
1550 nm	65.92	α	P_{05}	P_{75}	Q_{25}	Q_{50}	Q_{75}	Q_{99}	dns
1064 nm	66.48	P_{01}	P_{25}	$Q_{\max}$	$\bar{Q}$	Q_{10}	Q_{25}	Q_{99}	dns
532 nm	67.6	P_{10}	Q_{01}	Q_{50}	Q_{95}				
Single returns									
1550 nm	65.36	γ	P_{25}	P_{90}	P_{95}	$Q_{\min}$	$\bar{Q}$	Q_{10}	Q_{25}
1064 nm	64.25	γ	P_{10}	$Q_{\min}$	Q_{05}	Q_{10}	Q_{75}	Q_{99}	dns
532 nm	54.75	P_{25}	$Q_{\max}$	Q_{01}	Q_{50}	Q_{99}			

Table A.23. Distribution features of each channel selected using stepwise feature selection

Feature set	CR	Selected features						
All points								
1550 nm	40.22	P_{05}	P_{25}	P_{50}	P_{95}	dns		
1064 nm	37.99	P_{01}	P_{25}	P_{75}	P_{90}	dns		
532 nm	42.46	α	P_{05}	P_{75}	dns			
First returns								
1550 nm	39.66	P_{10}	dns					
1064 nm	35.75	P_{25}	P_{90}	dns				
532 nm	41.9	α	γ	P_{10}	P_{25}	P_{50}	P_{95}	dns
Single returns								
1550 nm	39.11	γ	P_{25}	dns				
1064 nm	40.22	α	P_{10}	P_{25}	P_{75}	dns		
532 nm	37.43	α	P_{25}	P_{50}	P_{75}	dns		

Table A.24. Structural distribution features of each channel selected using stepwise feature selection

Feature set	CR	Selected features						
All points								
1550 nm	64.25	$\bar{Q}$	Q_{05}	Q_{25}	Q_{75}	Q_{90}		
1064 nm	67.6	$Q_{\max}$	δ	Q_{05}	Q_{90}	Q_{95}	Q_{99}	
532 nm	55.31	Q_{25}	Q_{75}	Q_{90}				
First returns								
1550 nm	60.89	$\bar{Q}$	Q_{25}	Q_{50}	Q_{75}	Q_{99}		
1064 nm	67.6	$Q_{\max}$	$\bar{Q}$	δ	Q_{10}	Q_{50}	Q_{75}	Q_{99}
532 nm	56.42	δ	Q_{25}	Q_{95}	Q_{99}			
Single returns								
1550 nm	55.87	$\bar{Q}$	Q_{50}	Q_{75}	Q_{90}	Q_{95}		
1064 nm	56.42	$\bar{Q}$	δ	Q_{25}	Q_{75}			
532 nm	46.93	$Q_{\min}$	$\bar{Q}$	Q_{01}	Q_{50}			

Table A.25. Distribution spectral features of each channel selected using stepwise feature selection

Bibliography

Åge, P. J. (1983). *Mätning och tolkning i flygbilder för skogsinventering.* research report 1983:5. Lantmäteriverket, p. 44.

Axelson, H. and Nilsson, B. (1993). "Skoglig flygbildstolkning". In: *Flygbildsteknik och Fjärranalys.* Jönköping: Skogsstyrelsen, pp. 295–336. ISBN: 91-884 62-04-8.

Baltsavias, E. (1999). "Airborne laser scanning: basic relations and formulas". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 54.2, pp. 199–214. ISSN: 0924-2716. DOI: `10.1016/S0924-2716(99)00015-5`. URL: `http://www.sciencedirect.com/science/article/pii/S0924271699000155`.

Brandtberg, T. (2007). "Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 61.5, pp. 325–340. ISSN: 0924-2716. DOI: `10.1016/j.isprsjprs.2006.10.006`. URL: `http://www.sciencedirect.com/science/article/pii/S0924271606001274`.

Brandtberg, T. et al. (2003). "Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America". In: *Remote Sensing of Environment* 85.3, pp. 290–303. ISSN: 0034-4257. DOI: `10.1016/S0034-4257(03)00008-7`.

Cottin, A., Fleming, S., and Woodhouse, I. H. (2015). "The use of Multispectral LiDAR for Improved Forest Mapping". In: *Proceedings of SilviLaser 2015.* 14th conference on Lidar Applications for Assessing and Managing Forest Ecosystems. La Grande Motte, France, pp. 56–58.

Dalponte, M. et al. (2013). "Tree Species Classification in Boreal Forests With Hyperspectral Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 51.5, pp. 2632–2645. DOI: `10.1109/TGRS.2012.2216272`.

Fisher, R. A. (1936). "The Use of Multiple Measurements in Taxonomic Problems". In: *Annals of Eugenics* 7.2, pp. 179–188. DOI: `10.1111/j.1469-1809.1936.tb02137.x`.

Garczarek, U. M. (2002). "Classification Rules in Standardized Partition Spaces". PhD thesis. University of Dortmund.

Guyon, I. and Elisseeff, A. (2003). "An Introduction to Variable and Feature Selection". In: *Journal of Machine Learning Research* 3, pp. 1157–1182. ISSN: 1532-4435.

Holmgren, J. and Persson, Å. (2004). "Identifying species of individual trees using airborne laser scanner". In: *Remote Sensing of Environment* 90.4, pp. 415–423. DOI: `10.1016/S0034-4257(03)00140-8`.

Holmgren, J., Persson, Å., and Söderman, U. (2008). "Species identification of individual trees by combining high resolution LiDAR data with multi-

spectral images". In: *International Journal of Remote Sensing* 29.5, pp. 1537–1552. DOI: 10.1080/01431160701736471.

Koenig, K. and Höfle, B. (2016). "Full-Waveform Airborne Laser Scanning in Vegetation Studies—A Review of Point Cloud and Waveform Features for Tree Species Classification". In: *Forests* 7.9. ISSN: 1999-4907. DOI: 10.3390/f7090198.

Lantmäteriet (2016). *Quality description of National Elevation Model.* Version 1.1. URL: http://www.lantmateriet.se/globalassets/kartor-och-geografisk-information/hojddata/produktbeskrivningar/eng/quality_description_dem.pdf.

Lillesand, T. M., Kiefer, R. W., and Chipman, J. W. (2007). *Remote sensing and image interpretation.* 6th ed. John Wiley & Sons. ISBN: 81-265-3223-8.

Lin, Y. and Herold, M. (2016). "Tree species classification based on explicit tree structure feature parameters derived from static terrestrial laser scanning data". In: *Agricultural and Forest Meteorology* 216, pp. 105–114. DOI: 10.1016/j.agrformet.2015.10.008.

Lindberg, E. et al. (2014). "Delineation of Tree Crowns and Tree Species Classification From Full-Waveform Airborne Laser Scanning Data Using 3-D Ellipsoidal Clustering". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.7, pp. 3174–3181. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2014.2331276.

Lindberg, E. et al. (2015). "Multi-wavelength Airborne Laser Scanning for Characterization of Tree Species". In: *Proceedings of SilviLaser 2015.* La Grande Motte, France, pp. 271–273.

Matikainen, L., Hyyppä, J., and Litkey, P. (2016). "Multispectral Airborne Laser Scanning for Automated Map Updating". In: *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B3, pp. 323–330. DOI: 10.5194/isprsarchives-xli-b3-323-2016.

Næsset, E. et al. (2004). "Laser Scanning of Forest Resources: The Nordic Experience". In: *Scandinavian Journal of Forest Research* 19, pp. 482–499.

Nilsson, M. et al. (2017). "A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory". In: *Remote Sensing of Environment* 194, pp. 447–454. DOI: 10.1016/j.rse.2016.10.022.

St-Onge, B. and Budei, C. (2015). "Individual tree species identification using the multispectral return intensities of the Optech Titan lidar system". In: *Proceedings of SilviLaser 2015.* La Grande Motte, France, pp. 71–73.

Reese, H. et al. (2002). "Applications using estimates of forest parameters derived from satellite and forest inventory data". In: *Computers and Electronics in Agriculture* 37.1, pp. 37–55. DOI: 10.1016/S0168-1699(02)00118-7. URL: http://www.sciencedirect.com/science/article/pii/S0168169902001187.

Reese, H. et al. (2003). "Countrywide Estimates of Forest Variables Using Satellite Data and Field Data from the National Forest Inventory". In: *Ambio* 32.8, pp. 542–548. DOI: 10.1639/0044-7447(2003)032[0542:CEOFVU]2.0.CO;2.

Samuels, M. L., Witmer, J. A., and Schaffner, A. (2012). *Statistics For the Life Sciences.* Pearson. ISBN: 0-321-70967-5.

Sandström, J. et al. (2015). *Tillstånd och trender för arter och deras livsmiljöer – rödlistade arter i Sverige 2015*. Swedish. Artdatabanken rapporterar 17. Artdatabanken, SLU.

Skogsstyrelsen (2016). *Skogliga grunddata – produktbeskrivning*. @manual. URL: https://www.skogsstyrelsen.se/globalassets/sjalvservice/karttjan ster/geodatatjanster/produktbeskrivning-och-teknisk-beskrivning/ skogliga-grunddata%E2%80%94produktbeskrivning.pdf.

Teledyne Optech (2015). *Optech Titan Multispectral Lidar System*. @manual. URL: http://www.teledyneoptech.com/wp-content/uploads/Titan-Specsheet-150515-WEB.pdf.

Wagner, W. et al. (2008). "3D vegetation mapping using small-footprint full-waveform airborne laser scanners". In: *International Journal of Remote Sensing* 29.5, pp. 1433–1452. DOI: 10.1080/01431160701736398. URL: https://doi.org/10.1080/01431160701736398.

Weihs, C. et al. (2005). "klaR Analyzing German Business Cycles". In: *Data Analysis and Decision Support*. Ed. by Baier, D., Decker, R., and Schmidt-Thieme, L. Berlin: Springer-Verlag, pp. 335–343.

Wonnacott, T. H. and Wonnacott, R. J. (1990). *Introductory Statistics, 5th Edition*. Wiley. ISBN: 0471615188. URL: https://www.amazon.com/Introd uctory-Statistics-5th-Thomas-Wonnacott/dp/0471615188?Subscriptio nId=0JYN1NVW651KCA56C102&tag=techkie-20&linkCode=xm2&camp=2025& creative=165953&creativeASIN=0471615188.