Remote Sensing of Environment 140 (2014) 306–317

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data

Michele Dalponte [a,*], Hans Ole Ørka [b], Liviu Theodor Ene [b], Terje Gobakken [b], Erik Næsset [b]

[a] Department of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre, Fondazione E. Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy
[b] Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Ås, Norway

ARTICLE INFO

Article history:
Received 8 February 2013
Received in revised form 9 September 2013
Accepted 10 September 2013
Available online 1 October 2013

Keywords:
ALS
Hyperspectral
Tree species classification
Individual tree crowns
Delineation
Forest inventory
Post classification

ABSTRACT

Tree species classification accuracy at the individual tree crown (ITC) level depends on many factors, among which in this paper we analyzed: i) the remote sensing data used for the ITC delineation process carried out prior to the classification, and ii) the pixels considered inside each ITC during the classification process. These two factors were analyzed on the ITC level classification accuracy of boreal tree species (Pine, Spruce and Broadleaves), considering two remote sensing data types: hyperspectral and airborne laser scanning (ALS). ITCs were delineated automatically on ALS and on hyperspectral data. A manual ITC delineation was used as reference in the analysis. The pixel level classification was performed on the hyperspectral bands using a non-linear support vector machine. The classification at ITC level was obtained by applying a majority voting rule to the classified pixels confined by each ITC. The results showed that ITCs automatically delineated from hyperspectral data were usually smaller than those from ALS, and the tree detection rate for hyperspectral data was much lower compared to ALS data (28.4 versus 48.5%). Regarding the classification results, using only manually delineated ITCs a kappa accuracy of 0.89 was obtained, while using only automatically delineated ITCs from hyperspectral or ALS data reduced the kappa values to 0.79 and 0.76, respectively. Slightly different results were achieved using semi-automatic approaches based on both manual and automatically delineated ITC (0.81 and 0.74, respectively). A selection of only certain pixels inside each ITC improved the classification accuracy from 1 to 7 percentage points. A selection based on the spectral values of the pixels was found more influential than the one based on the ALS-derived canopy height model. The best results were obtained after a selection based on the spectral values in the bands in the blue region of the spectrum using either the Otsu method or an ad-hoc percentile-based thresholding method.

1. Introduction

During recent years, extensive research has been devoted to the use of remote sensing data for forestry applications. Many kinds of remote sensing sensors have been tested, from multispectral to hyperspectral, and from RADAR to airborne laser scanning (ALS). In this paper the attention was focused on ALS and hyperspectral data, that, regarding forest studies, represent the most advanced and promising data sources. Hyperspectral data provide detailed spectral information of tree's canopies due to their dense spectral sampling (e.g., hundreds of narrow wavelengths). These data have a high potential in tree species classification when the tree species have different spectral responses (Clark, Roberts, & Clark, 2005; Dalponte, Bruzzone, Vescovo, & Gianelle, 2009; Heinzel & Koch, 2012) and they have shown good performance for tree species discrimination in different types of forest environments, from tropical (Clark et al., 2005) to boreal (Dalponte, Ørka, Gobakken, Gianelle, & Næsset, 2013). Airborne laser scanning (ALS) data provide three-dimensional information of the forest canopy structure, and they have been widely used for estimation of biophysical forest properties such as stem volume and tree height (Næsset, 1997a, 1997b) as well as for tree species classification (Brandtberg, 2007; Heinzel & Koch, 2011; Holmgren & Persson, 2004; Ørka et al., 2009).

Although the focus in tree species classification using hyperspectral data has mainly been on the pixel-level classification and mapping, a classification map at the individual tree crown (ITC) level can be easily linked to the tree's biophysical properties and it can be used for practical forest management purposes. Studies on ITC level species classifications have been presented (Clark et al., 2005; Heinzel & Koch, 2012), and higher classification accuracies have been reported compared to pixel level (Clark et al., 2005; Dalponte et al., 2013; Yan, Maathuis, Xiangmin, & Dijk, 2007). ITC level information can be obtained using two main approaches: i) either a *direct classification* where each ITC is treated as one observation; or ii) a *post-classification* approach (Hutchinson, 1982) treating each pixel as one observation and aggregating the pixel level species classifications inside each ITC.

An outcome of an ITC classification based either on *direct classification* or *post-classification*, likely depends on the ITC delineation carried out prior to the classification. ITCs are often obtained from an automatic delineation method using various types of remote sensing data (e.g., ALS, multispectral and hyperspectral data). Concerning the hyperspectral data, common algorithms developed for image segmentation can be

* Corresponding author.
E-mail address: michele.dalponte@fmach.it (M. Dalponte).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.09.006

used, like marked-based watershed algorithm (Wang, Gong, & Biging, 2004) or region growing (Ardila, Bijker, Tolpekin, & Stein, 2012; Bunting & Lucas, 2006). Moreover, in the delineation process, the rich spectral information contained in these data can be exploited by using multiple bands (Bunting & Lucas, 2006) or features extracted from the hyperspectral bands (e.g., principal components, spectral indices; Wang et al., 2004; Ardila et al., 2012). Using ALS data, ITCs can be delineated with algorithms that can be applied to a rasterized canopy height model (CHM; Hyyppä, Kelle, Lehikoinen, & Inkinen, 2001; Popescu & Wynne, 2004; Solberg, Næsset, & Bollandsås, 2006; Ene, Næsset, & Gobakken, 2012) or directly to the three-dimensional coordinates (Li, Guo, Jakubowski, & Kelly, 2012; Rahman, Gorte, & Bucksch, 2009; Wang, Weinacker, & Koch, 2008).

Despite the abundance of studies comparing the performances of ITC delineation algorithms (Erikson & Olofsson, 2005; Kaartinen et al., 2012; Vauhkonen et al., 2012), only few of them analyzed the influence of using different data sources on the ITC delineation results. To the best of our knowledge, only a few studies that compared ITC delineated with ALS and aerial imagery data exist (Coops, Wulder, Culvenor, & St-Onge, 2004; Leckie et al., 2003), and we are not aware of any study which compared the performance of ALS and hyperspectral data in ITC delineation for classification purposes.

Furthermore, it is recognized that not all the pixels inside an ITC provide the same information for the classification (Clark et al., 2005; Heinzel & Koch, 2012; Puttonen, Litkey, & Hyyppä, 2009). Due to issues such as shadowing or errors in ITC delineation, some pixels may only contain noise. Shadowed pixels are a particular problem at northern latitudes where the solar zenith angle is always quite large and thus shadows are present even in the best acquisition periods. Removing shaded pixels by retaining only the sunlit ones has been suggested to account for the shadowing problem (Clark et al., 2005; Korpela, Heikkinen, Honkavaara, Rohrbach, & Tokola, 2011; Puttonen et al., 2009). Leckie et al. (2005) considered only the pixels above a threshold defined in the green channel while Clark et al. (2005) considered only pixels having in the 800 nm band a reflectance value higher than the median crown reflectance of that band. Heinzel and Koch (2012) excluded the border pixels among the ITCs automatically delineated with a watershed segmentation algorithm on a normalized digital surface model image based on the assumption that these pixels had a higher probability to be shadowed or noisy. Ardila et al. (2012) in an urban environment applied a threshold on the NDVI values discarding the left-hand tail side of the NDVI distribution with 10%, 15%, 20%, and 25% thresholds. The effects of the errors in the delineation of ITCs, such as parts of a neighboring tree of another species being included in a given ITC, are less studied. Furthermore, fusing remote sensing data from different acquisitions might also introduce errors due to misalignment of the data. To the best of our knowledge, no previous studies have analyzed the effects of different thresholding options on the classification accuracy at ITC level. Moreover, no previous study proposed an automatic and adaptive thresholding method for this purpose. An automatic and adaptive thresholding method can be very useful in this domain as it removes any subjectivity present in the previously cited methods, making the thresholding system adaptable to different datasets.

Thus, the aim of this research was to analyze the effects of i) the use of hyperspectral and ALS data in the automatic ITC delineation, and ii) applying various thresholding methods to the pixels inside each ITC, on tree species classification accuracy. This last analysis comprises the definition of a novel adaptive thresholding method that can remove any subjectivity to the thresholding process. This paper aims also to provide important indications to the local forest managers on the use of hyperspectral and ALS data for ITC level tree species classification. In particular, it addresses which data should be used for ITC delineation, the possibility of combining manual and automatic ITC delineations, and the selection of the most informative pixels inside each ITC. All these aspects are very important when dealing with practical applications of remote sensing in forestry.

2. Materials

2.1. Study area

The study area is located in the municipality of Aurskog-Høland, southeastern Norway, 40 km east of Oslo (59°50′N, 11°40′E, and 120–390 m above sea level). Approximately 75% of the total land area of 890 km^2 is characterized by managed, productive forests dominated by *Pinus sylvestris* L. (Scots pine; 50%), *Picea abies* (L.) Karst (Norway spruce; 35%), and deciduous tree species (15%), such as *Betula* spp. L. (birch) and *Populus tremula* L. (aspen).

2.2. Field data

During the fall of 2007 and winter of 2008, field data were collected on 23 circular sample plots, whereby 11 sample plots were located in spruce dominated stands (in basal area) and the remaining 12 in pine dominated stands. The sample plots were located in both mature and young productive forests. The size of the plots was 1000 m^2, except for one of the young forest plots where the plot size was reduced to 500 m^2 due to a very high stem density.

Within each sample plot, tree species, diameter at breast height (DBH), and tree coordinates were recorded for all trees with DBH $\geq$ 5 cm. Totally, 2363 trees were recorded in the 23 plots having a dominant species distribution of 57% spruce, 28% pine, and 15% broadleaves. Tree positions were determined by measuring the azimuth and distance of the tree stems from the plot center with a total station (Topcon Sokkia SET5F). Plot center coordinates were determined using differential Global Navigation Satellite Systems (GNSS) with two Topcon Legacy E+ receivers as base and rover receivers, respectively. Random errors reported from the post-processing indicated an average error of 12 cm for the planimetric coordinates of the plot centers.

The measured trees were divided into three classes according to the dominant species: i) Pine, ii) Spruce, and iii) Broadleaves (see Table 1). Henceforth we will refer to the classes using uppercase initials, and to the tree species using lowercases.

2.3. Remote sensing data

Airborne hyperspectral data were acquired on the 28th of July 2008 with the HySpex VNIR-1600 sensor mounted on a Piper Chieftain PA-31-350 fixed-wing aircraft. The average flying altitude was 1500 m above ground level, and the average flying speed was of 70 m/s. The spectral resolution of the sensor was 3.7 nm, and 160 bands ranging from 410 nm and 990 nm were recorded. Three hyperspectral images were acquired in order to cover all the 23 plots.

ALS data were acquired using a PA31 Piper Navajo fixed-wing aircraft at a mean flying altitude above ground of 800 m. The data were collected under leaf-on canopy conditions on the 6th of June 2006 with an Optech ALTM 3100 system operating with a pulse repetition frequency of 100 kHz, scan frequency of 70 Hz, and using a half-angle of 5°. Up to four echoes per pulse were recorded (approximately 1.4 echoes per emitted pulse on average) and the resulting number of first returns was 7.4 echoes per m^2.

Table 1
Data used in terms of pixels and ITCs obtained with the three delineation methods, and the number of trees measured on the ground for each class.

	Ground measured trees	M-ITCs		H-ITCs		L-ITCs	
		Pixels	Trees	Pixels	Trees	Pixels	Trees
Pine	642	9738	399	17,721	263	32,582	438
Spruce	1358	7850	502	16,747	340	34,732	580
Broadleaves	363	2137	101	4874	69	8778	129
Total	2363	19,725	1002	39,342	672	76,092	1147

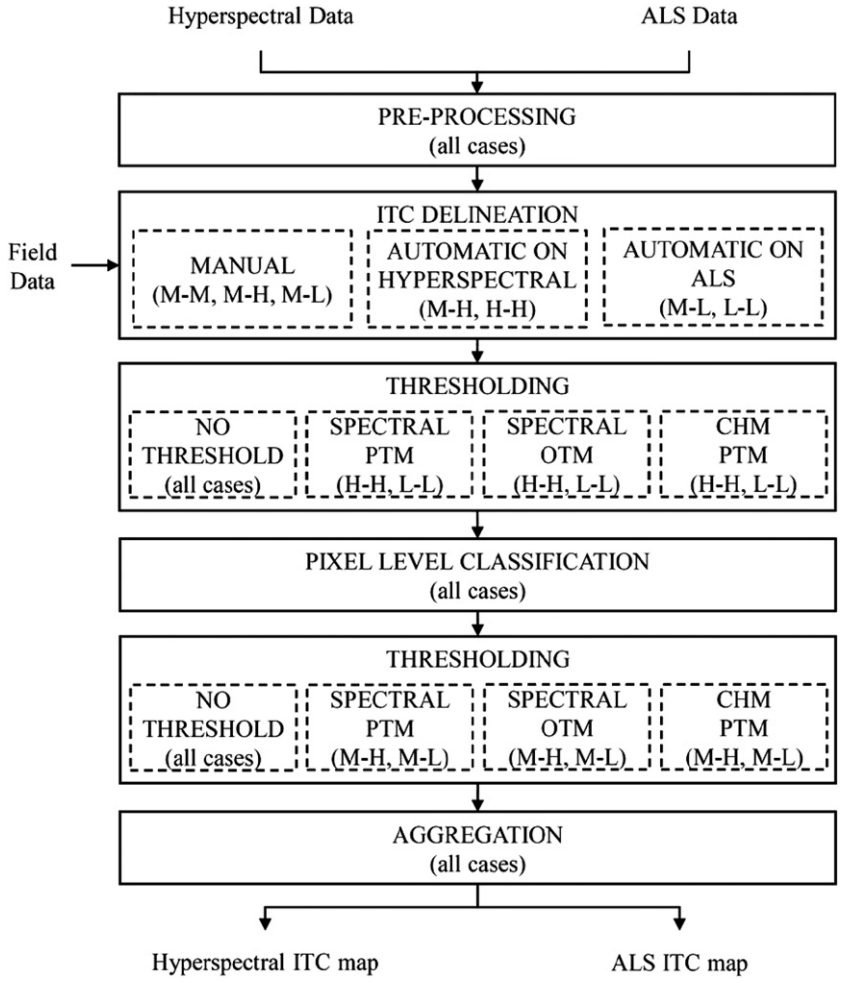

Fig. 1. Processing chain adopted in this study. The case in which the step is applied is indicated in brackets.

3. Methods

The main steps of the data processing work flow adopted in this study are presented in Fig. 1. The processing chain considered for the hyperspectral data follows the indications of a previous study on the same dataset (Dalponte et al., 2013).

3.1. Remote sensing data pre-processing

The hyperspectral images were orthorectified using a digital terrain model (DTM). The ground sampling distance of the orthorectified images was 0.4 m. A radiometric normalization of the pixels in the three images was performed by dividing each pixel value by the sum of the values of that pixel in all the bands (Yu, Ostland, Gong, & Pu, 1999). From a preliminary analysis the normalization process had the effect to improve the kappa accuracy of about 7%, especially for the Broadleaves class. The normalization had also an effect on the spectral signatures, in particular the range of spectral values in the NIR bands was reduced. In the preliminary analyses we applied also a generic atmospheric correction (i.e., QUAC algorithm; Bernstein, Sundberg, Levine, Perkins, & Berk, 2005), but, as it was not providing any improvement in the classification results, in the final experiments we preferred to use no corrections.

The pre-processing of the raw ALS data was performed by the data vendor. The ALS echoes were classified in ground and non-ground classes using the TerraScan software,[1] and a DTM was built from the ground echoes. The DTM values were subsequently subtracted from the height values of all the echoes recorded to obtain the heights above the ground level (relative heights) of the echoes. According to the vendor's technical report, the expected accuracy of the relative heights was ± 10 cm. The non-ground echoes were further used for creating canopy height models (CHMs) at a spatial resolution of 0.2 m for each of the 23 plots, using natural neighbor interpolation.

In order to have the best possible overlap among hyperspectral and CHM images, the two data were co-registered among each other. A separate co-registration for each plot was carried out. An RST transformation with about 10 ground control points per plot was applied using the commercial software ENVI.[2]

3.2. Delineation of ITCs

Three types of ITCs were used in this study: i) manually delineated (M-ITCs), ii) automatically delineated from hyperspectral data (H-ITCs), and iii) automatically delineated from ALS data (L-ITCs).

The M-ITCs were delineated on hyperspectral data, since we found easier to discriminate individual tree crowns due to their species variability. For the delineation we used a RGB false color composition, and we considered different RGB compositions according to the species present in the plots. One and only one of the three classes (Pine, Spruce or Broadleaves) was assigned to each ITC using the geo-referenced ground measured sample trees. In order to reduce the errors in the

[1] http://www.terrasolid.fi/en/products/terrascan

[2] http://www.exelisvis.com/ProductsServices/ENVI/ENVI.aspx

data set, we retained only the segments which i) contained trees of the same species, and ii) have ALS tree top height that matches the ground measured height.

The automatic delineations of the H-ITCs and L-ITCs were performed using an adaptive method based on a Poisson forest stand model (Ene et al., 2012), which demonstrated a good balance between omission and commission errors in a comparative study of different algorithms and test sites operating on a CHM derived from ALS data (Vauhkonen et al., 2012). The algorithm operates on a single raster band and utilizes the area-based stem density estimate for each plot as input parameter. Assuming that the stems are uniformly randomly distributed inside each plot, the stem density allowed to calculate the expected nearest neighbor distance between stems within each plot. Then, the size of the low-pass filter used to smooth the raster image (CHM or hyperspectral band) prior to segmentation was approximated to the expected nearest neighbor stem distance. The ITCs were delineated using a marker-based watershed algorithm on the part of the raster defined as canopy after a histogram thresholding. The markers were the local maxima in the smoothed image. The delineation algorithm provided also the estimated tree positions that were defined as the position of the highest ALS echo inside the ITC for the L-ITC delineation, while for the H-ITC it was the position of the brightest pixel in the raster image used for the delineation.

The approach for obtaining the area-based predictions of the stem number varied according to the remote sensing data used (hyperspectral and ALS). When using the hyperspectral images the method in Ene et al. (2012) was slightly modified. First, for each plot, a histogram thresholding operation using Otsu's method (Otsu, 1979) was applied on band 107 (810 nm). The Otsu method is a histogram-based thresholding method which aims in finding a thresholding value which splits the distribution of the pixel values of an image into two classes (foreground and background) such that the within class variances are minimized (or the between-class variance is maximized). The 810 nm band was previously used for the separation of vegetation and background in a previous study (Clark et al., 2005). Each plot image was then binarized into two categories: background ($\leq$threshold value) and canopy (>threshold value). The resulting canopy mask was further processed using a morphological opening with a structural element of three pixels in order to remove spurious object pixels appearing after thresholding. The size of the opening filter was chosen in order to be the smallest one, to be sure to eliminate only the commission errors of the watershed segmentation. An estimate of the canopy cover was also calculated as the proportion of object pixels covering the plot. Imposing the opened mask on the original images, textural features were extracted using gray level co-occurrence matrices (Haralick, Shanmugam, & Dinstein, 1973) calculated on four directions (0°,45°, 90°, and 130°), and considering three offsets (1, 2, and 3 pixels). The final textural features were obtained by averaging over all directions and offsets. The stem number estimates at plot level were obtained using a predictive generalized regression model (GLM; Nelder & Wedderburn, 1972). The model predictors considered were a) the average contrast feature produced by the gray level co-occurrence matrices, b) the canopy cover estimated as from the proportion of canopy pixels, and c) their product as interaction term. The root mean square error (RMSE) of the model was 26.5% and the linear correlation coefficient between the empirical and estimated stem numbers was $r = 0.700$. However, in order to mimic a practical situation when the model would be used for prediction on new data, the final plot-wise stem number estimates were obtained using a leave-one-out cross validation (RMSE = 28.8%, $r = 0.560$).

When using ALS data, the model for the area-based predictions of the stem densities (RMSE = 23.4%, $r = 0.730$) was developed from the first returns of the ALS data. Inside each plot, the echoes were classified into canopy and background using the minimum error thresholding (Kittler & Illingworth, 1986), as described in Ene et al. (2012). The variables included in the selected prediction model were a) the canopy coverage estimated as the proportion of canopy echoes to the total number of echoes, b) the proportion of echoes in the lowest interval obtained by dividing the range of canopy height echoes in ten equal intervals, and c) the interquartile range of the height quartiles. The estimates after leave-one-out cross validation (RMSE = 28.8%, $r = 0.610$) were used as input in the delineation algorithm.

3.3. Classification

The pixel level classification was carried out using the support vector machine classifier (SVM), having as input features all the hyperspectral bands acquired by the sensor. The SVM classifier has been proved to be effective in tree species classification (Dalponte, Bruzzone, & Gianelle, 2012; Jones, Coops, & Sharma, 2010; Koetz, Morsdorf, Vanderlinden, Curt, & Allgower, 2008). Further details on the SVM classifiers and their use with hyperspectral data can be found in Melgani and Bruzzone (2004). In this study, we used the R (R Development Core Team, 2008) implementation of SVM available in the package *kernlab* (Karatzoglou, Smola, Hornik, & Zeileis, 2004). In all cases, the classification at ITC level was obtained by aggregating the classified pixels inside each ITC according to a majority rule.

In the current study, five classification cases were analyzed: i) a fully manual case based on manually delineated ITCs (the M–M case), ii–iii) two fully automatic cases based on ITCs automatically delineated on hyperspectral data (the H–H case) and on ALS data (the L–L case), and iv–v) two semi-automatic cases that consider manually delineated ITCs in the training phase and ITCs automatically delineated on hyperspectral data (the M–H case) and on ALS data (the M–L case) in the validation phase.

3.4. ITC thresholding

The thresholding of the pixels confined by each ITC was applied prior to classification in the H–H and L–L cases (thus influencing the training phase of the classifier) and as a post-classification procedure in the M–H and M–L cases (without influencing the training phase of the classifiers). Two thresholding methods were tested: i) the automatic Otsu thresholding method (OTM; Otsu, 1979), and ii) a percentile-based thresholding (PTM). As we wanted to compare the optimal performances of each thresholding method, we tested the PTM and OTM thresholding over all the hyperspectral bands. For the PTM thresholding, we tested the percentiles ranging between the 5th and the 95th percentile, in steps of 5%. The OTM thresholding was tested also over the average of the bands in the blue spectrum used by satellite multispectral sensors (e.g., Landsat TM, GeoEye-2; 450–520 nm). OTM and PTM were also tested on the height values of CHM pixels contained by each ITC. They were applied as lower bound thresholding (removing pixels below the threshold value) for the CHM pixels, and as lower and upper bound thresholding (removing pixels below or above the threshold value) for the hyperspectral imagery. In both cases, the pixels which were not removed (called here selected pixels) were used further for classification. Finally, the results were assessed against the case of using all the ITC pixels (no thresholding), and two reference thresholding methods: i) the mean of the spectral values in the 810 nm band (mean IR; Clark et al., 2005); and ii) the mean NDVI in each ITC computed before the radiometric normalization (mean NDVI). It is worth noting that the analyses of PTM and OTM thresholding over all the spectral bands represent an exploratory analysis and they are not part of the proposed automatic thresholding method.

3.5. Accuracy assessment

The validation of the classification results was performed using a k-fold cross validation method where each plot was a fold, i.e., the classifier was trained with the trees belonging to $k - 1$ plots (where k is the number of plots; $k = 23$ in our study) and validated on the left-out plot. This process was repeated $k = 23$ times.

The results are presented in terms of overall accuracy (OA), kappa accuracy (KA), and producer's and user's accuracies (PA and UA, respectively) of the three classes (Spruce, Pine, and Broadleaves). The model selection of the SVM was performed using a five-fold cross validation on the training dataset of size N − 1.

The trees identified by both the delineation based on the hyperspectral data and on the ALS data were analyzed further. The relationship between the classification accuracy and the distributions of i) the DBH of the trees (measured on the ground), and ii) the crown area provided by the two delineation methods, was investigated using analysis of variance (ANOVA) and additional multiple comparison test of the differences in means (the Tukey's "Honest Significant Difference") implemented in the *stats*-package of R (R Development Core Team, 2008). The choice of this test was driven by the fact that it is more conservative with respect to, for example, a standard *t*-test. The effect of the classification case and of the thresholding were assessed using logistic regression fitted using the tree species as the response variable. The fitting and subsequent multi-comparison analyses were performed using the R-packages *stats* (R Development Core Team, 2008), *rms* (Harrell, 2012), and *multcomp* (Hothorn, Bretz, & Westfall, 2008). A confidence level of 95% was used in all the tests.

4. Results

4.1. Delineation of ITCs

The delineation based on the ALS data provided the highest number of ITCs among the tested methods (1147; 48.5% of the total trees measured on the ground), followed by the manual delineation (1002; 42.4%), and the hyperspectral (672; 28.4%). The average ITC area was 3.15 m^2 for the manually delineated ones, 9.37 m^2 for the hyperspectral ones, and 10.61 m^2 for the ALS. The largest ITC detected by M-ITC had an area of 22 m^2, while the largest one detected by H-ITC and L-ITC was of 40 m^2 and 58 m^2, respectively. As a consequence, the number of pixels extracted from the hyperspectral image in the L-ITCs was almost twice as high as the number of pixels in the H-ITCs, and almost four times the manual ones (Table 1). An example of ITCs obtained with the three delineation methods for a field plot is presented in Fig. 2.

From all the ITCs delineated by the two automatic methods, 586 trees were found by both of them (Table 2). The average location error, the distance of the estimated tree position from the real tree position, was twice as large for the hyperspectral ITCs compared to the ALS ones, and in particular this error was more than 2 pixels in the hyperspectral images (approximately 0.8 m). The main orientation of the shift, with respect to the real position, was along the south-east direction.

4.2. Classification accuracy

The accuracies obtained using the M-ITCs were the highest with a kappa value of 0.887 (Table 3). The H-ITCs provided slightly higher accuracies than the L-ITCs (Table 3). The semi-automatic cases (M–H and M–L) provided higher accuracies for H-ITCs, but not for L-ITCs when compared to H–H and L–L, respectively. The Broadleaves class was always poorly classified (less than 50% PA) when using automatically delineated ITCs (H–H and L–L cases). Conversely, using M-ITCs (the M–M case), the PA reached 71.3%. In the M–L case, the classification accuracy of the class Broadleaves improved noticeably compared to the fully automatic L–L case (about 7 percentage points).

4.3. Improving classification accuracy in the semi-automatic cases

The band used for the PTM influenced similarly the kappa accuracies obtained in the two semi-automatic cases (M–H and M–L) (see case M–H in Fig. 3). Applying a percentile thresholding as upper bound on the pixels of the blue bands improved the kappa accuracy in both cases (Table 4). The highest kappa accuracy obtained thresholding bands in the blue region of the spectrum was 0.860 for the 469 nm band thresholded with the 30th percentile as upper bound. In general, an upper bound PTM on the bands in the blue spectrum provided a higher accuracy than without any thresholding, irrespective of the threshold level. Considering the NIR spectrum, frequently used for thresholding, the peak of KA was around 0.850 when using a lower bound thresholding at the 75th percentile on the 764 nm wavelength.

The lower bound PTM thresholding applied to the CHM values inside each ITC did not improve significantly the classification accuracy. Considering the M–H case, the kappa accuracy ranged from a maximum of 0.800 to 0.824 depending on percentile threshold. Similar results were obtained in case M–L where the kappa accuracy ranges from 0.730 to 0.749 depending on the percentile threshold.

The KAs obtained using the OTM threshold in each band (Fig. 4 for the M–H case) were quite similar to the ones obtained using PTM thresholds on the same bands (Fig. 3). When using OTM, the highest accuracies were obtained with an upper bound thresholding in the blue region of the spectrum for both M–H and M–L cases (Table 4).

Considering the two reference methods, the mean IR thresholding provided identical results as the OTM thresholding on the mean blue

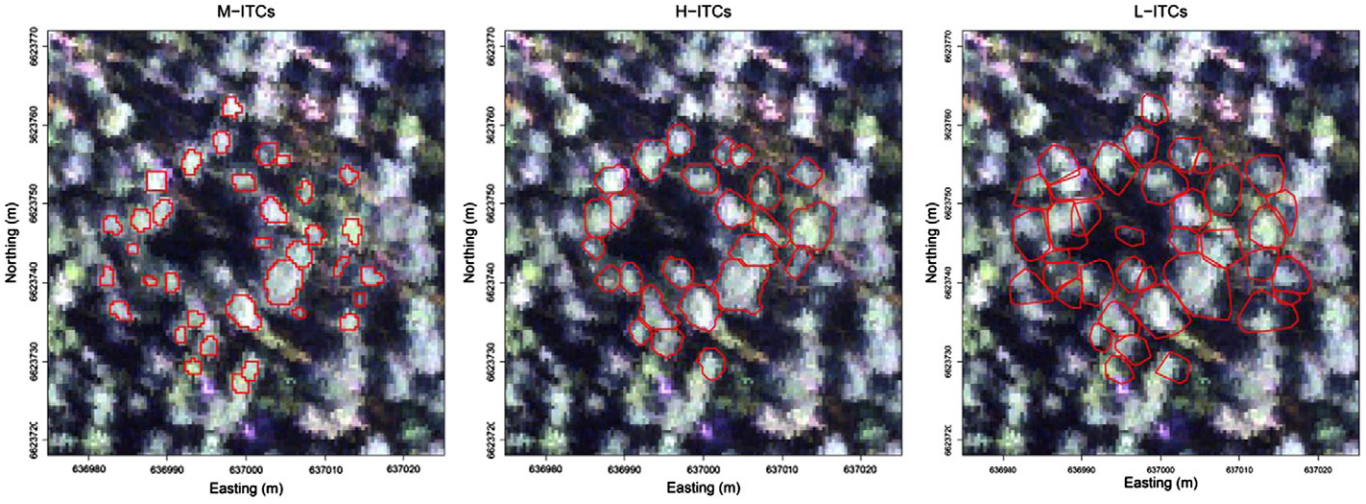

Fig. 2. Example of ITCs for a field plot obtained with the three delineation methods. The RGB images were generated using the following bands before applying the normalization: i) red: 640 nm; ii) green: 550 nm; and iii) blue: 458 nm.

Table 2
Crown area and position of the ground measured trees with respect to the detected ITC position for the common 586 trees detected by both hyperspectral and ALS delineations.

ITCs	Crown area (m^2)			Position of the ground measured tree	
	Min	Mean	Max	Radius (m)	Azimuth (°)
H-ITC	0.4	12.4	58.3	0.83	171.5
L-ITC	1.0	12.7	52.4	0.46	159.1

band, while the best PTM thresholding provided higher accuracies. Regarding the thresholding based on the mean NDVI, the accuracies were much lower. The pixels selected by the OTM thresholding on the mean blue band and by the mean IR thresholding were quite similar. For the H-ITCs, 51.8% of the pixels were selected by both methods, 21.1% were not selected by neither of them, and 27.1% of the pixels were selected only by one of the two methods. For the L-ITCs, 28.4% of the pixels were not selected by both methods, 52.6% of the pixels were selected by both methods, and 19.0% of the pixels were selected by only one of the two methods.

The kappa accuracies obtained with PTM and OTM in the blue region, and mean IR thresholding resulted significantly different from the no thresholding case at 95% confidence interval using a Z-test (Congalton & Green, 2008).

4.4. Improving classification accuracy in the automatic cases

The results of cases H–H and L–L after the application of the best PTM and OTM thresholds found for cases M–H and M–L, and after the application of the OTM thresholding over the mean blue band, and the mean IR thresholding appear in Table 5. For cases L–L and H–H all the thresholds tested improved the classification accuracies. The spectral PTM and OTM provided almost equivalent results with an improvement in kappa accuracy of approximately 2 percentage points, while when applied on the CHM they provided a slightly lower improvement (around 1 percentage points) concentrated mainly in the Broadleaves class (approximately 10 percentage points). Also in this case the mean IR thresholding provided very similar results to the mean blue band OTM thresholding, while the mean NDVI thresholding provided much lower results.

4.5. Analysing the spatial distribution of selected pixels within the ITCs

Each ITC was divided into eight sectors relative to the north direction (one sector every 45°). Inside each sector the number of selected pixels by the OTM thresholding over the mean blue bands, and the mean IR thresholding, was normalized to a value between 0 and 1 by dividing it by the total number of pixels in the respective sector. Similarly, the distance of each pixel from the tree position found by the delineation algorithm was normalized with respect to the largest distance in each sector. The results are showed in Fig. 5. For the H-ITCs the majority of the selected pixels were on the south-east side of the crown for both the considered thresholds. For the L-ITCs, this trend was weaker but still present (Fig. 5). In Fig. 6 an example of the spatial location of pixels selected by the OTM thresholding is showed. It is worth noting that as expected they were located in the shadowed areas of the ITCs (compare with Fig. 2).

Regarding the distance of each pixel from the tree position found by the delineation algorithm, Fig. 5 shows that the average distance decreased in each sector, indicating that the selected pixels were located mostly in the central part of the ITCs. The south-east trend was still present, as the average distance was higher in the south-east sectors compared to north-west.

4.6. The combined effect of ITC delineation and thresholding method on classification accuracy

The difference in DBH distribution of the automatically delineated ITCs was tested using ANOVA and Tukey-HSD tests for cases M–H and M–L. The analyses revealed that the ITCs classified correctly in both delineations had larger DBHs compared to the misclassified ones ($p < 0.012$) and that there were no differences between the DBH size of the ITCs classified correctly in case M–H and in case M–L ($p > 0.91$).

Regarding the crown area of the H-ITCs, the analyses showed that, when no thresholding was applied, there were significant differences in the ITCs area among four groups of ITCs ($p = 0.004$). The four groups were: 1) the ITCs always correctly classified, 2) the ITCs correctly classified only in H-ITCs, 3) the ITCs correctly classified only in L-ITCs, and 4) the ITCs always incorrectly classified. The same happened when applying the CHM thresholding ($p = 0.042$). In both cases (no thresholding and CHM thresholding) the ITCs classified correctly in the M–L case had larger crown areas provided by the H-ITCs compared to the trees classified correctly only in M–H case ($p < 0.039$). However, the areas were not significantly different at 5% significance level. When no thresholding was applied, the H-ITC area was larger for trees correctly classified in the case M–L compared to ITCs classified correctly both in cases M–L and M–H ($p = 0.017$).

The crown area provided by L-ITCs showed an opposite trend (compared to the area of the H-ITCs) which was found significantly different ($p < 0.05$) only when using the either PTM or OTM spectral thresholding. The L-ITC area was smaller for the M–L case compared to ITCs classified correctly in all cases ($p = 0.026$) using OTM, and for ITCs classified incorrectly in all cases ($p = 0.033$) when PTM on the spectral bands was applied.

A Generalized Linear Model (GLM) was fitted to provide information on the change in probability for a correct classification based on different cases (M–H, M–L, H–H, and L–L) and thresholding methods (PTM and OTM) applied to the hyperspectral imagery and CHM pixels. Both the case ($p < 0.001$) and the thresholding ($p = 0.015$) were significant, while the interaction term of case and thresholding method was tested and was not significant ($p > 0.05$). The Hosmer–Lemeshow goodness of fit test (Hosmer, Hosmer, Le Cessie, & Lemeshow, 1997) indicated an appropriate fit of the model ($p = 0.809$). The coefficients for case M–H were significantly higher than those for case H–H. Furthermore, the analysis showed that the spectral thresholding increased the probability of a correct classification. The cases L–L and M–L and the CHM thresholding were not significant in the model ($p > 0.063$). Multi-comparison test of cases revealed that there were differences between M–L and L–L ($p < 0.001$) and between M–L and M–H ($p < 0.001$). Furthermore, multi-comparison test for the thresholding revealed that there were significant differences between using the spectral PTM and not using any thresholding ($p = 0.015$).

Table 3
Classification results obtained in the five cases at ITC level.

Case	OA (%)	KA	PA (%)			UA (%)		
			Pine	Spruce	Broadleaves	Pine	Spruce	Broadleaves
M–M	93.5	0.887	97.5	94.8	71.3	90.3	96.6	92.3
H–H	87.9	0.786	97.3	92.4	30.4	83.4	92.9	77.7
L–L	86.3	0.759	93.8	92.4	33.3	81.4	91.2	79.6
M–H	89.1	0.811	96.2	91.5	50.7	85.5	95.3	72.0
M–L	85.4	0.739	85.4	95.3	40.3	88.4	83.3	86.7

5. Discussion

The ITC delineation using hyperspectral data identified almost half of the trees detected by the delineation based on ALS data. The limited capacity for ITC detection in spectral images compared to ALS was also mentioned in previous studies (Coops et al., 2004; Rossmann & Buecken, 2010). Two factors could have had a strong influence

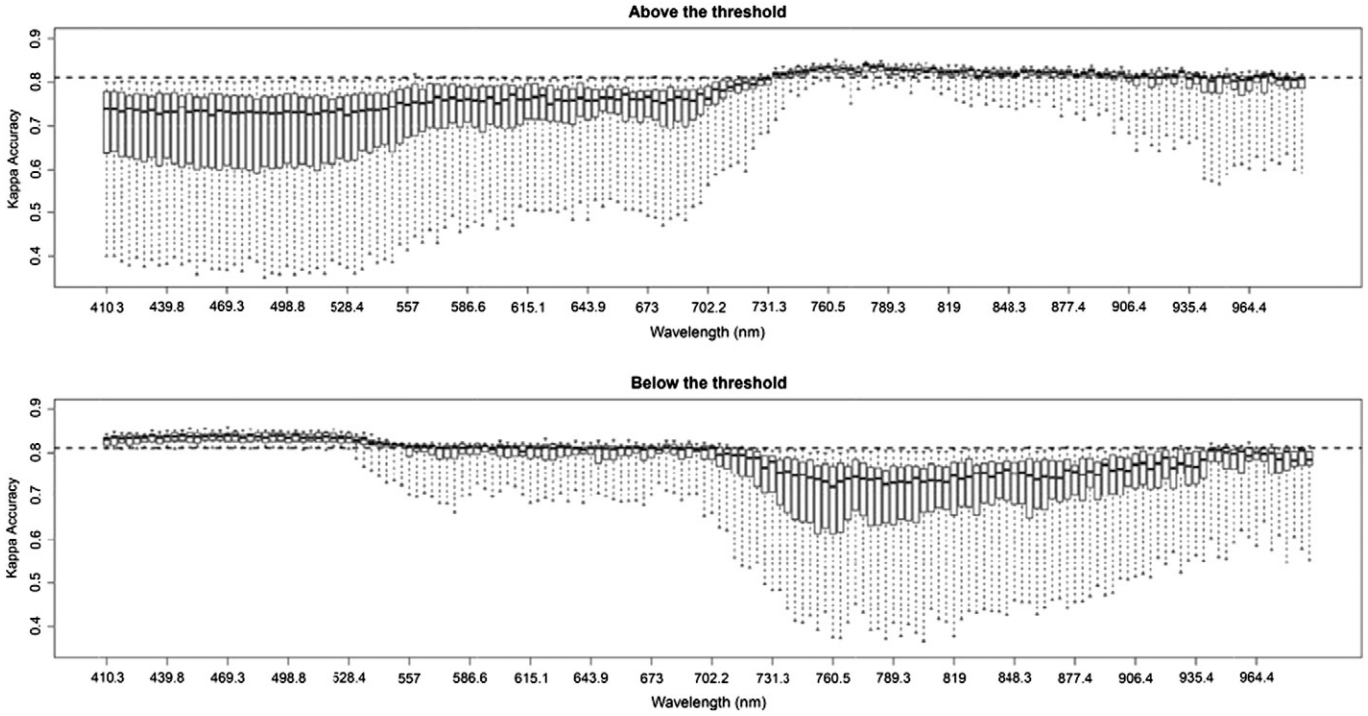

Fig. 3. Boxplot of the KA obtained with various PTM thresholds (from 5th to 95th percentile) on each hyperspectral band for case M–H. The dashed line represents the KA obtained without the use of any threshold. In the upper graph only the pixels whose values were "above" the threshold were considered in the aggregation, while in the lower graph only the pixels whose values were "below" the threshold were considered in the aggregation.

on the detection rate: i) the shadowing in the hyperspectral images, and ii) the coarser spatial resolution of the imagery product. Regarding the shadows in the hyperspectral images, they can be considered partially responsible of the fact that hyperspectral ITCs had a smaller range of sizes compared to ALS ones. ALS data have both smaller and larger ITCs compared to hyperspectral ones. Small trees can be in shadowed areas in hyperspectral data (an example is the central tree in Fig. 2, detected on the ALS data and not in the hyperspectral ones). Regarding the spatial resolution of the data used for the ITC delineation, in the current study, the L-ITC delineation was carried out at a higher spatial resolution (0.20 m) compared to the hyperspectral images (0.4 m). The decision to use a higher spatial resolution for the ALS data was driven by the fact that we wanted to compare the best delineation product possible with the available data. Ene et al. (2012) showed on the same dataset that the detection difference among 0.2 m and 0.4 m was only of 14 trees, but the accuracy index based on omission and commission errors was 3% higher with a resolution of 0.2 m. Hengl (2006) suggested having at least four pixels representing the smallest circular objects and at

Table 4
Best results obtained at ITC level applying the thresholds on the pixels inside the automatically delineated ITCs for the two semi-automatic cases.

Case	Threshold	OA (%)	KA	PA (%)			UA (%)		
				Pine	Spruce	Broadleaves	Pine	Spruce	Broadleaves
M–H	–	89.1	0.811	96.2	91.5	50.7	85.5	95.3	72.0
	30th percentile 469 nm band	91.8	0.860	96.2	92.4	72.5	91.0	96.9	71.4
	45th percentile CHM	89.9	0.824	96.7	92.1	52.2	85.9	96.3	72.0
	Otsu 462 nm band	90.5	0.836	97.0	91.5	60.9	87.0	97.2	71.2
	Otsu Mean blue bands	90.3	0.833	96.2	91.5	62.3	86.9	96.9	71.7
	Mean IR (Clark et al., 2005)	90.0	0.827	97.0	91.5	56.5	85.6	97.2	72.2
	Mean NDVI	89.6	0.818	95.1	93.2	50.7	85.9	95.4	71.4
M–L	–	85.4	0.739	85.4	95.3	40.3	88.4	83.3	86.7
	20th percentile 432 nm band	89.4	0.817	93.8	92.1	62.0	87.4	92.9	78.4
	70th percentile CHM	85.9	0.749	88.1	94.7	38.8	86.7	85.5	83.3
	Otsu 425 nm band	87.9	0.788	92.5	93.4	47.3	85.4	90.8	80.3
	Otsu Mean blue bands	87.7	0.785	92.4	93.0	48.1	85.0	90.8	80.5
	Mean IR (Clark et al., 2005)	88.1	0.792	92.4	93.8	48.1	86.5	90.1	82.7
	Mean NDVI	85.8	0.747	86.8	95.2	40.3	87.8	84.8	82.5

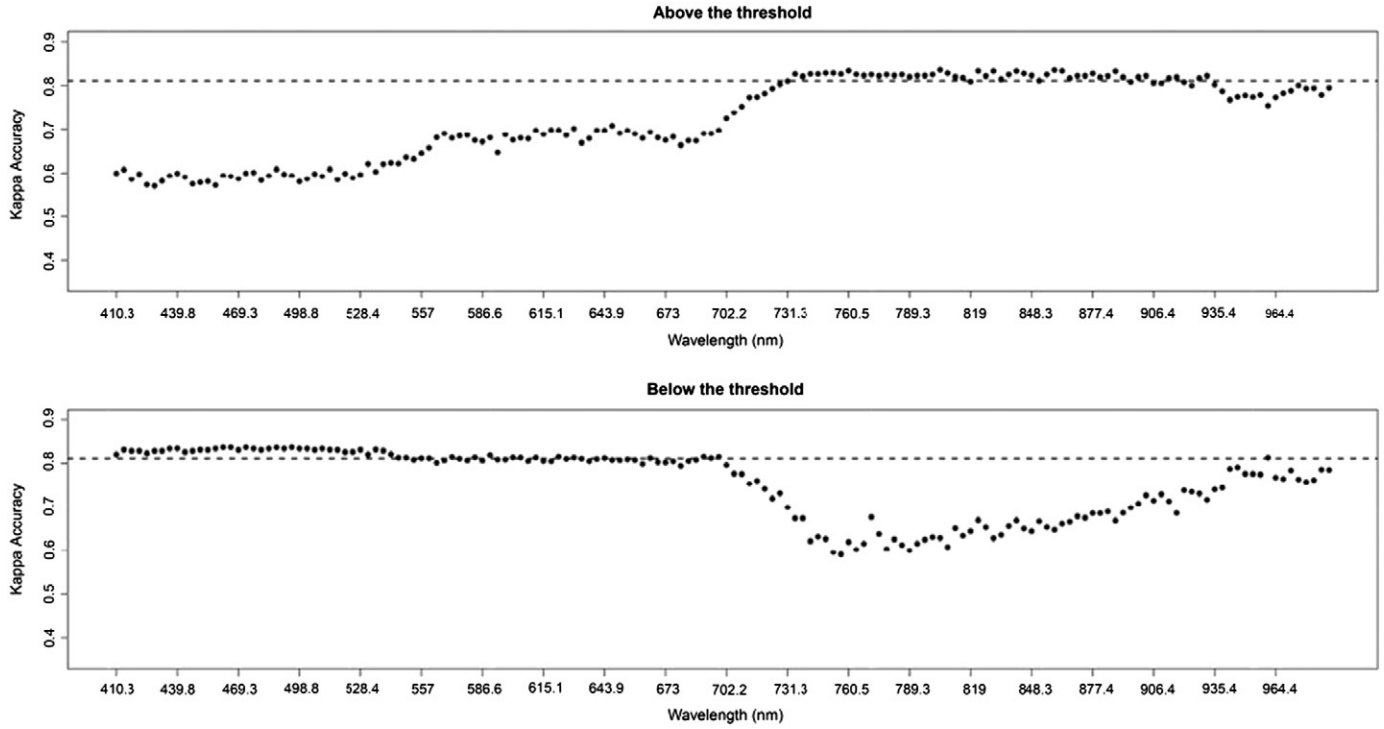

Fig. 4. KA obtained with the OTM thresholding on each hyperspectral band in case M–H. The dashed line represents the KA obtained without the use of any threshold. In the upper graph only the pixels whose values were "above" the lower bound threshold were considered in the aggregation, while in the lower graph only the pixels whose values were "below" the upper bound threshold were considered in the aggregation.

least two pixels representing the narrowest objects. According to this, the spatial resolution of the imagery used for the delineation should match the expected minimum size of the tree crowns across the scene. With hyperspectral data this is not always possible since the spatial resolution of the imagery depends on many factors, like the specific characteristics of the sensor and the flying altitude. On the other hand, high density ALS data can produce better detection rates because it allows creating high-resolution CHMs (e.g., at 0.10–0.20 m resolution). For instance, the smallest crown area measured in the field (0.6 m^2) would correspond to only 4 pixels in the hyperspectral imagery and to approximately 15 pixels in the CHM.

Both the H-ITC and L-ITC delineations detected a low number of trees, due to the fact that suppressed trees hardly can be seen in the remotely sensed data. This is a common problem using ITC delineation where mainly the dominant and co-dominant trees in the higher levels of the canopy can be detected (Falkowski et al., 2008). The species

Table 5
Results obtained at ITC level by applying the best thresholding methods of the semi-automatic case at the automatic cases.

Case	Threshold	OA (%)	KA	PA (%)			UA (%)		
				Pine	Spruce	Broadleaves	Pine	Spruce	Broadleaves
H–H	–	87.9	0.786	97.3	92.4	30.4	83.4	92.9	77.7
	30th percentile 469 nm band	89.0	0.806	96.2	93.2	40.6	85.8	93.5	73.7
	45th percentile CHM	88.2	0.791	97.7	92.9	29.0	83.4	94.0	71.4
	Otsu 462 nm band	89.0	0.806	96.6	92.9	40.6	85.5	93.8	73.7
	Otsu Mean blue bands	89.1	0.809	95.1	93.2	43.5	86.3	93.7	71.4
	Mean 810 nm (Clark et al., 2005)	89.1	0.808	98.0	93.2	34.7	84.6	94.3	77.4
	Mean NDVI	87.2	0.771	97.0	93.2	20.3	81.5	93.5	70.0
L–L	–	86.3	0.759	93.8	92.4	33.3	81.4	91.2	79.6
	20th percentile 432 nm band	87.4	0.778	89.5	94.0	50.4	88.3	87.5	81.3
	70th percentile CHM	86.7	0.766	91.3	94.3	37.2	84.9	88.9	78.7
	Otsu 425 nm band	87.8	0.785	92.7	94.0	43.4	85.7	89.9	83.6
	Otsu Mean blue bands	87.7	0.784	92.4	93.5	45.7	86.5	89.6	79.7
	Mean 810 nm (Clark et al., 2005)	87.7	0.783	91.5	94.1	45.7	86.9	88.8	83.1
	Mean NDVI	86.5	0.762	92.2	93.3	36.4	82.3	90.8	78.3

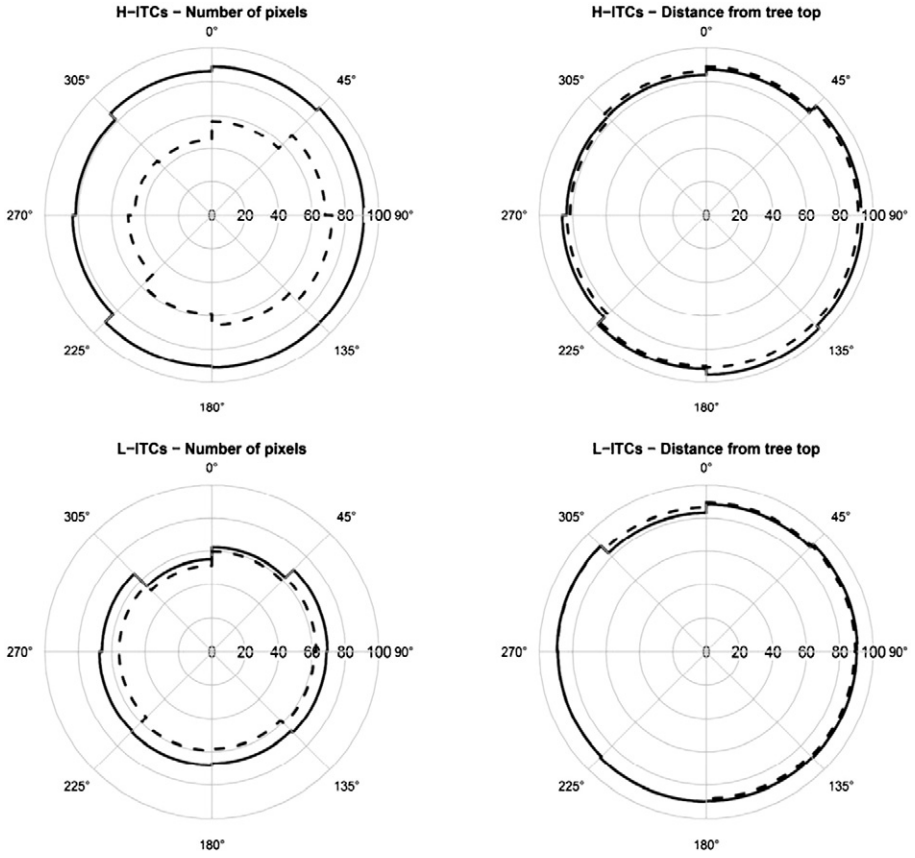

Fig. 5. Normalized number of pixels and normalized distance of the pixels from the tree position in the H-ITCs and L-ITCs for the pixels used in the classification after the mean IR thresholding (dashed line) and the mean blue band OTM thresholding.

composition produced using H-ITC and L-ITC was quite similar, but in both situations Spruce and Broadleaves classes were underrepresented. In general, pines are much less shadow tolerant than spruce, and thus they grow usually less clustered and are less often suppressed by taller trees. Moreover, in Norway, pines usually grow on poor sites forming stands with low stem densities, while spruce trees tend to grow on richer sites and at higher stem densities. All together, these factors improved the detection rate for pine compared to the other species.

Regarding the size distribution of the ITCs, our study indicated that the L-ITCs had in general wider range of ground covered than the ones delineated from the hyperspectral data (see Table 2). This might be explained by several factors. First, the hyperspectral imagery does not perform well in detecting the crown borders, especially in the overlapping areas between neighboring trees. This can produce the effect to have several trees within the same ITC (Leckie et al., 2003). Another factor that influenced the ITC size distribution was the spatial resolution of the data used for the ITC delineation. A coarser spatial resolution (larger pixel sizes), besides reducing the detection rate, could increase the average ITC area as trees smaller than the pixel size are not detected. Coops et al. (2004) compared ITC detected by multispectral and ALS data

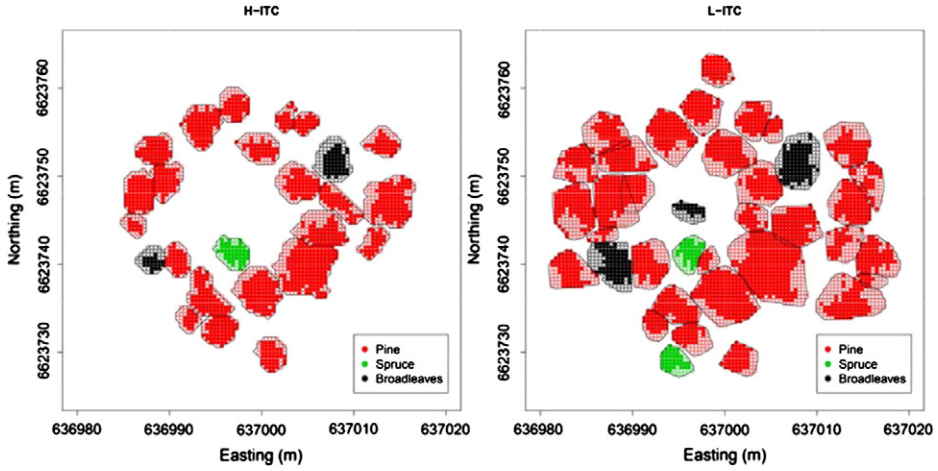

Fig. 6. Example of location of the pixels selected by the mean blue band OTM thresholding for H-ITC and L-ITC. The selected pixels are the filled circles.

and they obtained slightly different results than the ones we found. In particular, they underlined that the multispectral data overestimated the ITC size, while ALS data provided slightly smaller ITC as compared to the field-based estimates. However, this study and the one of Coops et al. (2004) differ mainly for two factors: i) the ALS data used in their study had a much lower point density compared to our data (1 point per square meter versus 7 points). A higher point density provides a more detailed CHM; and ii) they used multispectral and not hyperspectral data.

H-ITCs allow obtaining a higher classification accuracy compared to L-ITCs. This result is based on separate classifications using the two sets of ITCs separately. Considering only the common ITCs (586 trees) the difference in classification was not significant. It is worth noting that it is obvious that if a tree was detected only in ALS data, when that ITC was overlaid over hyperspectral data it provided probably not informative pixels, like shadowed ones. A clear example is the tree found in the center of Fig. 2. The better classification performances of H-ITCs compared to L-ITCs can also be due to the absence of co-registration errors, which can be present when using the L-ITCs. Moreover, the smaller size of the H-ITCs (Table 2) may have reduced the presence of noisy pixels at the ITCs' borders. This is confirmed by the analyses regarding the thresholding methods (see Fig. 5) which indicated that a lower number of pixels were removed by thresholding the H-ITCs, compared to L-ITCs. An important result was that for the classes Pine and Spruce, the accuracies were always high, reaching PAs and UAs over 90% after thresholding. These two main species represent the most valuable source of timber in Norway, and approximately 99% of the harvested volume in Norway is coming from these species (Statistics Norway, 2008). However, a better detection/classification of the broadleaves species can be important for e.g. ecological studies, and the H-ITCs seemed to be more suitable for the separation of the Broadleaves class compared to the ALS ones. In fact, it was possible to find differences in PAs between hyperspectral and ALS ITC classifications of more than 10 percentage points (Table 4).

All the five classification cases presented in this study (M–M, H–H, L–L, M–H and M–L) are interesting from both a researcher's and a practitioner's perspective. The M–M case is useful for defining a baseline of the classification results. Many papers base their analyses only on manually delineated ITCs (Clark et al., 2005; Dalponte et al., 2013; Puttonen et al., 2009). The H–H and L–L cases are fully automated and might be applicable in operational forest inventories. Providing the remote sensing data (hyperspectral and/or ALS) and the position of the ground measured trees is sufficient to obtain the classification map of the area at both pixel and ITC levels. Moreover, since in cases H–H and L–L the thresholding was applied prior the training of the SVM, it improved the training speed, as it reduced the number of training samples. The cases M–H and M–L are semi-automatic ones, combining the possibility of having a high classification accuracy (due to the training based on M-ITCs), with the possibility of covering large areas (due to the automatic ITC delineation). The semi-automatic cases provided classification accuracies very close to the ones obtained with the manual case (case M–M), especially for the Broadleaves class which was heavily penalized by the automatic cases (H–H and L–L). In the considered dataset Broadleaves was a minority class (all the plots were in coniferous dominated stands), thus differently than Pine or Spruce classes, where sample trees were mainly surrounded by trees of the same species, the Broadleaves trees were mainly surrounded by Pine and Spruce. In case of dense forest plots this creates spurious spectral signatures in the border pixels, and this can be the reasons why the thresholding had more effect on this class. Another important reason was that the average height of Broadleaves trees was lower than that of Pine and Spruce. Small trees are more subject to shadowing effects.

Another important result was that applying the pixel thresholding on bands in the blue spectrum (which is related to the shadows in the vegetation areas) was generally more effective and more robust compared to other regions of the spectrum (in terms of the fact that different threshold values applied to all the blue area bands provided similar results). It was noticed that in shadowed areas the average values of the pixels above the threshold in the blue-bands of the spectrum were higher than those in the green one. This general phenomenon was probably accentuated by the normalization process applied to our data that increased the variability in the visible part of the spectrum, and reduced it in the infrared. Assuming that the ITCs contain only vegetation pixels, this may indicate that either the removed pixels did not represent vegetation, or that they were shadowed. The results obtained with the NDVI thresholding clearly showed that the problem was not the presence of non-vegetation, but of shadows, as the effect of a threshold based on NDVI values was not significant. Thresholding in the blue band provided very similar results to the ones in the infrared, with the difference that in the blue area almost all the bands can be used to threshold, while in the infrared the results obtained were varying band by band (Figs. 3 and 4). The main innovation given by the OTM method compared to the ones already presented in the literature (e.g., Ardila et al., 2012; Clark et al., 2005) is that it is fully automatic and it lacks subjectivity in the selection of the threshold. Moreover, comparing the thresholds in the blue band computed with OTM, with the mean value in the same band in each ITC, we noticed that the values were significantly different at 95% confidence level (Kruskal–Wallis test).

Considering the threshold values of the OTM thresholding over the mean blue band, it emerged that the thresholds used for Spruce ITCs were higher than those for Broadleaves and Pines. In general, the shape of the ITCs of Spruce trees is different compared to Pines and Broadleaves.

Thresholding CHM heights did not provide significant improvements in the classification accuracy. This might be related to the different geometry of the two datasets and misalignment problems (e.g., nearly one pixel shift between tree positions in the hyperspectral and ALS data; see Table 2). Moreover, the height thresholding does not consider other important factors that influenced the spectral signature, such as the sun position or the influence of surrounding vegetation. Some studies in the literature report different results and they showed that a threshold based on the CHM values inside the crown improves the classification process (e.g., Heinzel & Koch, 2012).

By studying the pixels excluded from the classification after being removed by the thresholding it was possible to notice that they were mainly located in the north-western part of the ITCs, for either PTM, OTM, or mean IR thresholding. This is most probably due to the sun position at the moment of the image acquisition (south-east azimuth). This concurs with the conclusion of using the sunlit pixels in tree species classification pointed out in previous studies (Clark et al., 2005). Furthermore, the excluded pixels were not distributed uniformly along the ITC border, thus it is not possible to apply simply a shrinking of the ITC size, or a CHM thresholding. The tree size emerged also as an important factor for having a correct classification of the tree species. In general, correctly classified trees have a larger DBH and a larger crown area than incorrectly classified trees. This can be due to the fact that larger trees are easy to automatically delineate in hyperspectral data and also that the effect of misalignments on the classification is lower for trees with larger crowns.

6. Conclusion

The aim of this research was to analyze the effects of i) the use of hyperspectral and ALS data in the automatic ITC delineation, and ii) applying various thresholding methods to the pixels inside each ITC, on tree species classification accuracy. Regarding the first objective, it has been shown that ITC delineation on CHM derived from ALS data provided a higher tree detection rate than that on hyperspectral bands. L-ITCs covered also a wider range of sizes respect to H-ITCs. On the contrary, H-ITCs provided a better matching with hyperspectral data used in the classification phase. On the ITCs found by both delineations no significant difference in classification accuracy was found. Regarding the second objective of the paper, a proper thresholding of the pixels inside the ITCs emerged to be important to achieve a

good classification accuracy. Not all the thresholding methods tested provided significant improvements. The proposed methodology based on Otsu thresholding on the mean blue bands achieved accuracies significantly different than the ones achieved without any thresholding.

As a final remark, in a practical case of tree species mapping at ITC level the most suitable approach, that emerged from this study, is to use manually delineated ITCs for the training of the classifier and ALS delineated ITCs (with mean IR or mean blue bands OTM thresholding) for the species mapping. In this way it can be possible to obtain both a high classification accuracy, and a high tree detection rate. This is a feasible approach in a remote sensing based forest inventory, as ALS data are usually acquired for stem volume estimation, and the manual delineation of a small number of ITCs to use in the training of the classifier can take only few hours of work.

Acknowledgments

The research leading to these results received funding from the "SpectralLIDAR" project, the Flexwood ("Flexible Wood Supply Chain") project, and the GHG Europe project. "SpectralLIDAR" is funded by the Autonomous Province of Trento (Italy) and by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement Marie Curie 7th Framework Program — PCOFUND-GA-2008-226070, acronym "progetto Trentino". The "Flexible Wood Supply Chain" project received funding from the [European Union] [European Atomic Energy Community] Seventh Framework Programme ([FP7/2007-2013] [FP7/2007-2011]) under grant agreement no [245136]. GHG Europe is founded by the EU (EU contract no. 244122). We wish to thank Terratec AS, Norway for providing and processing the hyperspectral data and Blom Geomatics, Norway for providing and processing the ALS data. We also wish to thank Dr. Knut Marius Hauglin and Mr. Vegard Lien, both at the Norwegian University of Life Sciences, who were responsible for the fieldwork.

References

Ardila, J. P., Bijker, W., Tolpekin, V. a, & Stein, A. (2012). Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. *International Journal of Applied Earth Observation and Geoinformation*, *15*, 57–69. http://dx.doi.org/10.1016/j.jag.2011.06.005.

Bernstein, L. S., Sundberg, R. L., Levine, R. Y., Perkins, T. C., & Berk, A. (2005). A new method for atmospheric correction and aerosol optical property retrieval for VIS-SWIR multi- and hyperspectral imaging sensors: QUAC (QUick Atmospheric Correction). *IEEE IGARSS*, *Vol. 00*. (pp. 3549–3552).

Brandtberg, T. (2007). Classifying individual tree species under leaf-off and leaf-on conditions using airborne LiDAR. *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(5), 325–340. http://dx.doi.org/10.1016/j.isprsjprs.2006.10.006.

Bunting, P., & Lucas, R. (2006). The delineation of tree crowns in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data. *Remote Sensing of Environment*, *101*(2), 230–248. http://dx.doi.org/10.1016/j.rse.2005.12.015.

Clark, M., Roberts, D., & Clark, D. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, *96*(3–4), 375–398. http://dx.doi.org/10.1016/j.rse.2005.03.009.

Congalton, R. G., & Green, K. (2008). *Assessing the accuracy of remotely sensed data: Principles and practices* (2nd ed.): Taylor & Francis.

Coops, N. C., Wulder, M. a, Culvenor, D. S., & St-Onge, B. (2004). Comparison of forest attributes extracted from fine spatial resolution multispectral and LiDAR data. *Canadian Journal of Remote Sensing*, *30*(6), 855–866. http://dx.doi.org/10.5589/m04-045.

Dalponte, M., Bruzzone, L., & Gianelle, D. (2012). Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment*, *123*, 258–270. http://dx.doi.org/10.1016/j.rse.2012.03.013.

Dalponte, M., Bruzzone, L., Vescovo, L., & Gianelle, D. (2009). The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment*, *113*(11), 2345–2355. http://dx.doi.org/10.1016/j.rse.2009.06.013.

Dalponte, M., Ørka, H. O., Gobakken, T., Gianelle, D., & Næsset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(5), 2632–2645. http://dx.doi.org/10.1109/TGRS.2012.2216272.

Ene, L., Næsset, E., & Gobakken, T. (2012). Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-based stem number estimates. *International Journal of Remote Sensing*, *33*(16), 5171–5193.

Erikson, M., & Olofsson, K. (2005). Comparison of three individual tree crown detection methods. *Machine Vision and Applications*, *16*(4), 258–265. http://dx.doi.org/10.1007/s00138-005-0180-y.

Falkowski, M. J., Smith, A.M. S., Gessler, P. E., Hudak, A. T., Vierling, L. a, & Evans, J. S. (2008). The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using LiDAR data. *Canadian Journal of Remote Sensing*, *34*(S2), S338–S350. http://dx.doi.org/10.5589/m08-055.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *6*, 610–621.

Harrell, J. F. E. (2012). rms: regression modeling strategies. Retrieved from. http://biostat.mc.vanderbilt.edu/rms

Heinzel, J., & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, *13*(1), 152–160. http://dx.doi.org/10.1016/j.jag.2010.09.010.

Heinzel, J., & Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 101–110. http://dx.doi.org/10.1016/j.jag.2012.01.025.

Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences*, *32*(9), 1283–1298. http://dx.doi.org/10.1016/j.cageo.2005.11.008.

Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, *90*(4), 415–423. http://dx.doi.org/10.1016/S0034-4257(03)00140-8.

Hosmer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. (1997). A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, *16*(9), 965–980 (Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9160492)

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical journal. Biometrische Zeitschrift*, *50*(3), 346–363. http://dx.doi.org/10.1002/bimj.200810425.

Hutchinson, C. F. (1982). Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing*, *48*, 123–130.

Hyyppä, J., Kelle, O., Lehikoinen, M., & Inkinen, M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(5), 969–975. http://dx.doi.org/10.1109/36.921414.

Jones, T. G., Coops, N. C., & Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment*, *114*(12), 2841–2852. http://dx.doi.org/10.1016/j.rse.2010.07.002.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., et al. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, *4*(4), 950–974. http://dx.doi.org/10.3390/rs4040950.

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab — An S4 package for kernel methods in R. *Journal of Statistical Software*, *11*(9), 1–20.

Kittler, J., & Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, *19*, 41–47.

Koetz, B., Morsdorf, F., Vanderlinden, S., Curt, T., & Allgower, B. (2008). Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management*, *256*(3), 263–271. http://dx.doi.org/10.1016/j.foreco.2008.04.025.

Korpela, I., Heikkinen, V., Honkavaara, E., Rohrbach, F., & Tokola, T. (2011). Variation and directional anisotropy of reflectance at the crown scale — Implications for tree species classification in digital aerial images. *Remote Sensing of Environment*, *115*(8), 2062–2074. http://dx.doi.org/10.1016/j.rse.2011.04.008.

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., & Shreenan, R. (2003). Combined high-density LiDAR and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing*, *29*(5), 633–649.

Leckie, D.G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F. a, Cloney, E., et al. (2005). Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing*, *31*(2), 175–190. http://dx.doi.org/10.5589/m05-004.

Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M. (2012). A new method for segmenting individual trees from the LiDAR point cloud. *Photogrammetric Engineering and Remote Sensing*, *78*(1), 75–84.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(8), 1778–1790. http://dx.doi.org/10.1109/TGRS.2004.831865.

Næsset, E. (1997a). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*(2), 49–56. http://dx.doi.org/10.1016/S0924-2716(97)83000-6.

Næsset, E. (1997b). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, *61*(2), 246–253. http://dx.doi.org/10.1016/S0034-4257(97)00041-2.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A (General)*, *135*(3), 370–384.

Ørka, H. O., Næsset, E., & Bollandsås, O. M. (2009). Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment*, *113*(6), 1163–1174. http://dx.doi.org/10.1016/j.rse.2009.02.002.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics Part B, Cybernetics: a publication of the IEEE Systems, Man, and Cybernetics Society*, *9*(1), 62–66.

Popescu, S.C., & Wynne, R. H. (2004). Seeing the trees in the forest: Using LiDAR and multispectral data fusion with local filtering and variable window size

for estimating tree height. *Photogrammetric Engineering and Remote Sensing, 70*(5), 589–604.

Puttonen, E., Litkey, P., & Hyyppä, J. (2009). Individual tree species classification by illuminated—Shaded area separation. *Remote Sensing, 2*(1), 19–35. http://dx.doi.org/10.3390/rs2010019.

R Development Core Team (2008). *R: A language and environment for statistical computing. R foundation for statistical computing.* R Foundation for Statistical Computing (Retrieved from http://www.r-project.org)

Rahman, M. Z. A., Gorte, B. G. H., & Bucksch, A. K. (2009). A new method for individual tree measurement from airborne LiDAR. *Silvilaser 2009, October 14–16, 2009* (pp. 1–10). Texas, USA: College Station.

Rossmann, J., & Buecken, A. (2010). A comparison of single tree delineation based on LiDAR and optical data. *2010 RSPSoc and Irish Earth Observation Symposium "Visualising the World: From the sea-bed to the cloud-tops", 1st–3rd September, Cork, Ireland*.

Solberg, S., Næsset, E., & Bollandsås, O. M. (2006). Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing, 72*(12), 1369–1378.

Statistics Norway (2008). *Skogstatistikk 2008.* Oslo: Statistics Norway (Retrieved from http://www.ssb.no/nos_skogstat/nos_d430/nos_d430.pdf)

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., et al. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, *85*(1), 27–40. http://dx.doi.org/10.1093/forestry/cpr051.

Wang, L., Gong, P., & Biging, G. S. (2004). Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing, 3114*(3), 351–358.

Wang, Y., Weinacker, H., & Koch, B. (2008). A LiDAR point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. *Sensors*, *8*(6), 3938–3951. http://dx.doi.org/10.3390/s8063938.

Yan, G., Maathuis, B. H. P., Xiangmin, Z., & Dijk, P.M. Van (2007). Comparison of pixel based and object oriented image classification approaches — A case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*, 37–41 (June 2012).

Yu, B., Ostland, I. M., Gong, P., & Pu, R. (1999). Penalized discriminant analysis of in situ hyperspectral data for conifer species recognition. *IEEE Transactions on Geoscience and Remote Sensing, 37*(5), 2569–2577.