Remote Sensing of Environment 121 (2012) 210–223

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

3-D mapping of a multi-layered Mediterranean forest using ALS data

António Ferraz [a,b,c,*], Frédéric Bretar [d], Stéphane Jacquemoud [b], Gil Gonçalves [c,e], Luisa Pereira [f], Margarida Tomé [g], Paula Soares [g]

[a] Université Paris Est, IGN, Laboratoire MATIS, 73 avenue de Paris, 94165 Saint-Mandé Cedex, France
[b] Institut de physique du globe de Paris - Sorbonne Paris Cité, Université Paris Diderot, UMR CNRS 7154, Case 7011, 35 rue Hélène Brion, 75013 Paris, France
[c] Instituto de Engenharia de Sistemas e Computadores de Coimbra, rua Antero de Quental, no 199, 3000-033 Coimbra, Portugal
[d] CETE Normandie Centre, Laboratoire des Ponts et Chaussées, 10 chemin de la Poudrière, BP 245, 76121 Grand Quevilly, France
[e] Departamento de Matemática, Universidade de Coimbra, Apartado 3008, 3001-454 Coimbra, Portugal
[f] Universidade de Aveiro, Escola Superior de Tecnologia e Gestão de Águeda, apartado 473, 3754-909 Águeda, Portugal
[g] Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Centro de Estudos Florestais, Tapada da Ajuda, 1349-017 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 11 July 2011
Received in revised form 18 January 2012
Accepted 21 January 2012
Available online xxxx

Keywords:
Airborne laser scanning
LiDAR
Multi-layered forest
Unsupervised segmentation
Mean shift algorithm
Fuel mapping
Vertical stratification
Tree crown
3-D mapping
Ground vegetation
Understory
Overstory

ABSTRACT

This study presents a robust approach for characterization of multi-layered forests using airborne laser scanning (ALS) data. Fuel mapping or biomass estimation requires knowing the diversity and boundaries of the forest patches, as well as their spatial pattern. This includes the thickness of the main vegetation layers, but also the spatial arrangement and size of the individual plants that compose each stratum. In order to decompose the ALS point cloud into genuine 3-D segments corresponding to individual vegetation features, such as shrubs or tree crowns, we apply a statistical approach based on the mean shift algorithm. The segments are progressively assigned to a forest layer: ground vegetation, understory or overstory. Our method relies on a single biophysically meaningful parameter, the kernel bandwidth, which is related to the local forest structure. It is validated on 44 plots of a Portuguese forest, composed mainly of eucalyptus (*Eucalyptus globulus* Labill.) and maritime pine (*Pinus pinaster* Ait.) trees. The number of detected trees varies with the dominance position: from 98.6% for the dominant trees to 12.8% for the suppressed trees. Linear regression models explain up to 70% of the variability associated with ground vegetation and understory height.

1. Introduction

Forests, woodlands, and shrub formations are very important ecosystems because they provide foundations for life on Earth through their ecological functions: regulation of climate and water, habitat for animals, and supply of food and goods. They exhibit various canopy structures, from homogeneous to heterogeneous, and from single- to multi-layered (Landsberg & Gower, 1997). Today we know the horizontal structure that describes the patchiness in forest stands better than the vertical structure, which is difficult to quantify and yet is an important characteristic (Hall et al., 2011). The canopy layers (overstory, understory, and ground vegetation) are distinct from each other in their density, thickness, and water content. A better appraisal of this vertical arrangement, at high spatial resolution, would be valuable for many applications in forestry (Ares et al., 2010), carbon cycle studies (Moore et al., 2007), and ecology (Brokaw & Lent, 2000; Camprodon & Brotons, 2006). As an example, foresters use fuel models for predicting fire behavior (Pyne et al., 1996), and fire behavior models, such as FARSITE (Finney, 2004) or BehavePlus (Andrews et al., 2005), require information about vegetation strata thickness to detect areas where fire easily propagates and spreads (Anderson, 1982; Sandberg et al., 2001).

Airborne laser scanning (ALS) is an active remote sensing technique that provides georeferenced distance measurements between a remote sensing platform and the surface (Mallet & Bretar, 2009; Shan & Toth, 2009). In recent years, it has been applied over natural landscapes to extract terrain elevation (Bretar & Chehata, 2010; Kraus & Pfeifer, 1998), classify land cover (Antonarakis et al., 2008; Asner et al., 2008; Breidenbach et al., 2010; Hyyppä et al., 2008; Yoon et al., 2008), evaluate wildlife habitat (Clawges et al., 2008; Martinuzzi et al., 2009), estimate biomass (Asner et al., 2010; García et al., 2010; Zhao et al., 2009), and assess fuel characteristics (Andersen et al., 2005; Hollaus et al., 2006; Mutlu et al., 2008; Riaño et al., 2003). Depending on the nature of the target, a single pulse emission may induce one or

* Corresponding author at: Institut Géographique National, Laboratoire MATIS, 73 avenue de Paris, 94160 Saint Mandé, France.
E-mail address: antonio.ferraz@ign.fr (A. Ferraz).

0034-4257/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2012.01.020

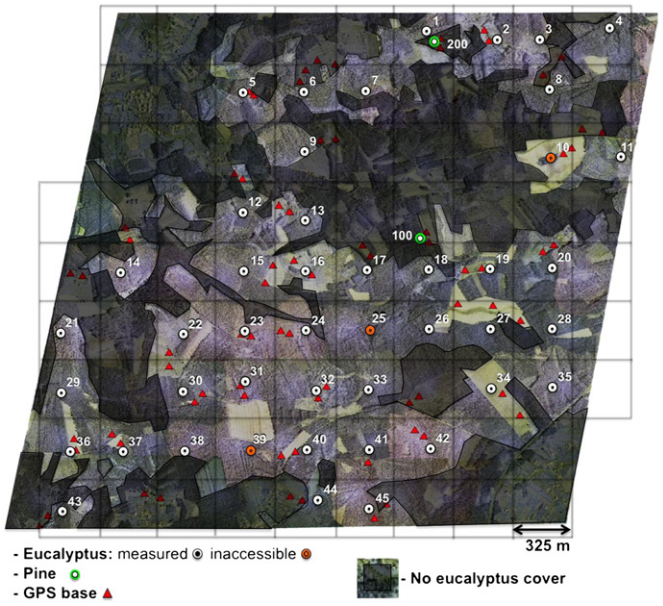

Fig. 1. Regular grid superimposed on the land cover map of the study area.

several backscattered echoes. As the laser beam penetrates down into the forest canopy layers, an unstructured 3-D point cloud that is a discrete model of the target is obtained. There are two main spatial scales for tackling the extraction of forest parameters from ALS data: at the plot scale, the biophysical variables are averaged over an area encompassing several trees (e.g. mean canopy height, biomass, stem density, leaf area index), while at the individual scale, they are estimated for a single tree (e.g. tree height, crown diameter, crown base height).

Vertical stratification has been assessed at the plot scale (Maltamo et al., 2005; Riaño et al., 2003, 2004; Zimble et al., 2003). Morsdorf et al. (2010) use the ALS intensity to discriminate different vegetation strata. They apply a supervised cluster analysis, assuming that some species have a better light reflection ratio than others. This method works fairly well in forest ecosystems made of mono-species strata. The intensity is somewhat difficult to analyze because it depends on the sensor as well as on the geometry, orientation, and optical properties of the target (leaves, branches, trunks). Some authors delineate vegetation strata by fitting continuous probability distributions, like the Weibull distribution or mixture models, to the ALS density profiles (Coops et al., 2007; Dean et al., 2009; Jaskierniak et al., 2010; Maltamo et al., 2004). However, plot-based methods are not the most appropriate means to describe the vertical stratification of complex ecosystems, such as Mediterranean forests that are characterized by an open dominant canopy and a lush undergrowth made of herbaceous and woody plants (Di Castri, 1981). These are often highly fragmented forests, the stratification of which varies locally due to small ownerships administered according to different management rules (EEA, 2008).

So far, single-tree based methods rely on a canopy height model (CHM), which is an oversimplified representation of reality in vertically heterogeneous canopies (Hyyppä et al., 2004; Morsdorf et al., 2004; Persson et al., 2004; Popescu & Wynne, 2004; Solberg et al., 2006). In order to investigate the spatial pattern of dominated trees, some authors developed multi-stage approaches. For instance, Richardson and Moskal (2011) first delineate groups of trees in the CHM and then calculate the number of trees by fitting a statistical relationship to the corresponding point cloud distribution. Reitberger et al. (2009) identify the taller trees within each group, determine the stem position, and apply a normalization-cut segmentation method to extract the smaller ones. Despite good performance, these approaches are site-dependent because they require several empirical parameters. Moreover, they do not properly address the issue of vertical stratification in multi-layered forests because, even if they delineate the topmost tree crowns, many ALS points corresponding to ground or understory vegetation remain unassigned.

Therefore, it seems that an approach that simultaneously segments vertical and horizontal structures of forest canopies is lacking. This paper validates a segmentation method based on the mean shift algorithm. This method has been tested on a 3-D point cloud acquired with a small-footprint ALS in a multi-layered Mediterranean forest. We first present the experimental data and the algorithm. The segmentation of the forest into different strata and the derivation of the geometry of individual vegetation features are then detailed.

Table 1
Biophysical characteristics of stand #30.

Height class (m)	Species	% Dominance	Mean height (m)	% Cover
0–2	Ferns	95	1.2	50
	Ulex	5		
2–8	*Acacia*	70	6.0	8
	Pinus	30		
>8	*Eucalyptus*	100	21.2	20

Table 2
Field inventory of ground vegetation and understory, all stands.

	Mean height (m)		% Cover	
	Ground vegetation	Understory	Ground vegetation	Understory
Minimum	0.15	0	2	0
Maximum	1.3	6	100	95
Mean	0.53	2.41	52.1	15.6
Standard deviation	0.3	1.64	33	20.2

2. Experiment

2.1. Study area

The study area is located near the city of Águeda in northwest Portugal (40°36′ N, 8°25′ W). It covers 9 km^2 and its altitude varies

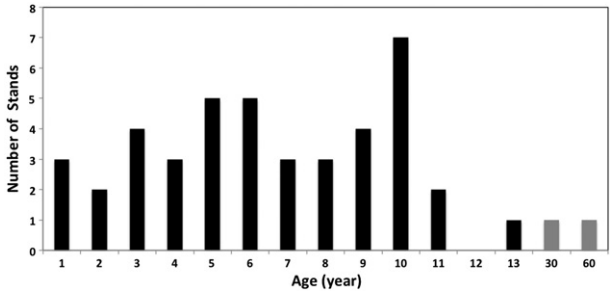

Fig. 2. Age class of the stands. The black bars correspond to the eucalyptus and the gray bars to the pines.

Table 3
Field inventory statistics for trees in mature eucalyptus and pine stands.

		DBH (cm)	CBH (m)	Total height (m)	Crown depth (m)	Atypical shape (%)
Eucalyptus	Minimum	1.5	2.5	3.7	0.4	17.2
	Maximum	70.0	22.5	35.4	14.2	
	Mean	9.7	9.4	13.3	3.9	
	Standard deviation	5.3	3.9	5.2	2.4	
Pine	Minimum	4.9	7.1	8	0.4	7.7
	Maximum	41.4	22.3	25	9.0	
	Mean	23.5	13.7	17.5	3.8	
	Standard deviation	8.3	4.8	5.0	1.5	

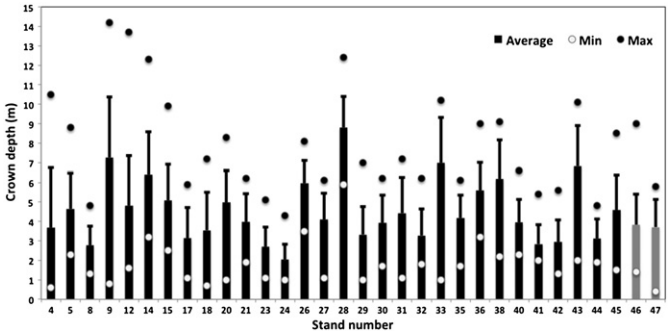

Fig. 3. Mean crown depth of dominant, codominant and dominated eucalyptus (black bars) and pine (gray bars) trees per stand. Standard deviations are plotted above each bar and the dots represent the minimum and maximum values.

from 70 m to 220 m, with slopes ranging from 2.5% to 34.2%. The landscape is predominantly composed of woodlands dominated by eucalyptus (*Eucalyptus globulus* Labill.) with some stands of maritime pine (*Pinus pinaster* Ait.). Shrublands are also present, as well as agricultural fields. The eucalypts grow in pure and mixed stands, the management of which is mainly done by 3–4 short rotations of about 10–12 years to supply raw materials to the Portuguese pulp and paper industry. Despite a limited spatial extension, the study area displays various kinds of stands and trees in terms of age and canopy structure. The lower strata are composed mainly of suppressed trees (eucalyptus, pine, acacia, and oak), gorse bush (*Ulex* spp.), heath (*Erica* spp., *Pterospartum* spp.), ferns, and herbaceous plants.

2.2. Field data collection

The forest inventory was performed in the framework of a Portuguese research project, in accordance with a field protocol recommended by the Portuguese National Forest Inventory (AFN, 2009). The superimposition of a 325 m × 325 m regularly spaced grid on a land cover map (DGRF, 2005) led to the selection of 45 plots covered mainly by eucalypts and 2 plots covered mainly by pines (plots #100 and #200, Fig. 1). The coordinates of the plot centers correspond to the grid cell centers: they were staked out in the field using GPS or, when the signal was too weak, traditional terrestrial surveying techniques. If the plot center was inaccessible due to dense shrubby vegetation, it was shifted to one of the eight points located at a distance of 50 m in all cardinal and intercardinal directions. Three eucalyptus plots could not be sampled. Each plot actually consists of two concentric circles, an outer (400 m^2) and an inner (200 m^2) circle, hereafter called plot and subplot. They were delimited using a decameter and the trees were numbered using a marker pen. The field operators defined different forest stands, i.e., uniform plant communities in terms of species, age, and spatial arrangement (Stokes et al., 1989). We use stand and substand to designate the forest stands corresponding to plot and subplot. If a plot contained more than one stand, only the stand coincident with the plot center was described. Among the forest biophysical variables measured during the field work, the vertical structure (at the stand level) and the size and shape of individual trees (at the substand level) were carefully investigated (Pereira et al., 2009).

The vertical structure of a stand was described by seven height classes (0–0.6 m, 0.6–1 m, 1–2 m, 2–4 m, 4–8 m, 8–16 m and >16 m) that could be aggregated in situ to better represent the vegetation strata (Table 1). The mean height and percent cover of each stratum were

Table 4
ALS acquisition parameters.

ALS sensor	RIEGL LMS-Q560
Wavelength	1550 nm
Scan angle	45°
Pulse rate	150 kHz
Effective measurement rate	75 kHz
Beam divergence	0.5 mrad
Ground speed	46.26 m/s
Flying height	600 m
Swath width	479 m
Swath overlap	70%
Nominal distance between two lines	150 m
Footprint diameter	30 cm
Single run density	3.3 pt/m^2
Expected final point density	9.9 pt/m^2

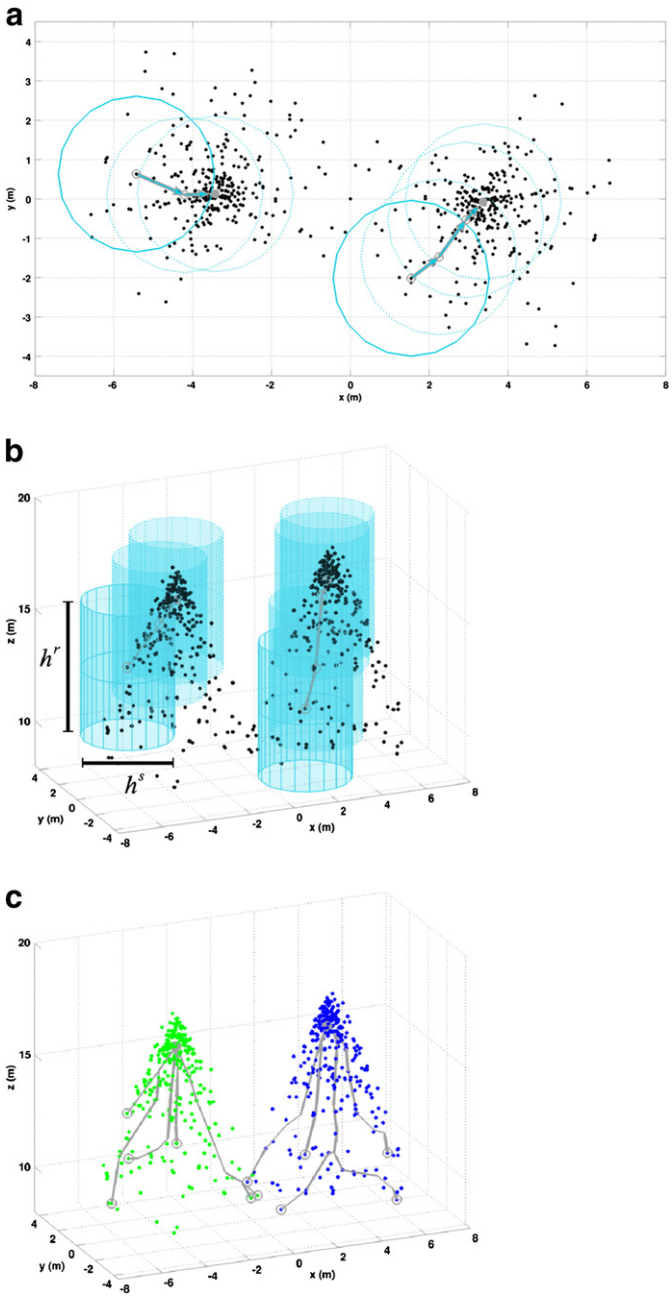

Fig. 4. Mean shift segmentation applied to (a) 2-D and (b) 3-D simulated tree crowns. The initial kernel bandwidths with different vertical and horizontal components are represented by cylinders. The mean shift vectors are represented by arrows that define the successive positions of the kernel bandwidth (dashed cylinders). All the data points that have converged to the same mode (filled gray sphere) are grouped together. The gray lines in (c) correspond to the trajectory of random points.

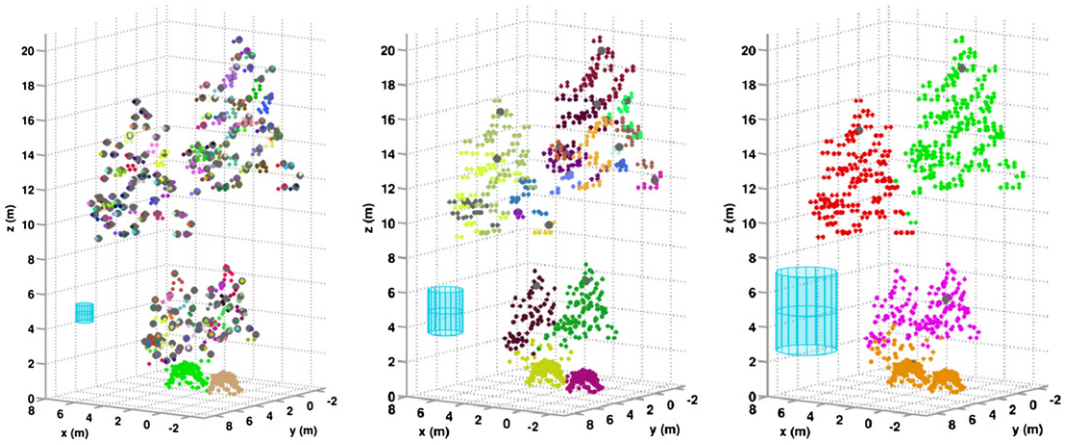

Fig. 5. Mean shift segmentation of a simulated forest scene using (left) $h^*=(1.3,2)$ m (middle) $h^*=(3.3,5)$ m and (right) $h^*=(6,9)$ m. The cylinders correspond to the different kernels and the gray spheres represent the calculated modes.

visually estimated by the field operators. Note that one tree may belong to several height classes.

All trees were assigned a class and a dominance position (dominant, codominant, dominated, and suppressed). Calipers gave a direct measurement of the diameter at breast height (DBH) whereas the total height and the crown base height (CBH) were measured using either a telescopic tape measure or a Vertex hypsometer. We only considered trees higher than 2 m with a DBH larger than 5 cm. Note that the forest inventory data are usually acquired with a lower geometric accuracy than the ALS data. To improve the accuracy, a local geodetic network made of 41 pairs of GPS-derived points was built in the same map projection as the ALS data (Fig. 1) to survey the tree positions using total stations (Gonçalves & Pereira, in press). All the data were subsequently integrated into a single three-dimensional geometry.

2.3. Characteristics of the stands

Table 2 sums up the main structural characteristics of ground vegetation and understory. The large range of percent cover indicates that the canopy varies from sparse to very dense.

The forest is highly variable in terms of tree age, architecture, and metrics. The eucalyptus stands are between 1 and 13 years old, while the two pine stands are 30 and 60 years old (Fig. 2). In total there are 12 plots with juvenile stands (1–4 years) and 32 plots with mature stands (> 4 years).

The plots contain one (59%) or more (41%) forest stands. They may display an intrinsic structural heterogeneity: the architecture of the trees differs depending on whether they grow in the middle of the forest or near roads and clearings. In open space areas, the trees tend to expand horizontally to search for light, reducing their apical

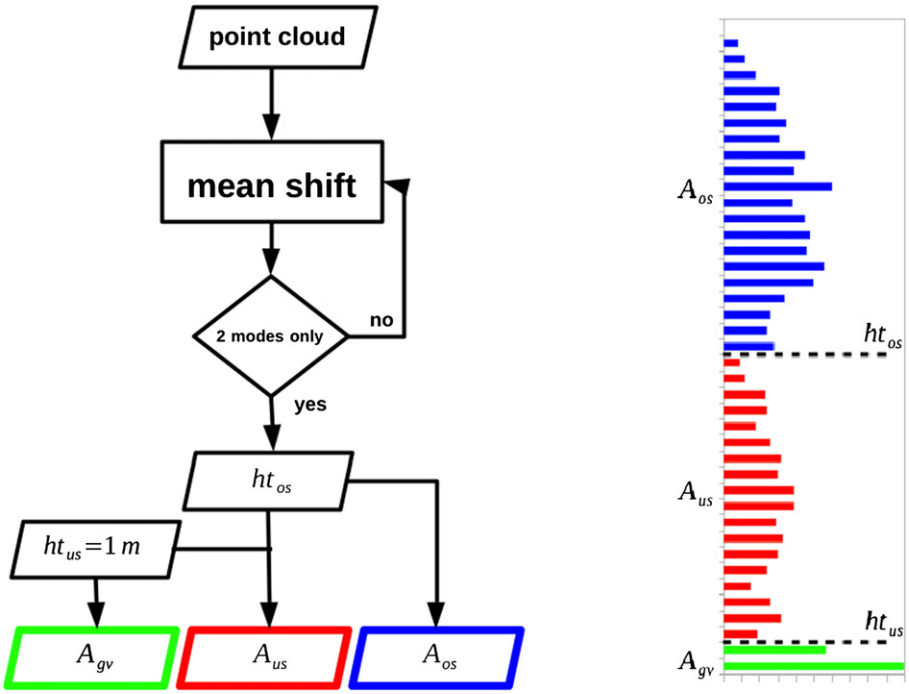

Fig. 6. Mean shift segmentation algorithm at the plot level (left) and subsequent histogram (right). ht_{us} and ht_{os} are the understory and overstory height thresholds. A_{gv}, A_{us} and A_{os} are the ground vegetation, understory, and overstory thicknesses.

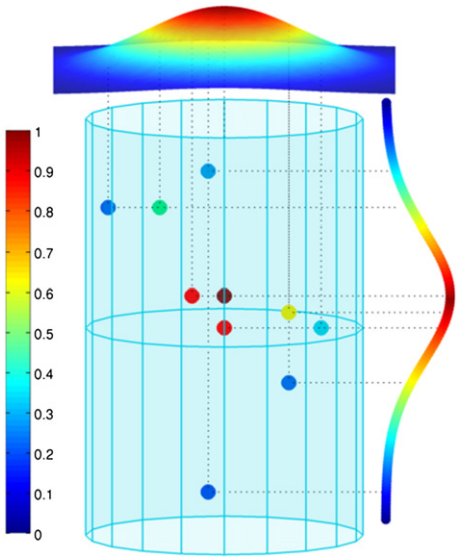

Fig. 7. Horizontal (g^s, Gaussian profile, surface) and vertical (g^r, Epanechnikov profile, curve) kernel profiles. The point and color bar indicates their weight in the calculation of the kernel barycenter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dominance. Heterogeneity also influences ground vegetation and understory since clearings let direct sunrays reach the lowest strata. About 50% of the measured stands are considered to be heterogeneous.

The stands can also be sorted according to three regeneration methods: forest planting produces the so-called high forests (eucalyptus and pine); coppicing, a traditional method of woodland management that consists in pruning trees to near the base, allows the stumps to regenerate over-vigorous coppiced trees (eucalyptus); and when after cutting, a stand contains trees that are left to grow to full height, it belongs to the category coppice-with-standards (pine). Twenty-five stands are allocated to high forest, 16 to coppice, and 3 to coppice-with-standards. Table 3 summarizes the main structural characteristics of mature eucalyptus and pine trees, as well as the percentage of trees with atypical shape: crooked, leaning, and broken trees. Specimens belonging to juvenile stands are not processed as individuals but as a forest stratum.

Fig. 3 details the crown depth in terms of minimum, maximum, mean, and standard deviation for each stand. Suppressed trees that are poorly represented in the point cloud are omitted.

2.4. Airborne laser scanning data

The data were acquired on July 14, 2008 in a full-waveform mode using a LiteMapper 5600 airborne LiDAR system (Table 4), which digitizes the waveform of the echo signal for every emitted laser pulse. The company in charge of the airborne measurements delivered both the raw and processed laser data. The digitized waveforms were converted into echo signals, each laser pulse giving rise to 1–5 ALS points (RiANALIZE, RIEGL, 2011a). The position and orientation of the platform, which are given by onboard GPS/IMU measurements, were corrected by analyzing overlapping laser strips from the calibration flight lines (TerraMatch, Burman & Soininen, 2004). These parameters, together with the GPS measurements acquired during the flight using a reference ground station, provided a point cloud in the WGS84/UTM zone 29N coordinate system for further processing (RiWORLD, RIEGL,

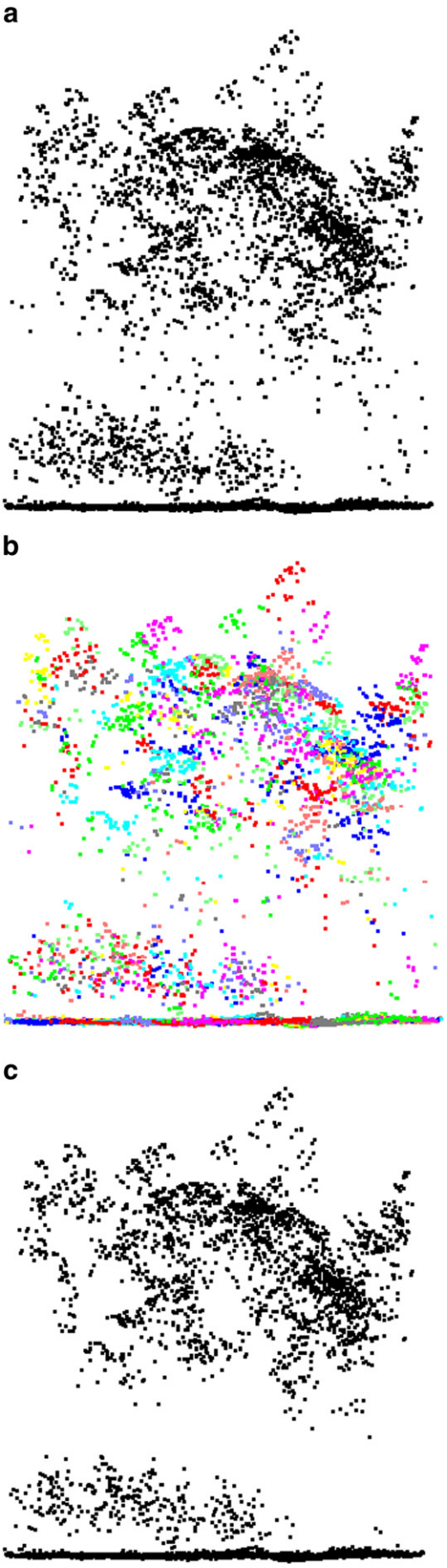

Fig. 8. (a) Original point cloud measured on plot #17; (b) MS algorithm applied using a radially symmetric kernel and a 3 m bandwidth; (c) MS segments corresponding to more than five ALS points.

2011b). Systematic height errors were finally removed by using ground control data spread over the study area.

The average point density within each plot is of 9.5 pt/m^2 (min = 4.7 pt/m^2, max = 15.5 pt/m^2, σ = 1.9 pt/m^2). To calculate the effective height of the objects in the scene, ground and vegetation points were separated (TerraScan, Soininen, 2010). A Delaunay triangulation was then generated to produce a 0.3 m × 0.3 m digital terrain model, which was used to normalize the point cloud. Note that the points filtered as ground were kept in the dataset.

3. Methodology: the mean shift algorithm

An ALS point cloud can be regarded as a multimodal distribution where each mode, here defined as a local maximum both in density and height, corresponds to a crown apex. In this study, we investigate the ability of the mean shift (MS) algorithm to extract the modes of the point cloud. Due to the complexity of the forest stands, which mix shrubs, suppressed trees, and dominant trees, a single kernel bandwidth is unsuitable. To improve the segmentation of individual vegetation features, we propose to apply a bottom-up iterative method based on an adaptive MS algorithm, which sequentially segments individual vegetation features.

3.1. Background

The mean shift has been primarily applied to image segmentation (Comaniciu & Meer, 2002). Here we explore its potential for segmenting a three-dimensional point cloud. The Parzen window (or kernel density estimation) technique is a method for estimating the probability density function (PDF) of a random variable X that is distributed in a d-dimensional space R^d. Each point X_i contributes to the PDF based on its distance from the center of the volume where the data are distributed. The estimated PDF is

$$\hat{f}_{h,K}(X) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{X-X_i}{h}\right) \tag{1}$$

where n is the number of samples of the random variable, K is the chosen kernel function, and h, called the bandwidth, is the smoothing parameter that determines the contribution of each sample. K is a non-linear function of the distance from the data points to X. We define a radially symmetric kernel that satisfies $K(X) = c_{k,d} \times k(||X||^2)$, where $c_{k,d}$ is a normalization constant, which makes K integrate to one, and k is called the kernel profile. The algorithm tries to determine local maxima of the density function $f(X)$, which correspond to the zeros of the gradient $\nabla f(X) = 0$. Assuming that g is the derivative of the kernel profile, $g(X) = -k'(X)$, and G the corresponding kernel defined by $G(X) = c_{g,d} \times g(||X||^2)$, where $c_{g,d}$ is another normalization constant, Comaniciu and Meer (2002) calculate the density gradient estimator as

$$\nabla\hat{f}_{h,K}(X) = \hat{f}_{h,G}(X)\frac{2c_{k,d}}{h^2 c_{g,d}} m_{h,G}(X) \tag{2}$$

with $m_{h,G}(X)$ the mean shift vector:

$$m_{h,G}(X) = \frac{\sum_{i=1}^{n} X_i\, g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)} - X \tag{3}$$

The mean shift is the difference between the weighted mean (G-distance), using the kernel G for weights, and X, the center of the kernel. $m_{h,G}(X)$ can be inferred from Eq. (2)

$$m_{h,G}(X) = \frac{h^2 c_{g,d}}{2c_{k,d}} \frac{\nabla\hat{f}_{h,K}(X)}{\hat{f}_{h,G}(X)}. \tag{4}$$

Eq. (4) shows that, at location X, the mean shift vector computed with kernel G is proportional to the normalized density gradient estimate obtained with kernel K. Thus, it always points toward the direction of the maximum slope of the density function. The procedure does not need to evaluate the density function $\hat{f}_{h,K}$ itself, but only the kernel profile g. In a multidimensional space, the kernel is usually split into two or more kernels. Here we separate the horizontal and vertical domains. The MS vector is then defined as

$$m_{h^*,G^*}(X) = \frac{\sum_{i=1}^{n} X_i g^s\left(\left\|\frac{X^s-X_i^s}{h^s}\right\|^2\right) g^r\left(\left\|\frac{X^r-X_i^r}{h^r}\right\|^2\right)}{\sum_{i=1}^{n} g^s\left(\left\|\frac{X^s-X_i^s}{h^s}\right\|^2\right) g^r\left(\left\|\frac{X^r-X_i^r}{h^r}\right\|^2\right)} - X \tag{5}$$

where the superscripts s and r refer to the horizontal and vertical domains. g^s and g^r are the two associated kernel profiles, h^s and h^r the two bandwidths, and X^s and X^r the two components of the vectors. At step t the iterative process can be written as

$$X^{t+1} \leftarrow X^t + m_{h^*,G^*}\left(X^t\right). \tag{6}$$

2-D and 3-D synthetic tree crowns were simulated to test the performance of the MS algorithm. Fig. 4a and b shows that points converge toward the modes. This procedure can be easily extended to a distance-based segmentation technique if all the data points that converge toward the same mode are grouped together (Fig. 4c). All the modes inscribed in a sphere with radius 1 m are considered as a single mode.

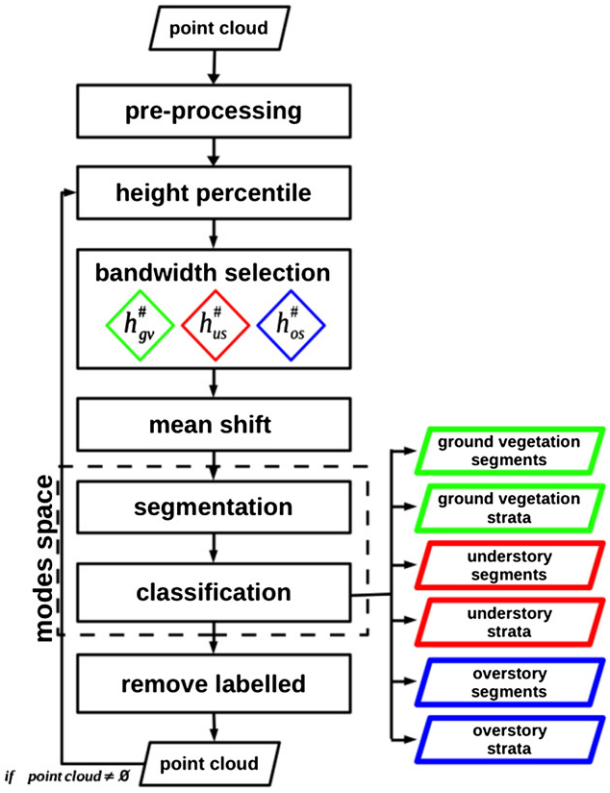

Fig. 9. Workflow of the adaptive mean shift algorithm.

3.2. Determination of the bandwidth

The choice of the kernel bandwidth is critical because it strongly impacts on the results. Setting a small value produces several distinct modes (local basins of attraction), while setting a large one aggregates small structures into larger ones (large basins of attraction). The determination of an optimal value is actually a major challenge. The thickness of the forest strata generally increases with height, i.e., scrubby vegetation is typically thinner than overstory. Three segmentations have been applied to a simulated scene using different bandwidths (Fig. 5). The smaller bandwidth that is optimal for ground vegetation tends to fragment the trees into numerous segments (Fig. 5a). Increasing the bandwidth definitely improves the segmentation of the understory without effect on the taller trees (Fig. 5b). Finally, the optimal bandwidth for the overstory causes under-segmentation of the scene (Fig. 5c). Worse yet, dense ground vegetation tends to attract a sparse understory, overestimating the thickness of this layer. Thus, using a single scale over the entire space is not suitable for the analysis of forest environments. The issue of bandwidth selection has been studied for the purpose of multiscale segmentation using either multispectral or hyperspectral images (Bo et al., 2009; Comaniciu, 2003; Huang & Zhang, 2008). Variable bandwidth MS has already been proved to converge, and even to surpass, fixed bandwidth MS (Comaniciu & Meer, 2002).

In order to properly segment individual vegetation features, a different bandwidth is assigned to each vegetation stratum. The thicker the forest layer the larger the bandwidth. Since vegetation volumes are better predicted if the stratum thickness is known, the first stage of the algorithm consists in plotting the height histograms of the forest plots in order to identify the strata: overstory, understory, and ground vegetation. A first pass of the MS algorithm is applied to the ALS point cloud to compute their basins of attraction. Eq. (5) is applied to the ALS points using the uniform kernel profile on both components:

$$g^s(X^s)=\begin{cases}1 & \text{if } \|X^s\|\le 1\\ 0 & \text{otherwise}\end{cases} \text{ and } g^r(X^r)=\begin{cases}1 & \text{if } \|X^r\|\le 1\\ 0 & \text{otherwise}\end{cases}. \quad (7)$$

Thus, in such a case, the ratio in Eq. (5) is simply the mean of the ALS points contained within a cylinder of radius h^s, height h^r, centered in X. To remove the influence of the horizontal coordinates, h^s

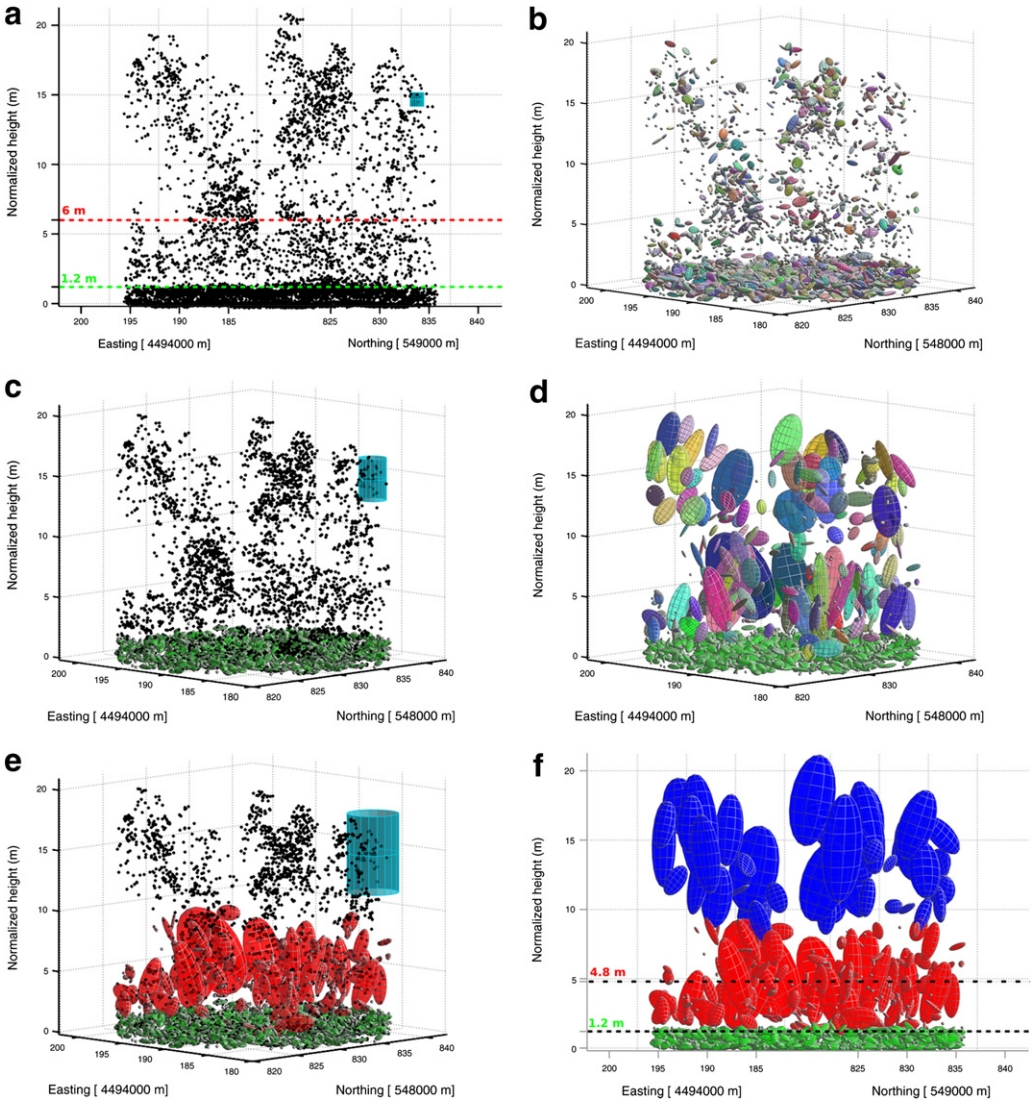

Fig. 10. Segmentation of plot #30 with $ht_{us}=1$ m and $ht_{os}=8$ m. The black dots correspond to the ALS points that remain unlabelled after an iteration and that are inputs for the next iteration. (a–b) First iteration: $w=0$ m and $h_{gv}{}^*=(1,1)$. (c–d) Second iteration: $w=2$ m and $h_{us}{}^*=(2.3,3.5)$. (e–f) Third iteration: $w=9.5$ m and $h_{os}{}^*=(4.3,6.5)$. The lines in (a) and (f) respectively correspond to the field-measured and ALS-derived mean height of ground vegetation (green) and understory (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is set to the plot diameter (~22 m) and h^r is defined as the value that forces the ALS points to converge toward two modes. We set $h^r = 1$ m as an initial estimate and increment it to obtain these two modes (Fig. 6a). The borderline between the basins of attraction of each mode defines the overstory height threshold, ht_{os} (Fig. 6b). We assume that a plot holds a single layer when $ht_{os} < 1$ m and two layers when $ht_{os} < 5$ m; otherwise a third layer may exist. In this case, the understory height threshold ht_{us} is set to 1 m. Afterwards one can easily compute the thickness of the overstory (A_{os}), understory (A_{us}), and ground vegetation (A_{gv}).

Finally, the kernel bandwidth $h^* = (h^s, h^r)$ corresponding to the crown segmentation is adapted to the vegetation architecture to account for the aspect ratio of tree crowns, so the vertical (h^r) and horizontal (h^s) components may be different (Morsdorf et al., 2004). Based on the current ALS dataset, we find that the tree crown height is at least two thirds larger than the crown diameter, while ground vegetation is spherical ($h^s_{gv} = h^r_{gv}$). Then, equalizing the two vertical bandwidths h^r_{os} and h^r_{us} to half the thickness of the layers avoids under-segmentation in bilayered forests (Eqs. 8–9). Since ground vegetation is always considered as a uniform layer, the bandwidth $h_{gv}{}^*$ is set to the corresponding thickness in both directions (Eq. 10).

$$h^*_{os} = \left(\frac{2h^r_{os}}{3}, \frac{A_{os}}{2}\right) \tag{8}$$

$$h^*_{us} = \left(\frac{2h^r_{us}}{3}, \frac{A_{us}}{2}\right) \tag{9}$$

$$h^*_{gv} = \left(A_{gv}, A_{gv}\right). \tag{10}$$

3.3. Adjustment of the kernel profile

We design a 3-D kernel profile as the product of two profiles to compute the modes of the point cloud, i.e. the crown apices. Whereas the horizontal profile searches for the local density maxima, the vertical one deals with the local height maxima. The horizontal kernel profile g^s follows a Gaussian function,

$$g^s(x) = \exp\left(-\gamma \|x\|^2\right) \tag{11}$$

with $\gamma = 5$. Isotropic kernels are standard in image segmentation where emphasis is put on bandwidth selection (Comaniciu, 2003; Singh & Ahuja, 2003). Asymmetric kernels have been used in video tracking to adapt to the structure of moving targets, e.g. an airplane or a human body (Wang et al., 2004; Yi et al., 2008; Yilmaz, 2007). In this study, an asymmetric kernel is applied to the vertical component in order to assign a higher weight to the highest points within the bandwidth (Fig. 7). Therefore the MS vector converges toward the local height maximum. Following Yilmaz (2007) and Yi et al. (2008) we first create a mask of the foreground object,

$$\text{mask}(X_i) = \begin{cases} 1 & \text{if } X^r - \frac{h}{4} \le X^r{}_i \le X^r + \frac{h}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

And the kernel value is the distance between one data point and the boundary of the mask:

$$\text{dist}(X_i) = \left\{ \begin{array}{ll} \min\left\{ \left\| \frac{\left(X^r - \frac{h^r}{4}\right) - X^r_i}{\frac{3h^r}{8}} \right\|, \left\| \frac{\left(X^r + \frac{h^r}{2}\right) - X^r_i}{\frac{3h^r}{8}} \right\| \right\} & \text{if mask}(X_i) = 1 \\ 0 & \text{otherwise} \end{array} \right\} \tag{13}$$

where $3h^r/8$ is a normalizing factor equal to half the bandwidth of the asymmetric kernel. Using an Epanechnikov profile, the weight of

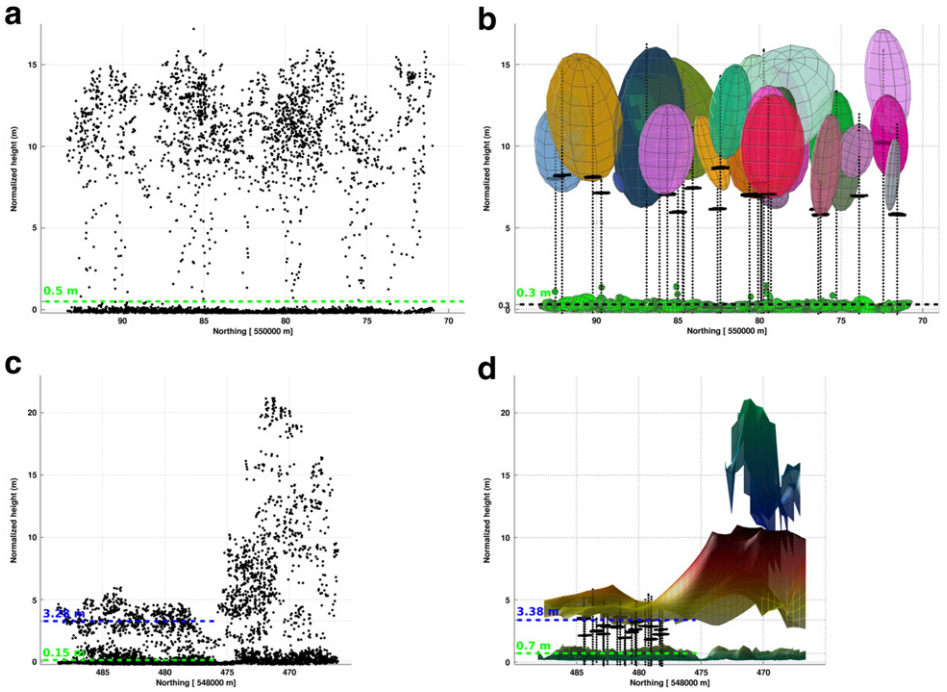

Fig. 11. Original point cloud for (a) plot #47 only composed of pine trees and (c) plot #16 made of two stands. Both plots do not display understory layers and the measured mean heights of ground vegetation (green) and overstory (blue) are represented by the lines in the figures. (b) MS individual vegetation features from (a). (d) Canopy height model of ground vegetation, understory, and overstory calculated from the individual vegetation features computed in (c). The surveyed tree metrics are also shown (line segments in black) in both figures. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Linear regression parameters for ALS-derived versus field-measured vegetation mean height. (*) The results only concern juvenile stands. Negative values mean an underestimation.

	Number of stands	Outliers	R^2	RMSE (m)	Δh (m)
Ground vegetation	44	3	0.70	0.15	0
Understory	32	5	0.68	0.96	0.44
Overstory (*)	10	2	0.92	0.31	−0.12

each point is calculated using

$$g^{ar}(X_i) = \begin{cases} 1-\|1-\mathrm{dist}(X_i)\|^2 & \text{if } \mathrm{mask}(X_i) = 1 \\ 0 & \text{otherwise} \end{cases}. \quad (14)$$

In the case of an asymmetric kernel, the MS vector in Eq. (5) can be then rewritten as

$$m_{h^s,G^s}(X) = \frac{\sum_{i=1}^{n} X_i g^s\left(\left\|\frac{X^s - X_i^s}{h^s}\right\|^2\right) g^{ar}(X_i)}{\sum_{i=1}^{n} g^s\left(\left\|\frac{X^s - X_i^s}{h^s}\right\|^2\right) g^{ar}(X_i)} - X. \quad (15)$$

Note that the profile is still radially symmetric (Eq. 14). The neighborhoods accounted for in the calculation of $m_{h^*,G^*}(X)$ are selected as a function of an asymmetric bandwidth. The weighted distance between points is the product of the two kernels, which makes the method more robust (Fig. 7). For instance, overlapped crowns may also correspond to local density maxima. Whereas the horizontal profile tends to converge to such zones, the vertical profile forces the MS vector to converge on the local height maximum, i.e. the crown apex. Conversely, when undergrowth and overgrowth vegetation interpenetrate, the vertical profile tends to converge toward the upper plants. In such a case, the horizontal profile helps the MS vector to stabilize on the crown apex of the lower plants, which is supposed to be denser than the crown base of the upper plants.

3.4. Pre-processing of the point cloud

In a forest canopy, the laser beams hit leaves, branches, and trunks. Since the point cloud is very scattered, keeping all points significantly overestimates the number of individual vegetation features, as well as the estimation of the stratum height. In order to identify the crown elements in the 3-D point cloud, the mean shift (Eq. 5) has

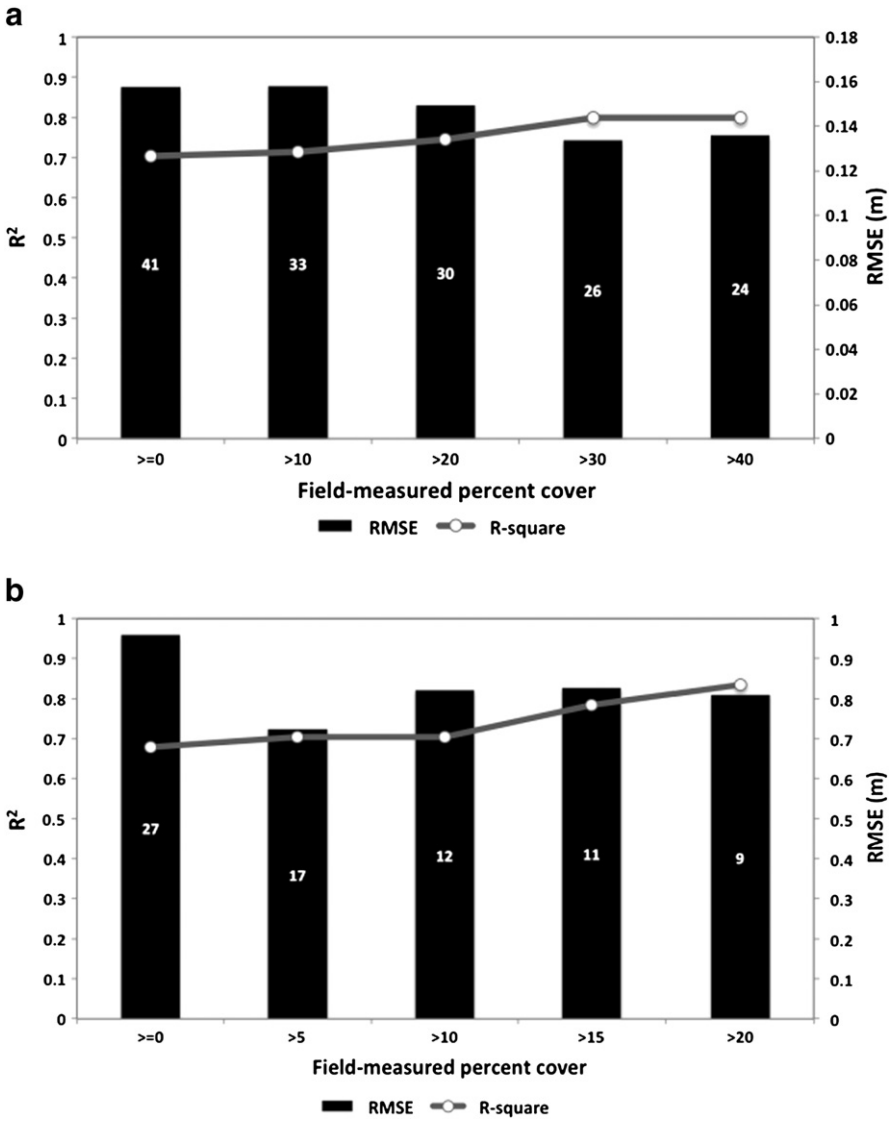

Fig. 12. Analysis of the R^2 (left axis) and the RMSE (right axis) for height estimation, as a function of the percent cover for (a) ground vegetation and (b) understory. The number of plots used to calculate these statistics is inscribed in the bars.

been applied to each plot using a uniform kernel (Eq. 7) and the bandwidth $h^* = (h^s, h^r)$, with $h^s = (3,3)$ m and $h^r = 3$ m. If all segments containing less than 5 points are removed from the data set, because of their poor topological structure, the bandwidth is large enough to keep the most significant vegetation features (Fig. 8). However, this technique may remove suppressed trees that are poorly represented in the point cloud due to occlusion that masks some parts of the canopy volume.

3.5. Extraction of individual trees and refinement of the forest strata

The algorithm involves two or three iterations (Fig. 9). It first computes a set of mean shift vectors using the ALS points (Eq. 15), which are all considered as seeds. The vectors search for the local highest density direction with the appropriate bandwidth. The latter is selected by calculating the 5th height percentile of the current point cloud, w. In the first iteration, the bandwidth is set to h_{gv}^* (Fig. 10a) since w always tends toward 0 m. A trajectory links every ALS point with a certain mode. A vegetation feature having a mode lower than ht_{us} is considered as ground vegetation (Fig. 10c, green ellipsoids). At the end of the first iteration, the corresponding ALS points are removed from the point cloud. The calculation of w in the second iteration defines the bandwidth and therefore the number of iterations (two or three). The bandwidth is h_{us}^* if $w < ht_{os}$ or/and ht_{os}^* if $w > ht_{os}$.

The second iteration extracts the understory, which corresponds to vegetation features with modes ranging between ht_{us} and ht_{os} (Fig. 10e, red ellipsoids). The third iteration identifies the overstory as vegetation features with modes higher than ht_{os} (Fig. 10f, blue ellipsoids). Applying a threshold to the mode space allows definition of fuzzy frontiers between the strata. This is physically meaningful compared to a simple vertical stratification based on height thresholds. After each iteration, removing points already assigned improves the segmentation by reducing the influence of the denser layers. Thus, when two regions of different densities are close together, the points belonging to sparser regions are likely to be aggregated by those belonging to the denser ones. This effect is obvious in Fig. 5b where the forest strata are either overestimated or underestimated.

4. Results

This section discusses the results of the algorithm over 44 plots. They are validated in terms of the forest vertical stratification, as well as the identification of individual trees.

4.1. Segmentation of forest strata

The mean height of ground vegetation is calculated as the 90th height percentile (Riaño et al., 2007) of the corresponding laser points (green ellipsoids of Figs. 10f and 11b). Unlike other approaches, we keep all the points, including ground reflections, which justify such a high value. The 50th height percentile is naturally used to calculate the mean heights of understory (Fig. 10f, red ellipsoids) and overstory (juvenile stands, Fig. 11d) (Peterson, 2005).

Linear regression analysis allows investigation of the strength of the relationship between the ALS-derived and field-measured heights of each forest stratum (Table 5). The outliers that represent about 7% and 16% of the plots in ground vegetation and understory, respectively, are identified after Huber (1981) and removed from the linear regressions. A linear model with a satisfactory RMSE explains 70% of the variability associated with ground vegetation height. Note the refinement accomplished by the algorithm: initially set to a 1 m threshold (Fig. 6), the computed height ranges from 0.15 m to 1.25 m. The number of retrieved layers is inherent to the forest pattern. Although all mature plots were initially divided into three strata, stands #9, #29, #45, #46 and #47 converge toward only two strata (Fig. 11a–b), which means that the echoes reflected by the trunks are successfully identified. Due to the lack of understory, the condition $w > ht_{us}$ is verified earlier in the second iteration and, consequently, the kernel bandwidth is immediately optimal for the overstory stratum. The MS algorithm also works on plots containing several stands, the vertical stratification of which varies radically (Fig. 11d). The mean height of the understory is overestimated. The linear model explains 68% of the variance (Table 5). This may be due to the assignment of suppressed trees to this layer, contrary to field measurements. These trees can be considered as understory since they grow below the canopy and do not receive direct sunlight. As expected, the estimates of overstory mean height are more accurate for the juvenile stands (Table 5).

Fig. 12 shows how the percent cover affects the estimation of ground vegetation and understory height. Ground vegetation is surprisingly not much affected with R^2 varying from 0.70 to 0.80 and RMSE lower than 0.02 m (Fig. 12a). As for the understory, the percentage of explained variance increases with the percent cover while the RMSE decreases (Fig. 12b). A higher percent cover indicates more plant material and a higher proportion of laser pulses hitting the canopy. Therefore, the discrete model of vegetation generates a better estimate of forest parameters. The understory height is more accurate when the percent

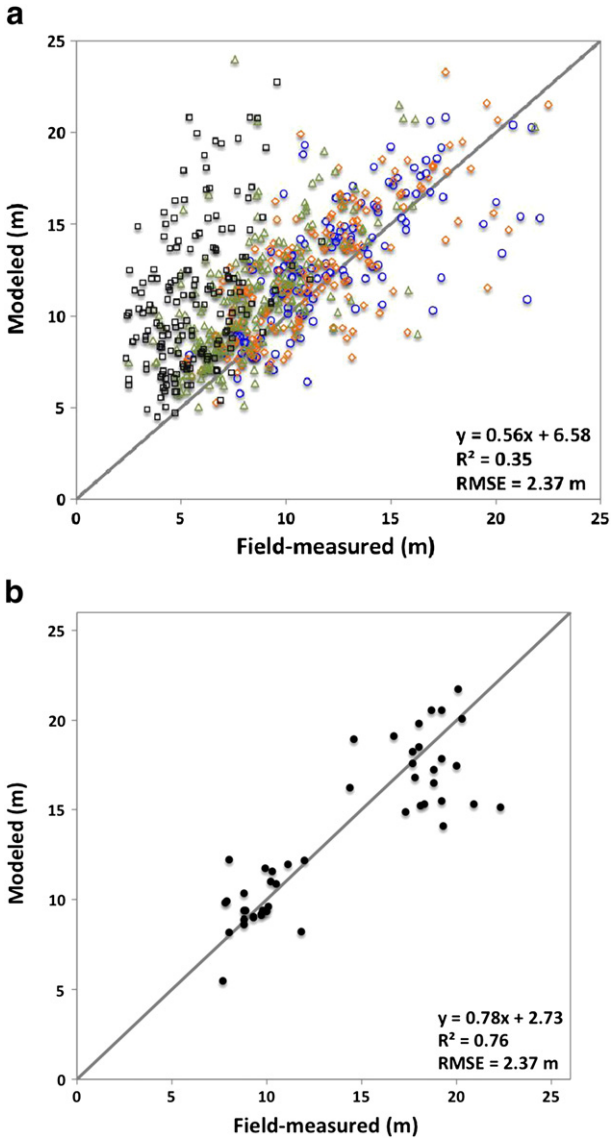

Fig. 13. Modeled vs. field-measured CBH for (a) eucalypts (◦ dominant, ◊ codominant, Δ dominated, □ suppressed) and (b) pine trees.

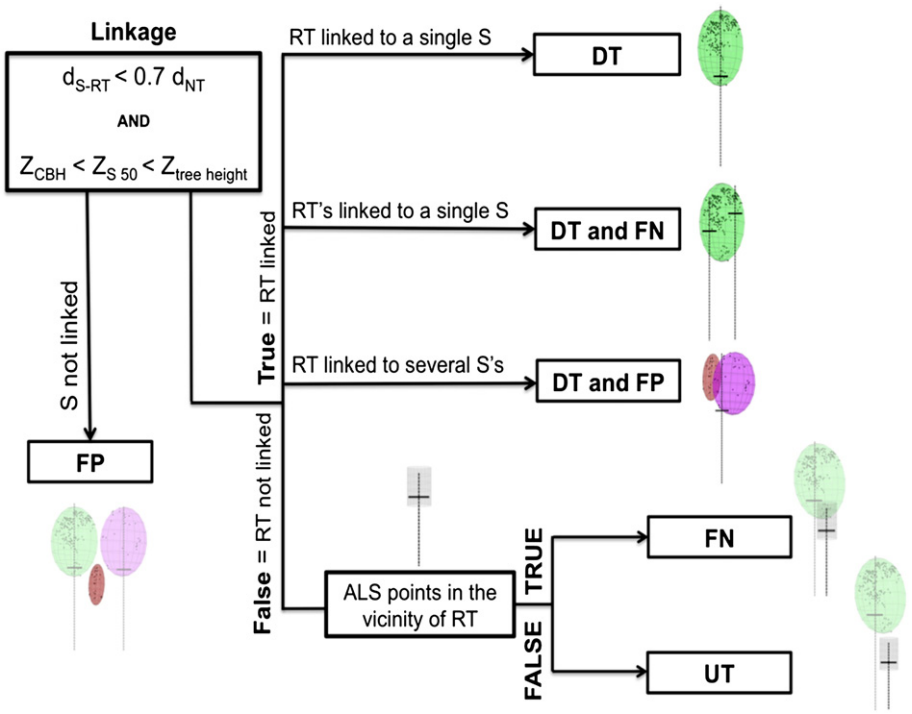

Fig. 14. Flowchart of the reference trees (RT) and ALS segments (S) linkage method.

cover exceeds 10%, thus a post-processing analysis for identifying sparse canopies may improve the results.

We are interested in comparing our results with CBH, which plays a greater role in forest stratification. Fig. 13 compares the field-measured CBHs with those modeled by selecting the lowest points sorted out as overstory in 0.3 m × 0.3 m areas (Fig. 10f and Fig. 11b, blue and colored ellipsoids). The missing pixels were generated using a Delaunay triangulation. Such a surface explains 76% of the variability of the pine CBH but it poorly characterizes the eucalyptus stands, which are more heterogeneous (Fig. 13).

4.2. Identification of individual tree crowns

As in Solberg et al. (2006) and Reitberger et al. (2009), the 3-D segmentation of individual tree crowns is validated by comparing field measurements with ALS segments (Figs. 11b and 14). A segment is linked with a reference tree provided that *i*) the distance $d_{S\text{-}RT}$ is lower than 70% of the mean distance d_{NT} between eight neighboring trees and, *ii*) the height values of at least 50% of the ALS points of S, $Z_{S\ 50}$, are contained between the CBH and the tree height.

If a segment is assigned to more than one reference tree, the farthest tree from the reference tree is considered a false negative (FN). In order to quantify the remaining omission errors, the neighborhood of unlinked reference trees was analyzed using a cylinder of radius 1.5 m. If there is at least one laser point linked with another reference tree within this volume, the current one is also called a false negative. Thus, the FN class means that the tree crown was detected by the ALS but the algorithm failed to see it as a tree. This is the case when two crowns were clustered in the same segment. If no laser point belongs to this buffer area, a reference tree is declared as an undetected tree (UT). Finally, segments linked with any reference tree are classified as false positive (FP). This class may contain vegetation features wrongly assigned to the overstory, e.g. tall shrubs, but also trees located outside the substand boundary when their crowns fall inside and are not surveyed. Thus, the detected trees (DT) quantify the performance of ALS in characterizing the forest (Table 6).

As expected, the detection rate decreases with dominance position. The estimation error of biomass or basal area should vary accordingly (Persson et al., 2002). To report the number of trees missed by the method, we can sum the omission errors introduced by the algorithm, i.e., DT−FN. They are actually low compared to those introduced by the ALS (0.7, 7.4, 4.3, 1.7, and 3.8 percentage points for dominant, co-dominant, dominated, suppressed, and pine, respectively). The percentage of FP, or commission error, equals 8.6%, which is in good agreement with other studies. In a forest mainly covered with Norway spruce, European beech, fir, and sycamore maples, Reitberger et al. (2009) detect 66% of the reference trees (upper layer 88%, intermediate layer 35%, lower layer 24%) with a commission error of 11%. In a Norway spruce forest, Solberg et al. (2006) announce a global detection rate of 66% (dominant trees 93%, codominant trees 63%, subdominant trees 38%, and suppressed trees 19%) with a commission error of 26%. It is unclear whether the omission errors reported by other studies are due to the inability of the ALS to characterize tree crowns or to the algorithm itself. Therefore, it is tricky to compare our results with the literature since the forest architecture and the ALS configuration both have an important effect on the accuracy of the different methods.

Although the present method searches for local density maxima in the point cloud, it is not affected by the point density variability because the MS is a kernel gradient estimator, i.e., it does not evaluate the density function itself but normalized local gradients. Thus, provided that the local density and height gradients point toward the crown apices, the point density at which the crowns are sampled

Table 6
Tree identification. (*) In total there are 167 suppressed reference trees but 50 that have been classified as understory are not taken into account.

Tree	Dominance position	Reference trees	Identified		FP
			DT	DT−FN	
Eucalyptus	Dominant	146	145 (99.3%)	144 (98.6%)	60 (9.2%)
	Codominant	176	163 (92.6%)	150 (85.2%)	
	Dominated	210	138 (65.7%)	129 (61.4%)	
	Suppressed*	117	17 (14.5%)	15 (12.8%)	
Pine		52	50 (96.1%)	48 (92.3%)	0
Total		701	513 (73.2%)	486 (69.3%)	60 (8.6%)

has only a slight impact on the mode search, i.e., on the identification of individual vegetation features.

4.3. Validation of tree height and CBH

Fig. 15 correlates the ALS-derived and field-measured tree height (Fig. 15a and 15c) and CBH (Fig. 15b and 15d) for the identified trees. Characterization of the CBH greatly improves in eucalyptus stands when individual trees are first extracted (Figs. 13a and 15b), while it is slightly better in pine stands (Figs. 13b and 15d). Table 7 shows that our method globally underestimates the tree height, with a limited influence of the dominance position. The slopes of the linear regressions almost equal 1, the R^2 vary between 0.91 and 0.95, and the RMSE between 0.75 m and 0.90 m. These results are comparable with those of other studies that show that ALS data tend to underestimate tree height (Gaveau & Hill, 2003; Hyyppä et al., 2008).

Our method overestimates the CBH of 1.29 m for eucalyptus, and a positive correlation with the dominance position is obvious. The linear regressions follow the same trends, with an R^2 increasing from 0.58 (dominant) to 0.71 (suppressed) and an RMSE decreasing from 2.80 m (dominant) to 1.30 m (suppressed). The crown base is not as well delineated for eucalyptus as for pine. Suppressed trees are more compact than taller trees, the shape of which is more complicated with small dead branches lying on the stems. Moreover, the reflection of the laser beam on a curved branch can be located under the field-measured CBH. This variable is actually difficult to survey because of its approximate definition: it can be viewed as the height of the first branch along the stem or as the height where the crown bulk density exceeds a critical threshold of 0.011 kg/m^3 (Scott & Reinhardt, 2001). The pine CBH is underestimated by 0.66 m, mainly because of dead branches that were not measured in the field. Many ALS points corresponding to trunks are also clustered together with crowns, particularly in the old stands. Compared to eucalypts and young pines, trunks of old pines are well represented in the point cloud. Other methods are more successful in removing their reflections (Popescu & Zhao, 2008) but it is unclear whether they would improve the CBH estimation. Our results agree with other studies: in a Scots pine forest, Riaño et al. (2004) claim that ALS overestimates the CBH and obtain R^2 values ranging from 0.65 to 0.68. In Norway spruce and Scots pine forests, Holmgren and Persson (2004) also notice an overestimation by 0.75 m ($R^2 = 0.84$, RMSE = 2.82 m). Popescu and Zhao (2008) extract the CBH of pines and deciduous trees with an RMSE of 2.08 m and an R^2 of 0.78.

5. Conclusion

This study demonstrates the ability of our method to provide genuine 3-D segments corresponding to individual vegetation features of the main forest layers: ground vegetation, understory, and overstory. Unlike other methods, our approach does not rely on a CHM and directly applies to the 3-D point cloud, which is an advantage in characterizing heterogeneous forests. Segmentation occurs in the mode space where vegetation features are more likely to be discriminated. Our maps allow local calculation of specific statistics for each vegetation layer and consequently accurate delineation of forest areas with similar horizontal and vertical structures, i.e., forest stands and consequently fuel types. Moreover, our approach introduces a robust discrimination between ground vegetation and taller plants.

We show that the mean shift algorithm is a reliable technique for finding the modes in the multi-modal point cloud distribution of a multi-layered Mediterranean forest. Due to the complex pattern of the forest environment, we established a multi-scale approach where modes are computed with an adaptive kernel bandwidth optimized for each stratum. However, so far it can only handle forest structures with a maximum of three layers. A more sophisticated method might be developed to deal with highly stratified environments.

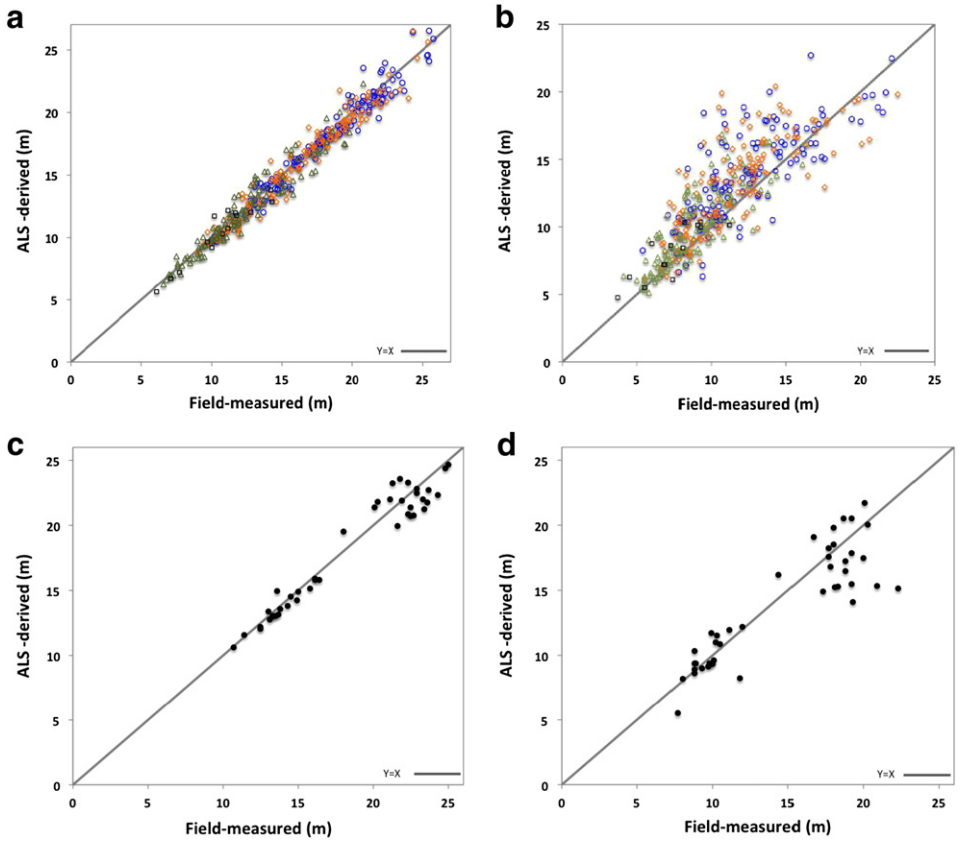

Fig. 15. ALS-derived vs field-measured tree height (a–c) and CBH (b–d) for eucalyptus (a–b, ○ dominant, ◊ codominant, Δ dominated, □ suppressed) and pine trees (c–d).

Table 7
Linear regression parameters for data displayed in Fig. 15. Negative values mean an underestimation while positive values mean an overestimation.

Tree	Dominance position	Δh (m)		R^2		RMSE (m)	
		TH	CBH	TH	CBH	TH	CBH
Eucalyptus	Dominant	−0.23	1.44	0.95	0.58	0.85	2.80
	Codominant	−0.27	1.45	0.95	0.61	0.87	2.70
	Dominated	−0.17	1.03	0.93	0.67	0.90	1.92
	Suppressed	−0.22	0.73	0.91	0.71	0.75	1.30
	All together	−0.23	1.29	0.96	0.69	0.86	2.48
Pine		−0.28	0.66	0.94	0.79	1.07	2.25

Our approach relies on only one parameter, the three-dimensional kernel bandwidth. Its vertical component is set as a function of the stratum depth and its horizontal component is defined in relation to the vertical one. Therefore, the kernel bandwidth has a biophysical meaning: the width of a crown depends on its length and the depth of a forest stratum on the length of the crowns. Note that these correlations may vary significantly depending on the tree species and the forest biome. Thus, it is necessary to determine the validity domain of these kernel bandwidth settings. The robustness of the method was assessed at four different levels:

a) Intra-plot. The method is able to depict the real nature of the strata, even when the vertical stratification varies within a plot (41% of the plots have more than one stand, Fig. 11d).
b) Intra-stand. The bandwidth settings apply well to crowns with different volumes, from suppressed to dominant trees (Fig. 3 and Table 6).
c) Inter-stand. The validated stands display structures with different arrangements, from little to lush ground vegetation, combined with either absent or luxurious understory that can co-exist with overgrowth vegetation at different growth stages (Fig. 2 and Table 2).
d) Inter-plot. Our forest is made up of many small properties that lead to a fragmented landscape. The method does a good job of handling the point density variability within the study area (Fig. 1 and Table 4).

Finally, the correlation between field measurements and ALS-derived structural characteristics of ground vegetation and understory depends on the forest type and the ALS configuration. Such values may be different in forests with more closed canopies or sparser point clouds.

Acknowledgments

This experiment is part of a PTDC/AGR-CFL/72380/2006 research project. A. Ferraz holds a fellowship (SFRH/BD/38390/2007) funded by the Portuguese Foundation for Science and Technology. Many thanks to Susan L. Ustin (U.C. Davis) for editing the paper. IPGP contribution no. 3257.

References

AFN (2009). *Instruções para o trabalho de campo do Inventário Florestal Nacional, IFN 2005/2009.* Direção de Unidade de Gestão Florestal, Divisão para a Intervenção Florestal, Lisboa, Portugal: Autoridade Florestal Nacional 62 pp.
Andersen, H. E., McGaughey, R. J., & Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using LiDAR data. *Remote Sensing of Environment*, *94*, 441–449.
Anderson, H. (1982). Aids to determining fuel models for estimating fire behavior. *USDA, forest service—intermountain experiment station* 22 pp.
Andrews, P., Bevins, C., & Seli, R. (2005). BehavePlus fire modeling system, version 3.0: User's guide revised. *USDA, forest service—rocky mountain research station* 132 pp.
Antonarakis, A. S., Richards, K. S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment*, *112*, 2988–2998.
Ares, A., Neill, A. R., & Puettmann, K. J. (2010). Understory abundance, species diversity and functional attribute response to thinning in coniferous stands. *Forest Ecology and Management*, *260*, 1104–1113.
Asner, G. P., Hughes, R. F., Vitousek, P. M., Knapp, D. E., Kennedy-Bowdoin, T., Boardman, J., et al. (2008). Invasive plants transform the three-dimensional structure of rain forests. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 4519–4523.
Asner, G. P., Powell, G. V. N., Mascaro, J., Knapp, D. E., Clark, J. K., Jacobson, J., et al. (2010). High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 16738–16742.
Bo, S., Ding, L., Li, H., Di, F., & Zhu, C. (2009). Mean shift-based clustering analysis of multispectral remote sensing imagery. *International Journal of Remote Sensing*, *30*, 817–827.
Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., & Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment*, *114*, 911–924.
Bretar, F., & Chehata, N. (2010). Terrain modelling from lidar range data in natural landscapes: A predictive and Bayesian framework. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 1568–1578.
Brokaw, N. V., & Lent, R. A. (2000). Vertical structure. In M. L. Hunter (Ed.), *Maintaining biodiversity in forest ecosystems* (pp. 373–399). : Cambridge University Press.
Burman, H., & Soininen, A. (2004). Available online at:. *TerraMatch user's guide.* http://www.terrasolid.fi/system/files/tmatch.pdf (accessed: 6/07/2011).
Camprodon, J., & Brotons, L. (2006). Effects of undergrowth clearing on the bird communities of the Northwestern Mediterranean Coppice Holm oak forests. *Forest Ecology and Management*, *221*, 72–82.
Clawges, R., Vierling, K., Vierling, L., & Rowell, E. (2008). The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment*, *122*, 2064–2073.
Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*, 603–619.
Comaniciu, D. (2003). An algorithm for data-driven bandwidth selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 281–288.
Coops, N. C., Hilker, T., Wulder, M. A., St-Onge, B., Newnham, G., Siggins, A., et al. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees—Structure and Function*, *21*, 295–310.
Dean, T. J., Cao, Q. V., Roberts, S. D., & Evans, D. L. (2009). Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand. *Forest Ecology and Management*, *257*, 126–133.
EEA (2008). European forests—ecosystem conditions and sustainable use. *EEA report no 3/2008*. Copenhagen (Denmark): European Environment Agency 105 pp.
DGRF (2005). 5° Inventario Florestal Nacional. Fotointerpretaçao. *Direcção Geral dos Recursos Florestais, Lisboa, Portugal* 12 pp.
Di Castri, F. (1981). Mediterranean-type shrublands of the world. In F. Di Castri, D. Goodall, & R. Specht (Eds.), *Ecosystems of the world: Mediterranean-type shrublands* (pp. 1–52). Amsterdam (The Netherlands): Elsevier Scientific Publications.
Finney, M. (2004). FARSITE: Fire area simulator-model development and evaluation. *USDA forest service, research paper RMRS-RP-4* 47 pp.
García, M., Riaño, D., Chuvieco, E., & Danson, F. M. (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, *14*, 816–830.
Gaveau, D., & Hill, R. (2003). Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, *29*, 650–657.
Gonçalves, G., & Pereira, L. (in press). A thorough accuracy estimation of DTM produced from airborne full-waveform laser scanning data of unmanaged eucalypt plantations. *IEEE Transactions on Geoscience and Remote Sensing*. doi:10.1109/TGRS.2011.2180911.
Hall, F. G., Bergen, K., Blair, J. B., Dubayah, R., Houghton, R., Hurtt, G., et al. (2011). Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment*, *115*, 2753–2775.
Hollaus, M., Wagner, W., Eberhöfer, C., & Karel, W. (2006). Accuracy of large-scale canopy heights derived from LiDAR data under operational constraints in a complex alpine environment. *ISPRS Journal of Photogrammetry and Remote Sensing, 60*, 323–338.
Holmgren, J., & Persson, A. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, *76*, 283–297.
Huang, X., & Zhang, L. (2008). An adaptive mean-shift analyses approach for object extraction and classification from urban hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *46*, 4173–4185.
Huber, P. J. (1981). *Robust statistics.* New York: Wiley 320 pp.
Hyyppä, J., Hyyppä, H., Litkey, P., Yu, X., Haggrén, H., Ronnholm, P., et al. (2004). Algorithms and methods of airborne laser scanning for forest measurements. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*, 82–89.
Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., & Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, *29*, 1339–1366.
Jaskierniak, D., Lane, P., Robinson, A., & Lucieer, A. (2010). Extracting LiDAR indices to characterize multi-layered forest structure using mixture distributions functions. *Remote Sensing of Environment*, *115*, 537–585.
Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*, 193–203.
Landsberg, J. J., & Gower, S. T. (1997). Forest biomes of the world. *Applications of physiological ecology to forest management* (pp. 19–50). San Diego: Academic Press.
Mallet, C., & Bretar, F. (2009). Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 1–16.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., & Vehmas, M. (2004). Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment*, *90*, 319–330.

Maltamo, M., Packalén, P., Yu, X., Eerikainen, K., Hyyppä, J., & Pitkanen, J. (2005). Identifying and quantifying structural characteristics of heterogeneous boreal forest using laser scanner data. *Forest Ecology and Management*, *216*, 41–50.

Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., et al. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, *113*, 2533–2546.

Moore, P. T., Van Miegroet, H., & Nicholas, N. S. (2007). Relative role of understory and overstory in carbon and nitrogen cycling in a southern Appalachian spruce-fir forest. *Canadian Journal of Forest Research*, *37*, 2689–2700.

Morsdorf, F., Meier, E., Kötz, B., Itten, K. I., Dobbertin, M., & Allgöwer, B. (2004). LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sensing of Environment*, *92*, 353–362.

Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., et al. (2010). Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. *Remote Sensing of Environment*, *114*, 1403–1415.

Mutlu, M., Popescu, S. C., Stripling, C., & Spencer, T. (2008). Mapping surface fuel models using lidar and multispectral data fusion for fire behavior. *Remote Sensing of Environment*, *112*, 274–285.

Pereira, L., Gonçalves, G., Soares, P., Cambra, S., Carvalho, S., & Tomé, M. (2009). Planning and acquisition of control data to validate forest inventory and the estimation of fuel variables derived from LiDAR data and high resolution CIR images. *Proc. 6° Congresso Florestal Nacional, Ponta Delgada- Açores, 6–9 Outubro 2009* 9 pp.

Persson, Å., Holmgren, J., & Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, *68*, 925–932.

Persson, Å., Holmgren, J., Söderman, U., & Olsson, H. (2004). Tree species classification of individual trees in Sweden by combining high resolution laser data with high resolution near-infrared digital images. *International Archives of Photogrammetry*, *36*, 204–207.

Peterson, B. (2005). *Canopy fuels inventory and mapping using large-footprint lidar*. PhD Thesis, University of Maryland (MD), 218 pp.

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*, 589–604.

Popescu, S. C., & Zhao, K. (2008). A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sensing of Environment*, *112*, 767–781.

Pyne, S. J., Andrews, P. L., & Laven, R. D. (1996). *Introduction to wildland fire* (2nd Edition). New York: John Wiley & Sons 808 pp.

Reitberger, J., Schnörr, C., Krzystek, P., & Stilla, U. (2009). 3D Segmentation of single trees exploiting full waveform LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 561–574.

Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., & Ustin, S. L. (2003). Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behaviour modeling. *Remote Sensing of Environment*, *86*, 177–186.

Riaño, D., Chuvieco, E., Condés, S., Gonzalez-Matesanz, J., & Ustin, S. L. (2004). Generation of crown bulk density for *Pinus sylvestris* L. from lidar. *Remote Sensing of Environment*, *92*, 345–352.

Riaño, D., Chuvieco, E., Ustin, S. L., Sala, J., Rodriguez-Perez, J. R., Ribeiro, L. M., et al. (2007). Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. *International Journal of Wildland Fire*, *16*, 341–348.

Richardson, J. J., & Moskal, L. M. (2011). Strengths and limitations of assessing forest density and spatial configuration with aerial LiDAR. *Remote Sensing of Environment*, *115*, 2640–2651.

RIEGL (2011). Available online at:. RiANALYZE. http://www.riegl.com/products/software-packages/rianalyze/ (accessed: 21/07/2011).

RIEGL (2011). Available online at:. RiWORLD. http://www.riegl.com/products/software-packages/riworld/ (accessed: 21/07/2011).

Sandberg, D. V., Ottmar, R. D., & Cushon, G. H. (2001). Characterizing fuels in the 21st century. *International Journal of Wildland Fire*, *10*, 381–387.

Scott, J. H., & Reinhardt, E. D. (2001). Assessing crown fire potential by linking models of surface and crown fire behaviour. *USDA forest service research paper RMRS-RP-29* (pp. 9–21). Fort Collins, CO: Rocky mountain research station.

Topographic laser ranging and scanning. Shan, J., & Toth, C. K. (Eds.). (2009). *Principles and processing*. : CRC Press 608 pp.

Singh, M., & Ahuja, N. (2003). Regression based bandwidth selection for segmentation using Parzen windows. *Proc. 9th IEEE International Conference on Computer Vision, Nice (France), 13–16 October 2003* (pp. 2–9).

Soininen, A. (2010). Available online at:. *TerraScan user's guide*. http://www.terrasolid.fi/en/users_guide/terrascan_users_guide (Accessed 6/07/2011).

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

Stokes, B. J., Ashmore, C., Rawlins, C. L., & Sirois, D. L. (1989). Glossary of terms used in timber harvesting and forest engineering. General technical report SO-73. *USAD, forest service*. New Orleans (LA): Southern Forest Experiment Station 33 pp.

Wang, J., Thiesson, B., Xu, Y., & Cohen, M. (2004). Image and video segmentation by anisotropic kernel mean shift. *Proc. European Conference on Computer Vision*, *vol. 2*. (pp. 238–249).

Yi, K. M., Ahn, H. S., & Choi, J. Y. (2008). Orientation and scale invariant mean shift using object mask-based kernel. *Proc. 19th International Conference on Pattern Recognition, Tampa (FL,) 8–11 December 2008* (pp. 1–4).

Yilmaz, A. (2007). Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection. *Proc. IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis (MN), 17–22 June 2007* (pp. 1–6).

Yoon, J. S., Shin, J. I., & Lee, K. S. (2008). Land cover characteristics of airborne LiDAR intensity data: A case study. *IEEE Geoscience and Remote Sensing Letters*, *9*, 463–466.

Zhao, K., Popescu, S., & Nelson, R. (2009). LiDAR remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment*, *113*, 182–196.

Zimble, D. A., Evans, D. L., Carlson, G. C., Parker, R. C., Grado, S. C., & Gerard, P. D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*, *87*, 171–182.