Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data

Crespo-Peremarch, Pablo [1,*], Tompalski, Piotr [2], Coops, Nicholas C.[2] and Ruiz, Luis Ángel [1]

[1] Geo-Environmental Cartography and Remote Sensing Group (CGAT), Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camí de Vera s/n, 46022, València, Spain.

[2] Integrated Remote Sensing Studio (IRSS), Department of Forest Resources Management, Forest Science Centre, 2424 Main Mall, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Keywords: lidar, airborne laser scanning, full-waveform, terrestrial laser scanning, understory, characterization, Mediterranean forest.

Abstract

The use of laser scanning acquired from the air, or ground, holds great potential for the assessment of forest structural attributes, beyond conventional forest inventory. The use of full-waveform airborne laser scanning (ALS_{FW}) data allows for the extraction of detailed information in different vertical strata compared to discrete ALS (ALS_D). Terrestrial laser scanning (TLS) can register lower vertical strata, such as understory vegetation, without issues of canopy occlusion, however is limited in its acquisition over large areas. In this study we examine the ability of ALS_{FW} to

* Corresponding author

Email address: pabcrepe@cgf.upv.es

characterize understory vegetation (i.e. maximum and mean height, cover, and volume), verified using TLS point clouds in a Mediterranean forest in Eastern Spain. We developed nine full-waveform metrics to characterize understory vegetation attributes at two different scales (3.75 m square subplots and circular plots with a radius of 15 m); with, and without, application of a height filter to the data. Four understory vegetation attributes were estimated at plot level with high R^2 values (mean height: R^2 = 0.957, maximum height: R^2 = 0.771, cover: R^2 = 0.871, and volume: R^2 = 0.951). The proportion of explained variance was slightly lower at 3.75 m side cells (mean height: R^2 = 0.633, maximum height: R^2 = 0.470, cover: R^2 = 0.581, and volume R^2 = 0.651). These results indicate that Mediterranean understory vegetation can be estimated and accurately mapped over large areas with ALS_{FW}. The future use of these types of predictions includes the estimation of ladder fuels, which drive key fire behaviour in these ecosystems.

1. Introduction

Understory vegetation is an essential component of forest ecosystems (Suchar and Crookston, 2010). The understory is critical for wildlife habitat, nesting and foraging (Hill and Broughton, 2009; Martinuzzi et al., 2009, Wing et al., 2012), impacts overstory regeneration (Royo and Carson, 2006), provides protection against soil erosion (Suchar and Crookston, 2010), as well as mediates microclimatic conditions below the canopy. The height, cover, and condition of the understory are also key drivers of fire behavior through fuel ladders, which drive crown fires (Molina et al., 2011). These types of fires are the most dangerous in terms of economic impacts and tree death (Molina et al., 2009).

Despite its importance, understory vegetation has conventionally been difficult to describe spatially, particularly over large areas (Wing et al., 2012). Traditional techniques, such as the line interception method (Canfield, 1941), often used in field surveys (Vierling et al., 2013), are very costly and only provide information over small spatial extents (Riaño et al., 2007). Airborne or satellite-borne passive optical remote sensing approaches can acquire data over large areas, but have limitations for characterizing vertical forest structure (Kerr and Ostrovsky, 2003; McDermid et al., 2005; Wulder and Franklin, 2012).

Active remote sensing techniques, such as Light Detection and Ranging (lidar), provide horizontal and vertical information of different canopy layers (Ruiz et al., 2018). Several studies have estimated characteristics of understory vegetation cover using discrete return airborne lidar, also known as discrete airborne laser scanning (ALS_D, Table 1). Most of these studies utilise classification approaches, where understory vegetation is classified based on a set of characteristics derived from point cloud data (Hill and Broughton, 2009; Martinuzzi et al., 2009; Morsdorf et al., 2010). Less common approaches involve regression, where understory characteristics are mapped in a continuous fashion (Wing et al., 2012). Martinuzzi et al. (2009) defined and classified two categories of understory cover (above and below 25%) using ALS_D in a mixed temperate coniferous forest in Northern Idaho with an overall classification accuracy of 0.83 and a kappa value of 0.66. In a temperate deciduous woodland in Cambridgeshire (England), Hill and Broughton (2009) predicted the presence and absence of understory using two separate leaf-on and leaf-off ALS flights, with a pulse density of 0.5 m^{-2} and 1 m^{-2}, respectively. The overall accuracy and kappa value of the classification were 0.77 and 0.53, respectively. Mosdorf et al. (2010) classified different vertical layer strata using height and intensity from ALS_D in a pine-evergreen oak woodland in the French Mediterranean region, resulting in an overall accuracy of 0.48 for the shrub layer. More recently, Wing et al. (2012) estimated understory cover in an interior ponderosa pine forest in Northeastern California using ALS_D with a mean density of 6.9 points m^{-2}. The authors introduced a new metric to characterize understory ALS points using a height and intensity filter, resulting in a proportion of explained variance of 0.74 and relative root mean square error (nRMSE) of 22%. Kobal et al. (2015) also used ALS_D and extracted a range of canopy gap and understory information such as canopy “sinkholes” and plant species richness beneath dense forest cover. Other studies estimated shrub height and cover in Central Portugal and the Spanish Mediterranean using ALS_D (Riaño et al., 2007; Estornell et al., 2011). However, these sites were dominated by shrubland, where there is little overstory, which reduces the impact of resulting of overstory occlusion.

As opposed to discrete return systems, full-waveform airborne laser scanning (ALS_{FW}) can register the returning pulse characteristics as they pass through the forest canopy, allowing for the extraction of additional information on forest structure (Hermosilla et al., 2014a). As the return pulse provides a full representation of the intercepted forest structure, it is likely an improved representation of understory vegetation (Anderson et al., 2016). This is because the vertical resolution is increased within each footprint and compared to a limited number of discrete points (Vierling et al., 2013). However, there are only few studies demonstrating the capability of ALS_{FW} to characterize understory vegetation (Table 1). Hancock et al. (2017) characterized voxelized understory cover in an urban area (Luton, England) using ALS_{FW} data. They proposed a new method to calibrate and validate results retrieved from ALS_{FW} using TLS as reference and obtained an understory cover accuracy of 24% at 1.5 m horizontal and 0.5 m vertical resolution. Harding et al. (2001) derived canopy height profiles (CHP) retrieved from a large-footprint ALS_{FW} such as Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) and ground-based measures.

Focusing on the understory strata, SLICER underestimated cover by 33% compared to ground-based measures. Comparing ALS_{FW} to ALS_D for more conventional forest inventory attribute estimation, Hermosilla et al. (2014a) found no statistical difference for many of the compared both technologies to estimate canopy fuel and structure attributes. Cao et al. (2014) used ALS_{FW} to estimate biomass components, finding that ALS_{FW} explained more variability for crown biomass than ALS_D, and that the combination of both datasets produced the best results. Fieber et al. (2015) applied a procedure based on Harding et al. (2001) to obtain the CHP, using small-footprint $ALS_{FW,}$ and observed a strong relationship between ALS and field data with a mean R^2 of 0.75. Lastly, Anderson et al. (2016) found that in an urban woodland landscape, canopy height estimated by ALS_D was more biased, and intensity less accurate, than that provided by ALS_{FW}.

Compared to ALS, terrestrial laser scanning (TLS) can produce a higher number of laser returns due to the close range nature of the technology (Vierling et al., 2013). This allows analysis of understory structure in much more detail (Vierling et al., 2013). TLS systems can register denser point clouds in lower vegetation (e.g. terrain, canopy base and understory) (Chasmer et al., 2006; Hilker et al., 2010; Crespo-Peremarch and Ruiz, 2017) and produce forest inventory information commensurate with field observations, registering data for more than 97% of the trees in deciduous, coniferous and mixed forests (Maas et al., 2008). However, despite its high accuracy, there is a lack of automatic algorithms to extract height and species from individual trees with TLS data (Liang et al., 2016). The highly detailed representation of the three-dimensional structure of the forest stand makes TLS point clouds an ideal dataset to characterize understory vegetation (Vierling et al., 2013; Greaves et al., 2015). TLS is often considered a much more efficient method than conventional field work, and it has successfully been proved as an effective and accurate approach to calibrate ALS-based models (Hopkinson et al., 2013; Hancock et al., 2017). However, because TLS is limited in its spatial coverage, it is restricted in its use as a forest management tool at broad spatial scales.

In this paper we explore the capacity of ALS_{FW} data to characterize understory vegetation in a Mediterranean forest ecosystem in Eastern Spain. In this region, understory structure depends on local climate and management practices. It is also a key variable in fire fuel assessment, which is a critical social and environmental issue for Eastern Spain. We first review existing ALS_{FW} metrics in the literature for overstory vegetation assessment and use these underlying principals to propose a set of new full-waveform metrics designed for understory vegetation assessment. These new metrics were derived using a voxel based approach and applied to estimate understory height, cover, and volume across a series of plots in the region. The metrics were validated using TLS data acquired simultaneously with ALS_{FW} point clouds. We conclude with an assessment of the ALS-based approaches and propose some recommendations for further development and testing of ALS_{FW} metrics.

2. Methods

1. Study area

The study area is located in the Natural Park of Sierra de Espadán, in the central Mediterranean region of Spain, about 50 km to the north of València (Fig. 1). The region is highly mountainous with steep hillsides, where elevation ranges from sea level to 1100 m within a few kilometers. Because of its topography and orientation, Sierra de Espadán Natural Park receives higher annual rainfall than its local surroundings, which combined with its unique geomorphology makes it a regional hotspot for biodiversity. The total area of the Natural Park is 31,000 ha, with our foci sites covering 12% (3,741.5 ha). The dominant species are Aleppo pine (*Pinus halepensis*), maritime pine (*Pinus pinaster*), cork oak (*Quercus suber*), and holm oak (*Quercus ilex*). The presence/absence and density of understory is very heterogeneous in the study area. Generally, forest stands dominated by maritime pine and cork oak have little or no understory (see Fig. 2), while stands dominated by Aleppo pine have much taller and denser understory (Fig. 2b and 3). The most common understory species are kermes oak (*Quercus coccifera*), tree heath (*Erica arborea*), brezo (*Erica multiflora*), flax-leaved daphne (*Daphne gnidium*), mastic (*Pistacia lentiscus*), aulaga (*Genista scorpius*), wild asparagus (*Asparagus acutifolius*), rosemary (*Rosmarinus officinalis*), Mediterranean buckthorn (*Rhamnus alaternus*), black hawthorn (*Rhamnus lycioides*), false olive (*Phillyrea angustifolia*), wild madder (*Rubia peregrina*), phoenicean juniper (*Juniperus phoenica*), common smilax (*Smilax aspera*), and thyme (*Thymus sp.*).

Table 1. Summary of existing studies about the characterization of understory using ALS with overstory presence.

Study	Study Area	Ecosystem	Definition of forest types	Target attributes	Data	Density (points.m^{-2})	No. of plots (plot size m^2)	Results
Martinuzzi et al. (2009)	Private industrial and experimental managed forest in Moscow Mountain in Northern Idaho, USA (30,000 ha)	Ponderosa pine (*Pinus ponderosa*), Douglas fir (*Pseudotsuga menziesii*), grand fir (*Abies grandis*), western red cedar (*Thuja plicata*) and western larch (*Larix occidentalis*)	Mixed temperate coniferous	Presence/absence of understory shrubs and snags (where cover > 25%)	ALS_D	-	83 (405)	Overall accuracy = 0.83 kappa = 0.66 (Classification)
Hill and Broughton (2009)	Monks Wood National Nature Reserve in Cambridgeshire, England (157 ha)	Ash (*Fraxinus excelsior*), English oak (*Quercus robur*), field maple (*Acer campestre*), silver birch (*Betula pendula*), aspen (*Populus tremula*) and small-leaved elm (*Ulmus carpinifolia*)	Temperate deciduous woodland	Presence/absence of understory combining data from leaf-on and leaf-off	ALS_D	Leaf-off: 1 pulse.m^{-2} Leaf-on: 0.5 pulse.m^{-2}	132 (400)	Overall accuracy = 0.77 kappa = 0.53 (Classification)
Morsdorf et al. (2010)	Experimental Mediterranean region of Lamanon, France (16.5 ha)	Aleppo pine (*Pinus halepensis*) and holm oak (*Quercus ilex*)	Mediterranean pine-evergreen oak woodland	Presence/absence of different vertical strata	ALS_D	3.7	63 (25)	Overall accuracy = 0.48 (Classification of shrub layer)
Wing et al. (2012)	Managed Blacks Mountain Experimental Forest in northeastern California, USA (4,358 ha)	Ponderosa pine (*Pinus ponderosa Dougl. ex P.* and *C. Laws*), fir (*Abies concolor (Gord.* And *Glend.*) *Lindl*), incense-cedar (*Calocedrus decurrens* (*Torr.*) *Florin*) and Jeffrey pine (*Pinus jeffreyi*)	Interior ponderosa pine	Understory vegetation cover	ALS_D	6.9	154 (40.5)	R^2 = 0.74 bias = 0 RMSE = 0.064 - 0.0735 nRMSE = 22% (Regression)
Hancock et al. (2017)	Luton, England (100 ha)	Woodland, scrubland, and parkland	Urban area	Understory vegetation cover	ALS_{FW}	0.5-4 pulses.m^{-2}	8 (subplot=1.5m)	nRMSE = 24% (Verification at voxel-level)

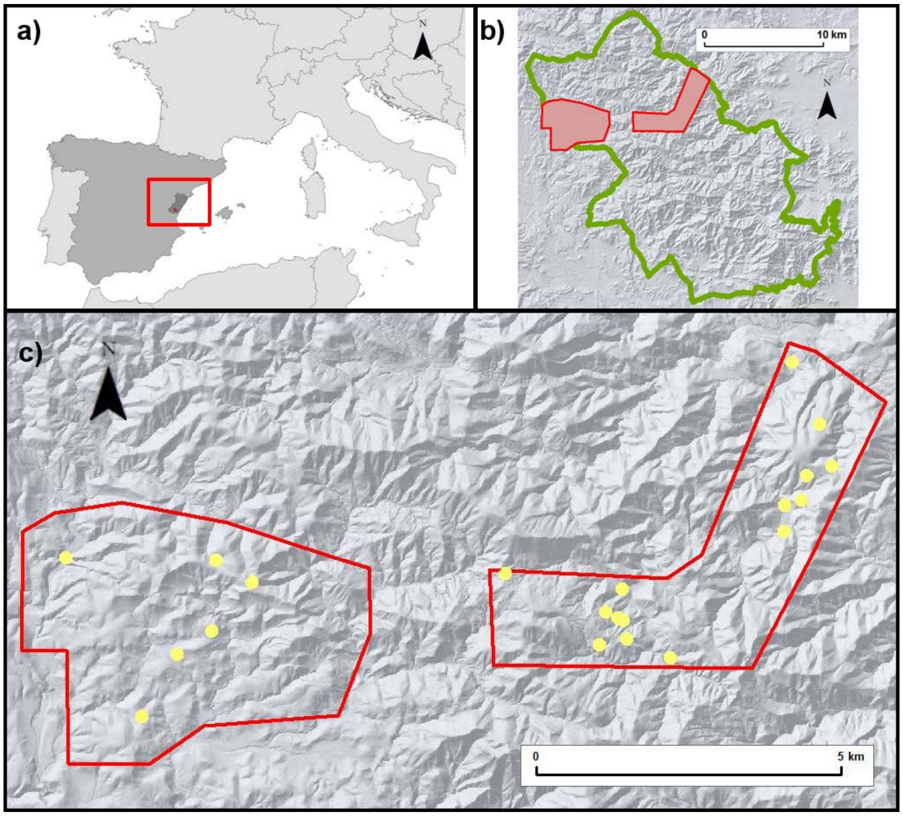


Fig. 1. Study area location in (a) South-Western Europe, (b) Natural Park of Sierra de Espadán (in green), and (c) plot locations (in yellow) within study area.

a) b)

Fig. 2. Field photographs from (a) a maritime pine dominant plot with absence of understory, and (b) an Aleppo pine dominant plot with high presence of understory.

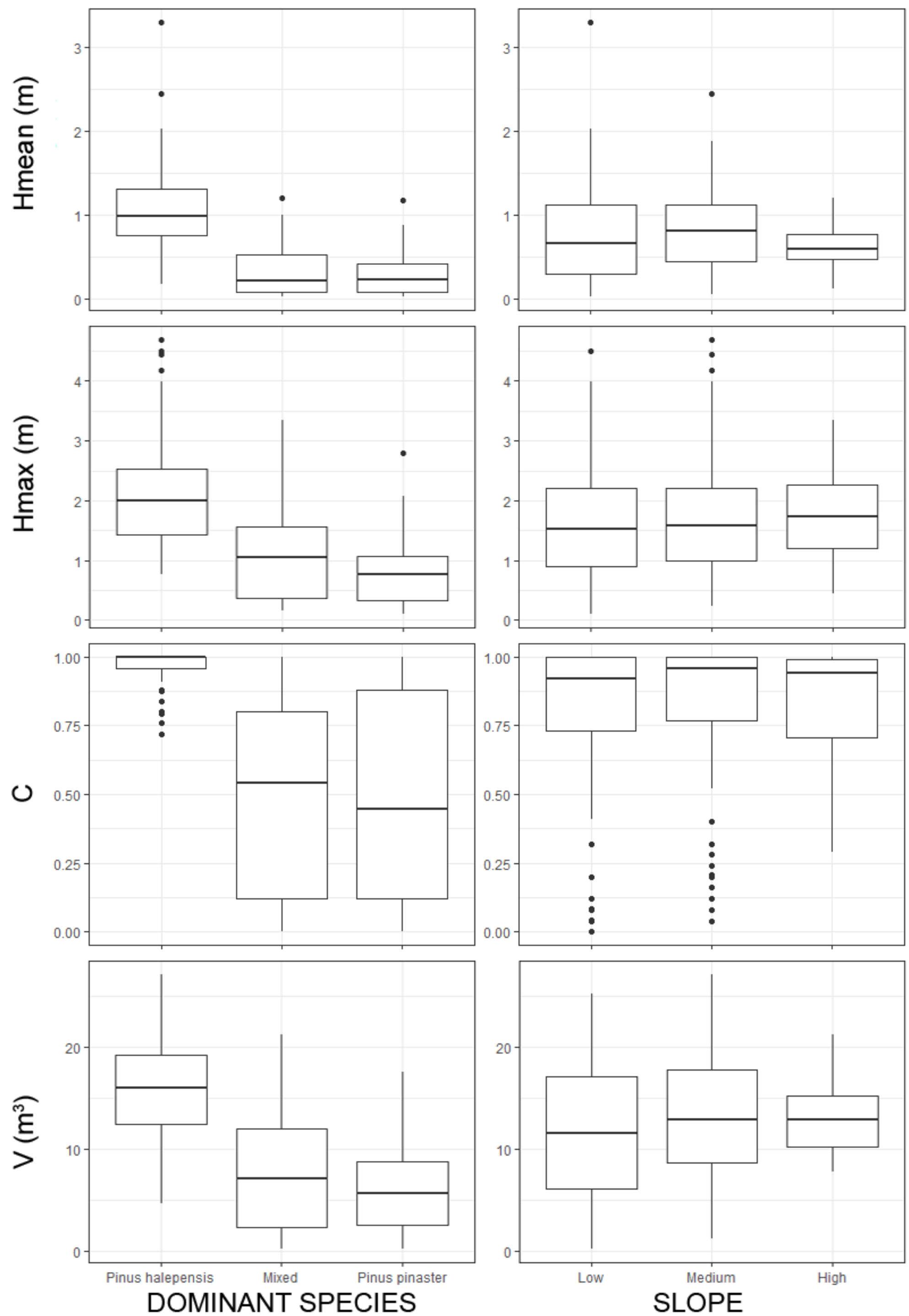


Fig. 3. Box and whiskers representing TLS understory metrics (mean height: H_{mean}, maximum height: H_{max}, cover: C, and volume: V) categorized by dominant species (Pinus halepensis, Mixed Pinus pinaster and Quercus suber, and Pinus pinaster) and slope (low, medium, and high) of the plot.

2. TLS data

TLS data acquisition was undertaken between September 29th 2015 and October 23rd 2015 using a FARO FOCUS 3D 120 phase-based laser scanner (Table 2). Data were acquired in 21, 15 m radius circular plots (area of 706.86 m^2). Plot centers were registered with a GPS Leica RTK 1200+ series receiver with an average accuracy of 0.40 m ± 0.27 m in XY dimension and 0.73 m ± 0.51 m in Z dimension. At each plot nine scans were acquired to minimize occlusion, with one scan in the plot center, four at the edge of the plot at each cardinal direction (N, E, S, W), and four at a distance of 7.5 m from plot center with the directions corresponding to NE, SE, SW, and NW. The total point count of the 9 co-registered point clouds was approximately 100 million, with each return consisting of XYZ coordinates, intensity, plot id, and scan id.

During TLS data acquisition the maximum height of the understory was also assessed at each site by trained forestry staff. This involved measuring the lower crown of the dominant and co-dominant trees, as well as the maximum height of the shrub and understory layer. This information was later used to provide the height threshold between understory and overstory in order to remove overstory point clouds from TLS data described in section 2.4.

Table 2. TLS data specifications.

Specification	Value
Sensor	FARO FOCUS 3D 120
Accuracy	±2 mm at 25 m
Range	0.6 – 120 m

Pulse frequency	97 Hz
Scan angle	Horizontal: 300º Vertical: 360º
Wavelength	905 nm
Beam divergence	0.19 mrad

3. ALS_{FW} data

ALS_{FW} data were acquired on September 16th 2015 over 7,465.53 ha using a LiteMapper 6800 with a pulse density of 14 pulses m^{-2}. Data were acquired at a flight altitude between 600 and 820 m above sea level, at 300 kHz pulse frequency, and with a scan angle of ±37°. The study area was flown over with contiguous flight stripe side-lap between 55% and 77%. After processing, waveforms were provided in a variable number of bins (80-160-240 bins) depending on what height the pulse intercepted the vegetation, with a temporal sample spacing of 1 ns (0.15 m) and a footprint size of 0.24 m. In addition to the ALS_{FW}, the data were also provided in a discrete format (ALS_D), which was later used to create the Digital Terrain Model (DTM). The vertical accuracy of the ALS_D, verified using a set of ground control points located in open and flat areas, was 4.3 cm (RMSE).

4. Data pre-processing

Point heights of the ALS and TLS datasets were normalized using DTMs derived from each of the point clouds. In the case of ALS, classified ground points were provided by the vendor. TLS ground points were classified using a variation of the Axelsson (2000) algorithm implemented in

LAStools (2017; version 171017). DTMs with a resolution of 0.3 m were generated and each dataset was then normalized.

TLS-based metrics characterising the understory require two additional pre-processing steps. First, points registered on tree trunks were removed using a combination of intensity filtering and manual point cloud editing. By examining the TLS point cloud intensity values we found that returns with intensity value higher than 170 can be flagged as tree trunks. Using a point cloud editor, TLS returns adjacent to the trunks were also removed to ensure points located on tree trunks were no longer included in the analysis. In the second pre-processing step returns located above the field-measured maximum understory height were removed (see Fig. 4).

A process described by Hermosilla et al. (2014b) was used to remove waveform noise present in the ALS_{FW} data. First, a noise threshold was defined as the mean plus four times the standard deviation of the waveform (Lefsky et al., 2005). This provided a lower threshold by which all lower waveform information was removed. Next, a Gaussian filter was applied to smooth and remove any remaining noise present in the waveform, where the kernel size was defined by the full width at half-maximum (Duong, 2010; Cao et al., 2014; Hermosilla et al., 2014b). This filter calculates the new amplitude value as the weighted average of the adjacent amplitude values, where the weights depend on the bell shaped Gaussian distribution. The new amplitude values slightly differ from the original ones (Hancock et al., 2015), however, the shape and proportion of the waveform is kept, and therefore ALS_{FW} data values are not highly influenced.

5. Voxelization

Voxelization offers a number of key benefits when dealing with huge amount of data, such as TLS and ALS_{FW} data. It allows the reduction of data volume by clustering lidar return pulses into voxels (e.g. rectangular prisms), and in the case of vegetation it allows the characterization of the amount of space vegetation occupies. The voxel size is defined by the user and depends on the density of the data and the desired level of abstraction. In our case the horizontal size of the voxels was based on the ALS footprint size and pulse density, while the vertical dimension was based on the temporal sample spacing (i.e. 0.15 m). Each voxel was also assigned a maximum amplitude value of the points located inside.

We voxelized both the TLS and ALS datasets using a 0.75 x 0.75 x 0.15 m voxel size (henceforth referred to as 0.75 m) in order to have the lowest number of empty voxels without a loss of accuracy (Crespo-Peremarch et al., 2016). We then characterized the understory at two spatial scales; 3.75 x 3.75 x 0.15 m (i.e. 5 x 5 columns of voxels; henceforth referred to as 3.75 m), which is denoted as "cell-level" of understory vegetation, as well as at the broader plot-level scale (15 m radius).

In case of the ALS_{FW} data, the voxelization had additional purpose and was used to derive pseudo-vertical waveforms (Hermosilla et al., 2014b). Pseudo-vertical waveforms are created using the voxel amplitude values in each column of voxels. These artificial waveforms are used to correct the spatial displacement produced by off-nadir scan angles. In the case of the TLS data, voxelized point clouds were used to calculate understory attributes described below.

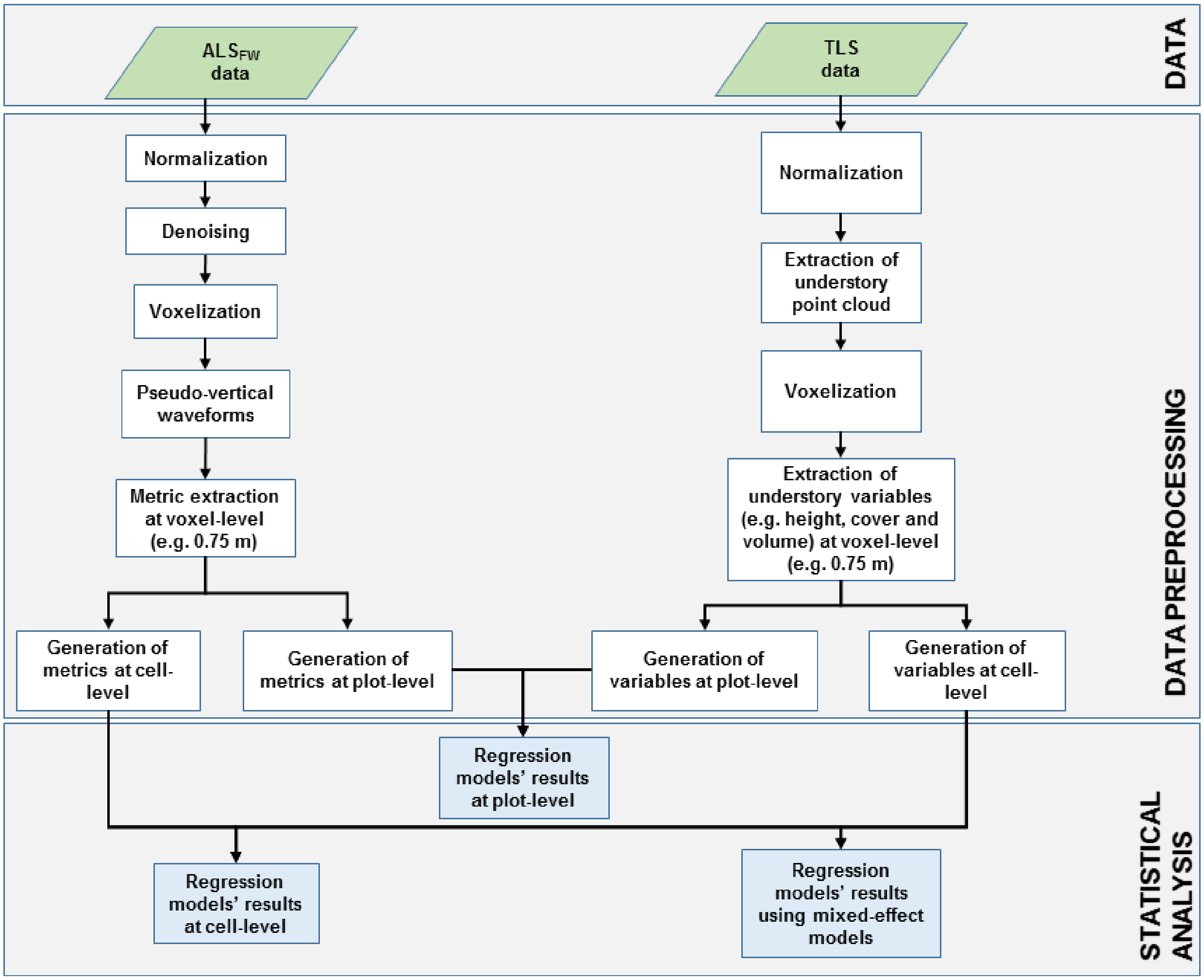


Fig. 4. Flowchart of ALS_{FW} and TLS data processing

1. TLS-based understory attributes

Four key variables describing the understory vegetation were extracted from the TLS voxels: mean understory height (H_{mean}), maximum understory height (H_{max}), understory canopy cover (C) and total volume, which is defined as three-dimensional space occupied by understory (V) (Fig. 4). These four understory attributes were used as the observed variables and modelled with ALS_{FW} derived predictors.

To calculate the H_{max}, we computed the 99% height of each 0.75 m voxel and then extracted the maximum within each 3.75 m side cells (cell-level). H_{mean} was defined as the average of the 99% heights of each 0.75 m across the 3.75 m cells. A proportion of filled voxel columns within each 3.75 m cell was used to describe C. A minimum threshold of 10 points was used to determine filled voxels in each column, and a minimum of one filled voxel was required to define a column as filled. A sum of all filled voxels in each column was used as an estimate of V. Fig. 3 shows these TLS variables categorized by the dominant species and the slope of the plot.

In addition to the cell level (3.75 m), all attributes were also calculated at plot-level (15 m) (see Fig. 4).

2. ALS_{FW} metrics

A suite of ALS_{FW} metrics were used to predict TLS-derived understory attributes. We examined 20 metrics previously described in the literature (Duong, 2010; Duncanson et al., 2010; Zhang et al., 2011) (Table 3) and computed nine more, potentially more suitable for characterizing the understory structure. The 20 previously applied ALS_{FW} metrics are based on (1) return energy, (2) elevation, or (3) extracted from Gaussian iterative decomposition (i.e. optimized using the Levenburg-Marquardt method) (Hofton et al., 2000). The nine new metrics we introduce focus on the lower part of the waveform and include: HFEV (Height at First Empty Voxel) and HFEVT (Height at First Empty Voxel from Threshold), EFEV (Energy to First Empty Voxel), nEFEV (normalized Energy to First Empty Voxel), FVU (Filled Voxel at Understory), NFVU (Number of Filled Voxels at Understory), BC (Bottom of Canopy), BCE (Bottom of Canopy Energy) and BCD (Bottom of Canopy Distance).

HFEV and HFEVT are related to the understory height and analyze the pseudo-vertical waveform in the vertical dimension from the ground upwards. HFEV is computed as the height from the ground to the first filled voxel (defined as an amplitude higher than five (Fig. 5a)).

To account for low shrubs close to the ground and a more open understory, the HFEVT calculates the height of the first filled voxel above 1 m (Fig. 5b). EFEV and nEFEV are related to the properties of the understory. The first attribute is the sum of amplitudes from the ground to the understory height, which corresponds to HFEV. The nEFEV is a relative measure, and is equal to the EFEV divided by the sum of amplitude of the whole waveform. FVU and NFVU are related to understory cover. FVU examines if there are any filled voxels between two given heights (Fig. 5c), and NFVU is the number of filled voxels divided by the number of voxels between these two heights (Fig. 5d). Based on the vegetation in our study site, the lower and upper thresholds were set to 0.15 m and 1 m, respectively.

Name	**Class**	**Description**	**Reference**
WD	Elevation	Waveform distance	Duong, 2010
ROUGH		Roughness of outermost canopy	
Hn		Height at nth percentile of energy	Kimes et al., 2006
RWE	Energy	Return waveform energy	Duong, 2010
MAX E		Maximum energy	Duncanson et al., 2010
VARIANCE		Variance of energy	
SKEWNESS		Skewness of energy	
HEIGHT Qn		Proportion of energy in nth elevation quarter	
ENERGY Qn		Proportion of energy in nth energy quarter	
N GS	Gaussian Iterative Decomposition	Number of Gaussian curves in the waveform	
N GS STARTPEAK		Number of Gaussian curves between the beginning of the waveform and the position of MAX E	
N GS ENDPEAK		Number of Gaussian curves between the position of MAX E and the end of the waveform	
CE		Canopy return energy extracted from canopy Gaussian curves	Zhang et al., 2011
GE		Ground energy extracted from ground Gaussian curve	
GRR		Ground return ration: GE divided by RWE	
CHn		Elevation of nth quarter of energy, excluding ground Gaussian curve	
Rn		CHn divided by WD	
AGS		Average Gaussian curve slope	
SGS		Standard deviation Gaussian curve slope	
MSGS		Modified standard deviation Gaussian curve slope	
HFEV	New	Height at first empty voxel	This study
HFEVT		Height at first empty voxel from threshold	

Table 3.

EFEV	Energy from beginning of the waveform to first empty voxel
nEFEV	Energy from beginning of the waveform to first empty voxel divided by RWE
FVU	Filled voxels at understory
NFVU	Filled voxels at understory divided by number of voxels
BC	Bottom of canopy: elevation of the first canopy Gaussian curve
BCE	Bottom of canopy energy: energy from the beginning of the waveform to BC
BCD	Bottom of canopy distance: distance from BC to the end of the waveform

Description of ALS_{FW} metrics used.

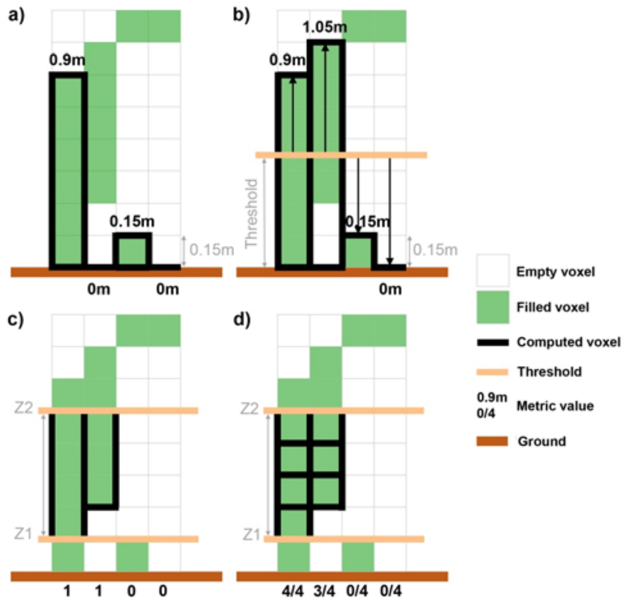


Fig. 5. Graphical representation of voxel transects to describe metrics (a) HFEV, (b) HFEVT, (c) FVU, and (d) NFVU. Voxel height is equal to 0.15 m and metric values for each column of voxels is written in black. Height thresholds in (b), (c), and (d) are user inputs.

Gaussian iterative decomposition metrics were designed by Zhang et al. (2011) for large-footprint lidar, and Hancock et al. (2015) showed that Gaussian iterative fitting was the most accurate method comparing energy values for large-footprint lidar. However, we decided to test the potential of these metrics as descriptors of the understory vegetation, since according to Hancock et al. (2015), energy differences for the Gaussian iterative method and small-footprint lidar were small as well (i.e. nRMSE = 1.37%). The new metrics (BC, BCE, and BCD) are based on Gaussian iterative decomposition described by Zhang et al. (2011) (Fig. 6). Once the derived boundary between the canopy and ground returns is calculated, BC metric is defined as the height from the ground to the first Gaussian curve above the boundary. BCE is the energy from the ground to BC, and BCD is the distance from BC to the top of the canopy. We assumed that the first energy peak excluding the ground must be related to either the understory or the canopy base.

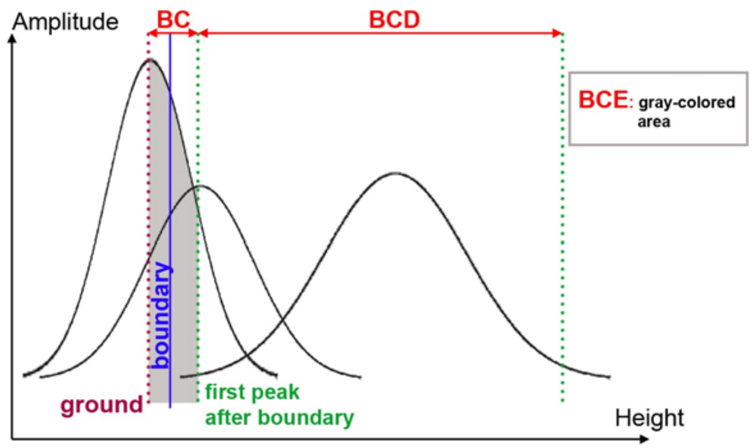


Fig. 6. Graphical representation of metrics BC, BCE, and BCD

To better understand if limiting the calculation of the pseudo-vertical waveform metrics to lower components of the canopy enhance estimations of understory vegetation, we applied a height filter to ALS_{FW} metrics. This height filter consisted of cutting off the pseudo-vertical waveform at a given height threshold. The height threshold for the whole study was computed as 99% height of understory heights extracted from TLS data. We therefore computed the 20, and 9 new, ALS_{FW} metrics on both the full pseudo-vertical waveform as well as a pseudo-vertical waveform limited to the height of the TLS understory height threshold.

As all these metrics were computed for each column of voxels, mean and standard deviation was calculated at the corresponding cell- and plot-level as variables for regression models explained in section 2.6.

6. Regression models

1. Linear regression

We used linear regression to develop predictive models of the four understory attributes, using ALS_{FW} metrics as independent variables. Attribute selection consisted of comparing the Akaike Information Criterion (AIC) (Akaike, 1973) of all possible model comparisons using a maximum of three ALS-derived variables in each model. Each plot was composed of 40 samples (i.e. cells). In order to reduce spatial autocorrelation we randomly sampled 10 samples per plot, which resulted in 210 samples at the cell-level and 21 for the plot-level analysis. A total of 16 model sets were tested (4 understory TLS metrics x 2 resolutions (cell- and plot-level) x 2 sets of ALS-derived metrics (with and without the TLS height filter)). Models were compared using the adjusted coefficient of determination (R^2), root mean square error (RMSE), normalized root mean square error (nRMSE; i.e. RMSE divided by the range of observed values) and coefficient of variation (CV; i.e. RMSE divided by the mean of the observed values). In the case of C, which is a bounded variable between 0 and 1, we replaced linear regression with Beta regression (Ferrari and Cribari-Nieto, 2004) where a pseudo-coefficient of determination (pR^2) was generated for these regression models.

2. Linear mixed effect models

To assess if the ability of ALS to predict the TLS metrics was site dependent, we also undertook a mixed effect modelling approach, which involved developing statistical models containing both fixed and random effects (Crawley, 2012). The two known variables from each plot, slope and dominant species, were used as categorical class variables since both can affect the understory (see

Fig. 3). We categorized the slope in three groups: low, medium, and high. The dominant species were split into three groups as well: H (*Pinus halepensis*), P (*Pinus pinaster*), and M (*Pinus pinaster* + *Quercus suber*). Beatty (1984) found that microrelief could affect nutrient content, making mounds poorer and pits richer in biodiversity. Barbier et al. (2008) found that understory vegetation was highly affected by overstory species, since a number of environmental factors (e.g. light and nutrients) highly influence species. We allowed both the model slope and intercept to vary (based on Gelman and Hill (2007)) while utilizing Nakagawa and Schielzeth's (2013) steps with an update of Johnson's (2014) to calculate two model estimators: marginal R^2 (R^2m) and conditional R^2 (R^2c) for model comparison, as well as standard RMSE and nRMSE for linear mixed effect models. These 24 models (4 TLS understory metrics x 2 full-waveform metrics datasets (with and without height filter) x 3 combination of categorical variables (slope, dominant species, and both)) plus the 16 models explained above, resulted in 40 models in total for this study.

7. Software used

We used LAStools (2017; version 171017) to extract the ground points from TLS and to generate the DTMs. R packages, lidR (Roussel and Auty, 2017) to manage TLS data, and lme4 (Bates et al., 2014) to generate mixed-effect models, were used. In addition, we also used our own software to process and generate ALS_{FW} data.

3. Results

The detection pR^2 of the understory cover (C) was 0.871. The R^2 values of the predicted understory attributes were 0.957, 0.771, and 0.951, for H_{mean}, H_{max}, and V, respectively.

Fig. 7 shows an example of the four TLS and ALS_{FW} derived metrics of the understory with a site photograph for three plots within the study area. These three characteristic plots demonstrate low, moderate, and high degrees of understory cover (i.e. plots id 28, 31, and 7, respectively).

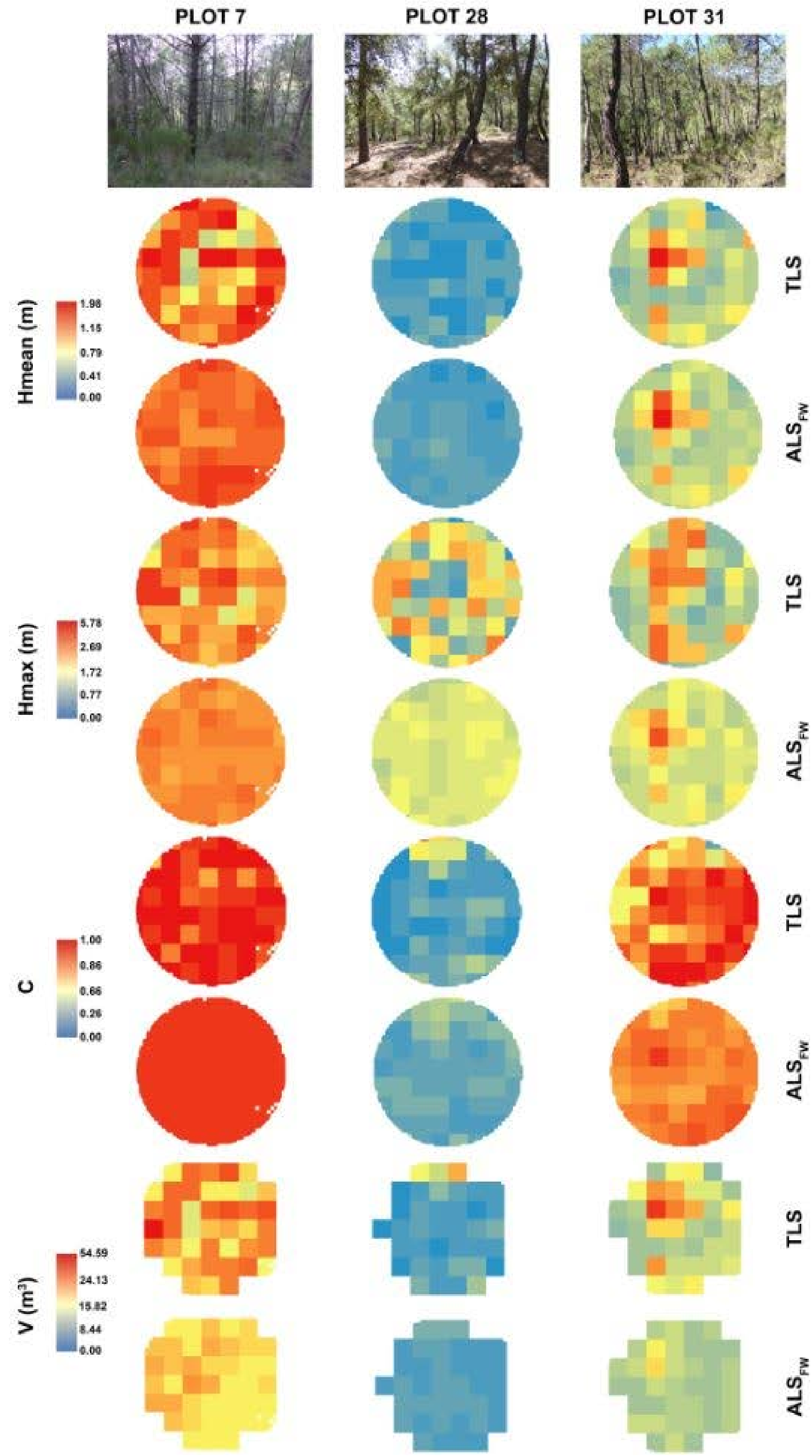
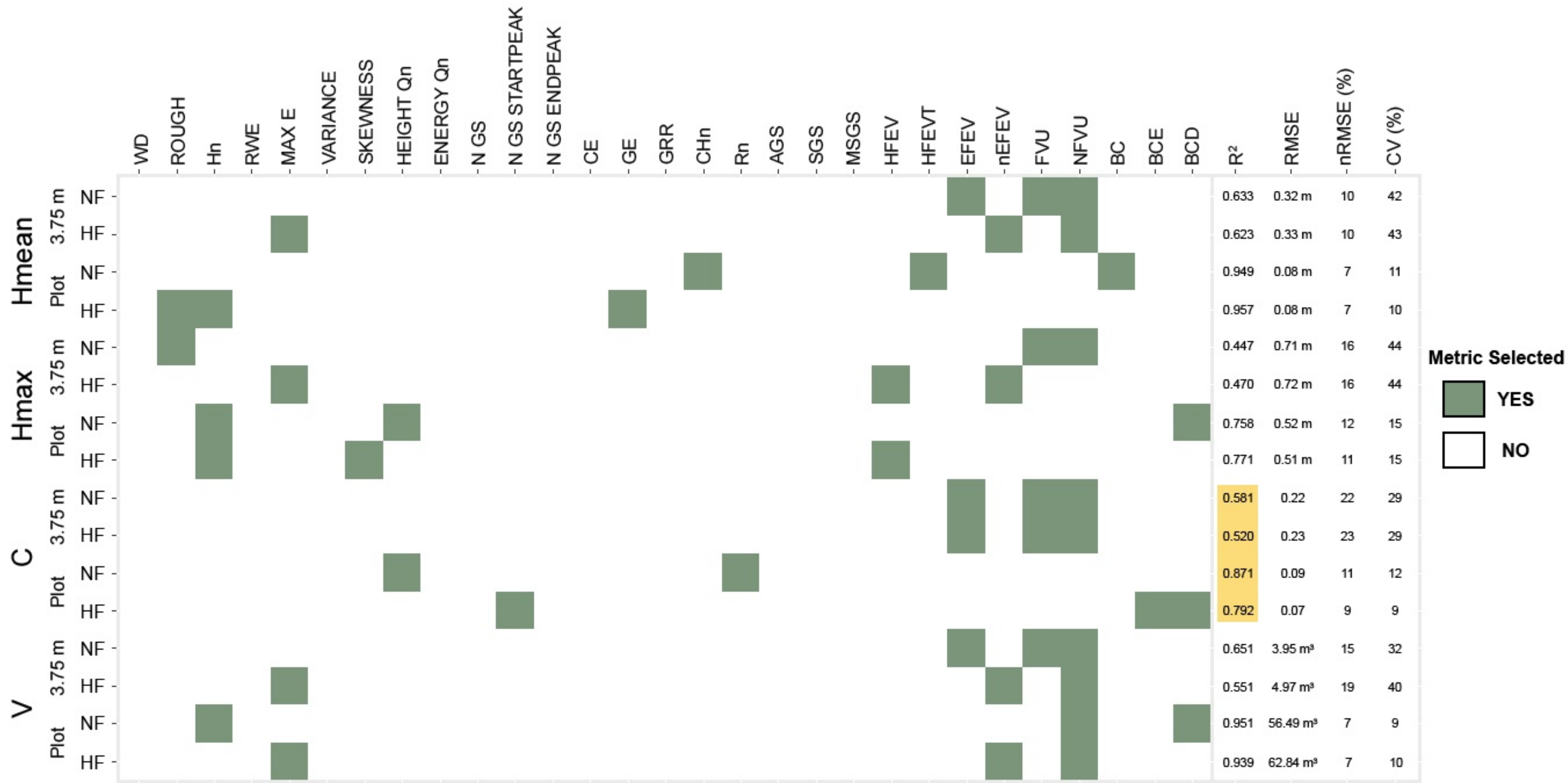


Fig. 7. TLS and ALS_{FW} derived four metrics (H_{mean}, H_{max}, C and V) and field photographs extracted from three plots (id 7, 28, and 31) with 15 m radius within the study area. Plots id 28, 31, and 7, represent low, moderate, and high degrees of understory cover, respectively.

Table 4 shows the ALS_{FW} metrics selected for the 16 regression models (4 understory TLS metrics x 2 resolutions (cell- and plot-level) x 2 set of full-waveform metrics (with and without the TLS height filter)) with corresponding R^2, RMSE, nRMSE and CV values. Results indicate that the best model for H_{mean} and H_{max} was developed at the plot-level using a height filter, and had R^2 values of 0.957 and 0.771, respectively. These models also had the lowest RMSE and nRMSE (0.08 m and 7% for H_{mean}; 0.51 m, and 11% for H_{max}, respectively). The best model for C was also developed at the plot-level, with similar results with and without a height filter. Model performance was characterized by R^2 = 0.871, RMSE = 0.09, nRMSE = 11%, CV = 12% when the height filter was used, and by R^2 = 0.792, RMSE = 0.07, nRMSE = 9%, CV = 9% without the height filter. Lastly, the plot-level model for V, without a height filter, was the most accurate and had R^2 = 0.951, RMSE = 56.49 m^3, nRMSE = 7%, and CV = 9%. Among all models, H_{max} modeled at cell-level had the lowest accuracy with a R^2 of 0.447.

The most frequently used metrics in the regression models included NFVU, FVU, nEFEV, EFEV, Hn, and MAX E, while WD, RWE, VARIANCE, ENERGY Qn, N GS, N GS ENDPEAK, CE, GRR, AGS, SGS, and MSGS were not included in any of the models.

Table 4. ALS_{FW} metrics selected for the estimation of the different variables (H_{mean}, H_{max}, C, and V) for cell- (3.75 m resolution) and plot-level (15 m radius) resolution, and for each height filter (NF: no filter, HF: height filter). The results from these regression models, as well as R^2 values and pseudo-R^2 (orange highlighted), are also included.



Results of the mixed-effect models that incorporated different combinations of categorical variables (slope, dominant species, and both) are shown in Table 5. These results indicate that the highest accuracy was achieved for H_{mean}, with a nRMSE of 9%, for the model that used both categorical variables, as well as for the model that used only the dominant species. For all the understory variables, using just the dominant species or both variables as categorical variables reached the best results.

When compared to the results of the linear regression models (Table 4), all understory variables were predicted with higher accuracy. The improvement in nRMSE was about 1% for H_{mean}, 2% for H_{max}, 7% for C, and 2-3% for V.

Table 5. Results of mixed-effect models for the estimation of the four understory variables (H_{mean}, H_{max}, C, and V).

Categorical Variable	Variable	Height Filter	R^2m	R^2c	RMSE	nRMSE (%)	CV (%)
Slope	Hmean	NO	0.271	0.847	0.31 m	10	41
		YES	0.625	0.627	0.33 m	10	43
	Hmax	NO	0.344	0.550	0.67 m	15	42
		YES	0.433	0.519	0.70 m	15	43
	C	NO	0.466	0.670	0.21	21	27
		YES	0.238	0.793	0.21	21	26
	V	NO	0.311	0.849	3.85 m^3	14	31
		YES	0.068	0.943	4.58 m^3	17	37
Dominant Species	Hmean	NO	0.394	0.666	0.30 m	9	40
		YES	0.526	0.606	0.31 m	10	41
	Hmax	NO	0.294	0.421	0.67 m	15	41
		YES	0.397	0.575	0.67 m	15	41
	C	NO	0.055	0.960	0.17	17	22
		YES	0.059	0.946	0.17	17	22
	V	NO	0.191	0.876	3.61 m^3	13	29

		YES	0.110	0.898	4.49 m^3	17	36
Slope + Dominant Species	Hmean	NO	0.232	0.791	0.30 m	9	39
		YES	0.260	0.780	0.31 m	9	41
	Hmax	NO	0.157	0.613	0.64 m	14	40
		YES	0.145	0.745	0.66 m	14	41
	C	NO	0.032	0.972	0.15	15	20
		YES	0.036	0.961	0.16	16	20
	V	NO	0.118	0.914	3.55 m^3	13	29
		YES	0.035	0.967	4.26 m^3	16	34

Fig. 8 shows scatter plots of the TLS-based observed and ALS-based predicted variables at cell- and plot-level, as well as using the mixed-effect models. Predictions of H_{mean}, H_{max}, and V to their respective observations were closer to 1:1 than C at the cell-level and when using mixed-effect models. Improvement between cell-level and mixed-effect models is especially visible for C. As demonstrated previously, results at the plot-level were more accurate than at the cell-level.

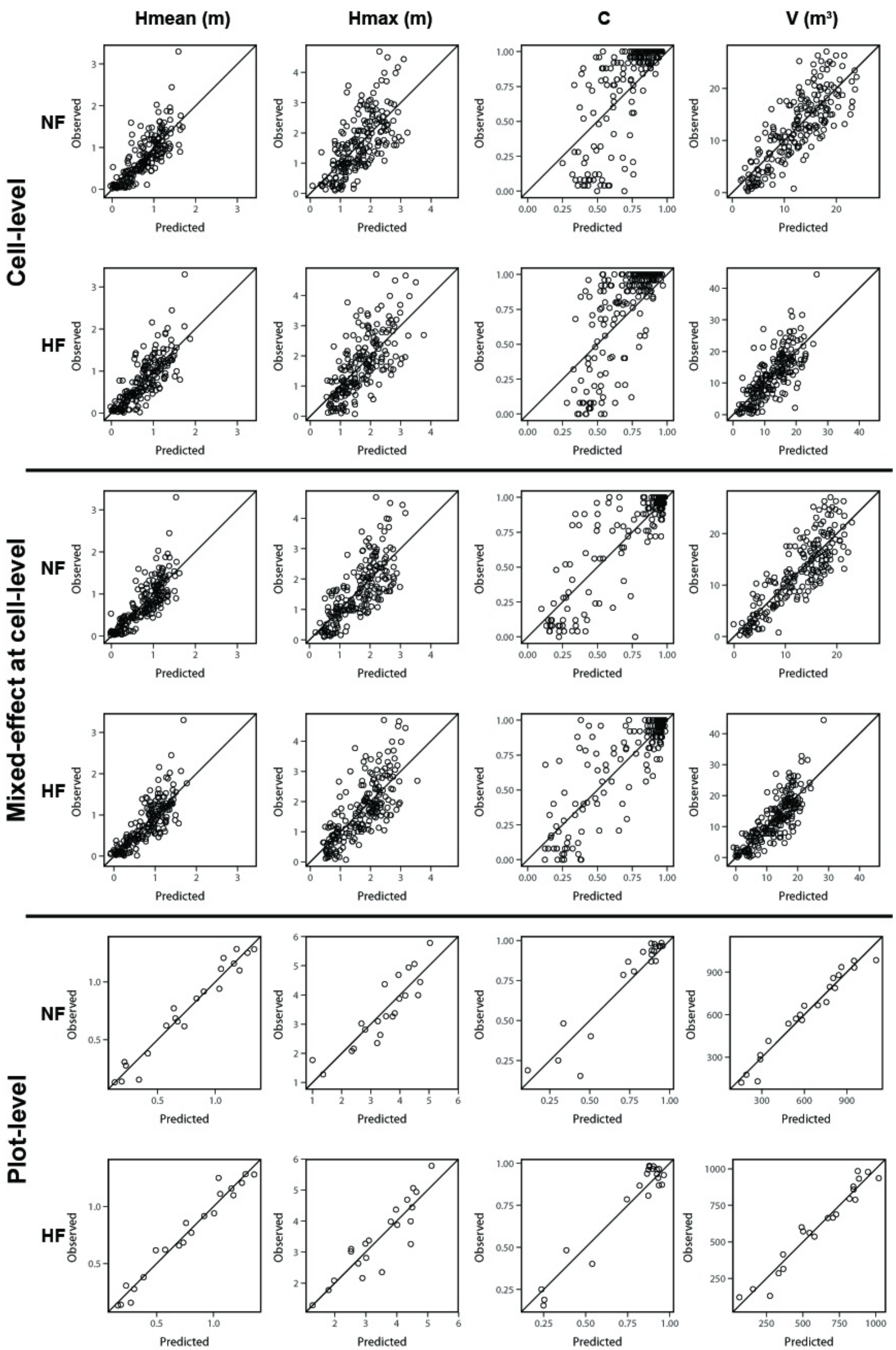


Fig. 8. Regression graphs for the estimation of the different variables (H_{mean}, H_{max}, C and V) for each resolution (cell-level, mixed-effect (cell-level) and plot-level (15 m radius)) and for each height filter (NF: no filter, HF: height filter). Solid line represents the 1:1 line.

4. Discussion

In this research we developed a new methodology to characterize vegetation understory from ALS data, verifying it with TLS data acquired at key plot locations. Key results from this study indicate that understory cover, height, and volume were accurately predicted from ALS_{FW} at both the cell and plot scale when compared to the reference.

Overall, the results showed a high performance of ALS_{FW} for estimating H_{mean}, H_{max}, C, and V, especially at plot-level. H_{mean} and V were modeled with highest accuracy, while poorer results were obtained for C and H_{max}. These results suggest that H_{mean} had a higher performance than H_{max} since mean values are smoother than maximum values, due to the latter being able to have extreme values. V results were close to H_{mean}, given that both variables are directly related. Most of the C training values were close to 1, hence not being a distributed sample, causing poorer estimates of C. A possible solution to improve C estimate results is to increase the number of plots with an intermediate understory cover. Results at the cell-level were poorer since estimates were more sensitive to small changes due to the finer scale. Although results were lower at cell-level, these values were acceptable having in mind its resolution.

A number of key findings were apparent. We applied a height filter in order to determine whether cutting off the pseudo-vertical waveform fragment that corresponds to understory enhanced estimations of understory vegetation characterization. Nevertheless, applying this filter to the ALS_{FW} prior to metric calculation did not result in an improvement in accuracy when predicting H_{mean} at cell-level, as well as C and V at both scales. In addition, in those cases where results from height filter tests were higher, improvements compared to no height filter tests were small. This is likely due to the fact that contrary to ALS_D, which has a limited number of digitized returns, ALS_{FW} can fully discriminate height strata through decomposing the waveform. As a result height thresholds for data processing are not needed.

Estimation results of understory cover, height, and volume improved when mixed-effect models were applied using just the dominant species as variable, or combined with the slope. These results suggest that terrain slope alone has little influence on the prediction of the understory variables, however when combined with dominant species it has a more significant effect.

With respect to the accuracy of the predictions, our results correspond to those of others (Martinuzzi et al., 2009; Hill and Broughton, 2009; Morsdorf et al., 2010; Wing et al., 2012; Hancock et al., 2017). Most of the studies to date (Martinuzzi et al., 2009; Hill and Broughton, 2009; Morsdorf et al., 2010) have estimated the presence or absence of understory by applying a classification based approach. Contrastingly, Wing et al. (2012) estimated understory cover using regression models and found a coefficient of determination (R^2) of 0.74, with a similar nRMSE as reported in our study (nRMSE = 22%), but used a resolution of 40.5 m^2 and applyed height and intensity filters. This study suggests that ALS_{FW} can be used to estimate understory cover with a similar nRMSE, but with a higher resolution (i.e. 3.75 m or 14.06 m^2) and without applying any filter. Alternatively, Hancock et al. (2017) obtained a similar accuracy (nRMSE = 24%) at finer scale (1.5 m horizontal and 0.5 m vertical resolution), but in an urban landscape. This suggests that understory cover can be extracted more accurately in urban environments, where vegetation is likely more intensively managed by humans.

Scaling from the cell-level to the full plot showed an increase in accuracy and decrease in error when compared to the reference TLS predictions. In the case of H_{mean}, the R^2 coefficient increased from 0.633 to 0.949, and from 0.447 to 0.758 for H_{max}. The R^2 coefficient for C increased from 0.581 to 0.871, and from 0.651 to 0.951 for V. From a modelling point of view, the most selected attributes were those developed in this research, especially at the finer scale. The newly created attributes were also used more frequently in the regression models at the plot scale, but they were selected by fewer models. Attributes from Gaussian iterative decomposition related to return energy were not selected, except for BCE. As Hancock et al. (2015) suggested, Gaussian iterative decomposition methods were poorer when extracting return energy from ALS_{FW} when a small-footprint is used because of the increased heterogeneity of the targets. Other methods such as the sum of waveform amplitude and spline may be used in further studies instead of the Gaussian iterative decomposition, since they are less time consuming and robust (Hancock et al., 2015).

H_{mean}, H_{max}, C, and V can be represented as four layers that can be used in three key ways for fire behavior assessment. First, fire models need understory height. These layers give an accurate height that, with the canopy base height measure, can be used to calculate the gap between understory and overstory. This gap is critically important for Mediterranean forests as it describes when a surface fire will likely become a crown fire (e.g. fuel ladder fires). Second, fire behavior depends on understory cover. Surface fire intensity is higher with larger amounts of understory, which is determined by cover and biomass. The latter of which was not able to be predicted in this study, since ground-based data from understory species registered by TLS were not available, as well as the lack of allometric equations for these species to predict biomass. Third, forest clearing in the Mediterranean for fire prevention consists of removing understory vegetation and creating controlled fires. Knowing the understory vegetation volume easily allows determination of how much volume will be removed during a fire, which can also be converted to biomass for other purposes.

5. Conclusions

This study presented a method to characterize the understory vegetation through ALS_{FW} data in a Mediterranean forest. Our results suggest that the use of ALS_{FW} provides an alternative to traditional or local techniques for understory characterization. ALS_{FW} is able to accurately estimate understory vegetation variables such as height, cover, and volume over large areas. These variables reached very high R^2 values at plot scale (mean height: R^2 = 0.957, maximum height: R^2 = 0.771, cover: R^2 = 0.871, and volume: R^2 = 0.951), but were slightly lower at cell-level (i.e. 3.75 m side) (mean height: R^2 = 0.633, maximum height: R^2 = 0.470, cover: R^2 = 0.581, and volume: R^2 = 0.651). The new proposed metrics proved to be decisive for a more accurate characterization of the understory vegetation. This is an advantage to traditional or TLS techniques, which can only be collected in small areas and tend to be very costly. The results presented in this study are particularly important for forest management, as well as fire prevention and prediction. Further studies must be conducted in different ecosystems in order to assess the potential use of ALS_{FW}

for various tree and shrub densities and types, as well as predicting other variables such as biomass, which is essential to analyze forest fire intensity.

6. Acknowledgments

This research was developed mainly in the Integrated Remote Sensing Studio (IRSS) of University of British Columbia (UBC) (Canada) as a result of the Erasmus+ KA-107 mobility grant. The authors thank the financial support provided by the Spanish Ministerio de Economía y Competitividad and FEDER, in the framework of the project CGL2016-80705-R.

7. References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Proceedings of the 2nd International Symposium on Information, BN Petrow, F. Czaki, Akademiai Kiado, Budapest.

Anderson, K., Hancock, S., Disney, M., Gaston, K.J., 2016. Is waveform worth it? A comparison of LiDAR approaches for vegetation and landscape characterization. Remote Sensing in Ecology and Conservation. 2(1), 5-15.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing. 33(B4/1; PART 4), 111-118.

Barbier, S., Gosselin, F., Balandier, P., 2008. Influence of tree species on understory vegetation diversity and mechanisms involved—a critical review for temperate and boreal forests. Forest ecology and management. 254(1), 1-15. https://doi.org/10.1016/j.foreco.2007.09.038

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. Linear mixed-effects models using Eigen and S4. http://CRAN.R-project.org/package=lme4

Beatty, S. W., 1984. Influence of microtopography and canopy species on spatial patterns of forest understory plants. Ecology. 65(5), 1406-1419. https://doi.org/10.2307/1939121

Canfield, R. H., 1941. Application of the line interception method in sampling range vegetation. Journal of Forestry. 39(4), 388-394.

Cao, L., Coops, N. C., Hermosilla, T., Innes, J., Dai, J., She, G., 2014. Using small-footprint discrete and full-waveform airborne LiDAR metrics to estimate total biomass and biomass components in subtropical forests. Remote Sensing. 6(8), 7110-7135. https://doi.org/10.3390/rs6087110

Chasmer, L., Hopkinson, C., Treitz, P., 2006. Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Canadian Journal of Remote Sensing. 32(2), 116-125. https://doi.org/10.5589/m06-011

Crawley, M. J., 2012. Mixed-Effects Models. The R Book, Second Edition, 681-714.

Crespo-Peremarch, P., Ruiz, L. A., Balaguer-Beser, A., Estornell, J., 2016. Analysis of the Side-Lap Effect on Full-Waveform LIDAR Data Acquisition for the Estimation of Forest Structure Variables. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 603-610. https://doi.org/10.5194/isprs-archives-xli-b8-603-2016

Crespo-Peremarch, P., Ruiz, L. A., 2017. Análisis comparativo del potencial del ALS y TLS en la caracterización estructural de la masa forestal basado en voxelización. Actas XVII Congreso de la Asociación Española de Teledetección. Nuevas plataformas y sensores de teledetección. 131-135.

Duncanson, L. I., Niemann, K. O., Wulder, M. A., 2010. Estimating forest canopy height and terrain relief from GLAS waveform metrics. Remote Sensing of Environment. 114(1), 138-154. https://doi.org/10.1016/j.rse.2009.08.018

Duong, V. H., 2010. Processing and application of ICESat large footprint full waveform laser range data (Doctoral dissertation, TU Delft, Delft University of Technology).

Estornell, J., Ruiz, L. A., Velázquez-Marti, B., 2011. Study of shrub cover and height using LiDAR data in a Mediterranean area. Forest Science. 57(3), 171-179.

Ferrari, S., Cribari-Nieto, F., 2004. Beta regression for modelling rates and proportions. Journal of Applied Statistics. 31(7), 799-815. https://doi.org/10.1080/0266476042000214501

Fieber, K.D., Davenport, I.J., Tanase, M.A., Ferryman, J.M., Gurney, R.J., Becerra, V.M., Walker, J.P., Hacker, J.M., 2015. Validation of Canopy Height Profile methodology for small-footprint full-waveform airborne LiDAR data in a discontinuous canopy environment. ISPRS Journal of Photogrammetry and Remote Sensing. 104, 144-157.

Gelman, A., Hill, J., 2007. Data analysis using regression and multilevel/hierarchical models. New York, NY, USA: Cambridge university press. https://doi.org/10.1017/cbo9780511790942.014

Greaves, H. E., Vierling, L. A., Eitel, J. U., Boelman, N. T., Magney, T. S., Prager, C. M., Griffin, K. L., 2015. Estimating aboveground biomass and leaf area of low-stature Arctic shrubs with terrestrial LiDAR. Remote Sensing of Environment. 164, 26-35. https://doi.org/10.1016/j.rse.2015.02.023

Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P., Disney, M., Danson, F.M., Strahler, A., Schaaf, C., Anderson, K., Gaston, K.J., 2015. Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. Remote Sensing of Environment. 164, 208-224. https://doi.org/10.1016/j.rse.2015.04.013

Hancock, S., Anderson, K., Disney, M., Gaston, K.J., 2017. Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. Remote Sensing of Environment. 188, 37-50. https://doi.org/10.1016/j.rse.2016.10.041

Harding, D.J., Lefsky, M.A., Parker, G.G., Blair, J.B., 2001. Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. Remote Sensing of Environment. 76(3), 283-297.

Hermosilla, T., Ruiz, L.A., Kazakova, A.N., Coops, N.C., Moskal, M., 2014a. Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. International Journal of Wildland Fire. 23(2), 224-233. https://doi.org/10.1071/wf13086

Hermosilla, T., Coops, N. C., Ruiz, L. A., Moskal, L. M., 2014b. Deriving pseudo-vertical waveforms from small-footprint full-waveform LiDAR data. Remote sensing letters. 5(4), 332-341. https://doi.org/10.1080/2150704x.2014.903350

Hilker, T., van Leeuwen, M., Coops, N. C., Wulder, M. A., Newnham, G. J., Jupp, D. L., Culvenor, D. S., 2010. Comparing canopy metrics derived from terrestrial and airborne laser scanning in a Douglas-fir dominated forest stand. Trees. 24(5), 819-832. https://doi.org/10.1007/s00468-010-0452-7

Hill, R. A., Broughton, R. K., 2009. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. ISPRS Journal of Photogrammetry and Remote Sensing. 64(2), 223-233. https://doi.org/10.1016/j.isprsjprs.2008.12.004

Hofton, M.A., Minster, J.B., Blair, J.B., 2000. Decomposition of laser altimeter waveforms. IEEE Transactions on Geoscience and Remote Sensing. 38(4), 1989-1996. https://doi.org/10.1109/36.851780

Hopkinson, C., Lowell, J., Chasmer, L., Jupp, D., Kljun, N., van Gorsel, E., 2013. Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. Remote Sensing of Environment. 136, 301-314. https://doi.org/10.1016/j.rse.2013.05.012

Johnson, P. C., 2014. Extension of Nakagawa and Schielzeth's R2GLMM to random slopes models. Methods in Ecology and Evolution. 5(9), 944-946. https://doi.org/10.1111/2041-210x.12225

Kerr, J. T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. Trends in Ecology & Evolution. 18(6), 299-305. https://doi.org/10.1016/s0169-5347(03)00071-5

Kobal, M., Bertoncelj, I., Pirotti, F., Dakskobler, I., Kutnar, L., 2015. Using Lidar data to analyse sinkhole characteristics relevant for understory vegetation under forest cover – case study of a high karst area in the Dinaric Mountains. PloS one. 10(3), e0122070. http://doi.org/10.1371/journal.pone.0122070

Kimes, D.S., Ranson, K.J., Sun, G., Blair, J.B., 2006. Predicting lidar measured forest vertical structure from multi-angle spectral data. Remote Sensing of Environment. 100(4), 503-511. https://doi.org/10.1016/j.rse.2005.11.004

LAStools, 2017. Efficient LiDAR Processing Software (version 171017, academic), obtained from http://rapidlasso.com/LAStools.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., Hunter, M.O., de Oliveira Jr., R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. Geophysical Research Letters. 32(22). https://doi.org/10.1029/2005gl023971

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., Vastaranta, M., 2016. Terrestrial laser scanning in forest inventories. ISPRS Journal of Photogrammetry and Remote Sensing. 115, 63-77. https://doi.org/10.1016/j.isprsjprs.2016.01.006

Maas, H. G., Bienert, A., Scheller, S., Keane, E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. International journal of remote sensing. 29(5), 1579-1593. https://doi.org/10.1080/01431160701736406

Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., Vierling, K. T., 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sensing of Environment. 113(12), 2533-2546. https://doi.org/10.1016/j.rse.2009.07.002

McDermid, G. J., Franklin, S. E., LeDrew, E. F., 2005. Remote sensing for large-area habitat mapping. Progress in Physical Geography. 29(4), 449-474. https://doi.org/10.1191/0309133305pp455ra

Molina, J. R., Rodríguez y Silva, F., Herrera, M. A., Zamora, R., 2009. A simulation tool for socio-economic planning on forest fire suppression management. Forest Fires Detection, Suppression and Prevention. Nova Science Publishers, New York, USA, 33-88.

Molina, J. R., y Silva, F. R., Herrera, M. A., 2011. Potential crown fire behavior in Pinus pinea stands following different fuel treatments. Forest Systems. 20(2), 266-267. https://doi.org/10.5424/fs/2011202-10923

Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., Allgöwer, B., 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(7), 1403-1415. https://doi.org/10.1016/j.rse.2010.01.023

Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution. 4(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Riaño, D., Chuvieco, E., Ustin, S. L., Salas, J., Rodríguez-Pérez, J. R., Ribeiro, L. M., Fernández, H., 2007. Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. International Journal of Wildland Fire. 16(3), 341-348. https://doi.org/10.1071/wf06003

Royo, A. A., Carson, W. P., 2006. On the formation of dense understory layers in forests worldwide: consequences and implications for forest dynamics, biodiversity, and succession. Canadian Journal of Forest Research. 36(6), 1345-1362. https://doi.org/10.1139/x06-025

Roussel, J-R., Auty, D., 2017. lidR: Airborne LiDAR data manipulation and visualization for forestry applications. http://CRAN.R-project.org/package=lidR

Ruiz, L. A., Recio, J. A., Crespo-Peremarch, P., Sapena, M., 2018. An object-based approach for mapping forest structural types based on low-density LiDAR and multispectral imagery. Geocarto International. 33(5), 443-457. https://doi.org/10.1080/10106049.2016.1265595

Suchar, V. A., Crookston, N. L., 2010. Understory cover and biomass indices predictions for forest ecosystems of the Northwestern United States. Ecological Indicators. 10(3), 602-609. https://doi.org/10.1016/j.ecolind.2009.10.004

Vierling, L. A., Xu, Y., Eitel, J. U., Oldow, J. S., 2013. Shrub characterization using terrestrial laser scanning and implications for airborne LiDAR assessment. Canadian Journal of Remote Sensing. 38(6), 709-722. https://doi.org/10.5589/m12-057

Wing, B. M., Ritchie, M. W., Boston, K., Cohen, W. B., Gitelman, A., Olsen, M. J., 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. Remote Sensing of Environment. 124, 730-741. https://doi.org/10.1016/j.rse.2012.06.024

Wulder, M., Franklin, S. E., 2012. Remote sensing of forest environments: concepts and case studies. Springer Science & Business Media.

Zhang, J., de Gier, A., Xing, Y., Sohn, G., 2011. Full waveform-based analysis for forest type information derivation from large footprint spaceborne lidar data. Photogrammetric Engineering & Remote Sensing. 77(3), 281-290. https://doi.org/10.14358/pers.77.3.281

List of figure captions

Name	**Caption**
Figure 1	Study area location in (a) South-Western Europe, (b) Natural Park of Sierra de Espadán (in green), and (c) plot locations (in yellow) within study area.

Figure 2	Field photographs from (a) a maritime pine dominant plot with absence of understory, and (b) an Aleppo pine dominant plot with high presence of understory.
Figure 3	Box and whiskers representing TLS understory metrics (mean height: H_{mean}, maximum height: H_{max}, cover: C, and volume: V) categorized by dominant species (Pinus halepensis, Mixed Pinus pinaster and Quercus suber, and Pinus pinaster) and slope (low, medium, and high) of the plot.
Figure 4	Flowchart of ALS_{FW} and TLS data processing
Figure 5	Graphical representation of voxel transects to describe metrics (a) HFEV, (b) HFEVT, (c) FVU, and (d) NFVU. Voxel height is equal to 0.15 m and metric values for each column of voxels is written in black. Height thresholds in (b), (c), and (d) are user inputs.
Figure 6	Graphical representation of metrics BC, BCE, and BCD
Figure 7	TLS and ALS_{FW} derived four metrics (H_{mean}, H_{max}, C and V) and field photographs extracted from three plots (id 7, 28, and 31) with 15 m radius within the study area. Plots id 28, 31, and 7, represent low, moderate, and high degrees of understory cover, respectively.
Figure 8	Regression graphs for the estimation of the different variables (H_{mean}, H_{max}, C and V) for each resolution (cell-level, mixed-effect (cell-level) and plot-level (15 m radius)) and for each height filter (NF: no filter, HF: height filter). Solid line represents the 1:1 line.

Table 4 ALS_{FW} metrics selected for the estimation of the different variables (H_{mean}, H_{max}, C, and V) for cell- (3.75 m resolution) and plot-level (15 m radius) resolution, and for each height filter (NF: no filter, HF: height filter). The results from these regression models, as well as R^2 values and pseudo-R^2 (orange highlighted), are also included.