Pablo
Crespo-Peremarch

Processing and analysis of airborne full-waveform laser scanning data for the characterization of forest structure and fuel properties

PABLO CRESPO-PEREMARCH

Advisor: **Luis Ángel Ruiz Fernández**

PROCESSING AND ANALYSIS OF AIRBORNE FULL-WAVEFORM LASER SCANNING
DATA FOR THE CHARACTERIZATION OF FOREST STRUCTURE AND FUEL PROPERTIES

2020

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

September 2020

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Processing and analysis of airborne full-waveform laser scanning data for the characterization of forest structure and fuel properties

---

Pablo Crespo-Peremarch

Advisor: Luis Ángel Ruiz Fernández

Geo-Environmental Cartography and Remote Sensing Group (CGAT)
Department of Cartographic Engineering, Geodesy and Photogrammetry
Universitat Politècnica de València (UPV)

PhD in Geomatics Engineering

València, September 2020

This PhD Thesis was developed as a compendium of the following scientific papers:

1. **Crespo-Peremarch, P.**, Tompalski, P., Coops, N.C., Ruiz, L.Á., 2018. Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data. Remote Sensing of Environment, 217, 400–413. https://doi.org/10.1016/j.rse.2018.08.033
2. **Crespo-Peremarch, P.**, Ruiz, L.Á., Balaguer-Beser, Á., Estornell, J., 2018. Analyzing the role of pulse density and voxelization parameters on full-waveform LiDAR-derived metrics. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 453–464. https://doi.org/10.1016/j.isprsjprs.2018.10.012
3. **Crespo-Peremarch, P.**, Ruiz, L.Á., 2020. A full-waveform airborne laser scanning metric extraction tool for forest structure modelling. Do scan angle and radiometric correction matter? Remote Sensing, 12, 292. https://doi.org/10.3390/rs12020292
4. **Crespo-Peremarch, P.**, Fournier, R.A., Nguyen, V.-T., van Lier, O.R., Ruiz, L.Á., 2020. A comparative assessment of the vertical distribution of forest components using full-waveform airborne, discrete airborne and discrete terrestrial laser scanning data. Forest Ecology and Management, 473, 118268. https://doi.org/10.1016/j.foreco.2020.118268

Regarding the author rights for scholarly purposes, the published scientific papers (1), (2) and (4) are re-used by the author of portions or excerpts in other works (i.e., the present PhD thesis), with full acknowledgement of the original publications of the scientific papers.

In addition, this PhD Thesis was completed with the following scientific papers or conference presentations:

1. **Crespo-Peremarch, P.**, Ruiz, L.Á., Balaguer-Beser, Á., Estornell, J., 2015. Análisis temporal de la estructural forestal mediante métricas derivadas de LiDAR full-waveform. Actas XVI Congreso de la Asociación Española de Teledetección. Teledetección, humedales y espacios protegidos, 387–390. 21-23 Oct., Sevilla, Spain.

2. **Crespo-Peremarch, P.**, Ruiz, L.Á., Balaguer-Beser, Á., 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 & Spatial Information Sciences, XLI-B8, 603–610. 12-19 July, Prague, Czech Republic. https://doi.org/10.5194/isprs-archives-XLI-B8-603-2016

3. **Crespo-Peremarch, P.,** Ruiz, L.Á., Balaguer-Beser, Á., 2016. A comparative study of regression methods to predict forest structure and canopy fuel variable from LiDAR full-waveform data. Revista de Teledetección, 45, 27-40. https://doi.org/10.4995/raet.2016.4066

4. **Crespo-Peremarch, P.**, Ruiz, L.Á., 2016. Estimating forest structure attributes from full-waveform LiDAR: comparative analysis of methodological parameters in two geographic areas. Proceedings 3rd EARSeL Workshop SIG on Forestry, p.82. 15-16 Sept., Krakow, Poland.

5. Ruiz, L.Á., **Crespo-Peremarch, P.**, Estornell, J., Balaguer-Beser, Á., 2016. Analizing some factors affecting the extraction of full-waveform LiDAR metrics and their effect in forest structure variable estimates. Proceedings 7th Edition of the International Scientific Conference ForestSAT 2016, 24–26. 15-17 Nov., Santiago de Chile, Chile.

6. Ruiz, L.Á., **Crespo-Peremarch, P.**, 2017. Optimizing operational parameters in a full-waveform LiDAR processing tool for forestry. 17th Symposium on Systems Analysis in Forest Resources (SSAFR 2017), 27-30 Aug., Suquamish, WA, USA.

7. **Crespo-Peremarch, P.**, Ruiz, L.Á., 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. 4-7 Oct., Murcia, Spain.

8. **Crespo-Peremarch, P.**, Ruiz, L.Á., 2018. Influence of lidar full-waveform density and voxel size on forest stand estimates. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium IEEE, 5895–5898. 23-27 July, València, Spain. https://doi.org/10.1109/IGARSS.2018.8517594

9. **Crespo-Peremarch, P.**, Ruiz, L.Á., 2019. Analysis of side-lap effect and characterization of understory vegetation using full-waveform ALS.

Proceedings of the II Geomatics Engineering Conference, 19, 6. 26-27
June, Madrid, Spain. https://doi.org/10.3390/proceedings2019019006

Acknowledgments

He aquí la sección que todos se leerán, algunos por ver si se les menciona y otros por mero cotilleo, y además porque es la más entendible por la gente, puesto que no tiene tecnicismos relacionados con el campo de la teledetección. Dándole vueltas a la cabeza no sé si es la sección más sencilla de escribir o la más compleja, puesto que en teoría aquí es donde uno abre su corazón y deja salir las emociones, donde todo debe fluir y, mis dedos, no ser más que una prolongación de mis sentimientos. Pensando sólo en eso debería ser la sección más sencilla, en la cual no pararía de escribir. No obstante, el proceso de desarrollo de la tesis no han sido unos días, han sido años, en los cuales han ocurrido tantísimas cosas, muchas buenas y muchas malas. Para poder mencionar todos esos momentos y personas debería haber llevado dos documentos en paralelo cada día: el de agradecimientos y la propia tesis. Pero sabemos que esto no es viable, y las cosas se han hecho así. Así que, pensándolo de este modo, donde recordar todo lo que ha ocurrido durante estos años, lo convierte en una tarea bastante compleja. Sin embargo, de lo que sí estoy seguro es que tras escribirla me quedaré muy a gusto, como si me hubiese quitado un gran peso de encima, al poder abrirme al resto de gente.

Lo primero, antes de empezar con las personas, es mencionar mi afinidad con el tema de la tesis. Es cierto que como a cualquier ser querido, hay momentos que odias a la tesis (por no decir otra expresión más fuerte que es la que se suele emplear) y le coges manía, pero en el fondo te gusta y la quieres, y no puedes estar sin ella. Desde pequeño he sentido esa conexión con la montaña y el mundo forestal gracias a las oportunidades que me han dado mis padres. Yo soy urbanita, he nacido y principalmente crecido en la ciudad, así que no me puedo considerar de campo por pasar en él algún verano o fin de semana, hacer alguna excursión, o incluso vivir un par de años. Pero tampoco llego al nivel de dominguero. Sin embargo, dentro de mí el peso que tiene la montaña y la naturaleza es mucho mayor, y no podría vivir sin ella. El disponer de una casa en el campo en Gilet, en plena Serra Calderona, me ha hecho ser como soy y amar la naturaleza desde muy pequeño. Desde entonces me he perdido por el monte, salía fuera a la parcela a jugar aunque hiciese mucho frío, y ese contacto con la tierra, el agua y los árboles me hizo enamorarme de la naturaleza. Hoy

en día sigo disfrutando de ella, aunque de otra manera. El poder ir allí ayuda a desconectar cuando he tenido días malos o de agobio. El salir fuera en esos días frescos, cuando aún queda algo de niebla y el rocío cubre las plantas, sólo con eso se respira diferente y se cargan los pulmones de energía. Por otro lado está el mundo de la cartografía, en el cual también incluyo la teledetección. Elegí esta carrera universitaria poco antes de presentar el formulario con las carreras que quería estudiar, y no me arrepiento para nada. Mi amor por los mapas viene también de bien pequeño. Yo soy ese niño que cogía un atlas y lo observaba durante horas, para conocer las ciudades, las fronteras, los ríos, los mares, las banderas, y al final toda esa información quedaba grabada en mi cabeza. Cuando estaba en Gilet intentaba recrear los mapas en el terreno, vamos lo que sería un replanteo para un topógrafo. Cogía los mapas de alguna zona, una escobilla para el coche, y con ella barría la tierra suelta del terreno para hacer las calles y carreteras. Después ya venía con los cochecitos y me ponía a jugar. Lo bueno es que cuando me aburría de una zona, podía barrerla y rehacerla de nuevo. Este fue mi mundo 2D, previo al 3D, en el cual recreaba monumentos simbólicos de ciudades con el barro que cogía de la parcela. Esta nueva moda surgía cuando visitaba con mis padres alguna ciudad como París o Madrid, y me volvía con un mapa de allí. Una vez en la parcela recreaba las calles y edificios, incluso cavaba un poco para recrear los ríos. Era la curiosidad de un niño que no se escondía detrás de un dispositivo electrónico. Mi tema de la tesis, en la que se han tocado temas de cartografía, teledetección y forestales, ha sido volver a conectar con ese niño dentro de mí que jugaba al aire libre recreando el mundo real y se perdía entre los árboles.

Pero por mucho que te guste un tema, no sirve de nada si no estás rodeado de la gente adecuada y que quieres. Esa gente que te saca de casa, te escucha, o simplemente te saca una sonrisa que lo arregla todo. Es muy complicado acordarse de toda esa gente que ha pasado por mi vida durante estos años, ha ayudado de algún modo directo o indirecto, o simplemente ha estado ahí. Lo primero es acordarme de mis padres, que me han criado y facilitado todo para que yo pueda llegar hasta aquí. Mis gustos no serían lo mismo sin ellos ni otra gente que me rodeó durante mi infancia, como familiares, amigos y profesores. Aparte de aportarme esa base para llegar hasta aquí, también me han dado esos empujones para mantenerme. No podía olvidarme tampoco en estos agradecimientos de Elena, mi compañera de viaje que tantísimo quiero, que

siempre me ha escuchado cuando necesitaba desahogarme y siempre tiene buenos consejos para saber cómo afrontar "laj" cosas. Aunque una simple sonrisa suya lo ha cambiado todo en más de una ocasión. Después de la familia no puedo dejar de mencionar a la familia que se elige, a los amigos. Dentro de la categoría amigos hay muchos grupos que en algún momento o lugar habéis estado presentes. No me puedo olvidar de Pau, Carles, Marta, Elena, Alba, Inés y Alicia, aunque cada vez estemos más separados, porque cada uno está en una parte del mundo y, por lo tanto, cuesta más verse, sois unos amigos increíbles y los momentos vividos con vosotros son inolvidables. Ojalá podamos volver a juntarnos pronto y seguir viviendo momentos. Aparte de ser compañeros de trabajo, Jesús, Carlos, Juan Pedro, Jaime, Chema, Marta y Elena, sois mucho más. El trabajar en el CGAT con vosotros es increíble y hacéis que se lleve de otra manera. Cada uno de nosotros somos diferentes y tenemos nuestra parte especial, pero el ambiente en el CGAT es una risa y muy ameno gracias a todos y a cada uno de vosotros. En el CGAT hemos tenido momentos de estrés, de silencio máximo (pocos), y de venirnos arriba con alguna partida de algún juego inventado como el "chair-hockey", especialmente por las tardes, cuando ya empezamos a desvariar. Qué paciencia han tenido que tener gente como Marta o Elena. Otro grupo de gente que no podía olvidar mencionar son los de Gilet: David, Tere, Javi, Laura, Jeni y Diana. Las responsabilidades y la edad hacen que cada vez nos podamos ver menos, pero cuando nos vemos siempre pasamos buenos momentos y nos reímos un rato, y encima en un marco que nos encanta como son las montañas de Gilet. Algo fundamental son las actividades que me han ayudado a desconectar en los momentos de mayor agobio, sin ellas no habría conseguido llegar hasta aquí, y estas actividades se han podido llevar a cabo gracias a otros grandes amigos. Siempre había tenido respeto (y le sigo teniendo) al mar. Cuando empiezas a practicar windsurf siempre tienes el miedo de lo que te podría pasar, pero todo cambia cuando empiezas a avanzar, y cuando ya uno se queda completamente enganchado es cuando empiezas a planear. Esa sensación de libertad y velocidad en medio del mar hace que te olvides de todo y grites para liberar la adrenalina, lo que hace que te quedes completamente enganchado a este deporte. Todo esto no habría sido posible sin Mompa. Él me enseñó y me fue metiendo esta afición por el windsurf. Él me ha permitido empezar y avanzar utilizando su material sin pedir nada a cambio, simplemente porque es un deporte que se disfruta mucho más yendo varias

personas. Hemos disfrutado días de mucho viento, quedándome alguno de ellos en la playa, y otros de mucha calma para enseñar a otras personas o simplemente para darse una vuelta hasta la playa de Pinedo con un bocadillo y tomárnoslo allí en la orilla. También ha ayudado en poder llevar a cabo esta tesis esos días de partido del Llevant con mi primo Hugo. Esos partidos en el que se han pasado nervios, pero al final han sido más risas que otra cosa. Aunque nos riésemos de nuestro propio equipo, hacen pasar un buen rato de distracción y desconexión. No puedo olvidarme de otra actividad que ha sido muy importante para mí durante estos años de tesis, y en la que pude conocer a mi pareja, el Lindy Hop. Nunca había bailado, y nunca hubiese pensado que una actividad así me enganchara tanto y me ayudara tanto. Pero sin lugar a duda lo más importante ha sido toda la gente maravillosa que he conocido. Ese grupo de amigos que se formó, y que desafortunadamente ya no nos vemos tanto por diversos motivos. Me olvidaré de mucha gente que he conocido estos años, pero en estas palabras me gustaría acordarme de David, Pablo, Marta, Albert, Davinia, José Luis, Natalia y Sergio. Esas tardes/noches de clase, cena, sofá y baile en la escuela fueron unos meses increíbles. Por último, no quiero olvidarme de mis gatas Ellie, Chuchi y Duna. Poca gente entenderá que incluya a unos animales en los agradecimientos, pero no podía olvidarme de quien me ha hecho compañía, dado cariño y extraño mucho cuando no están. Tampoco quiero olvidarme de mi otra familia que tantos buenos momentos me han dado, se han preocupado mucho por mí y tanto me han hecho reír. Gracias Llanos, Juanba, Ana, Darío, los dos Pablos y Rita. Puede que este párrafo haya sido muy extenso, pero no quería olvidarme de todos aquellos que han puesto su granito para que este documento sea posible. Gracias a todos.

No sólo hay que rodearse de buena gente para poder llevar a cabo algo satisfactoriamente, sino que alguien te debe dar la oportunidad de trabajar en lo que te gusta desde pequeño, y eso ha sido gracias a Luis Ángel. Para mí fue muy importante que se acordara de mí para una oferta predoctoral en Suiza habiendo coincidido sólo en una asignatura, y en parte por eso y por mi interés por la teledetección decidí que fuera mi codirector en la tesina de máster. Gracias a él he tenido la oportunidad de trabajar en este tema tan bonito, y sin su ayuda y conocimientos no habría podido presentar esta tesis. También quiero agradecer a Ángel lo mucho que me ha ayudado con los temas matemáticos y estadísticos, sin él tampoco se habrían podido llevar a cabo muchas cosas. Por

último, quiero agradecer al resto de miembros del CGAT, Javier, Jesús, Alfonso, Jorge y Josep por su ayuda y los buenos momentos vividos durante estos años. Las risas y el sentido del humor nunca han faltado. Sin el equipo del CGAT y la oportunidad que se me dio, todo esto tampoco se podría haber llevado a cabo.

My experience at the IRSS of UBC in Vancouver was unforgettable. I had the opportunity to work in a leading group in remote sensing headed by Nicholas Coops. I felt at home there since I was given the opportunity to attend meetings and conferences, go to the IRSS retreat, travel to some amazing places, go to drink beers on Friday afternoons, and play games such as ping-pong or laser chess. Nicholas made me feel like one more, and he taught me many things. Apart from Nicholas I met other wonderful people who made up the IRSS like Nacho, Giona, Marco, Joseph, Txomin, Tristan, Piotr, Doug, Riccardo, Aria, Chris, Bean and Max. They are an amazing research group. In addition, Vancouver and surroundings are such a wonderful place. After living there for six months, I realized that it was the nicest place I had ever stayed.

Grâce à mon stage à l'IRSS, j'ai pu rencontrer Richard Fournier et avoir aussi l'opportunité de faire un petit stage à l'Université de Sherbrooke. Richard m'a hébergé chez lui, où j'ai rencontré sa famille, qui m'a traité comme un d'entre eux. Avec lui, j'ai aussi pu apprendre beaucoup de choses du monde de la recherche. De plus, bien que mon stage ait été court, j'ai pu rencontrer de très bonnes personnes comme Bastien, Van-Tho, Naime, Batistin, Catherine et Aurélie; et en plus revenir au Québec, où j'ai vécu de si bonnes expériences en 2008-2009.

Aunque estos agradecimientos los escribiera hace tiempo, no podía olvidarme de esta situación tan extraña vivida durante el desarrollo final de la tesis : la pandemia y confinamiento de la COVID-19. ¿Quién nos hubiera dicho esto hace unos meses? Esto te hace ver lo cambiante que puede ser el mundo, aunque vivas en lo conocido como primer mundo. No ha sido fácil llevar a cabo la tesis estando siete semanas confinados en casa, más la posterior desescalada. El no estar en la oficina o poder salir a la calle lo ha complicado todo. Aunque parece que el cerebro humano intenta olvidar la situaciones más adversas, es algo que siempre estará ahí, una demostración de lo vulnerables que somos.

Esta tesis no ha sido todo un camino de rosas, pero lo importante es que me voy con mucho más de lo que vine, tanto profesional como personalmente. No soy la misma persona y he aprendido de muchas personas tan diversas. En definitiva, todo esto ha sido lo que ha hecho que yo haya podido llevar a cabo esta tesis, y espero que la disfrutéis. Así que, gracias a todos, gràcies a tots, thank you all, merci à tous.

A toda esa gente que me rodea, y a la que pueda disfrutar de la lectura.

List of abbreviations

Abbreviation	Explanation
ABA	Area-Based Approach
AGB	AboveGround Biomass
AGS	Average Gaussian curve Slope
AIC	Akaike Information Criterion
ALS	Airborne Laser Scanning
ALS$_{FW}$	Full-Waveform Airborne Laser Scanning
ALS$_D$	Discrete Airborne Laser Scanning
BC	Bottom of Canopy
BCD	Bottom of Canopy Distance
BCE	Bottom of Canopy Energy
C	Cover of understory vegetation
CBH	Canopy Base Height
CD	Canopy Distance
CE	Canopy Energy
CER	Canopy Energy Ratio
CFL	Canopy Fuel Load
CH	Canopy Height
CHP	Canopy Height Profile
CHn	Canopy Height quartile n
CV	Coefficient of Variation
DBH	Diameter at Breast Height
DTM	Digital Terrain Model
EFEV	Energy to First Empty Voxel
ENERGY Qn	ENERGY Quarter n
FS	Front Slope
FVU	Filled Voxels at Understory
GE	Ground Energy
GEDI	Global Ecosystem Dynamics Investigation
GIS	Geographical Information System
GLAS	Geoscience Laser Altimeter System
GPS	Global Position System
GRR	Ground Return Ratio
HEIGHT Qn	HEIGHT Quarter n
HFEV	Height at First Empty Voxel
HFEVT	Height at First Empty Voxel from Threshold
H$_{max}$	Maximum Height of understory vegetation
H$_{mean}$	Mean Height of understory vegetation
Hn	Height percentile n
HOME	Height Of Median Energy
HTMR	HeighT/Median Ratio
ICESat	Ice, Cloud, and land Elevation Satellite

ISS	International Space Station
LiDAR	Light Detection And Ranging
LVIS	Laser Vegetation Imaging Sensor
MAX E	MAXimum Energy
MPD	Minimum Pulse Density
MSGS	Modified Standard deviation Gaussian curve Slope
NFVU	Number of Filled Voxels at Understory
Nb	Number of beams occluded prior to reaching a given voxel
Nh	Number of hits
Nh^{ALS}	Number of hits from ALS
Nh^{ALSD}	Number of hits from ALS$_D$
Nh^{ALSFW}	Number of hits from ALS$_{FW}$
Nt	Number of theoretical laser beams crossing a given voxel
NP	Number of Peaks
nEFEV	Normalized Energy to First Empty Voxel
N GS	Number of Gaussian curve
nRMSE	Normalized Root-Mean-Square Error
PAD	Plant Area Density
PAD^{TLS}	Plant Area Density from TLS
pR^2	Pseudo-coefficient of determination
Rn	Ratio quartile n
R^2	Coefficient of determination
R^{2}c	Conditional coefficient of determination
R^{2}m	Marginal coefficient of determination
RMSE	Root-Mean-Square Error
ROUGH	ROUGHness
RWE	Return Waveform Energy
SGS	Standard deviation Gaussian curve Slope
SLICER	Scanning Lidar Imager of Canopies by Echo Recovery
sqrt	Square-root-transformed
TLS	Terrestrial Laser Scanning
TVar	Total Variation
V	Volume of understory vegetation
VDR	Vertical Distribution Ratio
WD	Waveform Distance
WoLFeX	Waveform Lidar for Forestry eXtraction

Table of contents

Chapter 6: Final conclusions   155

References   167

List of figures

List of tables

List of equations

Abstract

Forest ecosystems are an important source of life and economic use, since they are large stores of carbon and a renewable raw material. Nevertheless, a poor forest management would considerably reduce these capacities and increase the risk of forest fires. In the last decades, remote sensing techniques have proven their capacity for forest management. Airborne laser scanning (ALS) provides horizontal and vertical information of different canopy layers. In particular, full-waveform airborne laser scanning (ALS_{FW}), which registers the complete signal emitted by the sensor and backscattered, provides more information about the vertical forest structure than traditional or discrete airborne laser scanning (ALS_D). However, ALS_{FW} has received less attention than ALS_D, due to its larger amount and complexity of data, and the lack of processing tools available.

This thesis addresses the development of ALS_{FW} processing and analysis methods to characterize the vertical forest structure, in particular, the understory vegetation. To answer this overarching goal, a total of six specific objectives were established: Firstly, the influence of pulse density, voxel parameters (i.e., voxel size and assignation value) and regression methods on ALS_{FW} metric values and on estimates of forest structure attributes are analyzed. To do this, pulse density was randomly reduced and voxel parameters modified, obtaining ALS_{FW} metric values for the different parameter combinations. These ALS_{FW} metrics were used to estimate forest structure attributes with different regression methods. Secondly, a set of ALS_{FW} data processing and analysis methods are integrated in a new software named *WoLFeX* (Waveform Lidar for Forestry eXtraction), including clipping, relative radiometric correction, voxelization and ALS_{FW} metric extraction, and proposing new metrics for understory vegetation. Thirdly, the influence of the scan angle of ALS data acquisition and radiometric correction on the extraction of ALS_{FW} metrics and on modeling forest fuel attributes is assessed. To do this, ALS_{FW} metrics were extracted applying and without applying relative radiometric correction and using different scan angles. Fourthly, signal occlusion is characterized along the vertical forest structure using and comparing three different laser scanning configurations (ALS_{FW}, ALS_D and

terrestrial laser scanning: TLS), determining their limitations in the detection of vegetative material in two contrasted forest ecosystems: boreal and Mediterranean. To quantify signal occlusion along the vertical forest structure, a new parameter based on the percentage of laser beams blocked prior to reach a given location, the rate of pulse reduction, is proposed. Fifthly, the assessment of how understory vegetation density classes are detected and determined by different ALS configurations is done. Vertical distribution profiles at the lower strata described by ALS_{FW} and ALS_D are compared with those described by TLS as reference. Moreover, understory vegetation density classes are determined by applying the Lorenz curve and Gini index from the vertical distribution profiles described by ALS_{FW} and ALS_D. Finally, the new proposed voxel-based ALS_{FW} metrics are applied and evaluated, using TLS-based attributes as a reference, to estimate understory height, cover and volume in a Mediterranean ecosystem.

Results show that variations of ALS_{FW} metric values may be reduced by either using a minimum pulse density or increasing the voxel size and modifying the assignation value. Given that forest attributes are estimated by ALS_{FW} data, they are also influenced by pulse density, which may also be reduced by increasing voxel size and modifying the regression method. Additionally, ALS_{FW} metric values and estimates of forest fuel attributes are also influenced by scan angle. This influence may be reduced, but not completely removed, by applying the radiometric correction.

Detection of the vertical distribution was observed to be dependent on occlusion. The degree of occlusion may be quantified by the rate of pulse reduction along the vertical structure, and therefore, the reliability in the characterization of the vertical distribution may also be estimated. In this regard, ALS configurations (ALS_D and ALS_{FW}) demonstrated their capabilities to detect understory vegetation, although much more accuracy was obtained using ALS_{FW}. The latter demonstrated its potential to detect and determine understory vegetation density classes in a boreal and a Mediterranean forest by using the Gini index, and to estimate the height, cover and volume of understory vegetation in a Mediterranean forest.

The use of ALS$_D$ data is adequate for most of the current ALS applications. However, and contrary to ALS$_{FW}$, it presents some limitations in the identification and/or characterization of the intermediate and lower vegetation strata. These limitations in the detection of vertical strata can be identified and quantified by retrieving vertical profiles of rate of pulse reduction from different laser scanning configurations. On the other hand, the use of ALS$_{FW}$ data requires the identification of the appropriate parameters (i.e., optimal pulse density and voxel parameters) and the application of radiometric correction prior to any data processing. These results highlight the potential of ALS$_{FW}$ to replace TLS in the extraction of forest internal structure in wider areas. Nevertheless, despite ALS$_{FW}$ presents less limitations in the detection of intermediate and lower strata than ALS$_D$, its signal occlusion may be significant in lower strata of forests with high canopies and very dense vegetation and internal structure (e.g., tropical areas), resulting in a lack of vegetative material identification. These findings in the characterization of the vertical forest structure, in particular the understory vegetation, are relevant for forestry applications such as wildfire mitigation, modeling fire behavior, planning forest thinning and maintenance tasks for wildfire reduction and biodiversity preservation, among others.

Resumen

Los ecosistemas forestales son una importante fuente de vida y económica, por su capacidad para almacenar carbono y estar formados por una materia prima renovable. No obstante, la gestión forestal inapropiada puede reducir considerablemente estas capacidades y aumentar el riesgo de incendios forestales. En las últimas décadas, las técnicas de teledetección han demostrado su contribución a la gestión forestal. El láser escáner aerotransportado (ALS, por sus siglas en inglés) proporciona información horizontal y vertical de las diferentes capas del dosel arbóreo. En concreto, el láser escáner aerotransportado *full-waveform* (ALS_FW), que registra la totalidad de la señal emitida por el sensor que es retrodispersada, proporciona mayor información que el láser escáner aerotransportado tradicional o discreto (ALS_D) sobre la estructura vertical del bosque. Sin embargo, el ALS_FW ha recibido menor atención que el ALS_D, debido a la gran cantidad y complejidad de sus datos y a la falta de disponibilidad de herramientas para su procesado.

Esta tesis aborda el desarrollo de métodos de procesado y análisis de datos ALS_FW para la caracterización de la estructura vertical del bosque y, en particular, del sotobosque. Para responder a este objetivo general, se establecieron seis objetivos específicos: En primer lugar, se analiza la influencia de la densidad de pulso, de los parámetros de voxelización (tamaño de vóxel y valor de asignación) y de los métodos de regresión sobre los valores de las métricas ALS_FW y sobre la estimación de atributos de estructura del bosque. Para ello, se redujo aleatoriamente la densidad de pulsos y se modificaron los parámetros de voxelización, obteniendo los valores de las métricas ALS_FW para las diferentes combinaciones de parámetros. Estas mismas métricas ALS_FW se emplearon para la estimación de atributos de la estructura del bosque mediante diferentes métodos de regresión. En segundo lugar, se integran métodos de procesado y análisis de datos ALS_FW en una nueva herramienta llamada *WoLFeX* (*Waveform Lidar for Forestry eXtraction*) que incluye los procesos de recorte, corrección radiométrica relativa, voxelización y extracción de métricas a partir de los datos ALS_FW, así como nuevas métricas descriptoras del sotobosque. En tercer lugar, se evalúa la influencia del ángulo de escaneo utilizado en la adquisición de datos ALS y la corrección radiométrica en la extracción de

métricas ALS$_{FW}$ y en la estimación de atributos de combustibilidad forestal. Para ello, se extrajeron métricas ALS$_{FW}$ con y sin corrección radiométrica relativa y empleando diferentes ángulos de escaneo. En cuarto lugar, se caracteriza la oclusión de la señal a lo largo de la estructura vertical del bosque empleando y comparando tres tipos diferentes de láser escáner (ALS$_{FW}$, ALS$_D$ y láser escáner terrestre: TLS, por sus siglas en inglés), determinando así sus limitaciones en la detección de material vegetativo en dos ecosistemas forestales diferenciados: el boreal y el mediterráneo. Para cuantificar la oclusión de la señal a lo largo de la estructura vertical del bosque se propone un nuevo parámetro, la tasa de reducción del pulso, basada en el porcentaje de haces láser bloqueados antes de alcanzar una posición dada. En quinto lugar, se evalúa la forma en que se detectan y determinan las clases de densidad de sotobosque mediante los diferentes tipos de ALS. Se compararon los perfiles de distribución vertical en los estratos inferiores descritos por el ALS$_{FW}$ y el ALS$_D$ con respecto a los descritos por el TLS, utilizando este último como referencia. Asimismo, se determinaron las clases de densidad de sotobosque aplicando la curva Lorenz y el índice Gini a partir de los perfiles de distribución vertical descritos por ALS$_{FW}$ y ALS$_D$. Finalmente, se aplican y evalúan las nuevas métricas ALS$_{FW}$ basadas en la voxelización, utilizando como referencia los atributos extraídos a partir del TLS, para estimar la altura, la cobertura y el volumen del sotobosque en un ecosistema mediterráneo.

Los resultados muestran que las variaciones de los valores de las métricas ALS$_{FW}$ se pueden reducir empleando una densidad de pulso mínima o incrementando el tamaño de vóxel y modificando el valor de asignación. Debido a que los atributos forestales se estiman mediante datos ALS$_{FW}$, también se ven influenciados por la densidad de pulsos. Esta influencia también se puede reducir incrementando el tamaño de vóxel o modificando el método de regresión. Asimismo, los valores de las métricas ALS$_{FW}$ y la estimación de atributos de combustibilidad forestal también se ven influenciados por el ángulo de escaneo. Esta influencia se puede reducir, pero no eliminar por completo, aplicando la corrección radiométrica.

Por otro lado, se observó una dependencia en la detección de la distribución vertical con respecto a la oclusión. El grado de oclusión a lo largo de la estructura vertical se puede cuantificar mediante la tasa de reducción del pulso

y, por lo tanto, se puede estimar la fiabilidad en la caracterización de la distribución vertical. En este sentido, el ALS (ALS_D y ALS_{FW}) demostró ser útil en la detección del sotobosque, obteniendo mucha mayor precisión con el ALS_{FW}. Este último demostró su potencial para la detección y determinación de las clases de densidad de sotobosque en un bosque boreal y otro mediterráneo mediante el uso del índice Gini, así como para la estimación de la altura, la cobertura y el volumen del sotobosque en bosques mediterráneos.

La utilización de los datos ALS_D es suficiente para la mayoría de aplicaciones actuales del ALS. No obstante, y contrariamente al ALS_{FW}, presenta algunas limitaciones en la identificación y/o caracterización de los estratos intermedios e inferiores de la vegetación. Estas limitaciones en la detección de los estratos verticales se pueden identificar y cuantificar mediante la extracción de perfiles verticales de la tasa de reducción del pulso para los diferentes tipos de láser escáner. Por otro lado, para la utilización de datos ALS_{FW} es conveniente identificar los parámetros adecuados (densidad de pulso óptima y parámetros de voxelización) y aplicar una corrección radiométrica, como paso previo a cualquier procesado de datos. Los resultados destacan el potencial del ALS_{FW} como sustituto del TLS en la extracción de la estructura interna del bosque en áreas extensas. No obstante, a pesar de que el ALS_{FW} presenta menores limitaciones con respecto al ALS_D en la detección de los estratos intermedios e inferiores, la oclusión de la señal puede ser significativa en los estratos inferiores de bosques que presenten doseles arbóreos altos y una gran densidad de vegetación en su estructura interna (por ejemplo, áreas tropicales), conllevando una falta de identificación del material vegetativo. Estos hallazgos en la caracterización de la estructura vertical del bosque y, en particular, del sotobosque, son relevantes para su aplicación en la planificación del desbroce de los bosques y de las tareas de mantenimiento para la prevención de los incendios forestales, la modelización del comportamiento del fuego o la conservación de la biodiversidad, entre otras aplicaciones.

Resum

Els ecosistemes forestals són una important font de vida i econòmica, per la seua capacitat per emmagatzemar carboni i estar formats per una matèria prima renovable. No obstant això, la gestió forestal inapropiada pot reduir considerablement aquestes capacitats i augmentar el risc d'incendis forestals. En les darreres dècades, les tècniques de teledetecció han demostrat la seua contribució a la gestió forestal. El làser escàner aerotransportat (ALS, per les seues sigles en anglès) proporciona informació horitzontal i vertical de les diferents capes del cobricel arbori. En concret, el làser escàner aerotransportat *full-waveform* (ALS$_{FW}$), que registra la totalitat del senyal emès pel sensor que és retrodispersada, proporciona major informació que el làser escàner aerotransportat tradicional o discret (ALS$_D$) sobre l'estructura vertical del bosc. En canvi, l'ALS$_{FW}$ ha rebut menys atenció que l'ALS$_D$, a causa de la gran quantitat i complexitat de les seues dades i a la falta de disponibilitat d'eines per al seu processament.

Aquesta tesi aborda el desenvolupament de mètodes de processament i anàlisi de dades ALS$_{FW}$ per a la caracterització de l'estructura vertical del bosc i, en particular, del sotabosc. Per a respondre a aquest objectiu general, s'establiren sis objectius específics: En primer lloc, s'analitza la influència de la densitat de pols, dels paràmetres de voxelització (grandària de vóxel i valor d'assignació) i dels mètodes de regressió sobre els valors de les mètriques ALS$_{FW}$ i sobre l'estimació dels atributs d'estructura del bosc. Per a això, es reduí aleatòriament la densitat de polsos i es modificaren els paràmetres de voxelització, obtenint els valors de les mètriques ALS$_{FW}$ per a les diferents combinacions de paràmetres. Aquestes mètriques ALS$_{FW}$ s'empraren per a l'estimació d'atributs de l'estructura del bosc mitjançant diferents mètodes de regressió. En segon lloc, s'integraren mètodes de processament i d'anàlisi de dades ALS$_{FW}$ en una nova eina anomenada *WoLFeX* (*Waveform Lidar for Forestry eXtraction*) que inclou el processos de retallada, correcció radiomètrica relativa, voxelització i extracció de mètriques a partir de les dades ALS$_{FW}$, així com noves mètriques descriptores del sotabosc. En tercer lloc, s'avalua la influència de l'angle de escaneig emprat en l'adquisició de les dades ALS i la correcció radiomètrica en l'extracció de mètriques ALS$_{FW}$ i en l'estimació d'atributs de

combustibilitat forestal. Per a això, s'extragueren mètriques ALSFW amb i sense correcció radiomètrica relativa i emprant diferents angles d'escaneig. En quart lloc, es caracteritza l'oclusió del senyal al llarg de l'estructura vertical del bosc emprant i comparant tres tipus diferents de làser escàner (ALSFW, ALSD i làser escàner terrestre: TLS, per les seues sigles en anglès), determinant així les seues limitacions en la detecció de material vegetatiu en dos ecosistemes diferenciats: un boreal i un mediterrani. Per a quantificar l'oclusió del senyal al llarg de l'estructura vertical del bosc es proposa un nou paràmetre, la taxa de reducció del pols, basada en el percentatge de rajos làser bloquejats abans d'arribar a una posició donada. En cinquè lloc, s'avalua la manera en la qual es detecten i determinen les classes de densitat de sotabosc mitjançant els diferents tipus d'ALS. Es compararen els perfils de distribució vertical en estrats inferiors descrits per l'ALSFW i l'ALSD respecte als descrits pel TLS, emprant aquest últim com a referència. A més a més, es determinaren les classes de densitat de sotabosc aplicant la corba Lorenz i l'índex Gini a partir dels perfils de distribució vertical descrits per l'ALSFW i l'ALSD. Finalment, s'apliquen i avaluen les noves mètriques ALSFW basades en la voxelització, emprant com a referència els atributs extrets a partir del TLS, per a estimar l'alçada, la cobertura i el volum del sotabosc en un ecosistema mediterrani.

Els resultats mostren que les variacions dels valors de les mètriques ALSFW es poden reduir emprant una densitat de pols mínima o incrementant la grandària del vóxel i modificant el valor d'assignació. A causa de que els atributs forestals s'estimen mitjançant dades ALSFW, també es veuen influenciats per la densitat de polsos. Aquesta influència també es pot reduir incrementant la grandària del vóxel o modificant el mètode de regressió. Tanmateix, els valors de les mètriques ALSFW i l'estimació d'atributs de combustibilitat forestal també es veuen influenciats per l'angle d'escaneig. Aquesta influència es pot reduir, però no eliminar per complet, aplicant la correcció radiomètrica.

Per altra banda, s'observà una dependència en la detecció de la distribució vertical respecte a l'oclusió. El grau d'oclusió al llarg de l'estructura vertical es pot quantificar mitjançant la taxa de reducció del pols i, per tant, es pot estimar la fiabilitat en la caracterització de la distribució vertical. En aquest sentit, l'ALS (ALSD i ALSFW) demostrà ser útil en la detecció del sotabosc, obtenint molta major precisió amb l'ALSFW. Aquest últim demostrà el seu potencial per a la

detecció i determinació de les classes de densitat de sotabosc en un bosc boreal i un altre mediterrani mitjançant la utilització de l'índex Gini, així com l'estimació de l'alçada, la cobertura i el volum del sotabosc en un boscos mediterranis.

La utilització de les dades ALS$_D$ es suficient per a la majoria d'aplicacions actuals de l'ALS. En canvi, i contràriament a l'ALS$_{FW}$, presenta algunes limitacions en la identificació i/o caracterització dels estrats intermitjos i inferiors de la vegetació. Aquestes limitacions en la detecció dels estrats verticals es poden identificar i quantificar mitjançant l'extracció de perfils verticals de la taxa de reducció del pols per als diferents tipus de làser escàner. D'altra banda, per a la utilització de dades ALS$_{FW}$ és convenient identificar els paràmetres adequats (densitat de pols òptima i paràmetres de voxelització) i aplicar una correcció radiomètrica, com a pas previ a qualsevol processament de dades. Els resultats destaquen el potencial de l'ALS$_{FW}$ com substitut del TLS en l'extracció de l'estructura interna del bosc en àrees extenses. No obstant això, malgrat que l'ALS$_{FW}$ presenta menors limitacions respecte a l'ALS$_D$ en la detecció d'estrats intermitjos i inferiors, l'oclusió del senyal pot ser significativa en els estrats inferiors de boscos que presenten cobricels arboris alts i una gran densitat de vegetació en la seua estructura interna (per exemple, àrees tropicals), comportant una falta d'identificació del material vegetatiu. Aquestes troballes en la caracterització de l'estructura vertical del bosc i, en particular, del sotabosc, són rellevants per a la seua aplicació en la planificació del desbrossament dels boscos i de les tasques de manteniment per a la prevenció dels incendis forestals, la modelització del comportament del foc o la conservació de la biodiversitat, entre d'altres.

Chapter 1

Introduction

Quercus ilex

*"Con el sol del otoño toda el agua
de mi fontana vibra,
y noto que sacando sus raíces
huye de mí la encina."*

Federico García Lorca

1.1. Background and research justification

In ecology, the forest structure is defined as the horizontal and vertical distribution of the vegetation elements such as trees and shrubs (Kükenbrink et al., 2017). More specifically, these elements are composed by tree trunks, branches, twigs, leaves and deadwood, which interact with each other forming the forest structure. In terms of productivity, the forest structure may also be described by basal area, which is the area occupied by the cross-section of tree trunks at the diameter breast height (i.e., 1.3 m) by unit area, and tree height heterogeneity (Bohn and Huth, 2017). Forest structure may be altered by some disturbance regimes (Kimes et al., 2006), most of them related to anthropogenic processes. Some of these activities or occurrences are infrastructure building, timber production and wildfires (Guo et al., 2017). As a consequence, the affection gradient of disturbance regimes is variable along forested areas, and in addition, some species are more suitable to survive or regenerate under these conditions, while others will disappear (Devictor et al., 2008; Johnstone et al., 2016). This leads to a more heterogeneous horizontal and vertical forest structures. The significance of forest structure is also clearly visible on several factors such as biodiversity and wildfire modeling. Disturbance regimes may cause microclimatic patterns and processes affecting distribution of bird species and wildlife habitats (Zimble et al., 2003; Hyde et al., 2005; Hyde et al., 2006; Lesak et al., 2011; Guo et al., 2017), modify biological process such as tree competition and growth (Drake et al., 2002; Coops et al., 2007), and condition the spread and severity of wildfires (Agee, 1996; Pollet and Omi, 2002; Graham et al., 2004; D.L. Peterson et al., 2005; Hyde et al., 2006; Prichard et al., 2010).

1.1.1. Fire modeling

The fire environmental triangle is made up of three legs: topography, weather conditions and fuel (Countryman, 1972). Topography and weather conditions play a more decisive role in fire behavior (Pollet and Omi, 2002). However, fuel, which is related to forest structure, is the only leg that may be manipulated (Agee, 1996). Given that the other two legs of the fire environmental triangle may not be controlled, they must be considered prior to

fuel treatments (D.L. Peterson et al., 2005). For instance, fuel treatments vary according to elevation, aspect and slope (i.e., topography), and humidity, temperature and wind speed (i.e., weather conditions). Fuel treatment techniques have been used by forest managers for decades. However, not all forests are at risk of severe wildfires, but drier forests require fuel treatments to control fire hazard (Agee and Skinner, 2005). These fuel treatment techniques aim therefore to reduce intensity (i.e., heat release per unit distance and unit time) and severity (i.e., related to post-fire vegetation survival) of wildfires (Agee, 1996; Arkle et al., 2012). These techniques are even more necessary in the last years, since fire frequencies have decreased owing to fire exclusion, harvesting and different land use practices, leading to an increment of fuels potentially hazardous for large wildfires (Viedma et al., 2018). Fires may be divided into three types: ground, surface and crown fires (Werth et al., 2011). Ground fires are related to soil organic horizons, surface fires to low vegetation, woody fuel, moss, lichen and litter, and crown fires to canopies (Graham et al., 2004; Weise et al., 2018). Crown fires are the most severe fires, and they are the main threat to ecological and human values as well as challenging for fire management, hence reducing these fires facilitates the suppression and reduces the likelihood of having large wildfires (Lecina-Diaz et al., 2014). Crown fires depend on the arrangement of available fuel from the ground to the canopy (D.L. Peterson et al., 2005), and this available fuel between the two strata is known as ladder fuel (National Wildfire Coordinating Group, 2005). Therefore, more severe wildfires take place in spots where there is a high presence of understory and regeneration of trees connected to trees densely stocked, resulting in crown fires spread from crown to crown (Pollet and Omi, 2002). Regarding the existing fuel treatment techniques to reduce wildfires intensity and severity, two of the most used techniques are thinning and prescribed fires. Thinning allows for carrying out a more precisely planned forest structure, while prescribed fires burn imprecisely low vegetation and lower branches from overstory (Graham et al., 2004). The most effective approach to reduce crown fires events by fuel treatments is by increasing the height to live crown, reducing fuels and canopy bulk density, and increasing distance between canopies (Graham et al., 1999). To do this, the most adequate strategy is applying both fuel treatment techniques mentioned: thinning and then

prescribed fires (Graham et al., 1999; Pollet and Omi, 2002; Agee and Skinner, 2005). Thinning removes small diameter trees (Pollet and Omi, 2002) and modifies canopy and ladder fuels, and then lower branches from overstory, woody fuel, litter and plant remains from thinning may be removed with prescribed fires (Harrod et al., 2009). As a result of modifying forest structure, crown fire severity and intensity are reduced, since vertical (i.e., ladder fuel) and horizontal (i.e., fire crown to crown) continuities are disrupted (Arkle et al., 2012).

In order to assess ongoing or upcoming fuel treatments, wildfire behavior may be modeled. To do this, the three legs of the fire environmental triangle (i.e., fuel, topography and weather conditions) are inputs into the models to be predicted (Agee, 1996). Topography consists of information about slope, aspect and elevation, which may all be extracted from a Digital Terrain Model (DTM), a raster file where each pixel value represents its elevation. Weather conditions consists of temperature, humidity, wind direction and speed, and actual values from a given scenery or arbitrary data may be used. Fuel, which consists of estimating living and dead biomass, requires efforts, since it is dynamic and changing (Schmidt et al., 2016; Davis et al., 2017). There are three concepts that are often employed mistakenly: fire model, fuel type and fuel model. A fire model is a mathematical model which describes the evolution and behavior of the fire; fuel types are the result of clustering vegetation according to a set of attributes such as vegetation density, loading and height, which determine the fire spread; and fuel models are the values of the attributes describing each fuel type (Arroyo et al., 2008). Fire models are divided into four types: physical, physical-statistical, statistical, and probabilistic models (Albright and Meisner, 1999). Physical fire models, such as Albini (1986), are based on the physics of combustion, however, they require such amount of detailed data that they are not usually employed (Albright and Meisner, 1999). Physical-statistical fire models join statistical correlation and physics. Among these fire models, there are Rothermel's (Rothermel, 1972) and Fire Behavior Prediction (FBP) (Forestry Canada, 1992) fire models, which are two of the most widely employed. Statistical fire models are based on test fires to predict fire parameters such as fire intensity and rate of spread. Lastly, probabilistic fire models provide contingency tables, which are used to predict the rate of spread by means of the

fuel type, fuel moisture and wind speed (Albright and Meisner, 1999). These fire models are implemented in decision support systems (Andrews and Queen, 2001), providing a prediction of fire parameters such as rate of spread, fire intensity and fuel moisture (Arroyo et al., 2008). Several decision support systems have been developed in different countries and providing different outputs. Most of the decision support systems consist of fire danger systems, in other words they provide a prediction of fire parameters or a gradient of fire hazard from weather, topography and fuel data. The McArthur Grassland Fire Danger Rating System (McArthur, 1966; McArthur, 1967) from Australia, the National Fire Danger Rating System (NFDRS) (Deeming et al., 1972; Deeming et al., 1977) from USA and, the Canadian Forest Fire Danger Rating System (CFFDRS) (Stocks et al., 1989) from Canada, are developed from experimental wildfires and rate the risk of wildfires at broad scales. They use weather parameters or indices, such as the Canadian Fire Weather Index (FWI) (Wagner, 1985; Wagner, 1987) for the CFFDRS, and fuel information to predict wildfire danger and potential over large areas. One of the most used decision support systems is BEHAVE (Burgan and Rothermel, 1984; Andrews, 1986) and its subsequent update BehavePlus (Andrews, 2009; Andrews, 2014) from USA, which predicts fire parameters at a local scale. Other decision support systems such as NEXUS (Scott, 1999), FlamMap (Stratton, 2006) and Crown Fire Initiation and Spread (CFIS) (Alexander et al., 2006) use weather, topography and fuel information to assess crown fire spread and intensity through a number of indices predicted from a surface-crown fire system (Scott, 2006). In addition to BEHAVE, Fire Area Simulator (FARSITE) (Finney, 1998) is one of the most widely used decision support systems. FARSITE differs from the above mentioned decision support systems, since it provides a representation of the evolution of the fire perimeter over time in a Geographical Information System (GIS) format. As input data, it requires a more specific information from the tree crowns (i.e., canopy bulk density, canopy base height, canopy cover and canopy height), since models for surface and crown fire behavior are integrated. In addition to FARSITE, Wildfire Analyst (Ramírez and Monedero, 2011) also represents the evolution and intensity of the fire over time to analyze the firefighting capabilities of a wildfire. This tool implements the fire model proposed by Rothermel (1972) and further modified by Albini (1976), and

accepts Scott and Burgan (2005) fuel models as well as other custom fuel models. Decision support systems require fire models to predict fire hazards, but also fuel types to include the variability of tree species and vertical and horizontal forest structure. There are many fuel type classifications according to ecosystems and countries. Some of the most widely employed are NFDRS (Deeming et al., 1977), Northern Forest Fire Laboratory (NFFL) (Albini, 1976; Burgan and Rothermel, 1984) and Fuel Classification System (FCCS) (Sandberg et al., 2001) from USA; FBP system fuel types (Forestry Canada, 1992) from Canada; and Prometheus (Prometheus, 1999) from Europe, which is adapted from NFFL for Mediterranean ecosystems. These fuel type classifications are employed in the above mentioned decision support systems. For instance, NFDRS fuel type classification is employed by NFDRS, NFFL by BEHAVE and FARSITE, and FBP system fuel types by CFFDRS (Arroyo et al., 2008). Instead of fuel type classifications, which are clusters of forest attribute values, more specific values of some forest attributes may also be employed as inputs in some decision support systems to predict wildfire behavior (García et al., 2011) (hereafter referred as forest fuel attributes). These forest fuel attributes describe the vertical and horizontal forest structure as well as the fuel load, and are canopy bulk density, canopy fuel load, canopy height, canopy base height, and canopy cover (Cruz et al., 2003; Graham et al., 2004). Canopy bulk density corresponds with the amount of fuel per unit of volume (Keane et al., 2005). It is one of the most significant attributes to predict crown fire behavior, since it is related to the spread rate between crowns (Cruz et al., 2003; Riaño et al., 2004; Keane et al., 2005; Skowronski et al., 2011). Canopy fuel load is defined as the amount of fuel that may potentially be consumed per unit of area (Skowronski et al., 2011). Canopy bulk density and canopy fuel load may be extracted direct or indirectly. Direct methods are referred to destructive sampling to quantify biomass, which is costly and difficult to implement (García et al., 2011). Once direct methods are carried out, allometric equations may be generated using forest biometric measurements (i.e., diameter at breast height: DBH, height, and tree species) as independent variables and biomass from direct methods as dependent variables (García et al., 2011; Skowronski et al., 2011). These allometric equations facilitate estimation of fuel attributes at field campaigns. Canopy height may have different definitions. It may be defined as either the

average height of the 100 tallest trees in a hectare (Assmann, 1961; Assmann, 1970) or the maximum height where a minimum value of canopy bulk density is reached (Reinhardt et al., 2006). Canopy height influences on wind speed reduction and fuel moisture content, and therefore it indirectly affects on crown fire occurrence (Reinhardt et al., 2006). Canopy base height is the lowest height where fuel may potentially be consumed (Cruz et al., 2003; B. Peterson et al., 2005; García et al., 2011). This attribute is crucial to predict crown fires, it defines the gaps between understory vegetation and tree crowns (Graham et al., 2004; Keane et al., 2005). Some studies (Graham et al., 2004; Keane et al., 2005; García et al., 2011) mention canopy cover as another significant forest fuel attribute. It is defined as the proportion of land covered by tree crowns from a zenithal view (García et al., 2011). Canopy cover is related to the potential for fire spread between crowns (Graham et al., 2004; García et al., 2011). Therefore, there is a set of attributes (i.e., canopy bulk density, canopy fuel load, canopy height, canopy base height, and canopy cover) that may be used as inputs in decision support systems to predict wildfire spread more accurately, instead of fuel type classifications.

Fires not only spread horizontally, but also vertically among the different vertical strata. Nevertheless, fire models for the mentioned decision support systems (e.g., FARSITE, BEHAVE) represent the spread of wildfires in a two-dimensional space. Therefore, these tools do not include heterogeneity of crown fuels, which results in the most dangerous wildfires (Parsons et al., 2011); and only predict the wildfire behavior considering the interaction fuel-atmosphere (Mell et al., 2007), which involves that they only require the terrain, wind and fuel parameters as input. On the other hand, more recent tools for predicting wildfire behavior, such as WFDS (Wildland-Urban Interface Fire Dynamics Simulator) (Mell et al., 2007) and Firetec (Linn, 1997; Linn et al., 2002), use physical fire models based on computational fluid dynamics methods (Sullivan, 2009; Hoffman et al., 2016), which include 3D simulations of wildfire behavior. These tools not only consider the interaction fuel-atmosphere, but also the interactions fuel-fire and fire-atmosphere. These interactions consist of the generation of combustion gases and their fluxes and the reaction of fire and its plume to local winds, respectively, which have an influence on the fire spread (Mell et al., 2007). Therefore, these tools simulate fuels and wildfire behavior

with much more detail (Pimont et al., 2016). Nonetheless, given that these tools consider vegetation fuels as heterogeneous, they require more information from the different vertical strata (Mell et al., 2011), such as the three-dimensional distribution of fuels from individual trees and understory vegetation (Pimont et al., 2016).

However, this new generation of 3D fire behavior models needs more precise and detailed information related to the abovementioned fire fuel attributes, so new remote sensing techniques and systems, in particular those based on laser scanner (aerial and terrestrial) should be explored to fill the gap between the high 3D resolution required by the new fire behavior models and the current capabilities offered by large area remote sensing systems.

1.1.2. Estimation of fuel attributes using discrete airborne laser scanning

Forest fuel attributes have been traditionally estimated by field campaigns and destructive sampling, which are very costly and limited to small spatial extents (Hyde et al., 2005; Riaño et al., 2007). Therefore, mapping forest fuel attributes over large areas was a tedious and inaccurate task. Nevertheless, the use of remote sensing techniques has facilitated this task in the last decades. Remote sensing techniques, and more specifically laser scanning, have demonstrated their potential to estimate vertical and horizontal forest structure, and other forest attributes (Dubayah and Drake, 2000; Lim et al., 2003; Wulder et al., 2012; Hevia et al., 2016; Bottalico et al., 2017). Laser scanning consists of a sensor installed on a platform (i.e., airborne: ALS, spatial, terrestrial: TLS, drone, backpack, etc.) emitting thousands of laser pulses per second and registering the laser response from intercepted objects. Laser scanning data collected are stored in 3D point cloud format, providing X, Y, Z coordinates and an intensity value related to the physical properties of the intercepted object and the laser wavelength. These data are known as discrete laser scanning. Regarding ALS, many studies have demonstrated the capacity of discrete ALS (ALS$_D$) to estimate forest fuel attributes. Riaño et al. (2003) described the existing relationship between some ALS$_D$ metrics and forest fuel attributes canopy bulk density, canopy fuel load, canopy height, canopy base height and canopy cover.

Riaño et al. (2004) went a step further and estimated at plot- and tree-level the mentioned forest fuel attributes, except for canopy cover, by using a single ALS$_D$ metric as independent variable in regression models. Andersen et al. (2005) estimated the same forest fuel attributes, however, they used several metrics in the regression models instead of using a single one. Similar to Riaño et al. (2004), Morsdorf et al. (2006) only used a single ALS$_D$ metric, but in this case to estimate canopy cover, and using hemispherical photographs as field data. Popescu and Zhao (2008) estimated canopy height and canopy base height at tree-level using a new voxel-based approach. On the other hand, Hopkinson and Chasmer (2009) also estimated canopy cover by using hemispherical photographs as field data, and they carried out the analysis in seven different sites distributed in five Canadian ecozones. Erdody and Moskal (2010) used ALS$_D$ metrics alone and in combination with high resolution color near-infrared aerial imagery to estimate canopy bulk density, canopy fuel load, canopy height and canopy base height. They found that combination of ALS$_D$ and imagery performed the best results, however, accuracy of ALS$_D$ alone in estimating forest fuel attributes outperformed the use of imagery alone. Zhao et al. (2011) also estimated forest fuel attributes using ALS$_D$ data and analyzing how the accuracy varies when using different machine learning models. They observed that Support Vector Machine and Gaussian processes reached better results than traditional approaches such as linear regression. Finally, Hevia et al. (2016) observed an influence of thinned and unthinned plots on ALS$_D$ metrics. Overall, estimate results show a high correlation between ALS$_D$ metrics and forest fuel attributes, resulting in a coefficient of determination (R^2) around 0.81 ± 0.07, 0.80 ± 0.19, 0.93 ± 0.04, 0.79 ± 0.06 and 0.74 ± 0.01 for canopy bulk density, canopy fuel load, canopy height, canopy base height and canopy cover, respectively, in the studies analyzed. As previously mentioned, these attributes are key forest attributes used as inputs in decision support systems to predict wildfire behavior. Therefore, ALS$_D$ has demonstrated its potential to estimate key forest fuel attributes over large areas, being widely employed since it reduces considerably time and work with respect to traditional field campaigns.

1.1.3. Full-waveform airborne laser scanning

Among laser scanning technology, full-waveform laser scanning goes one step further than discrete laser scanning. Instead of a discrete 3D point cloud, full-waveform laser scanning registers the complete signal emitted by the sensor and it is stored in wave form (Mallet and Bretar, 2009). Waveform amplitude values depend on the wavelength and physical properties of the intercepted objects (Song et al., 2002; Guo et al., 2011; Hermosilla et al., 2014a), and angle of incidence (Kukko et al., 2008). Contrary to discrete laser scanning, full-waveform laser scanning provides a continuous response along the trajectory of the laser pulse, and hence more information is provided from the different vertical strata of the vegetation (Mallet and Bretar, 2009). For instance, understory vegetation, which is key for ladder fuels to spread fire from surface to crowns, is retrieved in more detail by full-waveform laser scanning (Anderson et al., 2016; Hancock et al., 2017). In the last decades, several studies have been carried out using airborne (ALS_{FW}) and spaceborne full-waveform laser scanning. For the latter, it is worth mentioning Geoscience Laser Altimeter System (GLAS), which was the first laser scanning for continuous global earth observation, onboard the Ice, Cloud, and land Elevation Satellite (ICESat) until 2008. This system had a large footprint size (~50-65 m), and was originally launched to study changes in ice sheets, atmospheric properties and clouds (Abshire et al., 2005; Mallet and Bretar, 2009). Nevertheless, it has also been successfully used to estimate forest fuel attributes thanks to registering the full-waveform (Lefsky et al., 2007; García et al., 2012). Regarding ALS_{FW} systems, they can be divided according to its footprint size. Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) and its improved version, Laser Vegetation Imaging Sensor (LVIS) (Blair et al., 1999), are two prototypes developed by NASA with a footprint size of 5-15 m and 5 m, respectively. These systems were designed to characterize the vertical forest structure and forest fuel attributes with promising results (Lefsky et al., 1999; Means et al., 1999; Harding et al., 2001; Drake et al., 2003; Hyde et al., 2005). Compared to laser scanning systems with a small footprint size (i.e., < 1 m), laser beams with a large footprint size are less intercepted by vegetation strata and reach the ground (Fieber et al., 2015). Nevertheless, their wide footprint size limits its

resolution in retrieving the vertical forest structure, and its high pulse energy and low pulse rate limit its spatial sampling (Wulder et al., 2012). Regarding ALSFW systems with a small footprint size, some studies have demonstrated the potential of ALSFW to estimate forest attributes (Cao et al., 2014; Hermosilla et al., 2014a), including some of the forest fuel attributes (i.e., canopy bulk density, canopy fuel load, canopy height and canopy base height). A few studies have assessed the accuracy differences between ALSD and ALSFW for the estimation of some forest fuel attributes, finding more accurate results for ALSFW (Anderson et al., 2016; Hancock et al., 2017). Nevertheless, despite of its potential to characterize understory vegetation and structure, ALSFW has received less attention than ALSD because of three main limitations: (*i*) ignorance of data, (*ii*) large amount of data, and (*iii*) lack of processing tools. Ignorance of data is reflected in users, researchers, forestry consulting companies and some companies in charge of data collection. This is due to the relative novelty of ALSFW data and the shortage of researchers working on this. Simultaneously, the latter is due to the large amount of data to deal with, which makes tests much longer and powerful processors are often required. Additionally, the lack of processing tools makes that most researchers must have programming skills to research on ALSFW applications. This situation may be starting to change with the recent launching of the new full-waveform laser scanning system, called Global Ecosystem Dynamics Investigation (GEDI), onboard the International Space Station (ISS) (Dubayah et al., 2020). Some new processing tools to process GEDI data, such as rGEDI for R (Silva et al., 2020) and Gedi for Python (Camacho, 2020), along with the availability of data, will allow more users to investigate the potential of full-waveform laser scanning to characterize the vertical forest structure and forest fuel attributes.

1.2.   Aims and objectives

The general objective of this thesis is the development of processing and analysis methods based on full-waveform airborne laser scanner data to characterize the vertical forest structure, in particular the understory vegetation. It arises from the need to reduce the limitations and to present the potential of ALSFW in forestry applications.

Simultaneously, six hypotheses are raised with their respective six specific objectives to reach the main objective. These hypotheses and specific objectives are as follows:

Hypothesis 1: Computation of ALS$_{FW}$ metrics through the voxelization procedure, and subsequently the prediction of forest structure attributes, are influenced by a variable pulse density. This influence may be reduced by modifying the voxel parameters (i.e., voxel size and assignation value).

Objective 1: To analyze the influence of pulse density, voxel parameters and regression methods on ALS$_{FW}$ metric values and on forest structure attributes estimates, identifying those parameters and quantifying their relations to be able to tune their values in order to considerably reduce this influence in practice.

Hypothesis 2: Understory vegetation in Mediterranean forests is a key strata in characterizing wildlife habitats and mitigation of forest fires. Metrics to detect and characterize understory vegetation are crucial for modeling forest fuel attributes and forest structure, and ALS$_{FW}$ is a more suitable technique to capture information from understory. There is a lack of ALS$_{FW}$ processing tools integrating these tasks and able to perform radiometric correction of data.

Objective 2: To compile a set of methods to process and analyze ALS$_{FW}$ data, including the relative radiometric correction of the data to reduce the effect of the different angles of incidence and local altitude variations during the data acquisition process, the extraction of most ALS$_{FW}$ metrics as proposed in the literature, as well as new metrics focused on understory vegetation, integrating them in a new software available to use for the scientific community.

Hypothesis 3: The scan angle has an influence on the amplitude values of the waveforms, and therefore on the ALS$_{FW}$ extracted. The application of radiometric correction is expected to reduce these differences in the amplitude values. As well as ALS$_{FW}$ metric values, forest fuel attributes estimated using

these metrics may also be influenced by scan angle differences, and radiometric correction is expected to reduce this effect.

Objective 3: To assess the influence of the scan angle of ALS data acquisition and the application or not of a radiometric correction on (*i*) the extraction of an ALS$_{FW}$ metric; and (*ii*) modeling three of the most relevant forest fuel attributes—canopy fuel load (CFL), canopy height (CH), and canopy base height (CBH).

Hypothesis 4: ALS and TLS configurations are limited to detect lower and upper strata, respectively, because of the position of the sensors. This limitation depends on the forest ecosystem and density.

Objective 4: Characterize the signal occlusion along the vertical forest structure using different laser scanning configurations (i.e., ALS$_D$, ALS$_{FW}$, and TLS) in contrasted ecosystems with different canopy covers to determine how reliable the resulting vertical distribution profiles are, based on the amount of occlusion and the lack of information.

Hypothesis 5: Given that ALS$_{FW}$ registers the complete signal going through the vertical forest structure, and ALS$_D$ is a simplification of ALS$_{FW}$, understory vegetation can be detected and its density determined more accurately with ALS$_{FW}$ than with ALS$_D$.

Objective 5: Determine how understory vegetation density classes can be detected and further determined by ALS configurations, and whether ALS$_{FW}$ allows the detection and determination to a level of detail beyond ALS$_D$ capability.

Hypothesis 6: The new described ALS$_{FW}$ metrics in objective 2 may be used to accurately characterize the height, cover and volume of understory vegetation in a Mediterranean ecosystem.

Objective 6: Apply and validate the new metrics described in objective 2 derived using a voxel based approach to estimate understory height, cover, and volume in a Mediterranean forest ecosystem, proposing some practical recommendations for further development and testing ALS$_{FW}$ metrics.

1.3.   Thesis outline

This thesis is divided into six chapters, being this first chapter an introduction to the state-of-the-art and to the topics discussed in more detail in chapters 2-5. These four chapters are based on edited versions of four international scientific publications, one national scientific publication, and eight published conference papers. Lastly, the sixth chapter compiles the conclusions of the thesis, as well as raises possible lines for future research taking the current thesis as reference.

Figure 1.1 shows the development of the different chapters of the thesis in chronological order and their relation to the different publications carried out. The first two publications (one conference paper and one national scientific publication) described the estimation of forest attributes using ALS$_{FW}$, and how the accuracy of these estimates depends on the regression models used. This leaded to find the existing influence of pulse density on ALS$_{FW}$ metrics and on the estimate of forest attributes, which is addressed in Chapter 2. The thesis is divided in a methodological part (Chapters 2 and 3), including the development of a software tool, and the application and analysis of ALS$_{FW}$ for the characterization and detection of understory vegetation (Chapters 4 and 5).

In Chapter 2, the problem of the influence of pulse density on ALS$_{FW}$ metrics and on the estimate of forest attributes is addressed. As previously mentioned, a side-lap effect due to pulse density variation was detected when estimating forest attributes. In this chapter, the variation of the most common ALS$_{FW}$ metrics and of the estimates of forest attributes when forcing the variation of the pulse density, voxel parameters (i.e., voxel size and assignation value) and regression methods are analyzed. This analysis allows for better understanding the side-lap effect and how it can be reduced.
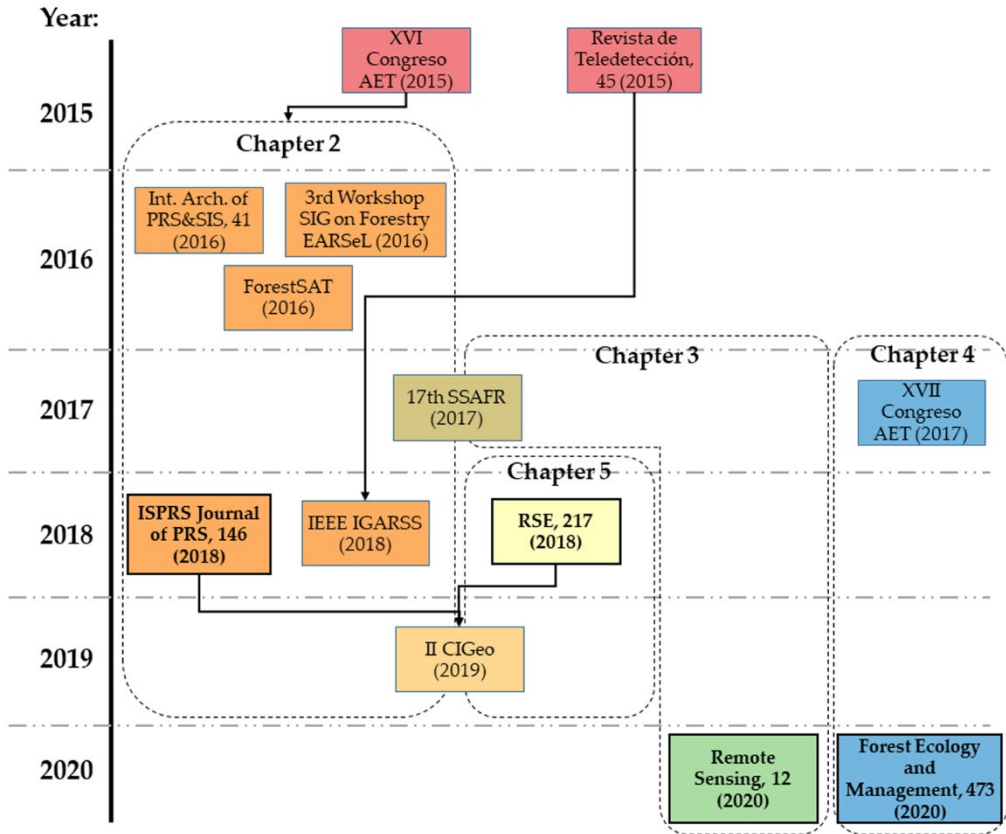
Figure 1.1. *Chronological diagram and structure of the research. Publications are colored according to the chapter they belong, and the four core publications are in bold and the edges highlighted.*

In Chapter 3, the influence of the scan angle and the relative radiometric correction on ALS_FW metrics and on the estimates of forest fuel attributes is assessed. This analysis and other ALS_FW data processing may be done in a new processing tool called *WoLFeX* (Waveform Lidar for Forestry eXtraction), which is also presented and described in this chapter. This processing tool is the basis for the application of methods in order to characterize understory vegetation presented in the following chapters.

In Chapter 4, the ability of two airborne platforms (ALS_D and ALS_FW) and one terrestrial platform (TLS) to characterize vertical forest structure is assessed, linking it with a new method to estimate signal occlusion in the different strata.

Moreover, the suitability of ALS_D and ALS_{FW} to determine understory vegetation density classes is also assessed, which is challenging for airborne platforms due to occlusion caused by overstory. All these analyses are carried out in two contrasted ecosystems (boreal and Mediterranean).

In Chapter 5, the line of Chapter 4 is followed, and it is focused on the characterization of understory vegetation in a Mediterranean forest using ALS_{FW}. The height, cover and volume of understory vegetation are estimated, and the results evaluated using TLS as reference data.

In Chapter 6, the conclusions of the different chapters are compiled to provide a global conclusion of the thesis. In addition to mentioning the most relevant points of the thesis, this chapter also discusses possible further research by continuing the line of the thesis.

Chapter 2

Influence of methodological parameters on full-waveform ALS metrics and forest attributes prediction

Rosmarinus officinalis

*"A la flor del romero,
romero verde,
si el romero se seca
ya no florece."*

Federico García Lorca

In this chapter, the effect that ALS$_{FW}$ pulse density, voxelization parameters (i.e., voxel size and assignation value), and regression methods have on ALS$_{FW}$ metric extraction and on the subsequent estimate of forest attributes are analyzed. This effect related to a variable pulse density due to differences in flight stripes overlap is called side-lap effect, and it is visually observed as a stripe where ALS$_{FW}$ metrics and forest attributes estimated differ from their neighborhood, although forest attributes are actually similar. Side-lap effect may lead to obtain wrong results when voxelizing ALS$_{FW}$ data. Hence, knowing the causes of side-lap effect and how to reduce it is fundamental to use properly ALS$_{FW}$ data for forestry applications by the voxelization procedure.

2.1.   Introduction

ALS data (i.e., ALS$_D$ and ALS$_{FW}$) are not homogeneous along a study area. Several factors such as scan angle, range from sensor to registered objects, topography or objects' structure prevent ALS point cloud from forming a regular mesh. Moreover, side-lap areas (i.e., where two or more flight stripes overlap), which are required to reduce occlusion, increase pulse density, and for flight stripes georeferencing, have a higher pulse density (see Figure 2.1). These pulse density variations affect metrics extracted from ALS$_{FW}$ and the subsequent forest attributes estimates. Therefore, ALS$_{FW}$ metrics extracted in two areas with identical forest features but different pulse densities may differ. In addition, given that forest attributes are estimated by using ALS$_{FW}$ metrics in regression models, the estimates are influenced by pulse density variations as well.

The influence of ALS$_D$ pulse density on estimates of forest attributes was analyzed in several studies (Table 2.1). All these studies present variations in estimates of forest attributes, however, since they were focused on different ecosystems and used different ranges of pulse densities, variations have different scales. Gobakken and Næsset (2008), Magnussen et al. (2010) and Jakubowski et al. (2013) observed that estimated attributes were not significantly affected by density until dropping 0.25 points·m^{-2} in the first study, and 1 pulse·m^{-2} in the last two. Analyzing specific groups of attributes, Magnussen et al. (2010), González-Ferreiro et al. (2012), Strunk et al. (2012),
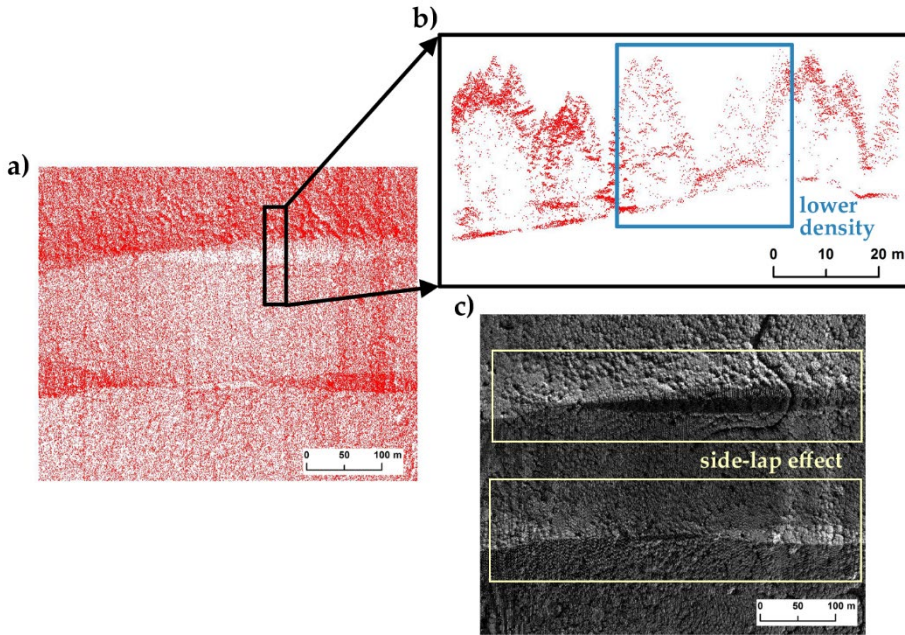
Figure 2.1. *Representation of the ALS density variation and subsequent side-lap effect in the ALSD point cloud from (a) a zenithal view and (b) a vertical transect, and (c) in the estimation of the ALSFW metric RWE (Return Waveform Energy).*

Treitz et al. (2012), Jakubowski et al. (2013) and Varo-Martínez et al. (2017) did not find significant influence of pulse density on attributes related to height, such as: mean, dominant, tree and Lorey's height, and mean height to live crown. According to Strunk et al. (2012) and Treitz et al. (2012), attributes related to tree density (i.e., number of stems and stem density) were not significantly affected either. However, Magnussen et al. (2010) observed on the reliability ratio that stem density was affected using low pulse densities. The reliability ratio was defined by Hansen et al. (2015) as the variance of a metric among sample plots divided by the total variance of the metric (i.e., the variance among sample plots plus the average variance within the plot). Regarding attributes related to trunk size, such as quadratic mean diameter (Treitz et al., 2012), DBH (Jakubowski et al., 2013), and basal area (Magnussen et al., 2010; González-Ferreiro et al., 2012; Strunk et al., 2012; Treitz et al., 2012; Jakubowski et al., 2013; Ruiz et al., 2014; Varo-Martínez et al., 2017), had no significant differences between different pulse densities, except for the basal area in a tropical forest in

a study carried out by Manuri et al. (2017). Among volume attributes (i.e., volume over bark, stem volume, gross total and merchantable volume), only volume over bark in González-Ferreiro et al. (2012) was significantly affected by pulse density variations. Additionally, Jakubowski et al. (2013) for shrub cover and height variables, Ruiz et al. (2014) for canopy cover, and Silva et al. (2017) for aboveground carbon, observed that they were not significantly affected either. Lastly, stem biomass and aboveground biomass were influenced by ALS_D pulse density in an Atlantic and tropical forest (González-Ferreiro et al., 2012; Manuri et al., 2017) but Treitz et al. (2012) did not find significant differences in aboveground biomass in a boreal forest using different densities. Overall, mass-related attributes such as aboveground biomass (AGB) are more influenced by pulse density than height attributes, although another factor affecting tree density, basal area and volume is the type of ecosystem.

The influence of pulse density on forest attributes estimated from ALS_D metrics has been widely studied in different ecosystems, however, less attention has been paid to how ALS_D metrics are influenced. Roussel et al. (2017) mentioned that even when the values of estimated attributes are stable for different pulse densities, ALS_D metrics are affected, since they are measures and are not statistically fitted. Gobakken and Næsset (2008) and other authors, such as Hansen et al. (2015) and Roussel et al. (2017), analyzed the effects of pulse density on ALS_D metrics. The first study computed height (e.g., percentiles, maximum, mean and coefficient of variation) and density metric differences between the initial point density (i.e., 1.13 points·m^{-2}) and thinned data (i.e., 0.25, 0.13 and 0.06 points·m^{-2}) at different sample sizes. They observed that the maximum height metric had large variations between point densities, these variations being even larger when point density decreased. The remaining metrics did not have a clear pattern. Hansen et al. (2015) computed seven ALS_D metrics: mean, maximum, variance, percentiles 10 and 90 of the above ground heights, and the proportion of points above the ground and above the mean. They observed that most of the metrics were not influenced by pulse densities, except for the maximum elevations that decreased with lower pulse densities. However, the reliability ratio increased for all the metrics as pulse density increased, until reaching a threshold where it remained stable. A possible

Table 2.1. *Summary of existing studies about the influence of ALSᴅ pulse density on estimates of forest attributes (adapted from Crespo-Peremarch et al. (2018a)).*

Study	Study Area	Ecosystem	Highest (HD)-lowest (LD) densities (pulses·m⁻²)	Estimated variables	Results (HD-LD)
Gobakken and Naesset (2008)	Våler, Southeastern Norway	Boreal forest	1.13-0.06 points·m⁻²	Hl: Lorey's height BA: basal area Vol: stand volume	Estimate differences: Hl ≈ 0.2–0.6 m BA ≈ 0.0–2.5 m²·ha⁻¹ Vol ≈ 5–30 m³·ha⁻¹
Magnussen et al. (2010)	Aurskog-Høland, Southeastern Norway	Boreal forest	2-0.25	Hl: Lorey's height BA: basal area V: volume over bark SD: stem density	R² (%): BA ≈ 79-72 V ≈ 85-80 Reliability ratio: Hl ≈ 1.0-0.9 BA ≈ 0.98-0.95 V ≈ 0.96-0.92 SD ≈ 0.96-0.81
González-Ferreiro et al. (2012)	Galicia, Northwestern Spain	Atlantic forest	8-0.5	Hm: mean height Hd: dominant height BA: basal area V: volume over bark Wcr: crown biomass Wst: stem biomass AGB: aboveground biomass	R² (%): Hm = 78.6-75.9 Hd = 84.6-86.5 BA = 67.8-69.2 V = 69.1-79.4 Wcr = 68.7-68.8 Wst = 73.2-82.7 AGB = 74.6-80.4
Strunk et al. (2012)	Western Washington State (USA)	Humid temperate – Pacific lowland mixed forest	3-0.05	ST: number of stems	nRMSE (%): ST ≈ 56-57

Table 2.1. (cont.) *Summary of existing studies about the influence of ALS$_D$ pulse density on estimates of forest attributes (adapted from Crespo-Peremarch et al. (2018a)).*

Study	Study Area	Ecosystem	Highest (HD)-lowest (LD) densities (pulses·m^{-2})	Estimated variables	Results (HD-LD)
Treitz et al. (2012)	Ontario, Canada	Boreal forest	3.2-0.5	Hm: mean height TH: tree height QMD: quadratic mean diameter BA: basal area GTV: gross total volume AGB: aboveground biomass SD: stem density	R^2 (%): Black spruce (BS), Intolerant hardwood (IH). Hm = 95.1-93.6 (BS); 77.3-76.7 (IH) TH = 92.3-90.3 (BS); 94.3-94.1 (IH) QMD = 86.3-83.8 (BS); 84.2-84.0 (IH) BA = 93.5-91.8 (BS); 83.7-82.3 (IH) GTV = 94.9-94.2 (BS); 83.7-82.3 (IH) GMV = 93.9-91.6 (BS); 87.7-87.3 (IH) AGB = 93.2-92.5 (BS); 78.8-77.5 (IH) SD = 88.8-86.1 (BS); 24.8-23.9 (IH)
Jakubowski et al. (2013)	Tahoe National Forest. Northern California (USA)	Mediterranean-climate forest	9-0.01	TH: tree height HTLCB: mean height lo live crown base BA: basal area DBH: diameter at breast height SC: shrub cover SH: shrub height	R^2 (%): TH = 86.8-52.4 HTLCB = 76.8-28.8 BA = 77.5-48.9 DBH = 59.7-38.0 SC = 53.1-11.9 SH = 45.9-29.0
Ruiz et al. (2014)	La Serranía de Cuenca, Central Spain	Mediterranean mountain forest	6-0.25 points·m^{-2}	V: volume AGB: aboveground biomass BA: basal area CC : canopy cover	R^2 (%): V ≈ 90.5-86.0 AGB ≈ 85.5-82.0 BA ≈ 87.0-83.0 CC ≈ 89.0

Table 2.1. (cont.) *Summary of existing studies about the influence of ALS$_D$ pulse density on estimates of forest attributes (adapted from Crespo-Peremarch et al. (2018a)).*

Study	Study Area	Ecosystem	Highest (HD)-lowest (LD) densities (pulses·m⁻²)	Estimated variables	Results (HD-LD)
Manuri et al. (2017)	Central Kalimantan, Indonesia	Tropical forest	2.8-0.01 points·m⁻²	AGB: aboveground biomass BA: basal area	R² (%): AGB ≈ (90.0)–(80.0,60.0) BA ≈ (90.0)-(70.0,40.0)
Silva et al. (2017)	Paraíba Valley, São Paulo, Brazil	Humid subtropical forest	10-5	AGC: aboveground carbon	R² (%): AGC = 82.17-81.79
Varo-Martínez et al. (2017)	Sierra de Los Filabres, Southeastern Spain	Semi-arid and Mediterranean forest	10-0.5	Hd: dominant height BA: basal area	R² (%): Hd = (97.0,94.0)-(95.0,93.0) BA = (92.0,88.0)-(93.0,87.0)

explanation for this might be that mean values of ALS_D metrics did not vary much due to pulse density. In contrast, the standard deviation increased for lower pulse densities, and hence the reliability ratio varied as well. Roussel et al. (2017) also analyzed how maximum height varied for different pulse densities. They concluded that metric variations were not only subject to pulse density, but additionally to ALS footprint size and canopy shape. The flatter the top canopy (i.e., fewer singularities), the lesser the difference between pulse densities. Pre-processing of ALS_D and ALS_{FW} data differ due to data characteristics. While ALS_D metrics can be recomputed by simply varying the number of points (i.e., pulse density), pre-processing of ALS_{FW} data is more complex and there are other parameters that may also be considered. This complexity may explain why the influence of pulse density on ALS_{FW} metrics and on estimate of forest attributes have received less attention. Furthermore, some published studies have analyzed the evolution of ALS_{FW} metrics by artificially reducing the pulse density. Crespo-Peremarch et al. (2016) observed side-lap effect in adjacent areas that were compared pairwise, with similar forest features but having different densities. It was found that ALS_{FW} metrics were influenced by density variations caused by flight stripe side-lap areas. A standard pre-processing method for ALS_{FW} metric extraction is voxelization (Hermosilla et al., 2014b). Once pseudo-vertical waveform is generated, ALS_{FW} metrics can be extracted. Changing the voxel size and the assignation value may diminish the side-lap effect without modifying the pulse density. Increasing the voxel size reduces the number of empty voxels, avoiding gaps in the pseudo-vertical waveform. On the other hand, changing the assignation value can avoid outliers from amplitude values, which is more likely when the voxel size increases.

Crespo-Peremarch et al. (2016) observed that the side-lap effect in ALS_{FW} metrics had an effect on forest attribute estimates as well, given that the latter are estimated through ALS_{FW} metrics. This influence was visually observed for a large area, resulting in a wrong mapping of forest attributes, with presence of side-lap effect due to pulse density variation. Therefore, correcting side-lap effect is essential to properly estimate forest attributes. Comparing ALS_{FW}

metrics obtained using different pulse densities may help to better understand how metrics are influenced and to reduce side-lap effect.

In this chapter, in section 2.2, the study area and the data used to carry out the analyses are described. In section 2.3, the different processing steps followed from raw data to the analyses of side-lap effect are enumerated. In section 2.4, the results of how side-lap effect influences on the ALS$_{FW}$ metrics and on the estimate of forest attributes are shown. In section 2.5, the discussion of the results is addressed. Finally, in section 2.6, the conclusions of this chapter are presented. This chapter sets out the Hypothesis 1: "Computation of ALS$_{FW}$ metrics through the voxelization procedure, and subsequently the prediction of forest structure attributes, are influenced by a variable pulse density. This influence may be reduced by modifying the voxel parameters (i.e., voxel size and assignation value)". To demonstrate this hypothesis, Objective 1 is addressed in this chapter: analyze the influence of pulse density, voxel parameters and regression methods on ALS$_{FW}$ metric values and on forest structure attributes estimates, identifying those parameters and quantifying their relations to be able to tune their values in order to considerably reduce this influence in practice.

2.2. Study area and data

The study area (2,258 ha) is located in Panther Creek (Oregon, USA) (Figure 2.2), in the Cascade mixed forest ecoregion (Bailey, 1980). The dominant species is Douglas-fir (*Pseudotsuga menziesii*) very occasionally mixed with other conifers such as western red cedar (*Thuja plicate*), western hemlock (*Tsuga heterophylla*) and grand fir (*Abies grandis*), and broad-leaved species such as bigleaf maple (*Acer macrophyllum*) and red alder (*Alnus rubra*). Tree heights are variable due to harvesting, being up to 60 m. Altitudes in the total extent of the study area range from 100 to 700 m.

ALS$_{FW}$ data were acquired in July 2010 using a Leica ALS60 over 3,264.51 ha, with a pulse density ranging from 2 to 168 pulses·m^{-2}, and an average of 10.4 pulses·m^{-2} (Figure 2.2c). Data were registered at an average flight altitude of 900 m above ground level, at 105 kHz pulse frequency, and with a scan angle of ±14°
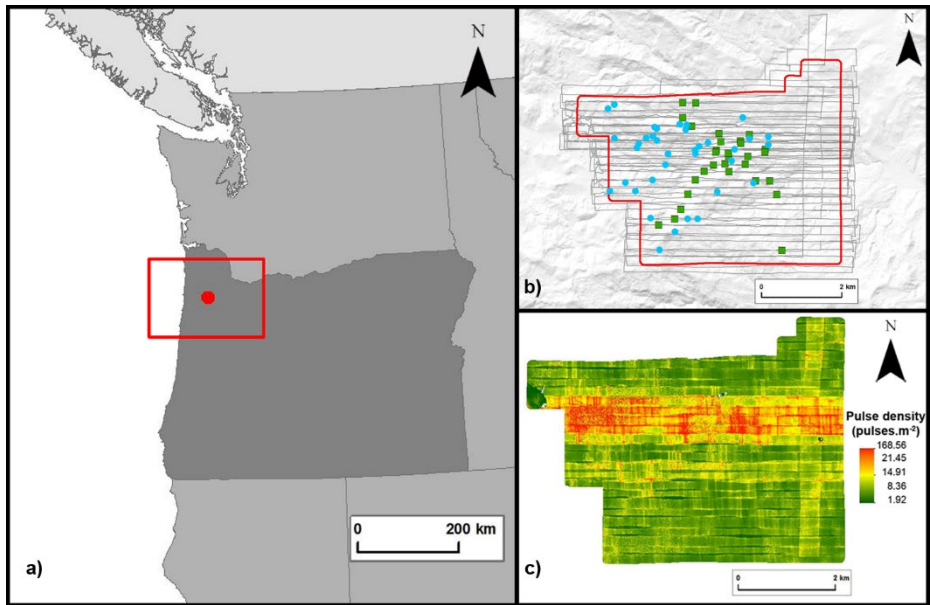
Figure 2.2. *(a) Study area location in the USA Pacific Northwest, (b) flight trajectories, sample (green) and plot (blue) locations within the study area limits (red), and (c) pulse density.*

from nadir. The study area was covered with flight stripe side-lap of ≥50% (≥100% overlap). Waveform amplitudes were recorded in 256 bins with a temporal sample spacing of 2 ns (i.e., 0.3 m) and a footprint size of ~0.25 m. In addition, a DTM with 1 m spatial resolution was provided by the company that registered ALS_FW data, and its vertical accuracy assessed using 33 GPS ground control points, obtaining a root-mean-square error (RMSE) of 0.19 m.

Regarding field works, a total of 84 circular plots with 16 m radius were measured. Within each plot, the dominant species and every tree with a DBH greater than 2.5 cm were registered. As a result, there were 47 plots where Douglas-fir was dominant, and 37 with mixed species. Afterwards, forest attributes were estimated using collected field data and allometric equations described by Standish et al. (1985).

2.3. Methods

Two main analyses were carried out: the influence of side-lap effect on (*i*) ALSFW metrics and (*ii*) on the estimate of forest attributes.

The overall strategy followed to analyze the influence of side-lap effect on ALSFW metrics and on the estimate of forest attributes is illustrated by the flow diagram of Figure 2.3. Six procedures were applied to carry out the analyses. As the first procedure (1), a radiometric correction and denoising were required prior to any analysis including ALSFW metrics. The second procedure (2) consisted of a pulse density reduction so as to simulate the pulse density variation of side-lap effect in samples. As the third procedure (3), ALSFW data were clustered into voxels, where different sizes and assignation values (i.e., voxel parameters) were tested to assess its influence on side-lap effect. The forth procedure (4) involved the extraction of the different ALSFW metrics from the voxelization carried out with different pulse densities and voxel parameters. As the fifth procedure (5), the first analysis is based on how the side-lap effect influences on the ALSFW metrics extracted in Procedure 4. The evolution of ALSFW metrics is analyzed as pulse density and voxel parameters are modified. The last procedure firstly involved estimation of forest attributes (6a) from ALSFW metrics extracted in Procedure 4. Lastly, the second analysis involved analyzing the evolution of accuracy in the estimation of forest attributes by also modifying pulse density and voxel parameters (6b).

2.3.1. Radiometric correction and waveform denoising

Prior to voxelization, the radiometric correction is a key pre-processing step for ALSFW data (Procedure 1), since the metrics extracted depend on the amplitude values of the waveform. These amplitude values registered vary according to sensor and atmospheric factors, as well as the local angle of incidence and the range from the sensor to the object intercepted. Thus, an object may be registered with different amplitude values depending on these factors. The goal of the radiometric correction is to reduce these differences of amplitude values to make amplitude values independent of these factors. Moreover, the radiometric correction also allows for converting digital numbers
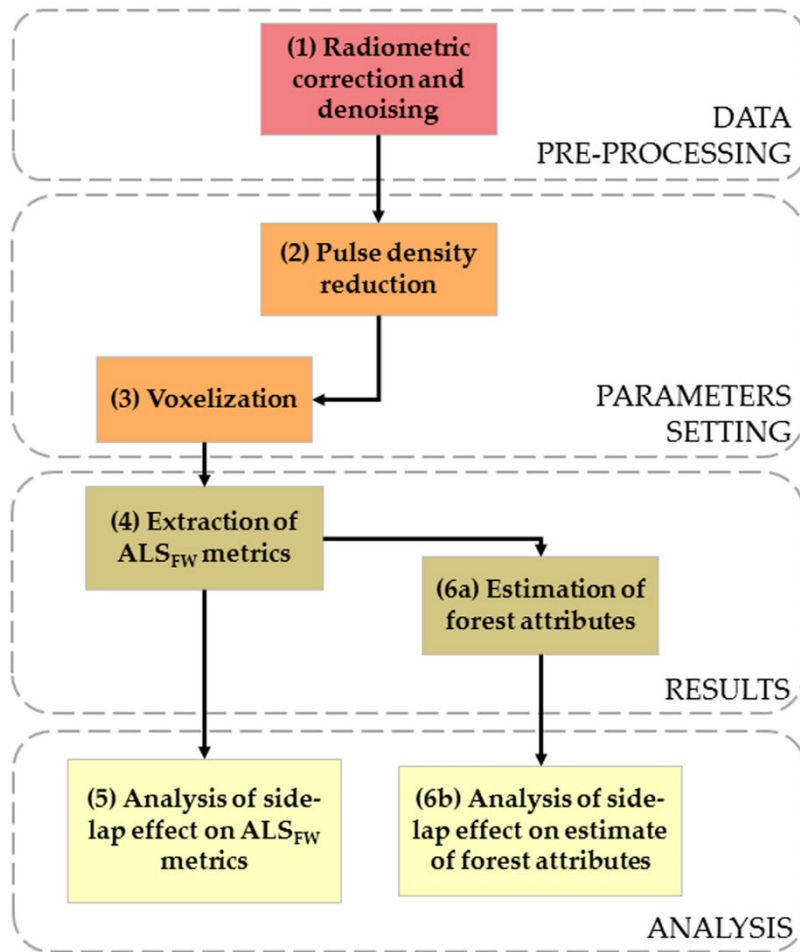
Figure 2.3. *Overall processing flowchart implemented to analyze the side-lap effect on ALS_FW metrics and on the estimate of forest attributes.*

of amplitude values to physical units such as reflectance by means of ground truth data from targets. Therefore, there are two main approaches of radiometric correction: relative and absolute. The relative radiometric correction reduces radiometric differences between flight stripes due to some factors such as the local angle of incidence or the range to the object registered. The absolute radiometric correction however also reduces radiometric differences between data acquired in different days and with different sensors (Wagner, 2010), and allows for converting amplitude values to physical units.

To deal with this correction, properties from targets distributed along the study area are required (e.g., targets or paved roads). In this research, the relative radiometric calibration was considered, since ground-truth data from targets were not available and there were no paved roads whose radiometry was known. Equation 1, described by Kashani et al. (2015), considers the distance from the sensor to the registered object and the angle of incidence. The angle of incidence depends on the scan angle, the slope and the aspect (Figure 2.4). Knowing these values, the angle of incidence may be calculated for those ALS$_{FW}$ returns corresponding to the ground. However, objects present on the ground, such as trees, do not follow the terrain slope. In general, trees grow vertically, and their branches almost horizontally. Since it is unfeasible to know the angle of incidence on tree leaves and branches without an accurate external data (e.g., TLS), for the computation of the angle of incidence the terrain slope is considered for ALS$_{FW}$ ground returns, and a null slope (i.e., 0%) for the rest of ALS$_{FW}$ returns.

$$A_C = A * \frac{R_i{}^n}{R_{ref}{}^n} * \frac{1}{\cos \alpha}$$

Equation 1

where A_C = corrected amplitude,

A = amplitude to be corrected,

R_i = range from the sensor to the object,

R_{ref} = reference range set to 1,000 m for this study,

n = power of the range,

α = local angle of incidence.

After the radiometric correction, waveforms still contain noise. The process followed to remove it was the one described by Hermosilla et al. (2014b) (Procedure 1). This process consists of first discarding noisy waveforms, and then removing noise from non-noisy waveforms. In this process, waveforms are tagged as noisy when all the amplitudes are below a threshold defined as the mean plus four times the standard deviation of the waveform amplitudes
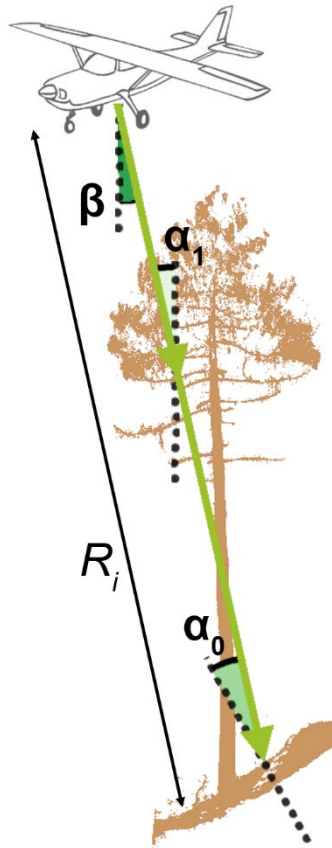
Figure 2.4. *Relation between the angle of incidence at the ground (α_0) and at a branch (α_1), the scan angle (β), and the range from the sensor to ALS$_{FW}$ return corresponding with the ground (R_i).*

(Lefsky et al., 2005). All these waveforms are removed, and only the rest of waveforms (i.e., non-noisy) are used for the next step. For the remaining waveforms, two denoising filters may be applied. First, noise, defined as 133% of the mode of the amplitudes, is subtracted from each amplitude value of the waveform. Then, a Gaussian filter is applied to eliminate the remaining noise. 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.

2.3.2. Pulse density reduction

For the first analysis (i.e., analyzing the influence of side-lap effect on ALS$_{FW}$ metrics), a total of 30 samples were selected from the study area where conifers were dominant (Figure 2.2b). These samples were located where pulse density was higher (i.e., side-lap areas) in order to be able to test a higher number of pulse density variations. Despite the fact that in side-lap areas scan angles are usually higher, we assumed the influence of scan angle negligible after the use of the radiometric correction. The polygon samples were square-shaped with an area of 804.25 m^2 each, this is the equivalent area of 16 m radius circular plots. For the second analysis (i.e., analyzing the influence of side-lap effect on estimate of forest attributes), a subsample of 36 conifer-dominant plots from the 84 measured was selected (Figure 2.2b) to work with plots having the highest pulse densities, and then testing a higher number of pulse density variations as well. In this thesis, we differentiate between the terms "sample" and "plot", requiring the latter ground-truth data from field works.

The analysis of the side-lap effect on ALS$_{FW}$ metrics requires forcing a pulse density reduction (Procedure 2) to assess its influence in samples whose forest attributes remain constant. In this way, the effect due to a pulse density variation may be assessed in more detail, since several pulse densities may be tested. This is a crucial processing step, since although some adjacent samples may have similar forest attributes, only identical forest attributes may be guaranteed using the same sample. Pulse density may be calculated by identifying the number of pulses in a sample and its area (Equation 2). The number of pulses in the sample may be easily calculated by counting the number of first or last returns in the sample. To force a pulse density reduction, it is needed to calculate the number of pulses in the sample by knowing the pulse density to reach and the area of the sample. Afterwards, the new number of pulses calculated are randomly selected among all the pulses in the sample (see Figure 2.5). This resulting new dataset is therefore voxelized in the next section to assess the side-lap effect at the new pulse density indicated.

Pulse densities tested differ according to the analysis. Samples from the analysis of the side-lap effect on ALS$_{FW}$ metrics were located in areas with higher

$$Pulse\ density = \frac{Number\ of\ pulses\ in\ the\ sample}{Area\ of\ the\ sample}$$

Equation 2
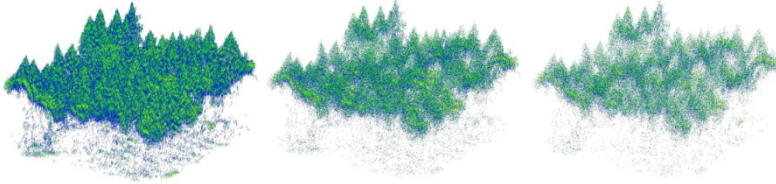
Figure 2.5. *Examples of gradual reduction of pulse density, representing values of 20, 4, and 2 pulses·m⁻² from left to right.*

pulse densities, therefore a wider range of pulse densities could be tested. The pulse density was reduced from 16 to 2 pulses·m⁻² with an interval of 1 pulse·m⁻², resulting in 15 different density values. The initial pulse density was selected considering the maximum and common pulse density value found in the 30 samples. Plots from the analysis of the side-lap effect on the estimate of forest attributes correspond with areas where field data are available. Because of this, the selection of plots for the analysis is restricted, and therefore a smaller range of pulse densities could be tested. In this case, the maximum and common pulse density value was set to 9 pulses·m⁻² to work with 36 plots. The pulse density was reduced from 9 pulses·m⁻² to 1 pulse·m⁻² every 0.5 pulses·m⁻².

2.3.3. Voxelization

Once ALS_FW data have been radiometrically corrected and noise removed, and selected according to pulse density, voxelization procedure (Procedure 3) may be carried out. However, prior to voxelization a height normalization is required to avoid mixing different vertical strata. The DTM described in section 2.2, and generated from the original pulse densities was used for height normalization. Up to this point, the terms "waveform" or "pulse" have been employed, since denoising and pulse density reduction was done at waveform-level. However, voxelization requires a point cloud, and this is obtained from the discretization of the waveforms. Waveforms are defined by a set of bins (Figure 2.6), which have an amplitude that is provided either in the LAS file or

in a WDP file. Nevertheless, XYZ coordinates of each bin must be calculated prior to voxelization. To do this, XYZ coordinates from at least one return belonging to the waveform, the return point waveform location of this return (i.e., offset in time units from the first bin of the waveform to the return point location), the temporal sample spacing value (i.e., distance between two consecutive bins), and the waveform line parameters Xt, Yt and Zt are required. All these data may be retrieved from the LAS file and are described in Figure 2.6. Firstly, the coordinates of the anchor point (i.e., the first bin of the waveform) (see Figure 2.6) must be calculated (Equation 3). To do this, the XYZ coordinates of the anchor point are calculated from the known coordinates of the return point, the waveform line parameters, and the offset between the anchor and the return point:

$$X_A = X_{Ri} - Xt * RPWL$$

$$Y_A = Y_{Ri} - Yt * RPWL$$

$$Z_A = Z_{Ri} - Zt * RPWL$$

Equation 3

where X_A, Y_A, Z_A = XYZ coordinates of the anchor point in meters,

X_{Ri}, Y_{Ri}, Z_{Ri} = XYZ coordinates of the return point i in meters,

Xt, Yt, Zt = waveform line parameters,

$RPWL$ = return point waveform location, which is the offset in picoseconds (10^{-12}) between the anchor and the return point i.

Secondly, XYZ coordinates of all the bins contained in the waveform may be calculated from the anchor point coordinates (Equation 4). To do this, the offset in each dimension between the anchor point and the corresponding bin is calculated using the waveform line parameters, the relative position of the corresponding bin in the waveform from the anchor point, and the distance between two consecutive bins. Next, this offset in each dimension is added to the corresponding coordinate of the anchor point:

$$X_i = X_A + Xt * i * TSS$$

$$Y_i = Y_A + Yt * i * TSS$$

$$Z_i = Z_A + Zt * i * TSS$$

Equation 4

where X_i, Y_i, Z_i = XYZ coordinates of current bin in meters,

X_A, Y_A, Z_A = XYZ coordinates of the anchor point in meters,

Xt, Yt, Zt = waveform line parameters,

i = bin position with respect to the anchor point (i.e., $i = 0$ for the anchor point),

TSS = temporal sample spacing in picoseconds (10^{-12}).

Following this procedure, XYZ coordinates from each bin of each waveform may be calculated. These coordinates together with its amplitude value yield to a discrete point cloud from ALS$_{FW}$ data, which may be further voxelized.

Voxelization is a procedure to drastically reduce the amount of data and facilitate the extraction of metrics. It consists of clustering data within voxels, which are 3D pixels or rectangular prisms (Figure 2.7a). A voxel must be defined by its XYZ dimensions. Usually, X and Y dimensions coincide, however Z dimension may differ, since horizontal and vertical accuracies differ in ALS systems. To process ALS$_{FW}$ data, XY dimensions are related to the footprint size of the laser pulse and the pulse spacing, while Z dimension is defined by the temporal sample spacing value (see Equation 5).

$$S = c \times \frac{TSS}{2}$$

Equation 5

where S = distance in meters between two consecutive bins,

c = speed of light in m·s^{-1},

TSS = temporal sample spacing in seconds.
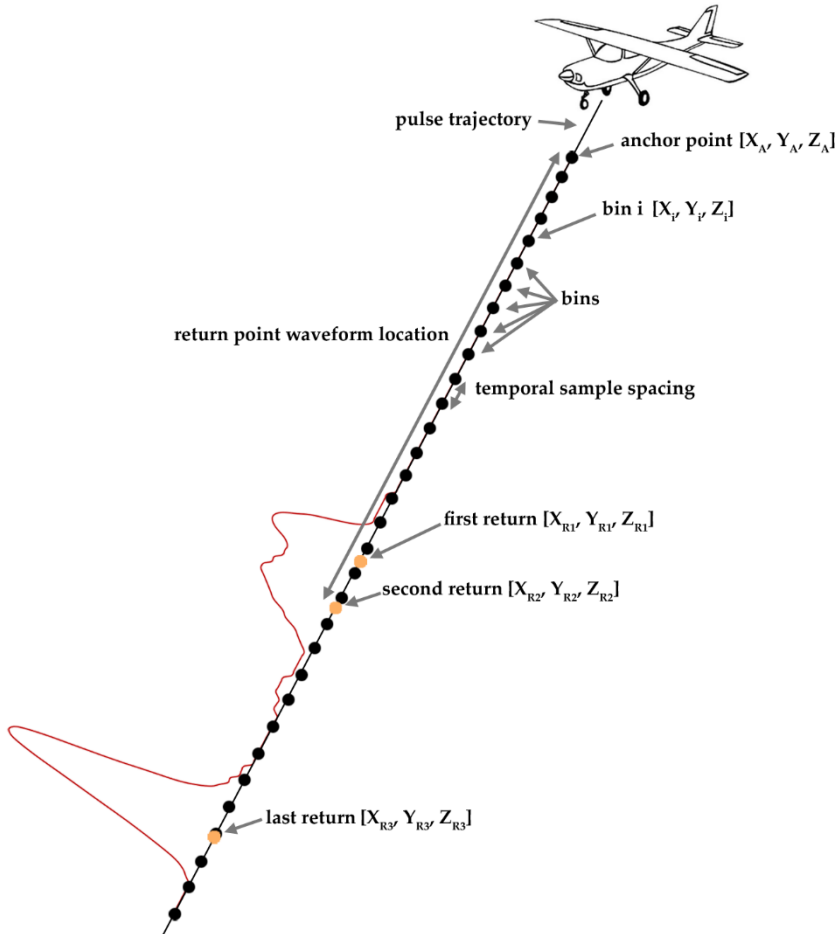
Figure 2.6. *Description of waveform elements.*

On the other hand, an assignation value must also be assigned for the voxelization. This is the statistical procedure (i.e., maximum, mean, median, etc.) employed to calculate the resulting voxel value from all the data contained inside (Figure 2.7b). Therefore, to carry out the voxelization procedure is necessary to set two voxel parameters: voxel size and assignation value. As a result, a regular voxel grid with their corresponding voxel values is obtained (Figure 2.7a).

Different voxel sizes and assignation values were tested in the analyses. In the analysis of side-lap effect on ALS_FW metrics, 14 voxel size variations in XY
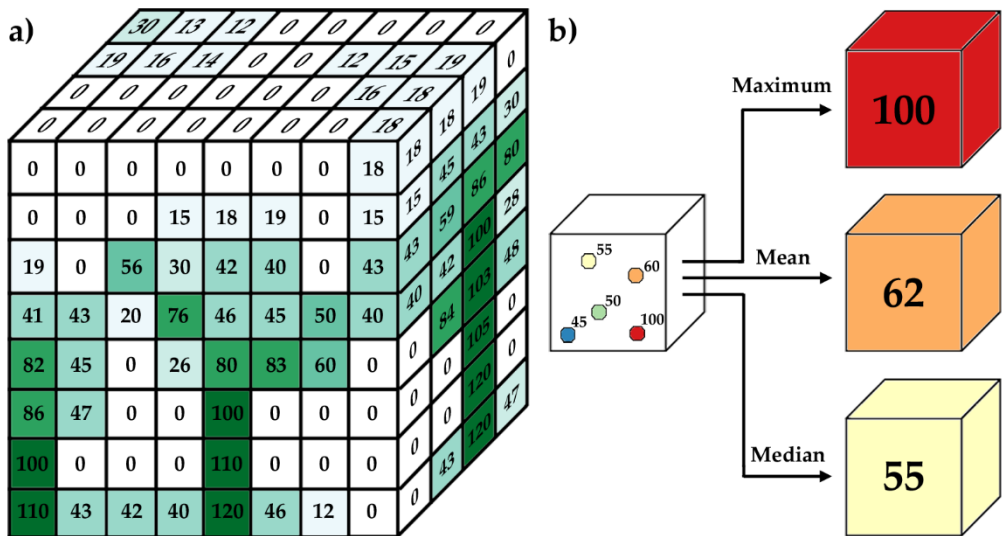
Figure 2.7. *Representation of (a) 3D matrix of voxels, and (b) voxelization of point cloud data within a voxel with different statistical procedures.*

dimensions were tested: 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95, 1.05, 1.15, 1.25, 1.35, 1.45 and 1.55 m. The minimum voxel size was equal to the footprint size. The voxel size in Z dimension was not modified, and the vertical distance between waveform bins, based on the temporal sample spacing of the ALS system, was respected. Therefore, the voxel size in Z dimension was 0.3 m, calculated from the temporal sample spacing (see Equation 5). In addition, the voxel value was computed using five different statistics (maximum, mean, median, percentiles 90 and 95) for all the waveform bins within each voxel. In the analysis of side-lap effect on the estimate of forest attributes, three voxel sizes were tested: 0.25, 0.5 and 1 m. As in the first analysis, Z dimension of voxels was calculated from the temporal sample spacing (i.e., 0.3 m) (see Equation 5). However, only maximum as the assignation value was tested for this second analysis to keep the significance of the maximum amplitudes from the original waveforms when generating the pseudo-vertical waveforms (Hermosilla et al., 2014b).

2.3.4. Extraction of ALS_FW metrics

Once voxelization has been carried out, pseudo-vertical waveforms must be generated. Although it is a step prior to the extraction of ALS_FW metrics, they are considered in the same procedure (Procedure 4). The generation of pseudo-vertical waveforms was proposed by Hermosilla et al. (2014b) to standardize the acquisition of ALS_FW metrics from different flight stripes and off-nadir scan angles. Most of the emitted pulses have off-nadir angles, and hence they have a horizontal displacement. In order to remove this horizontal displacement, the standardization of the pseudo-vertical waveform consists of generating new waveforms from the voxel values along a voxel column from the top of the canopy to the ground (Figure 2.8). As all these new waveforms have a vertical direction, they all represent the vertical profile of the vegetation from a nadir angle. As a result, a regular grid of waveforms with a vertical direction is generated, where each voxel column contains a pseudo-vertical waveform.
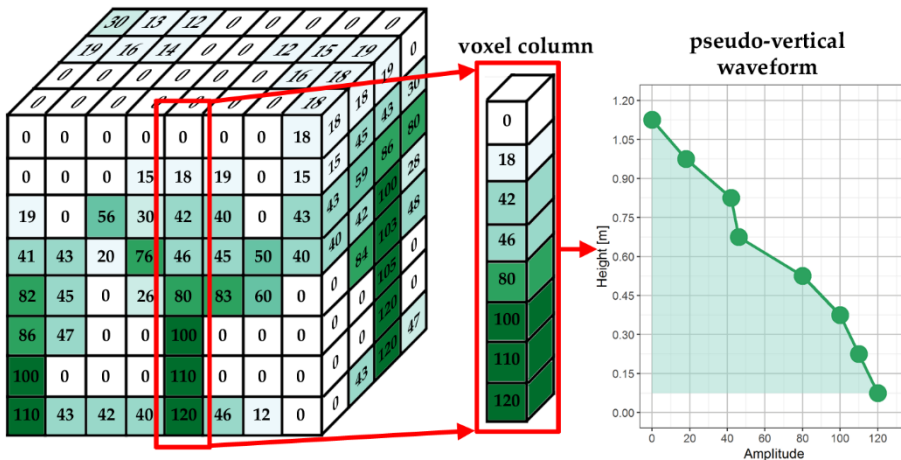
Figure 2.8. *Extraction of the voxel column and the corresponding pseudo-vertical waveform from the 3D matrix of voxels.*

Once waveforms from different scan angles have been standardized using pseudo-vertical waveforms, ALS_FW metrics may be extracted (Procedure 4). Some of the most employed metrics were proposed by Duong (2010) and further described by Cao et al. (2014). These metrics are HOME (Height Of Median Energy), WD (Waveform Distance), NP (Number of Peaks), ROUGH

(ROUGHness), HTMR (Height/Median Ratio), VDR (Vertical Distribution Ratio), RWE (Return Waveform Energy), and FS (Front Slope), which are described in Table 2.2 and Figure 2.9. All these metrics were extracted from each pseudo-vertical waveform, resulting in a regular grid where each voxel column contains the values of these metrics (Figure 2.10). To carry out an area-based approach (ABA) based on either plots, where field data have been acquired, or samples, to assess metric variations, metric values were obtained by calculating the mean and the standard deviation of each metric value within the area delimited by either the plot or the sample. As a result, 16 metrics were computed with an ABA from the initial eight metrics: HOMEμ, HOMEσ, WDμ, WDσ, NPμ, NPσ, ROUGHμ, ROUGHσ, HTMRμ, HTMRσ, VDRμ, VDRσ, RWEμ, RWEσ, FSμ, and FSσ. These resulting ALS$_{FW}$ metrics with an ABA were used for further assessment of metric variations and estimation of forest attributes.

Table 2.2. *Description of the ALS$_{FW}$ metrics.*

Metric	Description
HOME	Height of the waveform centroid
WD	Height of the waveform
NP	Number of peaks in the waveform
ROUGH	Offset between the waveform height and the height of the first peak
HTMR	HOME/WD
VDR	(WD-HOME)/WD
RWE	Integral of the waveform amplitudes
FS	Vertical angle from the beginning of the waveform to the amplitude of the first peak

2.3.5.   Analysis of the influence of side-lap effect on ALS$_{FW}$ metrics

In this section, the influence of side-lap effect of ALS$_{FW}$ metrics was analyzed. Firstly, it was analyzed how metric values vary as a function of pulse density; and secondly, as a function of voxel parameters (i.e., voxel size and assignation value).

2.3.5.1. Variation of ALS$_{FW}$ metrics related to pulse density

Once ALS$_{FW}$ metrics were computed for every sample, voxel size, assignation value and pulse density, its variation related to the pulse density was analyzed (Procedure 5). The goal was to analyze variations of ALS$_{FW}$ metrics modifying
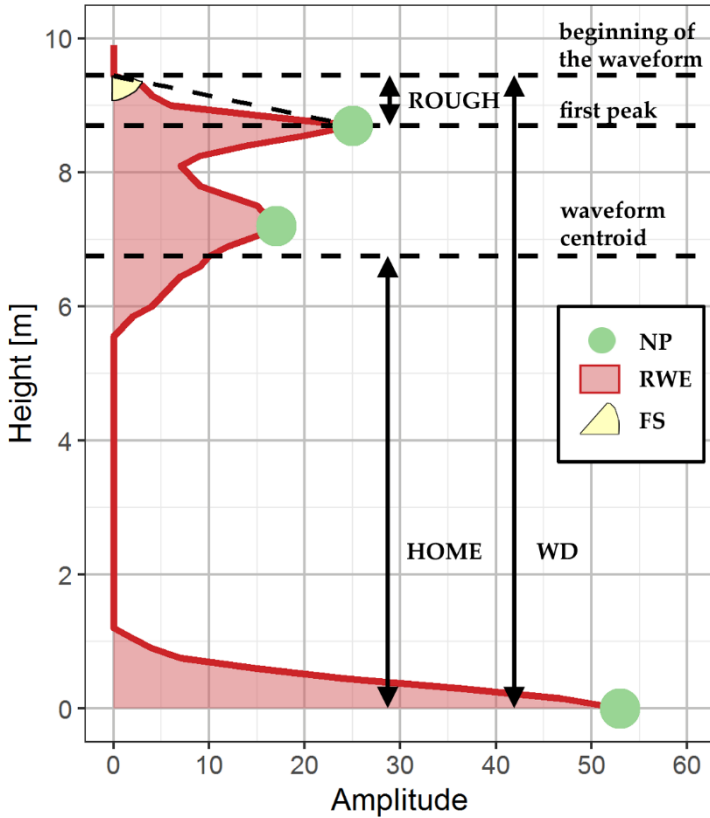
Figure 2.9. *Graphical description of six of the eight ALSFW metrics (HOME, WD, NP, ROUGH, RWE and FS) extracted from a pseudo-vertical waveform. The other two (HTMR and VDR) are computed as a combination of others.*

the three mentioned parameters (i.e., voxel size, assignation value and pulse density). Among the eight ALSFW metrics previously described, only six of them (HOME, WD, NP, ROUGH, RWE and FS) were analyzed in this analysis, since the other two (HTMR and VDR) depend on HOME and WD. Firstly, the variation related to pulse density for several samples at different voxel sizes and assignation values was observed. As this variation followed a negative exponential distribution, the least squares method was used to find the most appropriate parameter values, fitting a negative exponential model (Equation 6). In this model, based on the exponential semivariogram model (David, 1977), ALSFW metric values (y = dependent variable) tend to remain stable around a sill
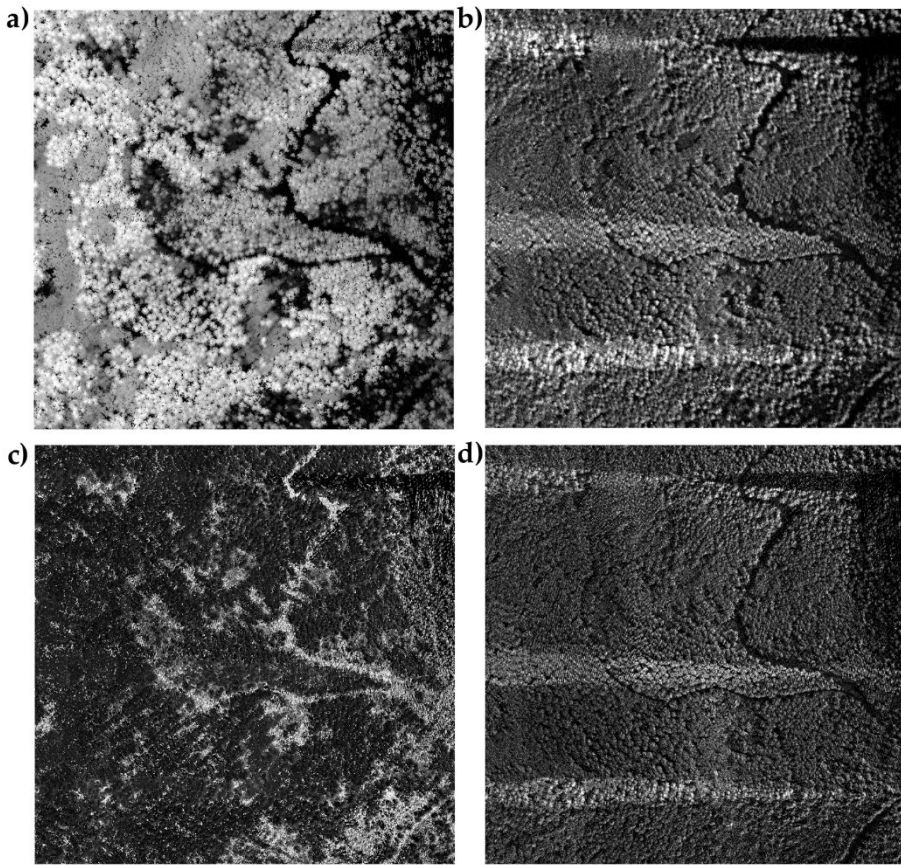
Figure 2.10. *Resulting values from the extraction of (a) HOME, (b) NP, (c) VDR and (d) RWE ALSFW metrics in a specific area of 500 × 500 m.*

with a slight positive slope at a given pulse density (x = independent variable). The formula of the negative exponential function is as follows:

$$y = a + c \times (1 - exp^{-\frac{3x}{b}})$$

Equation 6

where x = value of density in pulses·m^{-2},

y = value of the ALSFW metric,

a = value of y at which $x = 0$ in the negative exponential model,

b = value of x where y reaches the 95% of the sill value,

c = range of y between a and the value of y at which the function is stabilized then,

$a + c = y$ value of the sill.

On the other hand, each sample has different values for ALS$_{FW}$ metrics, due to vegetation variability. Therefore, with the aim of working with all 30 samples, a function was fitted for each sample separately. Instead, a function for each sample individually was fitted, and then the model results from the 30 samples clustered by ALS$_{FW}$ metric, voxel size and assignation value was averaged. As a result, 12,600 different models were computed (i.e., 30 samples × 6 ALS$_{FW}$ metrics × 14 voxel sizes × 5 assignation values) resulting 420 averaged results (i.e., 6 ALS$_{FW}$ metrics × 14 voxel sizes × 5 assignation values). Only negative exponential models with a convergence tolerance of $< 1 \times 10^{-5}$ in the iterative fitting process were used for the study. Validation was carried out using the Jackknife procedure described by Duda et al. (2012), which utilizes a leave-one-out procedure. Results were evaluated using the coefficient b, which shows the minimum pulse density where ALS$_{FW}$ metrics hardly vary, and the Jackknife bias, which shows the average of the deviations after removing one observation at each iteration.

2.3.5.2. Variation of ALS$_{FW}$ metrics according to voxel size and assignation value

As seen in the previous section, analyzing variability of ALS$_{FW}$ metrics as pulse density increases provides the minimum pulse density (MPD) where metrics stay steady, corresponding to the coefficient b of the negative exponential model. In addition, analyzing the variability using different voxel sizes and assignation values may help to diminish the influence of the pulse density (Procedure 5). Total Variation (TVar) (Equation 7) (Harten, 1983) can be used instead of the variability of ALS$_{FW}$ metric values for the different pulse densities explained in the previous section. The TVar computes the sum of differences between adjacent values. Hence, the lower the TVar value, the less variability the ALS$_{FW}$ metric has due to the pulse density. The formula of the TVar is as follows:

$$TVar = \sum_{pd=2}^{16-1} \left| y_{pd+1} - y_{pd} \right|$$

Equation 7

where y = value of the metric in a given pulse density (pd) and,

pd = pulse density.

Given that ALS$_{FW}$ metrics and assignation values have, in practice, a different range of values, ALS$_{FW}$ metrics were rescaled independently for each possible combination of metric and assignation type. A modified version of the feature scaling method was used (Equation 8) to standardize data. In our case, the minimum value was equal to zero to keep the minimum TVar value as zero:

$$y = \frac{x - min\,(x)}{max(x) - min\,(x)} \ / \ min\,(x) \ = 0$$

Equation 8

where y = standardization of the ALS$_{FW}$ metric value,

x = ALS$_{FW}$ metric value,

$min(x)$ = minimum ALS$_{FW}$ metric value grouped by ALS$_{FW}$ metric and assignation value, in our case modified to $min(x) = 0$,

$max(x)$ = maximum ALS$_{FW}$ metric value grouped by ALS$_{FW}$ metric and assignation value.

Afterwards, the TVar was computed from the 30 samples by averaging every ALS$_{FW}$ metric, voxel size and assignation value.

2.3.6. Analysis of the influence of side-lap effect on the estimate of forest attributes

The second analysis involved the assessment of side-lap effect when estimating forest attributes (Procedure 6b). To do this, accuracies of forest attribute estimates were compared by reducing pulse density in plots and testing different voxel sizes and regression methods. Prior to this analysis, the

estimation of forest attributes using ALS$_{FW}$ metrics as independent variables in the regression models was carried out (Procedure 6a). In this analysis, two forest attributes were estimated: AGB and canopy base height (CBH), being mass- and height-related attributes, respectively. Contrary to the previous analysis, the eight ALS$_{FW}$ metrics previously described were used as independent variables. The mean of each metric within each plot was computed and further used as variable in the regression models. Before generating the regression models, a selection of ALS$_{FW}$ metrics was performed only for the highest pulse density (i.e., 9 pulses·m^{-2}) and each regression method. The four regression methods tested were linear, square-root-transformed (sqrt), exponential and power, whose models are presented in Equation 9, Equation 10, Equation 11 and Equation 12, respectively. These selected metrics were used for all the pulse density reductions and voxel sizes in order to better compare how it affects estimates without using different explanatory variables. The process followed for the metric selection was to compare the Akaike Information Criterion (AIC) (Akaike, 1973) of all the possible models with a maximum of three ALS$_{FW}$ metrics as independent variables.

$$y = a_0 + a_1 \times x_1 + a_2 \times x_2 + \cdots + a_n \times x_n$$

Equation 9

$$\sqrt{y} = a_0 + a_1 \times x_1 + a_2 \times x_2 + \cdots + a_n \times x_n$$

Equation 10

$$y = exp^{(a_0 + a_1 \times x_1 + a_2 \times x_2 + \cdots + a_n \times x_n)}$$

Equation 11

$$y = a_0 \times x_1{}^{a_1} \times x_2{}^{a_2} \times \ldots \times x_n{}^{a_n}$$

Equation 12

where y = dependent variable (i.e., forest attribute to estimate),

$x_1, x_2, \ldots, x_n$ = independent variables (i.e., ALS$_{FW}$ metrics)

$a_0, a_1, a_2, \ldots, a_n$ = coefficients to be adjusted.

Once ALS$_{FW}$ metrics were selected, the four regression models described (i.e., linear, square-root-transformed, exponential and power) were generated for each pulse density and voxel size. Finally, regression models were evaluated by comparing R^2, and RMSE, and using leave-one-out cross-validation.

2.4.   Results

This section presents the results of the analysis of side-lap effect on ALS$_{FW}$ metrics and on the estimate of forest attributes. Firstly, variations in ALS$_{FW}$ metrics were analyzed by modifying pulse density and voxel parameters (i.e., voxel size and assignation value). Secondly, differences in evaluation parameters (i.e., R^2 and RMSE) of forest attribute estimates were analyzed when modifying pulse density, voxel size and regression methods.

2.4.1.   Analysis of the influence of side-lap effect on ALS$_{FW}$ metrics

Figure 2.11 shows how the pseudo-vertical waveform and the ALS$_{FW}$ metrics from the same voxel column vary modifying the pulse density, voxel size and assignation value. The lower the pulse density, the more null values and the less details appear in the pseudo-vertical waveform. However, changes in the waveform due to pulse density reduction seem to be less noticeable when voxel size increases to 1.25 m, except for the median assignation value. In addition, pseudo-vertical waveforms using the median assignation are smoother than those using the maximum assignation. Analyzing values of ALS$_{FW}$ metrics for the same voxel size, HOME, WD, ROUGH and FS do not show significant variations. On the contrary, NP and RWE are more variable.

2.4.1.1. Variation of ALS$_{FW}$ metrics related to pulse density

Figure 2.12 and Figure 2.13 show the variation of HOME, WD, NP, ROUGH, RWE and FS in one sample for the different pulse densities with the maximum assignation and voxel size of 0.25 and 0.75 m. In the case of 0.25 m (Figure 2.12a), the trend fits a negative exponential model for all the metrics. This does not occur using a voxel size of 0.75 m for metrics HOME and FS (Figure 2.12a and Figure 2.13c). The negative exponential function shows that metric values
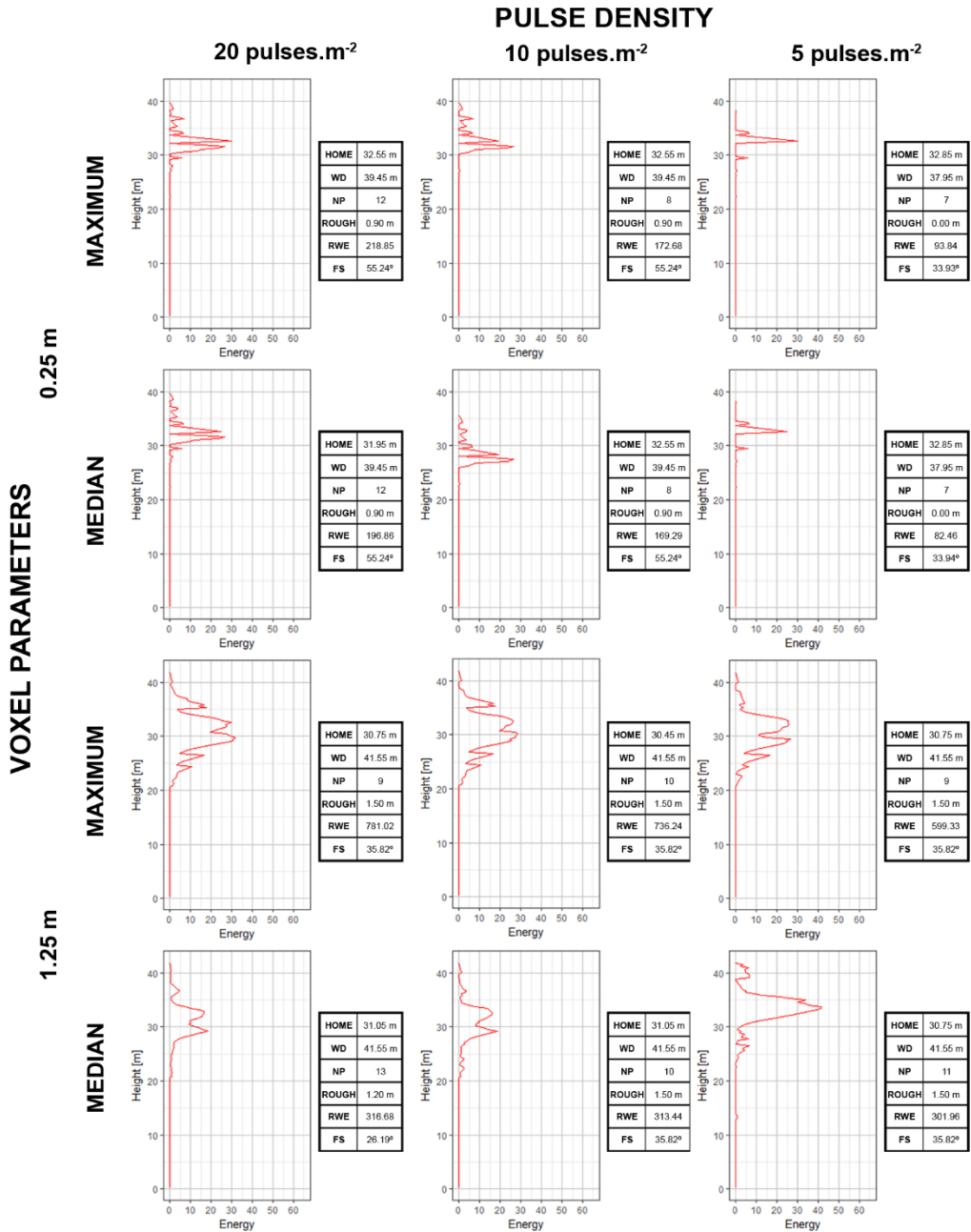
Figure 2.11. *Examples of pseudo-vertical waveforms at voxel column-level and ALS$_{FW}$ metric values for different pulse densities (20, 10 and 5 pulses·m^{-2}), voxel sizes (0.25 and 1.25 m) and assignation values (maximum and median).*
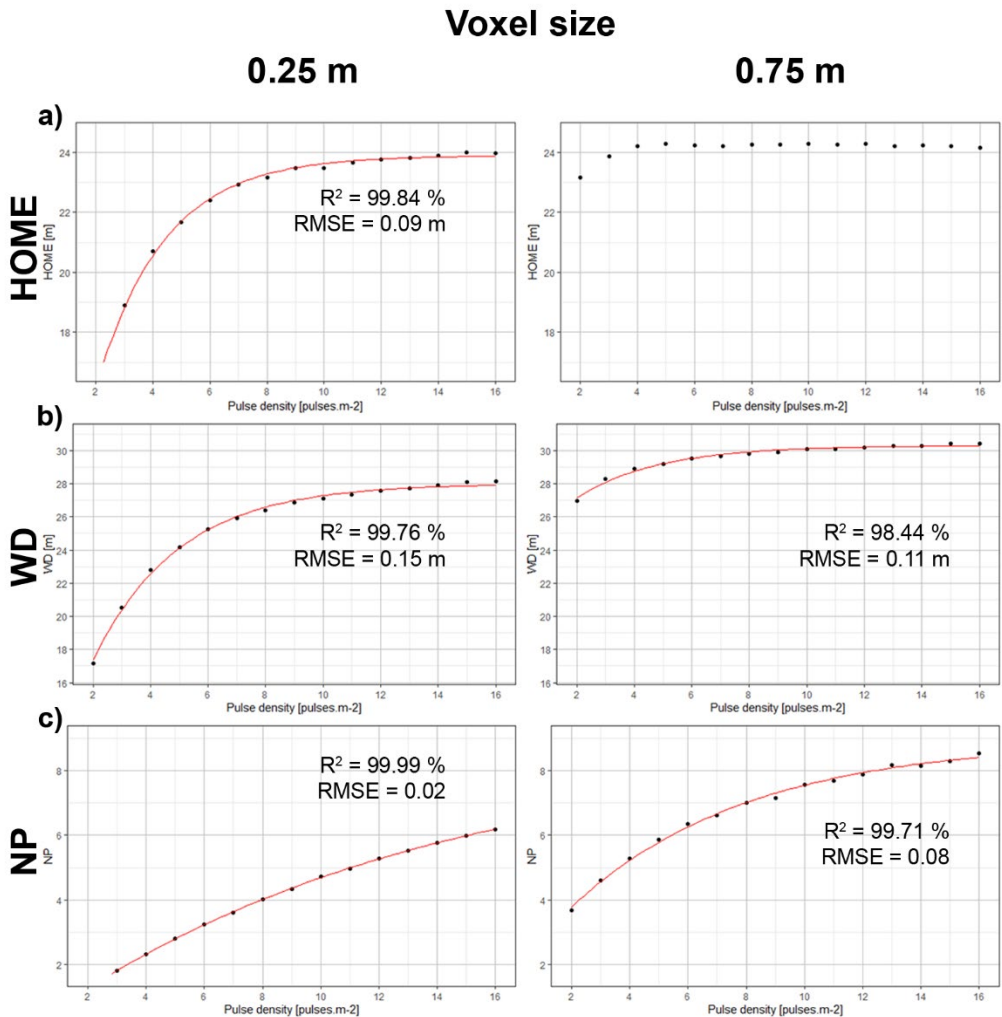
Figure 2.12. *Variation of (a) HOME in meters, (b) WD in meters and (c) NP, related
to pulse density in one sample for the maximum assignation value and a voxel size of
0.25 m (left column) and 0.75 m (right column). The black points represent the values
computed and the red curve the fitted negative exponential model. The values of HOME
in the right column do not fit a negative exponential model.*

progressively increase as pulse density increases, until they reach the sill of the
curve, being for example for HOME at 9-10 pulses·m⁻² (in this case the MPD was
7.11 pulses·m⁻²). However, the difference of metric values obtained at the lowest

and largest pulse density decreases with a voxel size of 0.75 m, therefore metric values become more stable increasing the voxel size.
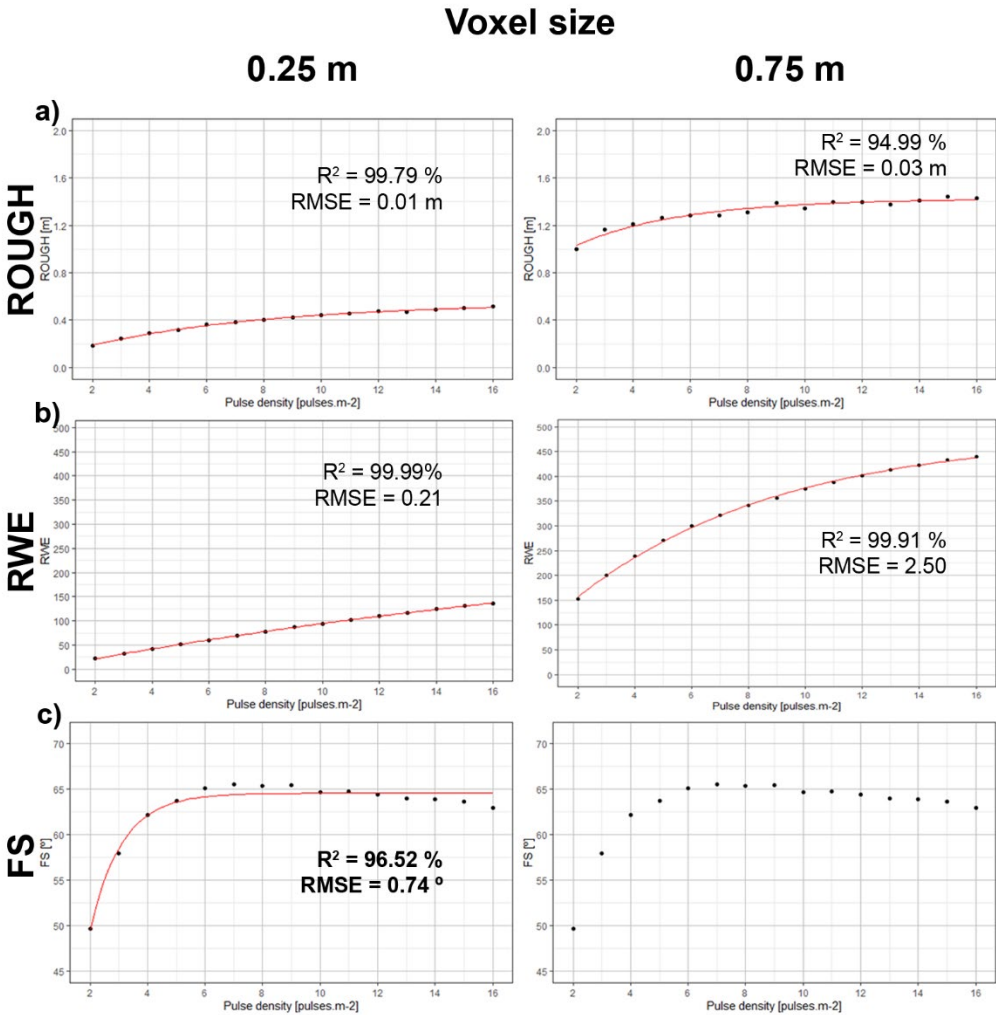
Figure 2.13. *Variation of (a) ROUGH in meters, (b) RWE and c) FS in degrees, related to pulse density in one sample for the maximum assignation value and a voxel size of 0.25 m (left column) and 0.75 m (right column). The black points represent the values computed and the red curve the fitted negative exponential model. The values of FS in the right column do not fit a negative exponential model.*

After generating the fitted models for every sample, Figure 2.14 shows the average of the adjusted MPD values from the 30 samples where the corresponding ALS$_{FW}$ metric remains stable (i.e., the *b* coefficients from the negative exponential models (see Equation 6)); and Figure 2.15 shows the standard deviation of the MPD for all samples. All the models obtained a Jackknife bias lower than 1.56×10^{-13} in the validation procedure for the three coefficients of the negative exponential model (i.e., *a*, *b* and *c*). This means that there were not outliers after applying the leave-one-out procedure. It is important to remark that the negative exponential models were generated using sample data from 2 to 16 pulses·m^{-2}. Hence, variation of ALS$_{FW}$ metric values estimated out of this range are extrapolations, and as such the resulting MPD values higher than 16 pulses·m^{-2} must be considered carefully. Additionally, empty cells in Figure 2.14 and Figure 2.15 correspond to combinations of metrics and voxel sizes that do not fit a negative exponential model. NP, ROUGH, and RWE are the metrics with highest MPD values (MPD ϵ [42.2, 46.2], MPD ϵ [18.7, 21.3], and MPD ϵ [60.2, 89.7] pulses·m^{-2}, respectively, for a voxel size of 0.25 m), while HOME, WD and FS have the lowest (MPD ϵ [7.1, 7.2], MPD = 9.6 and MPD ϵ [3.9, 4.1] pulses·m^{-2}, respectively, for a voxel size of 0.25 m). Every ALS$_{FW}$ metric remains asymptotically stable at lower pulse densities as voxel size increases. For instance, the MPD decreases from 7.1 to 3.4 pulses·m^{-2} for HOME; from 9.6 to 8.4 pulses·m^{-2} for WD; from 45.5 to 15.4 pulses·m^{-2} for NP; from 21 to 4.6 pulses·m^{-2} for ROUGH; and from 60.2 to 5.3 pulses·m^{-2} for RWE. However, WD has low values for voxel sizes of 0.35 and 0.45 m (MPD ϵ [8.4, 8.5]), but they increase again as the voxel size also increases (MPD = 13.5 pulses·m^{-2}). Results also show that for low MPD values (i.e., MPD ϵ [3.9, 5.6]), variation of ALS$_{FW}$ metrics does not fit a negative exponential trend for high voxel sizes. This behavior is observed with HOME, ROUGH, RWE and FS, except for ROUGH using the maximum assignation value. In these cases, ALS$_{FW}$ metric values tend to slightly decrease as pulse density increases.

Comparing different assignation values, HOME, WD, NP and FS have similar MPD values; however, ROUGH and RWE were influenced differently. Both ROUGH and RWE remain stable at lower pulse densities using the median as assignation value, but they present more variation using the maximum,

Voxel size (m)

Group	Metric	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1.05	1.15	1.25	1.35	1.45	1.55
HOME	Max.	7.2	5.8	5.1	4.6	4.3									
	Mean	7.1	5.7	4.8	4.2	3.7									
	Median	7.1	5.6	4.7	4	3.4									
	p90	7.2	5.8	5.1	4.5	4.2									
	p95	7.2	5.8	5.1	4.5	4.3									
WD	Max.	9.6	8.5	8.4	8.7	9.4	9.9	10.8	11.6	12.3	12.6	13	13.3	13.3	13.5
	Mean	9.6	8.5	8.4	8.7	9.4	9.9	10.8	11.6	12.3	12.6	13	13.3	13.3	13.5
	Median	9.6	8.5	8.4	8.7	9.4	9.9	10.8	11.6	12.3	12.6	13	13.3	13.3	13.5
	p90	9.6	8.5	8.4	8.7	9.4	9.9	10.8	11.6	12.3	12.6	13	13.3	13.3	13.5
	p95	9.6	8.5	8.4	8.7	9.4	9.9	10.8	11.6	12.3	12.6	13	13.3	13.3	13.5
NP	Max.	42.2	26.2	20.7	18.5	17.8	17.6	17.9	17.8	17.7	17.4	17	17	17.2	16.7
	Mean	45.5	29.5	23.6	20.7	18.9	17.7	17.1	16.5	16.2	15.9	15.7	15.6	15.7	15.4
	Median	46.2	30.9	25.5	22.7	20.9	19.6	18.8	18.1	17.8	17.1	16.9	16.6	16.6	16.5
	p90	42.9	27.2	22	20.2	20.1	20.4	21	20.7	20	19.5	18.8	18.5	18.3	17.8
	p95	42.6	26.8	21.4	19.3	18.8	18.9	19.6	20.1	20.7	21.1	20.7	20.2	19.9	19.2
ROUGH	Max.	21.3	17.5	14.6	12.2	11.4	10.9	10.5							
	Mean	19.4	13.9	9.2	5.5										
	Median	18.7	12.1	6.7											
	p90	21	16.6	12.7	9.3	7	4.6								
	p95	21.1	17	13.4	10.4	8.8	7.1	5.6							
RWE	Max.	89.7	48.9	33.6	26.7	23.2	21.3	20.1	19.5	19.1	18.8	18.6	18.5	18.4	18.4
	Mean	60.7	31.1	19.9	14.7	12	10.4	9.6	9.4	9.4	9.3	9.7	9.9	10.5	11
	Median	60.2	30.4	19.1	13.6	10.6	8.8	7.5	6.8	6.3	5.5	5.3			
	p90	82.4	44.3	30	23.3	19.7	17.5	16	15.1	14.5	14	13.7	13.6	13.5	13.5
	p95	85.9	46.6	31.8	25	21.5	19.3	17.9	17.1	16.3	15.6	15.2	14.9	14.7	14.6
FS	Max.	4.1													
	Mean	3.9													
	Median	3.9													
	p90	4.1													
	p95	4.1													

Figure 2.14. *Average minimum pulse density (MPD; i.e., coefficient b from the negative exponential model) from the 30 samples for different voxel sizes and assignation values. Empty cells correspond to combinations of metrics and voxel sizes that do not fit a negative exponential model. Values in bold correspond to MPD values higher than 16 pulses·m^{-2} (i.e., the maximum pulse density from sample data used to generate the negative exponential model). Smallest and highest values are represented by blue and red colors, respectively.*

percentiles 90 and 95. For instance, RWE has a MPD value of 5.3 pulses·m^{-2} using the median assignation and a voxel size of 1.25 m, while the MPD value was 18.6 using the maximum and the same voxel size.
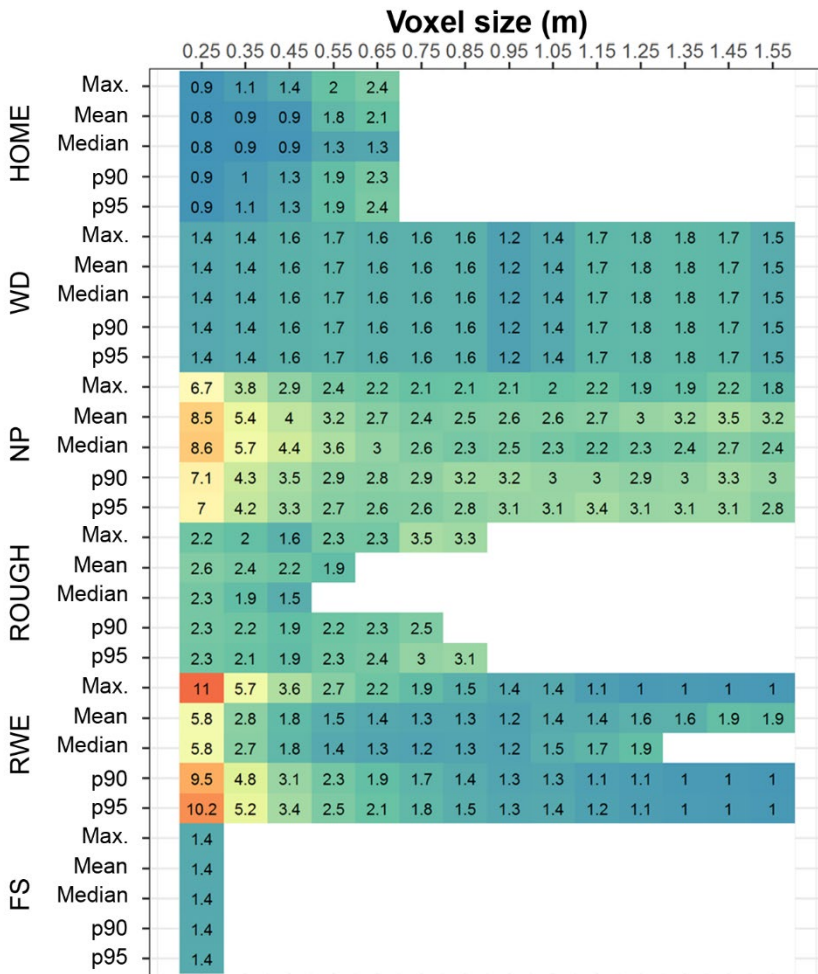
Figure 2.15. *Average standard deviation of minimum pulse density (MPD; i.e., coefficient b from the negative exponential model) from the 30 samples for different voxel size and assignation values. Empty cells correspond to combinations of metrics and voxel sizes that do not fit a negative exponential model. Smallest and highest values are represented by blue and red colors, respectively.*

Analyzing the average of the standard deviation of the MPD from the 30 samples (Figure 2.15), all the values are low (between 1 and 2.6 pulses·m⁻²) except for NP and RWE with small voxel sizes. These ALS$_{FW}$ metrics have large standard deviations for small voxel sizes ([6.7, 8.6] and [5.8, 11] pulses·m⁻², respectively), diminishing the values for larger voxel sizes ([1.9, 2.3] and [1.0,

1.2] pulses·m⁻², respectively). However, the standard deviation of ROUGH using the maximum assignation increases as voxel size increases. High standard deviation value of MPD are related to high MPD values.

2.4.1.2. Variation of ALS$_{FW}$ metrics according to voxel size and assignation value

Figure 2.16 shows the TVar value defined by Equation 7 and Equation 8, for every ALS$_{FW}$ metric computed at the different voxel sizes and assignation values. Overall, HOME, WD and FS present the lowest TVar values (TVar ϵ [0.03, 0.27], TVar ϵ [0.06. 0.28] and TVar ϵ [0.10, 0.28, respectively), while NP, ROUGH and RWE present higher values (TVar ϵ [0.24, 0.36], TVar ϵ [0.14, 0.52] and TVar ϵ [0.12, 0.45], respectively) using small voxel sizes. TVar values of HOME, WD, NP for maximum, RWE for mean and median, and FS, decrease as voxel size increases compared to the lowest voxel size (i.e., 0.25 m). These values range from 0.27 to 0.04 for HOME, from 0.28 to 0.06 for WD, from 0.33 to 0.30 for NP with the maximum assignation value; from [0.40, 0.45] to [0.12, 0.17] for RWE with the mean and median assignation values; and from [0.26, 0.28] to [0.10, 0.14] for FS. NP TVar values do not vary significantly as voxel size increases, the values being [0.24, 0.32] at 0.25 m, and [0.27, 0.30] the lowest TVar values at other voxel sizes. Regarding RWE, the TVar values are minimal at the lowest voxel size using the maximum, percentiles 90 and 95 as assignation values. Nevertheless, TVar values are particularly high at the lowest voxel size using the mean and median assignation value, and become low for the largest voxel sizes, especially with the median. In addition, TVar values from ROUGH steeply increase as voxel size increases, varying from [0.14, 0.23] at 0.25 m to [0.38, 0.52] at 1.55 m.

Regarding the assignation values, HOME and WD present little or no differences. However, NP, ROUGH and RWE have different TVar values depending on the assignation values. NP has the lowest value at 0.25 m for the median assignation value (TVar = 0.24 m). The lowest TVar values of ROUGH are reached using the maximum, percentiles 90 and 95. Finally, RWE TVar values have the largest differences between assignation values, the mean and median being completely different from the others.
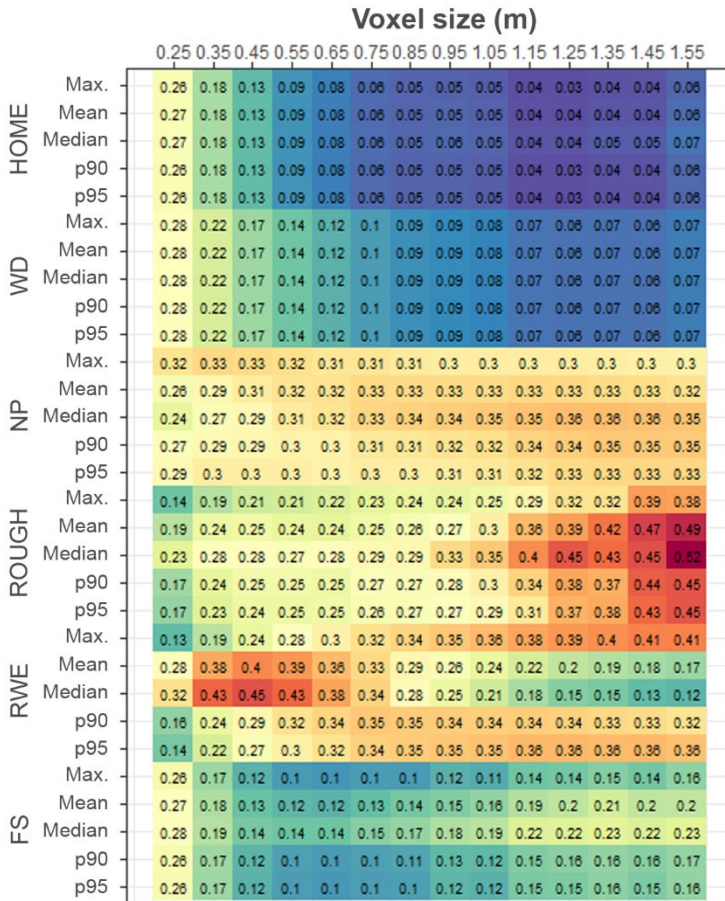
Figure 2.16. *TVar values for the different ALS$_{FW}$ metrics computed for the assignation values and voxel sizes. Smallest and highest values are represented by blue and red colors, respectively.*

2.4.2. Analysis of the influence of side-lap effect on the estimate of forest attributes

Figure 2.17 shows R² values obtained in AGB and CBH estimates for the different regression methods, pulse densities and voxel sizes. Regarding AGB, the voxel size of 1 m has the highest and less affected results by pulse density variations for all the regression methods. R² values range between 0.72-0.87,

0.81-0.88 and 0.83-0.89 for a voxel size of 0.25, 0.5 and 1 m, respectively. Additionally, for a voxel size of 1 m, R^2 variations along the different pulse densities with a same regression method are 0.01-0.04, while variations are 0.03-0.13 and 0.01-0.05 for a voxel size of 0.25 and 0.5 m, respectively. On the other hand, sqrt regression is also unaffected by pulse density at a lower voxel size (i.e., 0.5 m), varying up to 0.01. Overall, sqrt has the highest R^2 values, ranging between 0.87-0.88, and 0.88-0.89 for a voxel size of 0.5 and 1 m, respectively. On the other hand, R^2 values from CBH estimates stay steady until 1.5 pulses·m^{-2}, where they suddenly drop with a R^2 variation up to 0.11, except for 0.25 m with power and exponential regressions. For CBH estimates, the power regression is more constant for all the voxel sizes than the linear, sqrt and exponential regressions. R^2 variations for the power regression are between 0.02-0.04, while variations of linear, sqrt and exponential regressions range between 0.05-0.11, 0.10-0.15 and 0.05-0.05, respectively.

Table 2.3 shows the mean and standard deviation of the RMSE obtained from AGB and CBH estimates. Analyzing AGB results, the smaller the voxel size, the larger the RMSE and the more variability. For a voxel size of 0.25 m, mean and standard deviation values of RMSE range between 87.5-102.1 Mg·ha^{-1} and 2.9-11.1 Mg·ha^{-1}, respectively, while for a voxel size of 1 m, mean and standard deviation values of RMSE range between 80.6-91.1 Mg·ha^{-1} and 0.7-3.2 Mg·ha^{-1}, respectively. In addition, sqrt regression has the lowest RMSE, being 80.6 Mg·ha^{-1} with a standard deviation of 1.2 Mg·ha^{-1} for a voxel size of 1 m. Conversely, CBH has lower RMSE differences between voxel sizes and less variable. For all the voxel sizes, mean and standard deviation values are between 4.6-5.2 m and 0.1-0.3 m, respectively. Hence, CBH results do not depend on voxel size. In this case, power regression has the lowest RMSE, being 4.5 m with a standard deviation of 0.1 m for a voxel size of 1 m. In general, the behavior of RMSE is coincident with that of the R^2 values.

Figure 2.18 shows the estimation of AGB and CBH in a small area of 1.5 × 1.5 km with two different voxel sizes (0.25 and 0.75 m) and the maximum as assignation value. Estimation with a voxel size of 0.25 m shows that side-lap effect (highlighted with red rectangles) is present on both forest attributes. On

the contrary, the use of a larger voxel size (i.e., 0.75 m) shows that this side-lap effect is reduced or disappears.
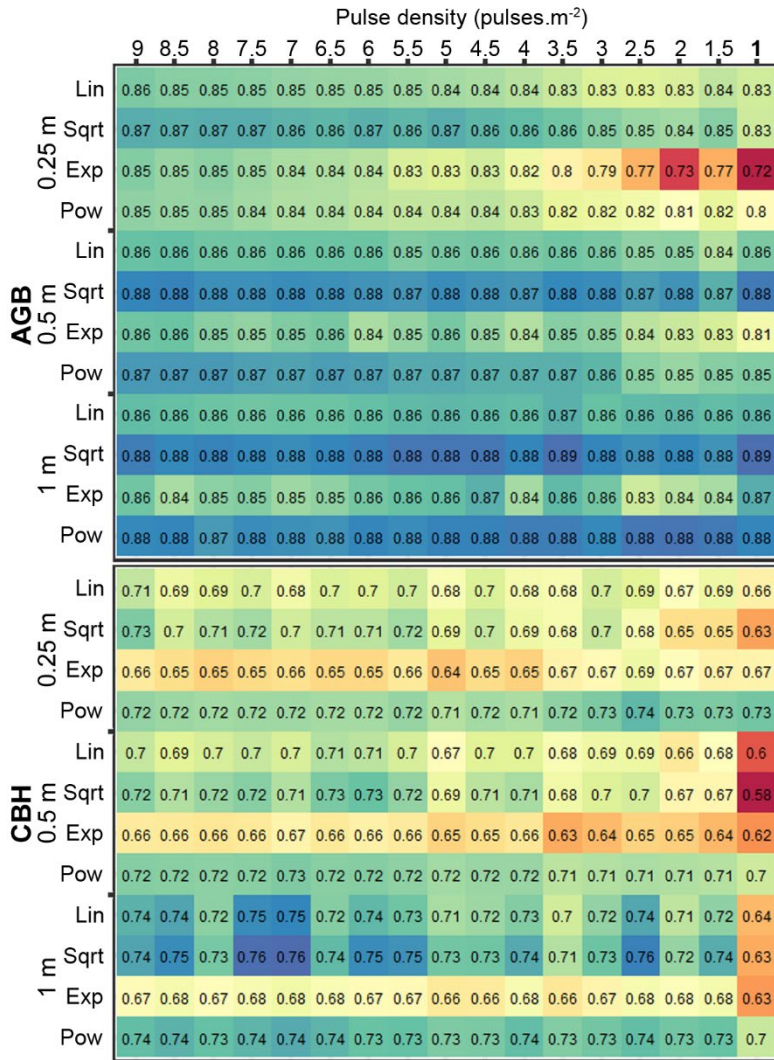
Figure 2.17. *Variation of R^2 values for AGB and CBH estimates for the different regression methods (Lin: linear, Sqrt: square-root-transformed, Exp: exponential, Pow: power), pulse densities and voxel sizes. Red- and blue-colored cells represent the lowest and highest R^2 values, respectively, for each attribute.*

Table 2.3. *Mean and standard deviation from the RMSE for the different forest attributes, voxel sizes and regression methods (linear, Sqrt: square-root-transformed, Exp: exponential, power).*

Methods	AGB (Mg·ha⁻¹)			CBH (m)		
	0.25 m	0.5 m	1 m	0.25 m	0.5 m	1 m
Lineal	92.6±2.9	88.4±1.6	86.6±0.7	4.8±0.1	4.8±0.2	4.5±0.2
Sqrt	87.5±3.2	82.1±1.2	80.6±1.2	5.0±0.3	4.9±0.3	4.6±0.2
Exp.	102.1±11.1	93.0±3.7	91.1±3.2	5.1±0.1	5.2±0.1	5.0±0.1
Power	97.4±4.0	87.2±2.8	82.4±1.1	4.6±0.1	4.7±0.1	4.5±0.1
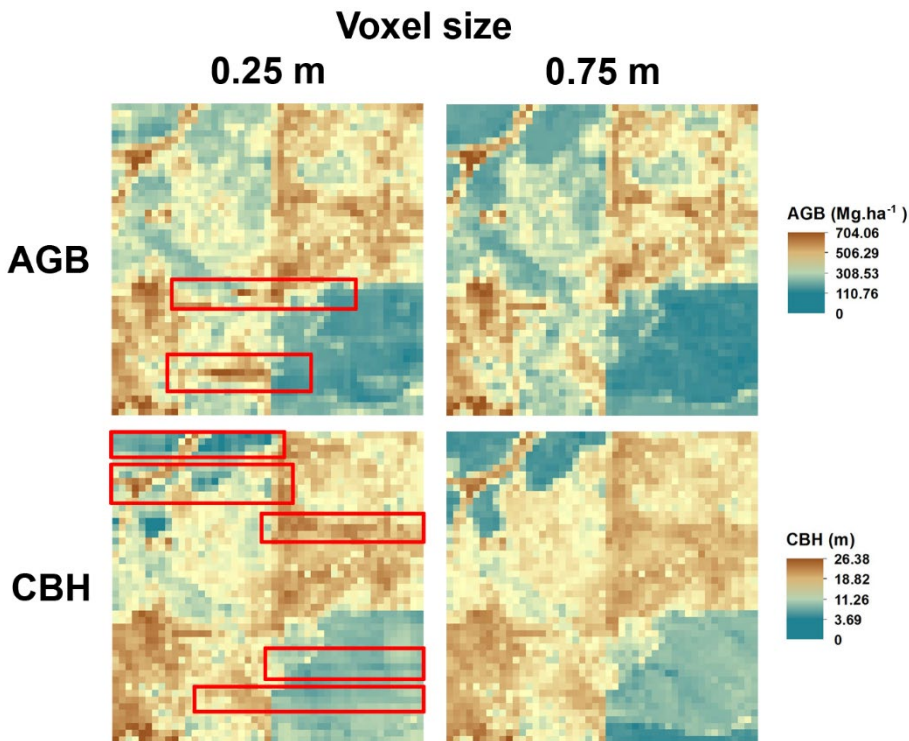
Figure 2.18. *Estimation of AGB and CBH in an area of 1.5 × 1.5 km using maximum assignation value and voxel sizes of 0.25 m and 0.75 m. Red rectangles highlight the side-lap effect on forest attribute values for a voxel size of 0.25 m.*

2.5. Discussion

Key results indicate that variations of ALS$_{FW}$ metrics due to pulse density differences can be foreseen, and therefore their impact reduced by setting a MPD, modifying the voxel size and/or the assignation value used. Additionally, forest attributes estimated from ALS$_{FW}$ metrics are apparently much less influenced by side-lap effect. However, this effect is present when estimating forest attributes in larger areas and must also be considered. Hence, modification of the voxel parameters may help to diminish the side-lap effect in a particular study area, and therefore to obtain a more accurate estimate of forest attributes.

Results showed that variations of ALS$_{FW}$ metrics related to pulse density have a negative exponential behavior, especially with small voxel sizes. Usually, there is a MPD from which metric values are stabilized. In new acquisitions of ALS$_{FW}$ data, this MPD should be the minimum pulse density value registered by the sensor to avoid the side-lap effect. However, the MPD is not constant for every ALS$_{FW}$ metric, voxel size or assignation values employed. Therefore, in practice, either the most affected ALS$_{FW}$ metrics should be avoided for estimation of forest attributes, the voxel size increased, the assignation value or the regression method modified.

On the other hand, when ALS$_{FW}$ data have already been acquired, pulse density cannot be increased, and therefore other strategies are required, such as modifying voxel parameters. Our results showed that increasing the voxel size and/or modifying the assignation value can make ALS$_{FW}$ metrics more stable. The probability that larger voxels are crossed by at least one waveform is higher, avoiding the gaps in the voxel columns that may alter values of ALS$_{FW}$ metrics. Therefore, side-lap effect is influenced by pulse density, which, together with the voxel size, is directly related to the number of waveforms crossing each voxel. Hence, a trade-off between increasing voxel size to reduce side-lap effect and a substantial loss of resolution should be considered. Regarding the assignation value, its effect on the stability of ALS$_{FW}$ metrics depends on the chosen metrics. Some standard ALS$_{FW}$ metrics, such as RWE, have unstable behavior, whereas some others, such as WD, have not. In general, the increment

of the voxel size and the change of the assignation value reduce the variation of ALS$_{FW}$ metrics.

MPD values determine the minimum pulse density required to obtain stable ALS$_{FW}$ metrics. However, the variation trend of some ALS$_{FW}$ metrics does not follow a negative exponential model. Additionally, in some metrics (e.g., WD) higher values of MPD do not correspond to higher values of TVar. Therefore, the introduction of TVar complements the MPD as an indicator of the variability of the ALS$_{FW}$ metric due to pulse density changes.

Regarding different behavior among ALS$_{FW}$ metrics, NP and RWE are more sensitive to pulse density changes than the rest. The lack of one or more voxel values means fewer peaks and a different sum of amplitudes in the wave. On the contrary, HOME, WD, ROUGH (at lower voxel sizes) and FS are less affected, since they are metrics that are related either to the height or to the top texture of the canopy, where the laser energy from airborne sensors arrives without occlusion. WD only requires a proper estimation of the height of the beginning of the waveform (top of the canopy), and it is well determined if the waveform intersects with the top of the trees. HOME calculation involves the beginning of the waveform as well as the height of the median energy. The latter is usually well registered, since it often corresponds to the densest vertical layer (see HOME values in Figure 2.11). ROUGH and FS calculation required the beginning of the waveform, and the position and amplitude of the first peak. Therefore, HOME, WD, ROUGH and FS vary if some voxel columns have no data due to a low pulse density. In order to avoid this, an increment of the voxel size is required.

In addition, there is remarkable disparity in values of ALS$_{FW}$ metrics using different assignation values. MPD and TVar values from WD do not vary, since the beginning of the waveform does not vary by modifying the assignation value. HOME has slight differences, since the height of the median energy may vary depending on the assignation employed. NP also presents minor variation, since the pseudo-vertical waveform has more singularities when the maximum assignation value is employed, ROUGH also has some differences due to possible variation of the first peak. RWE is the most variable ALS$_{FW}$ metric. As it is computed as the sum of amplitudes of a waveform, the sum of maximum

values may substantially differ from the sum of median values, for instance. A normalized metric may be used in order to avoid these differences. A possible approach could be to calculate a normalized RWE (nRWE) following Equation 8, where x is equal to RWE, and min(x) and max(x) are the minimum and maximum RWE values, respectively, for each assignation value. Thus, nRWE values from different assignation values would be comparable. Finally, FS may present small differences, since the amplitude and position of the first peak can vary as well.

To summarize, in order to reduce the side-lap effect in this scenario, the increment of the voxel size is recommended for HOME, WD, FS, and RWE for the mean and median assignation values, but not for ROUGH and RWE when maximum, percentiles 90 and 95 assignation values are used. Besides, depending on the voxel size, the selection of the assignation value has to be considered for RWE. According to results, NP might be discarded for estimating forest attributes because of its sensitivity to pulse density. Observing Figure 2.14 and Figure 2.16, MPD, voxel size and assignation values can be selected to minimize the side-lap effect in areas with similar vegetation types and densities. When planning a ALS project, a MPD around 10 pulses·m^{-2}, a voxel size of 0.75 m or similar, and the mean or median voxel assignation seem to optimize general performance. This combination of parameters provides the minimum values of MPD for most of the ALS$_{FW}$ metrics (Figure 2.14), except for NP. However, if ALS$_{FW}$ data are already available and the pulse density cannot be increased, the maximum assignation and a voxel size of about 0.75 m would be the most efficient option in terms of reduction of side-lap effect Figure 2.16.

Results are analogous to those of previous studies using ALS$_D$. In these studies, a similar tendency for R^2 (Jakubowski et al., 2013; Manuri et al., 2017), reliability ratio (Magnussen et al., 2010; Hansen et al., 2015) and maximum height metric (Roussel et al., 2017) was found. These values stabilize as pulse density increases.

On the other hand, estimate of forest attributes are also influenced by side-lap effect, given that influenced ALS$_{FW}$ metrics are used as independent variables in the regression models. Nevertheless, the influence of side-lap effect

is apparently lower on forest attributes, since the plots used for the analysis are located where field data are available, and therefore there might not be a variability in pulse density; and because forest attributes are statistically fitted and errors minimized by using regression models. However, side-lap effect is clearly present when forest attributes are estimated for large areas.

In the same way that the influence of side-lap effect depends on the ALS$_{FW}$ metrics, estimate of forest attributes are influenced differently. For instance, mass-related attributes (e.g., AGB) are more influenced by voxel size to reduce side-lap effect, while height-related attributes (e.g., CBH) are more influenced by the regression methods. Increasing the voxel size around 1 m makes that accuracy to estimate AGB does not vary as pulse density decreases. On the other hand, the use of the power and exponential regression methods makes the accuracy to estimate CBH more constant.

These results are coherent with those reported by González-Ferreiro et al. (2012) and Jakubowski et al. (2013) with ALS$_D$. They also observed a lower variation in precision for attributes related to height than for mass- and cover-related attributes as pulse density decreases.

Modeling variations of ALS$_{FW}$ metrics related to the pulse density is relevant to remove or reduce the side-lap effect when mapping or estimating metrics and forest attributes. Depending on the ALS data acquisition step, different strategies can be followed. First, if ALS$_{FW}$ data have not been acquired yet, a minimum pulse density that keeps ALS$_{FW}$ metrics stable may be set. Second, if ALS$_{FW}$ data have already been acquired, variation of ALS$_{FW}$ metrics can be reduced by increasing the voxel size to a certain extent, and/or using a specific assignation value. In this case, the pulse density cannot be increased, therefore voxel parameters that provide more stable metrics should be used. Finally, if some forest attributes do not respond to these strategies and reducing the side-lap effect is not possible, then either they should be avoided for further regression models or other regression methods should be employed, especially in height-related attributes.

2.6. Conclusions

In this chapter, the variation of ALS$_{FW}$ metrics and estimates of forest attributes according to the pulse density, voxel parameters and regression methods has been analyzed. This variation is common due to side-lap areas that are registered with a higher pulse density, and is known as "side-lap effect". Our results suggest that ALS$_{FW}$ metric variations related to pulse density can be modeled in most cases using a negative exponential model, and therefore there is a threshold at which their values stabilize. From this point, a minimum pulse density can be set to avoid the side-lap effect. In addition, modifying voxel parameters (i.e., voxel size and assignation value) reduces the side-lap effect when pulse density cannot be increased, e.g., when ALS$_{FW}$ data have already been acquired. Thus, an increment of the voxel size is recommended for HOME, WD, FS and RWE for the mean and median assignation values. Nevertheless, small voxel sizes make ROUGH and RWE for maximum, percentiles 90 and 95 more stable. On the other hand, the choice of the assignation value must be considered depending on the voxel used for RWE. However, NP is sensitive to pulse density variations and they cannot be reduced through voxel parameters, and therefore should be avoided for further analyses. Regarding estimates of forest attributes, mass-related attributes are more influenced by pulse density than height attributes. This influence however can also be reduced by increasing the voxel size. On the contrary, height-related attributes are influenced by the regression methods used for the estimate. The results presented in this chapter have practical relevance in order to avoid the side-lap effect when estimating forest attributes using ALS$_{FW}$ data. Further studies could focus on analyzing the effect of these parameters on different ecosystems with different dominant species, as well as the effect of the emitted pulse energy, footprint size, and vertical structure on ALS$_{FW}$ metrics, since they also influence the penetration of laser pulses.

Chapter 3

Creating a full-waveform airborne laser scanning metric extraction tool (*WoLFeX*), and analyzing the influence of scan angle and radiometric correction on metrics and forest fuel attributes estimates

Pinus halepensis

"Pino esbelto y tranquilo,
soledad de la tarde,
tan concreto en la libre
desolación del aire,
tan alto cuando todo
se confunde y abate
y huye el sol a tu copa
tibio y agonizante."

Dionisio Ridruejo

In this third chapter, a software tool (*WoLFeX*) that integrates methods for processing ALS$_{FW}$ data and extracting object-oriented metrics is presented. Additionally, the influence of the relative radiometric correction and the acquisition scan angle (*i*) on the ALS$_{FW}$ metric RWE, and (*ii*) on the estimation of three forest fuel attributes (canopy fuel load: CFL, canopy height: CH, and CBH) are tested and evaluated. The lack of ALS$_{FW}$ software and the fact that most of them do not include radiometric correction or computing specific metrics for understory vegetation, makes *WoLFeX* a significant tool to work with ALS$_{FW}$ data and to use it in the next sections of this thesis.

3.1. Introduction

As mentioned in Chapter 2, an essential pre-processing step before generating ALS$_{FW}$ metrics is the radiometric correction. Radiometric correction or calibration is a term widely used in remote sensing imagery. The goal of this correction is to reduce errors in the acquired digital values of the pixels due to atmospheric or sensor factors (Xu et al., 2019). This process is fundamental when dealing with images acquired from different sensors or in different days (Mafanya et al., 2018). Additionally, radiometric correction may involve converting digital numbers to physical units (Chander et al., 2009). This process is less extended for ALS data, however, it is an essential step for ALS$_{FW}$ due to the reliance between ALS$_{FW}$ metrics and amplitude values (Wagner, 2010). The use of ALS$_{FW}$ data without radiometric correction may lead to modified ALS$_{FW}$ metrics, and consequently modified estimates of forest fuel attributes. In this case, the goal of the radiometric correction is to provide amplitude values independent of the angle of incidence, range from the sensor to the target, as well as sensor and flight day atmospheric conditions. Attending to Briese et al. (2012), there are two main types of radiometric correction of ALS$_{FW}$: (*i*) correcting radiometric differences between flight lines without auxiliary data (i.e., relative correction) or (*ii*) using a surface whose approximate reflectance values are known or using measured reflectance values from ground targets (i.e., absolute correction). The use of any of these corrections, which depends on the available data and the presence of well-known surfaces, has been pointed

out by some authors as a relevant pre-processing step of ALS$_{FW}$ datasets (Wagner, 2010; Sevara et al., 2019).

There are several processing tools available to retrieve ALS$_D$ metrics, which are further used to predict forest fuel attributes. Two well-known software tools are *FUSION/LDV* (McGaughey, 2014) and *LAStools* (Isenburg, 2017). Additional libraries for processing ALS$_D$ data are available in different programming languages (e.g., *lidR* in R as used by Roussel and Auty (2017); *laspy* in Python [http://laspy.readthedocs.io/en/latest/]), such that users can customize their own tool for specific needs. However, due to the complexity of the use of ALS$_{FW}$ data and the non-standard level of its use and application, there is a very limited number of processing tools available for ALS$_{FW}$ data, both at open or commercial levels. Researchers working on ALS$_{FW}$ have developed their own tools for specific purposes, while others have made their tools available. Zhou and Popescu (2019) developed an R package named *waveformlidar* to process and visualize ALS$_{FW}$ data. This package allows for processing ALS$_{FW}$ data through two different strategies: (*i*) by means of deconvolution or decomposition of waveforms, therefore providing ALS$_D$ point clouds with more information (e.g., echo width); and (*ii*) generating dense point clouds from waveforms, self-named as "hyper point clouds". Furthermore, some commonly used ALS$_{FW}$ metrics, including those proposed by Duong (2010), may also be retrieved by voxelizing the hyper point clouds through the *waveformlidar* package. Miltiadou et al. (2019) created another open source software tool to process ALS$_{FW}$ and hyperspectral imagery data called *DASOS*. This tool visualizes a polygon representation from voxelized ALS$_{FW}$ data and computes ALS$_{FW}$ metrics at the voxel column level. Most of these metrics are related to height, distance between voxels, and number of empty/full voxels, except for the maximum and average intensity value of the voxel column. Another available tool for ALS$_{FW}$ data processing is *OPALS* (Orientation and Processing of Airborne Laser Scanning Data) (Pfeifer et al., 2014). This tool computes an ALS$_{FW}$ decomposition, but not specific ALS$_{FW}$ metrics from the whole waveform amplitudes. Apart from ALS$_D$, LAStools allows for the visualization of ALS$_{FW}$ data through *PulseWaves* (Isenburg, 2012) by representing the trajectories of the pulses. Among the software tools mentioned, only *waveformlidar* and *OPALS* include the radiometric correction, with the relative radiometric correction

being recently included in OPALS (Sevara et al., 2019). However, more tools are needed in order to cope a wider range of metrics able to better characterize forest structure (e.g., understory vegetation metrics), as well as to offer a more straightforward approach for the radiometric correction for raw ALS$_{FW}$ data.

In this chapter, in section 3.2, the different steps to process ALS$_{FW}$ data using the methodology based on voxelization and generation of pseudo-vertical waveforms to extract ALS$_{FW}$ metrics are described. In addition, a set of ALS$_{FW}$ metrics proposed in the literature and others in this study are enumerated. In section 3.3, it is shown the software tool *WoLFeX* and how ALS$_{FW}$ data processing and metric extraction described in previous sections may be carried out. In section 3.4, it is shown an application example of *WoLFeX* where the influence of scan angle and the relative radiometric correction of ALS$_{FW}$ data on one ALS$_{FW}$ metric and the estimate of three forest fuel attributes is analyzed. Finally, in section 3.5, the conclusions of this chapter are presented. This chapter sets out the Hypotheses 2 and 3: "Understory vegetation in Mediterranean forests is a key strata in characterizing wildlife habitats and mitigation of forest fires. Metrics to detect and characterize understory vegetation are crucial for modeling forest fuel attributes and forest structure, and ALS$_{FW}$ is a more suitable technique to capture information from understory. There is a lack of ALS$_{FW}$ processing tools integrating these tasks and able to perform radiometric correction of data" and "The scan angle has an influence on the amplitude values of the waveforms, and therefore on the ALS$_{FW}$ extracted. The application of radiometric correction is expected to reduce these differences in the amplitude values. As well as ALS$_{FW}$ metric values, forest fuel attributes estimated using these metrics may also be influenced by scan angle differences, and radiometric correction is expected to reduce this effect", respectively. To demonstrate these hypotheses, Objectives 2 and 3 are addressed in this chapter: (*i*) compile a set of methods to process and analyze ALS$_{FW}$ data, including the relative radiometric correction of the data to reduce the effect of the different angles of incidence and local altitude variations during the data acquisition process, the extraction of most ALS$_{FW}$ metrics as proposed in the literature, as well as new metrics focused on understory vegetation, integrating them in a

new software available to use for the scientific community, and (*ii*) assess the influence of the scan angle of ALS data acquisition and the application or not of a radiometric correction on the extraction of an ALS$_{FW}$ metric, and model three of the most relevant forest fuel variables—canopy fuel load (CFL), canopy height (CH), and canopy base height (CBH).

3.2. ALS$_{FW}$ data processing

One the most common methodologies proposed to process ALS$_{FW}$ data to study and model the three-dimensional structure of forests is the one based on voxelization and generation of pseudo-vertical waveforms (Hermosilla et al., 2014b), which has been described in more detail in sections 2.3.3 and 2.3.4.

The overall strategy followed to process ALS$_{FW}$ data and extract derived metrics is illustrated by the flow diagram of Figure 3.1. This strategy coincides with some procedures of the strategy followed in Chapter 2 (Figure 2.3). As the first procedure (1), radiometric correction and denoising were required prior to any treatment of ALS$_{FW}$ data. The second procedure (2) consists of voxelizing ALS$_{FW}$ data to generate pseudo-vertical waveforms and the third procedure (3) in extracting ALS$_{FW}$ metrics. The last procedure (4) computes ALS$_{FW}$ metrics from a pseudo-vertical waveform scale to an object scale (i.e., samples or plots).

As mentioned in the previous chapter, the first ALS$_{FW}$ processing step consists of the radiometric correction of the amplitude values contained in each waveform bin (Procedure 1). This is especially critical for ALS$_{FW}$, since the metrics extracted to create the models are directly related to the amplitude values. The approach followed to do the relative radiometric calibration is described in section 2.3.1.

After the radiometric correction, a denoising process is needed to remove the noise due to the system registration process (Procedure 1). The process followed is the one described in section 2.3.1.

Next step is the voxelization of ALS$_{FW}$ data (Procedure 2), which is described in section 2.3.3.
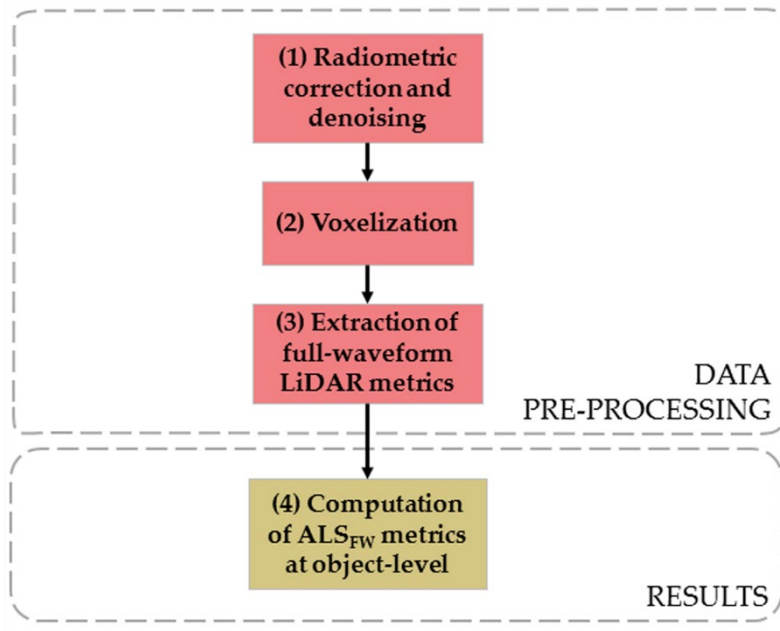
Figure 3.1. *Overall processing flowchart implemented to process ALS$_{FW}$ data and extract derived metrics.*

Once voxelization is carried out, pseudo-vertical waveforms are obtained based on the amplitude values of a same column of voxels (Procedure 3), correcting off-nadir waveforms. The description of how to generate the pseudo-vertical waveforms following Hermosilla et al. (2014b) and how to extract the derived metrics is described in more detail in section 2.3.4.

After generation of pseudo-vertical waveforms, ALS$_{FW}$ metrics may be extracted. Among the ALS$_{FW}$ metrics enumerated, we proposed 13 new metrics: kurtosis, height of the first empty voxel (HFEV), height of the first empty voxel from a threshold (HFEVT), energy to the first empty voxel (EFEV), normalized energy to the first empty voxel (nEFEV), filled voxels at the understory (FVU), normalized number of filled voxels at the understory (NFVU), bottom of canopy (BC), bottom of canopy distance (BCD), bottom of canopy energy (BCE), canopy distance (CD), canopy energy (CE), and the canopy energy ratio (CER) and perform an exhaustive compilation of those proposed by previous recent

studies (Kimes et al., 2006; Duncanson et al., 2010; Duong, 2010; Zhang et al., 2011). All of them are available in the software tool. The ALSFW metrics implemented can be divided into seven categories: height, energy, peaks, understory, percentiles, Gaussian decomposition, and others. Table 3.1 describes the different ALSFW metrics classified by category.

Table 3.1. *Description of ALSFW metrics available in WoLFeX (adapted from Crespo-Peremarch and Ruiz (2020)).*

Category	Name	Description	Units	Reference
Height	WD	Waveform distance	m	Duong (2010)
	ROUGH	Roughness of outermost canopy	m	
	HEIGHT Qn	Proportion of energy at the nth elevation quarter	-	Duncanson et al. (2010)
Energy	RWE	Return waveform energy	DN	Duong (2010)
	MAX E	Maximum energy	DN	Duncanson et al. (2010)
	VARIANCE	Variance of energy	DN^2	
	SKEWNESS	Skewness of energy	-	
	ENERGY Qn	Proportion of energy at the nth energy quarter	-	
	KURTOSIS	Kurtosis of energy	-	This study
Peaks	NP	Number of peaks	-	Duong (2010)
	START PEAK	Distance between the beginning of the waveform and the height of MAX E	m	Duncanson et al. (2010)
	PEAK END	Distance between the height of MAX E and the ground	m	
Understory	HFEV	Height of the first empty voxel	m	This study
	HFEVT	Height of the first empty voxel from a given threshold	m	
	EFEV	Energy from the ground to the first empty voxel	DN	
	nEFEV	Energy from the ground to the first empty voxel divided by RWE	-	
	FVU	Number of filled voxels at the understory	-	
	NFVU	Number of filled voxels at the understory divided by the total number of voxels	-	
Percentiles	Hn (H50 = HOME)	Height at the nth percentile of energy	m	Kimes et al. (2006)

* DN: Digital Number

Table 3.1. (cont.) *Description of ALSFW metrics available in WoLFeX (adapted from Crespo-Peremarch and Ruiz (2020)).*

Category	Name	Description	Units	Reference
Gaussian Decomposition	N GS	Number of Gaussian curves in the waveform	-	Duncanson et al. (2010)
	N GS STARTPEAK	Number of Gaussian curves between the beginning of the waveform and the height of the boundary	-	
	N GS ENDPEAK	Number of Gaussian curves between the height of the boundary and the ground	-	
	GE	Ground energy extracted from the ground Gaussian curve	DN	Zhang et al. (2011)
	GRR	Ground return ratio: GE divided by RWE	-	
	CHn	Elevation of the nth quarter of energy, excluding the ground Gaussian curve	m	
	Rn	CHn divided by WD	-	
	AGS	Average Gaussian curve slope	-	
	SGS	Standard deviation Gaussian curve slope	-	
	MSGS	Modified standard deviation Gaussian curve slope	-	
	BC	Bottom of canopy: elevation of the first canopy Gaussian curve	m	This study
	BCD	Bottom of canopy distance: distance from the beginning of the waveform to BC	m	
	BCE	Bottom of canopy energy: energy from the beginning of the waveform to BC	DN	
	CD	Canopy distance: distance from the beginning of the waveform to the boundary between ground and canopy	m	
	CE	Canopy energy: energy excluding GE	DN	
	CER	Canopy energy ratio: CE divided by RWE	-	
Others	HTMR	Height/median ratio: HOME divided by WD	-	Duong (2010)
	VDR	Vertical distribution ratio: WD minus HOME divided by WD	-	
	FS	Front slope: vertical angle from the beginning of the waveform to the amplitude of the first peak	degrees	

* DN: Digital Number

ALSFW metrics HOME, WD, NP, ROUGH, RWE and FS are described graphically in Figure 2.9 from section 2.3.4, and HTMR and VDR depend on HOME and WD. The other ALSFW metrics mentioned are described below in different figures, due to the large number of metrics.

Figure 3.2 shows the representation of MAX E, START PEAK and PEAK END. Generally, waveform peaks with the maximum energy correspond to the ground. However, if most of the energy is intercepted prior to reach the ground, maximum energy may be located in the denser vegetation strata. Therefore, these metrics not only provide information about the presence of vegetation and detection of the ground, but also about the densest strata. MAX E corresponds with the maximum energy of the waveform, which in this example is located at a height of 7.2 m. START PEAK is the distance between the beginning of the waveform and the height of the maximum energy, while END PEAK is the distance between the height of the maximum energy and the ground.
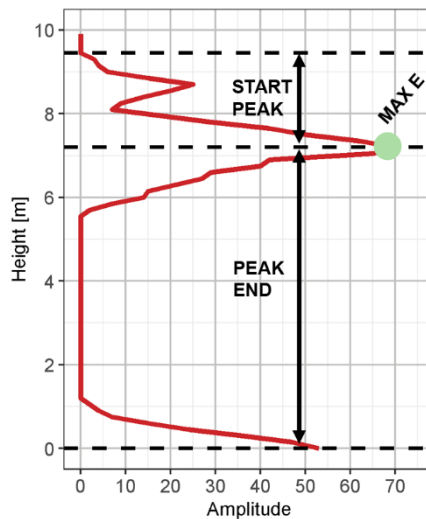
Figure 3.2. *Representation of ALSFW metrics MAX E, START PEAK and END PEAK. The waveform signal is represented in red, metric marks in black dashed lines, and MAX E with a green point.*

Figure 3.3 shows the amount of energy from the different height and energy quarters in Figure 3.3a and b, respectively. These quarters of height and energy are computed from the distance (i.e., WD) and the maximum energy of the

waveform (i.e., MAX E), respectively. Height quarters provide information about the distribution of the vegetation along the vertical structure. Energy quarters represent if the energy is focused on one peak or distributed in different strata. Afterwards, ALS_FW metrics HEIGHT Qn and ENERGY Qn are computed as the ratio of energy contained in the corresponding quarter divided by the total amount of energy (i.e., RWE).
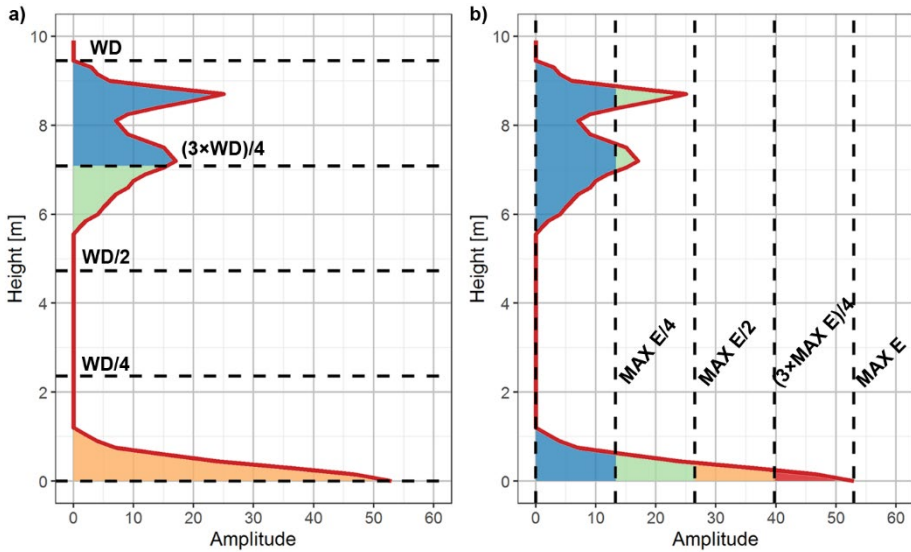
Figure 3.3. *Representation of the amount of energy from the different quarters of (a) height and (b) energy used for the computation of ALS_FW metrics HEIGTH Qn and ENERGY Qn, respectively. The waveform signal is represented in red, metric marks in black dashed lines, and height and energy quarters filled in blue, green, orange and red.*

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 a given number provided by the user (Figure 3.4a). To account for lower shrubs close to the ground and a more open understory, the HFEVT calculates the height of the first filled voxel above a given threshold (Figure 3.4b). FVU and NFVU are related to understory cover. FVU examines if there are any filled voxels between two given heights (Figure

3.4c), and NFVU is the number of filled voxels divided by the number of voxel between these two heights (Figure 3.4d).
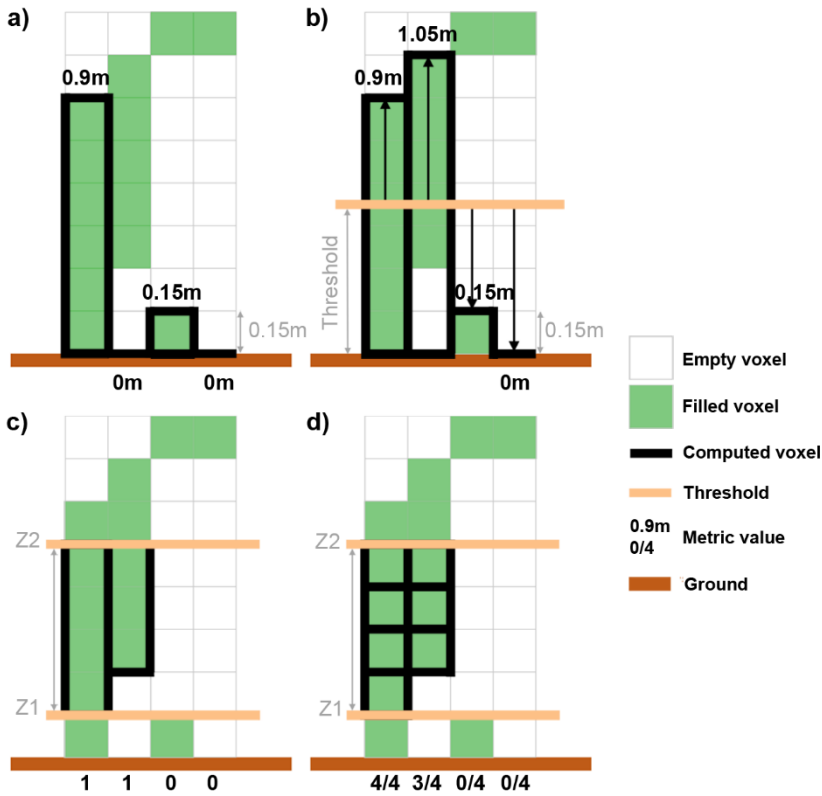
Figure 3.4. *Representation of voxel transects to describe ALSFW metrics (a) HFEV, (b) HFEVT, (c) FVU, and (d) NFVU. Voxel height is equal to 0.15 m and metric values for each column of voxels are written in black. Height thresholds in (b), (c), and (d) are user inputs.*

Following with the ALSFW metrics related to understory, EFEV and nEFEV are represented in Figure 3.5. These two metrics are related to the properties of the understory. EFEV 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 amplitudes of the whole waveform (i.e., RWE). Additionally, the ALSFW metric Hn is also represented in Figure 3.5 with different percentiles. To compute this metric, it

is necessary to check at which height a given percentile of energy is reached. The height of percentile 50 of energy (i.e., H50) is equal to HOME.
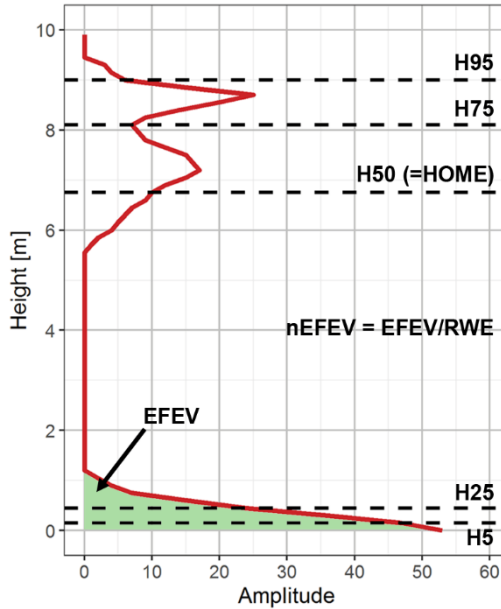
Figure 3.5. *Representation of ALS$_{FW}$ metrics EFEV, nEFEV and Hn. In this case percentiles 5, 25, 50, 75 and 95 are represented for the metric Hn. The waveform signal is represented in red, metric marks in black dashed lines, and EFEV filled in green.*

The following ALS$_{FW}$ metrics are based on the Gaussian iterative decomposition of the waveform. This decomposition consists of splitting the original waveform into different Gaussian curves, until the difference between the original waveform and the resulting Gaussian curves is below a given tolerance (Hofton et al., 2000). ALS$_{FW}$ metrics related to the Gaussian iterative decomposition were designed by Zhang et al. (2011) for large footprint ALS$_{FW}$, and Hancock et al. (2015) showed that Gaussian iterative fitting was the most accurate method comparing energy values for large footprint ALS$_{FW}$. However, in this study the potential of these metrics as descriptors for forestry applications was tested, since according to Hancock et al. (2015), energy differences for the Gaussian iterative method and small footprint ALS$_{FW}$ were small as well (i.e., nRMSE (normalized root-mean-square error) = 1.37%). In the

example represented in Figure 3.6, the original waveform is split into three Gaussian curves, therefore N GS is equal to three. After applying the Gaussian iterative decomposition to the waveform signal, the other ALSFW metrics represented in Figure 3.6 and Figure 3.7 require the computation of the boundary between ground and canopy curves (Zhang et al., 2011). The height of this boundary is defined as the height at the ground curve plus 1.5 times its standard deviation, being the ground curve the one with the highest amplitude located in the half of the waveform with the lowest heights. Once ground and canopy curves are identified, ALSFW metrics N GS STARPEAK and N GS ENDPEAK are computed as the number of Gaussian curves above and below the boundary, respectively (Figure 3.6). GE is equal to the amount of energy between the ground and the boundary, and GRR is the proportion between this amount of energy at the ground curve and the total amount of energy (i.e., RWE). At the canopy curve, CHn are computed as the height of the different energy quartiles from the canopy curve energy, and Rn are the proportion between the different CHn metrics and the total height of the waveform (i.e., WD).

Also from the Gaussian iterative decomposition Figure 3.7 represents ALSFW metrics BC, BCE, BCD, CD, CE and CER. BC 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. These metrics are related to understory, since it is assumed that the first energy peak excluding the ground must be related to either the understory or the canopy base. Regarding the overstory, CD is the distance between the beginning of the waveform and the boundary. CE is the energy from the beginning of the waveform to the boundary. CER is equal to CE normalized by the total energy of the waveform (i.e., RWE).

Finally, there is a set of ALSFW metrics (AGS, SGS and MSGS) related to overstory and proposed by Zhang et al. (2011). These metrics are also extracted from the Gaussian iterative decomposition, but only from the Gaussian curves above the boundary (i.e., the canopy). Prior to the computation of these metrics, the slope of each Gaussian curve must be calculated as the ratio between the amplitude and the standard deviation of each Gaussian curve (Equation 13).
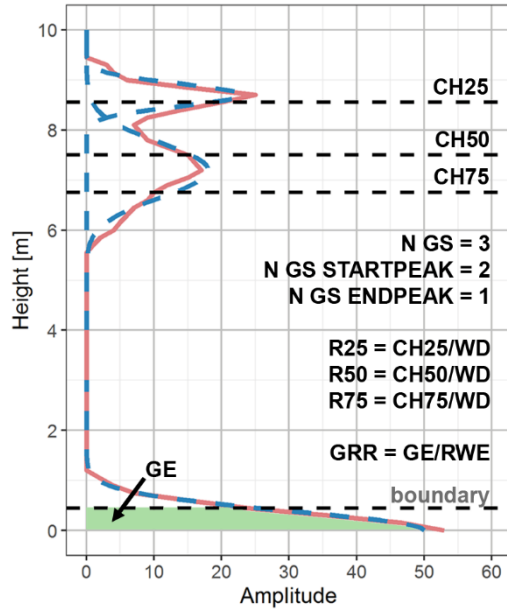
Figure 3.6. *Representation of ALSFW metrics N GS, N GS STARTPEAK, N GS ENDPEAK, GE, GRR, CHn and Rn, derived from the Gaussian decomposition. The waveform signal is represented in red, the resulting Gaussian iterative decomposition in blue dashed line, metric marks in black dashed lines, and GE filled in green.*
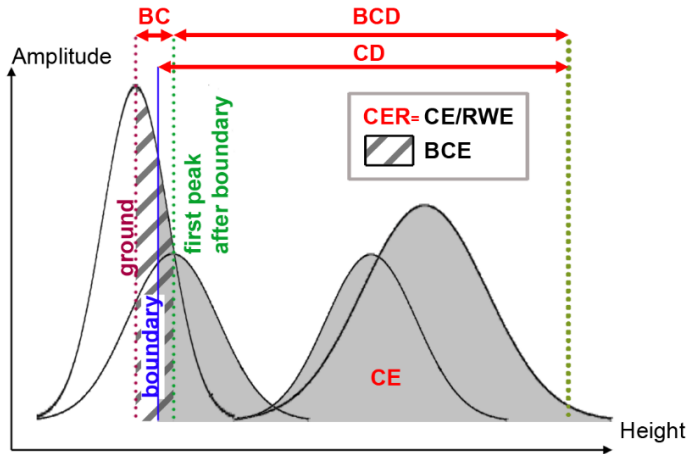
Figure 3.7. *Representation of ALSFW metrics BC, BCD, BCE, CD, CE and CER from the Gaussian iterative decomposition of the waveform signal.*

79

Next, AGS, SGS and MSGS metrics may be computed (Equation 14, Equation 15 and Equation 16, respectively).

$$S_i = \frac{A_i}{\sigma_i}$$

Equation 13

where i = current Gaussian curve above the boundary,

S_i = slope of the current Gaussian curve i,

A_i = amplitude of the current Gaussian curve i,

σ_i = standard deviation of the current Gaussian curve i.

$$AGS = \bar{S} = \sum_{i=1}^{n} S_i$$

Equation 14

where i = current Gaussian curve above the boundary,

S_i = slope of the current Gaussian curve i,

n = number of Gaussian curves above the boundary.

$$SGS = \Delta S = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (S_i - \bar{S})^2}$$

Equation 15

where i = current Gaussian curve above the boundary

S_i = slope of the current Gaussian curve i,

$\bar{S}$ = mean slope of the Gaussian curves above the boundary,

n = number of Gaussian curves above the boundary.

$$MSGS = \Delta S' = \sqrt{\sum_{i=1}^{n} \frac{E_i}{E_T}(S_i - \bar{S})^2}$$

Equation 16

where i = current Gaussian curve above the boundary

S_i = slope of the current Gaussian curve i,

$\bar{S}$ = mean slope of the Gaussian curves above the boundary,

E_i = energy of the current Gaussian curve i,

E_T = total energy of the Gaussian curves above the boundary (i.e., $\sum_{i=1}^{n} E_i$),

n = number of Gaussian curves above the boundary.

3.3. Software tool

The software tool *WoLFeX* is designed to perform all the processing steps described in previous sections, from the relative radiometric correction of ALS_FW data to the extraction of ALS_FW metrics for generating either regression or classification models, which can be further applied in larger study areas.

WoLFeX is divided into five sections (Figure 3.8): Inputs, Radiometric correction, Voxelization parameters, Metrics and Execution. In the Inputs section, the user selects the ALS_FW data files, typically LAS files (version 1.3 and point format 4), the DTM for the height normalization, and a workspace to save the outputs. In order to process a smaller area, data may be clipped using the limits saved in shapefile format with a polygon geometry. In addition, if the clip area is representing objects such as plots or segmented trees, an id field from the shapefile must be selected in the Inputs section to identify the different processed objects in the output statistics. To apply a radiometric correction of the data, the trajectory files related to the LAS files must be selected. This trajectory files can be in *.txt or *.trj format and they are needed to compute the
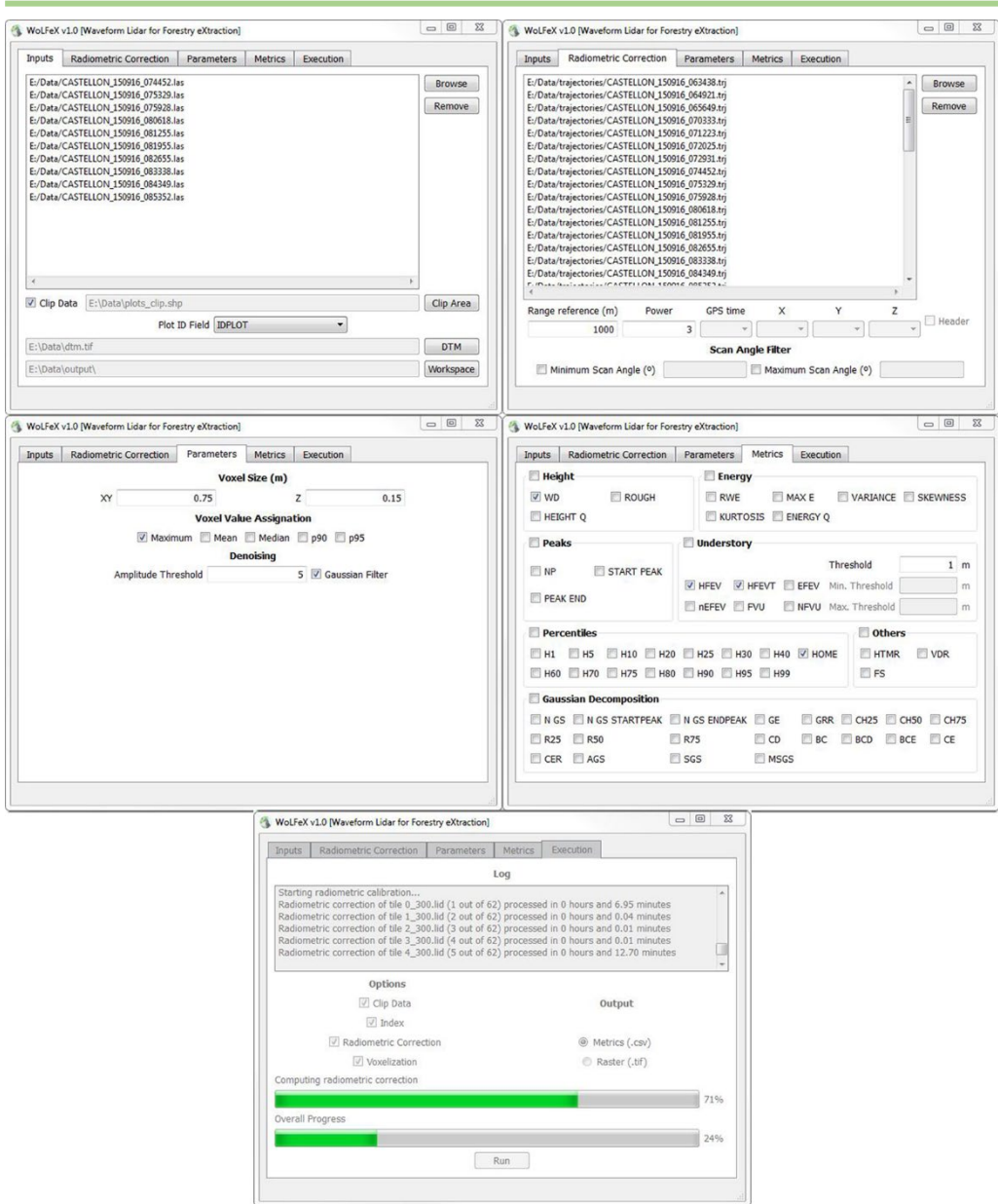
Figure 3.8. *Graphic user interface of WoLFeX and the five different sections: (a) Inputs, (b) Radiometric correction, (c) Voxelization parameters, (d) Metrics and (e) Execution.*

trajectory of each waveform. When the format is *.txt, the user should select the fields containing the GPS time, X, Y, and Z coordinates, and specify if there is a header in the text file. For both formats (*.txt and *.trj), the user should introduce

a range of reference and a power *n* as described in Equation 1 in section 2.3.1. On the other hand, *WoLFeX* also allows for filtering by scan angle intervals without selecting trajectory files. This option can be used to process a narrower range of scan angles to minimize the effect of a wide range of incidence angles on radiometric values. The third section allows for the selection of the *voxel size* and the *assignation value*. In the Metrics section, the user can select the specific ALSFW metrics to compute. Lastly, the Execution sections allows for the selection of the different processing steps that the user wants to execute, as well as the output format (*.csv or *.tif) for the metrics. In addition, this is the section where the completed steps or possible error messages are printed after the execution of the process.

3.4.   Case of study: Influence of radiometric correction and forest fuel modeling

3.4.1. Material and methods

In this application example, we tested the effect of the relative radiometric correction of ALSFW data on modeling forest fuel attributes for different scan angle intervals using the described software tool *WoLFeX*. The study area was located in the Natural Park of Sierra de Espadán, 30 km west of the Mediterranean Sea in eastern Spain (Figure 3.9a). The region is highly mountainous with steep hillsides, where elevation ranges from sea level to 1,100 m within 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*).

ALSFW data were acquired on September 16th 2015 over 7,465.53 ha using a LiteMapper 6800 with an average pulse density of 14 pulses·m^{-2}. Data were acquired at a flight altitude between 600 and 820 m above ground level, at 300
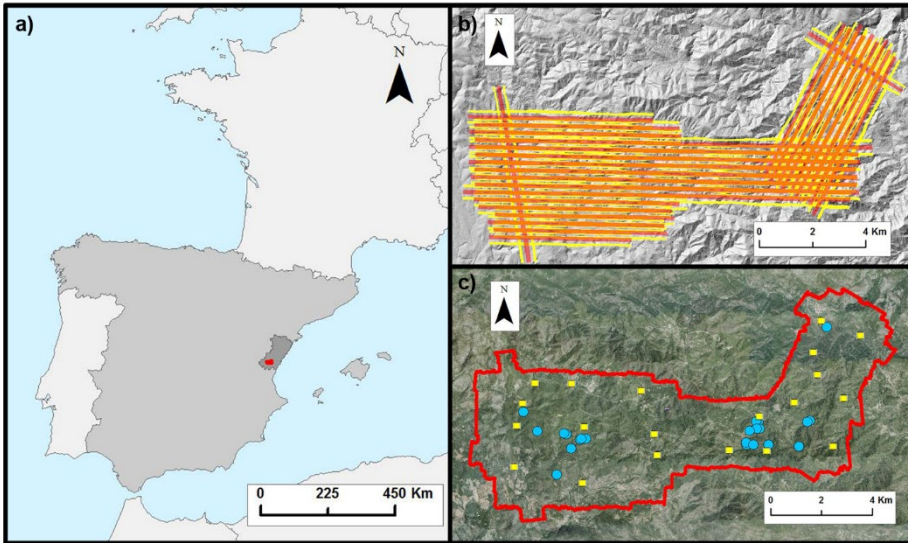
Figure 3.9. *Maps of (a) general location of the study area in Natural Park of Sierra de Espadán (Castellón, Spain), (b) flight stripes categorized by scan angle interval (0º-5º in orange and 15º-20º in yellow), and (c) sample and plot locations (square samples for analyzing radiometric differences in RWE are represented in yellow; and circular plots for analyzing estimation of forest fuel attributes in blue).*

kHz pulse frequency, and with a scan angle of ± 30º. 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, ALS$_{FW}$ data were discretized by the service provider (IMAO, France) using the Gaussian pulse estimation computation method to extract ALS$_D$ data, resulting in an average point density of 36 points·m^{-2}. The vertical accuracy of the ALS$_D$, verified using a set of ground control points located in open flat areas, was 4.3 cm (RMSE). Afterwards, ALS$_D$ data were used to create the DTM.

Samples differed according to the test. Firstly, the influence of radiometric correction and scan angle on the values of ALS$_{FW}$ metrics was analyzed in Test 1. To do this, 20 square samples of 75 m side (i.e., 5,625 m^2) were selected in areas registered from different scan angles but with similar pulse densities.

Secondly, 22 circular plots of 15 m radius (i.e., 706.86 m^2) from a total of 70 were selected for Test 2, where the influence of the application of the radiometric correction and scan angle on modeling forest fuel attributes was analyzed. Selected plots also needed to be registered from different scan angles and field data to estimate the forest fuel attributes. The locations of the 42 samples and plots for both tests is shown in Figure 3.9c. Ground-truth data collected from the 22 plots of Test 2 included DBH from trees with a value greater or equal to 5 cm, height and CBH from the seven trees with largest DBH, and tree species (see field survey sheet in Figure 3.10). Afterwards, allometric equations provided by Montero et al. (2005) were used to compute the reference data of three forest fuel attributes: CFL, CH, and CBH.

ALS$_{FW}$ metrics were extracted using *WoLFeX*, as described in section 3.2, for the different combinations of scan angle intervals and relative radiometric corrections. Radiometric correction reduces the effect of energy loss of the pulse due to different factors such as range (i.e., distance from the sensor to the target), attenuation (because of penetration of pulse through vegetation), and angle of incidence (slope and target orientation) (Kashani et al., 2015). Given that the RWE metric represents the sum of the waveform amplitudes from the beginning of the canopy to the ground, it is highly sensitive to pulse energy losses along the trajectory. For this reason, and in order to avoid redundancies in the test, only this metric was selected as a good indicator to evaluate the influence of scan angle and radiometric correction on ALS$_{FW}$ metrics. Hence, RWE metric was extracted for the samples of Test 1, while all the metrics from Table 3.1 were extracted for the plots of Test 2. The two scan angle intervals tested were 0º-5º and 15º-20º, in an attempt to differentiate between nadir and off-nadir pulses, respectively. Although the greater the scan angle, the greater the influence on radiometric values, we selected a maximum scan angle of 20º to have enough samples with enough size for the test. Four options were considered for the relative radiometric correction: uncorrected data, and corrected data varying the power *n* of Equation 1 in section 2.3.1 (i.e., *n*=2, *n*=3 and *n*=4). Hence, eight different datasets were computed (i.e., the combination of the two scan angle intervals and the four options for the relative radiometric correction) for the two

sets of samples and plots (i.e., Test 1 and 2) (Table 3.2).

Figure 3.10. *Example of field survey sheet from one of the plots used during the field campaign.*

Table 3.2. *Combination of datasets tested for Test 1 and 2.*

		Scan Angle Interval	
		0º-5º	15º-20º
Radiometric Correction	**Uncorrected**	(1)	(2)
	n=2	(3)	(4)
	n=3	(5)	(6)
	n=4	(7)	(8)

In Test 1, the mean values of RWE were computed for each sample at the different combinations. The RWE value variations were computed as the differences at sample-level between two combinations. The combinations compared had the same radiometric correction but different scan angle interval (i.e., (1)-(2), (3)-(4), (5)-(6) and (7)-(8) in Table 3.2). Additionally, the corrected data with a given scan angle interval were compared to their corresponding interval of uncorrected data (i.e., (3)-(1), (4)-(2), (5)-(1), (6)-(2), (7)-(1) and (8)-(2) in Table 3.2). Results were evaluated using the RMSE of these differences and the nRMSE, computed as the RMSE divided by the range of RWE values in the sample.

3.4.2. Results and discussion

Table 3.3 shows the RWE differences between the different scan angle intervals by means of RMSE and nRMSE. Results show that differences in RWE values between scan angle intervals decreased when relative radiometric correction was applied and as power (n) increases. For instance, uncorrected data had an RMSE of 262.29, while corrected data had a value of 117.41 and 93.25 for a power n=3 and n=4, respectively. This means that the influence of the scan angle on the metric value was smaller when using radiometrically corrected data, but it is not completely removed. On the other hand, results in Table 3.4 show that differences between uncorrected and corrected data increased as the power n increased, and the effect of the radiometric corrections on the metrics was more obvious at small scan angle intervals. For instance, differences between corrected data with a power n=2 and a scan angle interval of 0º-5º were equal to 183.86, while using the same scan angle interval and a power n=4 differences were equal to 299.29. Moreover, using the same power n, but a scan angle interval of 15º-20º resulted in differences of 90.74 and 164.43, respectively. This means that the larger the power n and the smaller the scan angle, the larger the correction that is applied to the uncorrected data. Analyzing Equation 1 in section 2.3.1, this occurred in this study since the range reference used was larger than the rest of ranges; otherwise it would be the opposite.

Table 3.3. *RWE differences between the different scan angle intervals (0º-5º and 15º-20º). RMSE: root-mean-square error; nRMSE: normalized RMSE.*

Radiometrically corrected data	Power *n*	Comparison (Table 3.2)	Differences between 0º-5º and 15º-20º	
			RMSE	nRMSE
No	-	(1)-(2)	262.29	15.40%
Yes	2	(3)-(4)	150.71	13.32%
Yes	3	(5)-(6)	117.41	11.21%
Yes	4	(7)-(8)	93.25	11.74%

Table 3.4. *RWE differences between the radiometrically uncorrected and corrected data.*

Scan angle interval	Power *n*	Comparison (Table 3.2)	Differences between Uncorrected and Corrected Data	
			RMSE	nRMSE
0º-5º	2	(3)-(1)	183.86	9.98%
15º-20º		(4)-(2)	90.74	5.39%
0º-5º	3	(5)-(1)	248.31	13.48%
15º-20º		(6)-(2)	128.57	7.63%
0º-5º	4	(7)-(1)	299.29	16.25%
15º-20º		(8)-(2)	164.43	9.76%

For the Test 2, the mean and the standard deviation of all the ALS$_{FW}$ metrics described in section 3.2 were computed at plot-level using *WoLFeX*. As a result, the software tool provided a *.csv file that was used as input file in statistical software. All possible combinations of linear regression models with a maximum of three metrics were computed, finally selecting the model with the minimum AIC. Among the selected ALS$_{FW}$ metrics, those proposed in the present thesis (i.e., KURTOSIS, HFEV, HFEVT, EFEV, nEFEV, FVU, NFVU, BC, BCD, BCE, CD, CE, and CER) were among the most selected, and therefore they had an influence on estimating forest fuel attributes. For instance, KURTOSIS and HFEVT were selected to estimate CFL and CBH, BC and CD for CH and CBH, EFEV for CFL, HFEV, nEFEV, NFVU and BCD for CH, and CE for CBH. Afterwards, a model was obtained for each of the three forest fuel variables (i.e., CFL, CH, and CBH), each combination of scan angle interval (i.e., 0º-5º and 15º-20º), and each radiometric correction type (i.e., uncorrected and corrected data with a power $n = 2$, $n = 3$, and $n = 4$). The linear regression models were evaluated

using leave-one-out cross-validation and computing the R^2, RMSE, nRMSE and coefficient of variation (CV). Table 3.5 shows the prediction results of the forest fuel attributes (i.e., CFL, CH, and CBH) using varied scan angle interval and the radiometric correction. The prediction of CFL was considerably improved when a radiometric correction was applied with a higher power *n* for both scan angle intervals, varying R^2 from 0.62 to 0.79 and from 0.68 to 0.85 for scan angle intervals of 0º-5º and 15º-20º, respectively. However, CBH prediction results did not improve, or even slightly worsened, when a radiometric correction was applied. In this case, differences were also smaller compared to CFL, and as in the CH predictions, results were similar for a scan angle interval of 15º-20º. The test shows that the influence of the radiometric correction was smaller predicting height attributes, such as CH and CBH, than in predicting mass-related attributes, such as CFL. Height attributes are fixed at a specific point on the waveform, usually a maximum or minimum, while mass-related attributes are described using the complete waveform profile. Therefore, the latter are more subject to radiometric values. Additionally, the difference between uncorrected and corrected data for the three forest fuel attributes was smaller when the scan angle interval was 15º-20º than when it was 0º-5º. A previous analysis (Morsdorf et al., 2008) found that parameters corrected by radiometric correction such as flying altitude and incidence angle have an influence on estimates of biophysical vegetation properties (i.e., tree height, crown width, fractional cover, and leaf are index). However, the influence of scan angle was not as apparent here, probably due to the use of small scan angles. Additionally, other studies also mentioned the radiometric correction as a key step in using backscattered measurements to estimate geophysical vegetation properties or similar analyses (Wagner, 2010; Sevara et al., 2019).

3.5.   Conclusions

In this chapter, we presented and described a software tool named *WoLFeX*, designed to process ALS_FW data, which includes a wide range of new proposed ALS_FW metrics. We assessed the influence of radiometric correction on ALS_FW metrics and on the estimates of forest fuel attributes through *WoLFeX*.

Table 3.5. *Prediction of forest fuel attributes (i.e., CFL, CH and CBH) using varied scan angle interval and radiometric correction.*

Forest fuel attribute	Scan angle interval	Radiometrically corrected data	Power n	R^2	RMSE	nRMSE	CV
CFL	0°–5°	No	-	0.62	4.61 Mg·ha⁻¹	0.15	0.20
		Yes	2	0.75	3.67 Mg·ha⁻¹	0.12	0.16
		Yes	3	0.77	3.55 Mg·ha⁻¹	0.11	0.15
		Yes	4	0.79	3.37 Mg·ha⁻¹	0.11	0.15
	15°–20°	No	-	0.68	4.13 Mg·ha⁻¹	0.13	0.18
		Yes	2	0.72	3.84 Mg·ha⁻¹	0.12	0.17
		Yes	3	0.83	3.04 Mg·ha⁻¹	0.10	0.13
		Yes	4	0.85	2.86 Mg·ha⁻¹	0.09	0.12
CH	0°–5°	No	-	0.89	1.09 m	0.08	0.08
		Yes	2	0.88	1.18 m	0.09	0.09
		Yes	3	0.93	0.89 m	0.07	0.07
		Yes	4	0.86	1.25 m	0.10	0.09
	15°–20°	No	-	0.91	1.02 m	0.08	0.08
		Yes	2	0.90	1.05 m	0.08	0.08
		Yes	3	0.91	0.99 m	0.08	0.07
		Yes	4	0.92	0.95 m	0.07	0.07
CBH	0°–5°	No	-	0.94	0.68 m	0.06	0.11
		Yes	2	0.93	0.73 m	0.06	0.12
		Yes	3	0.92	0.83 m	0.07	0.14
		Yes	4	0.89	0.89 m	0.07	0.15
	15°–20°	No	-	0.96	0.56 m	0.05	0.09
		Yes	2	0.92	0.74 m	0.06	0.12
		Yes	3	0.95	0.61 m	0.05	0.10
		Yes	4	0.95	0.59 m	0.05	0.10

This tool allows for clipping, radiometrically correcting, voxelizing the original ALS$_{FW}$ waveforms, creating pseudo-vertical waveforms and extracting an exhaustive set of object-oriented metrics. These metrics are saved into a *.csv file that can be used as an input file for generating either regression or classification models, such as forest fuel attributes or fuel types, respectively. Among these metrics, those related to the understory vegetation are the most remarkable, since they have not been considered by other processing tools so far, and they allow for the location and quantification of understory vegetation, which is a key parameter for the characterization of fire behavior in Mediterranean forests. Processing ALS$_{FW}$ is more challenging than ALS$_D$, since it registers the complete return of the signal, and therefore it allows for a better detection of the lower strata.

From the case of study of *WoLFeX* software, different models of forest fuel attributes (CFL, CH, and CBH) were generated, varying processing parameters related to radiometric correction and scan angle interval of ALS$_{FW}$ data acquisition. These tests showed that differences in metric values measured from nadir and off-nadir were reduced when a relative radiometric correction was applied. The improvement of the models obtained when the relative radiometric correction of the data was applied was noteworthy – from R^2=0.62 up to R^2=0.79 in the case of CFL. However, height attributes (i.e., CH and CBH) were less strongly influenced by a relative radiometric correction, presenting only subtle differences.

The software *WoLFeX*, freely available for download at [http://cgat.webs.upv.es/software/], is an alternative for processing ALS$_{FW}$ data in an integrated manner. It includes the relative radiometric correction of the data, which plays an important role in reducing radiometric differences between different scan angles and may be essential for estimating some forest fuel attributes. It also extracts multiple new and previously proposed metrics to generate models that characterize forest structure. Among these metrics, the most remarkable are those related to understory vegetation, due to the potential of ALS$_{FW}$ to register the complete vertical forest structure. This opens a wide range of applications in environmental sciences, forestry and fire ecology.

Chapter 4

Comparative assessment of the vertical distribution of forest components using full-waveform airborne, discrete airborne and discrete terrestrial laser scanning
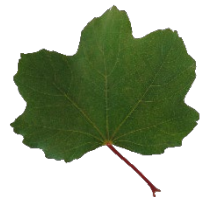
Acer opalus

"Caen a tierra en el bosque las hojas del arce:
palmas rojas de dioses con mensajes del cielo."

Jóhann Hjálmarsson

In the fourth chapter, three points are examined for a boreal and a Mediterranean forest with contrasted conifer canopy densities: (*i*) the characterization of the vertical distribution and occlusion from three laser scanning configurations: ALS_{FW}, ALS_D, and TLS; (*ii*) the comparison in the detection of understory vegetation by ALS_{FW} and ALS_D using TLS as reference; and (*iii*) the use of a methodological procedure based on the Gini index concept to determine understory vegetation density classes from both ALS_{FW} and ALS_D configurations. Estimating occlusion along the different vertical strata allows for a better prediction of the limitations of laser scanning configurations in registering forest structure. On the other hand, detection and determination of understory vegetation density classes with ALS configurations is relevant for several applications such as characterizing wildlife habitats, assessing timber productivity and improving silvicultural decision-making in support of wildfire mitigation, all of them over large areas.

4.1.   Introduction

Signal occlusion is the main limitation in acquiring fully comprehensive laser scanning datasets in forested environments. Signal occlusion occurs when the object to sample is partially or completely obscured by an intervening object. The presence and amount of signal occlusion found in a laser scanning dataset depends greatly on the scanning configuration (i.e., above-canopy for aerial or near-ground for terrestrial), vegetation cover and density, and its complexity (Watt and Donoghue, 2005). Signal occlusion can therefore limit the detection of forest's horizontal and vertical distribution of vegetation, particularly in dense forested environments. It is important to note that signal occlusion is not to be confounded with forest cover. While forest cover refers to the proportion of forest covered by the vertical projection of tree crowns, signal occlusion refers to the shadow (lack of laser signal) caused by canopy elements. Several strategies have been tested to reduce signal occlusion such as increasing the number of flyovers for ALS (Kükenbrink et al., 2017), combining data from multiple sensors (Giannetti et al., 2018), or sampling the plot with multiple scans from varying viewpoints with a TLS (Martin-Ducup et al., 2017). Another way

to deal with signal occlusion in TLS data is to divide the point cloud space into voxels and compute the Plant Area Density (PAD) for each voxel (Béland et al., 2014; Pimont et al., 2018). Analyzing signal occlusion and its effects on the estimation of forest structural attributes is essential in understanding the limitations of different laser scanning and sampling configurations, and therefore sampling designs to best minimize signal occlusion.

Although signal occlusion is present in all laser scanning datasets, ALS and TLS systems have nonetheless demonstrated their capability to characterize forest attributes with great precision and accuracy. As mentioned previously in the overall introduction, ALS_D and ALS_{FW} are now used operationally to estimate stand attributes for a wide range of forest ecosystems, and classify tree species and forest canopy fuels, receiving ALS_{FW} less attention due to the data's greater complexity and the current lack of processing tools. In comparison with ALS data, TLS data can provide a more detailed point cloud of a forest structure, albeit from a different viewpoint. For applications in forested environments, the useful portion of the TLS point cloud extent is often limited (10 – 30 m) with a hemispherical view around the sensor. Withstanding that, many studies have demonstrated the capabilities of TLS to estimate and extract forest stand attributes (Watt and Donoghue, 2005; Moskal and Zheng, 2011; Kankare et al., 2013; Srinivasan et al., 2015; Liang et al., 2016; Ravaglia et al., 2019), and fewer on the classification of tree species (Othmani et al., 2013; Lin and Herold, 2016; Torralba et al., 2018).

Since ALS and TLS sensors acquire data from differing positions relative to the forests canopy, different occluded forest strata can be observed in their point clouds. Consequently, it is of interest to compare their independent effectiveness to detect the forests horizontal vertical distribution, and estimate forest structural attributes. Several studies provide a base for comparison between three laser scanning configurations: ALS_D, ALS_{FW} and TLS. These studies generally show that CH estimations are more accurate using ALS than TLS (Hilker et al., 2010), while characterization of the foliage profile is estimated with more accuracy by TLS, especially in the lower strata (Chasmer et al., 2006; Hilker et al., 2010), where understory vegetation is found. On the other hand, other studies concur on a more accurate estimation of forest structural attributes

from ALS_{FW} than ALS_D for CH (Anderson et al., 2016), AGB (Nie et al., 2017), stand volume (Lindberg et al., 2012), and the classification of species composition (Torralba et al., 2018).

A common challenge in predicting forest structure from ALS data is finding associated reference data from which reliable error estimation is possible. Most studies on forest structure variable use a combination of field measurements and allometric relationships as reference data (González-Ferreiro et al., 2012; Treitz et al., 2012; Ruiz et al., 2014). However, in cases where 3D assessments of vegetative material are required, it can be beneficial to use TLS data as a reference as these can be difficult and often logistically challenging to quantify directly from field activities. TLS data often represent the best available information to describe forest elements (Martin-Ducup et al., 2016; Ravaglia et al., 2019), being sampled in much more detail when compared with ALS. However, there are important discrepancies between ALS and TLS point clouds that need to be acknowledged. Looking at a vertical profile of data points in relative terms, ALS_D point clouds have far more hits within the upper canopy and on the ground, while most of the hits for TLS point clouds are located at the lower crown, trunks-stems, understory, and ground. These differences in sampling capabilities lead to estimation divergences (Chasmer et al., 2006). In general, ALS_D point clouds tend to under-represent the lower strata. While correlation between ALS_D and TLS point clouds has been found to be 0.48 for heights below 20 m, it reached 0.87 when only the upper canopy ($z > 20$ m) was considered (Hilker et al., 2010). As for CH, estimation from ALS_D and TLS were generally similar: (*i*) an underestimation of 1 m by the TLS in a mixed forest in Ontario, Canada (Chasmer et al., 2006), (*ii*) a correlation of 0.94 between estimations from both datasets in a pine-dominated forest in South-Korea (Jung et al., 2011), and (*iii*) a correlation near 1 in a coniferous forest on Vancouver Island, BC, Canada (Hilker et al., 2010). Conversely, few studies found ALS_D more suited to estimate CH than TLS with an R^2 of 0.96 and 0.86, respectively in a lodgepole pine forest in Alberta, Canada (Hilker et al., 2012). Therefore, there are situations where ALS and TLS may not accurately estimate the entire vertical forest structure, primarily due to signal occlusion. This signal occlusion

problem is more severe for ALS than it is for the TLS because of the much smaller number of laser pulses. Fortunately, the beam width and the multiple return configuration mitigate this problem. Consequently, sampling understory vegetation is comparably far more comprehensive from TLS than it is from ALS. TLS data are therefore suitable reference data for the estimation of understory structural attributes and preferred over using traditional sampling techniques, which can be laborious and time consuming.

Within the vertical distribution of the forest, detection of the lower strata can be challenging from ALS sensors, especially from ALS_D as a results of signal occlusion from the overstory (Anderson et al., 2016). Nonetheless, ALS_D has proven to discriminate presence and absence of understory vegetation with promising accuracy, e.g., with $R^{2'}$s of 0.83 (Martinuzzi et al., 2009), 0.77 (Hill and Broughton, 2009), 0.74 (Wing et al., 2012), and 0.48 (Morsdorf et al., 2010), and has been demonstrated to be more accurately estimated by ALS_{FW} than ALS_D (Hancock et al., 2017; Torralba et al., 2018). In contrast, point clouds from TLS provide a large amount of detail on understory vegetation due to the position of the sensor (Liu et al., 2017). However, most studies have focused on the ability of TLS to characterize shrubs in ecosystems absent of overstory (Vierling et al., 2013; Olsoy et al., 2014; Greaves et al., 2015); few have focused on characterizing the understory vegetation from forested ecosystems (e.g., Chen et al. (2016)). Furthermore, few studies have estimated understory cover with ALS_{FW} using TLS as reference data and obtained a nRMSE of 24% (Hancock et al., 2017). Nevertheless, a limited amount of studies have compared the ability to estimate understory vegetation presence or distribution from three different laser scanning datasets (i.e., ALS_{FW}, ALS_D and TLS) (e.g., Hancock et al. (2017) and Torralba et al. (2018)). These studies confirmed that ALS_{FW} and TLS are both capable of estimating the spatial distribution of understory vegetation in more detail than using ALS_D data, e.g., with overall accuracies of 86.4% and 77.3%, respectively (Torralba et al., 2018). These results demonstrate the potential of combining ALS_{FW} and TLS data in a workflow that estimates the spatial distribution of the understory vegetation beyond what can be estimated from ALS_D data. However, these results were limited in their application as they were assessed in a fragmented urban forest and an open Mediterranean forest.

In this chapter, in section 4.2, the characteristics of both study areas and laser scanning configurations are described. In addition, the different processing steps followed to characterize the vertical distribution and occlusion, to compare the detection of understory vegetation from ALS_{FW} and ALS_D, and to determine the understory vegetation density through the Gini index and ALS data are described and enumerated. The results obtained following the different processing steps are shown in section 4.3. In section 4.4, the discussion of the results is addressed. Finally, in section 4.5, the conclusions of this chapter are presented. This chapter sets out the Hypotheses 4 and 5: "ALS and TLS configurations are limited to detect lower and upper strata, respectively, because of the position of the sensors. This limitation depends on the forest ecosystem and density" and "Given that ALS_{FW} registers the complete signal going through the vertical forest structure, and ALS_D is a simplification of ALS_{FW}, understory vegetation can be detected and its density determined more accurately with ALS_{FW} than with ALS_D", respectively. To demonstrate these hypotheses, Objectives 5 and 6 are addressed in this chapter: (*i*) characterize the signal occlusion along the vertical forest structure using different laser scanning configurations (i.e., ALS_D, ALS_{FW}, and TLS) in contrasted ecosystems with different canopy covers to determine how reliable the resulting vertical distribution profiles are based on the amount of occlusion and the lack of information, and (*ii*) determine how understory vegetation density classes can be detected and further determined by ALS configurations, and whether ALS_{FW} allows the detection and determination to a level of detail beyond ALS_D capability.

4.2. Material and methods

4.2.1. Study areas

Two study areas were selected based on their contrasting canopy densities and understory vegetation presence. Both sites are conifer dominated, albeit, structurally very different. Our first study area (111,257 ha) is located in a Boreal Shield Ecozone in western Newfoundland and Labrador, Canada (Figure 4.1).

The ecoregion is dominated (~70%) by forest land and is located within the most eastern boreal forest region of North America. Balsam fir (*Abies balsamea* (L.) Mill) is the dominant tree species of the regions followed by Black spruce (*Picea mariana* (Mill.) Britton, Sterns & Poggenb.). White birch (*Betula papyrifera* Marsh.), yellow birch (*Betula alleghaniensis* Britton), white spruce (*Picea glauca* (Moench) Voss) and eastern larch (*Larix laricina* (Du Roi) K. Koch) are present to a much lesser extent. The relief is gently undulating to hilly with elevation ranges between ~30 m and 640 m. Forest understory is extremely variable depending on stand density and age, soil conditions, status of regeneration and silvicultural treatments such as precommercial thinning (e.g., Figure 4.2 – Newfoundland, sparse understory). Understory vegetation can be composed of tree saplings and seedlings, ferns (e.g., *Dryopteris carthusiana* (Vill.) HP Fuchs) and to a lesser extent ericaceous shrubs (e.g., *Kalmia angustifolia* L., *Rhododendron groenlandicum* (Oeder) Kron & Judd, *Vaccinium* spp.).

The second study area is located in a Mediterranean forest in the Natural Park of Sierra de Espadán (Spain), and coincides with the one described in section 3.4.1. The presence of understory in this study area is very variable, mainly depending on the dominant species and soil properties. Understory vegetation is dominated by the following shrubs and flowering plants: rosemary (*Rosmarinus officinalis* L.), tree heath (*Erica arborea* L.), brezo (*Erica multiflora* L.), Mediterranean buckthorn (*Rhamnus alaternus* L.), kermes oak (*Quercus coccifera* L.) and mastic (*Pistacia lentiscus* L.). Figure 4.2 illustrates examples of different understory scenarios.

4.2.2. Forest plots

Circular plots were established with a radius of 11.28 m and 15 m for the Newfoundland and Spain sites, respectively. Plot center locations for both sites were measured with a GPS RTK with an average accuracy of ~0.40 m. As mentioned in section 3.4.1, tree species, living status, DBH, height and CBH were measured at all plot locations. For the Newfoundland site, 59 established experimental plots from Luther et al. (2019) were made available with associated ALS$_D$ and ALS$_{FW}$ data, while 70 established experimental plots were made available with similar data for the Spain site (Figure 4.1c and d). Among
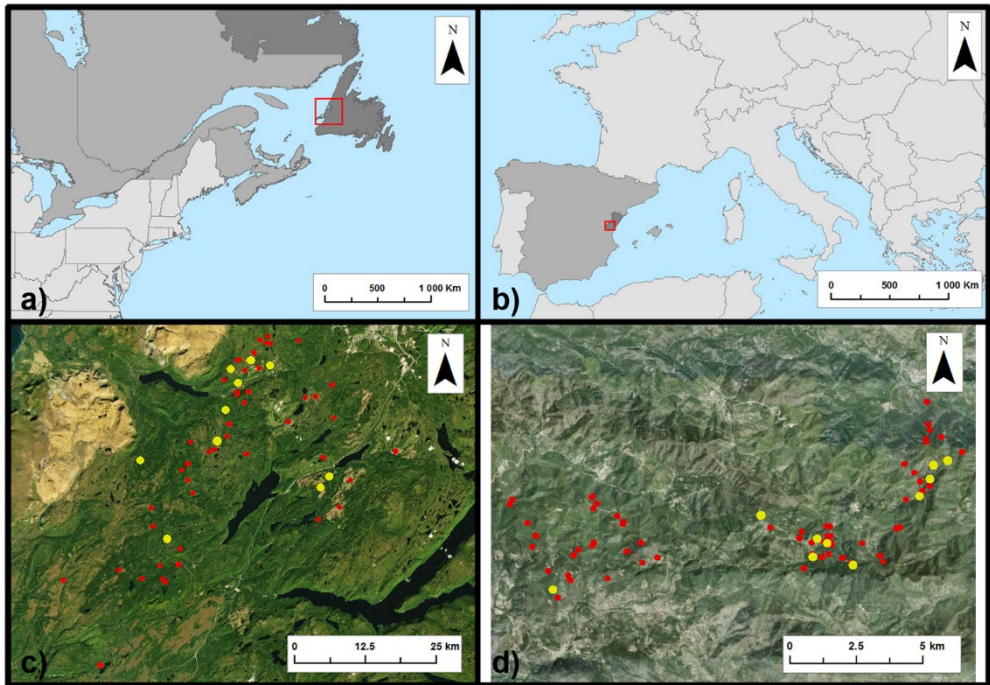
Figure 4.1. *Location of plots registered (red) and plots used in the current study (yellow) within each study area: (a, c) in western Newfoundland, Canada, and (b, d) in the Castellón province, Spain (Background imagery: PNOA and WorldView-2).*

these experimental plots, a structurally representative sample of ten plots was selected per site for TLS sampling by maximizing the variability of canopy cover and understory vegetation for analysis in this chapter. Regarding canopy cover, we estimated it all plot locations from the proportion of ALS$_D$ first hits to total hits above 2 m per McGaughey (2014). Then, plots with a percentage of first hits above 70% were classified as having dense canopy cover, between 40% and 70% as having a sparse canopy cover, and below 40%, as having a very sparse canopy cover. Furthermore, we assigned understory vegetation density classes at plot locations through field interpretations and classified plots as having dense, moderate, sparse, or absence of understory vegetation (Figure 4.2). Figure 4.3 illustrates the variability in structure from all conifer dominated plots and the structural representativeness of the retained plots.
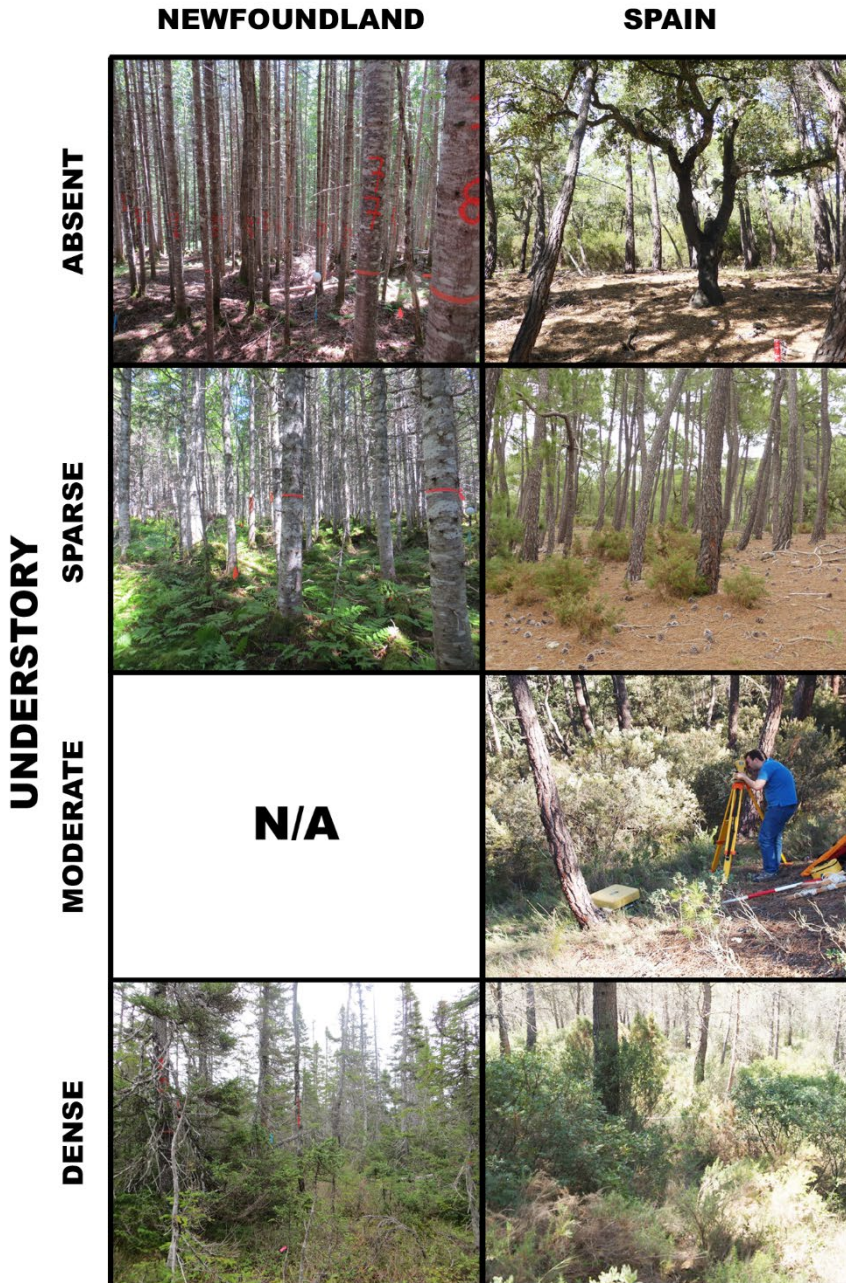
Figure 4.2. *Field photographs from the Newfoundland and Spain sites illustrating the varying densities of understory vegetation.*
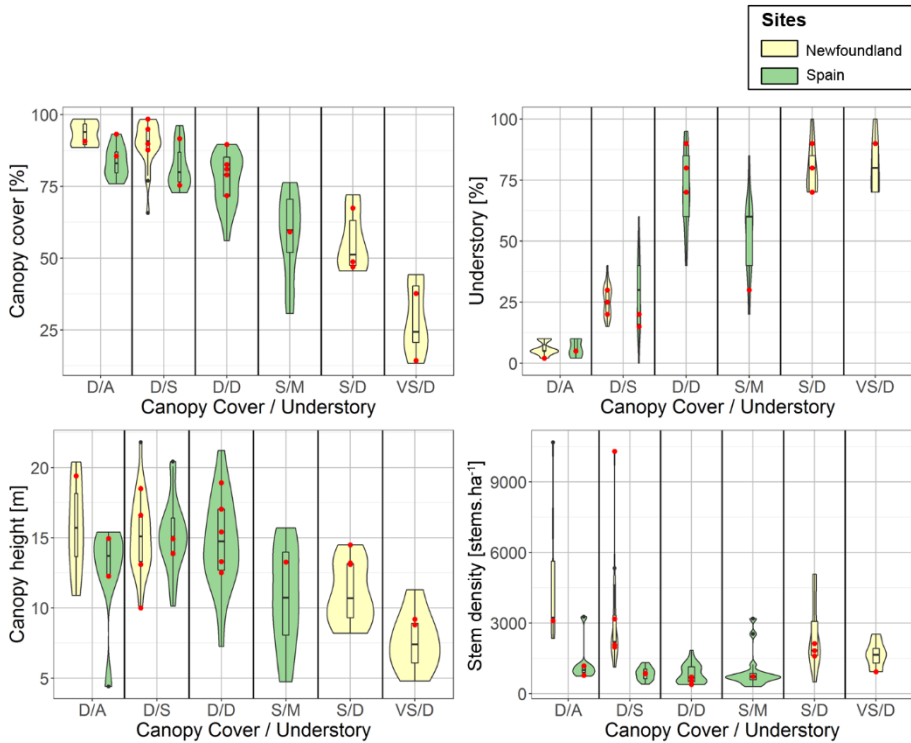
Figure 4.3. *Violin plots representing four structural attributes (canopy cover, understory, canopy height and stem density) from all available plots. Attribute values for plots retained for analysis are in red. Abbreviations: D-dense; M-moderate; S-sparse; VS-very sparse; A-absent.*

4.2.3. Laser scanning data

In this chapter, we analyzed laser scanning data obtained from three differing configurations, namely ALS$_D$, ALS$_{FW}$ and TLS. The ALS data obtained for the Newfoundland site were acquired between August 15th and September 24th 2016 with a Riegl LMS-Q680i. The approximate flight altitude was 1,000 m above ground level with an approximate speed of 100 knots. Data were acquired with a pulse frequency of 330 kHz and a scan angle range of ± 30º. Not excluding waterbodies, the overall average laser scanning pulse density was 7.34 pulses·m⁻

[2]. ALS$_{FW}$ data were discretized by the service provider (Leading Edge Geomatics, Canada) using the Gaussian pulse estimation computation method to extract ALS$_D$ data. Average point densities of 16 points·m^{-2} were observed at plot locations for ALS$_D$.

ALS data at the Spain site used in this chapter were the same used and described in section 3.4.1.

The TLS data were collected using a FARO FOCUS 3D 120 phase-based scanner (see specifications in Table 4.1) using a multi-scan configuration on both sites, recording only the first hit with an angular density between pulses of 0.0036 degree. TLS data for the Newfoundland site were acquired between June and August 2017 while the TLS data for the Spain site were acquired between September 29th and October 23rd, 2015. To minimize signal occlusion, each plot was scanned from nine positions: one at plot center, four at ~15 m from the center in each cardinal direction (i.e., N, W, S, E), and four at ~7.5 m and ~6 m from the center in each primary intercardinal direction (i.e., NW, SW, SE, NE) for the Spain and Newfoundland sites, respectively. Each scan identified a minimum of three co-registration spherical targets common with adjacent scans. Co-registration of the 9 scans was performed using FARO SCENE software version 6.2 (FARO, Lake Mary, FL). The resulting co-registered point cloud comprised, on average, 392 × 10^6 hits.

Table 4.1. *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

4.2.4. Overview of the methods

An overview of the methodological approach and associated procedures is presented in Figure 4.4. First, ALS_{FW} data were denoised and georeferenced in order to create an ALS_{FW} point cloud compatible with the ALS_D and TLS point clouds (Procedure 1). We then proceeded with the co-registration of the three laser scanning datasets: ALS_{FW}, ALS_D and TLS (Procedure 2). Once co-registered, all the point clouds were represented independently in voxel grids (Procedure 3a). Sampling of each voxel by the laser beams depends primarily on three variables: (*i*) the number of theoretical laser beams passing through the voxel (Nt), (*ii*) the number of these theoretical beams that were occluded prior to reaching the voxel (Nb), and (*iii*), the number of hits actually returned from these beams within the voxel (Nh). We then estimated these variables for each voxel of the three datasets (Procedure 3b) in order to derive the vertical distribution profiles from ALS (Nh^{ALS} for both ALS_D (Nh^{ALSD}) and ALS_{FW} (Nh^{ALSFW}) datasets) (Procedure 3c) as well as the Plant Area Density from TLS (PAD^{TLS}) (Procedure 3d), which is the projected surface of the vegetated materials (wood and leaves). In order to quantify and compare signal occlusion within the three laser scanning datasets, we computed for each the rate of pulse reduction of Nt as the proportion of beams blocked prior to reaching the voxel (Nb/Nt) (Procedure 3e). A height normalization was then applied to Nh^{ALS}, PAD^{TLS}, and the rate of pulse reduction from ALS_{FW}, ALS_D and TLS (Procedure 4). This created a coherent vertical leveling between these estimated variables from which we produced vertical profiles of Nh^{ALS}, PAD^{TLS} and the rate of pulse reduction (Procedures 5a and 5b). These vertical profiles were used to analyze the relationship between the detection of vegetative material in different strata from ALS and TLS configurations and the rate of pulse reduction. We computed the coefficient of correlation at lower strata ($0.5 \ m \leq z \leq 4 \ m$) to quantify the similarity of vertical profiles of Nh^{ALSFW} and Nh^{ALSD} with PAD^{TLS}. Afterwards, the lower strata of the Nh^{ALS} vertical profiles were compared to determine which ALS configuration (i.e., ALS_{FW} or ALS_D) depicts the understory vegetation in more detail. An application of the Lorenz curve (Lorenz, 1905) and the Gini index (Gini, 1912) calculated from the Nh^{ALSFW} and Nh^{ALSD} vertical

profiles of the lower strata determined the density of understory vegetation, which we compared with field observations.
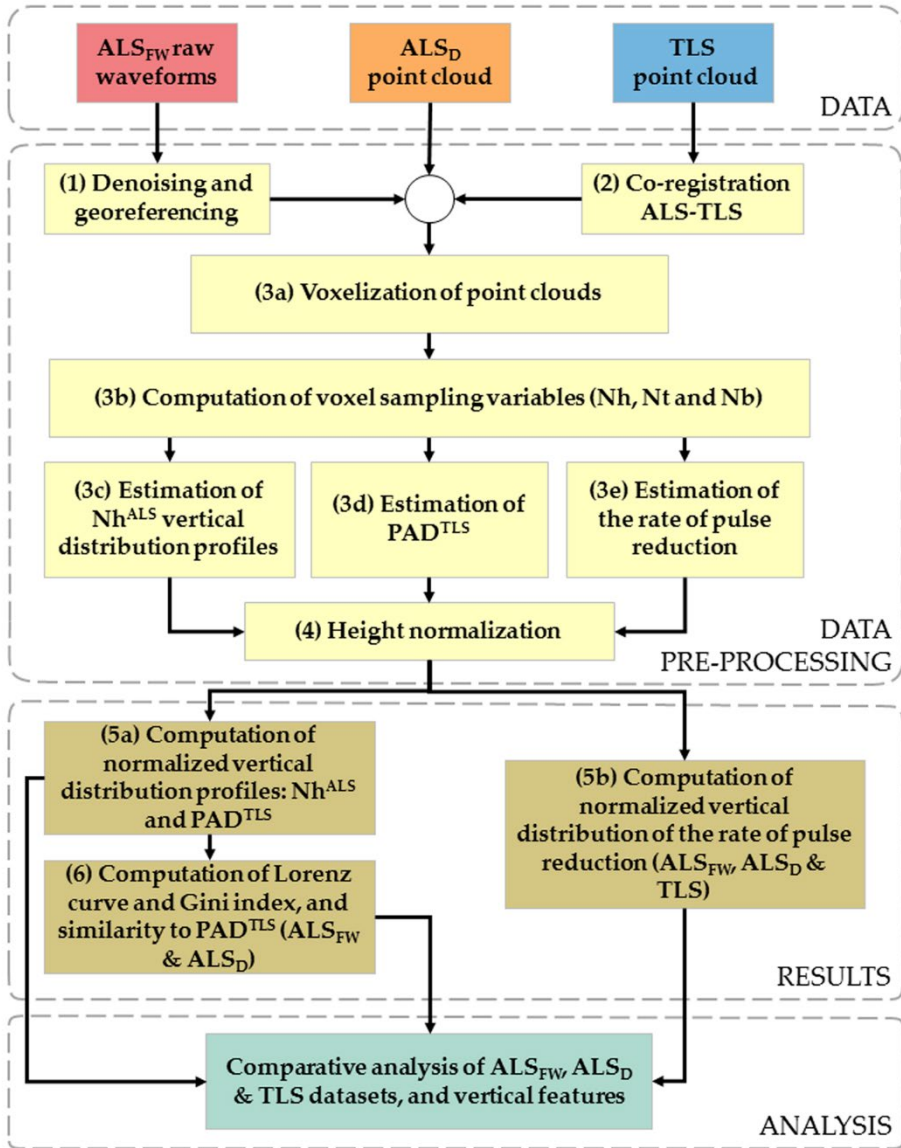
Figure 4.4. *Overview of the methodological approach.*

4.2.5. Data processing

4.2.5.1. Denoising

Initial data were available in point cloud format for the ALS_D and TLS configurations. Unlike ALS_D return sensors which record backscattered energy at precisely referenced points in time and space, ALS_{FW} sensors record backscattered energy as a nearly continuous signal in a full-waveform indexed bin. We therefore needed to create an ALS_{FW} point cloud compatible with the ALS_D and TLS point clouds. To do so, we removed the noise contained in the raw ALS_{FW} waveforms and georeferenced the remaining bins (Procedure 1). The denoising procedure followed was the one described in section 2.3.1. Afterwards, we computed each bin's XYZ coordinates following the procedure described in section 2.3.3. The resulting ALS_{FW} data contained only significant waveforms with the noise removed from which we were able to create a georeferenced point cloud compatible with those from the two other configurations: ALS_D and TLS.

4.2.5.2. XY co-registration

Considering that our main goal was to compare the data obtained from three different laser scanning configurations, it was necessary that all point clouds were co-registered in the same coordinate system (Procedure 2). Co-registration is a critical step to ensure that the three point clouds can be compared in our analysis. Georeferencing of the ALS data followed common practice and was done by registering the flight trajectory coordinates from the airborne GPS to a set of ground control points. These ground control points allowed for an accurate georeferencing of the flight lines, and therefore the resulting waveforms and point clouds. Given that ALS_D is derived from the ALS_{FW} data, co-registration between these two datasets was not necessary.

Co-registration of the TLS data to the ALS data was performed using the latter as reference. These data were co-registered on a plot-level basis. For each plot, Canopy Height Surfaces (CHS) were generated independently from both

ALS$_D$ and TLS data. The geometric distribution of tree crowns and canopy gaps guided the selection of homologous points from both CHS (see Figure 4.5). A 2D affine matrix transformation was then computed from the homologous point coordinates and applied to the TLS point clouds. Only translation in the horizontal plane and rotation around the vertical axis were applied since the distance values from both laser scanner systems needed to be maintained, therefore not altering the scale. The RMSE of the 2D affine transformation was 9 cm ± 4 cm and 7 cm ± 7 cm for the Newfoundland and Spain data, respectively.
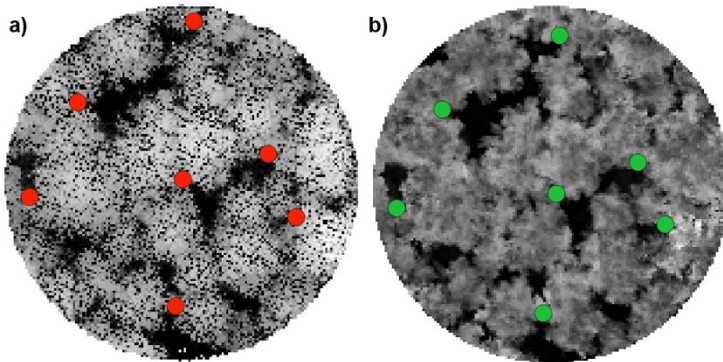
Figure 4.5. *Selection of homologous points in canopy height surfaces (CHS) computed from (a) ALS and (b) TLS datasets prior to 2D affine matrix transformation. Homologous points are represented in red and green color for ALS and TLS, respectively.*

4.2.6. Estimating voxel sampling variables and the rate of pulse reduction

The 3D space of the point clouds was discretized in voxels to produce vertical profiles (Procedure 3a) following section 2.3.3. The point density from the TLS point clouds was sufficiently high to allow the adoption of very small voxels (e.g., ~5 cm). However, assessing the capacity of each laser scanning configuration to detect understory vegetation required adopting a common voxel size: a trade-off between the fine vertical features of vegetation density and the availability of sufficient hits from laser scanning signal within a voxel. We therefore adopted a voxel size for all three datasets according to (*i*) the pulse

spacing of ALS in XY plane, (*ii*) the temporal sample spacing of ALS_{FW} in the Z axis, and (*iii*), avoiding empty voxels in either datasets. The most suitable voxel size was determined to be 0.5 m in X, 0.5 m in Y and 0.15 m in Z (vertical).

We first computed for all datasets the number of beams crossing the voxel (Nt), the number of hits within the voxel (Nh), and the number of beams blocked prior to reaching the voxel (Nb) (Procedure 3b) (see Figure 4.6). The computation approach of these three sampling variables differed between the TLS and ALS datasets due to their distinct sensor-signal-scene configurations. The approach taken for the TLS data used the one (first) hit per pulse of the TLS phase-shift technology. The approach taken for the ALS data assumed that all recorded hits were associated with an independent laser pulse having no cross-section (i.e., a vector with no divergence). This, however, is an abstraction. In fact, multiple hits originate from the same beam. Hence, the sampling variables derived from TLS data are not directly comparable with those derived from ALS data. We therefore assumed the number of hits in each voxel to be a representation of the vertical forest distribution only from ALS for both Nh^{ALSD} and Nh^{ALSFW} (Procedure 3c) while the vertical forest distribution from TLS was represented by the cumulative PAD^{TLS} (Procedure 3d). In addition to Nt, Nh and Nb, the path length of all pulses crossing the voxel was estimated from the TLS data. The three voxel sampling variables and the path length of all pulses crossing the voxel allowed calculating PAD^{TLS}, in $m^2 \cdot m^{-3}$, for each voxel according to mathematical framework proposed by Pimont et al. (2018). We used a minimum of five pulses reaching a voxel (Nt-Nb $\geq$ 5) as a threshold for calculating PAD^{TLS}, otherwise the voxel was assigned as being occluded. A negligible number of TLS voxels were tagged as occluded due to the large voxel size relative to point density.

In order to quantify signal occlusion caused by vegetation, we computed the rate of pulse reduction as the proportion of beams blocked prior to reaching the current voxel (Nb/Nt) (Procedure 3e). Knowing the rate of pulse reduction provides insight on the potential or drawbacks of the different laser scanning configurations.
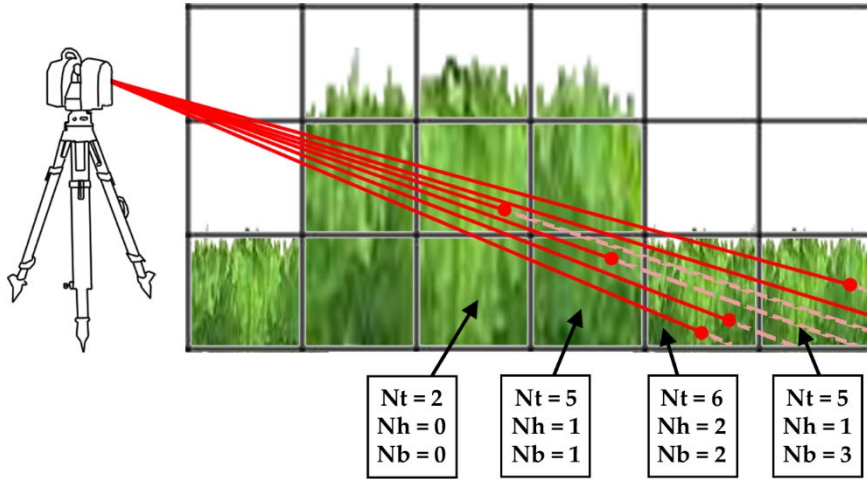
Figure 4.6. *Depiction of number of theoretical beams crossing a given voxel (Nt), number of hits (Nh), and number of hits blocked prior to reaching a given voxel (Nb) for the TLS configuration. Red dots represent the hits, red lines the laser beams prior to be blocked, and dashed pink lines the theoretical laser beams after being blocked.*

4.2.7.   Extracting vertical forest distribution and rate of pulse reduction profiles

In section 4.2.5, through the 2nd procedure, TLS data were co-registered only in XY (not in Z) to the ALSD data. Absolute heights of the canopy needed to be maintained in order to define the original laser pulse trajectories, which was critical in computing NhALSD, NhALSFW, PADTLS, and the rate of pulse reduction. However, in order to extract and make meaningful plot-level comparisons of the vertical profiles of these attributes (stored as 3D matrices of voxels), a co-registration in the Z-axis was necessary (Procedure 4). Co-registration ensured that the base of each column of voxels was set to a common Z reference system where all ground voxels were set to a height of 0 m. Both the ALS and TLS datasets had their respective DTMs created with a cell size of 0.5 m. The DTMs for ALS data were generated from ALSD data. Classification of ground points was done using the Axelsson algorithm (Axelsson, 2000) implemented in LAStools (Isenburg, 2017). The DTMs for the TLS data were produced for every plot using an open source ground classification algorithm in Computree

(Piboule et al., 2015). Height normalization of the 3D matrices was therefore done with their respective DTM; e.g., Nh^{ALSD} and Nh^{ALSFW} were normalized to its respective ALS-derived DTM as PAD^{TLS} was normalized to its respective TLS-derived DTM.

Next, we extracted vertical profiles of these 3D matrices to represent the vertical distribution of forest elements for each horizontal layer of voxels (i.e., a vertical bin with a height of 0.15 m). The value of each vertical bin was calculated as the sum of the voxel values of the corresponding horizontal layer for Nh^{ALSD}, Nh^{ALSFW} and PAD^{TLS} (Procedure 5a). The rate of pulse reduction was calculated as the average of the voxel values for that horizontal layer (Procedure 5b). Only information 0.5 m above the ground was considered therefore removing values associated to an understory zone strongly influenced by soil micro-relief and very low vegetation. This procedure provided the normalized vertical profiles of Nh^{ALSD}, Nh^{ALSFW} and PAD^{TLS}, and the rate of pulse reduction.

In order to assess how well we captured the vertical distribution profiles at different heights from ALS, we assessed the relationships between Nh^{ALS} and PAD^{TLS} by means of ratios. To do so, we computed, plotted and compared the ratios between Nh^{ALS} and PAD^{TLS} (i.e., Nh^{ALSD}/PAD^{TLS} and Nh^{ALSFW}/PAD^{TLS}). The resulting vertical profiles highlighted limitations in detecting the different vertical strata based on a unitless indicator. Higher values imply a higher detection of the vegetation, while lower values imply a more limited detection.

4.2.8. Classifying the vertical distribution of understory vegetation from ALS data

Once the vertical profiles were generated for Nh^{ALS}, both ALS configurations (ALS_D and ALS_{FW}) were compared to detect the understory vegetation. The height range for this comparison was set between 0.5 m and 4 m to detect the lower strata through Nh^{ALS} vertical profiles. The upper limit of 4 m deemed appropriate to capture high shrubs within our study sites. Detection of the understory vegetation was addressed through the characteristics of the Nh^{ALS} vertical profiles, whose curvature depends on the presence of understory

vegetation. To quantify the curvature of the Nh^{ALS} vertical profiles, we combined the fitting of the Lorenz curve (Lorenz, 1905) with the Gini index (Gini, 1912) (Procedure 6). The Gini index is a measure of statistical dispersion initially created to measure inequality of countries' wealth. It is computed as the area between the curve and the equality line (i.e., 1:1 line) (see area A in Figure 4.7a) divided by the area below the equality line and delimited by the main axes (see area B in Figure 4.7). Figure 4.7b and c show the two extreme cases, i.e., complete equality (i.e., Gini index = 0) and complete inequality (i.e., Gini index = 1), respectively. Consequently, the Gini index quantifies the curvature of a distribution, or in our application, a vertical profile. The Lorenz curve and Gini index have been widely used in economics, but also in some forestry applications. For instance, Valbuena et al. (2013) and Valbuena et al. (2014) proposed several indicators describing tree size inequality related to vertical forest structure. These indicators were based on the combined analysis of the Lorenz curve from ALS_D data, including the Gini index. In addition, the Gini index obtained from ALS_D was proposed to identify differences in structural complexity of forests (Valbuena et al., 2016).
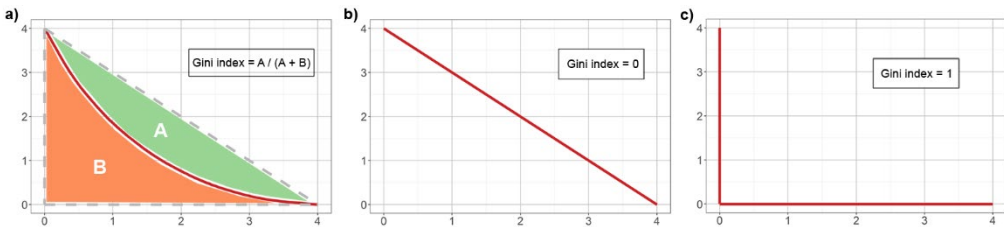
Figure 4.7. *Description of the Gini index for (a) a general case, and two examples showing (b) equality (i.e., Gini index = 0) and (c) inequality cases (i.e., Gini index = 1), respectively.*

In our study, each Nh^{ALS} vertical profile (Nh^{ALSD} and Nh^{ALSFW}) is represented by a Lorenz curve. The Gini index was estimated from this Lorenz curve through the ratio A / (A + B) (Figure 4.7a and Figure 4.8b). To do so, we applied the following three steps:

1. Nh^{ALS} vertical profiles were first filtered to exclude overstory following the same procedure for ALS_{FW} and ALS_D. We discarded the Nh^{ALS} values in the vertical profiles that had a height greater than the

first relative minimum of Nh^{ALS} (see Figure 4.8b). Consequently, high shrubs were included in the analysis but lower crowns were excluded from the computation of the Gini index (see Figure 4.8a).

2. Nh^{ALS} values were then normalized between 0 and 1 using Equation 17 (see Figure 4.8c). The normalization facilitated the comparison between plots since Nh^{ALS} values are variable according to the different plots and acquisition configurations.

$$Nh^{ALS}{}_{norm} = \frac{Nh^{ALS} - \min(Nh^{ALS})}{\max(Nh^{ALS}) - \min(Nh^{ALS})}$$

Equation 17

where $Nh^{ALS}{}_{norm}$ refers to the normalization (between 0 and 1) of Nh^{ALS}. Also, $min(Nh^{ALS})$ and $max(Nh^{ALS})$ are the minimum and maximum values, respectively, of Nh^{ALS} for the current plot between 0.5 and 4 m.

3. Finally, the Gini index was computed as the area between the curve and the equality line divided by the area below the equality line and delimited by the main axes (see Figure 4.8c).

The PAD estimation of the lower strata from the TLS data is a far less affected by signal occlusion because of the position of the scanner. Therefore the PAD^{TLS} vertical profiles was used as a reference to compare the ability of ALS_{FW} and ALS_D to detect understory vegetation. Nh^{ALS} vertical profiles were compared with PAD^{TLS} by calculating the correlation coefficient at lower strata ($0.5\ m \leq z \leq 4\ m$) and its significance using the Student's t-test (Gosset, 1908). This coefficient is a unitless quantity, and therefore allows comparing different datasets with different units and orders or magnitude, such as Nh^{ALSFW}, Nh^{ALSD} and PAD^{TLS} vertical profiles.
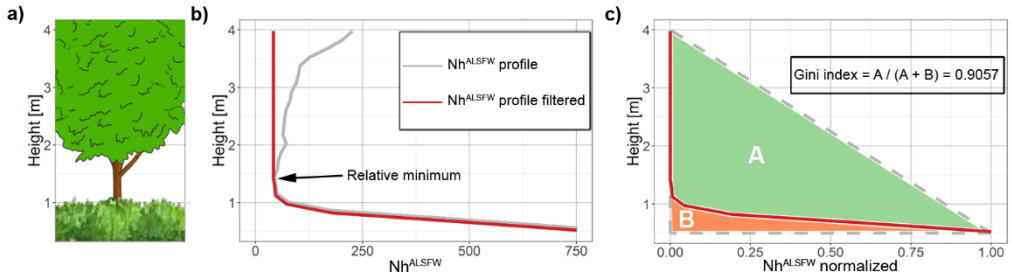
Figure 4.8. *Depiction of (a) separation between understory vegetation and overstory, (b) the filtering of NhALSFW vertical profiles corresponding to the lower strata from plot P3-NF and NhALSFW vertical profile, and (c) estimation of the Gini index from the resulting NhALS vertical profile.*

4.3. Results

4.3.1. Forest vertical distribution and rate of pulse reduction profiles

The data processing steps led to a representation of the vertical distribution and rate of pulse reduction profiles from the three laser scanning configurations, shown for a sample of plots from the Newfoundland and Spain sites in Figure 4.9 and Figure 4.10, respectively. Overall, results show that ALS, viewing the forest from the top-down, was more limited to sample the lower strata, while TLS, viewing the forest from bottom-up, was more limited to sample the top of the canopy. Despite these limitations, the vertical distribution profiles generally represented the forest's vertical structure, capturing components of the different vertical strata. For instance, plot P5-SP (Figure 4.10b) has a mixed presence of maritime pines and cork oaks. This heterogeneous vertical structure was represented by different peaks in associated PAD^{TLS} vertical profiles. Conversely, plots with a homogeneous vertical structure (e.g., plot P7-SP, Figure 4.10c) were represented by a single and well-defined peak in their associated vertical distribution profiles. A visual comparison between vertical distribution profiles (i.e., PAD^{TLS}, Nh^{ALSD} and Nh^{ALSFW}) and associated rate of pulse reduction with the one meter wide point cloud transects reveals an obvious correlation: both vertical profiles of element
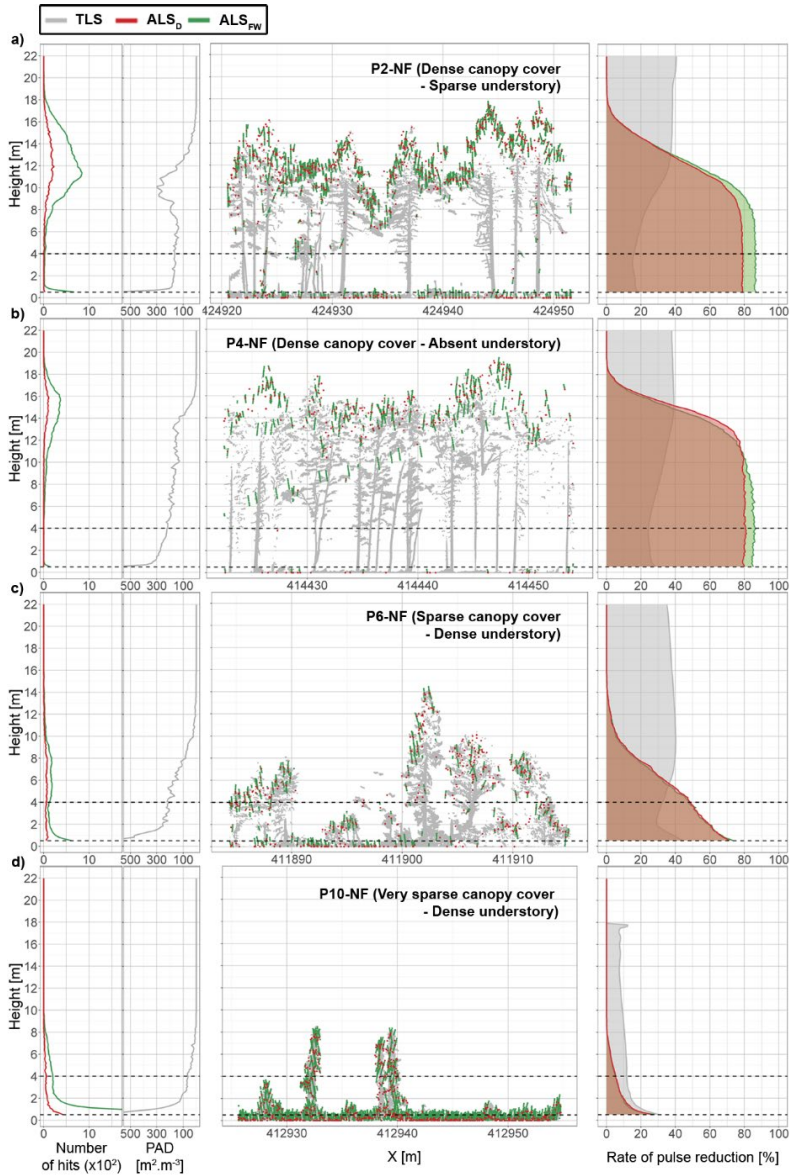
Figure 4.9. *Vertical profiles representing four plots of the Newfoundland site (a-d). The three figures from left to right represent: (i) the number of hits from ALS and cumulative Plant Area Density from TLS, (ii) a point cloud transect of one meter wide, and (iii) the rate of pulse reduction from the three configurations (i.e., TLS, ALS_D and ALS_{FW}). Dashed lines represent the limits of the lower strata (i.e., 0.5 and 4 m).*
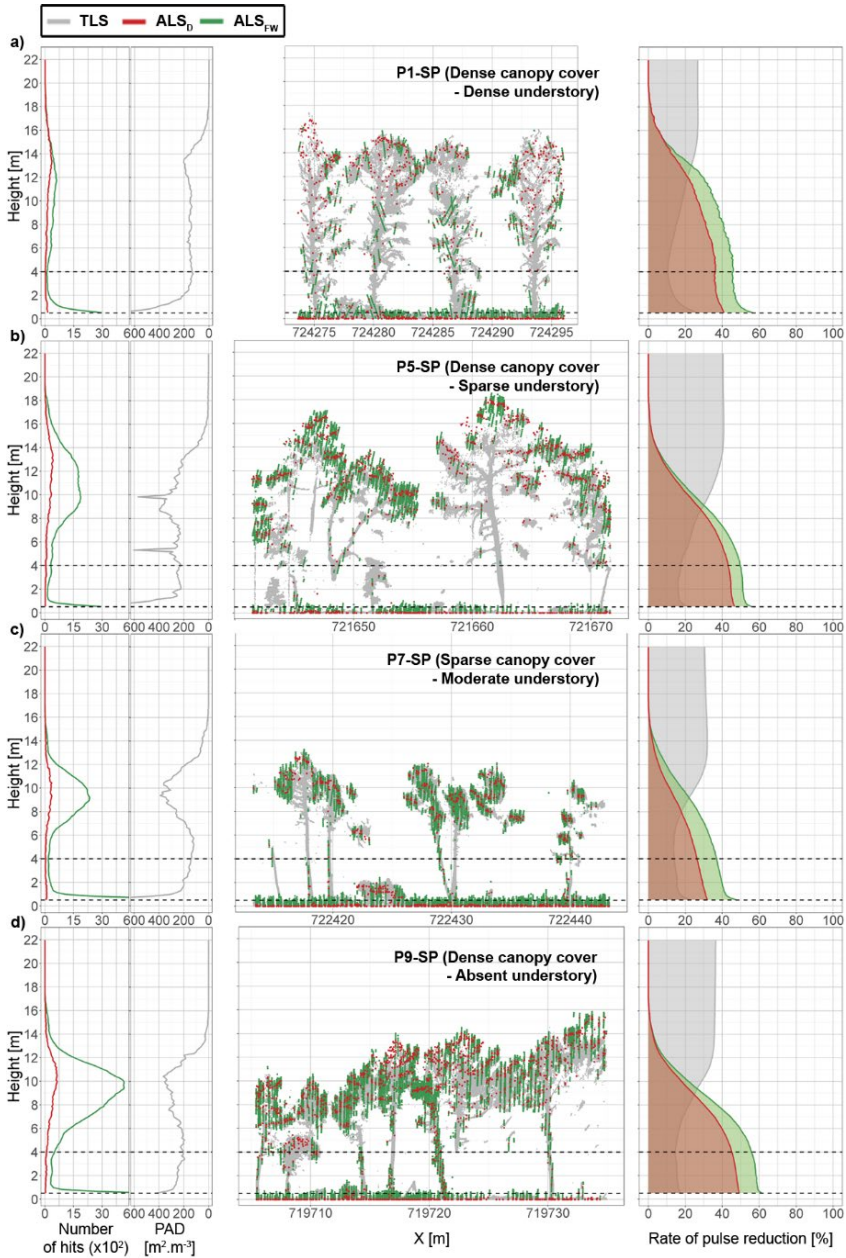
Figure 4.10. *Vertical profiles representing four plots of the Spain site (a-d). The figures from left to right represent: (i) the number of hits from ALS and cumulative Plant Area Density from TLS, (ii) a point cloud transect of one meter wide, and (iii) the rate of pulse reduction from the three configurations (i.e., TLS, ALSD and ALSFW). Dashed lines represent the limits of the lower strata (i.e., 0.5 and 4 m).*

distribution and point cloud density decreased as the rate of pulse reduction increased. The rate of pulse reduction profiles generally followed a distribution in the form of a sigmoid function or "S"-shaped curve, whose form, or increment of slope of the rate of pulse reduction, depended on the laser scanning configuration as well as the density of the upper canopy. For example, all plots with dense canopy cover (e.g., Figure 4.9a and b, Figure 4.10a, b and d) had rate of pulse reduction profiles following a clearly defined sigmoid distribution. However, converse trends were observed in the rate of pulse reduction from ALS and TLS: increasing rates of pulse reduction were associated with decreasing heights from ALS and increasing heights from TLS. Furthermore, in some instances, observed high values in the rate of pulse reduction were associated with very low values from the vertical distribution profiles (i.e., where PAD^{TLS} and/or Nh^{ALS} reached or approached 0). This occurred more frequently for Newfoundland plots where conifer species on this site grow dense, creating a dense upper canopy (e.g., P2-NF and P4-NF in Figure 4.9a and b, respectively). For these plots, the rate of pulse reduction profiles transition to high values (i.e., 35-40% for TLS; 80-90% for ALS) at heights of ~10-13 m. Conversely, for plots with sparsely distributed vegetation, the rate of pulse reduction was very low. In conditions of sparse vegetation, the rate of pulse reduction curve followed the typical sigmoid distribution, however did not reach the high values observed from denser plots (e.g., P7-SP in Figure 4.10c). Sparse and very sparse sites often displayed a relatively flat vertical line, terminating near ground by a steep high value (e.g., P6-NF and P10-NF in Figure 4.9c and d). Overall, but more specifically in dense canopy covers, an offset in the x-values from the midstory strata of PAD^{TLS} was observed relative to Nh^{ALS}. This offset coincides with the large discrepancy in the number of hits returned by ALS and TLS systems from the midstory strata, where ALS did not detect tree stems to the same degree as TLS. On the other hand, little variability was observed in Nh^{ALSD} vertical profiles in the lower strata, except in some cases where canopy cover was very sparse (e.g., P10-NF). The analysis of vertical distribution profiles from the lower strata generally showed lower Nh^{ALSD} when compared to Nh^{ALSFW}. In addition, similarity in overall shape, quantified and reported in section 4.3.2 by means of the coefficient of correlation, was greater

between Nh^{ALSFW} and PAD^{TLS} than between Nh^{ALSD} and PAD^{TLS}. Furthermore, Nh^{ALSFW} value has an exponential increment as they approached the ground when understory vegetation was absent. This trend was less obvious in plot P4-NF and most Newfoundland plots with higher rates of pulse reduction (> 80%).

Considering TLS as reference, the ratios between Nh^{ALS} and PAD^{TLS} in Figure 4.11 illustrate the limitations of ALS configurations in detecting the vertical distribution profiles at different heights based on a unitless indicator. Ratios were calculated as Nh^{ALS} divided by PAD^{TLS}, therefore implying higher detection of vegetation as this ratio value increases. The highest values for the ratio calculated from Nh^{ALSD} were observed in the upper strata for all plots (Figure 4.11). In most cases, ALS_D incoming pulses were blocked by the dominant strata, generating signal occlusion beneath. ALS_{FW} (Nh^{ALSFW}/PAD^{TLS}) and ALS_D (Nh^{ALSD}/PAD^{TLS}) ratio values were most similar in the upper strata. Nevertheless, ALS_{FW} ratio values below the dominant strata (i.e., intermediate and/or lower strata) remained high, while ALS_D ratio value dropped. Generally, ALS_D ratios dropped below 1 m, and in some cases below 2.5 m (e.g., P10-NF in Figure 4.11).

4.3.2. Understory characterization from ALS

In the previous subsection, it was observed that although the values of the vertical distribution profiles may differ between Nh^{ALS} and PAD^{TLS}, they remained similar in terms of shape, albeit with different units and order or magnitude. This similarity in terms of shape of the vertical distribution profiles was quantified between Nh^{ALS} and PAD^{TLS} by using the coefficient of correlation (see Table 4.2), which ultimately allows for determining whether understory vegetative material was detected. Coefficients of correlation were calculated between Nh^{ALS} (i.e., Nh^{ALSFW} and Nh^{ALSD}) and PAD^{TLS} vertical profiles from the lower strata. Null coefficient of correlation values (e.g., observed from ALS_D for plots P4-NF, P5-NF and P6-SP, and for plot P6-SP from ALS_{FW}) were due to the fact that no values were registered at the lower strata, and therefore the standard deviations of the corresponding vertical profiles were equal to zero. Coefficients of correlation between Nh^{ALSFW} and PAD^{TLS} were rarely below 90%, and all were considered to be significant correlations according to the Student's
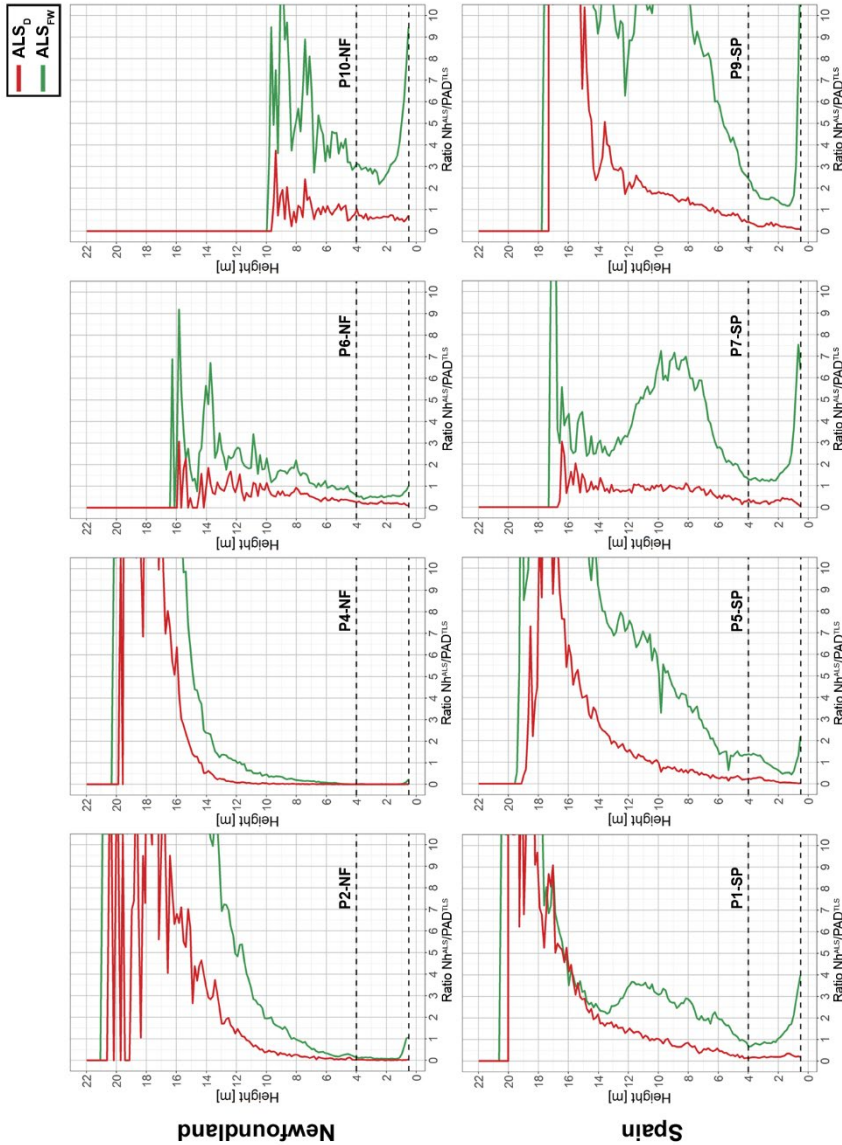
Figure 4.11. *Vertical profiles representing the ratio (Nh^{ALS}/PAD^{TLS}) between the Nh from ALS (i.e., ALS_{FW} and ALS_D) and cumulative PAD from TLS for a sample of plots. Dashed lines represent the limits of the lower strata (i.e., 0.5 and 4 m).*

t-test (Table 4.2). The range of coefficients of correlation for ALS_{FW} was between 53.48% and 99.58%, with an average value of 90.11% and an associated standard error of 3.04%. On the other hand, with the exception of plot P9-NF, all coefficients of correlation between Nh^{ALSD} and PAD^{TLS} were all lower or at par, with one correlation not being considered as significant (plot P7-SP). The range of coefficients of correlation for ALS_D was between 30.60% and 97.36%, with an average value of 82.57% and an associated standard error of 4.29%. Regarding summary statistics of the differences in coefficients of correlation between ALS_{FW} and ALS_D, the range was between 22.01% and -68.36% (being negative values when coefficients of correlation for ALS_{FW} were greater), with an average value of -9.50% and an associated standard error of 4.62%. Hence, although ALS_{FW} detected understory vegetation with a much larger number of hits than ALS_D (see Figure 4.9, Figure 4.10 and Figure 4.11), the latter still had a significant correlation with PAD^{TLS}. Remarkably, strong correlations between Nh^{ALSFW} and PAD^{TLS} were observed for dense canopy cover plots from the Newfoundland site, where the rate of pulse reduction was large in lower strata. For instance, plots P2-NF and P4-NF had rates of pulse reduction ~84% for ALS_{FW} at the lower strata and an associated coefficient of correlation with PAD^{TLS} equal to 98.50% and 93.39%, respectively. Nh^{ALSFW} and Nh^{ALSD} were equally correlated with PAD^{TLS} (i.e., < 1% of difference) in only a few plots having a dense canopy cover (plots P8-SP and P10-SP) and in a plot with very sparse canopy cover (plot P10-NF).

4.3.3. Understory vegetation density classification

Afterwards, the variation of the Nh^{ALS} vertical profiles at the lower strata was quantified by means of the Gini index, whose values for ALS_D and ALS_{FW} are presented in Table 4.3. Generally, sparse understory vegetation densities had large Gini indices (i.e., gradual increments of Nh^{ALS}), while dense understory vegetation had low Gini indices (i.e., steep increments of Nh^{ALS}). Despite considerable differences between the structure of boreal and Mediterranean forests, the Gini index values confirmed a coherent behavior for both sites as a vegetation density indicator. Specific Gini index ranges derived from the Nh^{ALSFW} vertical profiles were associated to understory vegetation density classes as follows: absent (91.63% ± 0.13), sparse (90.59% ± 2.23),

Table 4.2. *Coefficient of correlation values between* Nh^{ALS} *(i.e., Nh^{ALSFW} and Nh^{ALSD}) and PAD^{TLS} as reference at the lower strata.*

Site	Plot ID	Correlation Nh^{ALSFW}-PAD^{TLS} (%)	Correlation Nh^{ALSD}-PAD^{TLS} (%)
Newfoundland	P1-NF	63.25	56.21
	P2-NF	98.50	97.36
	P3-NF	92.95	88.09
	P4-NF	93.39	NA
	P5-NF	NA	NA
	P6-NF	95.66	67.40
	P7-NF	98.09	87.29
	P8-NF	96.06	90.46
	P9-NF	69.77	91.78
	P10-NF	96.64	96.70
Spain	P1-SP	96.89	92.31
	P2-SP	98.73	67.89
	P3-SP	99.58	91.13
	P4-SP	98.78	96.97
	P5-SP	94.96	84.02
	P6-SP	53.48	NA
	P7-SP	98.96	30.60*
	P8-SP	89.89	90.46
	P9-SP	84.37	82.16
	P10-SP	92.21	92.93

* Correlation deemed not significant from Student's t-test with a confidence level of 95%.

moderate (84.31% ± 0.00), and dense (75.45% ± 7.86). Similarly, from the Nh^{ALSD} vertical profiles, Gini index ranges were associated to understory vegetation density classes as follows: absent (97.69% ± 4.01), sparse (83.79% ± 12.31), moderate (53.38% ± 0.00), and dense (59.37% ± 16.71). Nevertheless, the Gini index class interval thresholds computed from ALS_D were fuzzier, implying more overlap between classes, than those from ALS_{FW}. Class intervals derived from computed Gini index values showed larger standard deviations for ALS_D than ALS_{FW}. Furthermore, misclassification between sparse and absent understory vegetation density classes occurred when derived with ALS_D data. Some plots with a moderate or sparse understory had a Gini index from Nh^{ALSD} lower than plots with a dense understory vegetation (e.g., P7-SP vs. P3-SP), which led to a misclassification. Similarly, plot P5-NF, with sparse understory vegetation, had a Gini index from Nh^{ALSFW} larger than that observed for plots with absent understory, which also led to a misclassification.

Table 4.3. *Gini index from Nh^{ALSFW} and Nh^{ALSD} vertical profiles for each plot from the Newfoundland and Spain sites. Plots are in ascending order according to the Gini index computed from Nh^{ALSFW}.*

Plot ID	Understory vegetation	Canopy cover	Gini index Nh^{ALSFW} (%)	Gini index Nh^{ALSD} (%)
P9-NF	Dense	Very Sparse	59.63	33.47
P6-NF	Dense	Sparse	65.10	45.97
P1-SP	Dense	Dense	72.65	55.04
P2-SP	Dense	Very Sparse	74.98	51.21
P7-NF	Dense	Very Sparse	75.74	89.94
P4-SP	Dense	Dense	78.88	60.13
P10-NF	Dense	Very Sparse	80.09	66.87
P10-SP	Dense	Dense	80.12	69.82
P8-NF	Dense	Sparse	83.03	45.25
P3-SP	Dense	Dense	84.27	76.04
P7-SP	Moderate	Sparse	84.31	53.38
P8-SP	Sparse	Dense	86.83	64.14
P2-NF	Sparse	Dense	90.55	79.33
P3-NF	Sparse	Dense	90.57	90.71
P5-SP	Sparse	Dense	90.61	77.81
P1-NF	Sparse	Dense	91.14	94.91
P4-NF	Absent	Dense	91.48	100.00
P9-SP	Absent	Dense	91.67	93.06
P6-SP	Absent	Dense	91.74	100.00
P5-NF	Sparse	Dense	93.81	95.83

4.4. Discussion

In this chapter, we assessed the ability of different laser scanning configurations to estimate vertical forest structure, linking it with a new method to estimate signal occlusion in the different strata. In addition, we also assessed and compared the suitability of ALS$_{FW}$ and ALS$_{D}$ to classify in understory vegetation density classes. Key results highlighted the limitations inherent to different configurations in estimating vertical forest structure and the importance of signal occlusion. More specifically, in the lower strata, which is highly occluded by ALS configurations, understory vegetation density was successfully assessed through vertical canopy density profiles. Moreover, the analysis of vertical profiles from our testing plots demonstrated that ALS$_{FW}$ improved understory identification and density determination over ALS$_{D}$.

Overall, our results confirmed the general trend largely accepted by the scientific community, which implies that laser scanning signal occlusion prevails in sectors blocked by dense canopy covers. Signal occlusion therefore depends largely on the laser scanning configuration: ALS, viewing the forest from the top-down, is more limited to sample the lower strata, while TLS, viewing the forest from bottom-up, is more limited to sample the top of the canopy (Hilker et al., 2010; Anderson et al., 2016). Regarding ALS configurations, canopy density of the upper layer is the single most important environmental factor in defining if sufficient airborne laser pulses reach the complete vertical range of the forest. Hence, signal occlusion can limit exhaustive sampling of the lower vertical strata with ALS data. In this sense, Maltamo et al. (2014) distinguished between signal occlusion and canopy cover as two different but related phenomena, both affecting the overestimation of CBH when using ALS. Conversely, for plots with sparsely distributed vegetation, the level of signal occlusion is very low. Consequently, estimation of the distribution of vegetative material is possible throughout the vertical range of the forest. In these cases, vertical distribution of forest materials can be estimated with high level of accuracy. LaRue et al. (2020) also observed that estimating canopy density of the lower strata is best achieved in open canopy covers for ALS configurations, because of the influence of signal occlusion caused by dense canopy covers.

In cases where significant signal occlusion exists, ALS configurations detection of lower strata density capabilities are limited. The reliability of vertical distribution profiles is directly dependent on the level of signal occlusion. In this regard, we proposed to use the rate of pulse reduction as an indicator of the amount of signal occlusion occurring at various heights in the forest. Inferring the ability to detect vertical distribution from our reference dataset (PAD^{TLS}) and the plotted values of rate of pulse reduction profiles led us to propose a threshold of the rate of pulse reduction from which the estimation of the distribution of vegetative material is no longer possible. This threshold is variable and related to the density of the canopy cover, which remains plot-specific. We noticed that Nh^{ALS} values between the height with a rate of pulse

reduction above ~80-90% and the ground do not provide reliable estimates of vegetation density. A similar principle, reversed vertically, applies to TLS datasets where PAD^{TLS} between the height with a rate of pulse reduction above ~35-40% and the top of the canopy. The slope of the rate of pulse reduction curve, which follows a sigmoid function, depends on the laser scanning configuration as well as the density of the upper canopy. Currently we suggest an approximate threshold for the rate of pulse reduction. The ability to define a more specific threshold may be tied to parameters of this sigma curve as a discriminating indicator of the overall signal occlusion and defining the vertical area where vegetation density can be estimated.

ALS_{FW} and ALS_D data come from the same signal, however, further processing steps before obtaining the final product make them different. Using the complete (i.e., ALS_{FW}) or the discretized (i.e., ALS_D) signal in the detection of the top of the canopy is not significantly different to assess vegetation density. Nonetheless, ALS_{FW} provided a definite advantage to detect vegetation density for intermediate and lower strata than ALS_D. Lower strata are generally occluded due to overstory blocking incoming laser pulses. Consequently, Nh^{ALSD} vertical profiles do not show much features in the lower strata, except in some cases where signal occlusion caused by overstory is low (e.g., very sparse canopy cover). Nevertheless, when using PAD^{TLS} as reference, Nh^{ALS} (i.e., Nh^{ALSFW} and Nh^{ALSD}) vertical profiles correlations are considered as significant, albeit Nh^{ALSFW} is more correlated. Although curve correlation is more accurate, and the number of hits much larger at the lower strata with ALS_{FW}, our results confirmed that understory vegetation was captured by ALS_D, albeit to a lesser extent and in plots with a high rate of pulse reduction (i.e., ~85%). Other studies have also found difficulties associated with ALS_D to detect the internal forest structure (Chasmer et al., 2006; Hilker et al., 2010; Hilker et al., 2012). The higher potential of ALS_{FW} when compared with ALS_D to detect and determine understory vegetation density classes was also found in several studies (Hancock et al., 2017; Torralba et al., 2018). Our results confirm the potential of both ALS configurations to detect non-occluded strata (i.e., top of the canopy) and demonstrated the increased capability of ALS_{FW} to detect strata with signal occlusion (i.e., intermediate and lower strata). Although ALS_D may be used to estimate understory vegetation at a plot-level, the signal is generally weak in

dense canopies and hence limits the estimation of density in the lower strata at such fine spatial scales.

Vertical profiles with Nh^{ALSFW} display a systematic artefact near the ground in the form of an exponential increment, even if understory vegetation is absent. This increment depends mainly on two factors: the large number of hits from the ground and the hits from the understory vegetation. The histogram of hits from the ground usually follows a Gaussian curve for which the upper side can be merged by the hit from understory, if present. When understory is present, generally variation of the Nh^{ALSFW} value relates to the understory vegetation density classes, whereas Nh^{ALSD} are not responsive, except for open canopies with a rate of pulse reduction below ~50%. Regardless, the number of hits from the ground dominates to the point of masking the understory signal in most situations. Hence, although Nh^{ALSFW} vertical profiles increase exponentially as they approach the ground for all the understory vegetation scenarios, variation of Nh^{ALSFW} increment can be used to identify and determine understory vegetation density classes.

We demonstrated the Gini index to be a useful and accurate indicator to determine understory vegetation density classes from either Nh^{ALS} vertical profiles. Our results demonstrated that understory vegetation density classes can be identified and further grouped by processing the ALS data in both dense and porous forests. Despite strong signal occlusion at lower heights in the forest, ALS and especially ALS_{FW}, the Gini index identifies understory densities. Additionally, Gini index thresholds established for understory vegetation densities coincide for both sites: boreal and Mediterranean. Thus, the understory vegetation density classes are represented by the following Gini value ranges from Nh^{ALSFW}: below a value of ~85% for dense understory vegetation density, between ~85% and ~90% for sparse, and above ~90% for absent. Fortunately, and contrary to ALS_D, misclassification from ALS_{FW} is not occurring between sparse and other understory vegetation density classes. In a related study, Valbuena et al. (2012) discriminated forest structural types by using an application of the Lorenz curve and the Gini index based on the basal area and the number of trees. Apart from a Gini index of 0% and 100%

representing the complete equality and inequality, respectively, they also found that a value of 50% was relevant. This value represents a uniform distribution of the basal area of the trees. However, with our datasets, Gini index values from Nh^{ALSFW} were all above 56%, since the Lorenz curves start increasing below the height considered as upper limit of the lower strata (i.e., 4 m), and therefore Gini index values are higher. Other indicators such as L-Skewness (Valbuena et al., 2017) and Shannon Index (Almeida et al., 2019) are complementary to Gini index. L-Skewness allows for quantifying the asymmetry of the Lorenz curve. This facilitates estimating mean height and absence of understory vegetation. Additionally, the Shannon index represents the diversity of the dataset using a variable as reference (e.g., species). Therefore, it would be feasible to use height thresholds as a variable to determine understory vegetation density classes. Nonetheless, when signal occlusion caused by overstory is important, the ability to use the Shannon or Gini indices is strongly compromised. It is therefore critical to estimate the level of signal occlusion by means of the rate of pulse reduction prior to carry out the analysis.

We proposed a simple way to identify the understory vegetation layer and exclude the overstory. The procedures most frequently adopted in the literature apply a threshold at 2 m height assuming that it covers the understory vegetation. This procedure is non-discriminant, not plot-specific, and therefore it may exclude shrubs or include lower crowns. Instead, we propose considering the vertical distribution of gaps in the density profiles to identify a local minimum separating overstory from understory vegetation. This results in different height values delimiting understory vegetation from overstory, which for our dataset varied between 0.525 and 3.975 m. The accuracy of this procedure depends on vegetation homogeneity. This vertical gap assessment was done at a plot-level, but it can also be applied at a finer scale (e.g., at voxel-column) if a minimum hit density is reached for all heights to avoid false gaps. The procedure we proposed is plot-specific and allows for an automatic height division of overstory and understory vegetation layers.

Some limitations in the application of the developed methodological procedure are noteworthy. Currently, PAD estimates are limited to being derived from TLS data, as the estimation of PAD from ALS configurations is

currently not possible. The unbiased estimation of PAD from the mathematical framework proposed by Pimont et al. (2018) decreases significantly the influence of signal occlusion for a reliable representation of vertical profiles. Therefore, it is a useful reference to represent distribution of vertical structure. PAD estimation from all the configurations would have allowed for a comparison in the detection of the vertical distribution profiles. Unfortunately, no methods were currently available to estimate PAD from multiple returns ALS data. Nevertheless, the number of hits in the ALS data in the non occluded areas can be used as a unitless indicator (i.e., coefficient of correlation) comparable with PAD^{TLS}. Furthermore, we note that the Gini index was successful in determining understory vegetation density classes from both airborne and terrestrial laser scanning data, but cannot be implemented with emerging full-waveform spaceborne laser scanning data such as GEDI (Dubayah et al., 2020). GEDI emits four laser beams with a large footprint (~19-25 m) that do not overlap. The method proposed in this chapter assumes that many laser beams cross a same voxel, and then the number of hits is used as a driving variable to calculate the Gini index. Nevertheless, other metrics based on return amplitude, and proposed in section 3.2 in Table 3.1, may be more suited to characterize understory vegetation on large areas with GEDI.

Knowing the limitations of laser scanning configurations is fundamental to use laser scanning point clouds for the estimation of forest canopy structure. The most severe limitation is caused by signal occlusion by vegetative elements, which can be quantified with the rate of pulse reduction. Despite high level of signal occlusion in the lower strata for ALS configurations, the proposed method allows estimating presence and density of understory vegetation in both dense and porous canopies of boreal and Mediterranean forests through the Gini index applied to ALS_{FW} data. This new indicator becomes one of the few options to characterize understory vegetation for ALS configurations, which has many implications for forest ecology and wildfire mitigation.

4.5.   Conclusions

In this chapter, the limitations and potentials of airborne and terrestrial laser scanning configurations to estimate the vertical forest structure have been assessed. We conclude that understory vegetation density classes can successfully be determined more accurately with ALS_{FW} than with ALS_D. More specifically, three key points stand out from this chapter. Firstly, the rate of pulse reduction profiles were demonstrated to be a good indicator to quantify occlusion along the vertical profile. This information can be used to determine the reliability of vegetation density estimates from different laser scanning configurations for specific vertical strata. Secondly, both ALS configurations (ALS_D and ALS_{FW}) showed their capability to detect understory vegetation, albeit significantly more accurately with ALS_{FW} due to the greater number of hits registered in lower strata. This considerably lower number of hits registered from ALS_D in the lower strata suggests that a forest plot would be the finest spatial scale (i.e., minimal mapping unit) for which understory vegetation can be successfully detected from ALS_D, and hence, our methods could be applied to. Finer scales would inevitably lack sufficient registered hits in understory vegetation for accurate understory characterization. Finally, and thirdly, we demonstrated the use of the Gini index as a way to determine understory vegetation density classes from both ALS configurations, again, more accurately with ALS_{FW}. Understory vegetation density classes (absent, sparse, moderate and dense) were defined through thresholds applied to the index for both ALS_{FW} and ALS_D. Computing the rate of pulse reduction and Gini index we characterized the vertical structure and understory vegetation of these structurally differing forests. The applications for which this contribution may be relevant are several, such as characterizing wildlife habitats, assessing timber productivity and improving silvicultural decision-making in support of wildfire mitigation. Further research is needed to better understand the relationships between estimates of PAD and vertical profiles of number of hits for ALS configurations, vertical profiles of rate of pulse reduction and classification of forest types, and the use of the Gini indicator to estimate presence and density of understory vegetation.

Characterization of understory vegetation using full-waveform airborne laser scanning and voxel-based metrics

Cistus albidus

"Que repiquen las campanas,
en las nupcias de la jara,
de tierno azahar se engalana,
su cándida vida, su trémula cara."

Nico Pozón Requejo

This chapter analyzes the characterization of understory vegetation by using ALS$_{FW}$ data in a Mediterranean forest. As described in chapter four, the use of these data allows for the extraction of detailed information in different vertical strata compared to ALS$_D$, since the complete signal emitted by the sensor is registered. The characterization of understory vegetation over large areas by means of ALS$_{FW}$ is a key factor to map and better describe the vertical structure and wildlife habitats, as well as for the estimation of ladder fuels, which are determinant in fire behavior in some ecosystems, such as the Mediterranean.

5.1.   Introduction

Understory vegetation is an essential component of forest ecosystems (Suchar and Crookston, 2010). 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 ladder fuels, 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 ALS, provide horizontal and vertical information of different canopy layers (Ruiz et al., 2018). Several studies have estimated characteristics of understory vegetation cover using ALS$_D$ (Table 5.1). Most of these studies utilize 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 pulses·m^{-2} and 1 pulse·m^{-2}, respectively. The overall accuracy and kappa value of the classification were 0.77 and 0.53, respectively. Morsdorf 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 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.

Although ALS$_{FW}$ provides a full representation of the intercepted forest structure, since it registers the complete signal emitted by the sensor, few studies have demonstrated its capability to characterize understory vegetation (Table 5.1). Hancock et al. (2017) characterized voxelized understory cover in 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 CHP retrieved from large

Table 5.1. *Summary of existing studies about the characterization of understory using ALS with overstory presence (adapted from Crespo-Peremarch et al. (2018b)).*

Study	Study Area	Ecosystem	Definition of forest types	Target attributes	Data	Density (points·m⁻²)	No. of plots (plot size m²)	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 pulses·m⁻²	132 (400)	Overall accuracy = 0.77 kappa = 0.53 (Classification)

Table 5.1. (cont.) *Summary of existing studies about the characterization of understory using ALS with overstory presence (adapted from Crespo-Peremarch et al. (2018b)).*

Study	Study Area	Ecosystem	Definition of forest types	Target attributes	Data	Density (points·m⁻²)	No. of plots (plot size m²)	Results
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 = 1.5 m)	nRMSE = 24 % (Verification at voxel-level)

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 forest 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 laser scanning and field data with a mean R^2 of 0.75. Lastly, Anderson et al. (2016) found that in an urban woodland landscape, CH estimated by ALS$_D$ was more biased, and intensity less accurate, than that provided by ALS$_{FW}$.

Compared to ALS, 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) and produce forest inventory information commensurate with field observations, registering data for > 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 chapter, in section 5.2, the understory characteristics of the study area and the datasets used are described. The different processing steps followed to

characterize understory vegetation using ALS$_{FW}$ and TLS as reference are enumerated in section 5.3. Results from the different tests to characterize understory vegetation are shown in section 5.4. In section 5.5, the discussion of the results is addressed. Lastly, in section 5.6, the conclusions of this chapter are presented. This chapter sets out the Hypothesis 6: "The new described ALS$_{FW}$ metrics in objective 2 may be used to accurately characterize the height, cover and volume of understory vegetation in a Mediterranean ecosystem". To demonstrate this hypothesis, Objective 6 is addressed in this chapter: apply and validate the new metrics described in objective 2 derived using a voxel based approach to estimate understory height, cover, and volume in a Mediterranean forest ecosystem, proposing some practical recommendations for further development and testing ALS$_{FW}$ metrics.

5.2. Study area and data

The study area is located in eastern Spain, in a Mediterranean forest in the Natural Park of Sierra de Espadán (see Figure 3.9 and Figure 4.1). This study area coincides with the one described in section 3.4.1, whose characteristics of understory vegetation are further described in section 4.2.1.

TLS data used in this chapter were previously described in section 4.2.3. Among the experimental plots registered by TLS, a structurally representative sample of 21 plots was selected (Figure 5.1) by maximizing the variability of dominant species and understory vegetation cover. Moreover, during TLS data acquisition, the maximum height of the understory was also assessed at each plot 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 5.3.4.

ALS$_{FW}$ data used in this chapter were previously described in section 3.4.1, and also further used in Chapter 4 as the second study area.
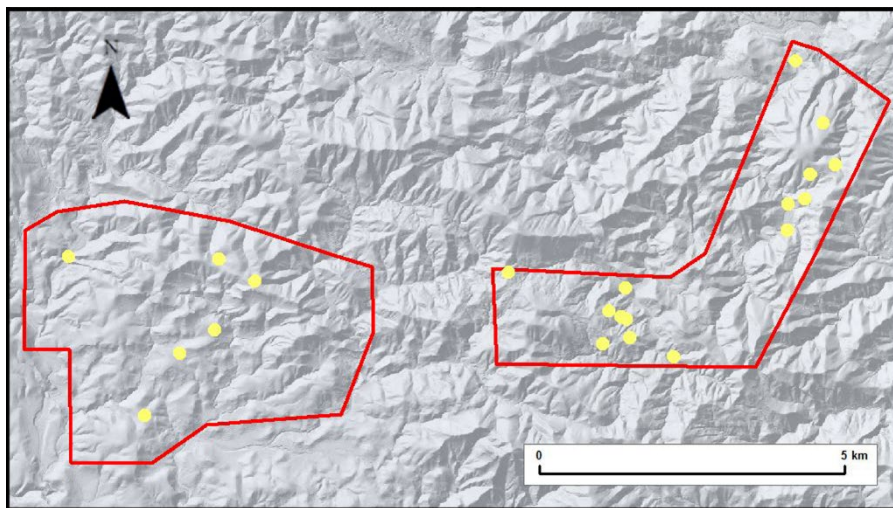
Figure 5.1. *Plot locations (in yellow) in the study area.*

5.3. Methods

In this chapter, we assess the capacity of ALS_FW to characterize understory vegetation using voxel-based metrics in a Mediterranean forest. The overall strategy followed in this chapter is illustrated by the flow diagram of Figure 5.2. As the first procedure (1), ALS_FW and TLS data must be co-registered to be in the same coordinate system in XY plane. In the second procedure (2a and 2b), heights from ALS_FW and TLS data are respectively normalized to work with the same Z origin and to avoid the effect of relief in the characterization of the understory strata. The third procedure (3) involves denoising of ALS_FW data to remove noise present in the signal registered. As the fourth procedure (4), the TLS point cloud is filtered to retrieve the understory strata, which is used as reference data in further procedures. The fifth procedure (5a and 5b) consists of the voxelization of ALS_FW and TLS data for the extraction of the corresponding metrics and attributes, respectively. As the sixth procedure (6), ALS_FW metrics are extracted from each pseudo-vertical waveform (i.e., voxel column), and then metrics are computed at cell- (6a) and plot-level (6b). The seventh procedure (7) consists of the extraction of understory attributes from the TLS voxelization used further to assess the characterization of understory vegetation from ALS_FW

data at plot- (7a) and cell-level (7b). Finally, metrics from ALS$_{FW}$ as independent variables and understory attributes from TLS as dependent variables are used in regression models at plot- (8a) and cell-level (8b). In addition, other plot characteristics such as slope and dominant species are used in mixed-effect models (9) to characterize understory vegetation.
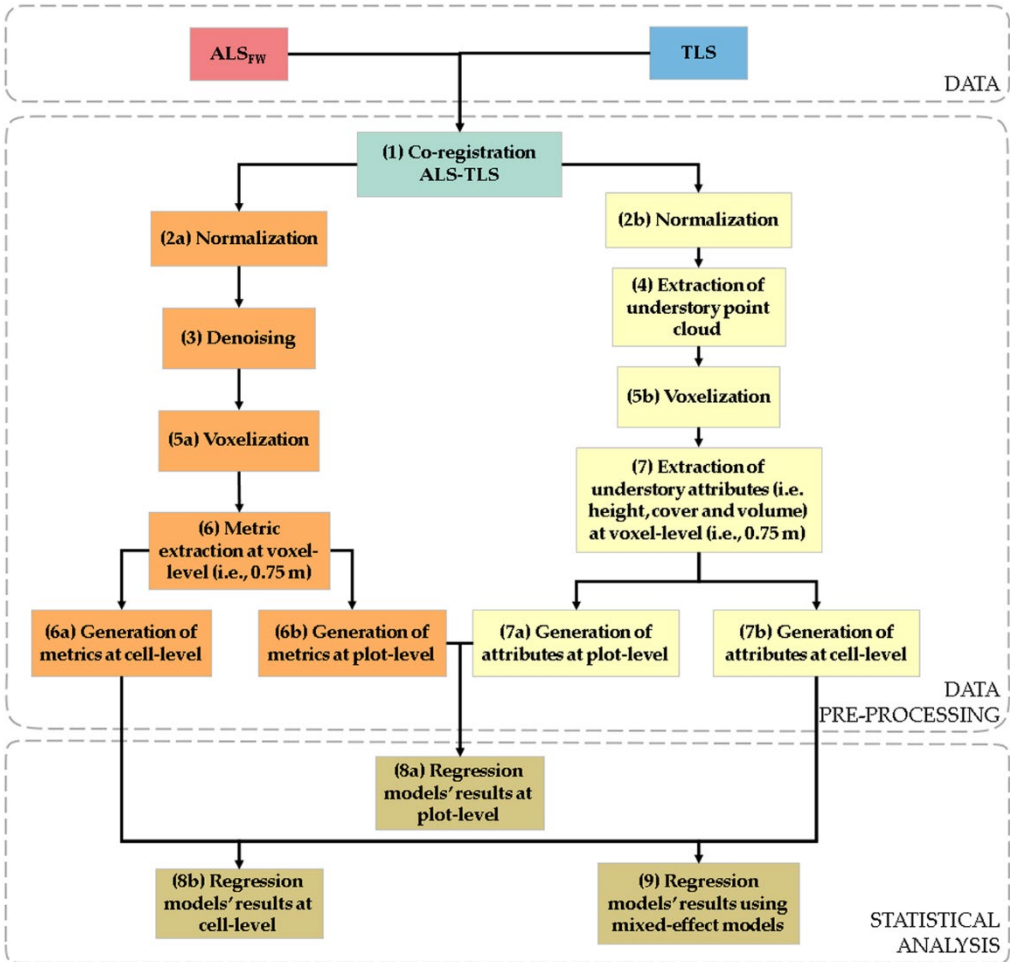
Figure 5.2. *Overall processing flowchart implemented to characterize understory vegetation from ALS$_{FW}$ and TLS data.*

5.3.1. XY co-registration of ALS$_{FW}$ and TLS datasets

Co-registration is a key step when dealing with different datasets from different sources. This co-registration in XY plane was done in this chapter by following the same procedure previously described in section 4.2.5.2. In this case, ALS data were also taken as the reference dataset, since its accuracy is higher than the one for TLS data.

5.3.2. Height normalization

Heights of the ALS and TLS datasets were normalized using DTMs derived from each of the point clouds (Procedures 2a and 2b, respectively) for a new co-registration in Z plane, being 0 m the reference surface. In the case of ALS, classified ground points were provided by the vendor. These classified points were used to generate the corresponding DTM by means of LAStools (Isenburg, 2017)(version 171017). TLS ground points were classified using a variation of the Axelsson (2000) algorithm implemented in LAStools. DTMs with a resolution of 0.3 m were generated and each dataset was then normalized. Height normalization of TLS point cloud was done by using LAStools, while normalization of ALS$_{FW}$ was done as described in Chapter 3, as a step of the ALS$_{FW}$ metric extraction implemented in *WoLFeX*.

5.3.3. Denoising

This procedure (Procedure 3) was the same followed and previously described in section 2.3.1.

5.3.4. Extraction of understory point cloud

TLS-based attributes characterizing the understory require two additional pre-processing steps (Procedure 4). 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 it was 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 (Figure 5.3).
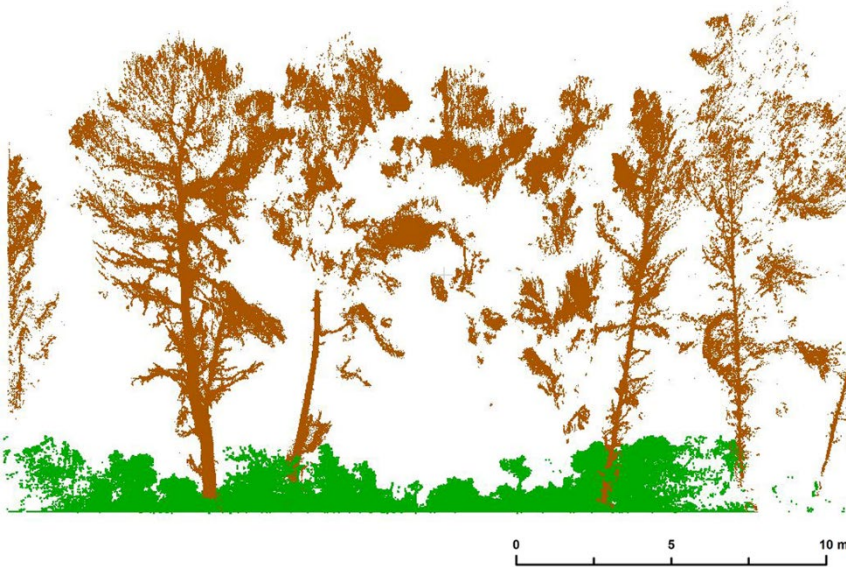
Figure 5.3. *Vertical transect showing the extraction from TLS point cloud of understory vegetation (in green) from overstory (in brown).*

5.3.5. Voxelization

This procedure (Procedures 5a and 5b) was the same followed and described in section 2.3.3. In this chapter, the horizontal size of the voxels was based on the results of Chapter 2 (i.e., 0.75 m), while the vertical dimension was calculated from the temporal sample spacing by using Equation 5 (i.e., resulting in 0.15 m). As assignation value, the maximum was the one selected. The goal of selecting this voxel size was to have the lowest number of empty voxels without a loss of accuracy.

5.3.6. Extraction of ALS_{FW} metrics

The extraction of ALS_{FW} metrics (Procedure 6) was done following the procedure previously described in section 2.3.4, where pseudo-vertical

waveforms are generated from voxelization and then metrics are extracted. Hence, ALS$_{FW}$ metrics were firstly extracted from the selected voxel size, i.e., $0.75 \times 0.75 \times 0.15$ m (henceforth referred to as 0.75 m). Then, the understory was characterized at two spatial scales; $3.75 \times 3.75 \times 0.15$ m (i.e., 5×5 columns of voxels; henceforth referred to as 3.75 m), which is denoted as "cell-level" of understory vegetation (Procedure 5a), as well as at the broader plot-level scale (15 m radius) (Procedure 5b).

The total number of ALS$_{FW}$ metrics extracted was 53, as described in section 3.2 and Table 3.1. In addition, to better understand if limiting the calculation of the pseudo-vertical waveform metrics to lower components of the canopy enhances estimations of understory vegetation, a height filter was applied to ALS$_{FW}$ metrics. This height filter consisted of cutting off the pseudo-vertical waveform at a given height threshold, then extracting ALS$_{FW}$ metrics from the resulting pseudo-vertical waveform. The height threshold for the whole study was computed as 99% height of understory heights extracted from TLS data. Therefore, all the ALS$_{FW}$ metrics were computed 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– (Procedure 6a) and plot-level (Procedure 6b) as variables for the regression models.

5.3.7. Extraction of understory attributes from TLS

Four key attributes 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 (V), which is defined as three-dimensional space occupied by understory (Procedure 7). These four understory attributes were used as the observed variables and modeled with ALS$_{FW}$ derived predictors.

To calculate the H_{max}, we computed the 99% height of each 0.75 m column of voxels and then extracted the maximum within each 3.75 m side cells (cell-

level) (Procedure 7b). H_{mean} was defined as the average of the 99% heights of each 0.75 m column of voxels 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 multiplied by the volume of the voxel was used as an estimate of V. A graphical description to compute these understory attributes at cell-level is showed in Figure 5.4. Moreover, Figure 5.5 shows these TLS attributes categorized by the dominant species and the slope of the plot. Dominant species had a high influence on all the understory attributes. Understory vegetation height, cover and volume were higher with presence of *Pinus halepensis*, and slightly lower with presence of *Pinus pinaster* compared to mixed *Pinus pinaster* and *Quercus suber*. These understory vegetation attributes were however less influenced by slope. In this case, median values were generally similar among the different slope categories, however, mean height and volume values were less dispersed as the slope increases.

In addition to the cell level (3.75 m) (Procedure 7b), all attributes were also calculated at plot-level (15 m) (Procedure 7a).

5.3.8. Linear regression models

Linear regression was used to obtain predictive models of the four understory attributes, using ALS_{FW} metrics as independent variables (Procedure 8). Attribute selection consisted of comparing the AIC of all possible model comparisons using a maximum of three ALS_{FW} metrics in each model. Each plot was composed of 40 samples (i.e., cells). In order to reduce spatial autocorrelation, 10 samples per plot were randomly sampled, which resulted in 210 samples at the cell-level (Procedure 8b) and 21 for the plot-level analysis (Procedure 8a). A total of 16 model sets were tested (4 understory TLS metrics × 2 resolutions (cell- and plot-level) × 2 sets of ALS_{FW} metrics (with and without the TLS height filter)). Models were compared using the R^2, RMSE, nRMSE and CV. In the case of C, which is a bounded variable between 0 and 1, we replaced linear regression with Beta regression (Ferrari and Cribari-Neto, 2004) where a
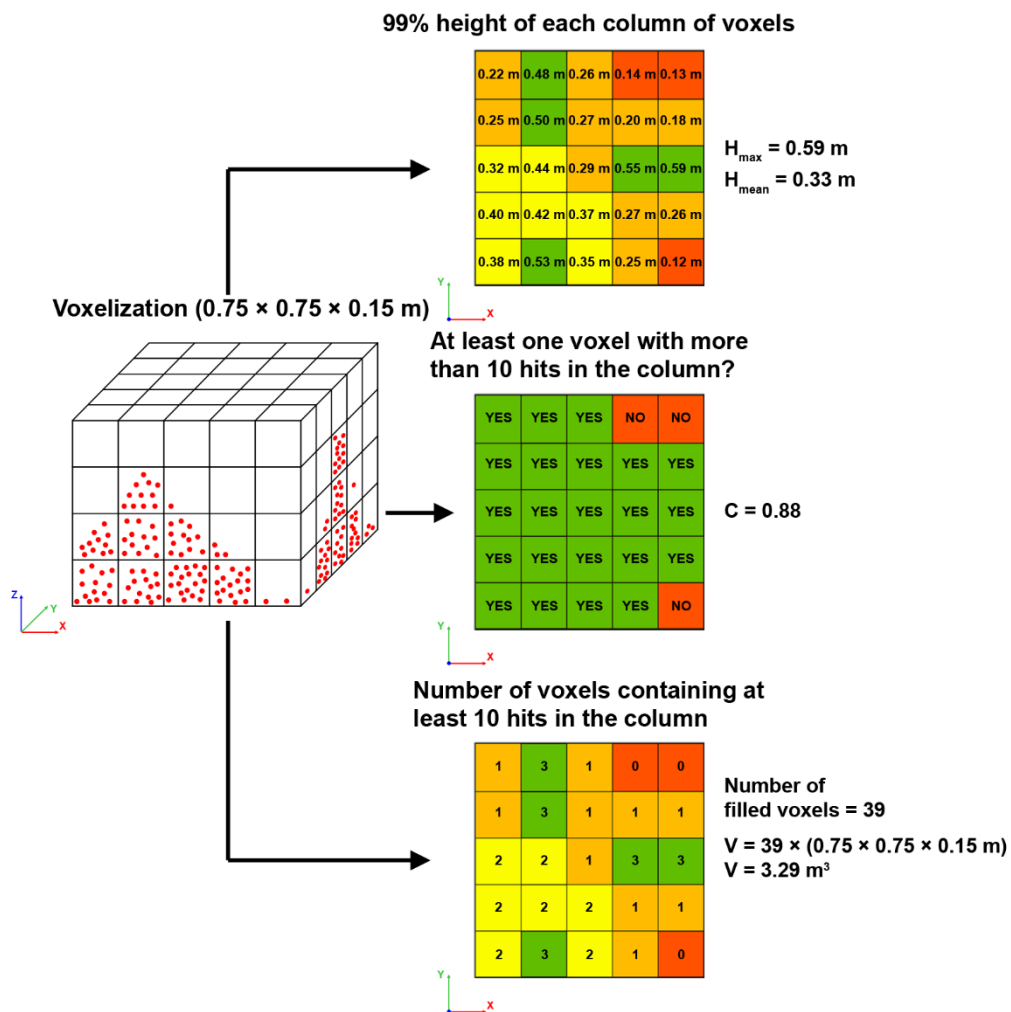
Figure 5.4. *Graphical description to compute the understory attributes (H_{max}, H_{mean}, C and V) at cell-level (i.e., 3.75 m) from a TLS point cloud.*

pseudo-coefficient of determination (pR^2) was generated for these regression models.
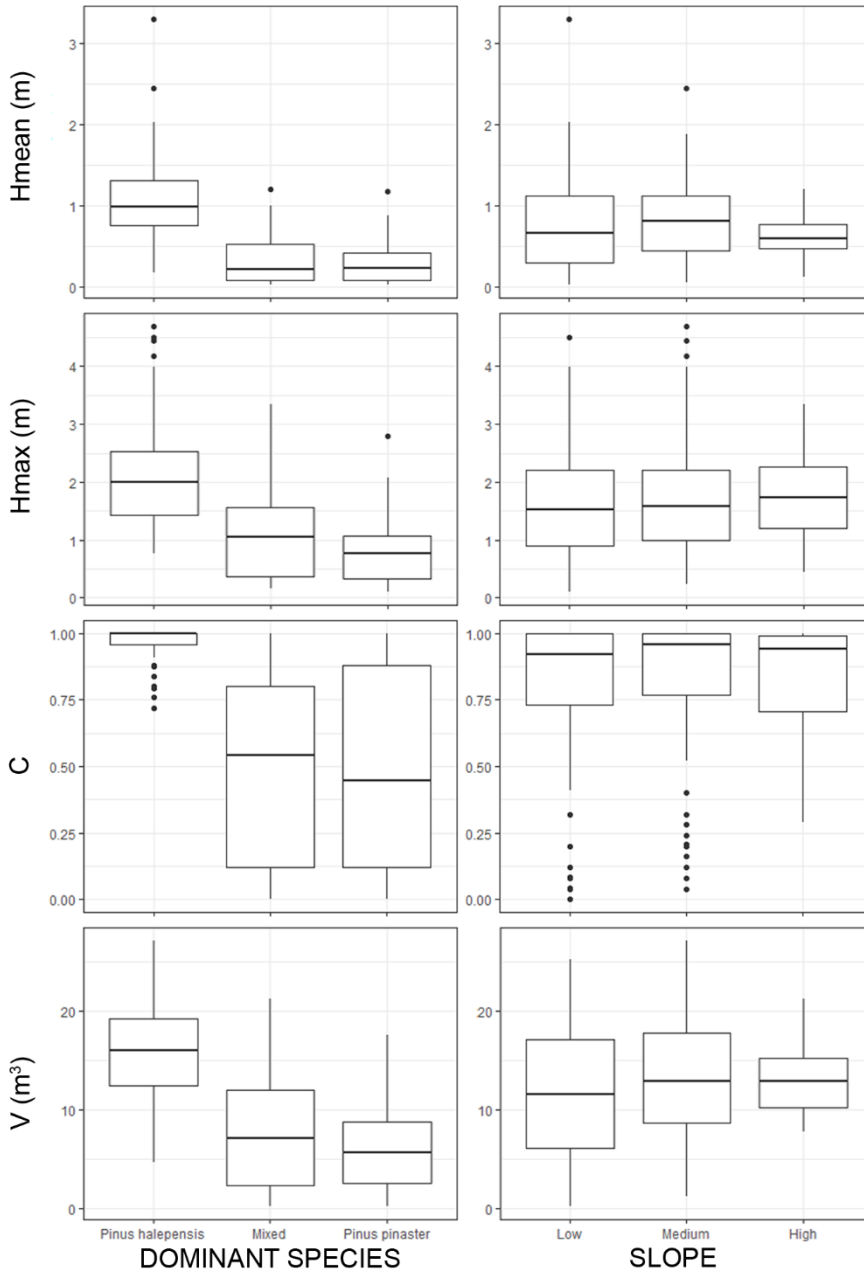
Figure 5.5. *Box and whiskers representing TLS understory attributes (mean height: H_mean, maximum height: H_max, cover: C, and volume: V) categorized by dominant species (Pinus halepensis, Mixed Pinus pinaster and Quercus suber, and Pinus pinaster) and slope (low, medium, and high) from the 21 plots.*

5.3.9. Linear mixed-effect models

To assess if the ability of ALS$_{FW}$ to predict the TLS attributes was site dependent, a mixed-effect modeling approach was also undertaken (Procedure 9), 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 Figure 5.5). The slope was categorized 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. Both the model slope and intercept were allowed to vary (based on Gelman and Hill (2007)) while utilizing Nakagawa and Schielzeth's (2013) steps with an update of Johnson (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 attributes × 2 ALS$_{FW}$ metric datasets (with and without height filter) × 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.

5.4. Results

The detection of 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.

Figure 5.6 shows an example of the four TLS and ALS$_{FW}$ derived attributes 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 P6-SP, P7-SP, and P2-SP, respectively).
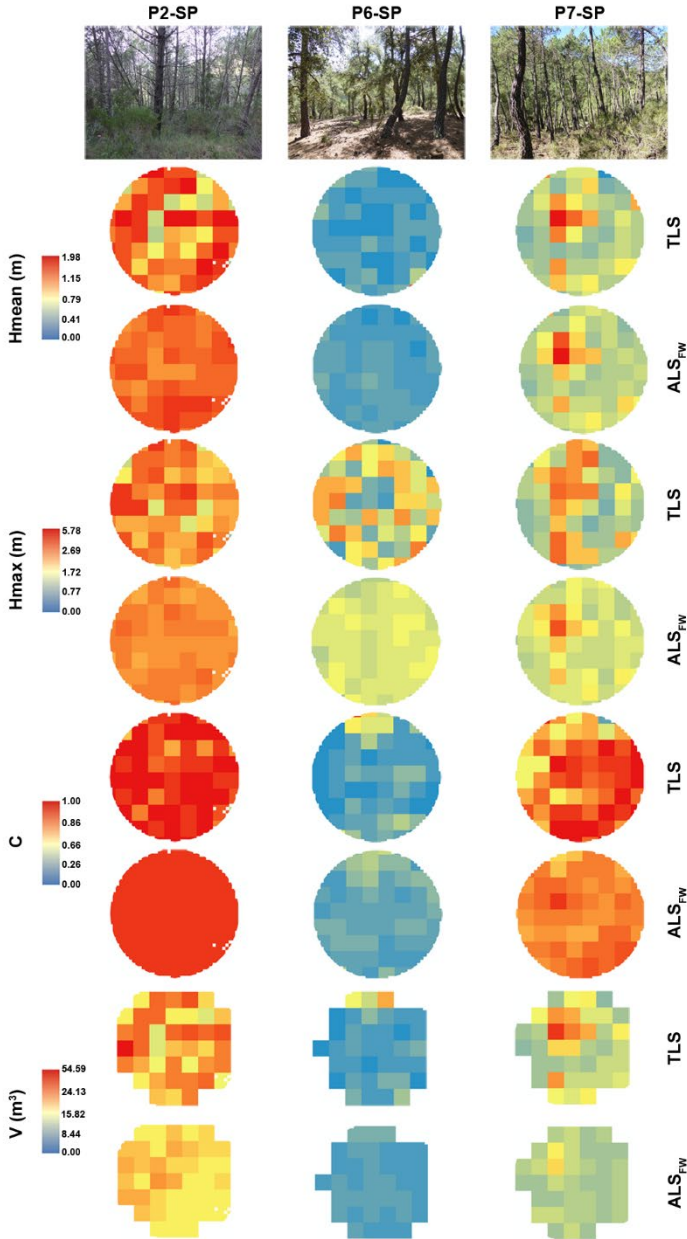
Figure 5.6. *TLS and ALS$_{FW}$ derived four attributes (H$_{mean}$, H$_{max}$, C and V) and field photographs extracted from three plots (P2-SP, P6-SP, and P7-SP) with 15 m radius within the study area. Plots P6-SP, P7-SP, and P2-SP, represent low, moderate, and high degrees of understory cover, respectively.*

Figure 5.7 shows a table with the ALS_{FW} metrics selected for the 16 regression models (4 understory TLS attributes × 2 resolutions (cell- and plot-level) × 2 set of ALS_{FW} 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 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.

Results of the mixed-effect models that incorporated different combinations of categorical variables (slope, dominant species, and both) are shown in Table 5.2. These results indicate that the highest accuracy was achieved for H_{mean}, with an 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 attributes, 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 (Figure 5.7), all understory attributes 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.

Figure 5.8 shows scatter plots of the TLS-based observed and ALS_{FW}-based predicted attributes at cell- and plot-level, as well as using the mixed-effect
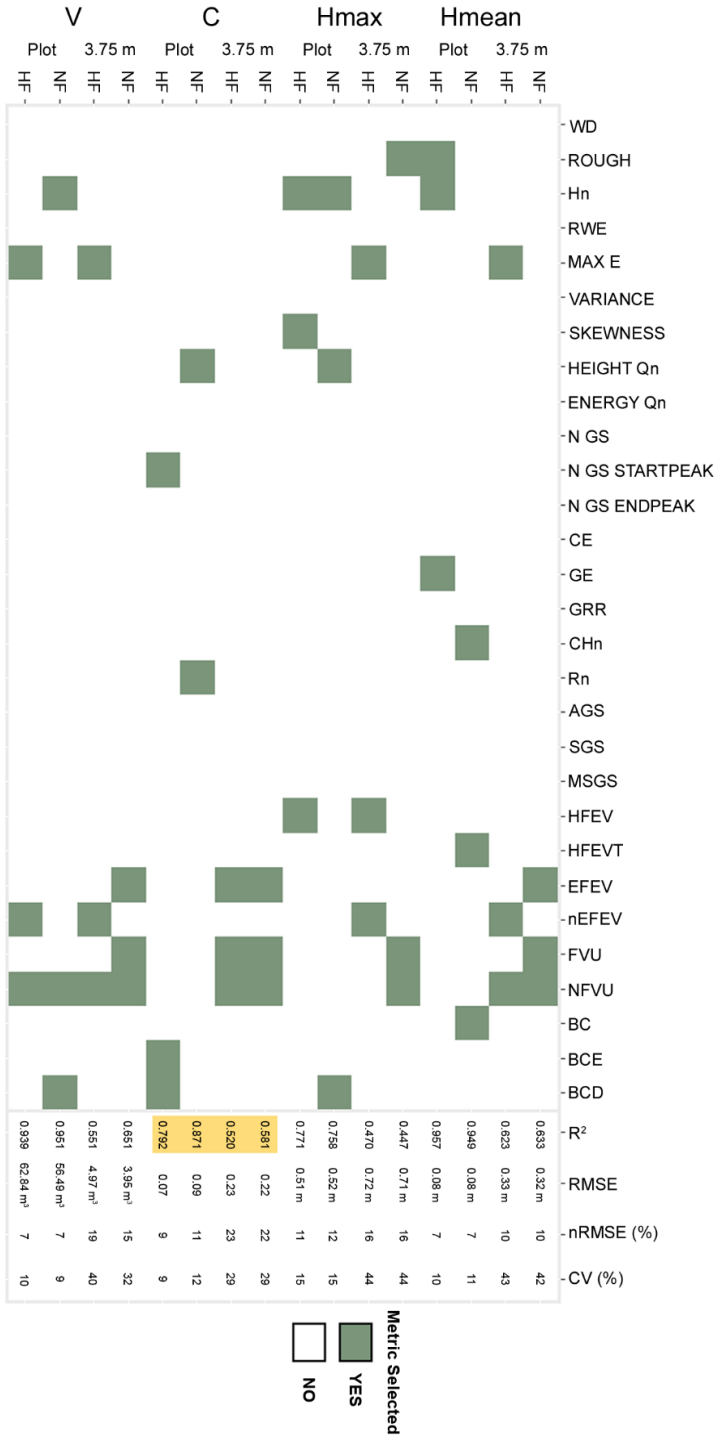
Figure 5.7. ALS$_{FW}$ metrics selected for the estimation of the different attributes (H_{mean}, H_{max}, C, and V) for cell- (3.75 m resolution) and plot-level (15 m radius) resolution, and for each height filter (NF: no filter, HF: height filter). The results from these regression models, as well as R^2 values and pseudo-R^2 (orange highlighted), are also included.

148

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.

Table 5.2. *Results of mixed-effect models for the estimation of the four understory attributes (H_{mean}, H_{max}, C, and V).*

Categorical variable	Attribute	Height filter	R²m	R²c	RMSE	nRMSE (%)	CV (%)
Slope	H_{mean}	NO	0.271	0.847	0.31 m	10	41
		YES	0.625	0.627	0.33 m	10	43
	H_{max}	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 m³	14	31
		YES	0.068	0.943	4.58 m³	17	37
Dominant species	H_{mean}	NO	0.394	0.666	0.30 m	9	40
		YES	0.526	0.606	0.31 m	10	41
	H_{max}	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 m³	13	29
		YES	0.110	0.898	4.49 m³	17	36
Slope + Dominant species	H_{mean}	NO	0.232	0.791	0.30 m	9	39
		YES	0.260	0.780	0.31 m	9	41
	H_{max}	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 m³	13	29
		YES	0.035	0.967	4.26 m³	16	34

5.5. Discussion

In this chapter, a new methodology to characterize understory vegetation from ALS_{FW} data has been described. This methodology was verified 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
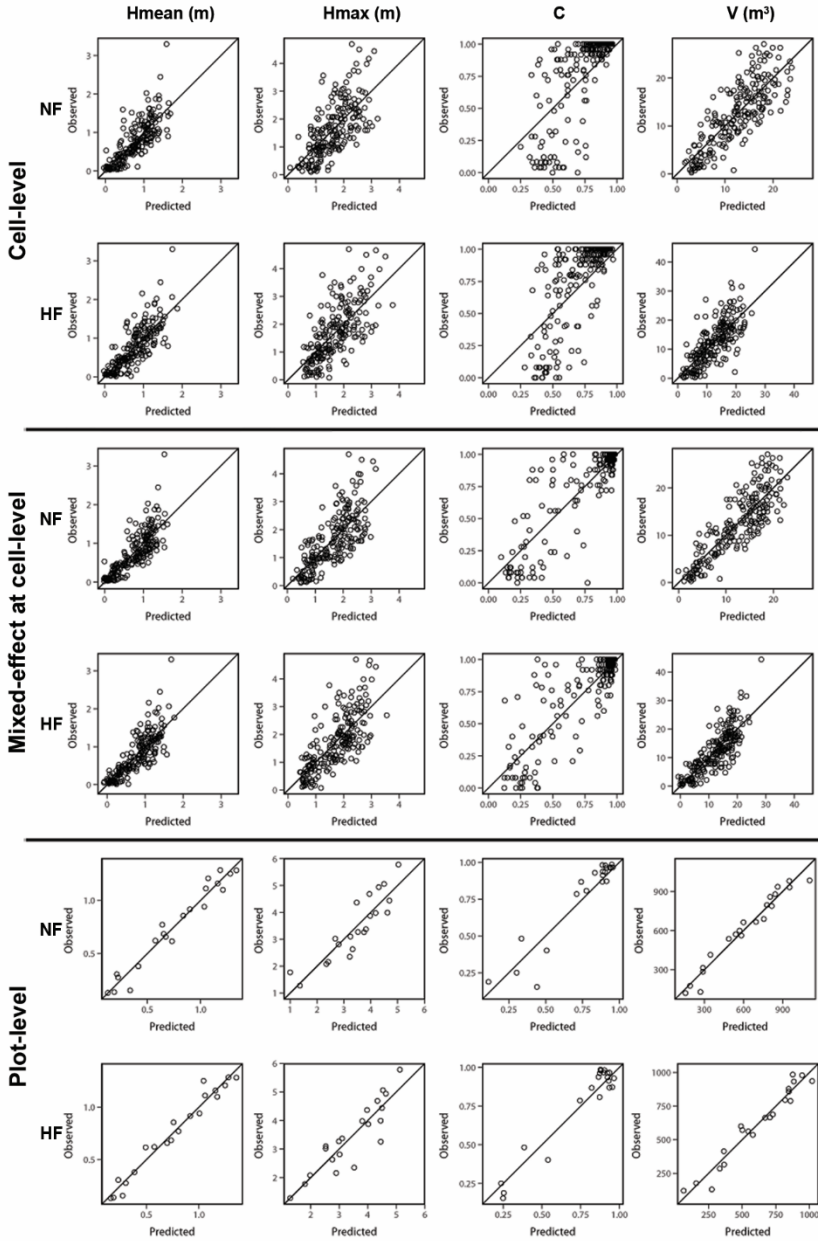
Figure 5.8. *Regression graphs for the estimation of the different attributes (H_{mean}, H_{max}, C and V) for each resolution (cell-level, mixed-effect (cell-level) and plot-level (15 m radius)) and for each height filter (NF: no filter, HF: height filter). Solid line represents the 1:1 line.*

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 attributes 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. A height filter was applied 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 hits, 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 (Hill and Broughton, 2009; Martinuzzi et al., 2009; Morsdorf et

al., 2010; Wing et al., 2012; Hancock et al., 2017). Most of the studies to date (Hill and Broughton, 2009; Martinuzzi et al., 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 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 applied 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 modeling point of view, the most selected attributes were those developed in this thesis, 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 increase 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 CBH 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., ladder fuel 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.

5.6. Conclusions

In this chapter, a method to characterize the understory vegetation through ALS$_{FW}$ data in a Mediterranean forest has been presented. 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 attributes such as height, cover, and volume over large areas. These attributes 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 in section 3.2 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 chapter 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 attributes such as biomass, which is essential to analyze forest fire intensity.

Chapter 6

Final conclusions

Olea europaea

"El olivo está triste,
languidece la florecilla blanca de su ramaje,
se desvanece hacia lechos de tierra
en su frondoso almagre."

Nieves Fernández Rodríguez

6.1.　Answers to the original research questions

This thesis addressed the development of ALS$_{FW}$ processing and analysis methods to characterize the vertical forest structure, in particular the understory vegetation.

ALS$_{FW}$ data have proven to be a powerful tool to characterize the entire vertical forest structure, since the complete signal emitted by the sensor and going through the different vertical strata is registered. The possibility of having all this information makes it possible for the user to analyze different vertical strata in more detail than with ALS$_D$, where data are reduced by the Gaussian decomposition technique. Focusing on the understory strata, its characterization is challenging for ALS configurations due to occlusion of overstory, and crucial for forest fire mitigation and mapping wildlife habitats. ALS$_{FW}$ has proven to detect and characterize with more detail these strata than ALS$_D$. Nevertheless, other aspects of ALS$_{FW}$ should be considered. The registration of the complete signal also involves a large amount of data that must be stored and managed, which implies availability of disk space and powerful processors. This drawback is disappearing, as more powerful computers with more storage capacity and cloud computing alternatives are being available. On the other hand, the creation and publication of new processing tools, such as *WoLFeX*, is expected to diminish the effect of the lack of ALS$_{FW}$ processing tools, which was a disadvantage until now.

The main conclusions for each of the specific objectives enumerated in section 1.2 are as follows:

Objective 1: To analyze the influence of pulse density, voxel parameters and regression methods on ALS$_{FW}$ metric values and on forest structure attributes estimates, identifying those parameters and quantifying their relations to be able to tune their values in order to considerably reduce this influence in practice.

- Values of ALS$_{FW}$ metrics extracted following the voxelization procedure are subject to pulse density. This means that ALS$_{FW}$ metric values vary if the pulse density changes. This entails a problem since pulse density is variable due to different factors such as objects'

occlusion, scan angle or flying speed. The variation of ALS$_{FW}$ metric values due to higher pulse density in side-lap areas is known as side-lap effect, and this is not the same for all the ALS$_{FW}$ metrics, therefore some are more influenced than others.

- Overall, this variation of ALS$_{FW}$ metric values follows a negative exponential distribution (Equation 6), being the variation lower as the pulse density increases. Therefore, variation of ALS$_{FW}$ metric values related to pulse density can be modeled using a negative exponential distribution to determine, through coefficient *b*, from which pulse density ALS$_{FW}$ metric values become stable. Hence, the coefficient *b* shows the minimum pulse density that should be used to avoid the side-lap effect.

- On the other hand, if a minimum pulse density cannot be set (e.g., ALS$_{FW}$ data have already been acquired), modifying voxel parameters (i.e., voxel size and assignation value) may be used to reduce the side-lap effect. Overall, an increment of the voxel size, and the modification of the assignation value for some metrics, results in reduction of the side-lap effect. This is due to the fact that the probability that larger voxels are crossed by at least one waveform is higher, avoiding the gaps in the voxel columns that may alter values of ALS$_{FW}$ metrics. Hence, a trade-off between increasing voxel size to reduce side-lap effect and a substantial loss of resolution should be considered.

- Regarding the different ALS$_{FW}$ metrics in more detail, the increment of the voxel size is recommended for HOME, WD, FS and RWE for the mean and median assignation values. Nevertheless, small voxel sizes make ROUGH and RWE for maximum, percentiles 90 and 95 more stable. On the other hand, the choice of the assignation value must be considered depending on the voxel size used for RWE. However, NP is sensitive to pulse density variations and it cannot be reduced through voxel parameters, and therefore should be avoided for further analyses.

- Forest attributes are estimated by using ALS$_{FW}$ metrics influenced by the side-lap effect; therefore, these attributes are also influenced by

the effect. Nevertheless, the influence of the side-lap effect on estimate of forest attributes is lower than on ALSFW metrics, since the former are statistically fitted, and errors minimized by using regression models.

- However, the increment of the voxel size also diminishes the side-lap effect on forest attributes. For instance, side-lap effect is lower in mass-related attributes when the voxel size is increased than in height-related attributes. In the latter, however, side-lap effect can be reduced more efficiently by changing the regression method used.

Objective 2: To compile a set of methods to process and analyze ALSFW data, including the relative radiometric correction of the data to reduce the effect of the different angles of incidence and local altitude variations during the data acquisition process, the extraction of most ALSFW metrics as proposed in the literature, as well as new metrics focused on understory vegetation, integrating them in a new software available to use for the scientific community.

- A new software tool named *WoLFeX* is presented and described. This software tool is designed to process ALSFW data, allowing for clipping, radiometrically correcting, voxelizing the original ALSFW waveforms, creating pseudo-vertical waveforms and extracting an exhaustive set of object-oriented metrics that can be used in regression and classification models as independent variables.
- A wide range of ALSFW metrics is included in *WoLFeX*. New metrics for the characterization of understory vegetation have been proposed. These new metrics allow for the identification and quantification of understory vegetation, which is a key parameter for forest fires mitigation.

Objective 3: To assess the influence of the scan angle of ALS data acquisition and the application or not of a radiometric correction on (i) the extraction of an ALSFW metric; and (ii) modeling three of the most relevant forest fuel variables—canopy fuel load (CFL), canopy height (CH), and canopy base height (CBH).

- Values of ALSFW metric RWE are influenced by the scan angle of the waveforms. Estimated values of RWE should be equal for a same sample; however, up to a nRMSE difference of 15.40% was observed in the estimate of RWE from waveforms acquired with extreme scan angles.

- These differences in RWE values related to scan angle decrease when relative radiometric correction is applied, but they are not completely removed. The efficiency of radiometric correction is affected by the correction parameter power *n* (see Equation 1). Thus, differences quantified by nRMSE can be reduced in more than 4% using radiometrically corrected data.

- Forest fuel attributes (CFL, CH and CBH) estimated from ALSFW metrics are also influenced by scan angle. Different values of forest fuel attributes are estimated for a same plot when using different scan angles. Nevertheless, estimates of forest fuel attributes are differently influenced by scan angles. For instance, mass-related attributes are more influenced than height attributes.

- On the other hand, differences in estimates of forest fuel attributes due to scan angle are also reduced, and its accuracy improved, by applying radiometric correction. Again, estimates of mass-related attributes are strongly influenced, and its accuracy improved, by applying radiometric correction, while height attributes are little influenced. Forest fuel attributes are differently influenced since height attributes are fixed at a specific point on the waveform, usually a maximum or minimum, while mass-related attributes are described using the complete waveform profile. Therefore, the latter are more subject to radiometric values.

Objective 4: Characterize the signal occlusion along the vertical forest structure using different laser scanning configurations (i.e., ALSD, ALSFW, and TLS) in contrasted ecosystems with different canopy covers to determine how reliable the resulting vertical distribution profiles are based on the amount of occlusion and the lack of information.

- Overall, results confirm the general trend largely accepted by the scientific community, which implies that laser scanning signal occlusion prevails in sectors blocked by dense canopies. Signal occlusion therefore depends largely on the laser scanning configuration: ALS, viewing forest from the top-down, is more limited to sample the lower strata, while TLS, viewing the forest from bottom-up, is more limited to sample the top of the canopy.
- The rate of pulse reduction, which is the ratio between the number of blocked laser beams prior to reach a given voxel and the number of theoretical laser beams that should cross a given voxel, is a good indicator of occlusion. Therefore, vertical profiles representing the rate of pulse reduction allows for quantifying the occlusion in the different vertical layers. This involves assessing the reliability of the vegetative material detected from the different laser scanning configurations at the different vertical strata.
- Comparing vertical profiles retrieved from the number of hits of ALS_{FW} and ALS_D, and taking PAD from TLS as reference, ALS_D concentrates most of its data at the upper strata, since the dominant strata generates occlusion underneath. Despite this, the detection of the top of the canopy by ALS_{FW} and ALS_D are not significantly different. Nevertheless, detection of strata below the dominant strata (intermediate and/or lower strata) is enhanced by ALS_{FW}.

Objective 5: Determine how understory vegetation density classes can be detected and further determined by ALS configurations, and whether ALS_{FW} allows for the detection and determination to a level of detail beyond ALS_D capability.

- Again, using the PAD from TLS as reference, vertical profiles at lower strata retrieved from the number of hits of ALS_D and ALS_{FW} data are both high correlated with vertical profiles of TLS. However, vertical profiles retrieved from ALS_{FW} are significantly more accurate. This implies that vegetative material at lower strata (i.e., understory vegetation) are generally well detected by ALS configurations, but especially by ALS_{FW}.

- Despite the correlations between the number of hits of ALS (i.e., ALS$_{FW}$ and ALS$_D$) and the PAD from TLS are high, this is a unitless indicator, and therefore, it only considers the shape but not the order of magnitude of the vertical profiles at the lower strata. In this regard, ALS$_{FW}$ detects understory vegetation with a much larger number of hits than ALS$_D$. This implies that lower strata, and therefore understory vegetation, is registered in more detail with ALS$_{FW}$.

- The considerably lower number of hits registered from ALS$_D$ in the lower strata with respect to those registered by ALS$_{FW}$ and TLS, but the significant correlation with TLS, suggests that understory vegetation can be successfully detected at plot-level but not at finer scales. The low density of ALS$_D$ point clouds cannot represent the complexity of each individual understory vegetation element. However, the significant correlation with TLS at plot-level shows that the presence of understory vegetation within the plot is generally detected.

- The variation of the vertical profiles of the number of hits of ALS configurations at the lower strata depends on the density of understory vegetation. From up-bottom, gradual variations are related to sparse understory vegetation, while steep variations are related to dense understory vegetation.

- The application of the Gini index quantifies how these variations in the vertical profile of the number of hits are. Therefore, this index can be used to determine the density of understory vegetation by using the ALS (i.e., ALS$_{FW}$ and ALS$_D$) vertical profiles of the number of hits at the lower strata as input. Again, the determination of understory vegetation density is more accurate with ALS$_{FW}$. Gini index class interval thresholds to determine understory vegetation density from ALS$_D$ are fuzzier, implying more overlap between classes, than those from ALS$_{FW}$.

- Despite strong signal occlusion at lower strata in the forest, ALS and, especially ALS$_{FW}$, can be processed to identify understory densities with the Gini index.

Objective 6: Apply and validate the new metrics described in objective 2 derived using a voxel based approach to estimate understory height, cover, and volume in a Mediterranean forest ecosystem, proposing some practical recommendations for further development and testing ALS$_{FW}$ metrics.

- Understory vegetation can be characterized very accurately in a Mediterranean ecosystem by using ALS$_{FW}$ metrics. Among the understory vegetation attributes used for the characterization, mean height and volume have the highest correlations, above the maximum height and cover.
- Contrary to ALS$_D$, and in order to enhance estimations of understory vegetation characterization, ALS$_{FW}$ data do not need to be cut off to exclude the waveform fragments that do not correspond to understory vegetation. This is likely due to the fact that contrary to ALS$_D$, which has a limited number of digitized hits, ALS$_{FW}$ can fully discriminate height strata through decomposing the waveform.
- The terrain slope alone has little influence on the prediction of understory vegetation cover, height, and volume. However, when combined with dominant species it has a more significant effect, and therefore, estimation results are improved.

Overall, the use of ALS$_D$ data is adequate for most of the current ALS applications. Nevertheless, we have demonstrated that ALS$_{FW}$ is capable of going one step further in the identification and/or characterization of the intermediate and lower vegetation strata. This is crucial for applications such as wildfire mitigation, modeling fire behavior, planning forest thinning and maintenance tasks for wildfire reduction and biodiversity preservation, or mapping biodiversity. However, the use of ALS$_{FW}$ data requires the identification of the appropriate parameters, such as the optimal pulse density and processing parameters (i.e., voxel size and assignation value), as well as the application of radiometric correction prior to undertake any data processing.

Analyzing the vertical forest structure and occlusion with terrestrial and airborne laser scanning configurations allowed us for a better understanding of a potential previously mentioned by other authors: the use of TLS to calibrate

ALSꜰᴡ, thus the latter could replace TLS to retrieve information from forest internal structure massively, at broader scales.

Nevertheless, ALSꜰᴡ still presents some limitations, such as signal occlusion due to dense vegetation. This may have particular importance in forests with high canopies and very dense vegetation and internal structure (e.g., tropical areas), where the energy from ALS pulses may not be sufficient to reach the ground or the lowest vegetation strata.

6.2. Further research

This research deepened into the use of ALSꜰᴡ for the characterization of the vertical forest structure in different ecosystems. Several aspects were addressed such as the influence of pulse density on ALSꜰᴡ metrics and on forest attributes estimates, and how this influence may be reduced; creation and description of a new processing tool to process ALSꜰᴡ data, including this tool new proposed ALSꜰᴡ metrics related to understory vegetation; the analysis of the influence of scan angle and the radiometric correction of ALSꜰᴡ metrics and on the estimate of forest fuel attributes; the characterization of the vertical distribution and occlusion for different laser scanning configurations (ALSᴅ, ALSꜰᴡ and TLS); the comparison in the detection and determination of understory vegetation density with ALSᴅ and ALSꜰᴡ in two contrasted ecosystems; and the characterization of understory vegetation with ALSꜰᴡ metrics in a Mediterranean ecosystem. This thesis may also be taken as the start point of further research to deepen even more into the use of ALSꜰᴡ for the characterization of the vertical forest structure.

Promising further research might be mainly focused on deepening into some aspects addressed in this thesis, such as the use of ALSꜰᴡ data to estimate useful inputs for 3D simulations of wildfire behavior (e.g., WFDS and Firetec). These new 3D physical fire behavior models include the fuel-fire-atmosphere interactions, which require very accurate forest structure and fuel moisture information at voxel-level. Thus, it is important to classify and map forest species by combining ALSꜰᴡ with other sensors (e.g., multispectral or hyperspectral), and to estimate forest structure parameters (canopy bulk

density, canopy base height, cover and height) at voxel-level following some of the methods presented in this thesis. On the other hand, it is also important to analyze the use of emerging ALS$_{FW}$ sensors and deep learning techniques, such as GEDI (Global Ecosystem Dynamics Investigation) and convolutional neural networks, respectively. Furthermore, new techniques for storing and processing ALS$_{FW}$ might also be addressed.

Several interests for further research are proposed below:

- Classification of forest species by using ALS$_{FW}$ and other sensors (e.g., multispectral or hyperspectral).
- Estimation of CBD at voxel-level using ALS$_{FW}$.
- Identification of understory vegetation and ladder fuels, and quantification of their structural properties (i.e., volume and biomass) at voxel-level using ALS$_{FW}$.
- Analysis of the side-lap effect on different ecosystems with different dominant species.
- Assessment of the effect of the emitted pulse energy and the footprint size on ALS$_{FW}$ values.
- Study of the relationships between estimates of PAD and vertical profiles of number of hits for ALS configurations.
- Analysis of the relationship between vertical profiles of rate of pulse reduction and classification of forest types and structure.
- Validation of the characterization of understory vegetation with ALS$_{FW}$ data in different ecosystems.
- During the development of this thesis, it has been observed that ALS$_{FW}$ data are not standardized. Firstly, this is reflected in how data are provided depending on the sensor and on the processing software used to generate LAS files from RAW files. For instance, line parameters of the waveforms (i.e., Xt, Yt and Zt), which describe the direction and location of the waveform, are sometimes represented differently. Secondly, given that LAS format was originally created for ALS$_D$ data, different hits belonging to one waveform are provided, which is redundant to represent the waveform. A standardization of a new LAS format for the

distribution of ALS$_{FW}$ data would be suitable. This new format would standardize the parameters present in the LAS file to represent the ALS$_{FW}$ data, as well as resulting in a reduction of memory and processing time.

- Exploration of a new denoising process for ALS$_{FW}$ data based on the Fourier transform. This technique decomposes a waveform into its constituent frequencies. Therefore, some of these frequencies may be identified as noise and removed.

- Comparison of new sources of full-waveform laser scanning data, such as GEDI (i.e., a space-borne laser scanning), with ALS data. In addition, given that ALS data are denser than those from GEDI, the latter could be calibrated using ALS data as reference. This would allow for the characterization of forest fuel and vertical forest structure in wider areas with time series using the wall-to-wall procedure, as it has been used for ALS$_D$ and Landsat data (Saarela et al., 2015).

- Exploration of the use of convolutional neural networks to characterize the forest fuel and vertical forest structure with ALS$_{FW}$ data for estimating forest attributes or for classifying tree species and fuel types, etc. The development of 3D segmentation methods to identify individual trees or tree elements (e.g., trunk, crown, etc.) can also be relevant for future work in forest structure.

References

A

Abshire, J.B., Sun, X., Riris, H., Sirota, J.M., McGarry, J.F., Palm, S., Yi, D., Liiva, P., 2005. Geoscience Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance. Geophys. Res. Lett. 32, L21S02. https://doi.org/10.1029/2005GL024028

Agee, J.K., 1996. The influence of forest structure on fire behavior, in: 17th Annual Forest Vegetation Management Conference. pp. 52–68.

Agee, J.K., Skinner, C.N., 2005. Basic principles of forest fuel reduction treatments. For. Ecol. Manage. 211, 83–96. https://doi.org/10.1016/j.foreco.2005.01.034

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in: B.N. Petrow, F.C. (Ed.), Proceedings of the 2nd International Symposium on Information, Akademiai Kiado, Budapest.

Albini, F.A., 1976. Estimating wildfire behavior and effects. Ogden, UT.

Albini, F.A., 1986. Wildland fire spread by radiation-a model including fuel cooling by natural convection. Combust. Sci. Technol. 45, 101–113. https://doi.org/10.1080/00102208608923844

Albright, D., Meisner, B.N., 1999. Classification of fire simulation systems.

Alexander, M.E., Cruz, M.G., Lopes, A.M.G., 2006. CFIS: A software tool for simulating crown fire initiation and spread. For. Ecol. Manage. 234, S133. https://doi.org/10.1016/j.foreco.2006.08.174

Almeida, D.R.A., Stark, S.C., Chazdon, R., Nelson, B.W., Cesar, R.G., Meli, P., Gorgens, E.B., Duarte, M.M., Valbuena, R., Moreno, V.S., Mendes, A.F., Amazonas, N., Gonçalves, N.B., Silva, C.A., Schietti, J., Brancalion, P.H.S., 2019. The effectiveness of lidar remote sensing for monitoring forest cover attributes and landscape restoration. For. Ecol. Manage. 438, 34–43. https://doi.org/10.1016/j.foreco.2019.02.002

Andersen, H.-E., McGaughey, R.J., Reutebuch, S.E., 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sens. Environ. 94, 441–449. https://doi.org/10.1016/j.rse.2004.10.013

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 Sens. Ecol. Conserv. 2, 5–15. https://doi.org/10.1002/rse2.8

Andrews, P.L., 1986. BEHAVE: fire behavior prediction and fuel modeling system-BURN Subsystem, part 1. https://doi.org/10.2737/INT-GTR-194

Andrews, P.L., Queen, Ll.P., 2001. Fire modeling and information system technology. Int. J. Wildl. Fire 10, 343–352. https://doi.org/10.1071/WF01033

Andrews, P.L., 2009. BehavePlus fire modeling system, version 5.0: Variables. https://doi.org/10.2737/RMRS-GTR-213

Andrews, P.L., 2014. Current status and future needs of the BehavePlus Fire Modeling System. Int. J. Wildl. Fire 23, 21. https://doi.org/10.1071/WF12167

Arkle, R.S., Pilliod, D.S., Welty, J.L., 2012. Pattern and process of prescribed fires influence effectiveness at reducing wildfire severity in dry coniferous forests. For. Ecol. Manage. 276, 174–184. https://doi.org/10.1016/j.foreco.2012.04.002

Arroyo, L.A., Pascual, C., Manzanera, J.A., 2008. Fire models and methods to map fuel types: The role of remote sensing. For. Ecol. Manage. 256, 1239–1252. https://doi.org/10.1016/j.foreco.2008.06.048

Assmann, E., 1961. Waldertragskunde — Organische Produktion, Struktur, Zuwachs und Ertrag von Waldbeständen. München, Germany.

Assmann, E., 1970. The principles of forest yield study: studies in the organic production, structure, increment and yield of forest stands.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. Int. Arch. Photogramm. Remote Sens. 33, 111–118.

B

Bailey, R.G., 1980. Description of the ecoregions of the United States. Washington DC, USA. https://doi.org/10.3133/ofr20161021

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. For. Ecol. Manage. 254, 1–15. https://doi.org/10.1016/j.foreco.2007.09.038

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

Béland, M., Widlowski, J.-L., Fournier, R.A., 2014. A model for deriving voxel-level tree leaf area density estimates from ground-based LiDAR. Environ. Model. Softw. 51, 184–189. https://doi.org/10.1016/j.envsoft.2013.09.034

Blair, J.B., Rabine, D.L., Hofton, M.A., 1999. The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. ISPRS J. Photogramm. Remote Sens. 54, 115–122. https://doi.org/10.1016/S0924-2716(99)00002-7

Bohn, F.J., Huth, A., 2017. The importance of forest structure to biodiversity–productivity relationships. R. Soc. Open Sci. 4, 160521. https://doi.org/10.1098/rsos.160521

Bottalico, F., Chirici, G., Giannini, R., Mele, S., Mura, M., Puxeddu, M., McRoberts, R.E., Valbuena, R., Travaglini, D., 2017. Modeling Mediterranean forest structure using airborne laser scanning data. Int. J. Appl. Earth Obs. Geoinf. 57, 145–153. https://doi.org/10.1016/j.jag.2016.12.013

Briese, C., Pfennigbauer, M., Lehner, H., Ullrich, A., Wagner, W., Pfeifer, N., 2012. Radiometric calibration of multi-wavelength airborne laser scanning data. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. I–7, 335–340. https://doi.org/10.5194/isprsannals-I-7-335-2012

Burgan, R.E., Rothermel, R.C., 1984. BEHAVE: fire behaviour prediction and fuel modeling system -- FUEL subsystem. Ogden, UT.

C

Camacho, J., 2020. Gedi 0.2 for Python.

Canfield, R.H., 1941. Application of the line interception method in sampling range vegetation. J. For. 39, 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 Sens. 6, 7110–7135. https://doi.org/10.3390/rs6087110

Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens. Environ. 113, 893–903. https://doi.org/10.1016/j.rse.2009.01.007

Chasmer, L., Hopkinson, C., Treitz, P., 2006. Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Can. J. Remote Sens. 32, 116–125. https://doi.org/10.5589/m06-011

Chen, Y., Zhu, X., Yebra, M., Harris, S., Tapper, N., 2016. Strata-based forest fuel classification for wild fire hazard assessment using terrestrial LiDAR. J. Appl. Remote Sens. 10, 046025. https://doi.org/10.1117/1.JRS.10.046025

Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.T., 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. Trees 21, 295–310. https://doi.org/10.1007/s00468-006-0119-6

Countryman, C.M., 1972. Fuel moisture and fire danger: some elementary concepts.

Crawley, M.J., 2012. Mixed-effects models, The R Book. ed.

Crespo-Peremarch, P., Ruiz, L.Á., Balaguer-Beser, Á., Estornell, J., 2015. Análisis temporal de la estructural forestal mediante métricas derivadas de LiDAR full-waveform. Actas XVI Congreso de la Asociación Española de Teledetección. Teledetección, humedales y espacios protegidos, 387–390. 21-23 Oct., Sevilla, Spain.

Crespo-Peremarch, P., Ruiz, L.Á., Balaguer-Beser, Á., 2016. A comparative study of regression methods to predict forest structure and canopy fuel variable from LiDAR full-waveform data. Revista de Teledetección, 45, 27-40. https://doi.org/10.4995/raet.2016.4066

Crespo-Peremarch, P., Ruiz, L.Á., Balaguer-Beser, Á., 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 & Spatial Information Sciences, XLI-B8, 603–610. 12-19 July, Prague, Czech Republic. https://doi.org/10.5194/isprs-archives-XLI-B8-603-2016

Crespo-Peremarch, P., Ruiz, L.Á., 2016. Estimating forest structure attributes from full-waveform LiDAR: comparative analysis of methodological parameters in two geographic areas. Proceedings 3rd EARSeL Workshop SIG on Forestry, p.82. 15-16 Sept., Krakow, Poland.

Crespo-Peremarch, P., Ruiz, L.Á., 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. 4-7 Oct., Murcia, Spain.

Crespo-Peremarch, P., Ruiz, L.Á., Balaguer-Beser, Á., Estornell, J., 2018a. Analyzing the role of pulse density and voxelization parameters on full-waveform LiDAR-derived metrics. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 453–464. https://doi.org/10.1016/j.isprsjprs.2018.10.012

Crespo-Peremarch, P., Tompalski, P., Coops, N.C., Ruiz, L.Á., 2018b. Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data. Remote Sensing of Environment, 217, 400–413. https://doi.org/10.1016/j.rse.2018.08.033

Crespo-Peremarch, P., Ruiz, L.Á., 2018. Influence of lidar full-waveform density and voxel size on forest stand estimates. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium IEEE, 5895–5898. 23-27 July, València, Spain. https://doi.org/10.1109/IGARSS.2018.8517594

Crespo-Peremarch, P., Ruiz, L.Á., 2019. Analysis of side-lap effect and characterization of understory vegetation using full-waveform ALS. Proceedings of the II Geomatics Engineering Conference, 19, 6. 26-27 June, Madrid, Spain. https://doi.org/10.3390/proceedings2019019006

Crespo-Peremarch, P., Ruiz, L.Á., 2020. A full-waveform airborne laser scanning metric extraction tool for forest structure modelling. Do scan angle and radiometric correction matter? Remote Sensing, 12, 292. https://doi.org/10.3390/rs12020292

Crespo-Peremarch, P., Fournier, R.A., Nguyen, V.-T., van Lier, O.R., Ruiz, L.Á., 2020. A comparative assessment of the vertical distribution of forest components using full-waveform airborne, discrete airborne and discrete terrestrial laser scanning data. Forest Ecology and Management, 473, 118268. https://doi.org/10.1016/j.foreco.2020.118268

Cruz, M.G., Alexander, M.E., Wakimoto, R.H., 2003. Assessing canopy fuel stratum characteristics in crown fire prone fuel types of western North America. Int. J. Wildl. Fire 12, 39. https://doi.org/10.1071/WF02024

David, M., 1977. Geostatistical ore reserve estimation, Developments in Geomathematics. Elsevier, Amsterdam, The Netherlands. https://doi.org/10.1016/C2009-0-14528-2

Davis, R., Yang, Z., Yost, A., Belongie, C., Cohen, W., 2017. The normal fire environment—Modeling environmental suitability for large forest wildfires using past, present, and future climate normals. For. Ecol. Manage. 390, 173–186. https://doi.org/10.1016/j.foreco.2017.01.027

Deeming, J.E., Lancarter, J.W., Fosberg, M.A., Furman, R.W., Schroeder, M.J., 1972. The national fire-danger rating system. Ogden, UT.

Deeming, J.E., Burgan, R.E., Cohen, J.D., 1977. The national fire danger rating system. Ogden, UT.

Devictor, V., Julliard, R., Jiguet, F., 2008. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. Oikos 117, 507–514. https://doi.org/10.1111/j.0030-1299.2008.16215.x

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., Chazdon, R.L., Weishampel, J.F., Prince, S., 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sens. Environ. 79, 305–319. https://doi.org/10.1016/S0034-4257(01)00281-4

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B., Hofton, M., 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Glob. Ecol. Biogeogr. 12, 147–159. https://doi.org/10.1046/j.1466-822X.2003.00010.x

Dubayah, R.O., Drake, J.B., 2000. Lidar remote sensing for forestry. J. For. 98, 44–46. https://doi.org/10.1093/jof/98.6.44

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Sci. Remote Sens. 1, 100002. https://doi.org/10.1016/j.srs.2020.100002

Duda, R.O., Hart, P.E., Stork, D.G., 2012. Pattern classification.

Duncanson, L.I., Niemann, K.O., Wulder, M.A., 2010. Estimating forest canopy height and terrain relief from GLAS waveform metrics. Remote Sens. Environ. 114, 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 Diss. TU Delft, Delft University of Technology. https://doi.org/10.1006/fsim.1997.0094

ℰ

Erdody, T.L., Moskal, L.M., 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sens. Environ. 114, 725–737. https://doi.org/10.1016/j.rse.2009.11.002

Estornell, J., Ruiz, L.Á., Velázquez-Martí, B., 2011. Study of shrub cover and height using LiDAR data in a Mediterranean area. For. Sci. 57, 171–179. https://doi.org/10.1093/forestscience/57.3.171

ℱ

Ferrari, S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. J. Appl. Stat. 31, 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 J. Photogramm. Remote Sens. 104, 144–157. https://doi.org/10.1016/j.isprsjprs.2015.03.001

Finney, M.A., 1998. FARSITE: Fire Area Simulator-model development and evaluation. Rocky Mountain Research Station.

Forestry Canada, 1992. Development and structure of the Canadian Forest Fire Behavior Prediction System. Ottawa, Canada.

𝒢

García, M., Danson, F.M., Riaño, D., Chuvieco, E., Ramirez, F.A., Bandugula, V., 2011. Terrestrial laser scanning to estimate plot-level forest canopy fuel properties. Int. J. Appl. Earth Obs. Geoinf. 13, 636–645. https://doi.org/10.1016/j.jag.2011.03.006

García, M., Popescu, S., Riaño, D., Zhao, K., Neuenschwander, A., Agca, M., Chuvieco, E., 2012. Characterization of canopy fuels using ICESat/GLAS data. Remote Sens. Environ. 123, 81–89. https://doi.org/10.1016/j.rse.2012.03.018

Gelman, A., Hill, J., 2007. Multilevel regression, in: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge, pp. 235–236. https://doi.org/10.1017/CBO9780511790942.014

Giannetti, F., Puletti, N., Quatrini, V., Travaglini, D., Bottalico, F., Corona, P., Chirici, G., 2018. Integrating terrestrial and airborne laser scanning for the assessment of single-tree attributes in Mediterranean forest stands. Eur. J. Remote Sens. 51, 795–807. https://doi.org/10.1080/22797254.2018.1482733

Gini, C., 1912. Variabilità e mutabilità.

Gobakken, T., Næsset, E., 2008. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. Can. J. For. Res. 38, 1095–1109. https://doi.org/10.1139/X07-219

González-Ferreiro, E., Diéguez-Aranda, U., Miranda, D., 2012. Estimation of stand variables in Pinus radiata D. Don plantations using different LiDAR pulse densities. Forestry 85, 281–292. https://doi.org/10.1093/forestry/cps002

Gosset, W.S., 1908. The Probable Error of a Mean. Biometrika 6, 1–25. https://doi.org/10.2307/2331554

Graham, R.T., Harvey, A.E., Jain, T.B., Tonn, J.R., 1999. Effects of thinning and similar stand treatments on fire behavior in western forests.

Graham, R.T., McCaffrey, S., Jain, T.B., 2004. Science basis for changing forest structure to modify wildfire behavior and severity. https://doi.org/10.2737/RMRS-GTR-120

Greaves, H.E., Vierling, L.A., Eitel, J.U.H., 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 Sens. Environ. 164, 26–35. https://doi.org/10.1016/j.rse.2015.02.023

Guo, L., Chehata, N., Mallet, C., Boukir, S., 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. ISPRS J. Photogramm. Remote Sens. 66, 56–66. https://doi.org/10.1016/j.isprsjprs.2010.08.007

Guo, X., Coops, N.C., Tompalski, P., Nielsen, S.E., Bater, C.W., John Stadt, J., 2017. Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data. Ecol. Inform. 38, 50–61. https://doi.org/10.1016/j.ecoinf.2017.01.005

H

Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P., Disney, M., Mark Danson, F., Strahler, A., Schaaf, C., Anderson, K., Gaston, K.J., 2015. Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. Remote Sens. Environ. 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 Sens. Environ. 188, 37–50. https://doi.org/10.1016/j.rse.2016.10.041

Hansen, E.H., Gobakken, T., Næsset, E., 2015. Effects of pulse density on digital terrain models and canopy metrics using airborne laser scanning in a tropical rainforest. Remote Sens. 7, 8453–8468. https://doi.org/10.3390/rs70708453

Harding, D.., Lefsky, M.., Parker, G.., Blair, J.., 2001. Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. Remote Sens. Environ. 76, 283–297. https://doi.org/10.1016/S0034-4257(00)00210-8

Harrod, R.J., Peterson, D.W., Povak, N.A., Dodson, E.K., 2009. Thinning and prescribed fire effects on overstory tree and snag structure in dry coniferous forests of the interior Pacific Northwest. For. Ecol. Manage. 258, 712–721.

https://doi.org/10.1016/j.foreco.2009.05.011

Harten, A., 1983. High resolution schemes for hyperbolic conservation laws. J. Comput. Phys. 49, 357–393. https://doi.org/10.1016/0021-9991(83)90136-5

Hermosilla, T., Ruiz, L.A., Kazakova, A.N., Coops, N.C., Moskal, L.M., 2014a. Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. Int. J. Wildl. Fire 23, 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 Sens. Lett. 5, 332–341. https://doi.org/10.1080/2150704X.2014.903350

Hevia, A., Álvarez-González, J.G., Ruiz-Fernández, E., Prendes, C., Ruiz-González, A.D., Majada, J., González-Ferreiro, E., 2016. Modelling canopy fuel and forest stand variables and characterizing the influence of thinning in the stand structure using airborne LiDAR. Rev. Teledetección 41–55. https://doi.org/10.4995/raet.2016.3979

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

Hilker, T., Coops, N.C., Newnham, G.J., van Leeuwen, M., Wulder, M.A., Stewart, J., Culvenor, D.S., 2012. Comparison of terrestrial and airborne LiDAR in describing stand structure of a thinned lodgepole pine forest. J. For. 110, 97–104. https://doi.org/10.5849/jof.11-003

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 J. Photogramm. Remote Sens. 64, 223–233. https://doi.org/10.1016/j.isprsjprs.2008.12.004

Hoffman, C.M., Canfield, J., Linn, R.R., Mell, W., Sieg, C.H., Pimont, F., Ziegler, J., 2016. Evaluating Crown Fire Rate of Spread Predictions from Physics-Based Models. Fire Technol. 52, 221–237. https://doi.org/10.1007/s10694-015-0500-3

Hofton, M.A., Minster, J.B., Blair, J.B., 2000. Decomposition of laser altimeter waveforms. IEEE Trans. Geosci. Remote Sens. 38, 1989–1996. https://doi.org/10.1109/36.851780

Hopkinson, C., Chasmer, L., 2009. Testing LiDAR models of fractional cover across multiple forest ecozones. Remote Sens. Environ. 113, 275–288. https://doi.org/10.1016/j.rse.2008.09.012

Hopkinson, C., Lovell, 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 Sens. Environ. 136, 301–314. https://doi.org/10.1016/j.rse.2013.05.012

Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., Walker, W., 2005. Mapping forest structure for wildlife habitat analysis using

waveform lidar: Validation of montane ecosystems. Remote Sens. Environ. 96, 427–437. https://doi.org/10.1016/j.rse.2005.03.005

Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. Remote Sens. Environ. 102, 63–73. https://doi.org/10.1016/j.rse.2006.01.021

I

Isenburg, M., 2012. PulseWaves: An open, vendor-neutral, stand-alone, LAS-Compatible full waveform LiDAR standar <http://pulsewaves.org>.

Isenburg, M., 2017. LAStools.

J

Jakubowski, M.K., Guo, Q., Kelly, M., 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sens. Environ. 130, 245–253. https://doi.org/10.1016/j.rse.2012.11.024

Johnson, P.C.D., 2014. Extension of Nakagawa & Schielzeth's R2 GLMM to random slopes models. Methods Ecol. Evol. 5, 944–946. https://doi.org/10.1111/2041-210X.12225

Johnstone, J.F., Allen, C.D., Franklin, J.F., Frelich, L.E., Harvey, B.J., Higuera, P.E., Mack, M.C., Meentemeyer, R.K., Metz, M.R., Perry, G.L., Schoennagel, T., Turner, M.G., 2016. Changing disturbance regimes, ecological memory, and forest resilience. Front. Ecol. Environ. 14, 369–378. https://doi.org/10.1002/fee.1311

Jung, S.-E., Kwak, D.-A., Park, T., Lee, W.-K., Yoo, S., 2011. Estimating crown variables of individual trees using airborne and terrestrial laser scanners. Remote Sens. 3, 2346–2363. https://doi.org/10.3390/rs3112346

K

Kankare, V., Holopainen, M., Vastaranta, M., Puttonen, E., Yu, X., Hyyppä, J., Vaaja, M., Hyyppä, H., Alho, P., 2013. Individual tree biomass estimation using terrestrial laser scanning. ISPRS J. Photogramm. Remote Sens. 75, 64–75. https://doi.org/10.1016/j.isprsjprs.2012.10.003

Kashani, A.G., Olsen, M.J., Parrish, C.E., Wilson, N., 2015. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. Sensors 15, 28099–28128. https://doi.org/10.3390/s151128099

Keane, R.E., Reinhardt, E.D., Scott, J., Gray, K., Reardon, J., 2005. Estimating forest canopy bulk density using six indirect methods. Can. J. For. Res. 35, 724–739. https://doi.org/10.1139/x04-213

Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. Trends Ecol. Evol. 18, 299–305. https://doi.org/10.1016/S0169-5347(03)00071-5

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

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, e0122070. https://doi.org/10.1371/journal.pone.0122070

Kükenbrink, D., Schneider, F.D., Leiterer, R., Schaepman, M.E., Morsdorf, F., 2017. Quantification of hidden canopy volume of airborne laser scanning data using a voxel traversal algorithm. Remote Sens. Environ. 194, 424–436. https://doi.org/10.1016/j.rse.2016.10.023

Kukko, A., Kaasalainen, S., Litkey, P., 2008. Effect of incidence angle on laser scanner intensity and surface data. Appl. Opt. 47, 986–992. https://doi.org/10.1364/AO.47.000986

L

LaRue, E.A., Wagner, F.W., Fei, S., Atkins, J.W., Fahey, R.T., Gough, C.M., Hardiman, B.S., 2020. Compatibility of aerial and terrestrial LiDAR for quantifying forest structural diversity. Remote Sens. 12, 1407. https://doi.org/10.3390/rs12091407

Lecina-Diaz, J., Alvarez, A., Retana, J., 2014. Extreme fire severity patterns in topographic, convective and wind-driven historical wildfires of Mediterranean pine forests. PLoS One 9, e85127. https://doi.org/10.1371/journal.pone.0085127

Lefsky, M.A., Harding, D., Cohen, W.., Parker, G., Shugart, H.., 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of Eastern Maryland, USA. Remote Sens. Environ. 67, 83–98. https://doi.org/10.1016/S0034-4257(98)00071-6

Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar remote sensing for ecosystem studies: lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular inte. Bioscience 52, 19–30. https://doi.org/10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2

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. Geophys. Res. Lett. 32. https://doi.org/10.1029/2005GL023971

Lefsky, M.A., Keller, M., Pang, Y., Camargo, P.B. de, Hunter, M.O., 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter

System waveforms. J. Appl. Remote Sens. 1, 013537. https://doi.org/10.1117/1.2795724

Lesak, A.A., Radeloff, V.C., Hawbaker, T.J., Pidgeon, A.M., Gobakken, T., Contrucci, K., 2011. Modeling forest songbird species richness using LiDAR-derived measures of forest structure. Remote Sens. Environ. 115, 2823–2835. https://doi.org/10.1016/j.rse.2011.01.025

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 J. Photogramm. Remote Sens. 115, 63–77. https://doi.org/10.1016/j.isprsjprs.2016.01.006

Lim, K., Treitz, P., Wulder, M.A., St-Onge, B., Flood, M., 2003. LiDAR remote sensing of forest structure. Prog. Phys. Geogr. 27, 88–106. https://doi.org/10.1191/0309133303pp360ra

Lin, Y., Herold, M., 2016. Tree species classification based on explicit tree structure feature parameters derived from static terrestrial laser scanning data. Agric. For. Meteorol. 216, 105–114. https://doi.org/10.1016/j.agrformet.2015.10.008

Lindberg, E., Olofsson, K., Holmgren, J., Olsson, H., 2012. Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. Remote Sens. Environ. 118, 151–161. https://doi.org/10.1016/j.rse.2011.11.015

Linn, R., Reisner, J., Colman, J.J., Winterkamp, J., 2002. Studying wildfire behavior using FIRETEC. Int. J. Wildl. Fire 11, 233. https://doi.org/10.1071/WF02007

Linn, R.R., 1997. A transport model for prediction of wildfire behavior. Los Alamos, NM. https://doi.org/10.2172/505313

Liu, L., Pang, Y., Li, Z., Si, L., Liao, S., 2017. Combining airborne and terrestrial laser scanning technologies to measure forest understorey volume. Forests 8, 111. https://doi.org/10.3390/f8040111

Lorenz, M.O., 1905. Methods of measuring the concentration of wealth. Publ. Am. Stat. Assoc. 9, 209–219.

Luther, J.E., Fournier, R.A., van Lier, O.R., Bujold, M., 2019. Extending ALS-based mapping of forest attributes with medium resolution satellite and environmental data. Remote Sens. 11, 1092. https://doi.org/10.3390/rs11091092

M

Maas, H. -G., Bienert, A., Scheller, S., Keane, E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. Int. J. Remote Sens. 29, 1579–1593. https://doi.org/10.1080/01431160701736406

Mafanya, M., Tsele, P., Botai, J.O., Manyama, P., Chirima, G.J., Monate, T., 2018. Radiometric calibration framework for ultra-high-resolution UAV-derived orthomosaics for large-scale mapping of invasive alien plants in semi-arid woodlands: Harrisia pomanensis as a case study. Int. J. Remote Sens. 39, 5119–

5140. https://doi.org/10.1080/01431161.2018.1490503

Magnussen, S., Næsset, E., Gobakken, T., 2010. Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. Remote Sens. Environ. 114, 700–712. https://doi.org/10.1016/j.rse.2009.11.007

Mallet, C., Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. ISPRS J. Photogramm. Remote Sens. 64, 1–16. https://doi.org/10.1016/j.isprsjprs.2008.09.007

Maltamo, M., Næsset, E., Vauhkonen, J., 2014. Foresty applications of airborne laser scanning. Springer, Netherlands.

Manuri, S., Andersen, H.-E., McGaughey, R.J., Brack, C., 2017. Assessing the influence of return density on estimation of lidar-based aboveground biomass in tropical peat swamp forests of Kalimantan, Indonesia. Int. J. Appl. Earth Obs. Geoinf. 56, 24–35. https://doi.org/10.1016/j.jag.2016.11.002

Martin-Ducup, O., Schneider, R., Fournier, R.A., 2016. Response of sugar maple (Acer saccharum, Marsh.) tree crown structure to competition in pure versus mixed stands. For. Ecol. Manage. 374, 20–32. https://doi.org/10.1016/j.foreco.2016.04.047

Martin-Ducup, O., Schneider, R., Fournier, R.A., 2017. A method to quantify canopy changes using multi-temporal terrestrial lidar data: Tree response to surrounding gaps. Agric. For. Meteorol. 237–238, 184–195. https://doi.org/10.1016/j.agrformet.2017.02.016

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 Sens. Environ. 113, 2533–2546. https://doi.org/10.1016/j.rse.2009.07.002

McArthur, A.G., 1966. Weather and grassland fire behaviour.

McArthur, A.G., 1967. Fire behaviour in eucalypt forests.

McDermid, G.J., Franklin, S.E., LeDrew, E.F., 2005. Remote sensing for large-area habitat mapping. Prog. Phys. Geogr. Earth Environ. 29, 449–474. https://doi.org/10.1191/0309133305pp455ra

McGaughey, R.J., 2014. FUSION/LDV: Software for LiDAR data analysis and visualization, Manual.

Means, J.E., Acker, S. a., Harding, D.J., Blair, J.B., Lefsky, M. a., Cohen, W.B., Harmon, M.E., McKee, W.A., 1999. Use of large-footprint scanning airborne Lidar to estimate forest stand characteristics in the western cascades of Oregon. Remote Sens. Environ. 67, 298–308. https://doi.org/10.1016/S0034-4257(98)00091-1

Mell, W., Jenkins, M.A., Gould, J., Cheney, P., 2007. A physics-based approach to modelling grassland fires. Int. J. Wildl. Fire 16, 1. https://doi.org/10.1071/WF06002

Mell, W., McNamara, D., Maranghides, A., McDermott, R., Forney, G., Hoffmanm, C., Ginder, M., 2011. Computer modelling of wildland-urban interface fires, in: Fire & Materials. San Francisco, CA, USA.

Miltiadou, M., Grant, M., Campbell, N.D., Warren, M., Clewley, D., Hadjimitsis, D.,

2019. Open source software DASOS: efficient accumulation, analysis, and visualisation of full-waveform lidar, in: Papadavid, G., Themistocleous, K., Michaelides, S., Ambrosia, V., Hadjimitsis, D.G. (Eds.), Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019). SPIE, p. 68. https://doi.org/10.1117/12.2537915

Molina, J.R., Rodríguez y Silva, F., Herrera, M.A., Zamora, R., 2009. A simulation tool for socio-economic planning on forest fire suppresion management. Forest fires: detection, suppression, and prevention. Nova Science Publishers, New York, USA.

Molina, J.R., Rodriguez y Silva, F., Herrera, M.A., 2011. Potential crown fire behaviour in Pinus pinea stands following different fuel treatments. For. Syst. 20, 266–277.

Montero, G., Ruiz-Peinado, R., Muñoz, M., 2005. Producción de biomasa y fijación de CO2 por los bosques españoles. Madrid, Spain.

Morsdorf, F., Kötz, B., Meier, E., Itten, K.I., Allgöwer, B., 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sens. Environ. 104, 50–61. https://doi.org/10.1016/j.rse.2006.04.019

Morsdorf, F., Frey, O., Meier, E., Itten, K.I., Allgöwer, B., 2008. Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning. Int. J. Remote Sens. 29, 1387–1406. https://doi.org/10.1080/01431160701736349

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 Sens. Environ. 114, 1403–1415. https://doi.org/10.1016/j.rse.2010.01.023

Moskal, L.M., Zheng, G., 2011. Retrieving forest inventory variables with terrestrial laser scanning (TLS) in urban heterogeneous forest. Remote Sens. 4, 1–20. https://doi.org/10.3390/rs4010001

n

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

National Wildfire Coordinating Group, 2005. Glossary of wildland fire terminology [WWW Document]. Natl. Interag. Fire Cent. URL https://www.nwcg.gov/ (accessed 6.20.19).

Nie, S., Wang, C., Zeng, H., Xi, X., Li, G., 2017. Above-ground biomass estimation using airborne discrete-return and full-waveform LiDAR data in a coniferous forest. Ecol. Indic. 78, 221–228. https://doi.org/10.1016/j.ecolind.2017.02.045

O

Olsoy, P.J., Glenn, N.F., Clark, P.E., Derryberry, D.R., 2014. Aboveground total and green biomass of dryland shrub derived from terrestrial laser scanning. ISPRS J. Photogramm. Remote Sens. 88, 166–173. https://doi.org/10.1016/j.isprsjprs.2013.12.006

Othmani, A., Lew Yan Voon, L.F.C., Stolz, C., Piboule, A., 2013. Single tree species classification from Terrestrial Laser Scanning data for forest inventory. Pattern Recognit. Lett. 34, 2144–2150. https://doi.org/10.1016/j.patrec.2013.08.004

P

Parsons, R.A., Mell, W.E., McCauley, P., 2011. Linking 3D spatial models of fuels and fire: Effects of spatial heterogeneity on fire behavior. Ecol. Modell. 222, 679–691. https://doi.org/10.1016/j.ecolmodel.2010.10.023

Peterson, B., Dubayah, R., Hyde, P., Hofton, M., Blair, J.B., Fites-kaufman, J., 2005. Use of LIDAR for forest inventory and forest management application, in: Proceedings of the Seventh Annual Forest Inventory and Analysis Symposium.

Peterson, D.L., Johnson, M.C., Agee, J.K., Jain, T.B., McKenzie, D., Reinhardt, E.D., 2005. Forest structure and fire hazard in dry forests of the Western United States. https://doi.org/10.2737/PNW-GTR-628

Pfeifer, N., Mandlburger, G., Otepka, J., Karel, W., 2014. OPALS – A framework for Airborne Laser Scanning data analysis. Comput. Environ. Urban Syst. 45, 125–136. https://doi.org/10.1016/j.compenvurbsys.2013.11.002

Piboule, A., Krebs, M., Esclatine, L., Hervé, J.-C., 2015. Computree: A collaborative platform for use of terrestrial lidar in dendrometry, in: International IUFRO Conference MeMoWood. Nancy, France.

Pimont, F., Parsons, R., Rigolot, E., de Coligny, F., Dupuy, J.-L., Dreyfus, P., Linn, R.R., 2016. Modeling fuels and fire effects in 3D: Model description and applications. Environ. Model. Softw. 80, 225–244. https://doi.org/10.1016/j.envsoft.2016.03.003

Pimont, F., Allard, D., Soma, M., Dupuy, J.-L., 2018. Estimators and confidence intervals for plant area density at voxel scale with T-LiDAR. Remote Sens. Environ. 215, 343–370. https://doi.org/10.1016/j.rse.2018.06.024

Pollet, J., Omi, P.N., 2002. Effect of thinning and prescribed burning on crown fire severity in ponderosa pine forests. Int. J. Wildl. Fire 11, 1–10. https://doi.org/10.1071/WF01045

Popescu, S.C., Zhao, K., 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sens. Environ. 112, 767–781. https://doi.org/10.1016/j.rse.2007.06.011

Prichard, S.J., Peterson, D.L., Jacobson, K., 2010. Fuel treatments reduce the severity of wildfire effects in dry mixed conifer forest, Washington, USA. Can. J. For. Res. 40, 1615–1626. https://doi.org/10.1139/X10-109

Prometheus, S.V., 1999. Management techniques for optimization of suppresion and

minimization of wildfire effects.

R

Ramírez, J., Monedero, S., 2011. New approaches in fire simulations analysis with Wildfire Analyst, in: 7th International Conference on Forest Fire Research.

Ravaglia, J., Fournier, R.A., Bac, A., Véga, C., Côté, J.-F., Piboule, A., Rémillard, U., 2019. Comparison of three algorithms to estimate tree stem diameter from terrestrial laser scanner data. Forests 10, 599. https://doi.org/10.3390/f10070599

Reinhardt, E., Scott, J., Gray, K., Keane, R., 2006. Estimating canopy fuel characteristics in five conifer stands in the western United States using tree and stand measurements. Can. J. For. Res. 36, 2803–2814. https://doi.org/10.1139/x06-157

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 behavior modeling. Remote Sens. Environ. 86, 177–186. https://doi.org/10.1016/S0034-4257(03)00098-1

Riaño, D., Chuvieco, E., Condés, S., González-Matesanz, J., Ustin, S.L., 2004. Generation of crown bulk density for Pinus sylvestris L. from lidar. Remote Sens. Environ. 92, 345–352. https://doi.org/10.1016/j.rse.2003.12.014

Riaño, D., Chuvieco, E., Ustin, S.L., Salas, J., Rodríguez-Pérez, J.R., Ribeiro, L.M., Viegas, D.X., Moreno, J.M., Fernández, H., 2007. Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. Int. J. Wildl. Fire 16, 341–348. https://doi.org/10.1071/WF06003

Rothermel, R.C., 1972. A mathematical model for predicting fire spread in wildland fuels.

Roussel, J.R., Auty, D., 2017. lidR: Airborne LiDAR data manipulation and visualization for forestry applications.

Roussel, J.R., Caspersen, J., Béland, M., Thomas, S., Achim, A., 2017. Removing bias from LiDAR-based estimates of canopy height: Accounting for the effects of pulse density and footprint size. Remote Sens. Environ. 198, 1–16. https://doi.org/10.1016/j.rse.2017.05.032

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. Can. J. For. Res. 36, 1345–1362. https://doi.org/10.1139/x06-025

Ruiz, L.A., Hermosilla, T., Mauro, F., Godino, M., 2014. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. Forests 5, 936–951. https://doi.org/10.3390/f5050936

Ruiz, L.Á., Crespo-Peremarch, P., Estornell, J., Balaguer-Beser, Á., 2016. Analizing some factors affecting the extraction of full-waveform LiDAR metrics and their effect in forest structure variable estimates. Proceedings 7th Edition of the International Scientific Conference ForestSAT 2016, 24–26. 15-17 Nov., Santiago

de Chile, Chile.

Ruiz, L.Á., Crespo-Peremarch, P., 2017. Optimizing operational parameters in a full-waveform LiDAR processing tool for forestry. 17th Symposium on Systems Analysis in Forest Resources (SSAFR 2017), 27-30 Aug., Suquamish, WA, USA.

Ruiz, L.Á., 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 Int. 33, 443–457. https://doi.org/10.1080/10106049.2016.1265595

S

Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyyppä, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. Remote Sens. Environ. 158, 431–440. https://doi.org/10.1016/j.rse.2014.11.020

Sandberg, D. V., Ottmar, R.D., Cushon, G.H., 2001. Characterizing fuels in the 21st Century. Int. J. Wildl. Fire 10, 381–387. https://doi.org/10.1071/WF01036

Schmidt, I.T., O'Leary, J.F., Stow, D.A., Uyeda, K.A., Riggan, P.J., 2016. Use of ultra-high spatial resolution aerial imagery in the estimation of chaparral wildfire fuel loads. Environ. Monit. Assess. 188, 697. https://doi.org/10.1007/s10661-016-5656-x

Scott, J.H., 1999. NEXUS: a system for assessing crown fire hazard. Fire Manag. Notes 59, 20–24.

Scott, J.H., Burgan, R.E., 2005. Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model. Fort Collins, CO. https://doi.org/10.2737/RMRS-GTR-153

Scott, J.H., 2006. Comparison of crown fire modeling systems used in three fire management applications. https://doi.org/10.2737/RMRS-RP-58

Sevara, C., Wieser, M., Doneus, M., Pfeifer, N., 2019. Relative radiometric calibration of airborne LiDAR data for archaeological applications. Remote Sens. 11, 945. https://doi.org/10.3390/rs11080945

Silva, C.A., Hudak, A.T., Klauberg, C., Vierling, L.A., Gonzalez-Benecke, C., de Padua Chaves Carvalho, S., Rodriguez, L.C.E., Cardil, A., 2017. Erratum to: Combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon in fast-growing Eucalyptus forest plantation using airborne LiDAR data. Carbon Balance Manag. 12, 14. https://doi.org/10.1186/s13021-017-0082-0

Silva, C.A., Hamamura, C., Valbuena, R., Hancock, S., Cardil, A., Broadbent, E.N., de Almeida, D.R.A., Junio, C.H.L.S., Klauberg, C., Garbow, B., Hillstrom, K., More, J., Craig, M., 2020. rGEDI: NASA's Global Ecosystem Dynamics Investigation (GEDI) data visualization and processing.

Skowronski, N.S., Clark, K.L., Duveneck, M., Hom, J., 2011. Three-dimensional canopy

fuel loading predicted using upward and downward sensing LiDAR systems. Remote Sens. Environ. 115, 703–714. https://doi.org/10.1016/j.rse.2010.10.012

Song, J.H., Han, S.H., Yu, K.Y., Kim, Y.I., 2002. Assessing the possibility of land-cover classification using lidar intensity data. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 34, 259–262.

Srinivasan, S., Popescu, S., Eriksson, M., Sheridan, R., Ku, N.-W., 2015. Terrestrial laser scanning as an effective tool to retrieve tree level height, crown width, and stem diameter. Remote Sens. 7, 1877–1896. https://doi.org/10.3390/rs70201877

Standish, J.T., Manning, G.H., Demaerschalk, J.P., 1985. Development of biomass equations for British Columbia tree species.

Stocks, B.J., Lynham, T.J., Lawson, B.D., Alexander, M.E., Wagner, C.E. Van, McAlpine, R.S., Dubé, D.E., 1989. Canadian Forest Fire Danger Rating System: an overview. For. Chron. 65, 258–265. https://doi.org/10.5558/tfc65258-4

Stratton, R.D., 2006. Guidance on spatial wildland fire analysis: models, tools, and techniques. https://doi.org/10.2737/RMRS-GTR-183

Strunk, J., Temesgen, H., Andersen, H.-E., Flewelling, J.P., Madsen, L., 2012. Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. Can. J. Remote Sens. 38, 644–654. https://doi.org/10.5589/m12-052

Suchar, V.A., Crookston, N.L., 2010. Understory cover and biomass indices predictions for forest ecosystems of the Northwestern United States. Ecol. Indic. 10, 602–609. https://doi.org/10.1016/j.ecolind.2009.10.004

Sullivan, A.L., 2009. Wildland surface fire spread modelling, 1990 - 2007. 1: Physical and quasi-physical models. Int. J. Wildl. Fire 18, 349–368. https://doi.org/10.1071/WF06143

7

Torralba, J., Crespo-Peremarch, P., Ruiz, L.Á., 2018. Assessing the use of discrete, full-waveform LiDAR and TLS to classify Mediterranean forest species composition. Rev. Teledetección 27–40. https://doi.org/10.4995/raet.2018.11106

Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., Etheridge, D., 2012. LiDAR sampling density for forest resource inventories in Ontario, Canada. Remote Sens. 4, 830–848. https://doi.org/10.3390/rs4040830

v

Valbuena, R., Packalén, P., Martn-Fernández, S., Maltamo, M., 2012. Diversity and equitability ordering profiles applied to study forest structure. For. Ecol. Manage. 276, 185–195. https://doi.org/10.1016/j.foreco.2012.03.036

Valbuena, R., Packalen, P., Mehtätalo, L., García-Abril, A., Maltamo, M., 2013. Characterizing forest structural types and shelterwood dynamics from Lorenz-

based indicators predicted by airborne laser scanning. Can. J. For. Res. 43, 1063–1074. https://doi.org/10.1139/cjfr-2013-0147

Valbuena, R., Vauhkonen, J., Packalen, P., Pitkänen, J., Maltamo, M., 2014. Comparison of airborne laser scanning methods for estimating forest structure indicators based on Lorenz curves. ISPRS J. Photogramm. Remote Sens. 95, 23–33. https://doi.org/10.1016/j.isprsjprs.2014.06.002

Valbuena, R., Maltamo, M., Packalen, P., 2016. Classification of forest development stages from national low-density lidar datasets: a comparison of machine learning methods. Rev. Teledetección 15–25. https://doi.org/10.4995/raet.2016.4029

Valbuena, R., Maltamo, M., Mehtätalo, L., Packalen, P., 2017. Key structural features of Boreal forests may be detected directly using L-moments from airborne lidar data. Remote Sens. Environ. 194, 437–446. https://doi.org/10.1016/j.rse.2016.10.024

Varo-Martínez, M.Á., Navarro-Cerrillo, R.M., Hernández-Clemente, R., Duque-Lazo, J., 2017. Semi-automated stand delineation in Mediterranean Pinus sylvestris plantations through segmentation of LiDAR data: The influence of pulse density. Int. J. Appl. Earth Obs. Geoinf. 56, 54–64. https://doi.org/10.1016/j.jag.2016.12.002

Viedma, O., Urbieta, I.R., Moreno, J.M., 2018. Wildfires and the role of their drivers are changing over time in a large rural area of west-central Spain. Sci. Rep. 8, 17797. https://doi.org/10.1038/s41598-018-36134-4

Vierling, L.A., Xu, Y., Eitel, J.U.H., Oldow, J.S., 2013. Shrub characterization using terrestrial laser scanning and implications for airborne LiDAR assessment. Can. J. Remote Sens. 38, 709–722. https://doi.org/10.5589/m12-057

W

Wagner, C.E. Van, 1985. Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Ottawa, Canada.

Wagner, C.E. Van, 1987. Development and structure of the Canadian forest fire weather index system. Ottawa, Canada.

Wagner, W., 2010. Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. ISPRS J. Photogramm. Remote Sens. 65, 505–513. https://doi.org/10.1016/j.isprsjprs.2010.06.007

Watt, P.J., Donoghue, D.N.M., 2005. Measuring forest structure with terrestrial laser scanning. Int. J. Remote Sens. 26, 1437–1446. https://doi.org/10.1080/01431160512331337961

Weise, D.R., Cobian-Iñiguez, J., Princevac, M., 2018. Surface to Crown Transition, in: Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires. Springer International Publishing, Cham, pp. 1–5. https://doi.org/10.1007/978-3-319-51727-8_24-1

Werth, P.A., Potter, B.E., Clements, C.B., Finney, M.A., Goodrick, S.L., Alexander, M.E., Cruz, M.G., Forthofer, J.A., McAllister, S.S., 2011. Synthesis of knowledge

of extreme fire behavior: volume I for fire managers.
https://doi.org/10.2737/PNW-GTR-854

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 Sens. Environ. 124, 730–741. https://doi.org/10.1016/j.rse.2012.06.024

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

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: A review. Remote Sens. Environ. 121, 196–209. https://doi.org/10.1016/j.rse.2012.02.001

𝒳

Xu, K., Gong, Y., Fang, S., Wang, K., Lin, Z., Wang, F., 2019. Radiometric calibration of UAV remote sensing image with spectral angle constraint. Remote Sens. 11, 1291. https://doi.org/10.3390/rs11111291

𝒵

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. Photogramm. Eng. Remote Sens. 77, 281–290. https://doi.org/10.14358/PERS.77.3.281

Zhao, K., Popescu, S., Meng, X., Pang, Y., Agca, M., 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. Remote Sens. Environ. 115, 1978–1996. https://doi.org/10.1016/j.rse.2011.04.001

Zhou, T., Popescu, S., 2019. waveformlidar: An R package for waveform LiDAR processing and analysis. Remote Sens. 11, 2552. https://doi.org/10.3390/rs11212552

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 Sens. Environ. 87, 171–182. https://doi.org/10.1016/S0034-4257(03)00139-1