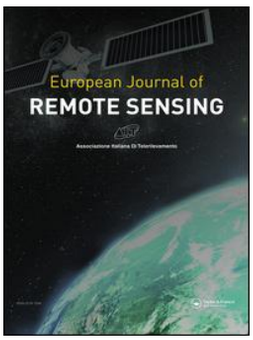

European Journal of Remote Sensing

ISSN: (Print) 2279-7254 (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20

Comparison of ALS based models for estimating aboveground biomass in three types of Mediterranean forest

Juan Guerra-Hernández, Eric Bastos Görgens, Jorge García-Gutiérrez, Luiz Carlos Estraviz Rodriguez, Margarida Tomé & Eduardo González-Ferreiro

To cite this article: Juan Guerra-Hernández, Eric Bastos Görgens, Jorge García-Gutiérrez, Luiz Carlos Estraviz Rodriguez, Margarida Tomé & Eduardo González-Ferreiro (2016) Comparison of ALS based models for estimating aboveground biomass in three types of Mediterranean forest, European Journal of Remote Sensing, 49:1, 185-204, DOI: 10.5721/EuJRS20164911

To link to this article: https://doi.org/10.5721/EuJRS20164911

Published online: 17 Feb 2017.

European Journal of Remote Sensing - 2016, 49: 185-204
doi: 10.5721/EuJRS20164911
Received 13/08/2015 accepted 24/03/2016

European Journal of Remote Sensing
An official journal of the Italian Society of Remote Sensing
www.aitjournal.com

Comparison of ALS based models for estimating aboveground biomass in three types of Mediterranean forest

Juan Guerra-Hernández[1]*, Eric Bastos Görgens[2], Jorge García-Gutiérrez[3], Luiz Carlos Estraviz Rodriguez[4], Margarida Tomé[1] and Eduardo González-Ferreiro[5,6,7]

[1]Grupo ForChange, Centro de Estudos Florestais (CEF), Universidade de Lisboa, Instituto Superior de Agronomía (ISA), Tapada da Ajuda, P-1349-017-Lisboa, Portugal
[2]Laboratório de Economia e Planejamento, Departamento de Engenharia Florestal, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Rodovia MGT 367, Km 583, nº 5000, Alto da Jacuba. B-39100-000-Diamantina, Brazil
[3]Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla, Escuela Técnica Superior de Ingeniería Informática, Reina Mercedes s/n, E-41012-Sevilla, Spain
[4]Grupo de Estudos em Tecnologias LiDAR (GET-LiDAR), Departamento de Ciências Florestais, Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Av. Pádua Dias, 11, B-13418-90-Piracicaba, Brazil
[5]Unidade de Xestión Forestal Sostible (UXFS), Departamento de Enxeñería Agroforestal, Universidade de Santiago de Compostela, Escola Politécnica Superior - R/Benigno Ledo, Campus Universitario, E-27002-Lugo, Spain
[6]Department of Forest Ecosystems and Society (FES), Oregon State University, 321 Richardson Hall, O-97331 Corvallis, Oregon, USA
[7]Laboratory of Applications of Remote Sensing in Ecology (LARSE), US Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, O-97331 Corvallis, Oregon, USA
*Corresponding author, e-mail address: juanguerra@isa.ulisboa.pt

Abstract

This study aimed to develop ALS-based models for estimating stem, crown and aboveground biomass in three types of Mediterranean forest, based on low density ALS data. Two different modelling approaches were used: (i) linear models with different variable selection methods (Stepwise Selection [SS], Clustering/Exhaustive search [CE] and Genetic Algorithm [GA]), and (ii) previously Published Models (PM) applicable to diverse types of forest. Results indicated more accurate estimations of biomass components for pure *Pinus pinea* L. (rRMSE = 25.90-26.16%) than for the mixed (30.86-36.34%) and *Quercus pyrenaica* Willd. forests (32.78-34.84%). All the tested approaches were valuable, but SS and GA performed better than CE and PM in most cases.

Keywords: Biomass components, remote sensing, airborne laser scanning, mediterranean forest, feature selection approaches.

Introduction

Mediterranean forest ecosystems provide multiple wood and non-wood forest products and services that are important for the socioeconomic development of rural areas. Current methods of estimating the variables of interest in this type of forest must be improved to

meet new demands for the type of information required to enable effective and sustainable forest management practises.

In the last few decades, ALS systems have become a viable alternative to traditional field surveys [e.g. Næsset, 2004], which are extremely labour intensive and expensive [Hall et al., 2005]. Countrywide collection of ALS data can reduce costs by encouraging multipurpose ALS flights (unit costs decrease as the scanned surface and number of goals for each flight increase), thereby reducing the overall costs of forest inventories [Nord-Larsen and Riis-Nielsen, 2010; González-Ferreiro et al., 2014; Vauhkonen et al., 2014].

In 2009, the need to obtain an accurate digital elevation model (DEM) led to a national ALS survey being carried out in Spain to map the country with a theoretical average point density data equal to 0.5 points m^{-2} (Plan Nacional de Ortofotografía Aerea: PNOA project). The amount of data available from ALS surveys is expected to increase in the next few years, as the PNOA project has scheduled ALS flights every 6 years. This further drives the need for robust ALS-based models for use in different surveys and with different sensor systems.

Discrete return ALS systems have been successfully used to estimate aboveground biomass at stand level over a wide range of forest types: Temperate [e.g Hall et al., 2005], Boreal [e.g. Næsset, 2002, 2004; Treitz et al., 2012], Atlantic [e.g. González-Ferreiro et al., 2012, 2014], Tropical [e.g. Asner et al., 2012; Cao et al., 2014], Alpine [e.g. Montaghi et al., 2013; Corona et al., 2014] and Mediterranean forests [García et al., 2010; González-Olabarria et al., 2012; Ruiz et al., 2014; Montealegre et al., 2016; Chirici et al., 2016]. Although several types of forest have been surveyed using ALS technology, this works present ALS-based models for estimating aboveground biomass that usually differ in terms of precision, form and the used ALS metrics [Li et al., 2008; Bouvier et al., 2015; Véga et al., 2016]. Some authors have suggested that most predictive ALS–based models should not include more than three variables that generally represent some form of three group metrics: (i) one related to height, (ii) one related to canopy cover, and (iii) one describing the variation in the height distribution [Lefsky et al., 2005; Li et al., 2008; White et al., 2013]. Recent ALS studies [Cao et al., 2014; Bouvier et al., 2015] also identified advantages over strata-specific prediction models. These advantages are more obvious in the wall-to-wall mapped area–based predictions [Latifi et al., 2015] and must be checked for Mediterranean forest structures, of which relatively few studies have been carried out.

Two possible approaches can be used when no models are available for the particular type of forest under consideration. One approach is to develop specific models for each forest type. In the present study, we developed linear models by using three different variable selection methods: Stepwise Selection (SS), the Clustering/Exhaustive search procedure (CE), and Genetic Algorithm regression (GA). The other approach is to take advantage of the relationships already determined in previously Published Models (PM) and which should be applicable to diverse types of forest with some modifications in the parameters. We applied the models of Lefsky et al. [2002], Li et al. [2008] and Zonete et al. [2010].

In this study, we aimed to develop parsimonious and robust ALS-based models to estimate aboveground biomass components for three types of Mediterranean forest not previously studied: pure *P. pinea* forest, mixed *P. pinea* forest, and *Q. pyrenaica* forest. For this purpose, we compared the performance of four different methods (SS, CE, GA and PM) in order to select the best predictors extracted from low density ALS data obtained by the PNOA project.

Materials and Methods

Study Site

This study was conducted in the 'Tudia y sus Faldas' forest (Fig. 1), located near the town of Monesterio in the province of Badajoz (southwest Spain). The forested area, classified as public utility forest number MUP1 (*Monte de Utilidad Pública número 1*), covers an area of 748.20 ha. The forest is representative of *P. pinea* forest in SW Spain, i.e. it is characterized by the dominance of pure *P. pinea* stands and mixed forest of *P. pinea* stands associated with *Pinus pinaster* Ait. and *Q. pyrenaica*. The forest also includes a small proportion of pure *Q. pyrenaica* stands. The study area is characterized by very steeply sloping terrain (average slope 25.5%) at an elevation ranging from 300 to 1100 m above sea level (Fig.1).

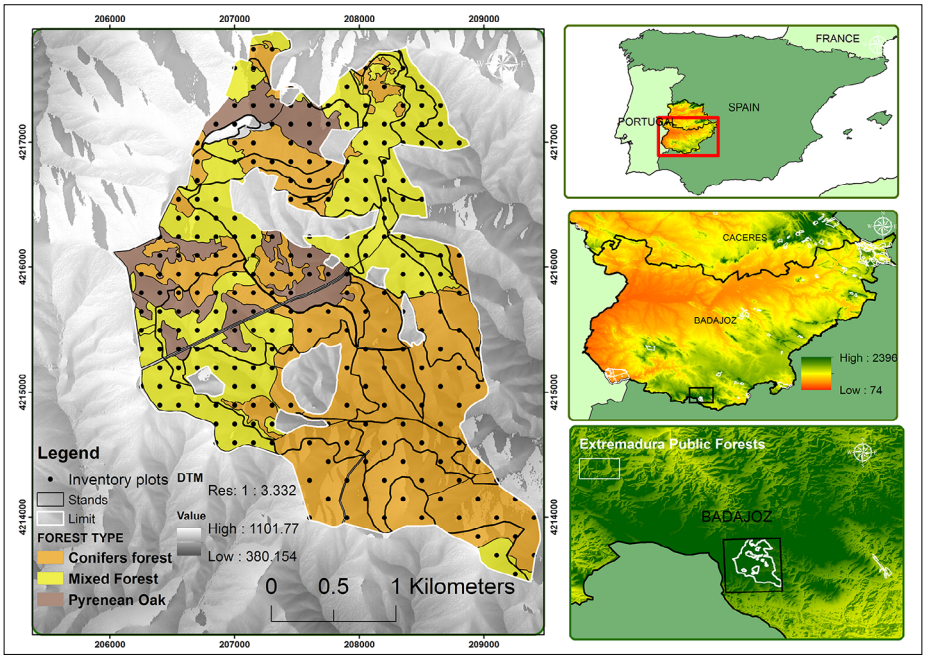

Figure 1 - Boundary of the 'Tudia y sus Faldas' forest study site (blank line) and the locations of the different types of forest plots (black dots).

Field data

Field data were obtained from a forest inventory carried out by the Extremadura Forest Service for forest management purposes. In total, 178 circular sample plots of radius 11 m (approx. 380 m^2) were measured in the study area, between July and August 2010. A LEICA GX1230 (dual frequency real time kinematic receiver with a planimetric precision of ±5 mm + 0.5 ppm and an altimetric precision of ±10 mm + 0.5 ppm) was used, along with a metal detector, to relocate the centre of each plot (marked with iron poles). At each point, GPS signals were logged using a roving receiver with an external antenna (ATX1230 GG), and the recordings were post-processed with correction data retrieved from the fixed base station in Llerena (Badajoz) (station number 355919, latitude 4236374.101 m, longitude 236491.045 m, ETRS89-30 Coordinate system and elevation 642.392 m), to yield the

plot positions. The average accuracy of the relative positioning of the field plots was approximately 0.32 m.

The forest type was considered dominant if the basal area of the dominant species represented more than 70% of the total basal area within the plot. Following this criterion, 120 plots were classified as pure *P. pinea* stands, 39 plots as mixed forest stands and 19 plots as pure *Q. pyrenaica* stands.

Species-specific allometric equations were used to estimate individual tree stem biomass, branch and foliage biomass, and aboveground biomass for *P. pinea*, *P. pinaster* [Ruiz-Peinado et al., 2011] and *Q. pyrenaica* [Ruiz-Peinado et al., 2012]. The aboveground biomass, stem biomass and crown biomass (branches and foliage) were determined by adding the values obtained for the individual trees.

The field measurements (heights and diameters) were used to estimate the following stand variables for each plot (on a per hectare basis): mean height (H_m), dominant height (H_0), stand basal area (G), stand volume (V), stand stem biomass (W_s), stand crown biomass (W_{cw}) and aboveground stand biomass (W_a) (Tab. 1).

Table 1 - Summary of the mean values and range of the main stand parameters and biomass components in the sample plots.

	Pure *P. pinea* n=120			Mixed forest n=39			Pure *Q. pyrenaica* n=19		
Stand descriptive variables	**Range**		**Mean**	**Range**		**Mean**	**Range**		**Mean**
N	26	1026	225	53	1605	360	26	1368	486
G	1.1	28.8	14.0	0.9	37.6	12.3	0.3	19.2	7.6
Vcc	4.1	141.8	63.1	2.4	193.4	57.7	2.0	96.4	38.8
Vsc	3.1	102.4	45.5	2.2	138.4	44.7	1.9	91.9	35.8
Hm	3.8	13.3	9.0	2.1	13.3	7.7	5.6	8.8	7.2
H_0	4.0	13.8	10.0	5.3	14.7	10.4	5.8	11.2	8.1
Stand biomass variables	**Range**		**Mean**	**Range**		**Mean**	**Range**		**Mean**
Ws	3.8	80.3	34.7	2.5	87.1	32.1	0.8	56.3	20.1
Wcw	2.2	62.5	28.9	2.4	43.4	19.3	0.4	21.9	10.2
Wa	4.1	142.9	63.6	4.8	126.8	51.4	1.1	75.2	30.3

N number of trees (trees ha^{-1}); *G*, Basal area; (m^2 ha^{-1}); *Vcc*, Volume over bark (m^3 ha^{-1}); *Vsc*, Volume under bark (m^3 ha^{-1}); H_0, Dominant height (m); *Hm*, Mean height (m); *Ws*, Stem biomass (Mg ha^{-1}); *Wcw*, Crown biomass (Mg ha^{-1}); *Wa,* Aboveground biomass (Mg ha^{-1}).

The ALS data and explanatory variables

ALS data were acquired between July and August 2010 for the PNOA project, funded by the Spanish Ministerio de Fomento (*Dirección General del Instituto Geográfico Nacional, IGN,* and *Centro Nacional de Información Geográfica, CNIG*). The laser equipment used was a LEICA ALS50 sensor operated with pulse repetition rate of 83 kHz, maximum scan frequency of 32.1 Hz, maximum scan angle of ± 50º and an average flying height of 2,866

m above sea level, which yielded a theoretical density of 0.5 first returns per square metre. The equipment operates at a wavelength of 1064 nm and is capable of registering up to 4 returns per pulse. Summary statistics of first return density per square metre within plots are as follows: average = 1.76, minimum = 1, maximum = 41 and standard deviation = 1.62.

Table 2 - Summary of ALS metrics extract by software FUSION for each plot. See McGaughey (2014) for more details of how to calculate each ALS metrics.

	ALS metrics	Description
(A) Height metrics		
(A.1) metrics expressing the central trend in ALS height distribution		
	h_{mean}	mean
	h_{mode}	mode
(A.2) metrics expressing the dispersion of ALS height distribution		
	h_{SD}	standard deviation
	h_{VAR}	variance
	h_{AAD}	absolute average deviation
	h_{IQ}	interquartile range
	h_{CV}	coefficient of variation
	$h_{max,}$ h_{min}	maximum and minimum
(A.3) metrics expressing the shape of ALS height distribution		
	h_{Skw}	skewness
	h_{Kurt}	kurtosis
	CRR	canopy relief ratio (((mean height- min height) / (max height– min height))
(A.4) percentiles of the ALS height distribution		
	h_{01}, h_{10}… $h_{95,}$ h_{99}	1^{th},5^{th}, 10^{th}, 20^{th}, 25^{th}, 30^{th}, 40^{th}, 50^{th}, 60^{th}, 70^{th}, 75^{th}, 80^{th}, 90^{th}, 95^{th}, 99^{th} percentiles
(B) Canopy cover metrics		
(B.1) fixed *HBT*		
	CC	percentage of first returns above 2.00/total all returns
	PARA2	percentage of all returns above 2.00/total all returns
	ARA2/TFR	ratio between all returns above 2.00 and total of first returns
(B.2) variable *HBT*		
	PFRAM	percentage of first returns above mean/total all returns
	PARAM	percentage of all returns above mean/total all returns
	PARAMO	percentage of all returns above mode/total all returns
	PFRAMO	percentage of first returns above mode/total all returns
	ARAM/TFR	ratio between all returns above mean and total of first returns
	ARAMO/TFR	ratio between all returns above mode and total of first returns

ALS metrics are descriptive structure statistics calculated from the raw ALS point cloud. The metrics for the 178 plots were calculated using the FUSION ALS Toolkit [McGaughey, 2014]. Possible outliers were removed from the ALS point cloud by returns from the dataset on the basis of standard deviation of the elevations. The ALS point clouds were then filtered and interpolated to generate a Digital Terrain Model (DTM) of cell size 1 m. ALS metrics were computed for each circular plot after normalising the data by subtracting the DTM. The ALS metrics were computed considering the first returns and all returns independently [Naesset, 2002]. The minimum height threshold (*MHT*), which is commonly specified as the lower boundary for calculating height metrics (central tendency, dispersion, shape and percentile statistics), was established as 2 m. The height break threshold (*HBT*), which is the limit for separating the point cloud data into two sets to separate canopy returns from the under canopy returns, in order to compute canopy cover metrics, was also established as 2 m following Næsset [2002]. In total, 36 metrics (including height, and canopy cover) were extracted from ALS pulses and used as regressors for statistical analyses. For further details of the procedure used to obtain the ALS metrics, see the steps outlined in González-Ferreiro et al. [2012]. The ALS metrics and the corresponding descriptions are summarised in Table 2.

Aboveground Biomass Modelling

The multiple linear regression model (MLR) used to establish empirical relationships between field measurements and ALS variables is defined as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon \quad [1]$$

where Y represents field variables, W_a (Mg ha^{-1}), W_s (Mg ha^{1}), W_{cw} (Mg ha^{1}); $X_1, X_2, \ldots, X_n$ are metrics derived from ALS data set; and Ɛ is a vector of true but unknown residuals whose elements have zero expected value and are independently and identically distributed. Four methods were used to select the ALS metrics to be used as independent variables in Equation 1: stepwise selection (SS), clustering and exhaustive search (CE), genetic algorithm (GA) and application of predictor metrics from three previously published general linear models (PM). A maximum of three explanatory variables was considered, in order to yield robust parsimonious models.

Comparison of the estimates for the selected models was based on the adjusted coefficient of determination (adj. R^2) and the relative Root Mean Square Error (rRMSE, see Equation 2). The residual normality was tested using the Shapiro-Wilk Test [Shapiro et al., 1968]. All statistical analyses were performed using R software [R Core Team, 2014], and the *leaps* package was also used for the SS and CE analyses.

$$rRMSE = \frac{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}}{\bar{y}} * 100 \quad [2]$$

where y_i is the observed value, $\hat{y}_i$ is the estimated value, $\bar{y}$ is the mean observed value and n is the number of observations.

Stepwise Selection (SS)

Stepwise selection is an automatic procedure commonly used to develop linear models based on ALS metrics [García et al., 2010; González-Ferreiro et al., 2012]. Stepwise selection fits the regression model by adding/dropping covariates one at a time on the basis of a specified criterion. At each stage, the improvement is tested by its significance in a sequence of tests, e.g. *F*-tests, *t*-tests, AIC and others. Collinearity between regressors was avoided by checking the condition index (CI) and the variance inflation factor (VIF) at the end of each stepwise procedure (Fig. 2). In this study, regressors with a CI above 30 or VIF above 10 were disregarded [Belsley et al., 2005; Stevens, 2012].

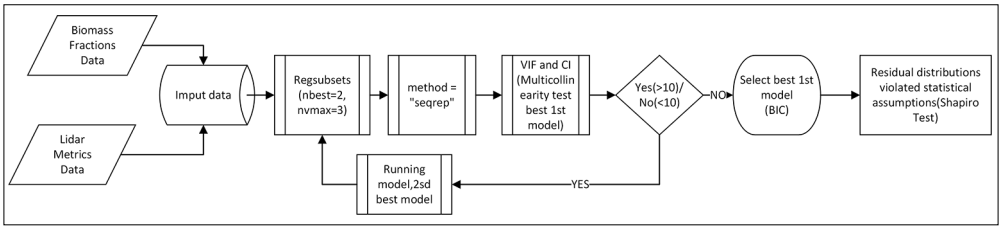

Figure 2 - Stepwise selection (SS) method for estimating biomass components at stand level.

Clustering and Exhaustive Search (CE)

This method was adapted from the statistical analysis carried out by Stephens et al. [2012]. Before fitting multiple regression models for biomass components and ALS metrics, the metrics were grouped into ten colinear groups on the basis of the correlation matrix. One variable from each group was selected as a regressor. The code performs an exhaustive search (or similar) for the best subsets of the variables in *X* for predicting *Y*, using the branch-and-bound algorithm. All possible combinations of selected metrics are regressed against biomass variables (Fig. 3). Regressors with CI > 30 or VIF > 10 were also disregarded.

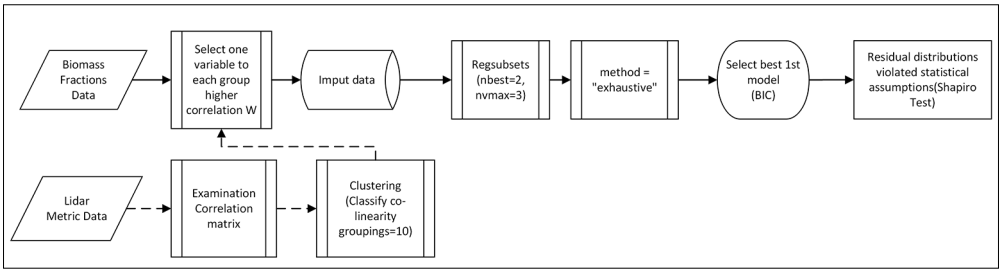

Figure 3 - Clustering/Exhaustive search method (CE) for estimating biomass components at stand level.

Genetic Algorithm (GA)

The third method used a genetic algorithm implemented with the Watchmaker framework [Dyer, 2006] to select the best metrics (Fig. 4). The GA started with a random population of possible individuals defined as a set of binary (0, 1) values associated with the possible predictors. For each generation, the algorithm selected the best individuals and combined

them to produce new individuals (offspring). A classical roulette selection technique [Holland, 1992] was used to choose two parents for each new individual, in order to generate each new offspring. This technique assigned a probability to every individual (potential parent) according to Equation [2], where is the fitness reached by the *i*th individual, and selected the first individual that caused the accumulated probability be greater than a random value between 0 and 1.

$$p_i = \frac{fi}{\sum_i^N fi} \quad [3]$$

This process was repeated until completion of the production of new offspring for each generation, taking into account that the two best individuals from the previous generation always survived (elitism). New individuals were the result of a uniform crossover operator [Holland, 1992]. After the crossover, a mutation operator was applied with a given probability, mutating or removing a randomly selected predictor. A fitness function evaluated each individual at the beginning of each generation. The fitness function assigned the goodness of fit of an individual according to the quality of an MLR model developed with the selected variables. Finally, collinearity between the explanatory variables was checked, and models with CI > 30 or VIF > 10 were disregarded.

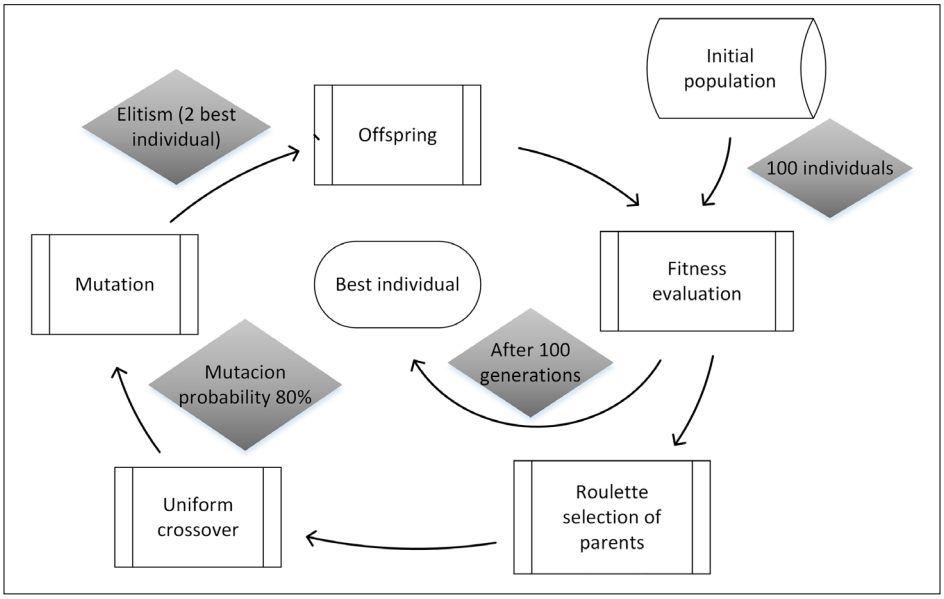

Figure 4 - Genetic algorithm (GA) method for estimating biomass components at stand level.

Previously Published Models (PM)

The empirical relationship between ALS characteristics and stand biomass suggests that common models may be widely applicable to diverse forest types. We used the same relationships (i.e. ALS explanatory metrics) determined in previously published models,

but adjusted the coefficients to our dataset. We evaluated the following PM:
I) *Lefsky model*
Lefsky et al. [2002] studied three distinct sites: boreal coniferous forest (dominated by *Picea mariana* (Mill.) B.S.P.), temperate coniferous forest (dominated by *Pseudotsuga menziesii* (Mirb.) Franco) and temperate deciduous forest (mixed deciduous forest with an overstory dominated by *Liriodendron tulipifera* L.). The reported model for estimating biomass was expressed as follows:

$$W = \beta_1 \cdot \left(h_{mean}\right)^2 + \beta_2 \cdot \left(CC * h_{mean}\right) + \varepsilon \quad [4]$$

where is the mean canopy height and *CC* is a cover metric. *CC* is calculated as follows: ((*Nv* > *HBT*)/*N*)/100, where: *Nv* = Number of first vegetation returns above a specific *HBT*, *N* = total number of returns.
II) *Li model*
Li et al. [2008] selected h_{mean}, h_{CV} and *CC* as the best predictive variables for three different types of forest. The study sites were located in the west of the State of Washington (US) (dominated by *Pseudotsuga menziesii* and *Tsuga heterophylla* (Rat) Sarg.), the Eastern Cascade Mountains in the State of Washington (US) (dominated by *Pinus ponderosa* Dougl. ex Laws), and on the Kenai peninsula, Alaska (US) (dominated by *Betula papyrifera* Marshal. and *P. mariana*). The reported model for estimating biomass was expressed as follows:

$$W = \beta_1 \cdot h_{mean} + \beta_2 \cdot h_{CV} + \beta_2 \cdot CC + \varepsilon \quad [5]$$

III) *Zonete model*
Zonete et al. [2010] proposed the use of h_{30} and h_{90} to simulate the biological condition concerning respectively site quality and stand density in a *Eucalyptus* spp. forest plantation in Brazil. The reported model for estimating biomass was expressed as follows:

$$W = \beta_1 \cdot h_{30} + \beta_2 \cdot h_{90} + \beta_2 + \varepsilon \quad [6]$$

Results

The models selected for each type specific regression model and forest type are shown in Tables 3, 4 and 5 respectively.

Pure P. pinea forests

Regression models for three biomass fractions (W_s, W_{cw}, W_a) in pure *P. pinea* stands yielded adj. R^2 values ranging from 0.73 to 0.74, with the SS and GA methods, and from 0.70 to 0.73, with the CE approach. In terms of rRMSE values, the range was slightly lower with the SS and GA approaches (25.89 to 26.16%) than with the CE procedure (27.02 to 27.30%).
The models selected from literature (PM) yielded adj. R^2 values ranging from 0.67 to 0.70 for the *Lefsky* model, from 0.64 to 0.73 for the *Li* model, and from 0.46 to 0.64 for the *Zonete* model. For biomass estimation, the methods (SS and GA) yielded slightly better results (as indicated by adj. R^2 and rRMSE) than the PM, except when using the *Li* model. The same predictors were selected with SS, GA and CE for all biomass components. The

best predictors with SS and GA were the mean height of the vegetation returns (h_{mean}), the proportion of first returns above 2 m (*CC)* and a metric expressing the shape of ALS height distribution (*CRR*). In W_{cw} modelling, the inclusion of h_{CV} was statistically significant ($p <$ 0.01) and improved the goodness-of-fit. The model proposed by Zonete et al. [2010] proved to be the least precise under these conditions, although h_{30} was significant for all biomass components. In the model proposed by Lefsky et al. [2002], the quadratic term in h_{mean} was also significant, suggesting a slightly nonlinear relationship between W_s and h_{mean}.

Table 3 - Summary of the biomass component prediction models and plot-level accuracy assessment obtained for pure *P. pinea* forest in each modelling method.

Method	Final model	R^2_{adj}	rRMSE (%)	Shapiro (*p*-value)
Dependent variable: W_s				
SS/GA	$-34.43^{***} + 4.59^{***} h_{mean} + 29.89^{**} CRR + 0.31^{***} CC$	0.74	26.16	0.146
CE	$-23.51^{***} + 5.55^{***} h_{mean} + 0.18^{**} PFRAMO + 0.39^{***} ARAM/TFR$	0.73	27.30	0.178
Lefky	$2.01^{*} + 0.16^{**} (h_{mean}{}^2) + 0.051^{***} (CC{*}h_{mean})$	0.71	27.67	0.043
Li	$-23.73^{***} + 5.82^{***} h_{mean} - 5.68^{ns} hcv + 0.30^{***} CC$	0.73	26.94	0.227
Zonete	$-11.45^{ns} + 4.20^{**} h_{30} + 2.34^{*} h_{90}$	0.59	32.90	0.289
Dependent variable: W_{cw}				
SS/GA	$-21.57^{***} + 2.45^{***} h_{mean} + 25.79^{**} CRR + 0.28^{***} CC$	0.73	25.89	0.071
CE	$-10.71^{***} + 3.16^{***} h_{mean} + 0.10^{**} PFRAMO + 0.38^{***} ARAM/TFR$	0.70	27.22	0.217
Lefsky	$5.70^{***} + 0.058^{ns} (h_{mean}{}^2) + 0.042^{***} (CC{*}h_{mean})$	0.71	27.03	0.057
Li	$-9.69^{***} + 3.46^{***} h_{mean} - 13.13^{*} h_{cv} + 0.26^{***} CC$	0.73	25.92	0.109
Zonete	$-0.068^{ns} + 3.58^{***} h_{30} + 0.77^{ns} h_{90}$	0.57	32.76	0.541
Dependent variable: W_a				
SS/GA	$-54.61^{***} + 7.18^{***} h_{mean} + 51.87^{**} CRR + 0.59^{***} CC$	0.74	25.90	0.075
CE	$-34.46^{***} + 8.75^{***} h_{mean} + 0.31^{**} PFRAMO + 0.74^{***} ARAM/TFR$	0.71	27.02	0.105
Lefsky	$7.94^{*} + 0.21^{*} (h_{mean}{}^2) + 0.093^{***} (CC{*}h_{mean})$	0.70	27.54	0.026
Li	$-34.99^{***} + 9.29^{***} h_{mean} - 13.15^{ns} h_{cv} + 0.57^{***} CC$	0.72	26.73	0.095
Zonete	$-12.49^{ns} + 7.30^{***} h_{30} + 3.52^{*} h_{90}$	0.57	33.12	0.349

Stepwise Selection (SS), Clustering/Exhaustive search (CE), Genetic Algorithm (GA); W_s (Mg ha^{-1}): stem biomass; W_{cw} (Mg ha^{-1}): crown biomass; W_a (Mg ha^{-1}): aboveground biomass. Pr(> |*t*|) *p* = ≤ 0.0001 '***' < 0.001 '**' < 0.01 '*' < 'ns'

Mixed forest

The adj. R^2 for mixed forest with the SS, GA and CE methods ranged from 0.68 to 0.79, from 0.65 to 0.79 and from 0.69 to 0.72, respectively, whereas the rRMSE ranged from 30.86 to 36.93%, from 30.83 to 38.72% and from 35.00 to 36.34%, respectively. The rRMSE of the fitted models was higher in mixed forest than in pure forest in all approaches. PM yielded adj. R^2 values ranging from 0.52 to 0.70 for the *Lefsky* model, from 0.64 to 0.74 for the *Li*

model, and from 0.46 to 0.64 for the *Zonete* model. For biomass estimation, SS and GA methods yielded slightly better results (as indicated by adj. R^2 and rRMSE) than PM.
In mixed forest, SS and GA indicated the same metrics for explaining W_s (h_{10}, h_{IQ}, *PFRAM*). As in pure stands, h_{mean} was also the best predictor for estimating the biomass fractions (W_s, W_a) with the CE method. In W_{cw} modelling, GA and CE indicated that the best models were a combination of h_{mode} with two variables related to a measure of height variation (h_{IQ}, h_{AAD}). SS and GA indicated that the best model for estimating W_a included a low height percentile (h_{10}), the coefficient of Kurtosis (h_{Kurt}) and two canopy cover metrics (*PARAM* and *CC*). Only in the cases of W_{cw} and W_a modelling h_{CV} was statistically significant ($p < 0.001$ and $p < 0.01$, respectively). PM showed that h_{30} (*Zonete* model) and $CC*h_{mean}$ (*Lefsky* model) were significant for all biomass components.

Table 4 - Summary of the models predicting biomass components and of the plot-level accuracy assessment obtained for mixed forest in each modelling method.

Method	Final models	R^2_{adj}	rRMSE (%)	Shapiro (*p*-value)
Dependent variable: W_s				
SS/GA	$-27.97^{***} + 7.02^{***} h_{10} + 4.19^{***} h_{IQ} + 0.60^{***} PFRAM$	0.78	32.50	0.523
CE	$-22.86^{***} + 5.90^{***} h_{mean} + 0.47^{**} PFRAM$	0.71	37.82	0.001
Lefsky	$-0.68^{ns} + 0.157^{*} (h_{mean}{}^2) + 0.055^{***} (CC* h_{mean})$	0.70	38.42	0.001
Li	$-20.01^{ns} + 6.22^{***} h_{mean} - 24.58^{ns} h_{cv} + 0.29^{**} CC$	0.72	36.74	0.005
Zonete	$-18.68^{*} + 4.16^{**} h_{30} + 3.01^{ns} h_{90}$	0.64	41.83	0.040
Dependent variable: W_{cw}				
SS	$-7.51^{ns} + 6.37^{***} h_{05} - 3.21^{**} h_{kurt} + 0.57^{***} PARAM$	0.68	36.93	0.564
CE	$-18.41^{**} + 1.78^{***} h_{mode} - 59.38^{**} h_{cv} + 3.01^{**} h_{IQ}$	0.69	36.34	0.395
GA	$-18.95^{**} + 1.60^{**} h_{mode} - 61.85^{**} h_{cv} + 6.05^{*} h_{ADD}$	0.65	38.72	0.168
Lefsky	$3.16^{ns} + 0.052^{ns} (h_{mean}{}^2) + 0.057^{**} (CC* h_{mean})$	0.52	43.61	0.720
Li	$-4.86^{ns} + 2.34^{***} h_{mean} - 41.78^{**} h_{cv} + 0.19^{**} CC$	0.64	39.59	0.124
Zonete	$-0.58^{ns} + 3.02^{**} h_{30} + 0.33^{ns} h_{90}$	0.46	46.81	0.546
Dependent variable: W_a				
SS	$-12.18^{*} + 15.25^{***} h_{10} - 12.19^{***} h_{kurt} + 0.60^{***} CC$	0.79	30.86	0.313
CE	$-5.36^{**} + 7.85^{***} h_{mean} - 59.49^{*} h_{cv} + 0.72^{**} PFRAM$	0.72	35.00	0.099
GA	$-10.37^{**} + 14.15^{***} h_{10} - 11.30^{***} h_{kurt} + 1.49^{***} PARAM$	0.79	30.83	0.265
Lefsky	$2.48^{ns} + 0.21^{ns} (h_{mean}{}^2) + 0.083^{***} (CC* h_{mean})$	0.67	38.51	0.011
Li	$-15.15^{ns} + 8.56^{***} h_{mean} - 66.36^{*} h_{cv} + 0.49^{**} CC$	0.74	33.66	0.018
Zonete	$-19.25^{ns} + 7.18^{**} h_{30} + 3.34^{ns} h_{90}$	0.61	41.47	0.355

Stepwise Selection (SS), Clustering/Exhaustive search (CE), Genetic Algorithm (GA); W_s (Mg ha^{-1}): stem biomass; W_{cw} (Mg ha^{-1}): crown biomass; W_a (Mg ha^{-1}): aboveground biomass. Pr(> |*t*|) $p \leq 0.0001$ '***' < 0.001 '**' < 0.01 '*' < 'ns'

Q. pyrenaica forest

GA yielded higher correlations (adj. R^2 = 0.82-0.83) in pure *Q. pyrenaica* forest than in mixed and pure *P. pinea* forests. GA and SS yielded improvements in adj. R^2 and rRMSE values, relative to CE in W_s and W_a. The GA method yielded the model that performed best in terms of adj. R^2 and rRMSE for W_{cw}. The PM yielded adj. R^2 values ranging from 0.51 to 0.54 for the *Lefsky* model, from 0.46 to 0.52 for the *Li* model, and from 0.55 to 0.62 for the *Zonete* model. The model proposed by Zonete et al. [2010] proved to be the most accurate for this type of forest.

Height metrics (h_{10}, h_{25}) calculated from the point cloud were best selected by GA for all the biomass components. Although h_{mean} and h_{30} were significant using the models published by Li et al. [2008] and Zonete et al. [2010], the proportion of variation explained by the regressions was lower (around 31 and 21%, respectively) than in, for example, the models selected by SS and GA in W_s modelling.

Table 5 - Summary of the biomass components prediction models and of the plot-level accuracy assessment obtained for pure *Q. pyrenaica* forest in each modelling method.

Method	Final models	R^2_{adj}	rRMSE (%)	Shapiro (*p*-value)
Dependent variable: W_s				
SS/GA	$-2.52^{ns} - 25.90^{***} h_{10} + 23.49^{***} h_{25}$	0.83	34.84	0.4044
CE	$119.24^{***} - 69.65^{*} h_{min} + 10.00^{**} h_{30}$	0.68	48.52	0.6195
Lefsky	$-2.51^{ns} + 0.24^{ns} (h_{mean}{}^2) + 0.039^{*}(CC^{*} h_{mean})$	0.54	57.81	0.2331
Li	$-21.86^{ns} + 6.45^{**} h_{mean} - 5.82^{ns} h_{cv} + 0.12^{ns} CC$	0.52	59.46	0.227
Zonete	$-21.45^{*} + 8.52^{***} h_{30} + 0.41^{ns} h_{90}$	0.62	52.78	0.2888
Dependent variable: W_{cw}				
SS	$-3.76^{ns} - 0.79^{*} PARAM + 0.71^{***} ARAM/TFR$	0.70	42.16	0.9814
CE	$2.14^{*} - 0.12^{**} PARAMO + 0.28^{***} ARAM/TFR$	0.72	38.65	0.2171
GA	$-2.63^{ns} - 8.90^{**} h_{10} + 8.08^{***} h_{25}$	0.82	32.78	0.0709
Lefsky	$-0.79^{ns} - 0.045^{ns} (h_{mean}{}^2) + 0.025^{**} (CC^{*} h_{mean})$	0.51	53.24	0.9643
Li	$-3.11^{ns} + 1.65^{*} h_{mean} - 9.61^{ns} h_{cv} + 0.09^{ns} CC$	0.46	56.36	0.983
Zonete	$-4.17^{ns} + 3.74^{**} h_{30} - 0.50^{ns} h_{90}$	0.55	46.81	0.5412
Dependent variable: W_a				
SS/GA	$-5.11^{ns} - 34.80^{***} h_{10} + 31.58^{***} h_{25}$	0.83	33.92	0.5381
CE	$7.95^{ns} - 0.54 PARAMO^{**} + 1.01^{***} ARAM/TFR$	0.70	45.15	0.6393
Lefsky	$-1.71^{ns} + 0.19^{ns} (h_{mean}{}^2) + 0.064^{*} (CC^{*} h_{mean})$	0.54	55.70	0.5452
Li	$-24.97^{ns} + 8.11^{**} h_{mean} - 15.44^{ns} h_{cv} + 0.21^{ns} CC$	0.50	57.70	0.7
Zonete	$-25.61^{*} + 12.25^{***} h_{30} - 0.09^{ns} h_{90}$	0.62	51.65	0.9245

Stepwise Selection (SS), Clustering/Exhaustive search (CE), Genetic Algorithm (GA); W_s (Mg ha^{-1}): stem biomass; W_{cw} (Mg ha^{-1}): crown biomass; W_a (Mg ha^{-1}): aboveground biomass. Pr(> |*t*|) $p \leq 0.0001$ '***' < 0.001 '**' < 0.01 '*' < 'ns'

Figure 5 shows the field-measured versus ALS-estimated (using the best linear model in terms of rRMSE) stem, crown, and aboveground biomass components.

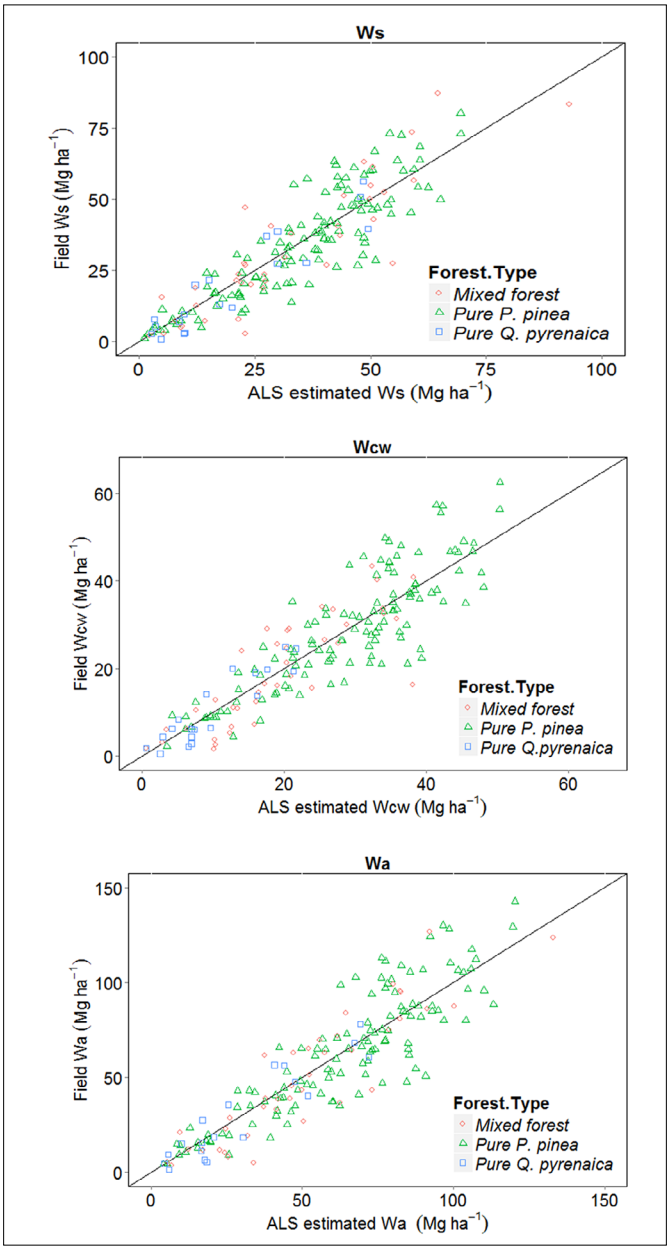

Figure 5 - Scatterplots of the field-measured biomass components against the most accurate model-estimated values of the biomass components in the sample plots for the different types of forest. W_s, Stem biomass. W_{cw}, Crown biomass, W_a, Aboveground biomass.

Discussion

The study findings show that biomass stocks in Mediterranean forest can be modelled using countrywide low density ALS data, with a precision comparable to that of other studies published in the relevant international literature. All models developed for estimating stand biomass performed similarly in terms of model precision in the case of pure *P. pinea* stands and mixed forest, except for the model proposed by Zonete et al. [2010], which performed less well. The study shows that in *Q. pyrenaica* stands, the use of PM representing all forest types may produce more bias than models developed for the specific forest types.

The SS and GA methods yielded the best model precision, although the differences between the latter approaches and the CE method were always equal or less than 5.32% in terms of rRMSE, for pure and mixed stands. García-Gutiérrez et al. [2014] showed that the non-parametric method (GA) performs better than SS, but often selecting a larger number of variables, which although not collinear would make models more complex for the final users. Thus, although GA may finally yield better results, no substantial differences were found when the number of variables was limited to three independent variables.

Regarding the PM models, the model reported by Li et al. [2008] explained a similar amount of variation as explained by SS, GA and CE for pure and mixed stands. The results of our study are consistent with the approaches used by other authors [Lefsky et al., 2005], in which measures of position, canopy cover and height variation were found to be suitable for explaining most of the variability in forest attributes. However, PM such as those reported by Lefsky et al. [2002] and Li et al. [2008] failed the test of normality for W_s and W_a in mixed forest.

ALS studies for assessing biomass components in many types of forest ecosystems report reliable results with acceptable uncertainty estimates [Maltamo et al., 2014]. The accuracy of the results depends on the difference in the field of view of the sensor and the laser point density (among other factors), as well as the slope and vegetation structure [Valbuena et al., 2011] and also plot size [Zolkos et al., 2013; Ruiz et al., 2014], field measurements and allometric equations [Zhao et al., 2012].

For pure *P. pinea* forests, the W_a model precision yielded by the methods compared was similar, in terms of adj. R^2 and rRMSE, to the levels of precision reported by González-Ferreiro et al. [2012] for *P. radiata* in an Atlantic forest in Spain (adj. $R^2 = 0.75$, rRMSE = 26.75%, plot size = 225 m^2) and by Stephens et al. [2012] for *P. radiata* forests in New Zealand ($R^2 = 0.81$, rRMSE = 22%, plot size = 600 m^2). The values obtained in the present study were slightly lower, in terms of adj. R^2 and rRMSE, than those reported by Treitz et al. [2012] for a range of boreal conifer forest types in Canada (adj. $R^2 = 0.93$, rRMSE = 11.05%, plot size = 400 m^2) and by Cao et al. [2014] for conifer species including *Pinus massoniana* Lamb. and *Cunninghamia lanceolata* (Lamb.) Hook.) (adj. $R^2 = 0.84$, rRMSE = 15.84%, plot size = 900 m^2).

The results of the present study showed that a metric expressing the central tendency of ALS heights (h_{mean}) was the best single predictor for all biomass components analysed. The *CC* also significantly improved the fit in two of the above-mentioned studies [Stephens et al., 2012, Cao et al., 2014]. Other authors [Lefsky et al., 2002; Lefsky et al., 2005; Ni-Meister et al., 2010; Asner et al., 2012; Bouvier et al., 2015] eliminated the search step to identify the best metric, but included h_{mean}, which is more sensitive to changes in both the vertical arrangement of canopy elements and the degree of canopy openness (tree

density). The inclusion of a third significant variable (*CRR*), related to the shape of ALS height distribution, explained 0.1% (W_s) or 0.2% (W_a) more of the variance than the model proposed by Li et al. [2008]. Our results demonstrate that no substantial improvement in fit was achieved by adding more than 2 variables in the final model for pure stands. According to the results obtained for other pure coniferous forests, the set of models confirmed that the combination of mean height and canopy cover represents a sufficient and concise quantitative description of a homogeneous vertical structure.

In the case of mixed forest, the best values for biomass component modelling, in terms of adj. R^2 (0.69-0.79), were slightly higher than those obtained for mixed forests by Cao et al. [2014] (adj. R^2 = 0.67-0.75, exclude root and foliage biomass). In W_a modelling, the adj. R^2 values achieved were similar to those reported by Bouvier et al. [2015], who obtained an adj. R^2 of 0.77 for mixed forests in north-eastern France, with a pulse density of 3.4 pulses m^{-2} and plot size of 706.86 m^2. Our models yielded similar values of adj. R^2 to those obtained by Treitz et al. [2012] (adj. R^2 = 0.71-0.78) for mixed boreal forest in Canada, with models fitted using small ALS datasets of 3.2, 1.6 and 0.5 returns m^{-2}, respectively. In terms of rRMSE, the results were similar to those reported by Gleason and Im [2012] (rRMSE = 32%, plot size = 380 m^2) for mixed forest (in the US). However, the results in these types of mixed forest [Treitz et al., 2012 (rRMSE = 15%); Cao et al., 2014 (rRMSE = 18.33%); Bouvier et al., 2015 (rRMSE = 18.6%)] were slightly better than those obtained in the present study.

All fitted models (Table 4) included at least one variable related to a measure of height position (h_{05}, h_{10}, h_{mean}, h_{mode}), a metric expressing the dispersion of ALS height distribution (h_{CV}, h_{IQ}, h_{ADD}) or the shape of height distribution (h_{kurt}), and most of the models include one variable related to cover metrics (*PFRAM, CC, PARAM*). The inclusion of variables that describe height variations (e.g. h_{CV}) and lower percentile heights (e.g. h_{05}, h_{10}) accounts for intermediate tree crowns in the mid and understory. Our results also demonstrate that a specific second metric related to the shape of ALS height distribution (h_{kurt}) is potentially useful for improving mixed forest models (Table 4).

As in other studies using low-density ALS data [Thomas et al., 2006; García et al., 2010; González-Ferreiro et al., 2012; González-Ferreiro et al., 2013], the crown component of biomass was the least accurately modelled of the components analysed. One of the reasons for this is the complexity of field estimation of foliage biomass and the limited penetrability of the low-density ALS in the canopy, particularly in mixed forest with relatively closed canopy.

The results obtained for pure *Q. pyrenaica* forest must be treated with caution, given the very small sample size, although they provide some information about the variables that potentially explain the biomass components. The values obtained (adj. R^2 = 0.82-0.83 and rRMSE = 32.78-34.44%) were similar, in terms of adj. R^2, to those obtained by Cao et al. [2014] (adj. R^2 = 0.84-0.87, exclude root and foliage biomass); however, the range of values, in terms of rRMSE, was higher than the values obtained in that study (RMSE = 15.4-16.8%). The results of this study are similar to the accuracy of W_a estimates (an average error of 31.0%) analysed by Zolkos et al. [2013] in 6 studies in temperate deciduous forests.

We suggest that the lower percentiles (h_{10}, h_{25}) should be included as explanatory variables in the models because of the former shape of the canopy, which starts close to the ground in *Q. pyrenaica*. Considering the physical model, the trees with lower branches displayed

substantially lower ALS returns than species trees with only canopy branches. This is consistent with findings reported by García et al. [2010] for *Juniper thurifera* L. forest and by Cao et al. [2014] for a deciduous forest in which the primary species are *Quercus acutissima* Carruth. and *Liquidambar formosana* Hance. In the present study, the model proposed by Zonete et al. [2010] showed that h_{30} was also a stable height metric for this type of forest and explained around 60% of the variability in biomass. *CC* was not significant in the model of *Li*, thus confirming the results obtained by Bouvier et al. [2015] for other deciduous forests in France.

Height and density metrics are commonly included in ALS-based models for estimating biomass; however, no consensus has been reached regarding their use [Görgens et al., 2015]. According to our results, the height metrics derived from first returns provided the greatest stability when used in the ALS-based models [Lim and Treitz, 2004; Thomas et al., 2006; Zonete et al., 2010, González-Ferreiro et al., 2012; Görgens et al., 2015; Chirici et al., 2016]; however, the use of canopy height metrics alone may omit some information in profiles with more vertical heterogeneity, such as in natural Mediterranean mixed and even pure forest. Conversely, density metrics contributed to estimating the biomass in most models, as shown in previous studies [Stephens et al., 2012, Cao et al., 2014; Bouvier et al., 2015; Montealegre et al., 2016; Véga el al., 2016], although we did not observe any great difference between using first-return and all-returns density metrics, as also reported by Hawbaker et al. [2010].

All methods compared enable confident and stable identification of ALS metrics, providing useful information for developing a more physically based approach for generalizing predictive models of biomass components in different types of forest. The present study findings also provide information about the biomass components in the Mediterranean forest structure, which has been lacking in previous discussions [García et al., 2010; González-Olabarria et al., 2012]. The methods applied in this study are effective tools for exploring regression relationships between ALS metrics and forest attributes and for improving biomass component estimates from ALS at stand level. Finally, this research also demonstrates that the relationships already published in previous models also produce similar estimation errors as models developed from scratch using training field data.

Acknowledgements

The study was supported by the ForEadapt project ‘Knowledge exchange between Europe and America on forest growth models and optimization for adaptive forestry’ (PIRSES-GA-2010-269257). The authors thank (i) the foresters of the Extremadura Forest Service for assistance with data collection, (ii) the Portuguese Science Foundation (SFRH/BD/52408/2013) for funding the research activities of Juan Guerra and (iii) the Galician Government and European Social Fund (Official Journal of Galicia – DOG nº 52, 17/03/2014 p. 11343, exp: POS-A/2013/049) for funding the postdoctoral research stays of Eduardo González-Ferreiro and iv) the anonymous Reviewers of the European Journal of Remote Sensing for their helpful feedback. The research was carried out in the Centro de Estudos Florestais: a research unit funded by Fundação para a Ciência e a Tecnologia (Portugal) within UID/AGR/00239/2013.

References

Asner G.P., Mascaro J., Muller-Landau H.C., Vieilledent G., Vaudry R., Rasamoelina M., Hall J.S., Van Breugel M. (2012) - *A universal airborne LiDAR approach for tropical forest carbon mapping*. Oecologia, 168: 1147-1160. doi: http://dx.doi.org/10.1007/s00442-011-2165-z.

Belsley D.A., Kuh E., Welsch R.E. (2005) - *Regression diagnostics: Identifying influential data and sources of collinearity.* John Wiley & Sons, New Jersey, pp. 310. doi: http://dx.doi.org/10.1002/0471725153.

Bouvier M., Durrieu S., Fournier R.A., Renaud J.P. (2015) - *Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data*. Remote Sensing of Environment, 156: 322-334. doi: http://dx.doi.org/10.1016/j.rse.2014.10.004.

Cao L., Coops N., Hermosilla T., Innes J., Dai J., She G. (2014) - *Using Small-Footprint Discrete and Full-Waveform Airborne LiDAR Metrics to Estimate Total Biomass and Biomass Components in Subtropical Forests*. Remote Sensing, 6: 7110-7135. doi: http://dx.doi.org/10.3390/rs6087110.

Chirici G., McRoberts R.E., Fattorini L., Mura M., Marchetti M. (2016) - *Comparing echo-based and canopy height model-based metrics for enhancing estimation of forest aboveground biomass in a model-assisted framework*. Remote Sensing of Environment, 174: 1-9. doi: http://dx.doi.org/10.1016/j.rse.2015.11.010.

Corona P., Cartisano R., Salvati R., Chirici G., Floris A., Di Martino P., Marchetti M., Scrinzi G., Clementel F., Travaglini D., Torresan C. (2012) - *Airborne Laser Scanning to support forest resource management under alpine, temperate and Mediterranean environments in Italy*. European Journal of Remote Sensing, 45: 27-37. doi: http://dx.doi.org/10.5721/EuJRS20124503.

Dyer D.W. (2006) - *Watchmaker Framework for Evolutionary Computation*. Available online at: http://watchmaker.uncommons.org/examples/biomorphs.php (Accessed 18 June 2006).

García-Gutiérrez J., González-Ferreiro E., Riquelme-Santos J.C., Miranda D., Diéguez-Aranda U., Navarro-Cerrillo R.M. (2014) - *Evolutionary feature selection to estimate forest stand variables using LiDAR*. International Journal of Applied Earth Observation and Geoinformation, 26: 119-131. doi: http://dx.doi.org/10.1016/j.jag.2013.06.005.

García M., Riaño D., Chuvieco E., Danson F.M. (2010) - *Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data*. Remote Sensing of Environment, 114: 816-830. doi: http://dx.doi.org/10.1016/j.rse.2009.11.021.

Gleason C.J., Im J. (2012) - *Forest biomass estimation from airborne LiDAR data using machine learning approaches*. Remote Sensing of Environment, 125: 80-91. doi: http://dx.doi.org/10.1016/j.rse.2012.07.006.

Görgens E.B., Packalen P., da Silva A.G.P., Alvares C.A., Campo O.C., Stape J.L., Rodriguez L.C.E. (2015) - *Stand volume models based on stable metrics as from multiple ALS acquisitions in Eucalyptus plantations*. Annals of Forest Science, 72: 489-498. doi. http://dx.doi.org/10.1007/s13595-015-0457-x.

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. doi: http://dx.doi.org/10.1093/forestry/cps002.

González-Ferreiro E., Miranda D., Barreiro-Fernández L., Buján S., García-Gutiérrez J., Diéguez-Aranda U. (2013) - *Modelling stand biomass fractions in Galician Eucalyptus globulus plantations by use of different LiDAR pulse densities*. Forest Systems, 22: 510-525. doi: http://dx.doi.org/10.5424/fs/2013223-03878.

González-Ferreiro E., Diéguez-Aranda U., Crecente-Campo F., Barreiro-Fernández L., Miranda D., Castedo-Dorado F. (2014) - *Modelling canopy fuel variables for Pinus radiata D. Don in NW Spain with low-density LiDAR data*. International Journal of Wildland Fire, 23: 350. doi: http://dx.doi.org/10.1071/WF13054.

González-Olabarria J.R., Rodríguez F., Fernández-Landa A., Mola-Yudego B. (2012) - *Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR measurements*. Forest Ecology and Management, 282: 149-156. doi: http://dx.doi.org/10.1016/j.foreco.2012.06.056.

Hall S.A., Burke I.C., Box D.O., Kaufmann M.R., Stoker J.M. (2005) - *Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests*. Forest Ecology and Management, 208: 189-209. doi: http://dx.doi.org/10.1016/j.foreco.2004.12.001.

Hawbaker T.J., Gobakken T., Lesak A., Trømborg E., Contrucci K., Radeloff V. (2010) - *Light detection and ranging-based measures of mixed hardwood forest structure*. Forest science, 56: 313-326. doi: http://dx.doi.org/10.1029/2008JG000870.

Holland J.H. (1992) - *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artific*. MIT Press, Cambridge, pp. 183.

Lefsky M.A., Cohen W.B., Harding D.J., Parker G.G., Acker S.A., Gower S.T. (2002) - *Lidar remote sensing of above-ground biomass in three biomes*. Global ecology and biogeography, 11: 393-399. doi: http://dx.doi.org/10.1046/j.1466-822x.2002.00303.x.

Lefsky M.A., Hudak A.T., Cohen W.B., Acker S.A. (2005) - *Patterns of covariance between forest stand and canopy structure in the Pacific Northwest*. Remote Sensing of Environment, 95: 517-531. doi: http://dx.doi.org/10.1016/j.rse.2005.01.004.

Lim K.S., Treitz P.M. (2004) - *Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators*. Scandinavian Journal of Forest Research, 19: 558-570. doi: http://dx.doi.org/10.1080/02827580410019490.

Li Y., Andersen H.E., McGaughey R. (2008) - *A comparison of statistical methods for estimating forest biomass from light detection and ranging data*. Western Journal of Applied Forestry, 23: 223-231. doi: http://dx.doi.org/10.1117/1.JRS.8.081598.

Lovell J.L., Jupp D.L.B., Newnham G.J., Coops N.C., Culvenor D.S. (2005) - *Simulation study for finding optimal lidar acquisition parameters for forest height retrieval*. Forest Ecology and Management, 214 (1): 398-412. doi: http://dx.doi.org/10.1016/j.foreco.2004.07.077.

Maltamo M., Næsset E., Vauhkonen J. (2014) - *Forestry applications of airborne laser scanning*. Springer, Dordrecht. doi: http://dx.doi.org/10.1007/978-94-017-8663-8.

McGaughey R. (2014) - *FUSION/LDV: software for LIDAR Data Analysis and Visualization*. In: US Department of Agriculture, F.S., Pacific Northwest Research Station, Seattle, USA, pp. 123.

Montaghi A., Corona P., Dalponte M., Gianelle D., Chirici G., Olsson, H. (2013) - *Airborne laser scanning of forest resources: an overview of research in Italy as a commentary*

case study. International Journal of Applied Earth Observation and Geoinformation, 23: 288-300. doi: http://dx.doi.org/10.1016/j.jag.2012.10.002.

Montealegre A.L., Lamelas M.T., de la Riva J., García-Martín A., Escribano F. (2016) - *Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest*. Forestry, cpw008. doi: http://dx.doi.org/10.1093/forestry/cpw008.

Næsset E. (2002) - *Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data*. Remote Sensing of Environment, 80: 88-99. doi: http://dx.doi.org/10.1016/S0034-4257(01)00290-5

Næsset E. (2004) - *Estimation of above-and below-ground biomass in boreal forest ecosystems*. International Archives of Photogrammetry, Remote Sensing Spatial and Information Sciences, 36: 145-148. doi: http://dx.doi.org/10.1016/j.rse.2008.03.004.

Ni-Meister W., Lee S., Strahler A.H., Woodcock C.E., Schaaf C., Yao T., Ranson K.J., Sun G., Blair J.B. (2010) - *Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing*. Journal of Geophysical Research: Biogeosciences, 115 (G2). doi: http://dx.doi.org/10.1029/2009JG000936.

Nord-Larsen T., Riis-Nielsen T. (2010) - *Developing an airborne laser scanning dominant height model from a countrywide scanning survey and national forest inventory data*. Scandinavian Journal of Forest Research, 25 (3): 262–272. doi: http://dx.doi.org/10.1080/02827581.2010.486000.

R Development Core Team (2014) - *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available online at: http://www.R-project.org (Accessed 27 October 2010).

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 (5): 936-951. doi: http://dx.doi.org/10.3390/f5050936.

Ruiz-Peinado R., del Rio M., Montero G. (2011) - *New models for estimating the carbon sink capacity of Spanish softwood species*. Forest Systems, 20: 176-188. doi: http://dx.doi.org/10.5424/fs/2011201-11643.

Ruiz-Peinado R., Montero G., Del Río M. (2012) - *Biomass models to estimate carbon stocks for hardwood tree species*. Forest systems, 21: 42-52. doi: http://dx.doi.org/10.5424/fs/2112211-02193.

Shapiro S.S., Wilk M.B., Chen H.J. (1968) - *A comparative study of various tests for normality.* Journal of the American Statistical Association, 63: 1343-1372. doi: http://dx.doi.org/10.1080/01621459.1968.10480932.

Stephens P.R., Kimberley M.O., Beets P.N., Paul T.S., Searles N., Bell A., Brack C., Broadley J. (2012) - *Airborne scanning LiDAR in a double sampling forest carbon inventory*. Remote Sensing of Environment, 117: 348-357. doi: http://dx.doi.org/10.1016/j.rse.2011.10.009.

Stevens J.P. (2012) - *Applied multivariate statistics for the social sciences*. Taylor & Francis, New York, pp. 652.

Thomas V., Treitz P., McCaughey J., Morrison I. (2006) - *Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density*. Canadian Journal of Forest Research, 36 (1): 34-47. doi: http://dx.doi.

org/10.1139/x05-230.

Treitz P., Lim K., Woods M., Pitt D., Nesbitt D., Etheridge D. (2012) - *LiDAR Sampling Density for Forest Resource Inventories in Ontario, Canada*. Remote Sensing, 4: 830-848. doi: http://dx.doi.org/10.3390/rs4040830.

Valbuena R., Mauro F., Arjonilla F.J., Manzanera J.A. (2011) - *Comparing airborne laser scanning-imagery fusion methods based on geometric accuracy in forested areas*. Remote Sensing of Environment, 115: 1942-1954. doi: http://dx.doi.org/10.1016/j.rse.2011.03.017.

Vauhkonen J., Maltamo M., McRoberts R.. E, Næsset E. (2014) - *Introduction to Forestry Applications of Airborne Laser Scanning. In Forestry applications of airborne laser scanning: Concepts and cases studies*. Springer, Dordrecht, pp. 1-16. doi: http://dx.doi.org/10.1007/978-94-017-8663-8.

Véga C., Renaud J.P., Durrieu S., Bouvier M. (2016) - *On the interest of penetration depth, canopy area and volume metrics to improve Lidar-based models of forest parameters*. Remote Sensing of Environment, 175: 32-42. doi: http://dx.doi.org/10.1016/j.rse.2015.12.039.

White J.C., Wulder M.A., Varhola A., Vastaranta M., Coops N.C., Pitt D., Cook B.D., Woods M. (2013) - *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach*. The Forestry Chronicle, 89: 722-723. doi: http://dx.doi.org/10.5558/tfc2013-132.

Zhao F., Guo Q., Kelly M. (2012) - *Allometric equation choice impacts lidar-based forest biomass estimates: A case study from the Sierra National Forest, CA*. Agricultural and forest meteorology, 165: 64-72. doi: http://dx.doi.org/10.1016/j.agrformet.2012.05.019.

Zolkos S.G., Goetz S.J., Dubayah R. (2013) - *A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing*. Remote Sensing of Environment, 128: 289-298. doi: http://dx.doi.org/10.1016/j.rse.2012.10.017.

Zonete F., Rodriguez L.C.E., Packalén P. (2010) - *Estimação de parâmetros biométricos de plantios clonais de eucalipto no sul da Bahia: uma aplicação da tecnologia laser aerotransportada*. Scientia Forestalis, 38: 225-223.