

Article

## Analysis of the Influence of Plot Size and LiDAR Density on Forest Structure Attribute Estimates

Luis A. Ruiz <sup>1,\*</sup>, Txomin Hermosilla <sup>1,2</sup>, Francisco Mauro <sup>3</sup> and Miguel Godino <sup>3</sup>

<sup>1</sup> Geo-Environmental Cartography and Remote Sensing Group, Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camino de Vera, s/n, Valencia 46022, Spain; E-Mail: txomin.hermosilla@live.forestry.ubc.ca

<sup>2</sup> Integrated Remote Sensing Studio, Department of Forest Resources Management, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

<sup>3</sup> School of Forestry, Technical University of Madrid, Ciudad Universitaria s/n, Madrid 28040, Spain; E-Mails: francisco.mauro@upm.es (F.M.); miguel.godino@upm.es (M.G.)

\* Author to whom correspondence should be addressed; E-Mail: laruiz@cgf.upv.es; Tel.: +34-963-87-75-50; Fax: +34-963-87-75-59.

Received: 1 January 2014; in revised form: 3 March 2014 / Accepted: 13 May 2014 /

Published: 16 May 2014

---

**Abstract:** This paper assesses the combined effect of field plot size and LiDAR density on the estimation of four forest structure attributes: volume, total biomass, basal area and canopy cover. A total of 21 different plot sizes were considered, obtained by decreasing the field measured plot radius value from 25 to 5 m with regular intervals of 1 m. LiDAR data densities were simulated by randomly removing LiDAR pulses until reaching nine different density values. In order to avoid influence of the digital terrain model spatial resolution, eight different resolutions were considered (from 0.25 to 2 m grid size) and tested. A set of per-plot LiDAR metrics was extracted for each parameter combination. Prediction models of forest attributes were defined using forward stepwise ordinary least-square regressions. Results show that the highest  $R^2$  values are reached by combining large plot sizes and high LiDAR data density values. However, plot size has a greater effect than LiDAR point density. In general, minimum plot areas of 500–600 m<sup>2</sup> are needed for volume, biomass and basal area estimates, and of 300–400 m<sup>2</sup> for canopy cover. Larger plot sizes do not significantly increase the accuracy of the models, but they increase the cost of fieldwork.

**Keywords:** forest inventory; LiDAR; plot size; data density; forest structure; forest attributes; remote sensing

---

## 1. Introduction

Strategic decisions about timing and allocation of forest operations require accurate and up-to-date spatial information of the forest status as well as its evolution [1,2]. Remote sensing devices allow for massive data acquisition, satisfying current demands for continuously precise data that accurately describe the territory and the forested areas. In addition to multispectral data collected from satellites or airborne sensors, LiDAR (Light detection and Ranging) systems produce point clouds representing the height distribution of the forest canopy, enable the characterization of the complex structure of forest ecosystems and provide valuable information to improve their management. Over the last decade, LiDAR data have been widely used to obtain more detailed knowledge for the characterization and creation of forest inventories [3] while contributing to a reduction of the associated cost [4].

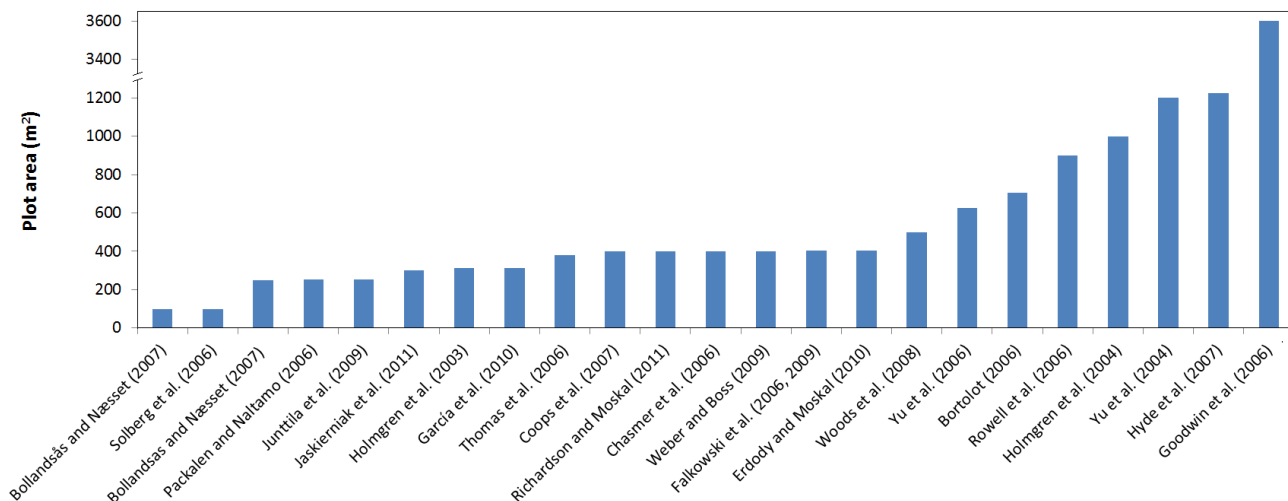
LiDAR-based forest inventory methods often employ a two-stage sampling approach with detailed measurement of a set of sampling plots, which are extended or extrapolated using surrogates such as forest type, age or structure [5]. The quality of the forest attribute estimates is thereby limited by the cost of establishing sufficient sample plots to represent existing variability [6]. For economic reasons, optimal specification of fieldwork, sensor and flight parameters for LiDAR data acquisition is important in practical forest inventory. A number of parameters for specifying airborne laser data acquisition must be decided upon prior to survey, and they may influence important properties such as point density, ability to derive forest structural information, and survey costs [7].

Plot size is an important design parameter in forest surveys, because it has the potential to either reduce or inflate the impact of the edge effects and co-registration error, affecting the forest stand estimation values [8]. Furthermore, field samples should be representative of the forest structure and density of the entire area of study in order to obtain reliable attribute models based on LiDAR metrics, and this can only be achieved by taking into consideration a minimum size and a proper distribution of the plots. As Figure 1 shows, there is some discrepancy when defining field sampling plot size when working with LiDAR data [5,9–31]. Sometimes, existing plot data are used and the size and shape of the plots are often constrained by the prerequisites of the standard techniques used for creating National Forest inventories [32]. In these cases, plot size and measurement protocols may not be optimized for LiDAR analyses.

A number of studies have been published, analyzing the effect of flight parameters [6,28,30] or LiDAR data density [5,6,15,33–35] on the forest stand attribute estimation accuracy. However, less effort has been made to assess the effect of fieldwork related parameters, such as plantation grid spacing [6], Global Positioning System (GPS) errors [36], or number of field samples required [34]. Regarding the plot size, Ene *et al.* [37] analyzed the effect of this parameter on the volume estimation considering sizes of 200, 400 and 600 m<sup>2</sup>. Similarly, Gobakken and Næsset [38] studied the effect of plot area by considering “small” (200 m<sup>2</sup>) and “large” plots (300 or 400 m<sup>2</sup>, depending on the age of the trees). A more detailed study was performed by Frazer *et al.* [8], where plot size and

co-registration effect for the estimation of biomass was simulated on synthetic data, considering circular plots with four different radii (10, 15, 20, 25 m), in addition to 25 × 25 m square plots. Results indicated that larger plot areas provide higher accuracy when estimating forest structure attributes, since the edge effect and co-registration errors are significantly reduced, but also demonstrated that an injudicious increment of the plot size would have a limited effect on the model adjustment quality.

**Figure 1.** Plot sizes used by different authors over the last decade to study and create forest attribute models based on LiDAR data.



In addition to plot size, LiDAR point density is another important parameter to be considered when estimating forest structure attributes. LiDAR point density is a function of the characteristics of the sensor system used, such as the pulse rate and energy emitted, the height and speed of flight, and is also dependent on the swath overlap. The height and speed of flight, as well as the swath overlap, have a great influence on the economic cost of data acquisition, particularly when working in large forest areas [39]. Magnusson *et al.* [33] found that low pulse density airborne laser scanner data could be cost efficient to use in inventory for estimation of forest attributes.

Gobakken and Næsset [40] studied the combined effect of plot size and LiDAR data density, with only two plot sizes tested, and data density was restricted to 1.13, 0.25, 0.13 and 0.06 points/m<sup>2</sup>, the results were limited to these specific conditions. However, they highlighted the importance of selecting the plot size depending on the qualitative type of strata considered (young or mature). Additionally, using a minimum point density is required, although the increment in this factor does not necessarily lead to significant improvements in the models. Magnussen *et al.* [35] found that a minimum point density is needed to reduce the replication variance in LiDAR derived predictions.

Since the fieldwork required for establishing large plots is expensive, as is reducing the flight speed or lowering the flight height in LiDAR data acquisition, among other variables, it is necessary to know specifically the behavior of the model adjustment quality and the errors as the plot size varies, and to relate them to the point density variations. This may enable an optimization of resources in forest inventory design by defining the plot size required to minimize the economical cost of fieldwork and facilitate accurate estimates of forest stand attributes for a particular LiDAR data density considered. This paper analyzes the combined effect of the plot size and LiDAR data density parameters to obtain

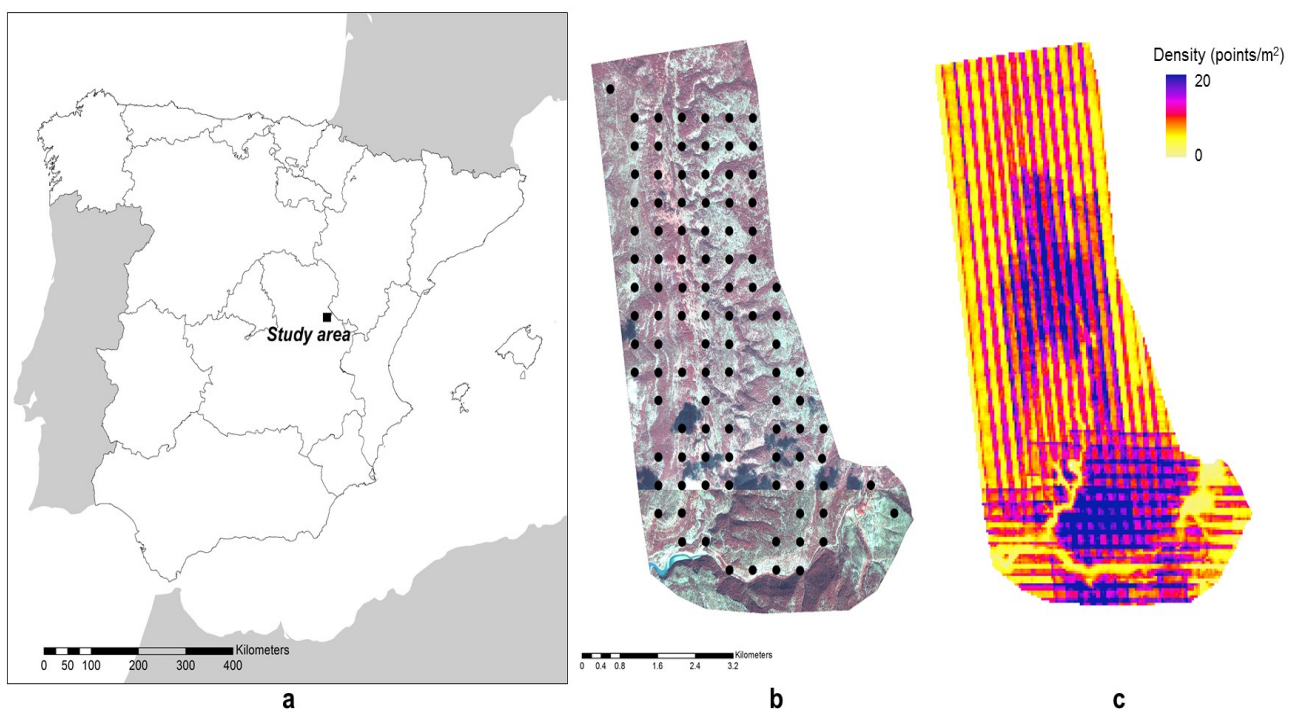
forest stand attribute estimates, testing an exhaustive range of possibilities for these two critical parameters. The paper is organized as follows: Section 2 fully describes the study area, LiDAR data, fieldwork performed, and methodology employed. Section 3 reports, analyses and discuss the statistical results obtained. Finally, Section 4 shows the overall conclusions.

## 2. Materials and Methods

### 2.1. Study Area

The study was carried out in “La Serranía de Cuenca”, a forested area of central Spain (see location in Figure 2a). The study area, characterized by very steep terrain slopes, covers an extension of approximately 4000 ha and it is dominated by European black pine (*Pinus nigra* Arn.) and scots pine (*Pinus sylvestris* L.). Both species are mixed in different proportions ranging from pure black pine stands to pure scots pine stands. The black pine is the main specie in approximately 80% of the study area and the rest is dominated by scots pine. In addition to these two species, other conifer (*Juniperus thurifera* L. and *Pinus pinaster* Ait.) and broadleaf (*Quercus ilex* L. and *Quercus faginea*) species are sparsely distributed over the study area. The canopy cover, defined as the percent of canopy overlying the forest floor with respect to the total area, is mainly low or medium (canopy cover <30% in approximately 70% of the study area) but small patches with high density are distributed over the entire study area.

**Figure 2.** (a) Location of the study area in central Spain, (b) distribution of the plots shown over a color infrared image composition of the area, and (c) average LiDAR point density computed using  $5 \times 5$  m cells.



## 2.2. Field Data

The field campaign was carried out in December 2008. A total of 102 circular plots of 25 m radius were distributed along a 500 m regular grid (see plot distribution in Figure 2b). The position of each plot center was staked out in the field with a GPS device observing the coarse acquisition code [36]. Afterwards, the relative coordinates (relative to the corresponding plot center) of each tree with a diameter at breast height (DBH) wider than 7 cm were obtained using a compass and a measuring tape. The expected accuracy of the relative positioning, based on previous experience, was approximately 0.5 m. The absolute coordinates of trees were obtained adding the relative coordinates of each tree to their correspondent plot center coordinates. Although the canopy cover was not very dense in the study area, a decrease in the GPS performance was expected. The GPS inaccuracy led to erroneous plot center locations and, therefore, to erroneous tree coordinates. This inaccurate location resulted in a lack of co-registration between field data and LiDAR data. This error was manually corrected for each plot using the normalized LiDAR point cloud and 0.25 m/pixel orthoimages. A photointerpreter used those datasets as references and manually moved the plots to obtain the best fit, thereby obtaining the new plot locations. This manual correction relied on the identification of at least seven different trees in both the field inventory dataset and the normalized point cloud, and the application of a translation, which made both datasets coincide.

Tree species, tree height (H), first branch height, defined as the height from the ground to the first live branch whorl, and the crown diameter of each tree with a DBH wider than 7 cm were measured in the field. Two perpendicular crown diameters were obtained and averaged, measured as the horizontal projection parallel and perpendicular to the direction defined by the plot center and the tree position. A Haglöf Vertex III hypsometer was used for measuring heights, DBH were determined with a caliper, and crown diameters were obtained with a measuring tape. The variables of interest were estimated at tree level using either standard methods or allometric models. Next, tree estimates were aggregated at plot level. Mean tree height, mean tree diameter, basal area, and density were determined for each plot. Stem volume was estimated for each tree using the non-linear models developed by the Spanish National Forest Inventory for each species, and for the region where the study area is located. These models predicted stem volume from DBH and tree height. Above ground tree biomass was estimated using the models developed by Montero *et al.* [41] for the main tree species in Spain. These models predict biomass from DBH and their general expression is in Equation (1):

$$\text{Biomass} = CF \times A \times (\text{DBH})^b \quad (1)$$

where  $CF$ ,  $A$  and  $b$  are coefficients specifically derived for each species.

The influence of plot size in the estimation of forest structure attributes was evaluated taking into consideration 21 different sizes created by decreasing the plot radius from 25 to 5 m with regular intervals of 1 m. For each plot, per hectare values for basal area, biomass and volume were computed. Additionally, canopy cover values were obtained using the position of each tree and the crown diameter measurements. Tree crowns were plotted and the area covered by crowns was graphically computed. Crown overlaps were discounted and parts of tree crowns exceeding plot limits were removed. Table 1 summarizes the forest structure attributes for the 102 plots in the study area.

**Table 1.** Summary statistics of plot-level forest structure attributes ( $r = 25\text{m}$ ).

Structure Attribute	Unit	Min.	Max.	Mean	Std Deviation
Volume	m <sup>3</sup> /ha	0.13	271.70	65.80	49.61
Total biomass	kg/ha	274.9	167,518.7	50,010.7	30,299.9
Basal area	m <sup>2</sup> /ha	0.06	34.52	11.41	6.86
Canopy cover	%	0.14	63.86	25.55	12.80

### 2.3. LiDAR Data

LiDAR data were collected in November 2008 using an Optech ALTM-1225 operating at 25 kHz and a scanning angle of  $\pm 18^\circ$ . The aircraft flew at an altitude of 610 m at 62 m/s resulting in a minimum nominal LiDAR posting density of 4 points/m<sup>2</sup>, a swath width of 550 m and considering four signal returns. The final point densities of the study area were variable, due to several overlapped scans during the acquisition. The final average point density was 11.4 points/m<sup>2</sup>.

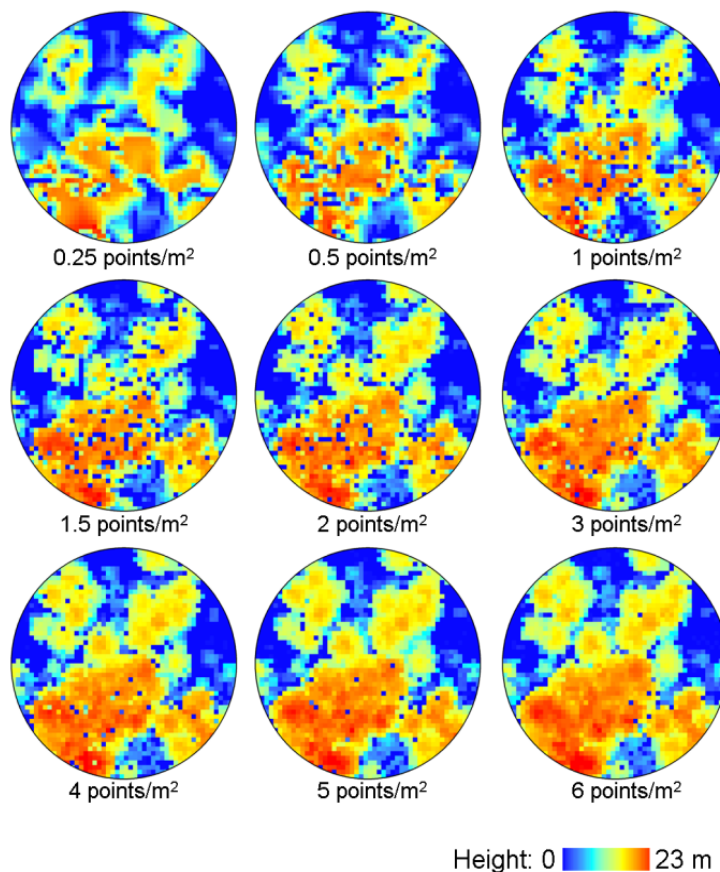
In order to analyze the effect of the LiDAR data density the actual density was reduced in a controlled manner. A homogeneous spatial density of data was obtained by dividing the area in regular  $5 \times 5$  m tiles, where LiDAR pulses were randomly removed until reaching the desired densities. After this thinning process, nine different point density values were obtained: 0.25, 0.5, 1, 1.5, 2, 3, 4, 5, and 6 points/m<sup>2</sup>. If any of the tiles in a plot presented as its maximum actual density a value lower than that required by its category, that plot was excluded from the analysis for that particular density category. Table 2 shows the total number of plots used for each of the nine density categories tested.

**Table 2.** Number of plots participating in the analysis as LiDAR data density varies.

Density (points/m <sup>2</sup> )	0.25	0.5	1	1.5	2	3	4	5	6
Number of plots	102	102	102	102	100	98	91	58	50

To avoid the possible influence of the digital terrain model's (DTM) spatial resolution on the accuracy of the attribute prediction models, eight different DTM spatial resolutions were considered and tested (0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 and 2 m grid spacing) for each combination of plot size and LiDAR point density, selecting the resolution that provided the best estimation results for each combination. The DTMs were created using an iterative algorithm that removes points belonging to any above ground objects based on the selection of height minima values on progressively smaller windows and the application of height thresholds, as described in [42]. A normalized point cloud was obtained by computing the height differences between the original LiDAR point cloud and the DTM. The canopy height model (CHM) was obtained as the difference between the digital surface model (DSM) and the DTM. Different DSM were computed using the spatial resolutions considered for the DTM. The value of each DSM cell was obtained as the maximum height value of all the points contained within the cell. The values for the cells without points were interpolated using the inverse distance weighting method. Figure 3 shows the CHM at 1 m spatial resolution of a sample plot created using the different simulated LiDAR data density categories considered in this study.

**Figure 3.** Canopy height models of a sample plot (radius = 25 m) obtained by a progressive variation of the simulated LiDAR density. The grid size of these models is always 1 m.



#### 2.4. Extraction of LiDAR Metrics and Model Generation for Forest Structure Estimates

Every plot resulting from all possible combinations of the parameters tested (9 LiDAR point densities, 21 plot sizes, and 8 spatial resolutions of DTM) was independently processed for the extraction of LiDAR metrics. Three sets of LiDAR-based descriptive features were computed: those derived from the CHM, from the normalized point cloud, and density profile based features. The features extracted from the CHM provide information about the maximum height values and their spatial distribution or texture. Some of them represent per-plot statistics: mean, standard deviation, range, kurtosis, and skewness. Additionally, the edgeness factor—a textural descriptor representing the density of edges present in the neighborhood [43]—and the mean and standard deviation values were computed by using an object-based approach, as described in [44]. Descriptive features derived from the normalized point cloud provide information about the internal structure of the vegetation provided by the height distribution of the LiDAR pulse returns. In addition to the mean, standard deviation, range, kurtosis, and skewness statistical descriptors of the normalized point cloud, the 25th, 50th, 75th and 100th height percentiles were computed. A deeper analysis of the normalized point cloud distribution is carried out using the density profiles, or per-plot height histograms [45]. Plots containing few trees present a large percentage of points at ground level. Alternatively, when several tree height strata are present in a plot, the density profile shows peaks corresponding to the different

tree species or ages. Five density profile features were computed: the percentage of ground returns, the height and the frequency of the first and second height peaks, obtained as the two consecutive frequency maxima in the density profile that are above 0.5 m. All the LiDAR-based descriptive features computed for the study are summarized in Table 3.

Ordinary least square multiple regressions were applied to create prediction models for volume, total biomass and basal area. In order to avoid an unrealistic over-fit of variables in the models, the criterion of including a maximum of three variables per model was followed. For each of the three forest structure attributes (volume, total biomass, and basal area) a total of 1512 equations were generated, corresponding to all possible combinations of the three parameters tested: LiDAR data density (21), plot size (9), and DTM spatial resolution (8). These equations were obtained by applying forward stepwise multiple regression with a maximum of three independent variables per model. The significance of the attributes was evaluated by accounting of their use in all the multiple regression models defined.

**Table 3.** Summary of the LiDAR based metrics computed.

Symbol	Descriptive Feature
$CHM_{\mu}$	Average of the CHM height distribution
$CHM_{\sigma}$	Standard deviation of the CHM height distribution
$CHM_{Range}$	Range of the CHM height distribution
$CHM_{Kurtosis}$	Kurtosis of the CHM height distribution
$CHM_{Skewness}$	Skewness of the CHM height distribution
$CHM_{\mu_{edge}}$	Average of the edgeness factor
$CHM_{\sigma_{edge}}$	Standard deviation of the edgeness factor
$H_{\mu}$	Average of the normalized point cloud height values
$H_{\sigma}$	Standard deviation of the normalized point cloud height values
$H_{Range}$	Range of the normalized point cloud height values
$H_{Kurtosis}$	Kurtosis of the normalized point cloud height values
$H_{Skewness}$	Skewness of the normalized point cloud height values
$H_{P25}$	25th percentile of the normalized point cloud height values
$H_{P50}$	50th percentile of the normalized point cloud height values
$H_{P75}$	75th percentile of the normalized point cloud height values
$H_{P100}$	100th percentile of the normalized point cloud height values
$DP_{GP}$	Percentage of ground points in the density profile
$DP_{1H}$	Height of the first peak of the density profile
$DP_{1P}$	Percentage of points at the first peak of the density profile
$DP_{2H}$	Height of the second peak of the density profile
$DP_{2P}$	Percentage of points at the second peak of the density profile

The fourth forest attribute tested, the canopy cover, was directly estimated by applying an automatic adaptive threshold to the density profile. Thus, the height threshold value between ground and tree canopy was located at the minimum value that precedes the first maximum in the density profile.

For every combination of the LiDAR density and plot size parameters, the maximum coefficient of determination ( $R^2$ ) and the coefficient of variation (CV), obtained using the different DTM spatial resolutions were computed.  $R^2$  is obtained as the square value of the Pearson's product-moment

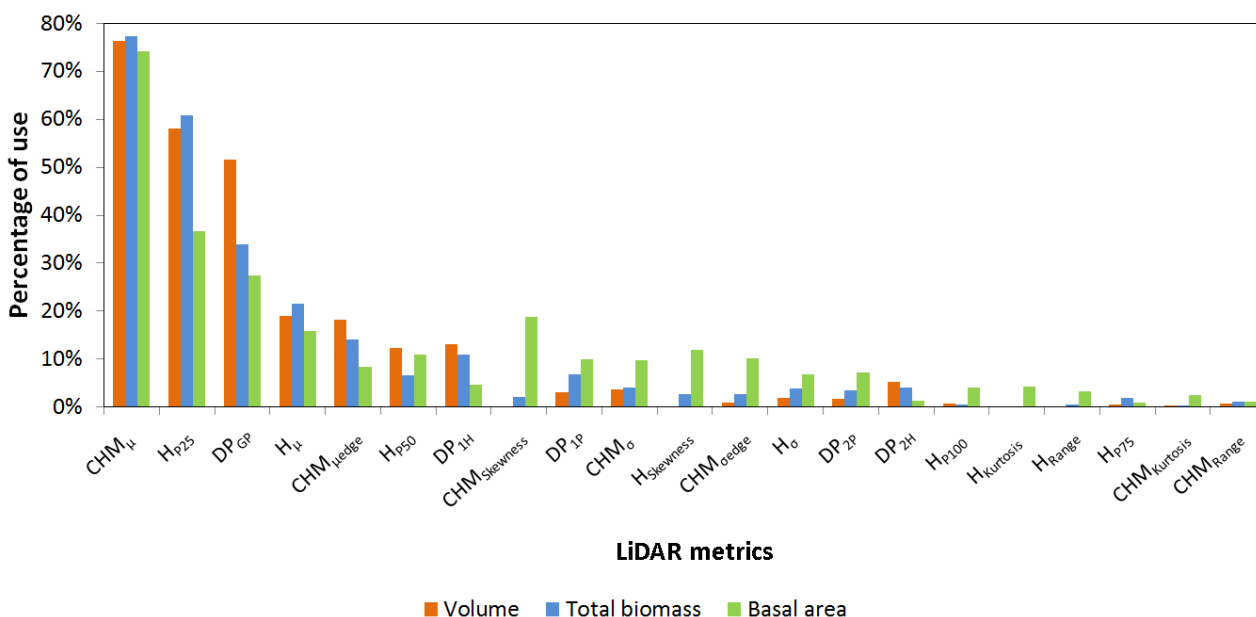


correlation coefficient. The CV is computed as the ratio of the root mean square error (RMSE) and the mean of the observed values. Accuracy values were computed using the leave-one-out cross-validation method, which uses a single observation from the original set as validation data, and the remainder as training data. This process is iterated until all observations in the sample set are used once as validation data [46]. As a result, graphical representations of the  $R^2$  distribution as the plot size and point density vary, and CV plots of the four different forest attribute estimates were obtained, and they are analyzed in the next section.

### 3. Results and Discussion

Regarding the relevance of LiDAR derived features in the generation of the prediction models, the use of the different features was registered for the 1512 possible combinations of parameters tested. Figure 4 shows the percentage of use of LiDAR metrics for three forest structure attributes: volume, total biomass and basal area.  $CHM_{\mu}$  is the most frequently used metric for describing the attributes tested in any parameter condition, being present in over 70% of the equations, followed by  $H_{P25}$  and  $DP_{GP}$ , which are significantly used for volume and total biomass attribute estimation. The remaining metrics are used 20% or less.  $H_{\mu}$ ,  $CHM_{\mu edge}$ , and  $CHM_{Skewness}$  features are included particularly in basal area models. These results illustrate the complementary nature of the three different groups of metrics—CHM, point cloud, and density profile—in describing and explaining these three forests stand attributes.

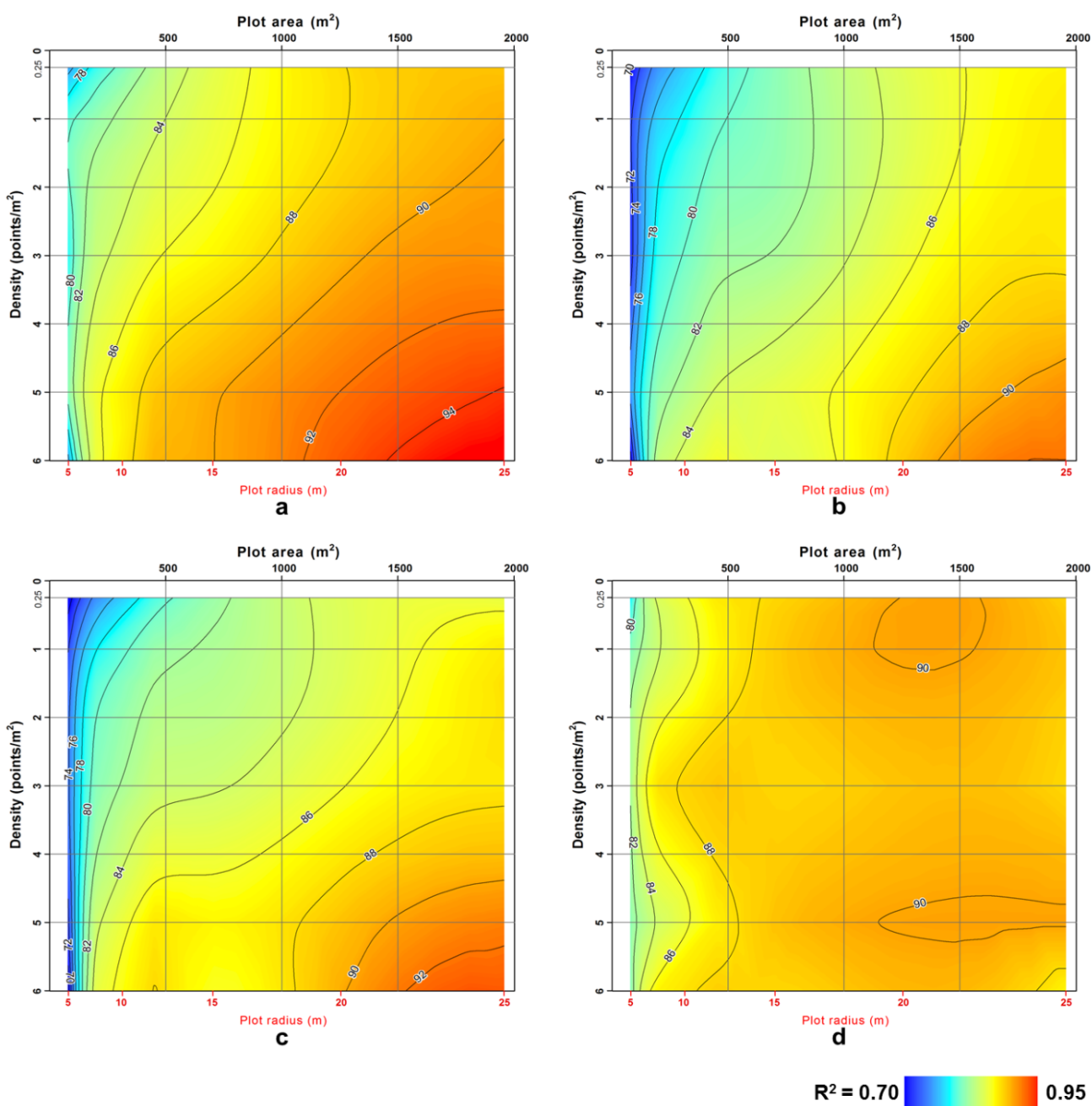
**Figure 4.** Percentage of use of LiDAR-based metrics in the prediction models of volume, total biomass, and basal area obtained using different parameter combinations.



The combined effect of LiDAR data density and plot size is represented in Figure 5, showing the  $R^2$  isolines as a function of these two methodological parameters. A fourth-order least-squares spline approximation was applied to the  $R^2$  surface before representation, in order to reduce experimental oscillations and facilitate the interpretation of main trends.

The overall behavior of the volume (Figure 5a), total biomass (Figure 5b), and basal area (Figure 5c) is similar, whereas the canopy cover (Figure 5d) presents a different trend. In the case of the first three variables, higher  $R^2$  values are reached as both plot size and LiDAR density increase, as was expected. Plot size, however, has a greater effect on the quality of forest stand estimations. In fact, for plots smaller than 200 m<sup>2</sup> (equivalent to a plot radius of 8 m), the variations of LiDAR density do not have a significant effect on  $R^2$  values, except when very low densities (less than 1 point/m<sup>2</sup>) are considered. For intermediate plot sizes, higher LiDAR data densities allow for a slight improvement of the models until a certain point density is reached (around 5 points/m<sup>2</sup>).

**Figure 5.** Variation of the coefficient of determination (%) as a function of the parameters plot size and LiDAR data density for forest stand attributes: (a) volume; (b) total biomass; (c) basal area; and (d) canopy cover.



As pointed out by the simulations performed by [8], plot size has a relevant effect on model performance. Thus, as the plot radius increases from 5 to 25 m an average increase of 0.13 in the  $R^2$  is produced in the case of volume estimates. The maximum increase occurs for a density of 6 points/m<sup>2</sup> ( $R^2$  increases from 0.76 to 0.95). The average increase of the coefficient of determination  $R^2$  for total biomass and basal area is 0.17, reaching the maximum increase for the highest LiDAR density tested ( $R^2$  ranges from 0.69 to 0.92 for total biomass, and from 0.69 to 0.93 for basal area).

Regarding the variations of  $R^2$  due to LiDAR data density, the increments produced are more moderate for the density range tested, with average increases of 0.05 for volume and total biomass, and 0.06 for basal area. The largest increases occur in plots with 8 m radius for both total biomass and basal area (from 0.75 to 0.85), and in plots with 9 m radius for volume estimation (from 0.80 to 0.87). Treitz *et al.*, [47] obtained similar results in different forest types in Ontario, where substantial differences were not found in  $R^2$  working with plot sizes of 400 m<sup>2</sup> (11.28 m radius) and testing three different pulse densities (3.2, 1.6, and 0.5 pulses/m<sup>2</sup>).

It is also remarkable how at intermediate and high LiDAR point densities, the increase of the rate of  $R^2$  is very steep for small plot sizes, up to 8 or 10 m radius, approximately. However, further increase of the plot size does not significantly improve model accuracy. This is even more noticeable for densities higher than 5 points/m<sup>2</sup>.

A different trend in the variation of  $R^2$  is obtained for the canopy cover (Figure 5d). LiDAR density has very little effect on the improvement of the estimates, which are uniquely influenced by the plot size. Thus, for plot areas  $\geq 500$  m<sup>2</sup>, then  $R^2 \geq 0.88$  are obtained. The smallest plot size tested produces an  $R^2$  value around 0.80, and presents very little  $R^2$  variation (0.78–0.82) for different LiDAR densities. The average increase of  $R^2$  reached by varying the plot size is 0.09, and the maximum increment (from 0.78 to 0.90) is given for the lowest LiDAR data density value tested (0.25 points/m<sup>2</sup>). These results show that the canopy cover attribute can be properly estimated using very low densities (0.25 or 0.5 points/m<sup>2</sup>), even when using reduced size plots. In general, plot sizes of 400 or 500 m<sup>2</sup> are sufficient to obtain good estimates. However, as pointed out by different authors [8,38], the co-registration errors expected between field plots and LiDAR data can be crucial in the final selection of plot sizes.

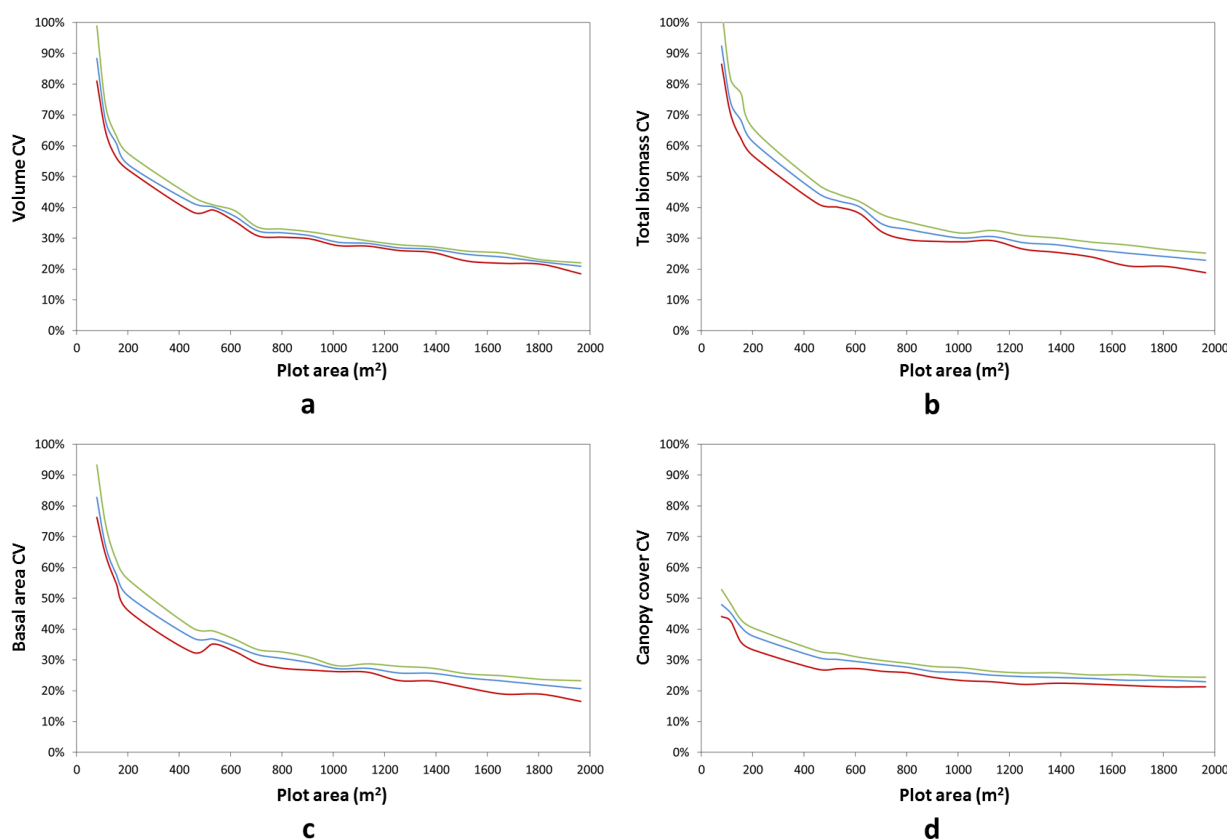
Figure 6 shows the variation of the CV in forest stand attribute estimates as a function of field plot size. The volume (Figure 6a), total biomass (Figure 6b) and basal area (Figure 6c) exhibit similar behavior with two marked trends. First, there is an exponential increase of the CV as the plot size decreases for plot areas smaller than 500–600 m<sup>2</sup>. For larger plots, the CV follows a slowly descending trend as the plot size increases. Increasing the plot size beyond 500–600 m<sup>2</sup> would not significantly improve LiDAR based prediction but it would significantly increase the number of trees to be measured in each plot. Therefore, an optimum cost-effective plot size can be expected to be near 500–600 m<sup>2</sup>.

The behavior of the canopy cover (Figure 6d) is somewhat different. The CV is stabilized at smaller plot sizes (300–400 m<sup>2</sup>) than the other three attributes tested. Under this size, there is a rapid increase of CV values as the plot area decreases. In general terms, the CV variation range for this attribute is not as pronounced as for the others. Therefore, the canopy cover seems to be less sensitive to variations in plot size than the volume, total biomass and basal area.

Plot sizes of approximately 400–500 m<sup>2</sup> seem to be efficient for diminishing prediction errors in forest attribute estimates and, at the same time, controlling the costs of field data collection. Many of

the studies of LiDAR based inventories found in the literature used similar plot sizes, in many cases due to field data acquisition standards of national forest inventories [39]. The type of forest attributes to be estimated, the expected positional precision of the co-registration of field plots and LiDAR data, and the LiDAR point density should also be considered in the definition of the plot size. In this study we applied a standard sampling design used in many local inventories. The context could be different in other countries and regions, where there may be other factors with a critical influence on the plot sample design, such as the trade-off between the distance from one plot to another and their size, which can have an impact in cases of large inventories or in areas in which field measurement is laborious due to difficult access.

**Figure 6.** Evolution of mean (blue), maximum (green) and minimum (red) coefficient of variation (CV) values as plot size varies for forest stand variables: (a) volume; (b) total biomass (c) basal area, and (d) canopy cover.



#### 4. Conclusions

The combined effect of plot size and LiDAR data density parameters on the quality of estimation of forest stand attributes (volume, total biomass, basal area, and canopy cover) has been evaluated and analyzed in this paper. Three different sets of LiDAR based descriptive features were extracted from 102 plots and used as metrics to create prediction models of the forest stand attributes, generating an exhaustive combination of different plot sizes (from 5 m radius to 25 m, with steps of 1 m) and simulated LiDAR data densities (0.25, 0.5, 1, 1.5, 2, 3, 4, 5, 6 points/m<sup>2</sup>). In addition, different DTM spatial resolutions were considered (0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 and 2 m grid spacing) to avoid

the possible influence of this parameter on the quality of the estimators. This large range of combinations of methodological parameters allowed for the analysis of their mutual behavior in terms of the  $R^2$  and CV of the prediction models. The results show that the quality of the models for volume, total biomass and basal area is affected by both parameters, but the plot size variation has a much greater effect than the LiDAR data point density. The best or most optimal performances are obtained when combining largest plot sizes and highest LiDAR densities. Nevertheless, the rate of improvement in model estimates decreases when using plot areas of 500–600 m<sup>2</sup> or more. Larger plot sizes do not significantly increase the accuracy of the models, but they clearly increase the cost of fieldwork. The accuracy of the canopy cover estimates mainly relies on the plot size parameter, and very slight variations were found when plot sizes greater than 300–400 m<sup>2</sup> were used. These results indicate that it is not necessary to use very large plot sizes to obtain quality forest stand attribute prediction models, especially when combining medium plot sizes with adequate LiDAR data densities. This may be particularly relevant considering that the cost of fieldwork is likely to rise while the acquisition cost of high-density LiDAR data is expected to decrease due to the technical advances in acquisition equipment. Nevertheless, the selection of plot size may also be affected by other specific factors, such as when working in forest types with a different mix or forest ages, or in cases of large inventories. These considerations are relevant in order to reduce the economic cost of the fieldwork, while preserving the quality of the estimated attribute values, and also to increase the frequency of updating some forest inventory variables using LiDAR data.

### Acknowledgments

The authors wish to thank the Spanish Ministry of Industry, Tourism and Trade, and the Spanish Ministry of Science and Innovation for the financial support provided in the framework of the projects InForest, and CGL2010-19591/BTE, respectively.

### Author Contributions

Luis A. Ruiz and Txomin Hermosilla conceived the work, designed the experiments and did the main writing. Miguel Godino and Francisco Mauro carried out the field work, field data compilation and analysis. Txomin Hermosilla: programming tasks. Luis A. Ruiz: coordination.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Mäkelä, H.; Pekkarinen, A. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. *For. Ecol. Manag.* **2004**, *196*, 245–255.
2. McRoberts, R.E.; Tomppo, E.O. Remote sensing support for national forest inventories. *Remote Sens. Environ.* **2007**, *110*, 412–419.
3. Van Leeuwen, M.; Nieuwenhuis, M. Retrieval of forest structural parameters using LiDAR remote sensing. *Eur. J. For. Res.* **2010**, *129*, 749–770.

4. Hyyppä, J.; Yu, X.; Hyyppä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H.; Jaakkola, A.; Vaaja, M.; Koskinen, J.; *et al.* Advances in forest inventory using airborne laser scanning. *Remote Sens.* **2012**, *4*, 1190–1207.
5. Junttila, V.; Kauranne, T.; Leppänen, V. Estimation of forest stand parameters from airborne laser scanning using calibrated plot databases. *For. Sci.* **2010**, *56*, 257–270.
6. Lovell, J.L.; Jupp, D.L.B.; Newnham, G.J.; Coops, N.C.; Culvenor, D.S. Simulation study for finding optimal LiDAR acquisition parameters for forest height retrieval. *For. Ecol. Manag.* **2005**, *214*, 398–412.
7. Gobakken, T.; Naesset, E. Assessing effects of laser point density on biophysical stand properties derived from airborne laser scanner data in mature forest. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *36*, 150–155.
8. Frazer, G.W.; Magnussen, S.; Wulder, M.A.; Niemann, K.O. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sens. Environ.* **2011**, *115*, 636–649.
9. Bollandsås, O.M.; Næsset, E. Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. *Scand. J. For. Res.* **2007**, *22*, 33–47.
10. Solberg, S.; Næsset, E.; Hanssen, K.H.; Christiansen, E. Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. *Remote Sens. Environ.* **2006**, *102*, 364–376.
11. Packalén, P.; Maltamo, M. Predicting the volume by tree species using airborne laser scanning and aerial photographs. *For. Sci.* **2006**, *52*, 611–622.
12. Jaskierniak, D.; Lane, P.N.J.; Robinson, A.; Lucieer, A. Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sens. Environ.* **2011**, *115*, 573–585.
13. Holmgren, J.; Nilsson, M.; Olsson, H. Estimation of tree height and stem volume on plots using airborne laser scanning. *For. Sci.* **2003**, *49*, 419–428.
14. García, M.; Riaño, D.; Chuvieco, E.; Danson, F.M. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sens. Environ.* **2010**, *114*, 816–830.
15. Thomas, V.; Treitz, P.; McCaughey, J.H.; Morrison, I. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using LIDAR: An examination of scanning density. *Can. J. For. Res.* **2006**, *36*, 34–47.
16. Coops, N.C.; Hilker, T.; Wulder, M.A.; St-Onge, B.; Newnham, G.; Siggins, A.; Trofymow, J.A.T. Estimating canopy structure of Douglas-fir stands from discrete-return LiDAR. *Trees Struct. Funct.* **2007**, *21*, 295–310.
17. Richardson, J.J.; Moskal, L.M. Strengths and limitations of assessing forest density and spatial configuration with aerial LiDAR. *Remote Sens. Environ.* **2011**, *115*, 2640–2651.
18. Chasmer, L.; Hopkinson, C.; Smith, B.; Treitz, P. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1359–1367.
19. Weber, T.C.; Boss, D.E. Use of LiDAR and supplemental data to estimate forest maturity in Charles County, MD, USA. *For. Ecol. Manag.* **2009**, *258*, 2068–2075.

20. Falkowski, M.J.; Smith, A.M.S.; Hudak, A.T.; Gessler, P.E.; Vierling, L.A.; Crookston, N.L. Automated estimation of individual conifer tree height and crown diameter via two dimensional spatial wavelet analysis of lidar data. *Can. J. Remote Sens.* **2006**, *32*, 153–161.
21. Falkowski, M.F.; Evans, J.S.; Martinuzzi, S.; Gessler, P.E.; Hudak, A.T. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sens. Environ.* **2009**, *113*, 946–956.
22. Erdody, T.; Moskal, L.M. Fusion of lidar and imagery for estimating forest canopy fuels. *Remote Sens. Environ.* **2010**, *114*, 725–737.
23. Jakubowski, M.K.; Guo, Q.; Kelly, M. Tradeoffs between lidar pulse density and forest measurement accuracy. *Remote Sens. Environ.* **2013**, *130*, 245–253.
24. Woods, M.; Lim, K.; Treitz, P. Predicting forest stand variables from LiDAR data in the Great Lakes-St. Lawrence forest of Ontario. *For. Chron.* **2008**, *84*, 827–839.
25. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37.
26. Bortolot, Z. Using tree clusters to derive forest properties from small footprint lidar data. *Photogram. Eng. Remote Sens.* **2006**, *72*, 1389–1397.
27. Rowell, E.; Seielstad, C.; Vierling, L.; Queen, L.; Shepperd, W. Using laser altimetry-based segmentation to refine automated tree identification in managed forests of the Black Hills, South Dakota. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1379–1388.
28. Holmgren, J. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand. J. For. Res.* **2004**, *19*, 543–553.
29. Yu, X.; Hyyppä, J.; Kaartinen, H.; Maltamo, M. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sens. Environ.* **2004**, *90*, 451–462.
30. Hyde, P.; Nelson, R.; Kimes, D.; Levine, E. Exploring LiDAR–Radar synergy predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR and InSAR. *Remote Sens. Environ.* **2007**, *106*, 28–38.
31. Goodwin, N.R.; Coops, N.C.; Culvenor, D.S. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sens. Environ.* **2006**, *103*, 140–152.
32. Bater, C.W.; Wulder, M.A.; Coops, N.C.; Nelson, R.F.; Hilker, T.; Nasset, E. Stability of sample-based scanning-LiDAR-derived vegetation metrics for forest monitoring. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2385–2392.
33. Magnusson, M.; Fransson, J.E.S.; Holmgren, J. Effects on estimation accuracy of forest variables using different pulse density of laser data. *For. Sci.* **2007**, *53*, 619–626.
34. Junttila, V.; Maltamo, M.; Kauranne, T. Sparse Bayesian Estimation of Forest Stand Characteristics from ALS. *For. Sci.* **2008**, *54*, 543–552.
35. Magnussen, S.; Naesset, E.; Gobakken, T. Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sens. Environ.* **2010**, *114*, 700–712.
36. Mauro, F.; Valbuena, R.; Manzanera, J.A.; García-Abril, A. Influence of global navigation satellite system errors in positioning inventory plots for tree-height distribution studies. *Can. J. For. Res.* **2011**, *41*, 11–23.

37. Ene, L.; Næsset, E.; Gobakken, T. Simulating sampling efficiency in airborne laser scanning based forest inventory. *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.* **2007**, *36*, 114–118.
38. Gobakken, T.; Næsset, E. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* **2009**, *39*, 1036–1052.
39. White, J.C.; Wulder, M.A.; Varhola, A.; Vastaranta, M.; Coops, N.C.; Cook, B.D.; Pitt, D.; Woods, M. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *For. Chron.* **2013**, *89*, 722–723.
40. Gobakken, T.; Næsset, E. 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.* **2008**, *38*, 1095–1109.
41. Montero, G.; Muñoz, M.; Ruiz-Peinado, R. *Producción de Biomasa y Fijación de CO<sub>2</sub> por los Bosques Españoles*; Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Ministerio de Educación y Ciencia: Madrid, Spain, 2005; p. 270.
42. Estornell, J.; Ruiz, L.A.; Velázquez-Martí, B.; Hermosilla, T. Analysis of the factors affecting LiDAR DTM accuracy in a steep shrub area. *Int. J. Digit. Earth* **2011**, *4*, 521–538.
43. Sutton, N.; Hall, E.L. Texture measures for automatic classification of pulmonary disease. *IEEE Trans. Comput.* **1972**, *21*, 667–676.
44. Ruiz, L.A.; Recio, J.A.; Fernández-Sarría, A.; Hermosilla, T. A feature extraction software tool for agricultural object-based image analysis. *Comput. Electron. Agric.* **2011**, *76*, 284–296.
45. Hermosilla, T.; Almonacid, J.; Fernández-Sarría, A.; Ruiz, L.A.; Recio, J.A. Combining features extracted from imagery and lidar data for object-oriented classification of forest areas. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2010**, *38*, 194–200.
46. Fukunaga, K. *Introduction to Statistical Pattern Recognition*; Academic Press: New York, NY, USA, 1990; p. 592.
47. Treitz, P.; Lim, K.; Woods, M.; Pitt, D.; Nesbitt, D.; Etheridge, D. LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sens.* **2012**, *4*, 830–848.