Document downloaded from:
http://hdl.handle.net/10251/53923
This paper must be cited as:
Estornell Cremades, J.; Velázquez Martí, B.; López Cortés, I.; Salazar Hernández, DM.; Fernández-Sarría, A. (2014). Estimation of wood volume and height of olive tree plantations using airborne discrete-return LiDAR data. GIScience and Remote Sensing. 51(1):17-29. doi:10.1080/15481603.2014.883209.


The final publication is available at
http://dx.doi.org/10.1080/15481603.2014.883209

Copyright Taylor \&amp; Francis: STM, Behavioural Science and Public Health Titles

Additional Information

# ESTIMATION OF WOOD VOLUME AND HEIGHT OF OLIVE TREE 

 PLANTATIONS USING AIRBORNE DISCRETE-RETURN LIDAR DATAJ. Estornell ${ }^{\text {a }}$, B.Velázquez-Martíb ${ }^{\text {b }}$ I. López-Cortés ${ }^{\text {c }}$, D. Salazar ${ }^{\text {c }}$, A. Fernández-Sarría ${ }^{\text {a }}$<br>${ }^{a}$ Dpto. Ing. Cartográfica, Geodesia y Fotogrametría, Universitat Politècnica de València. Camino de Vera s/n, 46022, Valencia (España).<br>${ }^{\mathrm{b}}$ Dpto. Ing. Rural y Agroalimentaria, Universitat Politècnica de València. Camino de Vera s/n, 46022, Valencia (España).<br>${ }^{\mathrm{c}}$ Dpto. Producción Vegetal, Universitat Politècnica de València. Camino de Vera $\mathrm{s} / \mathrm{n}, 46022$, Valencia (España).<br>Corresponding author: J. Estornell, email: jaescre@cgf.upv.es<br>Tel.: 00342843240


#### Abstract

The aim of this study is to analyze methodologies based on airborne LiDAR technology of low pulse density points $\left(0.5 \mathrm{~m}^{-2}\right)$ for height and volume quantification of olive trees in Viver (Spain). A total of 29 circular plots of radius 20 m were sampled and their volumes and height were obtained by dendrometric methods. For these estimations several statistics derived from LiDAR data were calculated in each plot. Regression models were calculated to predict volume and height. The results showed a good performance for estimating volume $\left(\mathrm{R}^{2}=0.70\right)$ and for total height estimations $\left(\mathrm{R}^{2}=0.67\right)$.


Keywords: biometrics, dendometry, remote sensing, digital terrain model

## 1. INTRODUCTION

Recently, new studies for orchard management are being based on the proportionality between wood biomass of the trees and several inputs and outputs in the crop system (Velázquez-Martí et al., 2011; Velázquez-Martí et al., 2012), such as yield, pruning residues, foliar area, soil shadow, water necessities or nutrients. Research on these topics support the contention that the amount of matter in the different structures maintain a balanced proportionality which would be characteristic of the species and cultivation practices (Diéguez., et al., 2003; VelázquezMartí et al., 2010). In addition, knowledge of total tree biomass allows determining the biomass available in the renewal of plantations due to: end of the productive life cycle, change of rootstock or restructuring of the plantation, and changing land use. Then, commercialization of residual biomass could mean additional income for farmers. Finally we must emphasize its influence on the balance of the $\mathrm{CO}_{2}$ and its effect on the environment as carbon stocks (Askew and Holmes, 2002). The study of total tree biomass can be very important from an ecological point of view since it is responsible of significant processes that affect the energy and material exchange between vegetation and atmosphere.

For the application of above studies in agriculture a morphometric characterization of fruit trees is needed. While several allometric relationships are known for predicting woody biomass in forest science, little research has been conducted in agriculture. Unlike forest trees where a significant fraction of woody biomass of the plant is contained by the trunk, for fruit trees the stem is very short and most of the biomass is concentrated in the crown. This fact inquires a particular adaptation of the forestry methods for estimating morphometric variables. Intuitively, we can state that different agronomic parameters may be related to total woody biomass of trees, such as requirements for pruning, fruit production, volume of pesticides
applied, etc. Then, the development of techniques for quantification of biomass opens a new line of work in the management of plantations. On the other hand height is an important parameter in pruning tasks and it should be within a certain range to facilitate the access to the fruit.

The development of new effective tools for the assessment of biomass has become a scientific challenge in order to perform maintenance and management actions of agricultural plantations. This fact entails the necessity to explore faster and less expensive methodologies as LIDAR data (Light Detection And Ranging). Several investigations have been performed successfully in forestry using these data (Hyyppä et al., 2001; Popescu et al., 2002; Yu et al., 2004; Reutebuch et al., 2005). In these applications two approaches are usually distinguished. On the one hand, plot and stand variables can be obtained such as height, volume, and biomass (Van Aardt et al., 2006; García et al., 2010; Estornell et al., 2011). On the other hand, the unit of study is the tree obtaining variables such as timber volume, crown diameter, stem diameter per tree (Popescu et al., 2007). When the scope of the study are plots, the methodology is commonly based on the calculation of regression models from the statistics derived from LiDAR data within each plot or stand and the field data not requiring the identification of individual trees. In the second approach individuals are often identified and extracted using algorithms based on the location of maximum heights in the canopy height model defined from LiDAR data. To apply this last approached, the density data is an important factor that must be considered. A point density lower than $4 \mathrm{~m}^{-2}$ may be insufficient to extract individuals according to previous research in forest areas (Нyyppä and Inkinen, 1999).

In agriculture little research has been done using airborne discrete-return LiDAR data. However some studies have been found using other LiDAR techniques. Morthy et al., (2011) used a Terrestrial Laser Scanner (TLS) to delineate olive tree crown. This technology registers a large amount of data for each individual from a station point increasing the accuracy in the
prediction of some morphometric variables such as crown and stem crown. Unlike of airborne LiDAR data, a TLS system shows a better accuracy to model individual tree characterizing the vertical distribution of vegetation structure. Nevertheless, it can result impractical for studying large areas.

Currently different governments, such as Spain, have provisioned for public use, LiDAR data of large regions. This availability favors the use of these data in agriculture at reasonable costs. However, this information provides the drawback of having a low point density, around 0.5 points $\mathrm{m}^{-2}$. This situation entails further studies adapted to these low densities, focused on estimating, inventory and resource management from orchards. The aim of this research is to develop tools to quantify the woody biomass of the olive trees using airborne discrete-return LiDAR data of low point density by plots.

## 2. MATERIALS AND METHODS

### 2.1 Study area

The study area (Fig. 1) is located in the municipality of Viver, in the province of Castellon (Spain) on a traditional and extensive farming area of olive trees (Olea europea L.). The area has a typical Mediterranean climate with hot and dry summers $\left(22^{\circ} \mathrm{C}\right)$ and mild winters $\left(7^{\circ} \mathrm{C}\right)$. The average annual rainfall is 550 mm . The plantations of the study area are in flat areas with an average elevation of 615 m above sea level.

### 2.2 Field data

The data of volume and height were obtained for 29 plots of radius 20 m randomly selected. In each plot, trees were classified according to the stem diameter into three categories: small (diameter $<25 \mathrm{~cm}$ ), medium (diameter between $25-50 \mathrm{~cm}$ ) and large (diameter> 50 cm ). The
table 1 shows statistics of measured trees. Subsequently, we measured the number of trees of each category, and then three trees per plot were selected, each one of them representative of a category. Finally, we measured by dendrometric methods each tree.


Figure 1. Distribution of the sampled plots used for the estimation of volume and height (circles) and plots used for the validation (triangles).

Table 1. Statistics of all trees used in the study

|  | Crown diameter (m) | Height (m) | Volume (m ${ }^{3}$ ) |
| :--- | :---: | :---: | :---: |
| Average | 4.31 | 3.11 | 0.207 |
| Stardard |  |  |  |
| deviation | 1.00 | 0.48 | 0.214 |
| Maximum | 7.15 | 4.4 | 1.246 |
| Minimum | 2.57 | 2.1 | 0.0152 |

### 2.3 Dendrometic analysis

The aim of the measurement process was to determine the biomass contained in whole trees (stem and crown). The calculation of stem volume is simple, applying methods fully developed in forest science such as measurements of diameter and length along it. The stem and crown diameters were measured using a diameter tape and the height from a metric pole. In contrast, the quantification of biomass contained in the crown is more complicated because the structure of crown in olive trees is latifoliate and measurement methods fully developed do not exist. For this, it was followed the methodology applied by Velázquez et al. (2012) for fruit trees that consisted in the conception of the tree crown as a theoretical forest stand, in which each branch was considered as an individual (a tree). Attending on this concept, for estimating crown biomass, a number of branches in each stratum of formation were sampled (main branches, secondary branches, etc.) and the volumes of branches were measured (Fig. 2). To measure easily the branches volume equations were calculated. Knowing the form parameters of the branches, statistical methods were applied to estimate the total biomass in the crown.


Figure 2. Different strata of the olive tree to measure the crown biomass All wood volume of the branches were measured (Fig 2) for the stratum 1 (first layer) applying volume equations. This stratum corresponds to the branches of the crown base. The number of
branches of this stratum is low (3-5 branches), being their diameters the greatest. The next stratum was sampled, selecting several representative branches (short and long branches). The number of branches in the stratum was counted to determine the volume of existing biomass. Then, the number of bud or ramifications in successive strata was also counted, sampling again several branches of them. The total volume of each stratum was calculated separately, multiplying the mean value of the branch volume by the number of occurrences. Generally, the last stratum contains very small branches. Because of this, it was not possible its measurement considering the field method previously described. In this case, an external central branch and another one from the top of the crown were extracted of each sampled tree, and their volumes were determined by submerging them into water in laboratory. Then, multiplying the obtained volume and the number of branches of this stratum, its total volume was calculated. In addition, some representative branches were stripped, obtaining the percentage of the leaf mass. The mean and standard deviation of the volumes of all plots were $3.788 \mathrm{~m}^{3}$ and $2.058 \mathrm{~m}^{3}$, respectively. For total height these parameters were 3.10 m and 0.34 m , respectively.

### 2.4 LiDAR data

LiDAR data used in this study are part of public data of the ©Institut Cartogràfic Valencià of the Valencia region (Spain) and they were acquired during a flight in November 2009, using the sensor Leica ALS60. The technical parameters were: average flight height 3070 m above sea level; pulse frequency 93.9 kHz ; scan frequency 33.7 Hz ; field of view (FOV) $50^{\circ}$; flight speed $70 \mathrm{~m} / \mathrm{s}$; nominal pulse density 0.5 points $/ \mathrm{m}^{2}$; The LiDAR data include the coordinates of the points (x, y, z) in reference system European Terrestrial 1989 (ETRS89) in UTM projection, Zone 30N.

To estimate the tree olive volume and height in plots of radius 20 m , several statistics were obtained from LiDAR data using FUSION software (McGaughey, 2008). They were potential
explicative variables in the regression models. For LiDAR data, the bare-earth surface elevation was first subtracted from each LiDAR point by using the DTM calculated with a spatial resolution of $1 \mathrm{x} 1 \mathrm{~m}^{2}$. This step was carried out by the same software. The DTM was evaluated by means of 62 ground-surveyed checkpoints measured with a GPS system (Leica System 1200) based on VRS (Virtual Reference Station) Internet RTK. Then the mean of the differences between the z values measured from the GPS system and the z values derived from the DTM was -0.02 m and the standard deviation of those differences 0.24 m . It must be clarified that this is an area with low complexity for selecting ground LiDAR points since trees are isolated and the variation in elevation of the ground is low. From these results, the accuracy of the DTM can be considered suitable to be applied to the LiDAR data in order to extract several statistics by plot.

The points with a height value less than 0.5 m were excluded to eliminate the data associated to the ground, herbs and sparse vegetation of the study area and from the remained data the following statistics were calculated by plot: maximum height, mean, standard deviation, coefficient of variation, kurtosis, skewness, interquartile distance and percentile values 5th ( $\mathrm{P}_{5}$ ), 20th $\left(\mathrm{P}_{20}\right)$, 40th $\left(\mathrm{P}_{40}\right)$, 50th $\left(\mathrm{P}_{50}\right)$, 60th $\left(\mathrm{P}_{60}\right), 80$ th $\left(\mathrm{P}_{80}\right), 95^{\text {th }}\left(\mathrm{P}_{95}\right)$. Furthermore, several measures of canopy density were derived (Means et al., 2000; Næsset 2004; van Aardt et al., 2006): $\mathrm{CH}_{0.5-1.5}$, as the proportion of laser hits above 0.5 m that belong to the height interval 0.5 m to $1.5 \mathrm{~m} ; \mathrm{CH}_{1.5-2.5}$, as the proportion of laser hits above 0.5 m that belong to the height interval 1.5 m to $2.5 \mathrm{~m} ; \mathrm{CH}_{2.5-3.5}$ as the proportion of laser hits above 0.5 m that belong to the height interval 2.5 m to $3.5 \mathrm{~m} ; \mathrm{CH}_{3.5-4.5}$ as the proportion of laser hits above 0.5 m that belong to the height interval 3.5 m to 4.5 m respect to all the laser hits above 0.5 m . These variables can describe the stratification of vegetation and foliage of the same. They may also indicate a relationship with the biomass of a plot. It is expected that the higher the percentage of points in intervals with greater heights, more biomass is found in a plot.

To estimate volume and height, we performed a stepwise regression analysis considering the variables above reported and the field data of 23 plots (circles in Fig.1). The goodness of the fit of the regression models was studied by the coefficient of determination $\left(\mathrm{R}^{2}\right)$, root mean square error (RMSE) and mean absolute error (MAE) of the residuals. In addition, it was analyzed if the residuals followed a Normal distribution. To do this, the Anderson-Darling test was applied using the significance level of $\alpha=0.05$. The null hypothesis $\mathrm{H}_{0}$ was the residuals follow a Normal distribution; the alternative hypothesis $\mathrm{H}_{\mathrm{a}}$ was that the residuals does not follow a Normal distribution. The models calculated were validated using a set of six additional plots (triangles in Fig. 1), which were not used for calculating the regression models for height and volume above explained. For these plots, the values of height and volume were measured at field. Then, a paired sample $t$-test was applied to determine whether there is a significant difference between the average of field values and estimated values. The latest values were obtained applying the coefficients of the regression models ( $\mathrm{n}=23$ ) and the LiDAR statistics of the six plots. The null hypothesis was that the difference in the mean values were zero using a confidence level of $\alpha=0.05$. Previously, it was verified that the two populations to be compared followed a normal distribution and they had the same variance.

## 3. RESULTS AND DISCUSSION

### 3.1 Models for tree volume calculation

The results of the stepwise regressions for estimating wood volume of olive tree by plots ( $\mathrm{n}=$ 23) are shown in Table 2. This model included five significant variables, three of them are statistics of the height distribution in each plot ( $P_{80}, P_{20}$, and Mean), and the other two correspond to measures of the canopy density $\left(\mathrm{CH}_{1.5-2.5}, \mathrm{CH}_{2.5-3.5}\right)$ and it gave $\mathrm{R}^{2}, \mathrm{RMSE}$, and MAE of $0.71,0.931 \mathrm{~m}^{3}$, and $0.82 \mathrm{~m}^{3}$, respectively. The p-values of the independent variables are less than 0.05 indicating that they are really different from 0 and have an effect on the volume estimation. Observing the coefficients of this regression model, it should be highlighted
the fact that the values of the $\mathrm{CH}_{1.5-2.5,} \mathrm{CH}_{2.5-3.5}$ coefficients were very close. This fact could indicate that each one of these variables practically explain the same percentage of the volume variability. Consequently, these two variables could be grouped into one variable that represents the percentage of points whose height is from 1.5 m to $3.5 \mathrm{~m}\left(\mathrm{CH}_{1.5-3.5}\right)$. So, it was proposed a new model to estimate volume olive tree by plot (Table 3). The results in terms of $R^{2}$, RMSE, and MAE were practically the same and the model included four significant variables ( $P_{80}, P_{20}$, Mean, and $\mathrm{CH}_{1.5-3.5}$ ).

Table 2. Parameters of the volume regression model

| Parameter | Estimate | Standard error | T Statistic | P-value | $\mathrm{R}^{2}$ | RMSE $\left(\mathrm{m}^{3}\right)$ | MAE $\left(\mathrm{m}^{3}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constant | -13.2348 | 4.53439 | -2.91875 | 0.0096 |  |  |  |
| $\mathrm{P}_{80}$ | 16.7159 | 4.98335 | 3.35436 | 0.0038 |  |  |  |
| $\mathrm{P}_{20}$ | 6.66144 | 2.2766 | 2.92605 | 0.0094 | 0.71 | 0.931 | 0.82 |
| Mean | -25.0727 | 7.74067 | -3.23908 | 0.0048 |  |  |  |
| $\mathrm{CH}_{1.5-2.5}$ | 0.197919 | 0.0517026 | 3.82803 | 0.0013 |  |  |  |
| $\mathrm{CH}_{2.5-3.5}$ | 0.188784 | 0.0559112 | 3.3765 | 0.0036 |  |  |  |
| Model | $\mathrm{V}=-13.235+16.716 \cdot \mathrm{P}_{80}+6.661 \mathrm{P}_{20}+-25.073$ Mean $+0.198 \mathrm{CH}_{1.5-2.5}+0.189 \mathrm{CH}_{2.5-3.5}$ |  |  |  |  |  |  |

$\bar{V}$ volume of biomass in $\mathrm{m}^{3}$ in circular plots of radius 0.20 m ; Independent variables derived from LiDAR data by plots: 80th percentile of the heights ( $P_{80}$ ), 20th percentile of the heights $\left(P_{20}\right)$, mean of the heights (mean); variables derived from the point distribution in height by plots (density metrics): percentage of points in a plot whose height is between 2.5 m and $3.5 \mathrm{~m}\left(\mathrm{CH}_{2.5-3.5}\right)$, percentage of points in a plot whose height is between 1.5 and $2.5 \mathrm{~m}\left(\mathrm{CH}_{1.5-}\right.$ 2.5); standard error of estimate in $\mathrm{m}^{3}(R M S E)$; mean absolute error in $\mathrm{m}^{3}(M A E)$.

The LiDAR data were initially grouped in four height intervals with the same range. The new regression model (Table 3) reveals that the calculation of the variables related to the density metrics for the species of this study should be calculated considering three intervals: percentage of points from 0.5 m to 1.5 m ; percentage of points from 1.5 m to 3.5 m ; percentage of points from 3.5 m to 4.5 m . The unique significant variable in the estimation of the volume and height was the one referred to the central interval whose width is twice compared to the rest of intervals.

Table 3. Parameters of the volume regression model grouping the variables

| Parameter | Estimate | Standard error | t Statistic | P -value | $\mathrm{R}^{2}$ | RMSE $\left(\mathrm{m}^{3}\right)$ | MAE $\left(\mathrm{m}^{3}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constant | -12.485 | 3.57771 | -3.48986 | 0.0026 |  |  |  |
| $\mathrm{P}_{80}$ | 16.326 | 4.66348 | 3.50088 | 0.0026 |  |  |  |
| $\mathrm{P}_{20}$ | 6.569 | 2.19468 | 2.99336 | 0.0078 | 0.70 | 0.933 | 0.83 |
| Mean | -24.886 | 7.51252 | -3.31264 | 0.0039 |  |  |  |
| $\mathrm{CH}_{1.5-3.5}$ | 0.195 | 0.04955 | 3.94156 | 0.001 |  |  |  |
| Model | $\mathrm{V}=-12.485+16.326 \cdot \mathrm{P}_{80}+6.569 \cdot \mathrm{P}_{20}+-24.886 \cdot \operatorname{Mean}+0.20 \cdot \mathrm{CH}_{1.5-3.5}$ |  |  |  |  |  |  |

$\overline{\mathrm{V}}$ volume of biomass in m 3 in circular plots of radius 0.20 m ; Independent variables derived from LiDAR data by plots: 80th percentile of the heights $\left(P_{80}\right)$, 20th percentile of the heights $\left(P_{20}\right)$, mean of the heights (mean); variables derived from the point distribution in height by plots (density metrics): percentage of points in a plot whose height is between 1.5 m and $3.5 \mathrm{~m}\left(\mathrm{CH}_{1.5-3.5}\right)$; standard error of estimate in $\mathrm{m}^{3}$ ( RMSE ); mean absolute error in $\mathrm{m}^{3}(M A E)$.

The ANOVA analysis shows a p-value of 0.0001 (Table 4) indicating there is a statistically significant relationship among the variables at $99 \%$ confidence level. It should be rejected the hypothesis of being null the coefficient of determination. Then, this model can explain the variability of the olive volume data by plots. Here, we must emphasize the importance of the canopy density metrics for olive volume prediction. We observed that when $\mathrm{CH}_{1.5-3.5}$ was removed from the stepwise model, the value of $\mathrm{R}^{2}$ decreased to 0.45 . These results cannot be compared to other agriculture studies as no published researches have been found on this topic. However, the relevance of this type of variables was reported in forest studies to estimate tree volume and biomass by plots (Næsset, 2004; Li et al., 2008; Kim et al., 2009). In our study it was observed a trend in which the higher values of $\mathrm{CH}_{1.53 .5}$, the larger values of volume were estimated. The same interpretation was obtained for the rest of the explicative variables of the model. LiDAR predicted versus field-measured volume showed a good linear relationship close to the 1:1 line (Fig. 3 left), which shows the absence of anomalous points.

Table 4. Analysis of variance of the volume regression model

| Source | Sum of Squares | Df | Mean square | F-ratio | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |


| Model | 47.5577 | 4 | 11.8894 | 10.69 | 0.0001 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Residual | 20.012 | 18 | 1.11178 |  |  |
| Total (Corr.) | 67.5697 | 22 |  |  |  |



Figure 3. Scatterplots of predicted versus observed volume (left) and normal probability plot of the volume residuals (right).

As far as the residuals are concerned it is remarkable that the values of standardized skewness and kurtosis are close to 0 indicating a good approximation to the Normal distribution. This fact was demonstrated by the results of the Anderson-Darling test as the computed p -value (Table 5) was greater than the significance level $\alpha=0.05$. Then, one should fail to reject the null hypothesis $\mathrm{H}_{0}$, in which the residuals follow a Normal distribution. The same conclusions were withdrawn from the normal probability plot (Fig. 3 right). It can be observed a nearly linear pattern of the data, which indicates that the normal distribution is a good model for the volume residuals.

Table 5. Statistics of the volume residuals

| Parameter | Value |
| :--- | ---: |
| Sample size | 23 |
| Average | $-9.49 * 10^{-7}$ |
| standard deviation | 0.953748 |


| Mínimum |  | -1.4385 |
| :--- | :--- | ---: |
| Máximum |  | 1.30993 |
| Standardized skewness |  | -0.0458772 |
| Standardized kurtosis |  | -1.44909 |
|  | Statistic |  |
| Anderson-Darling -test | A $^{2}$ | 0.629 |
|  | P-value | 0.089 |

### 3.2 Models for tree height calculation

The stepwise regression model for estimating total height of olive tree by plots $(\mathrm{n}=23)$ provided a good fit (Fig. 4 left), with values of $\mathrm{R}^{2}$, RMSE, and MAE of $0.67,0.19 \mathrm{~m}$, and 0.17 m , respectively (Table 6). Unlike for the volume estimation, the RMSE and MAE values are very low compared to the measured height values. This fact can be explained considering that LiDAR data provides height information directly. In contrast volume is a more indirect variable whose estimation is based on the relationships among the statistics derived from the distribution of the LiDAR data and the volume values influenced by the crown or stem diameters (Velázquez et al. 2012). The model for estimating height variable had four explanatory variables $P_{80}, P_{50}, C_{1.5-2.5}$, and $C_{2.5-3.5}$. In the same way as for volume estimation, it can be observed that the coefficients associated to the variables $\mathrm{CH}_{1.5-2.5}, \mathrm{CH}_{2.5-3.5}$ are very close what could indicate the possibility to group them into one $\left(\mathrm{CH}_{1.5-2.5}, \mathrm{CH}_{2.5-3.5}\right)$. It was proposed a new model to estimate height olive tree by plot (Table 7). The results in terms of $R^{2}, ~ R M S E$, and MAE were practically the same and the model included three significant variables ( $P_{80}, P_{50}$, and $\mathrm{CH}_{1.5-3.5}$ ). The p-values of them, less than 0.05 (Table 7), indicate that it is not necessary to remove any of them from the model.

Table 6. Parameters of the height regression model

| Parameter | Estimate | Standard error | t Statistic | P-value | $\mathrm{R}^{2}$ | RMSE (m) | MAE (m) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constant | 0.248388 | 0.930778 | 0.26686 | 0.7926 | 0.67 | 0.194 | 0.17 |


| $\mathrm{P}_{80}$ | 1.32855 | 0.470059 | 2.82635 | 0.0112 |
| :---: | :---: | :---: | :---: | :---: |
| $\mathrm{P}_{50}$ | -1.19849 | 0.412249 | -2.9072 | 0.0094 |
| $\mathrm{CH}_{1.5-2.5}$ | 0.0219148 | 0.00743176 | 2.9488 | 0.0086 |
| $\mathrm{CH}_{2.5-3.5}$ | 0.0269117 | 0.00715831 | 3.75951 | 0.0014 |
| Model | $\mathrm{H}=0.248+1.329 \cdot \mathrm{P}_{80}-1.198 \mathrm{P}_{50}+0.022 \mathrm{CH}_{1.5-2.5}+0.027 \mathrm{CH}_{2.5-3.5}$ |  |  |  |

$H$ average height in m of the trees in circular plots of radius 0.20 m ; Independent variables derived from LiDAR data by plots: 80th percentile of the heights $\left(P_{80}\right)$, 50th percentile of the heights $\left(P_{50}\right)$; variables derived from the point distribution in height by plots (density metrics): percentage of points in a plot whose height is between 2.5 m and $3.5 \mathrm{~m}\left(\mathrm{CH}_{2.5-3.5}\right)$, percentage of points in a plot whose height is between 1.5 and $2.5 \mathrm{~m}\left(\mathrm{CH}_{1.5-2.5}\right)$; standard error of estimate in $\mathrm{m}(R M S E)$; mean absolute error in $\mathrm{m}(M A E)$

Table 7. Parameters of the height regression model grouping the variables

| Parameter | Estimate | Standard error | t Statistic | P-value | $\mathrm{R}^{2}$ | RMSE (m) | MAE (m) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constant | -0.275479 | 0.603084 | -0.456784 | 0.6530 |  |  |  |
| $\mathrm{P}_{80}$ | 1.56022 | 0.348507 | 4.47686 | 0.0003 | 0.66 | 0.193 | 0.17 |
| $\mathrm{P}_{50}$ | -1.27911 | 0.393127 | -3.25369 | 0.0042 |  |  |  |
| $\mathrm{CH}_{1.5-3.5}$ | 0.0246351 | 0.00639784 | 3.85053 | 0.0011 |  |  |  |
| Model | $\mathrm{H}=-0.28+1.56 \cdot \mathrm{P}_{80}-1.28 \mathrm{P}_{50}+0.025 \mathrm{CH}_{1.5-3.5}$ |  |  |  |  |  |  |

$H$ average height in m of the trees in circular plots of radius 0.20 m ; Independent variables derived from LiDAR data by plots: 80th percentile of the heights $\left(P_{80}\right)$, 50th percentile of the heights $\left(P_{50}\right)$; variables derived from the point distribution in height by plots (density metrics): percentage of points in a plot whose height is between 1.5 m and $3.5 \mathrm{~m}\left(\mathrm{CH}_{1.5-3.5}\right)$, standard error of estimate in $\mathrm{m}(R M S E)$; mean absolute error (MAE)


Figure 4. Scatterplots of predicted versus observed height (left) and normal probability plot of the height residuals (right).

The ANOVA analysis shows a p-value of 0.0003 (Table 8) what means that there is a statistically significant relationship between the variables at the $99 \%$ confidence level and we can reject the hypothesis of being all model coefficients equal to 0 . In the same way as volume model, it is remarkable the importance of the canopy density metrics $\left(\mathrm{CH}_{1.5-3.5}\right)$ for height prediction. The value of $\mathrm{R}^{2}$ decreased to 0.40 when this variable was removed from the stepwise model. In the same way as volume estimation, we did not find any study on this topic in agriculture. In contrast, the ability of these data to estimate height variable has been demonstrated in forest studies, even in shrub areas where the vegetation is lower what makes more difficult its detection (Estornell et al., 2011b). Naesset (2004) used the percentile p80 and different canopy densities metrics to estimate the dominant height of mixtures of spruce, pines and deciduous species. The results obtained in our study confirm the feasibility of airborne LiDAR data to estimate height values of olive trees by plots.

Table 8. Analysis of variance of the height regression model

| Source | Sum of Suquares | Df | Mean square | F-ratio | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Model | 1.7356 | 3 | 0.578533 | 12.34 | 0.0001 |
| Residual | 0.890998 | 19 | 0.0468946 |  |  |
| Total (Corr.) | 2.6266 | 22 |  |  |  |

For the residuals, the results of the Anderson-Darling test ( $p$-value $>0.05$ ) indicates the failure to reject the null hypothesis in which the residuals follow a Normal distribution (Table 9). The nearly linear pattern of the residuals in the normal probability plot (Figure 4 right) confirms the normality of the residuals.

### 3.3 Validation

In order to validate the applicability of the selected model, we used an additional data set of six plots. For these, the values of height and volume were measured at field. Then, the coefficients of the regressions (tables 3 and 7) and LiDAR statistics computed for these six plots were used to calculate the values of height and volume. The means of volume and height
of the two sets of data (observed vs calculated) were compared applying a pair sample test being $\mathrm{H}_{0}$ : the difference between the means is equal to 0 . As the computed p -values for volume (Table 10) and height (Table 11) are greater than the significance level $\alpha=0.05$, there is no significant difference between these means. These results were corroborated observing that the mean difference of the volume and height values are included in the confidence interval at 95\% (Tables 10 and 11).

Table 9. Statistics of the volume residuals

| Parameter | Value |  |
| :--- | ---: | ---: |
| Sample size | 23 |  |
| Average | -0.00002 |  |
| standard deviation | 0.201246 |  |
| Mínimum | -0.36602 |  |
| Máximum | 0.301141 |  |
| Standardized skewness |  | -0.56504 |
| Standardized kurtosis |  | -0.98024 |
| Anderson-Darling -test | $\mathrm{A}^{2}$ Statistic | 0.386 |
|  | P -value | 0.362 |

Table 10. Results pair sample test for the validation of the height volume estimation

| 95\% confidence interval on the difference |  |
| :--- | :--- |
| between the means: ]-1.554; $2.452[$ |  |
| Difference | 0.454 |
| t (Observed value) | 0,584 |
| t (Critical value) | 2.571 |
| DF | 5 |
| p-value (Two-tailed) | 0.585 |
| Alpha | 0.05 |

Table 11. Results pair sample test for the validation of the height volume estimation

| $95 \%$ confidence interval on the difference |  |
| :--- | ---: |
| between the means:] -0.130; 0.303 [ |  |
| Difference | 0.087 |
| t (Observed value) | 1.028 |
| t (Critical value) | 2.571 |
| DF | 5 |
| p-value (Two-tailed) | 0.351 |
| alpha | 0.05 |

## 4. CONCLUSIONS

This study demonstrates the potential of airborne LiDAR data with low density to estimate the wood volume and height of olive trees by plots using low density LiDAR data. The obtained models explain around the $70 \%$ of variability for these parameters, which it could be acceptable to relate these parameters with production and pruning residues. The explicative variables are related with the point distribution in height. It was also shown the importance of adding the different canopy densities metrics, in particular the percentage of points whose height is from 1.5 m to 3.5 m , to explain the variability of volume and height.

The results obtained in this study could be improved using airborne LiDAR data with more density. These data may allow adopting a new approach based on the individual tree selection what can be more useful in the management of orchards in agriculture. These results can be very useful to be applied in biomass inventories in wide regions (e.g. these data are available in Spain), including the $\mathrm{CO}_{2}$ stored by plants in growing.

## ACKNOWLEDGMENTS

The authors appreciate material support (LiDAR data) provided by ©INSTITUT CARTOGRÀFIC VALENCIÀ (Spain). The authors appreciate the financial support provided by the Spanish Ministerio de Ciencia e Innovación in the framework of the project AGL201015334 and Vice-Rectorate for Research of Universitat Politècnica
de València (Grant PAID-06-12-3297; SP20120534).

## REFERENCES

Askew, M.F. and C.A. Holmes, 2002, "The potential for biomass and energy crops in agriculture in Europe, in land use, policy and rural economy terms," International Sugar Journal, 104(1247): 490-492.

Dieguez, U., Barrio, M., Castedo, F., Alvarez, M.F., Ruiz, A.D., Álvarez, J.G. and A., Rojo, 2003, "Dendrometría". Ed. Mundi-Prensa, Madrid, España.

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

Estornell, J., Ruiz, L.A., Velázquez-Martí, B. and A. Fernández-Sarría, 2011b, "Estimation of shrub biomass by airborne LiDAR data in small forest stands," Forest Ecology and Management, 262:1697-1703.

García, M., Riaño, D., Chuvieco, E. and F. M. Danson, 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.

Нууррä, J. and M.Inkinen, 1999, "Detecting and estimating attributes for single trees using laser scanner," The Photogrammetric Journal of Finland, 16:27-42.

Hyyppä, J., Kelle, O., Lehikoinen, M. and M. Inkinen, 2001, "A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners," IEEE Transactions on Geoscience \& Remote Sensing, 39:969-975.

Kim, Y., Yang, Z., Cohen, W.B., Pflugmacher, D., Lauver, C.L. and J.L. Vankat, 2009, "Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint LiDAR data," Remote Sensing of Environment, 113:2499-2510.

Li, Y., Andersen, H.E. and R. McGaughey, 2008, "A comparison of statistical methods for estimating forest biomass from light detection and ranging data," Western Journal of Applied Forestry, 23:223-231.

McGaughey, R.J., 2008, FUSION/LDV: Software for LIDAR data analysis and visualization. USDA Forest Service, Pacific Northwest Research Station.

Means, J. E., Acker, S. A., Brandon, J. F., Renslow, M., Emerson, L. and C. J. Hendrix, 2000, "Predicting forest stand characteristics with airborne scanning LiDAR," Photogrammetric Engineering \& Remote Sensing, 66:1367-1371.

Moorthy, I., Miller, J.R., Jimenez Berni, J.A., Zarco-Tejada, P., Hu, B., J. Chen, 2011, "Field characterization of olive (Olea europaea L.) tree crown architecture using terrestrial laser scanning data," Agricultural and Forest Meteorology, 151:204-214.

Næsset, E., 2004, "Accuracy of forest inventory using airborne laser scanning: evaluating the first Nordic full-scale operational project," Scandinavian Journal of Forest Research, 19: 554557.

Popescu, S.C., Wynne, R.H. and R.F. Nelson, 2002, "Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size," Computers and Electronics in Agriculture, 37:71-95.

Popescu, S.C., 2007, "Estimating biomass of individual pine trees using airborne LiDAR," Biomass and Bioenergy, 31:646-655.

Reutebuch, S. E., Andersen, H.-E. and R. J. McGaughey, 2005, "Light detection and ranging (LiDAR): An emerging tool for multiple resource inventory," Journal of Forestry, 103:286292.

Van Aardt, J.A.N., Wynne, R. H. and R. G. Oderwald, 2006, "Forest volume and biomass estimation using small-footprint LiDAR-distributional parameters on a per-segment basis," Forest Science, 52:636-649.

Velázquez-Martí, B., Fernandez-Gonzalez, E., Estornell, J. and L.A. Ruiz, 2010, "Dendrometric and dasometric analysis of the bushy biomass in Mediterranean forests," Forest Ecology and Management, 259:875-882.

Velázquez-Martí, B., Fernández-González, E., López-Cortes, I. and D.M. Salazar-Hernández, 2011, "Quantification of the residual biomass obtained from pruning of trees in Mediterranean olive groves," Biomass and Bioenergy, 35:3208-3217.

Velázquez-Martí B., Estornell J., López-Cortés I. and J. Martí-Gavila, 2012, "Calculation of biomass volume of citrus trees from an adapted dendrometry," Biosystems Engineering, 112: 285-292

Yu, X., Hyyppä, J., Kaartinen, H. and M. Maltamo, 2004, "Automatic detection of harvested trees and determination of forest growth using airborne laser scanning," Remote Sensing of Environment, 90:451-462.

