We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Abstract

Assessment and monitoring of forest biomass are frequently done with allometric functions per species for inventory plots. The estimation per area unit is carried out with an extrapolation method. In this chapter, a review of the recent methods to estimate forest above-ground biomass (AGB) using remote sensing data is presented. A case study is given with an innovative methodology to estimate above-ground biomass based on crown horizontal projection obtained with high spatial resolution satellite images for two evergreen oak species. The linear functions fitted for pure, mixed and both compositions showed a good performance. Also, the functions with dummy variables to distinguish species and compositions adjusted had the best performance. An error threshold of 5% corresponds to stand areas of 8.7 and 5.5 ha for the functions of all species and compositions without and with dummy variables. This method enables the overall area evaluation, and it is easily implemented in a geographic information system environment.

Keywords: QuickBird, multi-resolution segmentation, crown horizontal projection, forest inventory, regressions

1. Introduction

1.1. Role of forest inventories in biomass estimation

Forest covers a larger area on Earth and provides many products and services. Forest evaluation inventories were initiated when wood shortages arose. It can be said that they are the driving force to acquiring information on forest areas, stand composition and products. In the beginning
(Middle Ages), it was focused on timber volume and forest planning. With time demands for products and services changed, shifting their focus [1].

Forest inventories are based on a sampling design for a given threshold error, with ground plot assessment, providing data sets, which enable the forest evaluation. The need for measuring an increasing number of variables to evaluate products and services expected from forests turns them increasingly expensive and labour intensive. Remote sensing enabling the evaluation of some of those variables (e.g., areas, stand composition, crown cover) has been included in forests inventories. Though it cannot replace field work, it can rationalise it, allowing sample efficiency and reducing costs, error and labour intensity [1].

Biomass, both above and below ground, estimation, distribution and dynamics have been acquiring an increasing importance, especially in the last couple of decades. It was powered by the use of wood for bioenergy and by the evaluation of carbon stocks, sequestration and losses. The estimation of biomass can be grouped in two main methods: direct and indirect. The direct method, thought the most accurate, requires cutting trees, separating their different components (e.g., wood, bark, branches and leaves) and determining their dry weight. It is an expensive and labour demanding method, especially in what roots are concerned, which is feasible for small samples and prohibitive for large ones. The indirect method is traditionally based on mathematical relations between biomass (dependent variable) and one or a few easy-to-measure tree variables (independent variables). The biomass estimates attained with the former method are used to fit allometric functions at tree level, whose most frequent explanatory variables are the diameter at breast height and/or total height. Up to 2005 for Europe, North America and Australia, many functions were compiled by several authors [2–6]. More recent ones were found [7–12]. This large number of allometric functions stresses their species-specific and site-specific domain, reason why the selection of the best-suited function is of the utmost importance.

The tree-level biomass functions are frequently used to make estimations at plot scale (usually summing the biomass of all stems). Their estimation for a given area is based on the inventory plots (which depend on the sampling design and intensity for a certain error threshold) and an extrapolation method. In the literature, different extrapolation methods are described [1] with accuracy decreasing with the increase in the forest area, the variability of stand composition and structure, topography, soil and climate [13].

The forest inventories have cycles of 5 or 10 years. In between remote sensing can be used to evaluate the forest dynamics, in general, and biomass in particular. On the other hand, advantages can be gained as it can work at different scales and time frames; all area is under evaluation no requiring extrapolation methods; and maps can be produced in a geographical information system (GIS) environment.

A wide range of studies of biomass estimation derived from remote sensing can be found in the literature. In the next subchapter (2), a brief description of the methods and techniques is given. A case study (subchapter 3) will be used to illustrate the development of a methodology which was aimed to be simple and to be used either by researchers or by technicians. The challenge was to produce data from remote sensing images for two evergreen oaks (holm and
cork oak), native of the Mediterranean basin, and develop accurate above-ground biomass (AGB) allometric functions whose explanatory variable is crown horizontal projection, for pure and mixed stands of both species, in the latter analysing also the influence of the predictive ability of independent variables that differentiate species and composition.

The innovation of this study is the estimation of above-ground biomass considering or not stand composition, which is identified by the processing of high spatial resolution of satellite image data. In earlier studies, AGB estimation functions using satellite images did not consider the composition as independent variable. In the studies in which medium and low spatial resolution images are used, the pixel size frequently does not allow the separation of the species. High spatial resolution satellite images data enable the identification and delimitation of crowns of different tree species. Nonetheless, the studies with these satellite images seem to be focused in pure stands. The case study presented demonstrates that stand composition improves the accuracy of AGB estimation functions.

1.2. Contribution of remote sensing data to biomass estimation

Remote sensing was under a strong development in the last three decades due to the rapid advancement of remote sensing technology, increasing the availability of satellite imagery with different spectral, spatial, radioactive and temporal resolutions. Their sensor technologies enable a wide range of Earth surface monitoring scales [14] of forest areas distribution, species, and physic and biochemical properties [15]. The data can be derived from two sensors: the passive and the active.

In the **passive sensor**, the optical spectral reflectance of Earth surface is registered. This multi-spectral information is sensitive to vegetation characteristics such as their shadows and textures, tree density, leaf area index (LAI) and crown size [16]. Good correlations are found between the former and the latter, namely with biomass [17] for a wide range of geographical areas [18]. To develop biomass functions, several methods are used. The most frequent is the regression analysis using forest inventory biomass data and satellite generated data [19], such as spectral reflectance, crown diameter and crown horizontal projection [20–24], original bands and/or vegetation indices [25–28]. The green vegetation canopy has different interactions with specific electromagnetic spectrum ranges, especially the red (R) and the near infrared (NIR) regions, due to the leaves photosynthetic activity and thus with LAI and biomass. The middle infrared (MIR) region, because of its ability to normalise the canopy cover variability, plays also an important contribution [15]. High relations are established between spectral reflectance and/or vegetation indices, and canopy structure, LAI or biomass [16] from tropical [29, 30] to Mediterranean evergreen oaks and shrubland [17, 26, 31, 32] landscapes. Vegetation indices, condensing satellite imagery data in a quantitative numeric form, are related to some forest parameter's estimation, such as the number of trees per hectare [33], canopy cover [26, 34] and volume, basal area and biomass [35]. The aforementioned indices show a better sensitivity when compared with the spectral reflectance [36]. Apart from regression, other methods are described in the literature, such as k-nearest neighbour, neural network, regression tree, random forest and support vector machine [29, 37–40].
Passive sensor, satellite imagery data, depends on its spatial resolution to determine directly the working scale and can be divided into three groups: low, medium and high.

The **low spatial resolution** satellite imagery data are characterised by scenes covering large geographical areas. The first studies of biomass estimation at local, regional and global scales used the images from the National Oceanic and Atmosphere Administration (NOAA) with the Advanced Very High Resolution Radiometer (AVHRR) sensor, with 1.1 km of spatial resolution; Moderate Resolution Imaging Spectroradiometer (MODIS), with 500 m [23]; and Satellite Pour l’Observation de la Terre (SPOT) Vegetation with 1 km of spatial resolution [41]. The advantage that can be pointed out in the biomass estimation with this spatial resolution is that one scene covers large area allowing working at country and global scales. Conversely, the pixel size makes the image classification more difficult due to the different land cover types encountered in it, resulting frequently in low accuracy biomass estimations [42]. To surpass this disadvantage in some studies, spectral mixture analysis was employed, where each pixel is a physical mixture of the multiple components weighted by the dominant area and the mixture spectrum is a linear combination of the endmember reflectance spectra [43].

The **medium spatial resolution** (30 m) satellite imagery data more frequently used result from the processing of Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM +), Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) and Wide Field Sensors (WiFS) [23, 25, 44, 45]. The data set is usually combined with forest inventory sample plots data and low spatial resolution imagery. Examples of the first are the volume and biomass estimation of conifer species of boreal forest in northern Europe, with Landsat images and NOAA-AVHRR [46] and Landsat and WiFS [25]; example of the second is AGB estimation for local and regional scales [25, 39, 46].

The data of **high spatial resolution** (IKONOS, QuickBird, WordView 2) have had, more recently, a high availability increase. Improved accuracy in biomass estimation is reached when compared with the former two spatial resolutions. The main disadvantage derives from their spatial resolution, which makes the data processing more time-consuming, thus better suited for local or regional scales [27]. In the literature, some references were found for these data sets, but few for biomass estimation, most of which with functions developed with regression analysis methods. Examples with IKONOS image data are the estimation of wet and dry biomass of oil palm plantations in Africa, considering vegetation indices and individual reflectance bands [47] and allometric equations of diameter at breast height and AGB as a function of crown width in tropical forests [48]. Examples with QuickBird are a methodology to estimate and map AGB for black spruce stands in Canada with shadow fraction as explanatory variable [49], AGB functions whose independent variable is crown horizontal projection, applying multi-resolution segmentation and object-oriented classification methods, with regression analysis for *Eucalyptus* sp. and *Tectona grandis* L. [50] and *Quercus rotundifolia* [51].

The three referred spatial resolutions of the optical remote sensing imagery for the same region and their image processing techniques allow the monitoring of AGB with different degrees of detail which are able to decrease the processing time and costs [19, 27, 52].
The **active sensor** is based on a mechanism that transmits and receives a portion of energy (microwaves region), interacting with Earth surface, and whose relevance to modelling forest variables depends on how deep the signal penetrates in the canopy. The mostly used data are derived from the Space-borne synthetic Aperture Radar (SAR) and L-band ALOS PALSAR. Two groups of data can be defined: **RADAR** and **LiDAR**. The **RADAR** data’s main advantage is their independence of atmospheric conditions as wavelength of radiation is much larger that the atmospheric particles [53], reason why it has acquired a marked importance in regions with frequent cloud cover [15]. Several examples of AGB estimation with RADAR were found in the literature [54–59]. The **LiDAR** is mainly obtained by airborne discrete-return [60], airborne waveform and from ground-based [61] LiDAR. Several authors developed nonlinear functions for AGB with its derived variables [15, 62] for individual trees, plot and stands, depending on the type of LiDAR system used [16]. LiDAR derives forest structural information (canopy cover), and RADAR backscatter seems to be able to estimate AGB with good accuracy [63].

In summary, the estimation of forest AGB with remote sensing data has advantages and disadvantages. To obtain very good accuracy, these approaches need a heavy field work in order to attain training data sets. However, many studies present methodologies with very good results at local, regional and global scale for AGB estimation. There is a trend towards the combination of several types of remote sensing imagery data to generate vegetation parameters and their relation with forest biomass.

### 2. Case study

#### 2.1. Introduction

The data resulting from passive sensors, with low and medium spatial resolution, are usually used for the country or global scales. Due to their pixel size, it is not possible to identify and delimitate the tree crowns. Conversely, high spatial resolution satellite images enable it with good accuracy. The active sensors processing is more difficult than the passive sensors. Thus, when developing a methodology to be used both by researchers and by technicians, the passive sensors data processing is more straightforward than that of the active sensors. The objectives of this study are the development of allometric functions for the estimation of AGB with crown horizontal projection, obtained with high spatial resolution satellite image, as independent variable with linear regression for (i) cork oak pure stands, (ii) mixed cork oak and holm oak stands (from this point forward referred as evergreen oaks) and (iii) both pure and mixed stands (from this point forward referred as all). Though with similar stand parameters there are some differences between pure stands of cork oak or holm oak and mixed stands of both species. Thus it was considered that the model for both species might be improved with the inclusion of dummy variables for species and composition as independent variables.

Cork oak (*Quercus suber* L.) and holm oak (*Quercus rotundifolia* Lam.) are native of the Mediterranean basin, distributed from Portugal to Greece and Portugal to Syria [64]. These species occur usually in low-density stands in a silvopastoral system, called *montado*, with open heterogeneous canopies, where the main products are bark and fruit, associated with extensive
grazing, both in pure and in mixed stands. According to the Portuguese National Forest Inventory, cork oak and holm oak stands account for 22.5 and 13.7% of the forest area, respectively. The biomass for pure cork oak stands is 49.5 t/ha with an error of 5.2% and for holm oak 23.2 t/ha with an error of 7.2%. Mixed stands with cork oak or holm oak as dominant species correspond to 17% of their total area. The biomass estimation for mixed stands is 35.6 t/ha with an error of 13.7% [65]. They are similar in terms of habit, especially due to pruning to enhance fruit production and growth; both are slow growing [64].

2.2. Materials and methods

2.2.1. Study area

The study area is in southern Portugal, region of Mora, with approximately 80 km² (Figure 1). The area is mainly occupied by forest, composed predominantly by cork and holm oak in both pure and mixed stands. This region has a Mediterranean climate with a hot and dry summer and rainy winter with lower temperature. The terrain is characterised by small variations, with a mean elevation of about 200 m. The used QuickBird satellite image (August 2006) has a spatial resolution of 0.70 m, resulting from a fusion of panchromatic band with the four multi-spectral bands, blue (B), green (G), red (R) and near infrared (NIR).

2.2.2. Processing satellite image

The image was geometric and radiometric corrected using ENVI4.8 [66]. The geometric correction was based on ground control points obtained with Global Navigation Satellite System (GNSS) and geodetic vertices, identified on the ground and in the images, with a Root Mean Square Error (RMSE) of 0.49 m. The pixel values that are the digital numbers of the image were converted to top-of-atmosphere reflectance and finally to soil reflectance, using the dark object subtraction method [67]. The Normalized Difference Vegetation Index (NDVI) [68] was calculated with the latter and used to generate a vegetation mask, isolating tree crowns from other land cover types. For this propose, the multi-resolution segmentation method with the contrast split segmentation algorithm on eCognition, version 8.0.1 was applied [69]. Over this vegetation mask, forest species were identified using the nearest neighbour classification method, producing a map with the delimitation of trees crowns identified by species [70].

The area was divided into a square grid of 45.5 m× 45.5 m (2070.25 m²). Grids were classified according to the species present. The crown cover obtained with satellite data (CC) was defined as the percentage of area occupied by the crown horizontal projection per grid (CHP) in relation to the grid area. It was calculated with the vegetation mask per species using a geographic information system, ArcGIS, version 10 [71]. A grid was considered pure when crown cover of one species was equal or larger than 75% and mixed otherwise. The design of the forest inventory was a random stratified sampling by proportional allocation. Three strata were defined as function of crown cover: (i) 10–30%, (ii) 30–50% and (iii) >50%.
2.2.3. Forest inventory data and biomass estimation

Forest inventory data set is composed of 17 plots of pure holm oak, 11 plots of pure cork oak and 23 mixed plots of holm oak and cork oak, with a total sampled area of 10.6 ha. In these plots, for all individuals with a breast height diameter ≥6 cm, the diameter at breast height, total height and crown radii in four directions (North, South, East and West) were measured [72]. Each tree geographical location was recorded by GNSS. Tree crown horizontal projection (CHP) was calculated as the circle whose radius is the arithmetic mean of the four crown radii measured. The relation (in %) of sum of the tree crown horizontal projections to plot area was used to estimate plot crown cover (CCp). Ref. [8] developed a biomass allometric function per tree with simultaneous fitting for cork and holm oak. As the stands referred by the aforementioned authors were similar to those of this study, their functions were used to estimate above-ground biomass (AGB, Eq. (1)) as the sum of wood (ww), bark (wb) and crown (wc) biomass. Above-ground biomass per plot (AGBp) was defined as the sum of AGB of all the trees in the plot.

\[
AGB = \text{ww} + \text{wb} + \text{wc} = 0.164185 \times \text{dbh}^{2.011002} + \\
0.600169 \times \text{dbh}^{1.355957} + 1.909152 \times \text{dbh}^{1.200354}
\]  

(1)
2.2.4. Statistical analysis

Statistical analysis included correlation analysis and linear and multiple regression, implemented in R statistical software [73]. As AGB, CHP, and CHP did not have normal distribution (evaluated with Shapiro-Wilk normality test), the Spearman correlation test was used. The linear function was selected as it is especially suited for one layer or without crown closure stands or forests [26]. As it was assumed that null CHP corresponded to null AGB, the linear regression through the origin (Eq. (2), where $\beta$ is the slope) was used. The ordinary least square linear method was selected to fit the functions, both for the individual and for their cumulative values (Figure 2.) with CHP, as independent variable. Also, in order to understand whether AGB could be estimated with aforementioned models, a preliminary analysis was carried out for plot inventory data (CHP) with the same methodology for cork oak, evergreen oaks and all plots. The preliminary analysis for holm oak was done by [51], and the data were included in the evaluation for all plots in this study.

![Figure 2. Flow diagram of the models developed (where AGB is the above-ground biomass in kg, CHP is the crown horizontal projection in m$^2$, t is the tree, p is the plot values, i is the inventory data, c is the cumulative values, s is satellite image data and d is dummy variables).](image)
As the goal was to develop functions able to estimate AGB as function of CHP for the two evergreen oaks, in pure and mixed stands, dummy variables were defined to test the species composition contribution in the function predictive ability (Figure 2). Three dummy variables were defined. For each dummy variable, 1 indicates that the plot is pure holm oak (dQR), pure cork oak (dQS) or mixed holm oak and cork oak (dQRQS) and 0 otherwise. Multiple linear regression with stepwise method, using Akaike information criterion (AIC) as selection criteria, was used to fit the allometric functions (Eq. (3), where $\beta_i$ are the regression coefficients, $i = 0, 1, 2, 3$). Multicollinearity among explanatory variables was analysed with variance inflation factor (VIF), considering that values exceeding 10 are sign of serious multicollinearity [74, 75].

As suggested by several authors [76, 77], the models were studied by the sum of squares of the residuals (SQR), the coefficient of determination ($R^2$) and the adjusted coefficient of determination ($R^2_{adj}$). Validation tests entail an independent data set. To overtake the inexistence of an independent data set, Refs. [78–80] suggest using predicted residual error. The sum of its square values, PRESS (Eq. (4)), and the sum of its absolute values, APRESS (Eq. (5)), were used as the validation test. The closer to the null value of residuals, the better is the model. The percent value of the estimated and calculated AGB defined the error.

$$ AGB = \beta \times CHP $$

$$ AGB = \beta_0 \times CHP + \beta_1 \times dQS + \beta_2 \times dQR + \beta_3 \times dQRQS $$

$$ PRESS = \sum_{i=1}^{n} (y_i - \hat{y}_{i-i})^2 $$

$$ APRESS = \sum_{i=1}^{n} |y_i - \hat{y}_{i-i}|^2 $$

2.3. Results and discussion

2.3.1. Multi-resolution segmentation and object-oriented classification

Figure 3 presents the vegetation mask resulting from the multi-resolution segmentation and object-oriented classification process (yellow line), over the QuickBird image with false colour composite for an area of Mora. The vegetation index NDVI enables a spectral signature sufficiently different to obtain a good distinction of the two forest species (cork and holm oak), with an agreement between the classification and ground truth using Kappa statistic [81, 82], of 78% and a global precision of 89%. Thus, the simultaneous analysis of these two evergreen oak species in pure and mixed stands can be useful for the estimation of AGB and also because,
though they are visually similar, the methodology allows a mapping per species (cork and holm oak) with a good classification accuracy.

Figure 3. QuickBird image with false colour composite (RGB = red, NIR and blue) and illustration of the results of multi-resolution segmentation process (yellow line).

2.3.2. Individual trees and plot characteristics and above-ground biomass functions with inventory data

The descriptive statistics: minimum ($\text{min}$), maximum ($\text{max}$), arithmetic mean ($\text{mean}$), standard variation ($\text{SD}$) and coefficient of variation ($\text{CV}$) for all plots ($\text{All}$), evergreen oaks mixed plots ($\text{QRQS}$) and cork pure plots ($\text{QS}$) of the data are presented in Table 1. The density measures, number of trees per hectare ($\text{N}$), basal area per hectare ($\text{G}$), crown cover ($\text{CC}$) and above-ground biomass per hectare ($\text{AGB}$) were considered for the inventory data, while crown horizontal projection per plot ($\text{CHP}_h$) and crown cover ($\text{CC}$) were considered for satellite data. In Table 1, the differences between the three groups of plots can be observed. Comparing the descriptive statistics of cork oak pure plots with those of evergreen oak plots (Table 1), higher densities and $\text{AGB}$ are observed for the former. A more detailed analysis at tree level reveals that in the mixed plots about 54% of the trees are holm oak and 46% are cork oak. The differences are, at least partially, explained by the individual tree dimensions. While in the mixed plots mean crown radius is equal for both species (3.1 m) and similar to the cork oak pure plots (3.0 m), the mean total height and the mean quadratic diameter ($d_q$) are different:
6.8 m, 7.1 m, 35.8 cm and 37.6 cm, respectively. This is also reflected in the AGB, with a mean larger for the cork oak pure plots (463.9 kg tree\(^{-1}\)) than for the mixed plots (424.0 kg tree\(^{-1}\)). When the two species are analysed in the evergreen oak plots, the cork oak individuals have, on average, larger dimensions (7.1 m of total height and 38.4 cm of \(d_t\)) than the holm oak ones (6.3 and 33.6 cm), which is also reflected in the mean AGB (481.7 and 382.3 kg tree\(^{-1}\), respectively). The diameter distribution of the evergreen oak plots covers a wide span indicative of stands of several cohorts (or age classes). The species spatial distribution visual analysis points out towards an individual distribution both by species and by cohort. This might be probably due to the mutualistic and competitive interactions [83] and also by the sociability and compatibility [84] of these oaks, with similar growth rates, both slow, and rotation length. The holm oak pure plots [51] have values between the former two compositions. Also, trees with similar crown sizes may have quite different AGB\(_t\). This is expectable considering two factors. First, these species have weak epinastic control, but when in group tend to reduce their lateral development due to branch abrasion or to the effects of high and low shading [83], especially because they are shade intolerant. As in montado tree spatial distribution is sparse, though irregular, trees can be found isolated or in clusters and in the latter is the competition phenomenon that explains, at least partially, the variability in crown dimensions. Second, in the montado system pruning to enhance fruit production is frequent [64] in mature stands and this reverses temporarily the rate between CHP\(_t\) and AGB\(_t\).

<table>
<thead>
<tr>
<th>Variables</th>
<th>All plots</th>
<th>Evergreen oaks mixed plots</th>
<th>Cork oak pure plots</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(All)</td>
<td>(QRQS)</td>
<td>(QS)</td>
</tr>
<tr>
<td>(N) (trees ha(^{-1}))</td>
<td>19 140 70 28 40.5</td>
<td>19 140 61 28</td>
<td>46.5 53 135 88 24 27.7</td>
</tr>
<tr>
<td>(G) (m(^2) ha(^{-1}))</td>
<td>2.6 15.4 7.1 2.7 38.1</td>
<td>2.6 13.1 6.2 2.1</td>
<td>34.3 5.2 15.4 9.8 3.0 31.2</td>
</tr>
<tr>
<td>CC(_p) (%)</td>
<td>8.7 49.3 25.3 9.2 36.5</td>
<td>8.7 49.3 20.9 8.2</td>
<td>39.3 18.7 42.3 27.5 7.4 26.8</td>
</tr>
<tr>
<td>CC(_s) (%)</td>
<td>13.7 70.5 35.0 14.7 41.9</td>
<td>14.3 58.8 28.5 11.2</td>
<td>39.2 23.6 70.5 44.9 13.5 30.0</td>
</tr>
<tr>
<td>CHP(_p) (m(^2))</td>
<td>284.2 1460.2 725.3 303.7 41.9</td>
<td>296.9 1216.7 589.3 230.9</td>
<td>39.2 489.0 1460.2 928.7 278.7 30.0</td>
</tr>
<tr>
<td>AGB (tha(^{-1}))</td>
<td>10.4 62.7 30.1 11.1 36.8</td>
<td>10.4 51.5 26.0 8.6</td>
<td>33.1 25.8 62.7 40.7 11.9 29.2</td>
</tr>
</tbody>
</table>

N: number of trees per hectare, G: basal area per hectare, CC\(_s\): crown cover calculated with satellite data, CHP\(_p\): plot crown horizontal projection evaluated with satellite data, AGB: above-ground biomass per hectare, min: minimum, max: maximum, mean: arithmetic mean, SD: standard variation, CV: coefficient of variation.

Table 1. Descriptive statistics.

The strongest positive correlations are found for pure plots of cork oak between \(AGB\_t\) and \(CHP\_p\) (0.827), expected due to the smaller variability (Table 1). Strong positive correction for pure cork oak plots between \(AGB\_t\) and \(CHP\_p\) was attained (0.627). The difference between \(AGB\_t\) and \(CHP\_p\) and \(CHP\_p\) can be, at least partially, explained by the stems spatial distribution, the shade casted by trees and understory. In tree clusters, multi-resolution segmentation process is not able to isolate tree crowns, including mixed pixels of vegetation, soil and/or shadow inside the polygon that delimitates the cluster. Some types of understory vegetation, near tree crowns, have similar spectral signature, and when delimiting the crown those pixels are
included [85, 86]. The weakest correlations were observed for the evergreen oak plots between \( AGB_p \) and \( CHP_p \) (0.482) and \( CHP_s \) (0.552), which is expected due to the variability increase in those parameters. When all the plots are considered, strong positive correlations are found between \( AGB_p \) and \( CHP_p \) (0.671) and \( CHP_s \) (0.746).

<table>
<thead>
<tr>
<th>Model</th>
<th>Plots</th>
<th>Equation</th>
<th>SQR</th>
<th>R²</th>
<th>R²adj</th>
<th>PRESS</th>
<th>APRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>QS</td>
<td>( AGB_{pi, QS} = 14.7087 \times CHP_{pi, QS} )</td>
<td>23576825</td>
<td>0.972</td>
<td>0.969</td>
<td>0.00000019</td>
<td>0.00101252</td>
</tr>
<tr>
<td>2</td>
<td>QRQS</td>
<td>( AGB_{pi, QRQS} = 11.5889 \times CHP_{pi, QRQS} )</td>
<td>74401387</td>
<td>0.899</td>
<td>0.894</td>
<td>0.000000065</td>
<td>0.00313966</td>
</tr>
<tr>
<td>3</td>
<td>All</td>
<td>( AGB_{pi, All} = 11.4139 \times CHP_{pi, All} )</td>
<td>187602462</td>
<td>0.916</td>
<td>0.915</td>
<td>0.0000113</td>
<td>0.00361091</td>
</tr>
<tr>
<td>4</td>
<td>QS</td>
<td>( AGB_{pic, QS} = 14.8889 \times CHP_{pic, QS} )</td>
<td>16220130</td>
<td>0.9995</td>
<td>0.999</td>
<td>0.00000002</td>
<td>0.00034940</td>
</tr>
<tr>
<td>5</td>
<td>QRQS</td>
<td>( AGB_{pic, QRQS} = 11.75374 \times CHP_{pic, QRQS} )</td>
<td>29287434</td>
<td>0.997</td>
<td>0.997</td>
<td>0.00000004</td>
<td>0.00049897</td>
</tr>
<tr>
<td>6</td>
<td>All</td>
<td>( AGB_{pic, All} = 11.75374 \times CHP_{pic, All} )</td>
<td>306889502</td>
<td>0.998</td>
<td>0.998</td>
<td>0.00000001</td>
<td>0.00049770</td>
</tr>
<tr>
<td>7</td>
<td>QS</td>
<td>( AGB_{ps, QS} = 8.847 \times CHP_{ps, QS} )</td>
<td>40991932</td>
<td>0.953</td>
<td>0.948</td>
<td>0.00000016</td>
<td>0.00106716</td>
</tr>
<tr>
<td>8</td>
<td>QRQS</td>
<td>( AGB_{ps, QRQS} = 8.5093 \times CHP_{ps, QRQS} )</td>
<td>72831734</td>
<td>0.901</td>
<td>0.896</td>
<td>0.00000061</td>
<td>0.00317184</td>
</tr>
<tr>
<td>9</td>
<td>All</td>
<td>( AGB_{ps, All} = 8.1309 \times CHP_{ps, All} )</td>
<td>163946315</td>
<td>0.927</td>
<td>0.925</td>
<td>0.0000120</td>
<td>0.00656152</td>
</tr>
<tr>
<td>10</td>
<td>QS</td>
<td>( AGB_{pSc, QS} = 9.2671 \times CHP_{pSc, QS} )</td>
<td>43379144</td>
<td>0.999</td>
<td>0.998</td>
<td>0.0000001</td>
<td>0.00025905</td>
</tr>
<tr>
<td>11</td>
<td>QRQS</td>
<td>( AGB_{pSc, QRQS} = 9.037 \times CHP_{pSc, QRQS} )</td>
<td>104841889</td>
<td>0.999</td>
<td>0.999</td>
<td>0.0000005</td>
<td>0.00068011</td>
</tr>
<tr>
<td>12</td>
<td>All</td>
<td>( AGB_{pSc, All} = 8.85869 \times CHP_{pSc, All} )</td>
<td>1722352935</td>
<td>0.999</td>
<td>0.999</td>
<td>0.0000001</td>
<td>0.00050963</td>
</tr>
</tbody>
</table>

Satellite image data with dummy variables

<table>
<thead>
<tr>
<th>Model</th>
<th>Plots</th>
<th>Equation</th>
<th>SQR</th>
<th>R²</th>
<th>R²adj</th>
<th>PRESS</th>
<th>APRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>All</td>
<td>( AGB_{dp, ps} = 4.836 \times CHP_{ps} ) + 2191.3997 \times dQR + 3944.2222 \times dQS + 2532.8074 \times dQSR</td>
<td>104975100</td>
<td>0.953</td>
<td>0.949</td>
<td>0.00000151</td>
<td>0.00061458</td>
</tr>
<tr>
<td>14</td>
<td>All</td>
<td>( AGB_{dpSc, ps} = 8.6029 \times CHP_{ps} ) + 6425.9261 \times dQR + 6105.1513 \times dQS + 4827.1112 \times dQSR</td>
<td>1235713930</td>
<td>0.991</td>
<td>0.999</td>
<td>0.0000007</td>
<td>0.00070861</td>
</tr>
</tbody>
</table>

AGB: above-ground biomass in kg, CHP: crown horizontal projection in m²; p: plot; i: inventory data; c: cumulative values; s: satellite data; QS: cork oak, QSQR: evergreen oaks, All: all plots, d: dummy variable.

Table 2. Properties of the fitted models.

The statistical properties and the PRESS and APRESS statistics of the linear models for plot \( CHP_p \) (models 1–3) and \( CHP_s \) (models 4–6) are indicative of their good performance (Table 2).
When comparing the models for the four sets of plots, the statistical properties of the pure plots are better than those of the evergreen oak plots and all plots (Table 2), due to the increase in variability in the latter two. The results attained by [51] for holm oak pure plots are also better when compared with the evergreen oak and all plots of this study. Nonetheless, smaller differences in model performance are found for models 4–6 when compared with models 1–3. The differences might be partially due to the small differences in habit of the two species, the spatial distribution pattern of the individuals of both species and silvicultural practices, especially pruning where focus is put in the crown enlargement more for holm oak than for cork oak. This is also reflected in the overall model error. Larger overall errors are found for the models 2 and 3 (24.3% and 23.7%, respectively) when compared with model 1 (13.6%), and models 5 and 6 (9.8% and 9.2%, respectively) when compared with model 4 (4.4%).

2.3.3. Above-ground biomass allometric functions with satellite image data

The statistical properties and the validation statistics of the linear models for $CHP_{ps}$ and $CHP_{psc}$ for pure cork oak plots (models 7 and 10), for evergreen oak plots (models 8 and 11) and for all plots (models 9 and 12, Figure 4a, b) are indicative of their good performance (Table 2). These results are in accordance with [26] that states that linear regression behaves well in stands with one layer, while [47] refers that in multi-layered stands nonlinear functions are better suited. Noteworthy is that models 10–12 have better statistical properties than the models 7–9. When comparing the models for evergreen oaks with those for pure plots, the statistical properties of the former are worse than of latter (Table 2). This, as already referred, is due to the larger variability observed in the mixed plots. When comparing the models fitted with inventory and satellite image data, the performance is better for the former, which can be partially explained by the differences in the calculation of $CHP$ with the sets of data. Models for cumulative values have slightly better performance for the inventory (models 4–6) than for the satellite image data (models 10–12), corroborated by the overall error of 3.7, 4.4, 9.8 and 9.2%, and 13.4, 4.8, 6.4 and 5.8%, respectively.

Model 10, for cork oak pure stands, has larger errors up to 4000 m$^2$ and an irregular trend afterwards. For error thresholds of 10 and 5% $CHP$, areas larger than 4000 and 5000 m$^2$, respectively, are needed. Considering a mean $CHP$ area of 4486 m$^2$ha$^{-1}$, the corresponding stand area is 0.9 and 1.1 ha, respectively. Model 11, for evergreen oak stands, shows higher errors than the former models for $CHP$ areas up to 1000 m$^2$, with a continuous decrease, in absolute value, up to 3000 m$^2$ and stabilisation afterwards. Errors smaller or equal than 10 and 5% are attained for $CHP$ areas of 3000 and 7000 m$^2$, respectively. The mean $CHP$ area for these stands is 2847 m$^2$ha$^{-1}$, which corresponds to the above-referred errors to stand areas of 1.1 and 2.5 ha, respectively. Model 12, for both pure and mixed stands, shows the same error trend as model 11. Errors below the aforementioned are attained for $CHP$ areas of 2000 and 8000 m$^2$. As the mean $CHP$ area is 3503 m$^2$ha$^{-1}$, the corresponding stand areas are 0.6 and 2.3 ha. As 92% of the holm oak stands and 89% of the cork oak stands have areas bigger than 2 ha [65], this allometric function can be used at local and regional level.
The inclusion of independent variables identifying the plot composition originated a generalised improvement in the statistical properties of the models (Table 2) for \( CHP_{ps} \) (model 13) and \( CHP_{psc} \) (model 14), when compared with those without composition variables (models 9 and 12). In fact, plot composition has revealed important as for both models the three dummy variables were included. Also, multicollinearity did not present any problem as VIF values...
were always lower than 10, being for $CHP$, $dQR$, $dQS$ and $dQRQS$, 8.5, 3.8, 3.6 and 3.2 for model 13, and 4.6, 1.8, 3.0 and 1.8 for model 14, respectively. The models accuracy increases from the models with only $CHP$ (9 and 12) to those that include the dummy variables (models 13 and 14). Noteworthy is also the decrease in $SQR$ of 36% from models 9 to 13, and of 39% from models 12 to 14, with the corresponding increase in the adjusted regression coefficients. There is a considerable improvement in the overall error from models 9 to 13, from 23.1 to 5.2% and from models 12 to 14, from 5.8 to 0.1%. The error for model 14 shows an irregular trend with error smaller than 10 and 5% for cumulative crown projection for areas larger than 1500 and 2000 m$^2$, corresponding to stand areas of 0.4 ha and 0.6 ha, respectively (Figure 5). Considering the threshold of 5% for model 14, -1.7 ha is needed than for model 12.

All the fitted models (models 7–14) have a high goodness of fit, with more than 90% of the variability explained. It seems that crown horizontal projection as better predictive abilities with linear functions than the shadow fraction processed from QuickBird images [49]. The models developed in this study show also better performances than multiple regression functions with crown diameter and total height as independent variables, derived from a panchromatic band of QuickBird satellite [50] or than exponential functions with reflectance of band 3 or NDVI for IKONOS image [47]. Likewise, biomass estimation functions derived from medium spatial resolution satellite image data do not show as good performance as those obtained in this study. This is probably related to dissimilarity between ground and satellite data [27]. The former was confirmed in a study with Landsat 5 TM and MODIS images for $Pinus pinaster$ stands, whose independent variable is vegetation indices [17].

2.4. Conclusions

Remote sensing gives a good contribution to estimate the above-ground biomass. The high spatial resolution satellite image allows mapping with high accuracy the above-ground biomass at local and regional scale.

The overestimation of crown horizontal projection per plot using high spatial resolution satellite image, when compared with that calculated using inventory data, is related to the inclusion of mixed pixels in the boundary of tree crown delimitation in the multi-resolution segmentation and trees’ spatial distribution. Though the cork and holm oak are visually similar, their spectral signature is sufficiently different to obtain a very good classification by forest specie.

In general, above-ground biomass allometric functions for plot and their cumulative values showed a good performance, though better for pure stands than for mixed stands, which might be explained by the larger variability observed in the latter. However, the inclusion of dummy variables, which reflect the differences between the species and the stand structure, originates a generalised improvement in the functions performance. The same threshold errors of 10 and 5% are attained by the latter with less 33 and 70% of the stand areas. Considering that to a 10% error correspond stand areas of 1.1 ha, it includes about 90% of the area occupied by these species in Portugal [65]. This method has the advantage of enabling the overall area evaluation, not requiring forest inventory or extrapolation procedures. It can be used when species composition is not (models 9 and 12) and is (models 13 and 14) differentiated.
Acknowledgements

The authors would like to thank the forest producers for allowing plot installation and to Paulo Mesquita for satellite image processing. The work was financed by Programa Operativo de Cooperación Transfronteriza España—Portugal (POCTEP); project Altercexa—Medidas de Adaptación y Mitigación del Cambio Climático a Través del Impulso de las Energias Alternativas en Centro, Alentejo y Extremadura (Ref. 0317_Altercexa_I_4_E and 0406_ALTERCEXA_II_4_E); TrustEE—innovative market based Trust for Energy Efficiency investments in industry (Project ID: H2020—696140); and National Funds through FCT—Foundation for Science and Technology under the Project UID/AGR/00115/2013.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGB</td>
<td>Above-Ground Biomass per tree (kg)</td>
</tr>
<tr>
<td>AGB_t</td>
<td>Above-Ground Biomass per tree (kg)</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>All</td>
<td>all plots</td>
</tr>
<tr>
<td>ALOS</td>
<td>Advanced Land Observing Satellite</td>
</tr>
<tr>
<td>APRESS</td>
<td>Absolute Predicted Residual Error Sum</td>
</tr>
<tr>
<td>ASTER</td>
<td>Advanced Spaceborne Thermal Emission and Reflection Radiometer</td>
</tr>
<tr>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometer</td>
</tr>
<tr>
<td>c</td>
<td>cumulative values</td>
</tr>
<tr>
<td>CC</td>
<td>Crown Cover (%)</td>
</tr>
<tr>
<td>CHP</td>
<td>Crown Horizontal Projection (m²)</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>d</td>
<td>dummy variable</td>
</tr>
<tr>
<td>dbh</td>
<td>diameter at breast height, measured at 1.30 m height (cm)</td>
</tr>
<tr>
<td>dg</td>
<td>mean quadratic diameter (cm)</td>
</tr>
<tr>
<td>G</td>
<td>basal area per hectare (m²ha⁻¹)</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>i</td>
<td>inventory data</td>
</tr>
<tr>
<td>ETM+</td>
<td>Enhanced Thematic Mapper +</td>
</tr>
<tr>
<td>LAI</td>
<td>Leaf Area Index</td>
</tr>
<tr>
<td>max</td>
<td>maximum value</td>
</tr>
<tr>
<td>mean</td>
<td>arithmetic mean value</td>
</tr>
<tr>
<td>min</td>
<td>minimum value</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>N</td>
<td>number of trees per hectare (tree ha⁻¹)</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalized Difference Vegetation Index</td>
</tr>
</tbody>
</table>
NOAA National Oceanic and Atmospheric Administration

\( p \) plot

PALSAR Phased Array type L-band Synthetic Aperture Radar

PRESS Predicted Residual Error Sum of Squares

QR holm oak (\textit{Quercus rotundifolia}) pure plots

QRQS mixed plots of holm and cork oaks (evergreen oaks)

QS cork oak (\textit{Quercus suber}) pure plots

\( R^2 \) coefficient of determination

\( R_{adj}^2 \) adjusted coefficient of determination

RADAR RAdio Detection And Ranging

RMSE Root Mean Square Error

s satellite data images

SAR Space-borne synthetic Aperture Radar

SD Standard Deviation

SQR Sum of Squares of the Residuals

t tree

TM Thematic Mapper

VIF Variance Inflation Factor

WiFS Wide Field Sensor

\( \text{ww} \) wood biomass (kg)

\( \text{wb} \) bark biomass (kg)

\( \text{wc} \) crown biomass (kg)

Author details

Adélia M. O. Sousa*, Ana Cristina Gonçalves and José R. Marques da Silva

*Address all correspondence to: asousa@uevora.pt

Department of Rural Engineering, School of Sciences and Technology, Institute of Mediterranean Agricultural and Environmental Sciences (ICAAM), Institute of Research and Advanced Information (IIFA), University of Évora, Évora, Portugal

References


