iForest
Biogeosciences and Forestry

Research Article
doi: 10.3832/ifor1989-010
vol. 10, pp. 590-596

Estimation of aboveground forest biomass in Galicia (NW Spain) by the combined use of LiDAR, LANDSAT ETM+ and National Forest Inventory data

Enrique Jiménez [(1)], José A Vega [(1)], José M Fernández-Alonso [(1)], Daniel Vega-Nieva [(2)], Luis Ortiz [(3)], Pablito M López-Serrano [(2)], Carlos A López-Sánchez [(2)]

Assessing biomass is critical for accounting bioenergy potentials and monitoring forest ecosystem responses to global change and disturbances. Remote sensing, especially Light Detection and Ranging (LiDAR) data combined with field data, is being increasingly used for forest inventory purposes. We evaluated the feasibility of the combined use of freely available data, both remote sensing (LiDAR data provided by the Spanish National Plan for Aerial Ortophotography - PNOA - and Landsat vegetation spectral indices) and field data (from the National Forest Inventory) to estimate stand dendrometric and aboveground biomass variables of the most productive tree species in a pilot area in Galicia (northwestern Spain). The results suggest that the models can accurately predict dendrometric and biomass variables at plot level with an R^2 ranging from 0.49 to 0.65 for basal area, from 0.65 to 0.95 for dominant height, from 0.48 to 0.68 for crown biomass and from 0.55 to 0.82 for stem biomass. Our results support the use of this approach to reduce the cost of forest inventories and provide a useful tool for stakeholders to map forest stand variables and biomass stocks.

Keywords: Biomass Maps, Forest Inventory, LiDAR, Landsat Vegetation Indices

Introduction

Measurement and mapping of aboveground forest biomass at diverse scales is critical for estimating global carbon storage and assessing those ecosystem responses to climate change and anthropogenic disturbance (Ni-Meister et al. 2010). Estimation of forest biomass is of great importance because of the increasing value of this renewable resource for energy production (González-Ferreiro et al. 2013). Furthermore, biomass estimation and mapping are used to quantify fuel loads for fire risk assessment and fire prevention planning purposes (González-Olabarria et al. 2012).

The method most commonly used to estimate aboveground forest biomass is the forest inventory based on plot data. However, given the high costs and operational difficulties associated with this method, the use of remotely sensed data in combination with field-work is becoming increasingly popular for forest inventory purposes (Ji et al. 2012). Moreover, remote-sensing methods enable the estimation of stand dendrometric and biomass values at each pixel location, rather than estimation of average or total biomass within a given area (Jenkins et al. 2001).

Most studies using remote-sensing methods have been based on medium-to-high resolution sensors, especially the Landsat-TM sensor, given the good compromise between spectral and temporal resolution of the data achieved (Ji et al. 2012, López-Serrano et al. 2015). Remote-sensing indices have been used as explanatory variables for models of forest biomass and productivity (Manes et al. 2010). LiDAR (Light Detection and Ranging) technology has been recognized as a much more efficient, accurate and cost effective approach to sensing aboveground biomass. This technology is generally regarded as a more accurate method because LiDAR sensors provide information about vertical height of individual pulses returns, which can then be used to predict canopy attributes (Chen 2013). Previous studies have demonstrated the success of LiDAR estimates of aboveground biomass based on the relationship between LiDAR metrics and field measurements of biomass obtained from allometric models (Zhao et al. 2009). In this study we estimated aboveground forest biomass in Galicia (NW Spain) by the combined use of LiDAR, LANDSAT ETM+ and field data.

Two main types of statistical approaches are applied to estimate biomass from remotely sensed data: one focused on the data and the other on algorithms (Breiman 2001). The first assumes a stochastic model (linear, nonlinear regressions) used both to predict responses of population units not in the sample and future responses (Ji et al. 2012, González-Ferreiro et al. 2012, 2013). The second approach (generally called machine-learning models) has been reported for implicitly inferring unknown relationships underlying a given dataset, being versatile enough to uncover complicated nonlinear relationships (Cortés et al. 2014). In

(1) Centro de Investigación Forestal - Lourizán, PO Box 127, 36080 Pontevedra (Spain); (2) Facultad de Ciencias Forestales - Universidad Juárez del Estado de Durango (México) Río Papaloapan, Valle del Sur, 34120 Durango, Dgo. (México); (3) Departmento de Ingeniería de Recursos Naturales y Medio Ambiente, Universidad de Vigo, Campus A Xunqueira, Pontevedra, 36005 (Spain)

@ Enrique Jiménez (cordogaita@gmail.com)

Received: Jan 20, 2016 - Accepted: Mar 12, 2017

Citation: Jiménez E, Vega JA, Fernández-Alonso JM, Vega-Nieva D, Ortiz L, López-Serrano PM, López-Sánchez CA (2017). Estimation of aboveground forest biomass in Galicia (NW Spain) by the combined use of LiDAR, LANDSAT ETM+ and National Forest Inventory data. iForest 10: 590-596. - doi: 10.3832/ifor1989-010 [online 2017-05-15]

Communicated by: Piermaria Corona

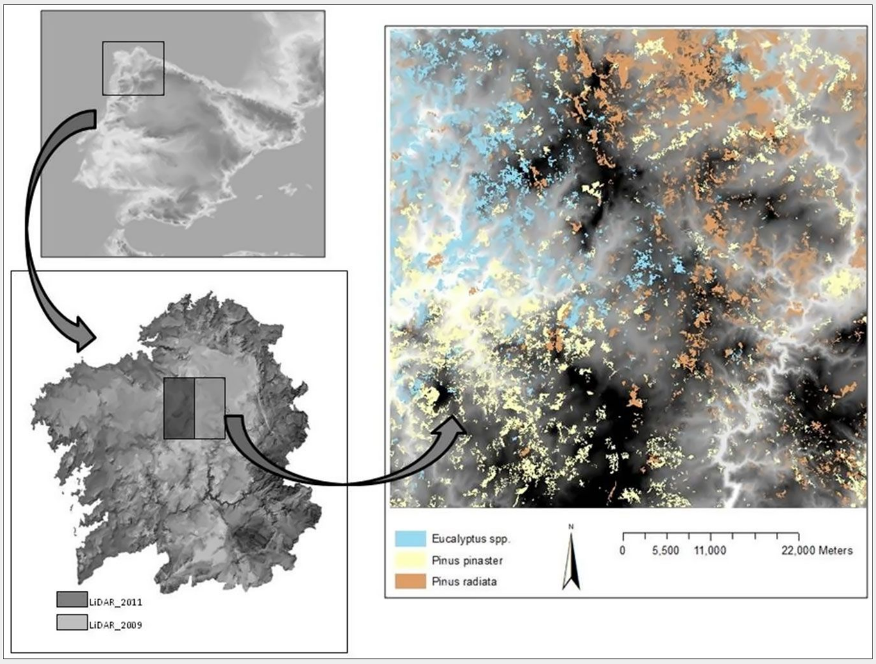

Fig. 1 - Location of the study area and National Forest Map polygons dominated by *Pinus pinaster, Pinus radiata* and *Eucalyptus* spp.

the present study the first approach was chosen.

Forest inventories based on remotely-sensed data can be carried out at two different spatial scales: the individual tree level (associated with spatially dense LiDAR data, > 5 points m^{-2} – Kaartinen et al. 2012) and the area-based, which can also be used with low density LiDAR data (Zhao et al. 2009, González-Ferreiro et al. 2012, González-Ferreiro et al. 2013). In this study, we focused on area-based forest inventories, as we used low density LiDAR data provided by the Spanish National Plan for Aerial Ortophotography (PNOA) for the Spanish territory (0.5 pulse m^{-2}), being higher point densities required for the estimation at individual tree level (Shendryk et al. 2014). The use of LiDAR for aboveground biomass estimations is still considered a challenging task, especially when using low density scanning systems, due to its lower accuracy in comprehensively reflecting the canopy structure (Lin et al. 2010). However, low density LiDAR has been shown to be quite effective for predicting biomass and other variables in temperate forests (Naesset et al. 2011, Nord-Larsen & Schumacher 2012, McRoberts et al. 2013).

For large areas covered by different types of vegetation, models based on remote sensing data are usually developed for each vegetation type, as this kind of models is greatly influenced by species and stand structure (Naesset et al. 2011, Nord-Larsen & Schumacher 2012, Lafiti et al. 2015). Implementation of models for mapping stand dendrometric and biomass variables requires an accurate classification of the area (Naesset et al. 2011, Nord-Larsen & Schumacher 2012, Lafiti et al. 2015).

Galicia has one of the greatest potential forest production rates in Europe, varying from 6 to 12 m^3 ha^{-1} $year^{-1}$ depending on the particular species (Chas-Amil 2007) and producing 45% of Spain's timber and 4.5% of Europe's. The most productive and fast-growing tree species in Galicia are *Eucalyptus* spp. (*Eucalytus globulus* Labill. and *Eucalyptus nitens* Deane et Maiden), *Pinus pinaster* Ait. and *Pinus radiata* D. Don, all of which are grown in extensive commercial forest plantations to produce panelboard, sawlog and pulpwood (Merino et al. 2005). Stands dominated by these tree species account for 133,224,356 m^3 of standing timber (69.1% of the total tree standing timber volume) in Galicia, and cover an area of 896,342 ha (62.9% of the total tree-covered area) in the region (MARM 2011).

The objective of this study was to evaluate the feasibility of the combined use of freely available remote sensing (low density LiDAR and Landsat) and National Forest Inventory data combined with the National Forest Map (based on PNOA aerial ortophoto, with a scale of 1:25,000 and minimum mapping unit of 1 ha) to estimate stand dendrometric and aboveground biomass variables of the most productive tree species in a pilot area in Galicia. This is the first study using this freely available information for these tree species in the NW Spain. The results may also be applicable to similar areas covered by similar sensors.

Material and methods

The study was conducted in the municipality of Palas de Rei (Lugo), in northwestern Spain (42° 52′ 23″ N; 07° 52′ 08″ W – Fig. 1). The study area consisted of a 60 × 60 km square, centered in Palas de Rei encompassing parts of the provinces of Lugo, Pontevedra and A Coruña. The mean elevation of the area is 525 m (range 147-1179 m) and the mean slope is 11.3%. The climate is Mediterranean, with a continental influence. The average precipitation is 1188 mm $year^{-1}$, and the mean annual temperature is 12.3 °C, with maximum temperatures in August (19.0 °C) and minimum temperatures in January (6.8 °C). Forests (more than 20% of tree cover) occupy 150,396.6 ha of the total surface area. The main tree species in the area are *Pinus pinaster* (covering 27.8% of the total forest area), *Pinus radiata* (22.7%), *Quercus robur* (30.7%) and *Eucalyptus* spp. (mainly *E. nitens* and *E. globulus* – 13.8%).

Field data

The Fourth Spanish National Forest Inventory (SNFI-4) plots were used as the source of field data for this study. Plots within the study area dominated by the three selected tree species, with tree cover > 20% and presence of trees with diameter at breast height (dbh) > 7.5 cm, were selected for the study (MARM 2011). The plots were established at the intersections of a 1 × 1 km grid, totaling 873 plots over the whole study area. In 749 of these, tree cover was > 20%, with presence of trees with dbh > 7.5 cm. In total, 159 plots were dominated by *Pinus pinaster*, 142 by *Pinus radiata* and 109 by *Eucalyptus* spp. The measurements were carried out during 2008 and 2009 (Tab. 1).

Sample plots consisted of four concentric circles of radii 5, 10, 15 and 25 m, in which dbh and total height were measured in all trees of dbh > 7.5, 12.5, 22.5 and 42.5 cm, respectively. Diameter at breast height was measured to the nearest 0.1 cm with a graduated caliper, and tree height was measured to the nearest 0.1 m with a hypsometer. The number of stems per hectare, basal area and dominant height (mean height of the 100 thickest trees per

Tab. 1 - Mean (± standard error) tree density, basal area (G), dominant height (H_0), crown biomass (W_{cr}), stem biomass (W_{st}) per sample plot of the National Forest Inventory within the study area for each species.

Main species	Density (trees ha^{-1})	G (m^2 ha^{-1})	H_0 (m)	W_{cr} (Mg ha^{-1})	W_{st} (Mg ha^{-1})
Eucalyptus spp. (n=109)	783 ± 48	17.0 ± 1.4	25.5 ± 1.1	17.1 ± 1.6	89.9 ± 10.0
Pinus pinaster (n=159)	579 ± 39	17.3 ± 1.1	20.6 ± 0.6	19.6 ± 1.2	58.4 ± 4.1
Pinus radiata (n=142)	646 ± 36	19.9 ± 1.1	20.4 ± 0.6	20.0 ± 1.1	70.8 ± 4.8

hectare) were calculated from the tree variable measurements by the use of expansion factors. These factors express the number of trees per hectare that each measured tree represents in the inventory in relation to the subplot radius.

Aboveground biomass per plot was estimated by applying existing allometric models for each measured tree within the plot. Specific allometric models were used for each species, and models constructed for the same ecoregion (NW Spain) were selected. The explanatory variables included in the models were tree variables measured in the National Forest Inventory (dbh and tree height). The allometric models by the following authors were used for the different species: Jiménez et al. (2013) and Gómez-Vázquez et al. (2013), for *Pinus pinaster;* Balboa-Murias et al. (2006), for *Pinus radiata;* Brañas et al. (2000) for *Eucalyptus globulus;* and Pérez-Cruzado & Rodríguez-Soalleiro (2011) for *Eucalyptus nitens*. For other tree species inside the plots, we used the model by Balboa-Murias (2005) for *Quercus robur* and those reported by Montero et al. (2005) for other species. After applying the models, we estimated the biomass values for following components in each tree measured: leaves, stem, fine branches (diameter < 2 cm) and coarse branches (diameter ≥ 2 cm). To obtain the aboveground biomass at plot level, we used the above-mentioned expansion factor. Aboveground biomass at plot level was subdivided in two components: crown biomass (leaves and branches) and stem biomass (Tab. 1).

The National Forest Map, associated with the National Forest Inventory was obtained from the PNOA aerial ortophoto (scale 1:25,000, minimum mapping unit of 1 ha) and used to classify vegetation types and spatially define the polygons dominated by each of the species under study (Fig. 1).

LiDAR data

The LiDAR data were provided by the PNOA. The study area was surveyed at two different times: once in 2009 (province of Lugo) and then in 2011 (provinces of Pontevedra and A Coruña). Data were delivered in 2 × 2 km tiles of points in LAS binary files. The resulting LiDAR point density for the study area was 0.5 pulse m^{-2}, with a vertical accuracy greater than 0.2 m. A total of 961 LAS files were required in order to cover the study area.

The LiDAR data were processed using the software FUSION LDV 3.50 (McGaughey 2009). After eliminating the noise from the point cloud, the Digital Elevation Model (DEM) of the study area was obtained. This was done by first filtering the ground returns (using the "GroundFilter" tool) by implementing a filtering algorithm, adapted from Kraus & Pfeiffer (1998). The DEM (1 m spatial resolution) was then generated using these returns through the "GridSurfaceCreate" command and used to normalize the heights of the point cloud. The LiDAR height and intensity statistics for each sample plot were obtained using the "ClipData" and "CloudMetrics" commands and the plot boundaries (25 m radius). A predefined threshold of 2 m above the ground was applied in order to exclude returns not corresponding to crowns (*e.g.*, understory, rocks, shrubs).

Landsat ETM+ data

We obtained a cloud-free Landsat ETM+ scene corresponding to a date close to the National Forest Inventory and LiDAR survey (1 June 2009). Digital numbers were converted to radiometric values by using the specific gain and offset of the sensor. Reflectance was obtained using the method of NASA (2011). The images (SLC-off) were fused using the method described by Scaramuzza et al. (2004). Reflectance values were used to obtain several vegetation indices (Tab. 2) for each plot sampled.

Regression models

Linear, power function and exponential models were used to estimate how plot values (stand variables: basal area and dominant height; aboveground biomass: crown and stem biomass) were related to LiDAR variables (height and intensity metrics) and Landsat vegetation indices. Models were obtained for each dominant species (*Pinus pinaster, Pinus radiata* and *Eucalyptus* spp.) and for each LiDAR survey area (2009 and 2011). Separate models were constructed for different species because they have different structural characteristics (*e.g.,* tree architecture, canopy stratification, canopy density, etc.) that generate differences in the models (Chen 2013, Cortés et al. 2014), and because the accuracy of prediction depends on the type of forest (Chen 2013). The LiDAR surveys were separated because the quality of laser-derived data depends on flight height, scan angle, point density and footprint size, among other factors. In the case of *Eucalyptus* spp., we only used the LiDAR_2011, as only a few of the National Forest Inventory plots dominated by this species were surveyed in 2009. The model expressions are as follows (eqn. 1, eqn. 2, eqn. 3):

$$\text{Linear: } Y=\alpha+\sum_{i=1}^{m}\beta_i X_i+\varepsilon$$

$$\text{Power: } Y=\alpha+\prod_{i=1}^{m}X_i^{\beta_i}+\varepsilon$$

$$\text{Exponential: } Y=\exp\left(\alpha+\sum_{i=1}^{m}\beta_i X_i\right)+\varepsilon$$

where *Y* represents the field values (stand dendrometric and aboveground biomass variables), X_i represents a set of *m* independent variables (height and intensity LiDAR variables and Landsat vegetation indices), α and β_i (i = 1, …, m) are parameters to be estimated, and ε is the error term.

Linear models were fitted by stepwise regression. Power and exponential models were fitted by nonlinear regression after prior linearization (taking natural logarithms) to select (by linear regression) the most significant subset of independent variables to include in the model. To correct for bias in log-transformed allometric equations, we adopted a correction factor (Sprugel 1983). Heterocedasticity and multicollinearity among explanatory variables were checked. The presence of multicollinearity among variables was evaluated by the condition number (Belsley 1991), selecting the models with condition number smaller than 10 (collinearity is not a major problem). For each dominant tree species and LiDAR survey, the selected model included the combination of independent variables with the largest coefficient of determination (R^2, defined as the square correlation coefficient between the measured and estimated values), the smallest values of the Akaike's information criterion (AIC – Burnham & Anderson 1998) and the root mean squared error of the estimate (RMSE).

Raster files were obtained for each explanatory variable (for LiDAR variables through the "CSV2Grid" FUSION command) with a spatial resolution similar to the Landsat image (30 m). The constructed models were used for spatial extrapolation of stand dendrometric and aboveground biomass variables to the study area, and the Spanish Forest Map was used to identify polygons dominated by the species considered. To avoid the inclusion of bias in the spatial extrapolation using model predictions, model-assisted estimators were

Tab. 2 - Landsat vegetation spectral indices employed in the study. (ρ_{NIR}): Near infrared; (ρ_R): Red; (ρ_B): Blue.

Index	Equation	Reference
Simple Ratio (SR)	$SR = \rho_{NIR} / \rho_R$	Jordan (1969)
Normalized Difference Vegetation Index (NDVI)	$NDVI = (\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R)$	Rouse et al. (1973)
Soil-Adjusted Vegetation Index (SAVI)	$SAVI = (1+0.5) \cdot (\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R + 0.5)$	Huete (1988)
Normalized Ratio Vegetation Index (NRVI)	$NRVI = (\rho_R/\rho_{NIR} - 1)/(\rho_R/\rho_{NIR} + 1)$	Baret & Guyot (1991)
Enhanced Vegetation Index (EVI)	$EVI = 2.5 \cdot (\rho_{NIR} - \rho_R)/(\rho_{NIR} + 6 \cdot \rho_R + 7.5 \cdot \rho_B + 1)$	Liu & Huete (1995)

Tab. 3 - Results of basal area (G), dominant height (H_0), crown biomass (W_{cr}), stem biomass (W_{st}) modelling. (RMSE): root of mean squared error; (AIC): Akaike's information criterion; (LIN): linear models; (EXP): exponential models; (Elmax): Maximum height; (1st Ret above 2): Number of first return above 2 m; (IntL3): L3 moments of intensity; (Int L4): L4 moments of intensity; (Elp90): Height 90th pertencile value; (Elp95): Height 95th pertencile value; (Elskewness): Height skewness; (Perc Ret above mode): Percentage of all returns above the mode height; (Ret above mean): number of returns above the mean value; (Ret above 2): number of returns above 2 m; (Elp10): height 10th percentile value; (ElCURTmenaCUBE): Cubic mean; (Ret2): Number of second returns; (IntLskewness): Moment ratio of intensity skewness; (ElCV): Coefficient of variation of height; (Ret3): Number of third returns; (Elp30): height 30th pertencile value; (Elp40): Height 40th pertencile value; (Intmode): Intensity mode; (Elp75): Height 75th pertencile value; (Elp01): height 1st pertencile value; (Int01): Intensity 1st pertencile value; (SR): Landsat Simple Ratio; (SAVI): Landsat SAVI Index; (ElL4): Moment L4 of height.

Main species	Variable	Model	R^2	RMSE	AIC
Eucalyptus spp. LiDAR_2011 (n=85)	G (m^2 ha^{-1})	EXP (Elmax - 1st Ret above 2 - IntL3)	0.49	8.7	334
	H_0 (m)	EXP (Elmax - IntL4)	0.65	5.9	273
	W_{cr} (Mg ha^{-1})	EXP (Elmax)	0.48	10.2	1390
	W_{st} (Mg ha^{-1})	LIN (Elmax - Elp90 - Elp95 - Elskewness)	0.55	58.0	1657
Pinus pinaster LiDAR_2009 (n=54)	G (m^2 ha^{-1})	LIN (Elp95 - 1st Ret above 2)	0.62	9.2	168
	H_0 (m)	LIN (Elp95 -Perc Ret above mode - Elp90)	0.95	1.3	27
	W_{cr} (Mg ha^{-1})	LIN (Elp95 - Ret above mean)	0.59	10.8	677
	W_{st} (Mg ha^{-1})	EXP (Elp95 -Ret above 2 - Elp10)	0.65	33.7	761
Pinus pinaster LiDAR_2011 (n=105)	G (m^2 ha^{-1})	EXP (ElCURTmeanCUBE - Ret2 - IntLskewness - ElCV)	0.61	7.2	406
	H_0 (m)	LIN (Elmax - Elskewness)	0.73	3.9	283
	W_{cr} (Mg ha^{-1})	EXP (Elmax - Ret3 - IntLskewness - Elp30 - IntL4)	0.58	8.4	1822
	W_{st} (Mg ha^{-1})	EXP (Elmax - Ret2 - IntLskewness - Elp40)	0.60	27.4	2056
Pinus radiata LiDAR_2009 (n=82)	G (m^2 ha^{-1})	LIN (ElCURTmeanCUBE - Intmode - Elp75)	0.61	7.6	322
	H_0 (m)	LIN (Elmax - Elp01 - Elp99 - Ret above 2)	0.86	2.4	147
	W_{cr} (Mg ha^{-1})	LIN (ElCURTmeanCUBE - Intmode - Elp75 - Intp01)	0.64	7.6	1388
	W_{st} (Mg ha^{-1})	LIN (Elp99 - ElCURTmeanCUBE)	0.62	34.2	1616
Pinus radiata LiDAR_2011 (n=60)	G (m^2 ha^{-1})	LIN (ElCURTmeanCUBE - SR)	0.65	5.5	158
	H_0 (m)	EXP (Elmax - SR)	0.68	3.7	123
	W_{cr} (Mg ha^{-1})	LIN (ElCURTmeanCUBE - SR)	0.68	5.2	761
	W_{st} (Mg ha^{-1})	EXP (Elmax - SR - SAVI - ElL4)	0.82	17.0	869

employed to include a correction for estimated bias (McRoberts et al. 2013). This estimate is adjusted for deviations between the model predictions and the observed values in the sample (eqn. 4):

$$MA\ mean=\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i-\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)$$

where *MA mean* is the model-assisted regression estimator of means, *N* is the population size, $\hat{y}_i$ is obtained from the models using the model parameters estimates and ε = 0. The first term in the equation is the mean of the model predictions ($\hat{y}_i$) for all population units, and the second term is an estimate of bias calculated over the sample units and compensates for systematic model prediction errors.

Results

The models obtained for each dominant species and LiDAR survey, the goodness-of-fit-statistics for the most significant model constructed and the stand dendrometric and aboveground biomass variables are shown in Tab. 3. Linear and exponential models were selected on the basis of their performance. *Eucalyptus* spp. stands produced the least quality of fit in terms of R^2, with models providing values of 49% for basal area, 65% for dominant height, 48% for crown biomass and 55% for stem biomass. By contrast, *Pinus radiata* stands (LiDAR_2011 data) produced the greatest quality of fit for basal area (65%), crown biomass (68%) and stem biomass (82%) (Fig. 2). The greatest quality of fit for dominant height was obtained for *Pinus pinaster* (LiDAR_2009 data), with a value of 95%.

Independent variables related to the height distribution metrics were included in all models (Tab. 3). Intensity metrics appeared as explanatory variables in *Eucalyptus* spp. (basal area and dominant height), *Pinus pinaster* (LiDAR_2011 data: basal area, crown biomass and stem biomass) and *Pinus radiata* (LiDAR_2009 data: basal area and crown biomass – Tab. 3). Landsat derived vegetation indices were explanatory variables in *Pinus radiata* (LiDAR_2011 data), with Simple Ratio (SR) in basal area, dominant height and crown biomass models, and Soil-Adjusted Vegetation Index (SAVI) in the stem biomass model. The inclusion of Landsat resulted in an increase in the R^2 values (from 0.64 to 0.65 for basal area, from 0.64 to 0.68 for dominant height, from 0.66 to 0.68 for crown biomass and from 0.60 to 0.82 for stem biomass), and reductions in RMSE (from 5.6 to 5.5 m^2 ha^{-1} for basal area, from 3.9 to 3.7 m for dominant height, from 5.4 to 5.2 Mg ha^{-1} for crown biomass and from 25.5 to 17.0 Mg ha^{-1} for stem biomass) and AIC values (from 161 to 158 for basal area, from 126 to 123 for dominant height, from 767 to 761 for crown biomass and from 899 to 869 for stem biomass).

Tab. 4 displays the model-assisted estimates for each species of the mean values of basal area and dominant height, crown and stem biomass which were obtained by applying the models achieved to the polygons dominated by each analyzed species in the whole study area. The standard errors for estimates of the means were small, ranging from 0.5 to 0.8 m^2 ha^{-1} for basal area, from 0.3 to 0.7 m for dominant height, from 0.6 to 0.8 Mg ha^{-1} for crown

Tab. 4 - Estimates mean values (± standard error) of basal area ($\hat{G}$) and dominant height ($\hat{H}_0$), crown biomass ($\hat{W}_{cr}$) and stem biomass ($\hat{W}_{st}$) for each species for the whole study area.

Main species	$\hat{G}$ (m^2 ha^{-1})	$\hat{H}_0$ (m)	$\hat{W}_{cr}$ (Mg ha^{-1})	$\hat{W}_{st}$ (Mg ha^{-1})
Eucalyptus spp.	12.8 ± 0.8	19.0 ± 0.7	11.4 ± 0.8	32.6 ± 5.2
Pinus pinaster	13.9 ± 0.7	14.3 ± 0.4	11.0 ± 0.6	29.5 ± 2.7
Pinus radiata	13.6 ± 0.6	14.7 ± 0.3	14.5 ± 0.6	29.9 ± 2.5

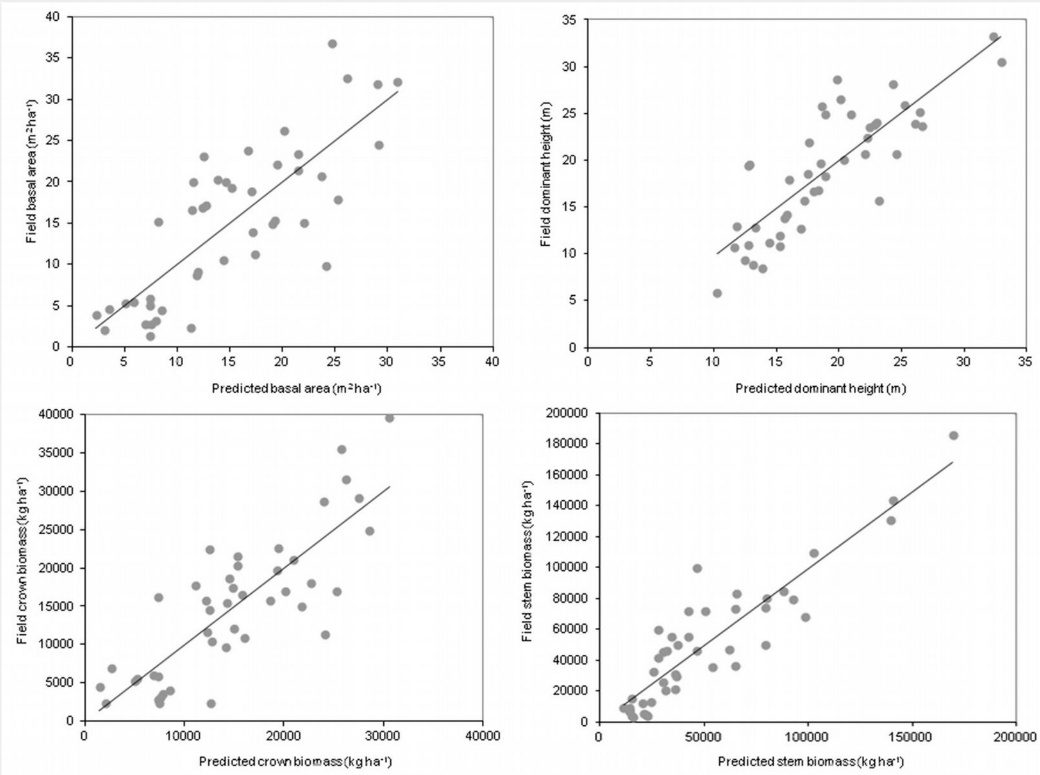

Fig. 2 - Field *vs.* predicted measured values of basal area (a), dominant height (b), crown biomass (c) and stem biomass (d) for *Pinus radiata*.

biomass and from 2.5 to 5.2 Mg ha^{-1} for stem biomass. Bias estimates for the model assisted estimator were also small, from -0.08 to 0.08 m^2 ha^{-1} for basal area, from -0.01 to 1.40 m for dominant height, from -0.01 to 0.01 Mg ha^{-1} for crown biomass and from -0.01 to 0.03 Mg ha^{-1} for stem biomass.

Discussion

The findings of this study demonstrate that stand dendrometric and biomass variables can be predicted with reasonable precision by using low density LiDAR variables in combination with Landsat data. The importance of linear and exponential models has previously been reported in studies relating LiDAR metrics (combined or not with spectral data) and stand dendrometric and biomass variables (González-Ferreiro et al. 2012, 2013), even for shrub vegetation (Estornell et al. 2012). The R^2 values obtained in the present study were of the same order of magnitude or slightly smaller than those previously reported for the same species (González-Olabarria et al. 2012, González-Ferreiro et al. 2012, 2013), probably as a consequence of the use of lower density LiDAR data, or (in the present study), the use of field data from the National Forest Inventory rather than specifically measured field data.

Previous studies have found that height metrics are closely correlated with stand dendrometric and biomass variables for the same species (González-Ferreiro et al. 2012, 2013) and others (Laurin et al. 2014). However, these studies were carried out with higher density LiDAR data and specifically measured field data. The most frequent height metric variables observed in our models are maximum height, the cubic mean of the height and diverse percentiles of height distribution. This is consistent with previously reported correlations between stand dendrometric and/or biomass variables and maximum height (Lim et al. 2003) and also between the former and height percentiles (González-Ferreiro et al. 2012, 2013, Laurin et al. 2014). The most appropriate height metrics reported in the literature widely differ as a likely consequence of differences in the vegetation structure and data processing procedures used (Chen 2013). The close correlation between aboveground biomass values and maximum height and the higher LiDAR percentiles observed in our study may be a result of the regular structure of the forest plantations surveyed (González-Ferreiro et al. 2013).

Although intensity metrics may not always selected as explanatory variables (González-Ferreiro et al. 2013), some studies on pine stands have also reported that the inclusion of intensity variables may improve the predictive power of the models, or may even be decisive if used in combination with density or height LiDAR values (González-Olabarria et al. 2012, González-Ferreiro et al. 2012).

In this study the least quality of fit observed for *Eucalyptus* spp. compared to *Pinus* species is consistent with previous findings (Cortés et al. 2014) and is explained by the sparser canopies of *Eucalyptus* trees, which result in less accurate digital canopy models.

Although there were not previous studies combining LiDAR and satellite data for these species in NW Spain, an improvement in model accuracy by combining both types of data has been previously observed in other tree species (Chen et al. 2012, Cortés et al. 2014). Some authors only used optical imagery for the first step in vegetation classification in order to account for the dependence of stand dendrometric and biomass estimation on vegetation types (Chen et al. 2012). In the present study, this classification was implemented by using the Spanish Forest Map based on the analysis of aerial photographs. We used Landsat derived vegetation indices as explanatory variables in combination with LiDAR variables to improve the stand dendrometric and biomass models (Cortés et al. 2014). In the present study Landsat derived vegetation indices contributed to significantly improve the quality of fit of the model to the data for *Pinus radiata* (LiDAR_2011 data). The spectral data alone had a small explanatory power, as previously observed for other tree species (Estornell et al. 2012, Laurin et al. 2014). However, the improvement in model performance for *Pinus radiata* (LiDAR_2011 data) by the inclusion of Landsat derived vegetation indices is consistent with previous studies in which forest stand structure metrics were predicted using a combination of LiDAR and remote sensing imagery (Cortés et al. 2014, Laurin et al. 2014), although the improvement was relatively small in our case.

The standard error of the model-assisted regression estimates were smaller than those obtained by field sampling (simple random sampling estimates – Tab. 1), confirming that remotely sense data made substantial contributions with greater precision, as compared with estimates based only on plot observations (McRoberts et al. 2013). Mean values of the model-assisted regression values were smaller than those of the plot observations (Tab. 1), likely because field plots were characterized by tree cover > 20% and the presence of trees with dbh > 7.5 cm, while the area covered by remotely sensed data also included locations with smaller tree cover and/or without the presence of trees with dbh > 7.5 cm.

Conclusions

Our findings confirmed the potential of the combined use of freely available remote sensing and regional forest inventories to establish relationships that allow to determine the spatial distribution of both stand dendrometric and aboveground biomass variables of stands dominated by different species in a large area (60 × 60 km) in Galicia. This is the first study combining this freely available information for these species in this area. This information is critical for calibrating and validating biogeochemical models, quantifying carbon fluxes and supporting the United Nations Framework Convention on Climate Change program (Chen 2013).

The future periodicity and availability of this information will enable spatial estimation of stand, biomass and carbon temporal evolution in different vegetation types. The findings of this study support the use of this approach to reduce the cost of forest inventories, thus providing a useful tool enabling stakeholders to map forest stand variables and biomass stocks.

Acknowledgements

This research was funded by Projects of the *Conselleria de Medio Rural e do Mar Xunta de Galicia* (FEADER 2013/38) and INIA RTA 2014-00011-C06 (FEDER). This work was also co-financed by the INIA and European Social Fund (*via* a grant awarded to E. Jiménez). We also thank the anonymous reviewer for the useful comments and suggestions, which helped to improve an earlier version of the manuscript.

References

Balboa-Murias MA (2005). Biomasa arbórea y estabilidad nutricional de los sistemas forestales de *Pinus pinaster* Ait., *Eucalyptus globulus* Labill. y *Quercus robur* L. en Galicia [Aboveground biomass and nutritional stability of *Pinus pinaster* Ait., *Eucalyptus globulus* Labill. and *Quercus robur* L. forest systems in Galicia]. PhD Thesis, Universidad de Santiago de Compostela, Spain, pp. 256. [in Spanish]

Balboa-Murias M, Rodríguez-Soalleiro R, Merino A, Alvarez-González JG (2006). Temporal variations and distribution of carbon stocks in aboveground of radiata pine and maritime pine pure stands under different silvicultural alternatives. Forest Ecology and Management 237: 29-38. - doi: 10.1016/j.foreco.2006.09.024

Baret F, Guyot G (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of Environment 35: 161-173. - doi: 10.1016/0034-4257(91)90009-U

Belsley DA (1991). Conditioning diagnostics, collinearity and weak data regression. Wiley Series in Probability, John Wiley and Sons, New York, USA, pp. 396.

Brañas J, González-Río F, Merino A (2000). Contenido y distribución de nutrientes en plantaciones de *Eucalyptus globulus* del Nordeste de la Península Ibérica [Nutrients content and distribution in *Eucalyptus globulus* plantations in northwestern Iberian peninsula]. Investigación Agraria: Sistemas y Recursos Forestales 9: 316-335. [in Spanish]

Breiman L (2001). Statistical modeling: the two cultures. Statistical Science 16: 199-301. - doi: 10.1214/ss/1009213726

Burnham KP, Anderson DR (1998). Model selection and inference. Springer-Verlag, New York, USA, pp. 515.

Chas-Amil ML (2007). Forest fires in Galicia (Spain): threats and challenges for the future. Journal of Forest Economics 13: 1-5. - doi: 10.1016/j.jfe.2007.02.001

Chen Q, Laurin GV, Battles JJ, Saah D (2012). Integration of airborne LiDAR and vegetation types derived from aerial photography for mapping aboveground live biomass. Remote Sensing of Environment 121: 108-117. - doi: 10.1016/j.rse.2012.01.021

Chen Q (2013). Lidar remote sensing of vegetation biomass. In: "Remote Sensing of Natural Resources" (Weng Q, Wang G eds). CRC Press, Taylor and Francis Group, Boca Raton, FL, USA, pp. 399-420. [online] URL: http://books.google.com/books?id=wIDNBQAAQBAJ

Cortés L, Hernández J, Valencia D, Corvalán P (2014). Estimation of above-ground biomass using Landsat ETM+, Aster GDEM and LiDAR. Forest Research 3: 117.

Estornell J, Ruiz LA, Velázquez B, Hermosilla T (2012). Estimation of biomass and volume of shrub vegetation using LiDAR and spectral data in a Mediterranean environment. Biomass and Bioenergy 46: 710-721. - doi: 10.1016/j.biombioe.2012.06.023

Gómez-Vázquez I, Crecente-Campo F, Diéguez-Aranda U, Castedo-Dorado F (2013). Modelling canopy fuel variables in *Pinus pinaster* Ait. and *Pinus radiata* D. Don stands in Northwestern Spain. Annals of Forest Science 70: 161-172. - doi: 10.1007/s13595-012-0245-9

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

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

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

Huete AR (1988). A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25: 259-309. - doi: 10.1016/0034-4257(88)90106-X

Jenkins JC, Birdsey RA, Pan Y (2001). Biomass and NPP estimation for the mid-Atlantic Region (USA) using plot level forest inventory data. Ecological Applications 11: 1174-1193. - doi: 10.1890/1051-0761(2001)011[1174:BANEFT]2.0.CO;2

Ji L, Wylie BK, Nossov DR, Peterson B, Waldrop MP, McFarland JW, Rover J, Hollingsworth TN (2012). Estimating aboveground biomass in interior Alaska with Landsat data and field measurements. International Journal of Applied Earth Observation and Geoinformation 18: 451-461. - doi: 10.1016/j.jag.2012.03.019

Jiménez E, Vega JA, Fernández-Alonso JM, Vega-Nieva D, González JG, Ruiz-González AD (2013). Allometric equations for estimating canopy fuel load and distribution of pole-size maritime pine trees in five Iberian provenances. Canadian Journal of Forest Research 43: 149-158. - doi: 10.1139/cjfr-2012-0374

Jordan CF (1969). Derivation of leaf area index from quality of light in the forest floor. Ecology 50: 663-666. - doi: 10.2307/1936256

Kaartinen H, Hyyppa J, Yu X, Vastaranta M, Hyyppa H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing 4: 950-974. - doi: 10.3390/rs4040950

Kraus K, Pfeiffer N (1998). Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53: 193-203. - doi: 10.1016/S0924-2716(98)00009-4

Lafiti H, Fassnacht FE, Hartig F, Berger C, Hernández J, Corvalán P, Koch B (2015). Stratified aboveground forest biomass estimation by remote sensing data. International Journal of Applied Earth Observation and Geoinformation 38: 229-241. - doi: 10.1016/j.jag.2015.01.016

Laurin GV, Chen Q, Lindsell JA, Coomes DA, Frate F, Guerriero L, Pirotti F, Valentini R (2014). Above ground biomass estimation in an Africa tropical forest with LiDAR and hyperspectral data. ISPRS Journal of Photogrammetry and Remote Sensing 89: 49-58. - doi: 10.1016/j.isprsjprs.2014.01.001

Lim K, Treitz P, Wulder MA, St-Onge B, Flood M (2003). LiDAR remote sensing of forest structure. Progress in Physical Geography 27: 88-106. - doi: 10.1191/0309133303pp360ra

Lin Y, Jaakkola A, Hyyppä J, Kaartinen H (2010). From TLS to VLS: biomass estimation at individual tree level. Remote Sensing 2: 1864-1879. - doi: 10.3390/rs2081864

Liu HQ, Huete AR (1995). A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Transactions on Geoscience and Remote Sensing 33: 457-465. - doi: 10.1109/36.377946

López-Serrano PM, López-Sánchez CA, Díaz-Varela RA, Corral-Rivas JJ, Solís-Moreno R, Vargas-Larreta B, González JG (2015). Estimating biomass of mixed and unevenaged forests using spectral data and a hybrid model combining regression trees and linear models. iForest - Biogeosciences and Forestry 9 (2): 226-234. - doi: 10.3832/ifor1504-008

Manes F, Ricotta C, Salvatori E, Bajocco S, Blasi C (2010). A multiscale analysis of canopy structure in *Fagus sylvatica* L. and *Quercus cerris* L. old-growth forests in the Cilento and Vallo di Diano National Park. Plant Biosystems 144 (1): 202-210. - doi: 10.1080/11263500903560801

MARM (2011). Cuarto Inventario Forestal Nacional. Comunidad Autónoma de Galicia [Fourth National Forest Inventory. Galicia]. Dirección General del Medio Natural y Política Forestal, Galicia, Madrid, Spain, pp. 52. [in Spanish]

McGaughey R (2009). FUSION/LDV: software for LIDAR data analysis and visualization. USDA Forest Service, Pacific Northwest Research Station, Seattle, WA, USA, pp. 123.

McRoberts RE, Naesset E, Gobakken T (2013). Inference for lidar-assisted estimation of forest growing stock volume. Remote Sensing of Environment 128: 268-275. - doi: 10.1016/j.rse.2012.10.007

Merino A, Balboa MA, Rodríguez-Soalleiro R, Alvarez-González JG (2005). Nutrient exports under different harvesting regimes in fast-growing forest plantations in southern Europe. Forest Ecology and Management 207: 325-339. - doi: 10.1016/j.foreco.2004.10.074

Montero G, Ruiz-Peinado R, Muñoz M (2005). Producción de biomasa y fijación de CO_2 por los bosques españoles [Spanish forest biomass production and CO_2 fixing]. Monografías INIA, Serie Forestal no. 13, Madrid, Spain, pp. 265. [in Spanish]

Naesset E, Gobakken T, Solberg S, Gregoire TG, Nelson R, Stahl G, Weydahl D (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: a case study from a boreal forest area. Remote Sensing of Environment 115 (12): 3599-3614. - doi: 10.1016/j.rse.2011.08.021

NASA (2011). Landsat 7 science data users handbook. Landsat Project Science Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA, pp. 186.

Ni-Meister W, Lee S, Strahler AH, Woodcock AH, Schaaf C, Ranson J, Sun G, Blair JB (2010). Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from vegetation lidar. Journal of Geophysical Research 115 (G2): 2156-2202. - doi: 10.1029/2009JG000936

Nord-Larsen T, Schumacher J (2012). Estimation of forest resources from a country wide laser scanning survey and national inventory data. Remote Sensing of Environment 119: 148-157. - doi: 10.1016/j.rse.2011.12.022

Pérez-Cruzado C, Rodríguez-Soalleiro R (2011). Improvement in accuracy of aboveground biomass estimation in *Eucalyptus nitens* plantations: effect of bole sampling intensity and explanatory variables. Forest Ecology and Management 261: 2016-2028. - doi: 10.1016/j.foreco.2011.02.028

Rouse J, Haas R, Schell J, Deering D (1973). Monitoring vegetation system in the great plains with ERTS. In: Proceedings of the "3rd ERTS Symposium". NASA SP-351, NASA, Washington, DC, USA, pp. 309-317.

Scaramuzza P, Micijevic E, Chander G (2004). SLC gap-filled products, phase one methodology. Landsat Technical Notes, pp. 5. [online] URL: http://landsat.usgs.gov/sites/default/files/documents/SLC_Gap_Fill_Methodology.pdf

Shendryk I, Hellström M, Klemedtsson L, Kljun N (2014). Low-density LiDAR and optical imagery for biomass estimation over boreal forest in Sweden. Forests 5: 992-1010.

Sprugel DG (1983). Correcting for bias in log-transformed allometric equations. Ecology 64: 209-210. - doi: 10.2307/1937343

Zhao K, Popescu S, Nelson R (2009). LiDAR remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment 113: 182-196. - doi: 10.1016/j.rse.2008.09.009