Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/forestsystems
http://dx.doi.org/10.5424/fs/2013223-03878

Forest Systems 2013 22(3): 510-525
ISSN: 2171-5068
eISSN: 2171-9845

Modelling stand biomass fractions in Galician *Eucalyptus globulus* plantations by use of different LiDAR pulse densities

E. Gonzalez-Ferreiro[1]*, D. Miranda[1], L. Barreiro-Fernandez[1], S. Bujan[1], J. Garcia-Gutierrez[2] and U. Dieguez-Aranda[3]

[1] *Laboratorio del Territorio. Departamento de Ingeniería Agroforestal. Escuela Politécnica Superior. Universidad de Santiago de Compostela. C/ Benigno Ledo. Campus Universitario, s/n. 27002 Lugo, Spain*
[2] *Departamento de Ciencias de la Computación, Lenguajes y Sistemas. Escuela Técnica Superior de Ingeniería Informática. Universidad de Sevilla. C/ Reina Mercedes, s/n. 41012 Sevilla, Spain*
[3] *Unidad de Gestión Forestal Sostenible. Departamento de Ingeniería Agroforestal. Escuela Politécnica Superior. Universidad de Santiago de Compostela. C/ Benigno Ledo. Campus Universitario, s/n. 27002 Lugo, Spain*

Abstract

Aims of study: To evaluate the potential use of canopy height and intensity distributions, determined by airborne LiDAR, for the estimation of crown, stem and aboveground biomass fractions.

To assess the effects of a reduction in LiDAR pulse densities on model precision.

Area of study: The study area is located in Galicia, NW Spain. The forests are representative of *Eucalyptus globulus* stands in NW Spain, characterized by low-intensity silvicultural treatments and by the presence of tall shrub.

Material and methods: Linear, multiplicative power and exponential models were used to establish empirical relationships between field measurements and LiDAR metrics.

A random selection of LiDAR returns and a comparison of the prediction errors by LiDAR pulse density factor were performed to study a possible loss of fit in these models.

Main results: Models showed similar goodness-of-fit statistics to those reported in the international literature. R^2 ranged from 0.52 to 0.75 for stand crown biomass, from 0.64 to 0.87 for stand stem biomass, and from 0.63 to 0.86 for stand aboveground biomass. The RMSE/MEAN · 100 of the set of fitted models ranged from 17.4% to 28.4%.

Models precision was essentially maintained when 87.5% of the original point cloud was reduced, *i.e.* a reduction from 4 pulses m^{-2} to 0.5 pulses m^{-2}.

Research highlights: Considering the results of this study, the low-density LiDAR data that are released by the Spanish National Geographic Institute will be an excellent source of information for reducing the cost of forest inventories.

Key words: Eucalypt plantations; airborne laser scanning; aboveground biomass; carbon stocks; remote sensing; forest inventory.

Introduction

Aboveground biomass is the total amount of biological material (usually oven-dried to remove water) present above the soil surface in a specified area. Because almost 50% of plant biomass is carbon, estimates of the total aboveground biomass in forest ecosystems are critical for carbon dynamics studies at multiple scales (Drake *et al.*, 2003). Estimation of the biomass and carbon stock of trees has gained importance in recent years, especially since the Kyoto Protocol of the United Nations Framework Convention on Climate Change (UNFCCC) entered into force on 16 February 2005. Signatory countries must estimate carbon stocks in 1990 and any changes since 1990 from all afforestation, reforestation and deforestation activities (UNFCCC, 1997). In addition, the continued growth in energy demands in technologically developed societies and the requirement to reduce the substantial use of fossil fuels have made it necessary to diversify means of energy production. In this sense, the forest biomass in established plantations is attracting great interest as a renewable resource for biomass energy production, because of the following advantages: reduction in net warming and environmental

* Corresponding author: edu.g.ferreiro@gmail.com
Received: 19-12-12. Accepted: 18-07-13.

pollution, decreased risk of sacrificing natural areas, enhanced regional welfare, and avoidance of competition with food production (Freppaz *et al.*, 2004; Field *et al.*, 2008). Therefore, the measurement of forest biomass provides an indication of carbon sequestration in trees and also an estimate of cellulosic material as a potential source of renewable energy (Popescu, 2007).

Eucalyptus spp. are the most productive tree species in Galicia (NW Spain). Eucalypt plantations, which occupy 215,899 ha in the region, accumulated approximately 52% of the harvested volume in the period 2005-2010 (3,574,500 m^3 harvested in 2010, with an increment of 15% in the period 2009-2010) (Confemadera, 2010). The most widespread management regime for these plantations is short rotation forestry (SRF), with initial densities of 1,000 to 2,400 stems ha^{-1} in single stem stands at first rotation and subsequent replanting or coppicing (Pérez-Cruzado *et al.*, 2011). The plantations are generally managed, without thinning, for wood pulp production (and to a lesser extent, chip-board), and they provide small wood that is not used by the sawnwood industry. Bioenergy production is one potential use for these plantations (Pérez-Cruzado *et al.*, 2011).

Despite the high productivity and potential of eucalyptus forest as carbon sinks, there are important deficiencies in the characterization and quantification of the biomass resource available in Galicia. This has arisen because forest information is usually acquired through conventional forest inventory methods, which often use detailed measurements taken from a small set of sampling plots distributed randomly or systematically over the forest area. The quality of the estimates is thus limited by the cost of establishing sufficient sample plots to measure the existing variability (Lovell *et al.*, 2005; Musk and Osborn, 2007; Rombouts *et al.*, 2008). Such inventories are extremely labour intensive and expensive (Hall *et al.*, 2005), especially in small areas of privately owned land, and in practice they do not allow full inventory coverage of large areas on the ground or can extend over several years (Boudreau *et al.*, 2008).

Airborne Light Detection and Ranging (LiDAR) is currently considered the most promising remote sensing technology for forest inventory and biomass estimation (Boudreau *et al.*, 2008). It is capable of offering detailed tri-dimensional information regarding the size and structure of the forest canopy (Reitberger *et al.*, 2008; Wagner *et al.*, 2008), which is closely related to aboveground biomass (Lim *et al.*, 2003b). This has potential for timely and accurate measurement of tree biomass components and carbon stored over time (Naesset and Gobakken, 2008; Rosenqvist *et al.*, 2003) for the whole area, with an equal or better accuracy than other remote sensing techniques (Bortolot and Wynne, 2005).

Two main approaches are used for biomass estimation with commercial small-footprint discrete return laser scanning data. The "single-tree" approach used with spatially dense LiDAR data is based on laser detection and on the measurement of individual tree parameters, usually tree height and crown diameter (Bortolot and Wynne, 2005; Popescu *et al.*, 2003; Popescu, 2007). The "stand-level" approach used with low resolution LiDAR data establishes empirical relationships between plot-level stand measurements and the height distribution obtained from laser returns or other laser measurements related to canopy density. This approach has been used in regression analysis to estimate forest biomass and carbon stock levels across a range of forest types (Lim *et al.*, 2003a; Lim and Treitz, 2004a, 2004b; Patenaude *et al.*, 2004; Andersen *et al.*, 2005; Hall *et al.*, 2005; Thomas *et al.*, 2006; Naesset and Gobakken, 2008; Sherrill *et al.*, 2008; García *et al.*, 2010; Treitz *et al.*, 2010; González-Ferreiro *et al.*, 2012) as well as shrub biomass (Estornell *et al.*, 2011).

Metrics derived from LiDAR data are highly dependent on the tree species involved (Heurich and Thoma, 2008) and are also correlated with field assessment of various aspects of vegetation structure, which may influence the derived relationships (Goodwin *et al.*, 2006). Small footprint discrete return LiDAR has been used in eucalyptus forests to estimate a variety of stand variables (Wack *et al.*, 2003; Tesfamichael *et al.*, 2010; Gonçalves-Seco *et al.*, 2011), and to explore the effects of airborne LiDAR acquisition parameters on vegetation structural assessment (Goodwin *et al.*, 2006). However, there are no reports of biomass estimation of Atlantic eucalyptus plantations of medium to high density, characterized by very small crowns with sparse leaves that may facilitate laser penetration.

Laser pulse density and vegetation structure are the factors with the greatest effects on the height accuracy of the laser-derived Digital Terrain Model (DTM). If the ground elevation or the uppermost portion of the forest canopy is not well detected, the normalised heights of the trees and the Digital Crown Model (DCM) obtained will be underestimated (Hyyppä *et al.*, 2008). Goodwin *et al.* (2006) argued that point density is even more important than footprint size or flight altitude in determining certain forest variables such as crown area and volume. Magnusson (2006) observed a significant increase in the root mean square error for mean height and stand volume estimates when the laser pulse density was greatly decreased, while other authors (Maltamo *et al.*, 2006; Thomas *et al.*, 2006; Gobakken and Næsset, 2007; García *et al.*, 2010; Treitz *et al.*, 2010; González-Ferreiro *et al.*, 2012) argued that the estimation of most forest stand variables may not be affected. This should be verified in different forests and regions to determine whether LiDAR data can be collected at a lower average point spacing for a more cost-effective ground coverage, while maintaining the accuracy of the estimates (Thomas *et al.*, 2006).

The aim of this study was to estimate biomass fractions in *Eucalyptus globulus* Labill. plantations from height and intensity data gathered with a small footprint discrete return LiDAR system. The effect of a reduction in pulse density on model precision was examined, in light of the availability of low-density LiDAR data (0.5 pulses m^{-2}) for most of the Spanish territory.

Materials and methods

Study area

The study area is a 1 × 4 km rectangle located in the municipality of Vilapena (Galicia, NW Spain) (see Fig. 1). This area is characterized by a wide variety of landform types and an elevation range of 150 to 530 m. The forests in this area are representative of *Eucalyptus globulus* stands in NW Spain, characterized by low-intensity silvicultural treatments and by the presence of tall shrub. The canopy of mature *Eucalyptus globulus* plantations is usually high and sparse, thus enabling sunlight to penetrate to the ground, which is often aggravated by severe defoliation caused by the eucalyptus weevil *Gonipterus scutellatus* Gyll. Such low resistance of the canopy to light penetration allows growth of herbaceous sub-shrub cover dominated by shade-intolerant fruticose species and both shade-intolerant and shade-tolerant herbs below the canopy, which often produce a large amount of biomass (about 2.5 to 3.2 Mg ha^{-1} total dry matter) (Silva-Pando *et al.*, 1993). The already well-known loss of precision in DTMs in forest areas, together with steep slopes and the high variability and quantity of understory, poses a serious challenge to the accuracy of LiDAR technology (Raber *et al.*, 2002; Hodgson *et al.*, 2003, 2005).

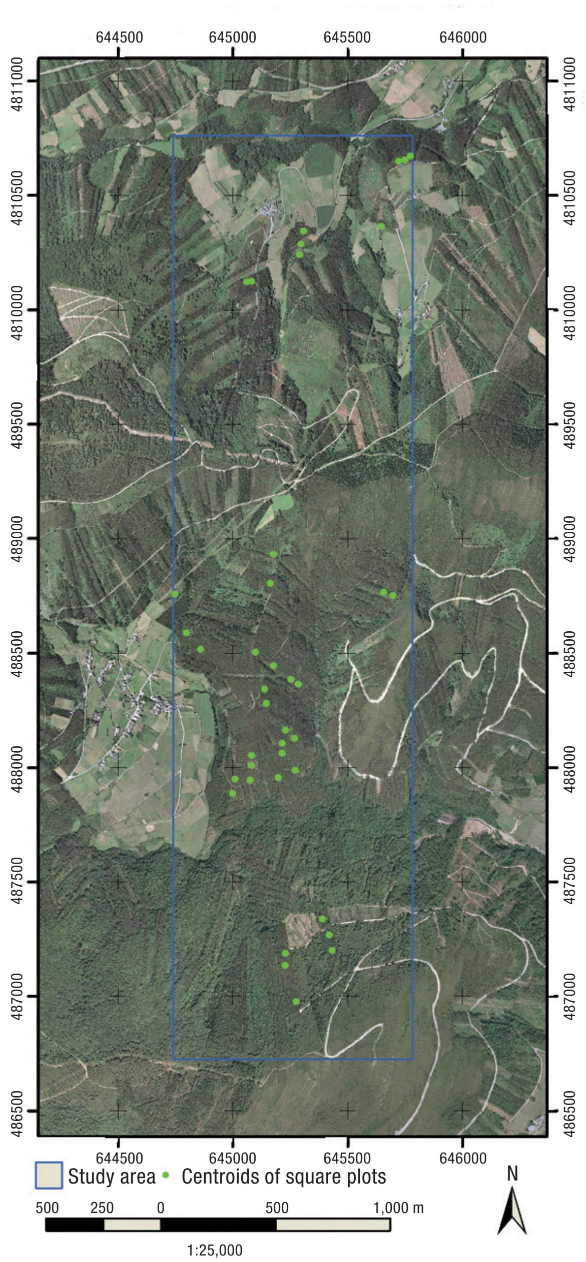

Figure 1. Inventory plots.

LiDAR data

The LiDAR data were acquired in November 2004 with an Optech Airborne Laser Terrain Mapper (ALTM) 2033 sensor (www.optech.ca) operated at a laser wavelength of 1,064 nm from a flight altitude of 1,500 m above sea level. The beam divergence was 0.3 mrad, the pulsing frequency 33 kHz, the scan frequency 50 Hz, and the maximum scan angle ±10∫. The first and last return pulses were registered. The whole study area was flown in 18 strips and each strip was flown three

Table 1. Forestry stand parameters of the sample plots (n = 39)

Variable	Average	Minimum	Maximum	Standard deviation
N	1,663	622	3,378	770
d	13.5	9.2	19.3	2.7
H_m	17.4	11.8	22.3	2.4
H_d	23.5	13.3	36.5	4.1
G	25.8	7.2	47.0	9.6
V	229.4	39.9	511.1	105.1
W_{cr}	17.4	4.5	35.6	7.0
W_{st}	114.6	19.9	276.9	53.9
Wabg	132.0	24.4	312.4	60.7

N: number of stems per hectare. d: diameter at breast height outside bark (1.3 m above ground, cm). H_m: mean height (m). H_d: dominant height (m). G: stand basal area (m^2 ha^{-1}). V: stand volume over bark (m^3 ha^{-1}). W_{cr}: stand crown biomass (Mg ha^{-1}). W_{st}: stand stem biomass (Mg ha^{-1}). W_{abg}: stand above ground biomass (Mg ha^{-1}).

times, which gave an average measurement density of about 4 pulses m^{-2}.

Field data

A total of 39 square plots of 225 m^2 were located and measured in the *Eucalyptus globulus* plantations in the study area, between February and March 2005. The plots were subjectively selected to represent the existing range of ages, stand densities and sites in the regions. Topographical surveys were carried out using total stations and GPS to determine the location of the four corners and the position of every tree within the plots. First, a Trimble® 5800 GPS (Trimble, Sunnyvale, CA, USA, www.trimble.com) (dual-frequency real-time kinematic receiver with a planimmetric precision of ±5 mm + 0.5 ppm and a altimetric precision of ±5 mm + 1 ppm) was used to obtain the coordinates of a densified geodetic network for the study area by applying real-time kinematics. Based on the network established with the GPS, a topographic survey of the plots was conducted using a Trimble® 5603 Robotic Total Station (Trimble, Sunnyvale, CA, USA, www.trimble.com) with a precision in distances measurement of ±2 mm + 2 ppm and a precision in angles measurement of 3 to 5”. Observations of the four corners of each plot and of the position of each tree within the plot were made during the survey.

For all the trees in each sample plot, two measurements of diameter at breast height (1.3 m above ground level) were made at right angles, with a tree calliper. Measurements were made to the nearest millimetre, and the arithmetic mean of the two measurements was calculated. Total tree height was measured to the nearest decimetre with a Vertex III hypsometer (Haglöf Sweden AB, Långsele, Sweden, www.haglof.se).

The dry weight of the biomass fractions of each tree were estimated from the following equations for *Eucalyptus globulus* in Galicia reported by Diéguez-Aranda *et al.* (2009):

$$w_w + w_{b7} = 0.01308\ d^{1.870} h^{1.172} \quad [1]$$

$$w_b = 0.01010\ d^{2.484} \quad [2]$$

$$w_{b2-7} + w_{b0.5-2} = 0.003685\ d^{2.654} \quad [3]$$

$$w_{b0.5} = 0.01258\ d^{1.705} \quad [4]$$

$$w_l = 0.02949\ d^{1.917} \quad [5]$$

where w_w is stem wood biomass (kg), w_{b7} is wood and bark biomass on branches with 7 cm minimum top diameter (kg), w_b is bark biomass on stem (kg), w_{b2-7} is wood and bark biomass on branches with 7 cm maximum butt diameter and 2 cm minimum top diameter (kg), $w_{b0.5-2}$ is wood and bark biomass on branches with 2 cm maximum butt diameter and 0.5 cm minimum top diameter (kg), $w_{b0.5}$ is wood and bark biomass on branches with 0.5 cm maximum butt diameter (kg), w_l is needles biomass (kg), d is diameter at breast height outside bark (1.3 m above the ground level, cm), and h is total tree height (m).

Finally, crown biomass (w_{cr}), stem biomass (w_{st}) and aboveground biomass (w_{abg}) were calculated from the sum of the biomass fractions included:

$$w_{cr} = w_{b2-7} + w_{b0.5-2} + w_{b0.5} + w_l \quad [6]$$

$$w_{st} = w_w + w_{b7} + w_b \quad [7]$$

$$w_{st} = w_w + w_{b7} + w_b \quad [8]$$

The field measurements (heights and diameters) and the estimated dry weight of the biomass fractions were used to estimate the following stand variables in each plot, on a per hectare basis: stand crown biomass (W_{cr}), stand stem biomass (W_{st}), and stand aboveground biomass (W_{abg}). The estimates were used to develop models to derive these stand variables from LiDAR data.

Preparation of LiDAR data

The LiDAR data provided by the contractor only provided information, for each laser pulse emitted by the sensor, about return type (first and last), X, Y and Z coordinates, and intensity values. Single returns, *i.e.* those that hit a solid surface, were recorded twice in the original dataset with the same information but different return type code. The theoretical nominal density of the LiDAR data was 4 pulses m^{-2}, which implied collecting 8 returns m^{-2}.

Original LiDAR data with different resolutions, i.e. obtained from flights carried out in the same area, at the same time, and with the same flight parameters but at different LiDAR pulse densities, are the most desirable for assessing the influence of density on the extraction of forest information. However, the lack of such data can be overcome by artificially reducing the original LiDAR point cloud. Although this approach does not fully mimic the actual reduction that would be obtained with different flights, it allows study of the most important factor in canopy height modelling. For this purpose, several alternatives have been proposed, as follows: (i) a random reduction over the entire dataset (Anderson *et al.*, 2006; Liu and Zhang, 2008; Puetz *et al.*, 2009), (ii) a systematic reduction in each scan line (Gueudet, 2004; Raber *et al.*, 2007; Magnusson *et al.*, 2007; Treitz *et al.*, 2010), and (iii) a reduction by randomly maintaining one return within a grid cell of a specific size (Magnusson, 2006; Gobakken and Næsset, 2007). The first alternative may not preserve the order and regularity of the original LiDAR data. The second requires information about the scan line and the order in which points were recorded. The third option overcomes the possible lack of regularity and allows for the use of data without scan line information.

Because of the limited information provided by the data used in this study, random selection of LiDAR returns in a grid cell of 1 m^2 was carried out (note that this procedure does not allow for the simulation of other potentially important parameters such as flight height or scan angle). Examination of the original point cloud showed that many cells included more than 8 returns m^{-2}. Therefore, to obtain a regular distribution of LiDAR returns, and to investigate the effect of the LiDAR point cloud density on the estimation of stand variables, two datasets were generated: one with a final density of 1 return m^{-2} and other with 8 returns m^{-2} (equivalent to 0.5 and 4 pulses m^{-2}). Fig. 2a and 2b show the two LiDAR datasets generated for the plot number 3.

Intensity is a radiometric constituent of LiDAR data (Singh *et al.* 2010), and it is recorded by the sensor as the amount of energy backscattered from objects or earth's surface. The intensity values recorded by the sensor remain unchanged under different conditions of illumination caused by, *e.g.*, shadows or occlusions (Donoghue *et al.,* 2007; Höfle and Pfeifer, 2007), but they are affected by other factors, such as properties of the terrain, topography, flight and sensor characteristics, and atmospheric conditions (Donoghue *et al.,* 2007; Höfle and Pfeifer, 2007; Mazzarini *et al.,* 2007).

Some authors have recommended considering range, incidence angle and atmospheric attenuation for intensity normalization (Höfle and Pfeifer, 2007; Gross *et al.,* 2008; Jutzi and Gross, 2010). However, such data are not always available, and other authors recommend normalizing intensity values to a user-defined standard range in order to remove the range

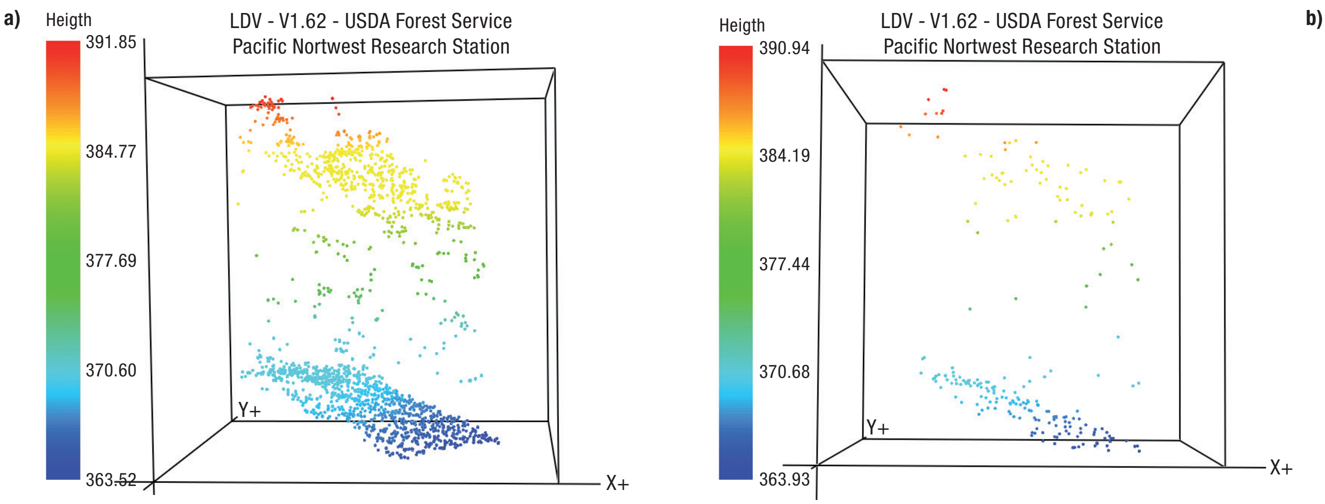

Figure 2. LiDAR cloud for the plot number 3 with (a) 4 pulses m^{-2} and (b) 0.5 pulses m^{-2}. Scale bars show ellipsoidal height of the LiDAR returns for the plot number 3.

dependency of the intensity signal (Donoghue *et al.*, 2007; Mazzarini *et al.,* 2007; García *et al.,* 2010). In this approach, the normalized intensity values (I') is obtained by multiplying the raw intensity value (I) by the quotient of the range of each return (rg), calculated as the difference between the average flying height and the height of each return, *i.e.* the difference between the sensor-target distance, and the standard range (rg_s) (*e.g.* 1,000 m) (Eq. [9]):

$$I' = I\frac{rg^2}{rg_s^2} \qquad [9]$$

This method eliminates the path length variations in the intensity recorded by the system, by providing values equivalent to the intensity that would have been recorded if all points were at the same range (García *et al.*, 2010).

In the present case, range and scan angle data were not available for every return. The terrain was very steep, with slopes above 40° in some areas and a topographic range of 150 to 530 m above mean sea level, and therefore the range was normalized to a user-defined standard range using Eq. [9]. For each return, the range (m) was estimated as the difference between the average altitude of the flight (1,500 m above sea level) and the ellipsoidal altitude of the return (m). This approach should not cause large errors in range computation because of the small scan angle (±10°) (García *et al.*, 2010). Fig. 3a and 3b show the 0.5 pulses m^{-2} LiDAR dataset for the plot number 3, before and after the intensity normalization process.

Extraction of LiDAR variables

For the generated datasets (0.5 and 4 pulses m^{-2}), FUSION software (McGaughey, 2009) was used to perform filtering, interpolation and DTM/DCM generation operations, as well as to compute the following variables related to the height and return intensity distributions metrics within the limits of the 39 field plots: mean, maximum and minimum values, mode, standard deviation, variance, interquartile distance, coefficients of skewness and kurtosis, average absolute deviation, and percentiles. The percentage of returns above a specific height threshold was also calculated.

The following steps were carried out with several processing programmes implemented in the FUSION LIDAR Toolkit (McGaughey, 2009). First, ground returns were extracted from the LiDAR point cloud with the GroundFilter tool, which implements a filtering algorithm adapted from Kraus and Pfeifer (1998) and based on linear prediction (Kraus and Mikhail, 1972). Second, these returns were used to generate a DTM grid with the GridSurfaceCreate tool, which computes the elevation of each grid cell from the average elevation of all points within the cell; if the cell does not contain any points, it is filled by interpolation from the neighbouring cells; the cell size value was 1 m^2. Third, the normalized LiDAR point cloud was obtained by subtraction of the ellipsoidal height of the DTM from the Z coordinate of each LiDAR return with the ClipData tool (the *height* switch in combination with the previously generated DTM were used); this

a)

b)

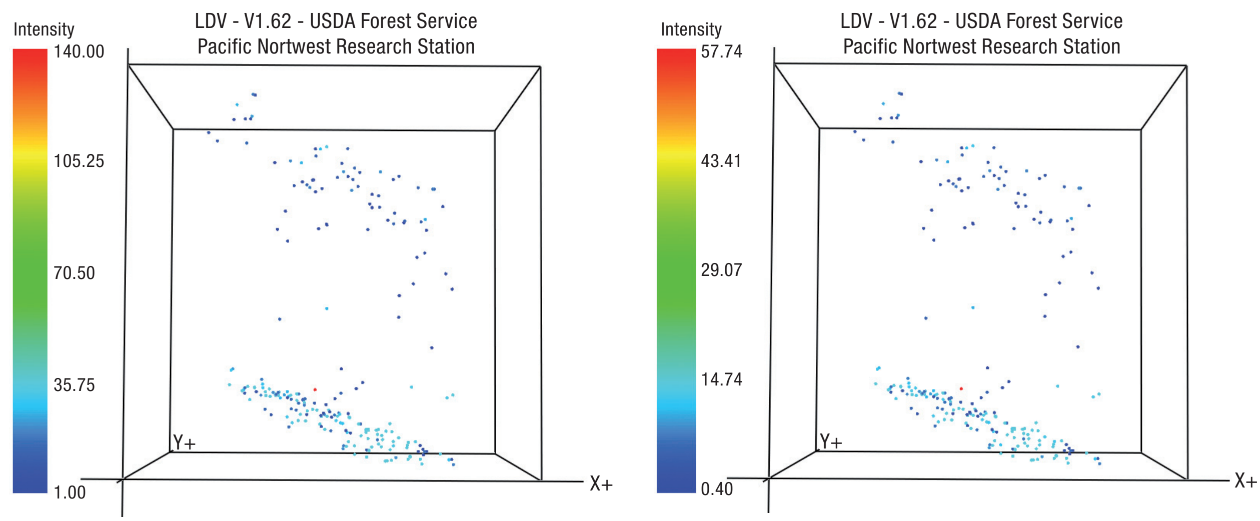

Figure 3. 05 pulses m^{-2} LiDAR dataset for the plot number 3. Scale bars show the intensity of the LiDAR returns: (a) before normalization and (b) after normalization.

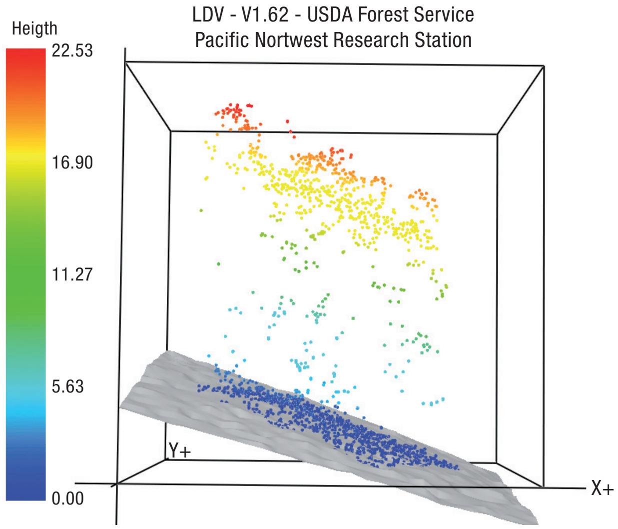

Figure 4. 4 pulses m^{-2} LiDAR dataset for the plot number 3 and the generated DTM. Scale bar shows the normalized height of the LiDAR returns.

tool was used also to exclude returns below a normalised height of 2 m (*zmin* switch = 2), which were considered as not belonging to tree crowns (*e.g.* hits on shrubs, rocks and logs). Fourth, the normalised LiDAR point cloud was clipped using the boundaries of the 39 field plots, which were previously stored as polygons in ESRI™ shape files (ESRI, 1998). The PolyClipData tool allowed extracting the LiDAR cloud within the limit of each plot area, so that, an independent LiDAR cloud file was created, covering each plot area. Fifth, the metrics of heights and return intensity distributions of these 39 clipped and normalized point clouds were computed with the CloudMetrics tool. Fig. 4 shows the 4 pulses m^{-2} LiDAR dataset for the plot number 3 with the normalized height of the LiDAR returns and the generated DTM.

Regression models

Linear, (multiplicative) power function and exponential models were used to establish empirical relationships between field measurements and LiDAR variables. Their respective general expressions are as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon \qquad [10]$$

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \ldots X_n^{\beta_n} + \varepsilon \qquad [11]$$

$$Y = \exp\left(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n\right) + \varepsilon \qquad [12]$$

where Y are field values of W_{cr} (kg ha^{-1}), W_{st} (kg ha^{-1}), and W_{abg} (kg ha^{-1}), and X_1, X_2..., X_n may be variables related to the metrics of heights and return intensity distributions or measurements related to canopy closure. All variables were computed from the datasets with resolutions of 4 and 0.5 pulses m^{-2}. The following variables can be related to height distribution: h_{min}, h_{max}, h_{mean}, h_{median}, h_{mode}, h_{SD}, h_{skw}, h_{kurt}, h_{ID}, h_{AAD}, h_{05}, h_{10}, h_{20},..., h_{90}, h_{95}, h_{25} or h_{75}, which are the minimum, maximum, mean, median, mode, standard deviation, coefficients of skewness and kurtosis, interquartile distance, average absolute deviation, percentiles, and first and third quartiles values of the height distribution of laser returns for each plot (m), respectively. The variables related to return intensity distribution are based on the same statistics as the variables related to height distribution, but in this case they are denoted with the letter i instead of h. The variables related to canopy closure can be either r_2, which is the number of returns above 2 m height for each plot, or $c_{2\text{-}FP}$, which is the ratio of the number of laser hits above 2 m height to the number of first returns for each plot, expressed as a percentage. The additive error term ε is assumed to be normally, independent and identically distributed with zero mean.

Model fitting and selection

Linear models were fit by ordinary least squares, by applying the REG procedure of SAS/STAT® (SAS Institute Inc., 2004). Power function and exponential models were fit by nonlinear regression, by use of the Gauss-Newton method implemented in the NLIN procedure of the same statistical package. In a previous step, the latter two types of models were linearized by taking natural logarithms from both sides of Eqs. [11] and [12], in order to select the best subset of independent variables to be included in each and to obtain initial estimates of the parameters using the linear regression technique (Myers, 1990, p: 444).

Once the complete linear form of the models was specified, the models were examined to determine whether all terms should be retained in the final regression equations. This involved fitting a number of subset models and comparing the relative performance of each (Clutter *et al.*, 1983, p: 318). Although there are different approaches for selecting the subset models to be fit by linear regression (Draper and Smith, 1998, chapter 15), the Mallows' C_p selection method of the

REG procedure, which performs all possible subset regressions and lists the models in ascending order of C_p, was used. Heteroscedasticity was examined visually, by plotting residuals as a function of predicted values, but any discernible trend or heteroscedasticity was found. Multicollinearity among the explanatory variables was also checked with the condition index. Regressions with a condition index above 30 were disregarded, as recommended by Belsley (1991). Finally, only models in which all the parameter estimates were significant at 5% level were considered.

Comparison of the estimates for the different models (linear, power function and exponential) was based on numerical and graphical analyses of the residuals. The following three statistics were calculated: the coefficient of determination (R^2), the root mean square error (RMSE), and the Bayesian Information Criterion (BIC) proposed by Schwarz (1978). Although BIC was used as the final criterion for model selection (Peña, 2002, p: 570), it does not provide an intuitive idea of model precision. R^2 (also referred to as pseudo-R^2 when applied in nonlinear regression) indicates the proportion of the total variance of the dependent variable that is explained by the model. Although there are several shortcomings associated with use of R^2 in nonlinear regression, the general usefulness of some global measure of model adequacy appear to override some of those limitations (Ryan, 1997, p: 424). The RMSE provides an idea of the precision of the estimates in the same units as the dependent variable.

Effects of density reduction

Using a similar procedure to that proposed by Treitz *et al.* (2010), but considering only laser pulse density as the factor of interest, prediction errors from the best linear, power function, and exponential models developed for each dependent variable were calculated for each plot:

$$e_{ij} = Y_i - \hat{Y}_{ij} \qquad [13]$$

where e_{ij} is the prediction error of the ith plot (from 1 to 39) associated with the jth LiDAR pulse density (0.5 pulses m^{-2} and 4 pulses m^{-2}), Y_i is the stand variable obtained from field measurements for the ith plot, and $\hat{y}_{ij}$ is the corresponding predicted value for the ith plot and the jth LiDAR pulse density.

The nonparametric Kruskal-Wallis one-way analysis of variance by ranks (Kruskal, 1952; Kruskal and Wallis, 1952) was used to compare the prediction errors distributions by LiDAR pulse density factor because of the non-normal distribution of the prediction errors for each model type and dependent variable, as suggested by the Shapiro-Wilk test for normality (Shapiro and Wilk, 1965, 1968) and the examination of Quantile-Quantile plots. If the computed value of the test suggests rejecting the null hypothesis, there is a high likelihood that the two samples represent populations with different median values (Sheskin, 2004, pp: 757-761).

Results

Regression models

The parameter estimates and goodness-of-fit statistics of the best model developed by type (linear, power function and exponential), dependent stand variable (W_{cr}, W_{st}, and W_{abg}) and generated dataset (0.5 and 4 pulses m^{-2}) are shown in Tables 2-4. For all the dependent variables and datasets, exponential models performed best (as shown by the BIC values), followed by linear and power function models. In W_{cr} modelling, exponential models provided R^2 values of 75.3% and 71.8% for the 4 and 0.5 pulses m^{-2} datasets, respectively (Table 2); the respective differences between the best and worst (power function) models, in terms of R^2, were 11.7% and 19.6%. In W_{st} modelling, exponential models provided R^2 values of 86.6% and 84.1% for the 4 and 0.5 pulses m^{-2} datasets, respectively (Table 3); the respective differences between the best and worst (power function) models, in terms of R^2, were 9.2% and 19.7%. In W_{abg} modelling, exponential models provided R^2 values of 85.7% and 83.1% for the 4 and 0.5 pulses m^{-2} datasets, respectively (Table 4); the respective differences between the best and worst (power function) models, in terms of R^2, were 9.7% and 19.9. Only independent variables related to the metrics of height distribution proved to be reliable statistics for predicting the three dependent variables W_{cr}, W_{st}, and W_{abg} (Supplementary data includes plot information of the values of dependent and explanatory variables used in the final models presented in the Tables 2, 3 and 4). Further analyses were only carried out for exponential models.

Figs. 5a, 5b, 6a, 6b, 7a and 7b show the field-measured *versus* predicted values (using exponential models) of crown, stem and aboveground biomass fractions for the 0.5 and 4 pulses m^{-2} datasets, respectively.

Table 2. Results of W_{cr} modelling for 0.5 and 4 pulses m^{-2} datasets

Pulses m^{-2}	Model	Independent variable	Parameter estimate	Standard error[1]	*t*-value	$p > \|t\|$	R^2	RMSE (kg ha^{-1})	BIC
0.5	Linear	β_0 parameter	–20,897	4,086	–5.11	< 0.0001	0.708	3,812	650.5
		h_{90}	1,916	202.3	9.47	< 0.0001			
	Power function	β_0 parameter	108.2	101.4	1.07	0.2929	0.522	4,879	669.8
		h_{60}	1.770	0.319	5.55	< 0.0001			
	Exponential	β_0 parameter	7.544	0.264	28.56	< 0.0001	0.718	3,745	649.1
		h_{75}	0.1164	0.0129	9.02	< 0.0001			
4	Lineal	β_0 parameter	–17,319	3,610	–4.80	< 0.0001	0.719	3,736	648.9
		h_{75}	1,835	188.4	9.74	< 0.0001			
	Power function	β_0 parameter	51.81	43.76	1.18	0.2421	0.636	4,256	659.1
		h_{60}	2.009	0.285	7.05	< 0.0001			
	Exponential	β_0 parameter	7.576	0.231	32.78	< 0.0001	0.753	3,504	643.9
		h75	0.1126	0.011	10.24	< 0.0001			

[1] For the power function and exponential models, the standard errors are approximates values.

Effects of density reduction

For all the pairwise comparisons, the computed chi-square approximation of the Kruskal-Wallis test statistic (H) was below the tabled critical 0.05 chi-square value ($\chi^1_{05} = 3.84$) for 2-1 degrees of freedom (*df*). The alternative hypothesis (no equality of medians) was not supported at the 0.05 level in any of the cases ($H < \chi^1_{05}$, $\alpha = 0.05$ and $p > 0.05$; see Table 5). Therefore, the results suggest that the reduction in the LiDAR point cloud has no effect on exponential model fit.

Comparison in terms of R^2 can provide a more intuitive idea of the non loss of fit caused by reduction of LiDAR density. Variations in R^2 of 3.5, 2.5 and 2.6% were obtained for W_{cr}, W_{st}, and W_{abg} exponential models, respectively. The selected exponential models for W_{cr}, W_{st}, and W_{abg} include the same independent variable (h_{75}) for the two datasets (4 and 0.5 pulses m^{-2}) (Tables 2-4).

Discussion

A statistical approach based on regressors, which were calculated directly from the previously normalized laser-derived canopy height and intensity distribu-

Table 3. Results of W_{st} modelling for 0.5 and 4 pulses m^{-2} datasets

Pulses m^{-2}	Model	Independent variable	Parameter estimate	Standard error[1]	*t*-value	$p > \|t\|$	R^2	RMSE (kg ha^{-1})	BIC
0.5	Linear	β_0 parameter	–16,3950	23,804	5,424	< 0.0001	0.801	24,732	800.0
		h_{75}	15,802	1,319	11.98	< 0.0001			
		h_{skw}	1,2920	5,425	2.38	0.0226			
	Power function	β_0 parameter	103.8	107.0	0.97	0.3383	0.644	32,594	817.9
		h_{60}	2.434	0.349	6.97	< 0.0001			
	Exponential	β_0 parameter	8.785	0.240	36.60	< 0.0001	0.841	21,773	786.4
		h_{75}	0.1493	0.0114	13.10	< 0.0001			
4	Linear	β_0 parameter	–169,209	23,446	–7.22	< 0.0001	0.803	24,261	794.9
		h_{75}	15,017	1,223	12.28	< 0.0001			
	Power function	β_0 parameter	39.57	33.29	1.19	0.2422	0.774	25,991	800.2
		h_{60}	2.745	0.282	9.73	< 0.0001			
	Exponential	β_0 parameter	8.902	0.200	44.51	< 0.0001	0.866	19,985	779.7
		h_{75}	0.1406	0.0093	15.12	< 0.0001			

[1] For the power function and exponential models, the standard errors are approximates values.

Table 4. Results of W_{abg} modelling for 0.5 and 4 pulses m^{-2} datasets

Pulses m^{-2}	Model	Independent variable	Parameter estimate	Standard error[1]	t-value	$p > \|t\|$	R^2	RMSE (kg ha^{-1})	BIC
0.5	Linear	β_0 parameter	–206,767	30,719	–6.73	< 0.0001	0.771	29,453	810.0
		h_{95}	16,422	1,471	11.16	< 0.0001			
	Power function	β_0 parameter	154.8	157.2	0.98	0.3312	0.632	37,319	828.5
		h_{60}	2.345	0.344	6.82	< 0.0001			
	Exponential	β_0 parameter	9.010	0.241	37.39	< 0.0001	0.831	25,337	798.2
		h_{75}	0.1450	0.0116	12.50	< 0.0001			
4	Linear	β_0 parameter	–186,527	26,791	–6.96	< 0.0001	0.797	27,723	805.3
		h_{75}	16,852	1,398	12.06	< 0.0001			
	Power function	β_0 parameter	60.83	51.08	1.19	0.2413	0.760	30,120	811.7
		h_{60}	2.646	0.281	9.42	< 0.0001			
	Exponential	β_0 parameter	9.115	0.202	45.12	< 0.0001	0.857	23,269	791.6
		h_{75}	0.1371	0.0095	14.43	< 0.0001			

[1] For the power function and exponential models, the standard errors are approximates values.

tions, was used for forest attribute estimation. Therefore, the model fit will be influenced by the precision of the point cloud processed into canopy heights and intensities.

Errors in DTM will lead to errors in the normalized canopy height distribution. In addition to errors caused by the sensor (*e.g.* variation in scanning angles due to different and multiple flight lines) and the methods and algorithms used (*e.g.* noise threshold algorithms used to identify the first and last returns), the quality of a laser-derived DTM is affected by data characteristics, such as point density, first/last pulse, footprint size, flight height or scan angle, and by errors caused by the complexity of the target, such as type and flatness of terrain, density of the canopy or amount and height of understory (Hyyppä *et al.*, 2008; Thomas *et al.*, 2006). However, relatively good canopy height information can be collected with various parameter configurations. Among the above factors, pulse density can be considered the most influential (Hyyppä *et al.*, 2008).

The results presented in this study demonstrate that descriptive variables, which are essential to biomass assessment, can be modelled with reasonable precision with medium- and low-density laser data obtained from Atlantic *Eucalyptus globulus* plantations. The modelling results for W_{cr} ($R^2 = 0.75$, RMSE = 3.50 Mg ha^{-1} for 4 pulses m^{-2}; $R^2 = 0.72$, RMSE = 3.75 Mg ha^{-1} for

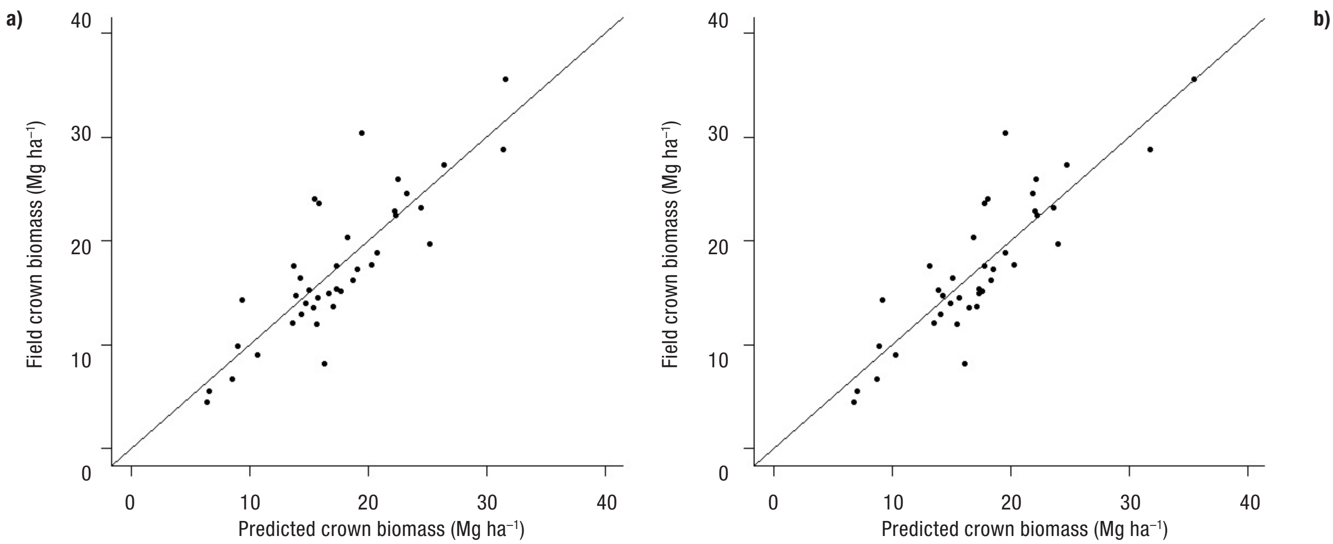

Figure 5. a) Field-measured *versus* predicted W_{cr} for 0.5 pulses m^{-2} dataset ($R^2 = 0.718$; p-value < 0.005). Line shows 1:1 relationship. b) Field-measured *versus* predicted W_{cr} for 4 pulses m^{-2} dataset ($R^2 = 0.753$; p-value < 0.005). Line shows 1:1 relationship.

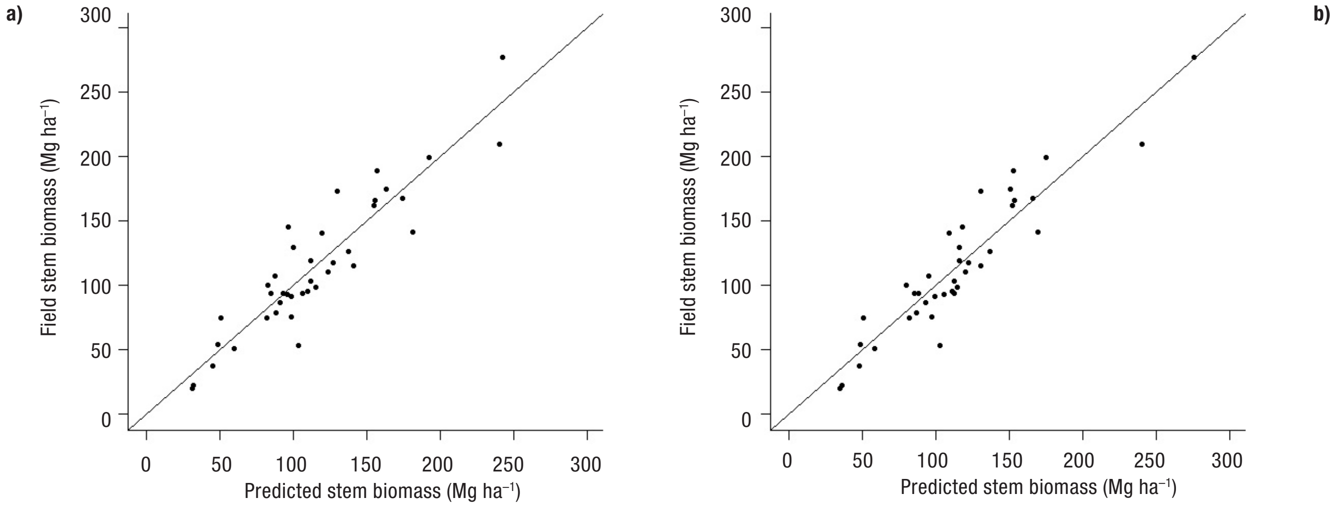

Figure 6. a) Field-measured *versus* predicted W_{st} for 0.5 pulses m^{-2} ($R^2 = 0.841$; p-value < 0.005). Line shows 1:1 relationship. b) Field-measured *versus* predicted W_{st} for 4 pulses m^{-2} ($R^2 = 0.866$; p-value < 0.005) dataset. Line shows 1:1 relationship.

0.5 pulses m^{-2}), W_{st} ($R^2 = 0.87$, RMSE = 20.0 Mg ha^{-1} for 4 pulses m^{-2}; $R^2 = 0.84$, RMSE = 21.8 Mg ha^{-1} for 0.5 pulses m^{-2}) and W_{abg} ($R^2 = 0.86$, RMSE = 23.3 Mg ha^{-1} for 4 pulses m^{-2}; $R^2 = 0.83$, RMSE = 25.3 Mg ha^{-1} for 0.5 pulses m^{-2}) (see Tables 2-4) are similar to those reported in the international literature. Carbon stocks can be easily calculated as a 45.1% of the aboveground biomass estimations; this value corresponds to the weighted average carbon concentration for *Eucalyptus globulus* in Galicia reported by Diéguez-Aranda *et al.* (2009, p: 241).

Aboveground biomass estimates from exponential models (see Table 4) were similar in terms of R^2 to those reported by Lim and Treitz (2004a) and Thomas *et al.* (2006) in Canada. Both studies used LiDAR-derived height quantiles as independent variables. The first, in uneven-aged mature to overmature tolerant hardwood forests, used the 25th percentile as the best explanatory variable; the second one, in mixedwood boreal forests, used the 50th percentile as the best explanatory variable. Modelling results were slightly better in terms of R^2 than those reported by Lim *et al.* (2003a) and Lim and Treitz (2004b) in Canada, by Hall *et al.* (2005) in USA, and by González-Ferreiro *et al.* (2012) in Spain. The first, in deciduous forest ecosystems composed predominantly of *Acer saccharum* Marsh.

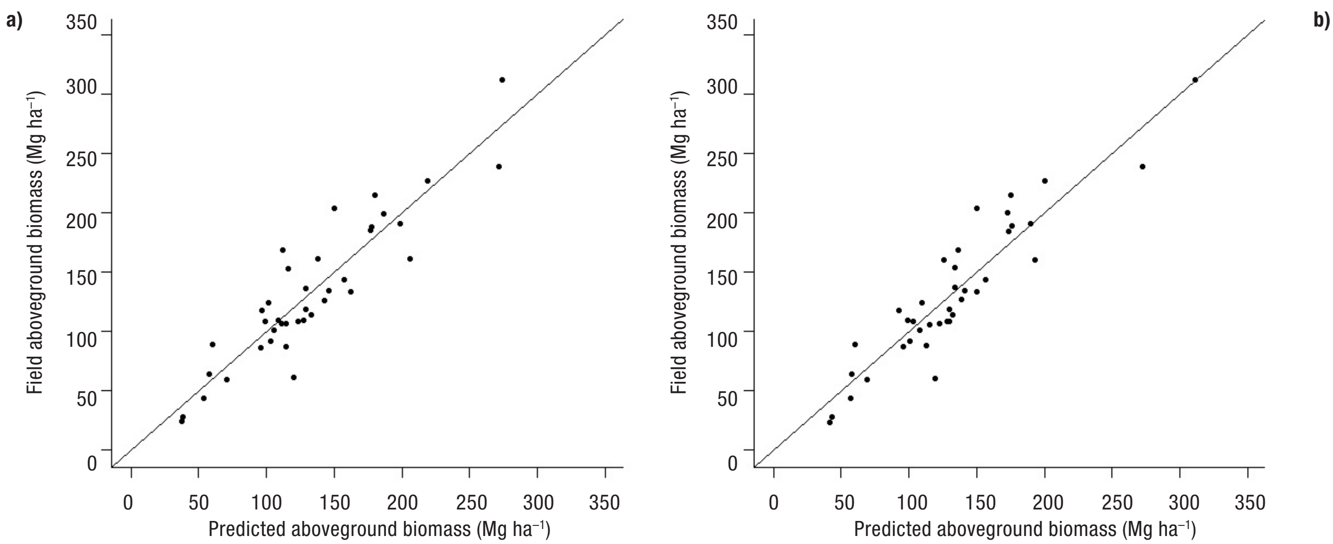

Figure 7. a) Field-measured *versus* predicted W_{abg} for 0.5 pulses m^{-2} ($R^2 = 0.831$; p-value < 0.005). Line shows 1:1 relationship. b) Field-measured *versus* predicted W_{abg} for 4 pulses m^{-2} ($R^2 = 0.857$; p-value < 0.005). Line shows 1:1 relationship.

Table 5. Comparison of the prediction errors for 0.5 pulse m^{-2} and 4 pulses m^{-2} datasets by use of the test of Kruskal-Wallis one-way analysis of variance by ranks

Dependent variable	Model	*H*	*p*-value
Wcr	Exponential	0.0042	0.9482
W_{st}	Exponential	0.0182	0.8927
W_{abg}	Exponential	0.0072	0.9323

and *Betula alleghaniensis* Britton, found that the best explanatory variable was the mean of laser height calculated from LiDAR returns filtered based on a threshold applied to the intensity return values; the second, in mature and homogeneous *Pseudotsuga menziesii* (Mirb.) forests, used the deciles of the distribution of the laser canopy heights as explanatory variables; the third, in *Pinus ponderosa* Dougl. ex Laws. forests, found that the best predictors were the mean height of the highest vegetation return in each m^2, the proportion of ground returns that are also 1st returns, and the proportion of 1st returns that are also ground returns; the last, in *Pinus radiata* D. Don forests, found that the skewness, the standard deviation, and the 5th percentile of height distribution were the best explanatory variables when a 0.5 pulses m^{-2} LiDAR dataset was used, while the maximum intensity, the number of returns above 2 m height, the skewness and the 80th percentile of height distribution were the best explanatory variables when a 8 pulses m^{-2} LiDAR dataset was used. Finally, the R^2 values for aboveground biomass modelling were considerably lower than those obtained by Garcia *et al.* (2010) in Spain and by Treitz *et al.* (2010) in Canada. The former, fitted species-specific models on *Pinus nigra* Arn., *Juniper thurifera* L. and *Quercus ilex* L. and also a general model in which the percentage of intensity of the height percentile 25 and the height percentile 50 were the explanatory variables; the latter, for *Picea mariana* (Mill) B.S.P. forests, used the mean height (for a LiDAR dataset of 1.6 pulses m^{-2}), the mean height and the cumulative proportions of LiDAR returns found in the intervals 6 and 9 of heights (for a LiDAR dataset of 3.2 pulses m^{-2}), and the mean height, the number of first divided by all returns, and the 40th percentile of the height distribution (for a LiDAR dataset of 0.5 pulses m^{-2}).

Several studies based on small-footprint LiDAR data systems have found that height percentiles are highly correlated with stand biomass (Lim and Treitz, 2004b; Patenaude *et al.*, 2004; Treitz *et al.*, 2010; García *et al.*, 2010; González-Ferreiro *et al.*, 2012). The selected exponential models for W_{cr}, W_{st}, and W_{abg} include the 75th percentile of the height distribution (h_{75}) as a single regressor to estimate the different biomass fractions (Tables 2-4), which is consistent with the findings of Gobakken and Næsset (2007), who reported that intermediate and upper height percentiles (h_{50} to h_{90}) remain very stable after a reduction in LiDAR pulse density. This demonstrates the potential of the height percentiles for estimation of several biomass fractions in Atlantic *Eucalyptus globulus* plantations, the stand structure characteristics of which are clearly different from those in previous studies. Similar to Hall *et al.* (2005), none of the regression models finally selected included intensity-derived metrics as explanatory variables; this is in contrast with the results of García *et al.* (2010), who found that intensity-derived variables were more strongly related to biomass than height-related variables. This may be because the intensity-derived explanatory variable was weighted by the mean point density. González-Ferreiro *et al.* (2012) found that independent variables related to return intensity distributions and measurements related to canopy closure may add some valuable information for predicting biomass fractions. In conclusion, but remaining cautious —as some of the papers did not consider the use of the intensity as a predictor— it is expected that height-derived variables are able to explain most of the variability of biomass fractions, while intensity-derived variables only explain a small part of the observed variability, although they may improve models (*e.g.* González-Ferreiro *et al.,* 2012) or be decisive if used in combination with density or height LiDAR values (*e.g.* Lim *et al.,* 2003a; García *et al.,* 2010).

The selected models for W_{cr}, W_{st}, and W_{abg} were very stable after thinning LiDAR pulse density (see Table 5), which is consistent with the findings reported by Treitz *et al.* (2010) and by González-Ferreiro *et al.* (2012), who did not find any evidence, at significance levels of respectively 10% and 5%, that a reduction in LiDAR density affected model precision for stand aboveground biomass. García *et al.* (2010) concluded that reduction of point density had little effect on the results of general models for mixed stands of *Pinus nigra* Arn., *Juniper thurifera* L., and *Quercus ilex* L.; with the exception of the species-specific equations fitted for *Pinus nigra* Arn. and *Quercus ilex* L. Thomas *et al.* (2006) asserted, on the basis of Q-Q analysis and comparison of R^2 values, that models performed similarly for low and high density LiDAR data Despite the stability to reduction demonstrated in the present study, all R^2 values associated with the models tested were higher when the variables were modelled with full density data (see Tables 2-4).

Exponential models performed better for all the fitted variables, using both 4 and 0.5 pulses m^{-2} datasets. Differences in R^2 of more that 5% were obtained between the "best" and "worst" model for each dependent variable, which confirms the importance of model selection. For all predicted variable power function models (one of the most widely used types of model for predicting forest stand variables with LiDAR: see *e.g.* Næsset, 1997, 2002, 2004; Næsset and Bjerknes, 2001; Næsset and Økland, 2002; Hollaus *et al.*, 2007) provided the poorest results in terms of R^2 for both datasets. Although the difference between power function and exponential models may be subtle, it becomes gradually more evident as data accumulates.

Conclusions

The results suggest that LiDAR is a valuable tool for estimating biomass fractions in Atlantic *Eucalyptus globulus* plantations. Exponential regression models performed better as regards estimating all biomass fractions, although they must be tested in different types of forest, regions and data ranges in order to verify their general applicability. Low-density LiDAR data (*e.g.* 0.5 pulses m^{-2}) can be used without significant loss of information, but usual density variation across areas should be considered when requesting airborne LiDAR surveys from commercial companies. The possibility of using low-density LiDAR data for retrieving stand biomass fractions is extremely important for inventory of remote and inaccessible areas as well as for monitoring long-term changes in aboveground biomass and carbon changes in the context of the Kyoto Protocol. Furthermore, low-density data will facilitate regional and national forest inventories, because it will reduce monetary costs along with computation, storage and handling efforts (Thomas *et al.*, 2006). Considering the results of this study, the low-density LiDAR data (0.5 pulses m^{-2}) that are released by the Spanish National Geographic Institute (*Instituto Geográfico Nacional* IGN) will be an excellent source of information for reducing the cost of forest inventories, and will thus have implications for forest management. Furthermore, IGN LiDAR data and the proposed models will be available to all stakeholders and the research community, thus facilitating mapping efforts. Subsequently, valuable knowledge about spatial variability in the different fractions of biomass and carbon stocks in Galicia will be added.

Acknowledgements

Galician Government, *Xunta de Galicia, Dirección Xeral de Montes* (09MRU022291PR); Norvento (Multinational energy company) (PGIDT09REM023E); Galician Government, *Dirección Xeral de Ordenación e Calidade do Sistema Universitario de Galicia (Consellería de Educación e Ordenación Universitaria)* and European Social Fund (Official Journal of Galicia – DOG nº 9, p. 2246, exp. 2011/14).

References

Andersen HE, McGaughey RJ, Reutebuch SE, 2005. Forest measurement and monitoring using high-resolution airborne LIDAR. In: Productivity of western forests: a forest products focus (Harrington CA, Schoenholtz SH, eds). Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, USA. pp: 109-120.

Anderson E, Thompson J, Crouse D, Austin R, 2006. Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM. Geoderm 132: 406-415.

Belsley D, 1991. Conditioning diagnostics: collinearity and weak data in regression, first ed. John Wiley & Sons, Inc, New York, USA. 396 pp.

Bortolot ZJ, Wynne RH, 2005. Estimating forest biomass using small footprint LiDAR data: an individual tree-based approach that incorporates training data. ISPRS J Photogramm Remote Sens 59: 342-360.

Boudreau J, Nelson RF, Margolis HA, Beaudoin A, Guindon L, Kimes DS, 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Quebec. Remote Sens Environ 112: 3876-3890.

Clutter JL, Forston JC, Pienaar LV, Brister GH, Bailey RL, 1983. Timber management: a quantitative approach, first ed. John Wiley & Sons, Inc, New York, USA. 333 pp.

Confemadera, 2010. Informe de resultados: industria de la madera de Galicia, 2010 [online]. Available in: http://vifoga.org/portal/images/stories/Presentaci%C3%B3n%20de%20resultados2010OK.pdf [19 Dec 2012].

Diéguez-Aranda U, Rojo Alboreca A, Castedo-Dorado F, Álvarez González JG, Barrio-Anta M, Crecente-Campo F, González González JM, Pérez-Cruzado C, Rodríguez Soalleiro R, López-Sánchez CA, Balboa-Murias MA, Gorgoso Varela JJ, Sánchez Rodríguez F, 2009. Herramientas selvícolas para la gestión forestal sostenible en Galicia, first ed. Xunta de Galicia, Santiago de Compostela, Spain. 259 pp.

Donoghue DNM, Watt PJ, Cox NJ, Wilson J, 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. Remote Sens Environ 110: 509-522.

Drake JB, Knox RG, Dubayah RO, Clark DB, Condit R, Blair JB, Hofton M, 2003. Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Global Ecol Biogeogr 12: 147-159.

Draper NR, Smith H, 1998. Applied regression analysis, 3rd ed. John Wiley & Sons, Inc, New York, USA. 736 pp.

ESRI, 1998. ESRI shapefile technical description. White paper, first ed [online]. Environmental Systems Research Institute, Inc, Redlands, USA. 34 pp. Available in: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf [19 Dec 2012].

Estornell J, Ruiz L, Velázquez-Martí B, Fernández-Sarría A, 2011. Estimation of shrub biomass by airborne LiDAR data in small forest stands. For Ecol Manage 262: 1697-1703.

Field CB, Campbell JE, Lobell DB, 2008. Biomass energy: the scale of the potential resource. Trends Ecol Evol 23: 65-72.

Freppaz F, Minciardi R, Robba M, Rovatti M, Sacile R, Taramasso A, 2004. Optimizing forest biomass exploitation for energy supply at a regional level. Biomass Energy 26: 15-25.

Garcia M, Fiano D, Chuvieco E, Danson FM, 2010. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. Remote Sens Environ 114: 816-830.

Gobakken T, Næsset E, 2007. Assessing effects of laser point density on biophysical stand properties derived from airborne laser scanner data in mature forest. Int Arch Photogram Rem Sens Spatial Inform Sci 36: 150-155.

Gonçalves-Seco L, González-Ferreiro E, Diéguez-Aranda U, Fraga-Bugallo B, Crecente R, Miranda D, 2011. Assessing attributes of high density *Eucalyptus globulus* stands using Airborne Laser Scanner data. Int J Remote Sens 32: 9821-9841.

González-Ferreiro E, Diéguez-Aranda U, Miranda D, 2012. Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. Forestry 85: 281-292

Goodwin NR, Coops NC, Culvenor DS, 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. Remote Sens Environ 103: 140-152.

Gross H, Jutzi B, Thoennessen U, 2008. Intensity normalization by incidence angle and range of full-waveform lidar data. Int Arch Photogram Rem Sens Spatial Inform Sci 37: 405-412.

Gueudet P, 2004. The influence of post spacing density of DEMs derived from LIDAR on flood modelling. Doctoral thesis. University of Texas at Austin.

Hall S, Burke I, Box D, Kaufmann M, Stoker J, 2005. Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. For Ecol Manage 208: 189-209.

Heurich M, Thoma F, 2008. Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. Forestry 81: 645-661.

Hodgson ME, Jensen JR, Schmidt L, Schill S, Davis B, 2003. An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. Remote Sens Environ 84: 295-308.

Hodgson ME, Jensen J, Raber G, Tulils J, Davis BA, Thompson G, Schuckman K, 2005. An evaluation of LIDAR-derived elevations and terrain slope in leaf-off conditions. Photogramm Eng Remote Sens 71: 817-23.

Höfle B, Pfeifer N, 2007. Correction of laser scanning intensity data: data and model-driven approaches. ISPRS J Photogramm Remote Sens 62: 1415-1433.

Hollaus M, Wagner W, Maier B, Schadauer K, 2007. Airborne laser scanning of forest stem volume in a mountainous environment. Sensors 7: 1559-1577.

Hyyppä J, Hyyppä H, Leckie D, Gougeon F, Yu X, Maltamo M, 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. Int J Remote Sens 29: 1339-1366.

Jutzi B, Gross H, 2010. Investigation on surface reflection models for intensity normalization in Airborne Laser Scanning (ALS) data. Photogramm Eng Remote Sens 76: 1051-1060.

Kraus K, Mikhail EM, 1972. Linear least squares interpolation. Photogramm Eng 38: 1016-1029.

Kraus K, Pfeifer N, 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS J Photogramm Remote Sens 53: 193-203.

Kruskal WH, 1952. A nonparametric test for the several sample problem. Ann Math Stat 23: 525-540.

Kruskal WH, Wallis WA, 1952. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47: 583-621.

Lim KS, Treitz PM, 2004a. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. Scand J Forest Res 19: 558-570.

Lim KS, Treitz PM, 2004b. Estimation of aboveground forest biomass using airborne scanning discrete return LIDAR in Douglas-fir. Int Arch Photogram Rem Sens Spatial Inform Sci 36: 149-152.

Lim K, Treitz P, Baldwin K, Morrison I, Green J, 2003a. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Can J Remote Sens 29: 658-678.

Lim K, Treitz P, Wulder M, St-Onge B, Flood M, 2003b. LiDAR remote sensing of forest structure. Progress in Physical Geography 27: 88-106.

Liu X, Zhang Z, 2008. Lidar data reduction for efficient and high quality dem generation. Int Arch Photogram Rem Sens Spatial Inform Sci 37: 173-178.

Lovell JL, Jupp DLB, Newnham GJ, Coops NC, Culvenor DS, 2005. Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. For Ecol Manage 214: 398-412.

Magnusson M, 2006. Evaluation of remote sensing techniques for estimation of forest variables at stand level. Doctoral thesis. Swedish University of Agricultural Sciences, Umeå.

Magnusson M, Fransson J, Holmgren J, 2007. Effects on estimation accuracy of forest variables using different pulse density of laser data. For Sci 53: 619-626.

Maltamo M, Eerikäinen K, Packalén P, Hyyppä J, 2006. Estimation of stem volume using laser scanning-based canopy height metrics. Forestry 79: 217-229.

Mazzarini F, Pareschi MT, Favalli M, Isola I, Tarquini S, Boschi E, 2007. Lava flow identification and aging by means of lidar intensity: mount Etna case. J Geoph Research 112, B02201. doi:10.1029/2005JB004166.

McGaughey R, 2009. FUSION/LDV: software for LIDAR Data Analysis and Visualization, US Department of Agriculture, Forest Service, Pacific Northwest Research Station, Seattle, USA. 123 pp.

Musk RA, Osborn JE, 2007. Calibrating LiDAR derived canopy metrics to account for data aquisition parameters and forest condition in Radiata pine plantations [online]. In: School of Geography and Environmental Studies Conference, Hobart, Tasmania. pp: 1-6. Available in: http://eprints.utas.edu.au/5548/ [19 Dec 2012].

Myers RH, 1990. Classical and modern regression with applications, 2nd ed. Duxbury Press, Belmont, CA, USA. 488 pp.

Næsset E, 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sens Environ 61: 246-253.

Næsset E, 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens Environ 80: 88-99.

Næsset E, 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scand J Forest Res 19: 164-179.

Næsset E, Økland T, 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sens Environ 79: 105-115.

Næsset E, Bjerknes K, 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. Remote Sens Environ 78: 328-340.

Næsset E, Gobakken T, 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. Remote Sens Environ 112: 3079-3090.

Patenaude G, Hill RA, Milne R, Gaveau DLA, Briggs BBJ, Dawson TP, 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. Remote Sens Environ 93: 368-380.

Peña D, 2002. Regresión y diseño de experimentos. Alianza Editorial SA, Madrid, España. 744 pp.

Pérez-Cruzado C, Merino A, Rodríguez-Soalleiro R, 2011. A management tool for estimating bioenergy production and carbon sequestration in *Eucalyptus globulus* and *Eucalyptus nitens* grown as short rotation woody crops in north-west Spain. Biomass and Bioenergy 35: 2839-2851.

Popescu SC, 2007. Estimating biomass of individual pine trees using airborne Lidar. Biomass and Bioenergy 31: 646-655.

Popescu SC, Wynne RH, Nelson RF, 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can J Remote Sens 29: 564-577.

Puetz A, Olsen R, Anderson B, 2009. Effects of lidar point density on bare earth extraction and DEM creation. In: Laser radar technology and applications (Turner MD, Kamerman GW, eds) . XIV Proceedings of the SPIE 7323. pp: 1-8.

Raber GT, Jensen JR, Schill SR, Schuckman L, 2002. Creation of digital terrain models using an adaptive lidar vegetation point removal process. Photogram Eng Remote Sens 68: 1307-1315.

Raber G, Jensen J, Hodgson M, Tullis J, Davis B, Berglund J, 2007. Impact of lidar nominal post-spacing on dem accuracy and flood zone delineation. Photogramm Eng Remote Sens 73: 793-804.

Reitberger J, Krzystek P, Stilla U, 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. Int J Remote Sens 29: 1407-1431.

Rombouts J, Ferguson IS, Leech JW, 2008. Variability of LiDAR volume prediction models for productivity assessment of radiata pine plantations in South Australia (Hill RA, Rosette J, Suárez J, eds). In: Proceedings of Silvilaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory Edinburgh, UK. pp: 39-49.

Rosenqvist A, Milne A, Lucas R, Imhoff M, Dobson C, 2003. A review of remote sensing technology in support of the Kyoto Protocol. Environ Sci Policy 6: 441-455.

Ryan TP, 1997. Modern Regression Methods, first ed. John Wiley & Sons, Inc, New York, USA. 515 pp.

SAS Institute Inc, 2004. SAS/STAT® 9.1 User's Guide. SAS Institute Inc, Cary, NC, USA.

Schwarz G, 1978. Estimating the dimension of a model. Ann Stat 6: 461-464.

Shapiro SS, Wilk MB, 1965. An analysis of variance test for normality (complete samples). Biometrika 52: 591-611.

Shapiro SS, Wilk MB, 1968. Approximations for the null distribution of the W statistic. Technometrics 10: 861-866.

Sherrill KR, Lefsky MA, Bradford JB, Ryan MG, 2008. Forest structure estimation and pattern exploration from discrete-return lidar in subalpine forests of the central rockies. Can J For Res 38: 2081-2096.

Sheskin DJ, 2004. Handbook of parametric and nonparametric statistical procedures, 3rd ed. Chapman & Hall/CRC, Boca Ratón, FL, USA. 1193 pp.

Silva-Pando FJ, González-Hernández MP, Prunell Tuduri A, 1993. Prácticas agroforestales en pinares y eucaliptales atlánticos. I. Producción del sotobosque II (Silva-Pando, FJ, Vega G, eds). In: Actas del I Congreso Forestal Español (Lourizán, España). pp: 637-642.

Singh KK, Vogler JB, Meentemeyer RK, 2010. Estimation of land-use in an urbanized landscape using lidar intensity data: a regional scale approach. Int Arch Photogram Rem Sens Spatial Inform Sci 38; 1-4.

Tesfamichael SG, Van Aardt JAN, Ahmed F, 2010. Estimating plot-level tree height and volume of *Eucalyptus grandis* plantations using small-footprint, discrete return lidar data. Prog Phys Geog 34: 515-540.

Thomas V, Treitz P, McCaughey JH, Morrison I, 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. Can J For Res 36: 34-47.

Treitz P, Kevin L, Murray W, Doug P, Nesbitt D, Etheridge D, 2010. LiDAR data acquisition and processing protocols for forest resource inventories in Ontario, Canada (Koch B, Kendlar G, eds). In: Silvilaser 2010: the 10th International Conference on LiDAR Applications for Assessing Forest Ecosystems Freiburg, Germany. pp: 1-10.

UNFCCC, 1997. Kyoto Protocol to the United Nation Framework Convention on Climate Change [online]. Available in: http://unfccc.int/kyoto_protocol/items/2830.php [19 Dec 2012].

Wack R, Schardt M, Lohr U, Barrucho L, Oliveira T, 2003. Forest inventory for Eucalyptus plantations based on airborne laser scanner data. Int Arch Photogram Rem Sens Spatial Inform Sci 34: 40-46.

Wagner W, Hollaus M, Briese C, Ducic V, 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. Int J Remote Sens 29: 1433-1452.