Estimation of aerial biomass using discrete-wave LiDAR data in combination with different vegetation indices in plantations of *Pinus radiata* (D. DON), Región del Maule, Chile.

Diego Valencia Delgado[1,], Jaime Hernández Palma[1], Fabian Fassnacht[2], Lissette Cortés Serey[1], Javier Lopatin Fourcade[1], Patricio Corvalán Vera[1]

1- Laboratorio de Geomática y Ecología del Paisaje. Universidad de Chile. Email: diegovalencia@ug.uchile.cl

2- Chair of Remote Sensing and Landscape Information Systems, University of Freiburg, Freiburg, 79085, Germany

Abstract

The aerial biomass of *Pinus radiata* plantations in the Región del Maule, Chile, was estimated from linear models using databases of LiDAR and multispectral LANDSAT ETM+. Six descriptive height variables were obtained from the LiDAR point cloud; the 25%, 50%, 75%, 95% and 100% percentiles and the mean height. Two variables associated with the density of points were also obtained, which relate the returns between fixed weighted intervals calculated as a function of the observed biomass. For multispectral variables we used NDVI, corrected NVDI (NDVIc) and the "Tasseled Cap" components brilliance, greenness and humidity. The results showed coefficients of determination (R^2) between 0.801 and 0.814, with errors between 36.07 and 36.11 ton ha^{-1} for the models generated using height percentiles, and R^2 from 0.807 to 0.823 with errors between 36.06 and 36.84 ton ha-1 for transformed LiDAR data. Finally, the stepwise model using all available variables had R2 of 0.821-0.835 with errors of 34.28 - 36.31 ton ha-1.

Key words: ALS, forest above ground biomass, point cloud density, LiDAR, NDVIc.

Resumen

La biomasa aérea en bosques de pino insigne en la región del Maule, Chile, fue estimada utilizando modelos lineales sobre la base de datos LiDAR y multiespectrales de LANDSAT ETM+. De la nube de puntos LiDAR se obtuvo un total de seis variables descriptivas de altura, los percentiles 25% , 50%, 75% , 95% , 100% y la altura promedio, y dos variables asociadas a la densidad de puntos, las cuales relacionan los retornos entre intervalos fijos ponderadores calculados en función de la biomasa observada. Para las variables multiespectrales, se utilizó: El NDVI, el NDVI corregido (NDVIc) y los componentes "Tasseled Cap" Brillantez, Verdor y Humedad. Los resultados mostraron coeficientes de determinación (R^2) entre 0,801 y 0,814 con errores entre 36,07 y 36,11 ton ha^{-1} para los modelos generados a partir de percentiles de altura y R^2 entre 0,807 y 0,823 con errores entre 36,06 y 36,84 ton ha^{-1} para datos de transformaciones de información LiDAR. Finalmente, el modelo "Stepwise" que involucra todas las variables disponibles tiene un ajuste de R^2 entre 0,821 y 0,835 con errores entre 34,28 y 36,31 ton ha^{-1}.

Introduction

The use of the LiDAR (Light Detection and Ranging) system in forestry application has important advantages to improve the management of natural resources both in the public and private sectors. Use of the LiDAR system in forest resource management is increasing, especially in the estimation of forest parameters such as tree height, DAP, standing volume, crown width, basal area, aerial biomass and dead trees, among others (Hudak et al. 2002, Lim and Treitz 2004, Popescu 2007; Falkowski et al. 2009, Martinuzzi et al. 2009). LiDAR analysis also allows characterization of the elements of the crown in three dimensions (Reitberger et al. 2009), providing good estimations in the vertical plane, especially in conifers (Popescu 2007, Sherrill et al. 2008). In recent years there have been attempts to combine LiDAR data with other biophysical variables, such as the Leaf Area Index (LAI), the Normalized Difference Vegetation Index (NDVI) and crown cover to generate combined models (Naesset 1997, 2002, Lim et al. 2003, Riaño et al. 2004, Morsdorf et al. 2006, Sasaki et al. 2008, Ioki et al. 2009, among others).

Discrete wave systems are considered to have a small footprint, with a size from 0.1 to 3.0 meters (Ioki et al. 2009). The distance is calculated by the time to return or the intensity of the

reflected signal (Lefsky et al. 2002), which are strongly influenced by the crown conditions and also by the size of the footprint and the scanning angle (Koch and Dees 2008). This system is the most utilized to obtain information on forest parameters, both at the stand level (Lim and Treitz 2004, Koch and Dees 2008, Ioki et al. 2009, Kim et al. 2009) and for individual trees (Popescu 2007, Sherrill et al. 2008, Koch and Dees 2008, Suratno et al. 2009, Forzieri et al. 2009, Zhao et al. 2009). This is due to the good spatial resolution that the small footprint size has and the high rate of repetition of these systems, up to 33,000 repetitions per second, which together can produce the point density of the sample. This ability to aggregate data on data of the specific areas and scales during the data analysis allows studying the locations in the field, characterizing a given forest, a plot sample or the crown of an individual tree (Lefsky et al. 2002). These characteristics allow the integration of the spectral information from passive sensors (*i.e*. LANDSAT ETM+ or QuickBird) in the estimation of crown sizes and confirmation of the information in the field (Hudak et al. 2002, Zhao et al. 2009, Latifi et al. 2010).

Biomass estimation calculated at the stand level is generally expressed in surface units (*i.e*. ton ha^{-1}), and is produced by summing the biomass values in plots and then extrapolating the to the entire study area (Popescu 2007, Hudak et al. 2008). Lim and Treitz (2004) estimated biomass in mixed forests in northern Canada using LiDAR data, finding adjusted R^2 values of 0.80 to 0.89 for their models. Ioki et al. (2009) compared 21 models of biomass estimation in evergreen forests in Japan; the best adjusted R^2 was 0.75. Popescu et al. (2004) reported adjusted R^2 values of 0.82 for a coniferous forest in northern Virginia, USA.

The objective of this study was to estimate the aerial biomass of plantations of *Pinus radiata* in south-central Chile in different stages of growth, constructing models with information on heights, LiDAR point densities and multispectral LANDSAT TM+ satellite information using indices and tasseled cap transformations. The study was performed on a 400 ha surface area. In Chile 20.7% of the national territory is forested; 14.1% of this area is composed of forest plantations (CONAF 2011). Studies of biomass estimation began in the 1980s, first for *Pinus radiata* plantations (Caldentey 1989) and later for native species (Garfias 1994, Caldentey 1995, Schlegel 2001, Herrera and Waisberg 2002, Schmidt et al. 2009). However, estimation of aerial biomass in Chile continues to use traditional measurements in most cases (Anuchin 1960, Bitterlich 1984, Husch et al. 1993; Avery and Burkhart 1994; Caldentey 1995; Prodan et al. 1997; Schlegel 2001; Herrera and Waisberg 2002; Schmidt et al.. 2009). This study promotes the use of remote technologies to estimate biomass in Chile.

Methods

Study area: the 400 ha study area is located in the Maule Region in a zone with coastal influence in south-central Chile. The area was originally dominated by vegetative formations composed mainly of deciduous *Nothofagus* forests called 'Bosque Maulino' (Gajardo 1994). It is currently covered with plantations of *P. radiata*; there are only a few remnant fragments of the Bosque Maulino, mostly secondary forests of *Nothofagus glauca* (Hualo), *N. Alessandri* (ruil) and *N. dombeyi* (Coihue) (Estades and Temple 1999).

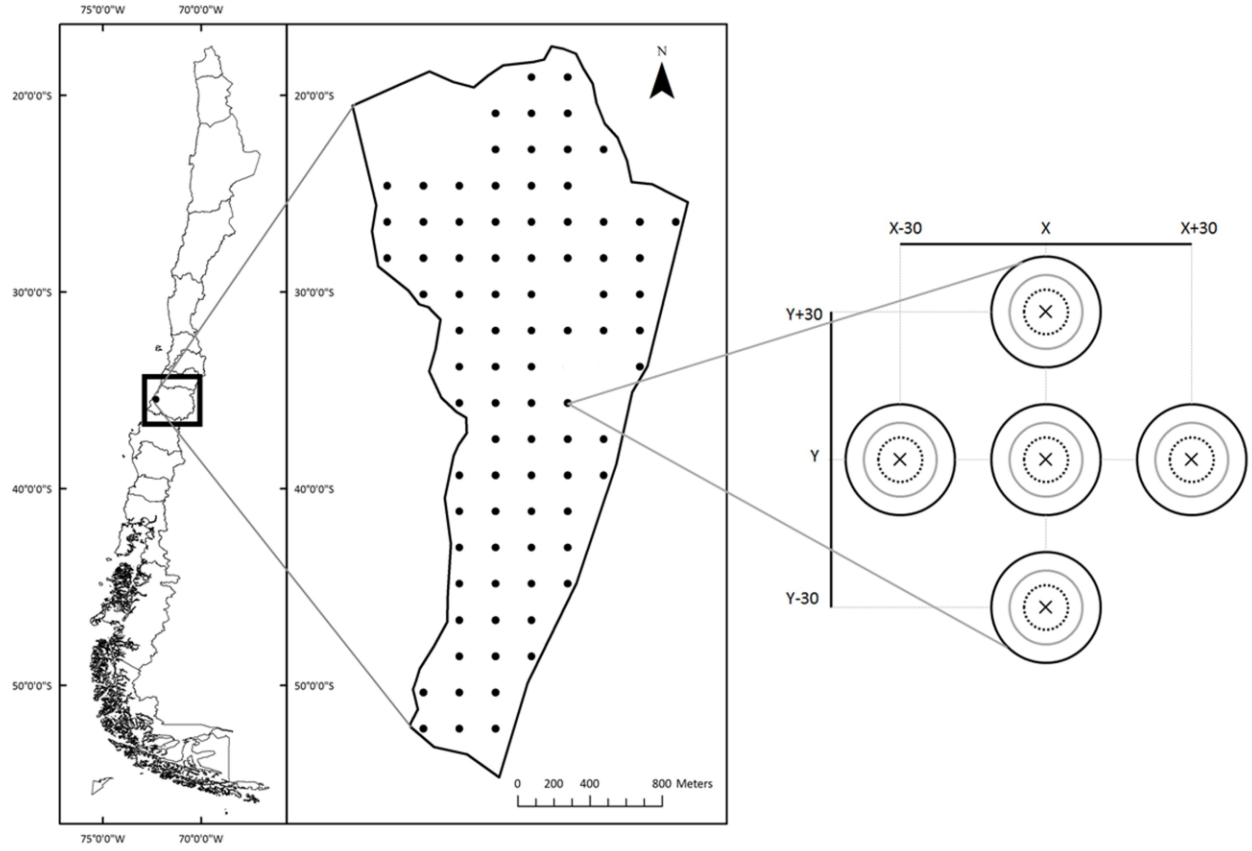

Fig. 1. Study area in Maule Region, South-central Chile. Central positions of clusters are overlays. The right figure represents the inside configuration of each cluster, and shows the relative position of subplots.

Data collection in the field: reference data was obtained in February, 2012 using a systematic sampling (200 x 200 m grid) in concentric circular plots of 2, 4 and 8 m radius, including trees with DAP ≥ 5, 10 and 20 cm, respectively. Each unit was constituted as a cluster of five sub-plots, one

central and four separated from it by 30 m in each of the cardinal directions. In each unit we recorded DAP (cm), species, height (m), start of live crown (m) and mean height of understory (m). The localization of the central plots used a GPS navigator.

LiDAR data and satellite images: LiDAR data were acquired in March, 2012, using a Harrier 54/G4 Dual System sensor mounted on a Piper PA-24 Comanche airplane (Digimapas Chile Ltda). Overflights were conducted at a mean height of 580 m above ground level at a velocity between 180 and 210 km/h, with a density of 4.64 points per m^2 (p m^{-2}). The beam divergence of the laser was 0.5 mrad, which produced a laser footprint of 29 cm. The pulse frequency and scanning frequency were both 100 Hz. The angle of the field of vision was approximately 22.5º, with a wave length of the laser pulse of 1550 nm. Aerial photos (VIS) with an ortho-rectified spatial resolution of 50 cm were obtained simultaneously with LiDAR measurements. Finally, we acquired a free unprocessed LANDSAT ETM+ image from the Earth Explorer Website of the United States Geological Survey (USGS) in February, 2012.

Table 1. Summary of information of plantations in the study site

P. radiata plantations (Mean years of age)	Biomass (ton ha^{-1})*	Basal Area (m^2 ha^{-1})*	Trees ha^{-1}*
Adult (25 años)	147.15	3.93	488.73
Young (10 años)	46.57	3.55	1798.9
Young (5 años)	13.93	2.7	1362.1
Total	98.81	3.61	953.32

* mean values of 141 plots

Biomass data: Biomass data were obtained using the methodology proposed by Corvalán and Hernández (2011), in which total aerial biomass is divided into two components, stem biomass and non-stem biomass. This method was used since local equations for pine aerial biomass are not yet available in Chile.

Stem biomass. We used the allometric function proposed by Madgwick (1994):

$$B_a = e^{(1.028*\ln(((DAP+1.9)^2*H)-4.892)} \quad (1)$$

Where Ba is the aerial biomass of the individual tree in kg, DAP is in cm and H is the height in m.

Non-stem biomass. Corvalán and Hernández (2011) expressed non-stem biomass as the percentage complement of the total aerial biomass as a function of the stand age, mean distance between trees and basal area. The final equation of the percentage of non-stem biomass is obtained from the

$$B_{\%} = 51.71 * e^{\left(\frac{-98^{0.78}}{X}\right)} \quad (2)$$

$$X = E * \sqrt{1000/N} * AB \quad (3)$$

Where $B_{\%}$ is the stem aerial biomass expressed as a percentage of the total biomass, E is the stand age, N is the number of trees per ha and AB is the basal area expressed in $m^2\ ha^{-1}$. Finally, the percentage of non-stem biomass ($NTB_{\%}$) is:

$$NTB_{\%} = 100 - (B_{\%} - 10\%\ cortex) \quad (4)$$

The total biomass per ha in each plot was calculated by summing the stem biomasses of individual trees multiplied by their respective expansion factors plus the non-stem biomass expressed as an aggregated function of the stem biomass over total biomass as a function of stand age, mean distance between trees and basal area.

The LiDAR point cloud was normalized using the TREESVIS software (Weinacker *et al.*. 2004). This process creates a digital terrain model (DTM) which allows transforming the points from elevation above sea level to meters above the ground. The information was extracted from the normalized point cloud in the R software (R Core Team 2014). Two selection radii were used: one of 8 meters, equivalent to the radius proposed in the sampling phase, and one of 12 m, to reduce the uncertainty of location of the plot in the data collection phase.

Processing of multi-spectral images: we obtained two Landsat Enhanced Thematic Mapper Plus (ETM+) images with acquisition dates 8 December 2011 and 9 January 2012. These images were rectified to UTM projection; then radiometric, atmospheric and topographic corrections were applied. For radiometric correction radiance was first converted to digital levels using the specific gain and offset of the sensor. Atmospheric correction was then applied using the dark object subtraction method (Chávez 1988); topographic correction using the method of Civco (1989) and reflectance using a simplified model and the correction parameters given in Chuvieco (2010). Since the images are SLC-off we applied the correction proposed by Scaramuzza *et al.*. (2004). After the corrections, the tasseled cap transformation was obtained from the spectral bands (Crist and Cicone 1984), along with two vegetation indices (NDVI, NDVIc) calculated as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (5)$$

$$NDVI_{C=}\frac{NIR - Red}{NIR + Red} * \left[1 - \frac{nIR - nIR_{min}}{nIR_{max} + nIR_{min}}\right] \quad (6)$$

Where NIR, red and mIR correspond to the reflectance in the near infrared, red and mid-infrared, respectively (Rouse *et al.* 1973; Nemani *et al.* 1993).

Extraction of variables: after data processing was completed, variables were extracted in order to generate aerial biomass estimation models using height percentiles and point density data. The first group included the values which divide the height distribution of each sampling point into percentage functions; we obtained the 25% (H_{25}), 50% (H_{50}), 75% (H_{75}), 95% (H_{95}) and 100% (H_{100}) percentiles and the mean height (H_{mean}). The second group corresponds to information on the number of points at a given height. The procedure was to divide the information of the point cloud of each plot into fixed strata based on the maximum height (H100), which yielded two density variables:

$$D_1 = \alpha_1 * N°_{[100\%-80\%]} * \frac{H_{100}}{20} + \alpha_2 * N°_{[80\%-60\%]} * \frac{H_{100}}{20} + \alpha_3 * N°_{[600\%-40\%]} * \frac{H_{100}}{20} + \alpha_4 * N°_{[40\%-20\%]} * \frac{H_{100}}{20} + \alpha_5 * N°_{[200\%-0\%]} * \frac{H_{100}}{20} \quad (7)$$

$$D_2 = D_1(\frac{N_T}{1000})^{-1} \quad (8)$$

Where the density parameter D_1 is a function of the number of points between defined intervals with respect to the maximum height (H_{100}). The weights α_1, α_2, α_3, α_4 and α_1, were obtained by an iterative process which searches for the best relative weights in relation to the observed aerial biomass. To select the weights for the density model (D_1) we used 3,200,000 combinations of five numbers between 0.1 and 2.0 (20^5 repetitions), which were submitted to an automatic selection process in which the relative weights which generate the best fit were selected iteratively. D_2 is simply a transformation of D_1 which adds the total number of points as a new density constant.

Statistical analysis: according to a number of studies (Naesset 1997, 2004; Lim *et al.*. 2003; Ioki *et al.*. 2009; Sasaki *et al.*. 2008; Latifi *et al.*. 2010), aerial biomass may be estimated using the following expression:

$$BT = \beta_1 * h^{\beta_2} * d^{\beta_3} \quad (9)$$

Where *BT* is the biomass per ha, *h* is the height variable determined in this case by the height percentiles, d is the point density variable or spectral information from satellite images and β_1, β_2 and β_3 are the regression coefficients. These same authors indicated that this expression may be log transformed easily to obtain a linear expression without altering the regression parameters. This linear expression is:

$$\ln BT = \ln(\beta_1) + \beta_2 \ln(h) + \beta_3 ln(d) \quad (10)$$

This type of equation allows estimations with individual variables and with the set of variables, and new parameters may be added easily:

$$\ln BT = \ln(\beta_0) + \sum_{i=1}^{n} \beta_1 ln(V_i) \quad (11)$$

Where *ln (BT*) is the logarithm of the biomass per hectare, Ln (V_i) are all the parameters extracted in the study and β_0 and the β_i are all the regression coefficients. A stepwise multiple regression was used to find the best model fit for all the data of this study.

To compare models and select the best we used the adjusted coefficient of determination R^2 and the root mean square error (RMSE) (12) and the corresponding percentage (RMSE%) (13), considering a correction factor produced by using logarithmic transformation (Baskerville 1972).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)}{n}} \quad (12)$$

$$RMSE\% = \frac{RMSE}{\sum_{i=1}^{n} x_1} \quad (13)$$

where X_i are the observed data X_1 are the data predicted by the model and n is the total number of data.

Results

Height percentiles: the total aerial biomass of the 141 pine plots was fitted using linear regressions, with the height percentiles from the LiDAR information as predictor variables. Table 2 shows the fits for different information extraction radii of the point cloud. With a radius of eight

meters the adjusted R^2 varied from 0.369 to 0.801, with RMSE in the range of 35.59 – 60.28 ton ha^{-1}, corresponding to percentage errors between 38.6% and 75.0%. The fits with best correlation were those of H_{95} (R^2 = 0.801), and H_{100} (R^2 = 0.793). The best fit in terms of percentage errors was found with H_{mean} (RMSE% = 38.6), with an R^2 of 0.769. With a 12 meter radius the fits had adjusted values between 0.372 and 0.814, with RMSE in the range of 35.10 – 58.88 ton ha^{-1}; the percentage errors ranged from 37.7% to 73.0%. As with the smaller radius, the best fits were found with H_{95} (R^2 = 0.814) y con H_{100} (R^2 = 0.811), and the lowest RMSE% was that of (37.7%, R^2 = 0.793).

Multi-spectral variables: a total of 5 models were fit using multi-spectral information; Table 2 summarizes the results of the fits for the different indices and transformations used as predictor variables. The fits had R^2 values between 0.043 and 0.462, with RMSE in the range 59-29 – 78.38 ton ha^{-1}. It should be mentioned that the origin of the information to produce the fits came from pixels with a size of 900 m^2 for estimations performed in an area of 200 m^2. The best fit was obtained using the NDVIc fit, which had an adjusted R^2 of 0.462 and RMSE = 59.29 ton ha^{-1}, or RMSE% = 73.2%.

Using this weighted method, for a radius of selection of 8 meters we obtained an adjusted R^2 of 0.807 with an RMSE of 36.84 ton ha^{-1}; the percentage error was 38.72%. For the 12 m radius the adjusted R^2 was 0.823, with RMSE = 36.06 ton ha^{-1} and percentage error 38.03%.

Stepwise Regression with height percentiles: table 4 shows that for an 8-meter radius the adjusted R2 was 0.814, with RMSE = 35.91 ton ha-1 and 38% error. For the 12-m radius the values were R^2 = 0.826, RMSE = 35.35 ton ha^{-1} and RMSE% = 37.3%. In both cases the values selected by the stepwise regression were H_{75}, H_{95} and H_{mean}; the first two were highly significant ($p < 0.001$).

Stepwise regression with height percentiles and multi-spectral information: for a radius of 8 meters the adjusted R^2 was 0.818, RMSE = 34.28 ton ha^{-1} and RMSE% = 36.5% (Table 4); for 12 meters the values were R^2 = 0.830, RMSE = 35.36 ton ha^{-1} and RMSE% = 37.4%. In both cases the most significant variable ($p < 0.001$) was the H_{95} percentile.

Stepwise regression with point densities and multi-spectral information: the stepwise regression for the 8-meter radius (Table 4) had R^2 = 0.818, RMSE = 37.33 ton ha^{-1} and RMSE% = 39%. For the 12-meter radius these values were R^2 = 0.836, RMSE = 36.41 ton ha^{-1} and RMSE% = 38.2%. For both radii, the significance level of the point density was greater ($p < 0.001$) than for the multi-spectral variables. Within the latter (tasseled cap) variables, the component related to brilliance (TC1) was the first selected for both radii; however, the analysis with 8-meter radius selected variable TC2 (greenness), while in the analysis with 12-meter radius the selected variable was TC3, which is related to humidity.

Table 2. Fits of linear models using the variables (Vi) separately

		Modelo : $ln(B) = ln(b_0) + b_1 * ln(V_i)$					
	Radius	V_i	$ln(b_0)$	b_1	R^2	RMSE	RMSE %
Height Percentiles	8	*H25*	4.188^{***}	0.447^{***}	0.369	60.28	75.0
	8	*H50*	3.085^{***}	0.629^{***}	0.645	38.44	38.4
	8	*H75*	2.012^{***}	0.905^{***}	0.679	38.74	45.7
	8	*H95*	0.278^{+}	1.413^{***}	0.801	36.11	38.8
	8	*H100*	-0.813^{***}	1.664^{***}	0.793	39.00	42.2
	8	*Hmean*	2.118^{***}	1.059^{***}	0.769	35.59	38.6
	12	*H25*	4.205^{***}	0.454^{***}	0.372	58.88	73.0
	12	*H50*	2.980^{***}	0.670^{***}	0.678	37.29	42.9
	12	*H75*	1.875^{***}	0.954^{***}	0.715	36.91	42.8
	12	*H95*	0.156^{+}	1.448^{***}	0.814	36.07	38.7
	12	*H100*	-1.248^{***}	1.778^{***}	0.811	39.23	42.1
	12	*Hmean*	2.036^{***}	1.093^{***}	0.793	35.10	37.7
Density of points	8	*D1*	-2.262^{***}	0.907^{***}	0.723	43.62	47.0
	8	*D2*	3.697^{***}	1.250^{***}	0.807	36.84	38.7
	12	*D1*	-3.645^{***}	1.018^{***}	0.708	46.43	50.5
	12	*D2*	3.948^{***}	1.520^{***}	0.823	36.06	38.0
Multiespectral	-	*NDVI*	5.273^{***}	1.329_{***}	0.073	76.93	123.1
	-	*NDVIc*	5.623^{***}	1.156^{***}	0.462	59.29	73.2
	-	*TC1*	36.598^{***}	-7.144^{**}	0.164	71.91	110.5
	-	*TC2*	2.264^{**}	0.547^{**}	0.042	78.38	127.3
	-	*TC3*	0.246^{+}	0.999^{***}	0.129	74.55	113.8

Significance level: $^{+}\ \alpha < 0.1$, $^{*}\ \alpha < 0.05$, $^{**}\ \alpha < 0.01$, $^{***}\ \alpha < 0.001$

Table 3. Estimated weights for the point density equation

Weight	α_1	α_2	α_3	α_4	α_5
Radius 8 m	1.8	1.2	0.7	0.1	0.5
Radius 12 m	1.8	0.3	1.2	0.1	0.4

All the variables studied: For the 8-meter radius the values obtained were $R^2 = 0.821$, RMSE = 34.28 ton ha^{-1} and RMSE% = 36.2%; for the 12-meter radius they were $R^2 = 0.835$, RMSE = 36.31 ton ha^{-1} and RMSE% = 38.18% (Table 4). For the 8-meter radius, the stepwise regression selected the

H_{95} percentile, the point density (D2) and component TC2, while for the 12-meter radius the method found density (D2) to be the most significant ($\alpha < 0.001$) but none of the height percentiles contributed significantly to the regression. Component TC3 should also be mentioned as part of the selected multi-spectral information.

Discussion

The first part of this study examined the information from height percentiles obtained by processing the information from a discrete-wave laser system to estimate aerial biomass. Good correlations and highly significant models were obtained with these variables (Table 2), with results similar to those reported from other countries (Naesset 1997, 2002, 2004, Holmgren 2004, Lim and Treitz 2004, Popescu *et al.*. 2004, Ioki *et al.*. 2009). Figure 2 is a graph of the two best models; the fits, errors and heteroscedasticity may be seen, which indicate that the information does not have a homogeneous distribution in these models.

The explanation of this result is found in the data used, since the study area does not have the age classes necessary to respond to a model which takes into account all size classes and heights.

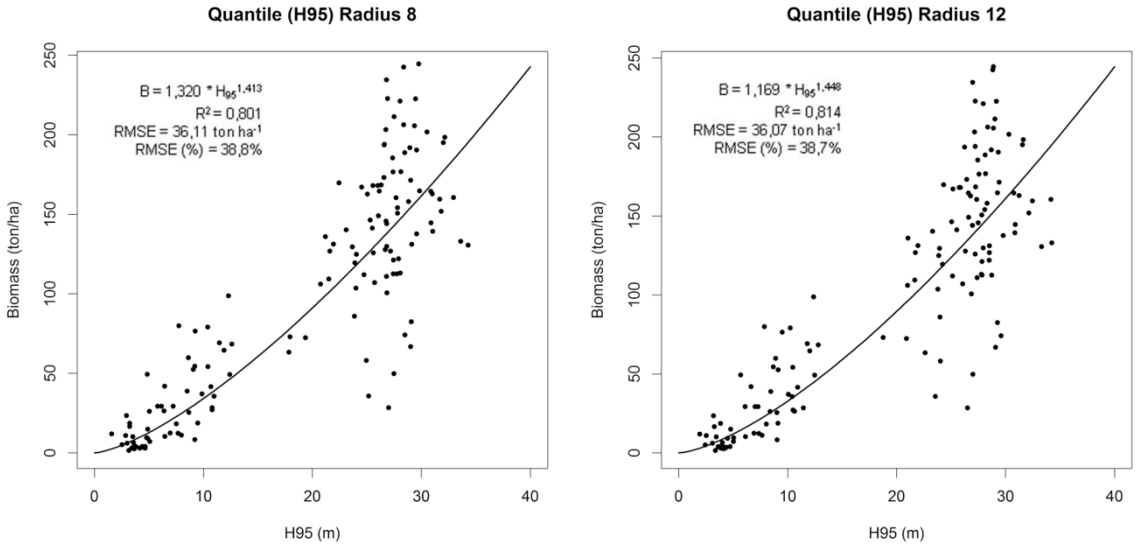

Fig. 2. Fit of the model of aerial biomass with the H_{95} percentile. Left: radius of 8 meters; Right: radius of 12 meters.

Figure 3 illustrates the different conditions present in the study site, in which there were three size classes: (A) stands of 25-year old adult trees, with good laser return percentages from the ground, which is usual in plantations of large, spaced trees with little undergrowth; (B) young stands of age 10 years in unfavorable stand conditions; and (C) very young stands of age 5 years in which the ground is not visible, with heights of no more than seven meters and with low biomass (*i.e.* 10 ton ha^{-1}). This analysis of the study area indicates that these models will never represent a general model; they are more local models which can characterize and provide good estimates of aerial biomass of plantations. The majority of studies have focused on this type of analysis. For example, Naesset (2004), who studied an area of 6500 ha, proposed a stratification methodology using young stands, adult stands in good sites and adult stands in poor sites. He also used a sub-stratification using the density of ground returns as an indicator of the vegetation density present in each sampling unit. With this he obtained very specific models at a local level, which he used to analyze not only biomass, but also basal area, volume, DAP, number of trees and dominant height.

Lim and Treiz (2004) obtained R^2 for fits similar to those obtained in this study; however, their RMSE values were much lower due to the segmented design of their experiment, which included plots in areas with different silvicultural treatments. Holmgren (2004) obtained higher correlations due to transformations and different utilization of the data of height percentiles. These were not used directly; however, as in the present study he obtained a high level of significance ($p< 0.001$) with height percentiles, especially H_{95}.

We also explored transformations of LiDAR data, from which a new variable of point density was constructed which has not been previously reported. The construction of this type of information was inspired by the studies of Ioki et al. (2009) y Sasaki et al. (2008); both studies used transformations which combine relations between ground points and aerial points, assuming some type of relational proportion to the density of the vegetation. The entrance data for this study did not include information on the return number, which is necessary to use the models of these two groups; thus we chose a combined formula which weights defined segments according to the maximum height of the plot in order to relate them to the biomass. The weights were chosen by an iteration process which searches for the best fit with the available information on aerial biomass. As well as providing a good prediction of biomass, the density data provide information on the vertical structure of the plots; the weights may be considered is indicators of the relation between the biomass and the vegetation structure. For example, our results indicate a much more significant relation of biomass with the upper and middle sections of the plantations.

The second part of this study combined LiDAR information with multi-spectral variables into the same models. First we analyzed the fit of the spectral information from LANDSAT ETM+ images, using the same logarithmic model used for fitting LiDAR data. The results showed poor fits

and RMSE which in most cases were above 100%; however, all variables utilized were highly significant. Zheng *et al.* (2004) obtained biomass fits using the NDVIc index as predictor variable; they reported an R^2 of 0.86, considerably greater than those found in this study. It should be noted that these authors used 55 sampling plots greater than 60 m in radius, distributed in a surface of 400 ha. In spite of the differences in sampling design, in both cases NDVIc was a good predictor of aerial biomass; this is due to the characteristics of this index, which is a transformation of NDVI that takes into account the effects of the undergrowth in stands which are not too closed (Nemani et al. 1993).

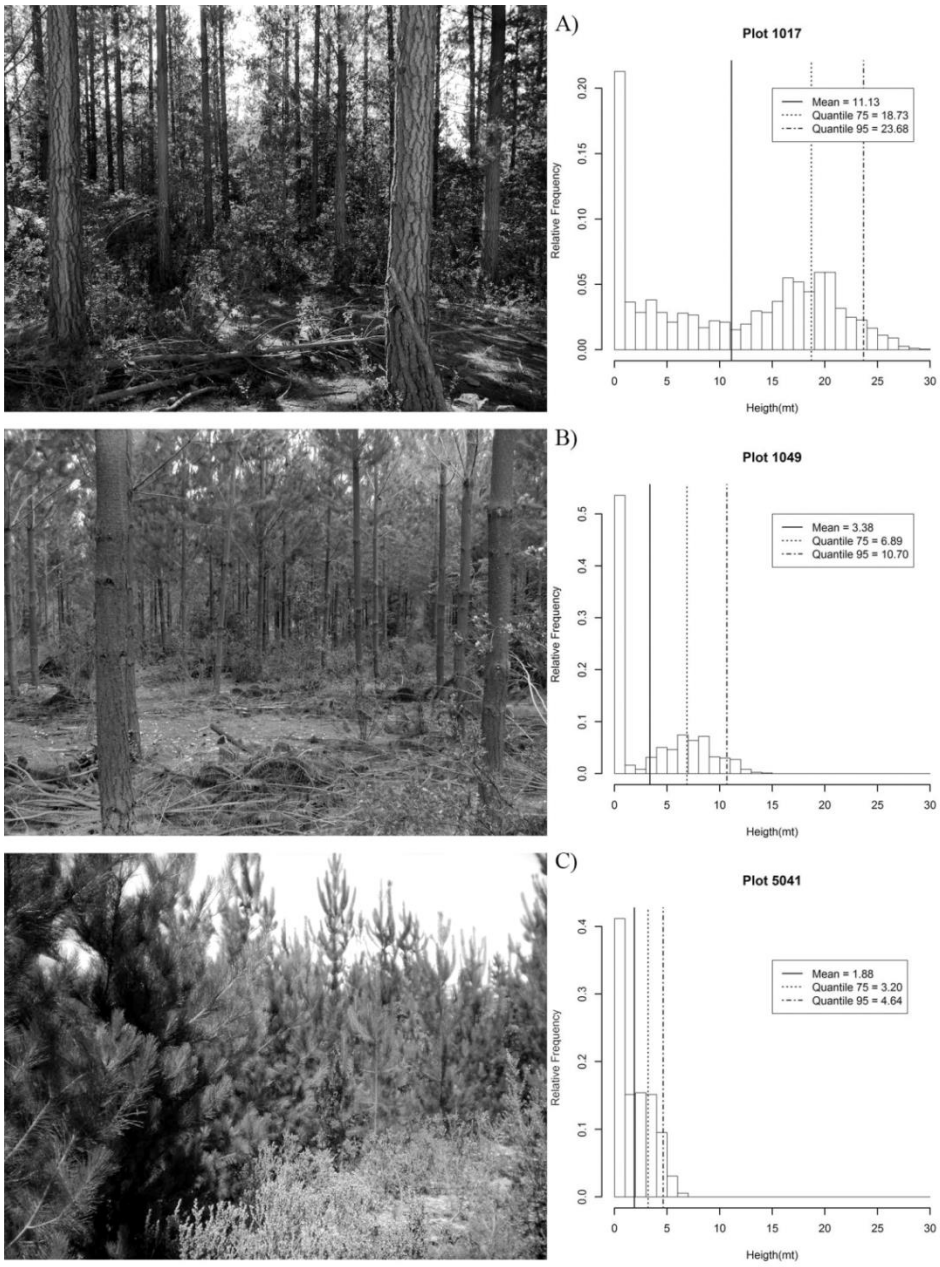

Fig. 3. Photographs and histograms of the different situations found in the study area: (A) Plot 1017; stand of adult pines planted in 1986. (B) Plot 1049; stand of pines planted in 2001. (C) Plot 5041; stand of pines planted in 2005.

Table 4. Results of stepwise regression of the variables: a) height percentiles. b) height percentiles and multi-spectral variables. c) point density and multi-spectral information. d) all variable used in the study.

Modelo: $Ln(B)= Ln(\beta_0)+\sum_{i=1}^{N} \beta_i Ln(V_i)$

		Radio	V_i	$ln(b_0)$	b_1	b_2	b_3	b_4	b_5	b_6	R^2	RMSE	RMSE %
Stepwise	a	8	$H_{75,}\ H_{95,}\ H_{mean}$	0.449	-0.605***	1.424***	0.638*	-	-	-	0.814	35.91	38.0
		12	$H_{75.}\ H_{95.}\ H_{mean}$	0.495	0.762***	-0.638***	1.341*	-	-	-	0.826	35.35	37.3
	b	8	$H_{95.}$ *NDVI. NDVIc. TC1. TC2*	-8.655*	1.447***	-2.622*	2.918**	2.990*	-1.002**	-	0.818	34.28	36.5
		12	$H_{95.}$ *NDVI. NDVIc. TC1. TC2*	-9.405*	1.556***	-2.095*	2.601**	3.134*	-1.022**	-	0.830	35.36	37.4
	c	8	$D_{2.}$ *NDVI. NDVIc. TC1. TC2*	-4.769	1.269***	-2.588*	2.948**	2.904*	-0.944*	-	0.818	37.33	39.0
		12	$D_{2.}$ *NDVI. NDVIc. TC1. TC3*	-5.284	1.586***	-4.568**	3.761**	2.843*	-0.741*	-	0.836	36.41	38.2
	d	8	$H_{95.}$ $D_{2.}$ *NDVI. NDVIc. TC1. TC2*	-7.246	0.761**	0.629**	-2.616*	2.953**	3.063*	-0.980**	0.821	34.28	36.2
		12	$D_{2.}$ *NDVI. NDVIc. TC1. TC3*	-4.26	1.592***	-1.937*	2.207*	2.603*	-0.731*	-	0.835	36.31	38.18

Significance level: $\alpha< 0.1$. * $\alpha< 0.05$. ** $\alpha< 0.01$. *** $\alpha< 0.001$

The third part of this study was to use a stepwise multiple regression on these variables, which produced a combined model with the best possible fit. The results indicated high correlations among the variables derived from LiDAR data; these are the variables which best explain the real aerial biomass. By contrast, variables derived from satellite images contribute to the models in a small but still significant proportion. Ioki *et al.* (2009) also found low correlations using NDVI from a digital camera obtained at the same time the LiDAR data were obtained; they reported an R^2 of 0.073.

This shows the importance of LiDAR data to generate biomass models, since it is the most precise current technology to obtain three-dimensional information. For this reason the combination of multi-spectral elements may not be necessary in some cases; however, it is in the use of this information in which precision silviculture makes sense, since any reduction in the errors and increase in the goodness of fit, however small, becomes significant at a small scale. Also, compared to LiDAR, multi-spectral information is more economical, even free if LANDSAT is used to generate the variables; thus it should be considered and used beginning in the sampling stage.

Finally, the general model using all the variables obtained in this study found significant contributions of data from the LiDAR system and from transformation of the satellite images. However there was a high correlation (R = 0.98) between the data of height percentiles and point densities. Due to this, using a selection radius of 8 meters selected both types of variables, while using a selection radius of 12 meters the density variables were selected instead of the percentile variables. This indicates that using the point cloud to obtain variables related to aerial biomass is the best approach to the different levels of information that may be obtained, compared to the utilization of height percentiles to generate models with greater precision.

Conclusion

In conclusion, this study examined LiDAR data using three focuses: (1) height percentiles, which as in previous studies in other countries proved to be a rapid estimator of aerial biomass in pine plantations in Chile, providing simple and easy to obtain models. (2) Variables constructed from the point clouds, which also provide useful information since they include in the analysis other factors which also influence the biomass in addition to height. (3) Combinations of the models with multi-spectral variables, which although they did not improve the models statistically, should be used since they provide information on physiology and/or photosynthetic vigor which cannot be obtained with LiDAR data. The combination of these three focuses is a contribution to precision silviculture,

although the processes should be improved, other technologies should be introduced and new research areas should be projected.

Acknowledgements

This study was financed by the BIOCOMSA Consortium (CORFO-INNOVA) as part of their project "Evaluation and monitoring of existing Aerial Biomass and optimization of supply networks to biofuel plants. We also thank the Laboratorio de Geomática y Ecología del Paisaje (GEP) of the Universidad de Chile.

References

Anuchin, N.P. 1960. Forest Mensuration. 2nd edn. Goslesbumizdat. Moskova- Leningrado, 454 pp.

Avery, T. and H. Burkhart, 1994. Forest Measurements. Fourth edition. McGraw-Hill, 408 pp

Baskerville, G. 1972. Use of logarithmic regression in the estimation of plant biomass. Canadian Journal Forestry Research 2: 49-53.

Bitterlich, W. 1984. The Relascope Idea. Relative Measurements in Forestry. Common wealth Agricultural Bureaux, 241 pp

Caldentey, J. 1989. Beziehungen zwischen Klimaelementen und der Produktivitat von Pinus radiata Plantagen in Chile. in Dissertation zur Erlangung der Doktorwurde der Forestwissenschaftlichen Fakultat der Ludwig-Maximilians. Universitat Munchen, 131pp Caldentey, J. 1995. Acumulación de biomasa en rodales naturales de Nothofagus pumilio en Tierra del Fuego, Chile. Investigación Agraria: Sistemas y Recursos Forestales 4: 165-175.

Chávez, P.S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data, Remote Sensing of Environment 24: 459-479.

Chuvieco, E. 2010.Teledecteccion Ambiental. La observación de la tierra desde el espacio. 3rd edn. Ariel Ciencia, 590 p.

Civco, D.L. 1989. Topographic normalization of landsat thematic mapper digital imagery. Photogrammetric Engineering and Remote Sensing 55: 1303-1309.

CONAF (Corporación Nacional Forestal, CL) 2011. Catastro de los Recursos Vegetacionales Nativos de Chile. Monitoreo de Cambios y Actualizaciones: Período 1997–2011.

CONAF-CONAMA-BIRF. 1999. Catastro y evaluación de recursos vegetacionales nativos de Chile. In: Informe nacional con variables ambientales. Santiago, Chile, 90 pp.

Corvalán, P. and J. Hernández, 2011. Tablas de estimación de biomasa aérea bruta en pie para plantaciones de pino insigne en chile. Facultad de ciencias Forestales y de la Conservación de la Naturaleza. Departamento gestión forestal y su medio ambiente. 100 pp.

Crist, E. and R. Cicone, 1984. Applications of the tasseled cap concept to simulated Thematic Mapper Data. Photogrammetric Engineering of Remote Sensing 50: 343-352.

Estades, C. and S. Temple, 1999.Temperate-forest bird communities in a fragmented landscape dominated by exotic pine plantations. Ecological Applications 9: 573-585.

Falkowski, M., J. Evans, S. Martinuzzi, P. Gessler, and A. Hudak, 2009. Characterizing forest succession with LiDAR data: An evaluation for the Inland Northwest, USA. Remote Sensing of Environment 113: 946-956.

Forzieri, G., L. Guarnieri, E. Vivoni, F. Castelli, and F. Preti, 2009. Multiple attribute decision making for individual tree detection using high-resolution laser scanning. Forest Ecology and Management 11: 2501-2510.

Gajardo, R. 1994. La Vegetación Natural de Chile. Clasificación y Distribución Geográfica. Editorial Universitaria, 165 pp.

García, M., K. Prado, D. Riaño, E. Chuvieco, and F. Danson, 2009. Ajuste Planimétrico de datos LiDAR para la estimación de características en el Parque Natural del Alto Tajo. GeoFocus 9: 184-208.

Herrera, S .and R. Waisberg, 2002. Estimación de carbono almacenado en el Tipo Forestal Roble-Raulí–Coigüe (Nothofagus obliqua Nothofagus alpina–Nothofagus dombeyi), para determinar los beneficios ambientales de someterlo a sumidero. Memoria de título Ingeniería de Ejecución en ambiente. Departamento de Ingeniería Geográfica, Universidad de Santiago de Chile. 36 pp

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research 19: 543-553.

Hudak, A., M. Lefsky, W. Cohen, and M. Berterretche, 2002. Integration of LiDAR and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sensing of Environment 82: 397-416.

Hudak, A.T., N.L. Crookston, J.S. Evans, D.E. Hall, and M.J. Falkowski, 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112: 2232–2245.

Husch, B., C. Miller, and T. Beers, 1993. Forest Mensuration. Krieger Publishing Company, 3rd edn. Malabar, 402 p

Ioki, K., I.J. Imanish, T. Sasaki, Y. Morimoto, and K. Kitada, 2009. Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach. Landscape and Ecological. Engineering 6: 29-36.

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

Koch, B. and M. Dees, 2008. Forest Applications: LiDAR Data. In Advances in photogrammetry, remote sensing and spatial information sciences. Z. Li, J. Chen and E. Baltsavias.(eds) Taylor and Francis Group, London 439-465 pp.

Latifi, H., A. Nothdurft, and B. Koch, 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. Forestry 83: 395-407.

Lefsky, M., W. Cohen, and D. Harding, 2002. LiDAR Remote Sensing for Ecosystem Studies. Bio Science 52: 19-30.

Lim, K.S. and P.M. Treitz, 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. Scandinavian Journal of Forest Research 19: 558-570.

Lim, K., P. Treitz, K. Baldwin, I. Morrison, and J. Green, 2003.LiDAR remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29: 658–678.

Madgwick, H.A.I. 1994. Pinus radiata: Biomass, Form and Growth. Ed. H.A.I. Madgwick, Rotura, New Zealand. 428 pp.

Martinuzzi, S., L. Vierling, W. Gould, M. Falkowski, J. Evans, A. Hudak, and K. Vierling, 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sensing of Environment 113:, 2533-2546.

Morsdorf, F., B. Kotz, E. Meier, K.I. Itten,. and B. Allgower, 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sensing of Environment 104: 50-61.

Naesset, E. 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61: 246-253.

Naesset, E. 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment 79: 105-115.

Naesset, E. 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research 19: 164-179.

Nemani, R., L. Pierce, S. Running, and L. Band, 1993. Forest ecosystem processes at the watershed scale: Sensitivity to remotely sensed leaf area index estimates. International Journal of Remote Sensing 14: 2519 – 2534.

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

Popescu, S., R. Wynne, and J. Scrivani, 2004. Fusion of small-footprint LiDAR and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. Forest Science 50: 551–565.

R Core Team. 2014. R: A language and environment for statistical computing. http://www.r-project.org/ (accessed on 10 July, 2014).

Reitberger, J., C. Schnörr, P. Krzystek, and U. Stilla, 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 64: 561-574.

Riaño, D., F. Valladares, S. Condés, and E. Chuvieco, 2004. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. Agricultural and Forest Meteorology 124: 269–275.

Rouse, J.W., R.H. Haas, J.A Schell, and D.W. Deering. 1973. Monitoring vegetation system in the great plains with ERTS. In Third ERST Symposium, NASA SP- 351 12 pp.

Schlegel, B. 2001. Estimación de la biomasa y carbono en bosques del tipo forestal siempreverde. In Simposio internacional medición y monitoreo de la captura de carbono en ecosistemas forestales 18-20 Octubre 2001. Valdivia, Chile. 13 pp

Schmidt, A., M. Poulain, D. Klein, K. Krause, K. Peña-Rojas, H. Schmidt, and A. Schulte, 2009. Allometric above-belowground biomass equations for Nothofagus pumilio natural regeneration in the Chilean Patagonia. Annals of Forest Sciences 66: 315-323.

Sasaki, T., J. Imanishi, K. Ioki, Y. Morimoto, and K. Kitada, 2008. Estimation of leaf area index and canopy openness in broad-leaved forest using an airborne laser scanner in comparison with high-resolution near-infrared digital photography. Landscape and Ecological Engineering 4: 47-55.

Sherrill, K., Lefsky, M., Bradford, J. and Ryan, M. 2008. Forest structure estimation and pattern exploration from discrete-return LiDAR in subalpine forests of the central Rockies. *Canadian Journal of Forest Research* 38(8), 2081-2096.

Scaramuzza, P., E. Micijevic, and G. Chander, 2004. SLC gap-filled products, phase one methodology. *http://landsat.usgs.gov/data_products/slc_off_data_products/documents/SLC_Gap_Fill_Methodology.pdf* (accessed on 10 July, 2014).

Suratno, A., C. Seielstad, and L. Queen, 2009. Tree species identification in mixed coniferous forest using airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing 64: 683-693.

Weinacker, H., B. Koch, and R. Weinacker, 2004. TREESVIS: A software system for simultaneous ED-real-time visualisation of DTM, DSM, laser raw data, multispectral data, simple tree and building models. In Proceedings of the ISPRS Working Group VIII/2. Freiburg, 3–6 October 2004. 90–95 pp.

Zhao, K., S. Popescu, and R. Nelson, 2009. LiDAR remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment 113: 182-196.

Zheng, D., J. Rademacher, J. Chen, T. Crow, M. Bresee, J. Le Moine, and S.R. Ryu, 2004. Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. Remote Sensing of Environment 93: 402–411.