Research Article **Open Access**

Estimation of Above-Ground Forest Biomass Using Landsat ETM+, Aster GDEM and Lidar

Lissette Cortés*, Jaime Hernández, Diego Valencia and Patricio Corvalán

Faculty of Forestry and Nature Conservation, University of Chile, Chile

Abstract

The measurement of above-ground biomass is important to understand carbon flow between trees and the atmosphere; remote sensing plays an important role in making this possible for extensive and hard-to-reach areas. This study compared above-ground forest biomass estimation models using data from different sources, including Landsat ETM+, Aster GDEM, ALS (LiDAR) and forest inventories. Two sets of predictors were established: the first included variables extracted from Landsat ETM+ and Aster GDEM, while the second included variables from Landsat in combination with LiDAR products (Digital Terrain Model, Digital Surface Model and Canopy Height Model). The Random Forest algorithm was used to build all models; this method explicitly returns the importance of each predictor and therefore allows the selection of the best set of variables. Estimations were made separately by forest cover for *Pinus radiata*, *Eucalyptus globulus* and second-growth *Nothofagus glauca*. Better results were obtained using the combination Landsat-LiDAR than those using Landsat-Aster GDEM data. Also, the results were better when applying the model to pine cover (pseudo R2 77.22%).

Keywords: Above-ground biomass; Landsat ETM+; LiDAR; Random forest.

Introduction

The study of forests is a growing area of research due to their importance in different natural processes; the carbon cycle is one of the main processes. This is due to the fact that terrestrial ecosystems are the greatest contributors to the global carbon balance [1]; they contain close to 60% of the carbon stored in vegetation and 50% of that stored in the soil [2]. Carbon in forests may be divided into three components: biomass in living plants; plant remains and soil [3].

Most studies related to carbon content in forests are performed on the above-ground biomass, defined as the weight or equivalent estimation of organic material growing above the ground [4]; the different methods of estimating this content may be classified by the type of information they utilize.

The method most used to estimate above-ground forest biomass is that of forest inventories. These use allometric equations which relate easily-measured field data such as tree height and diameter at breast height (DBH) with above-ground biomass [5]. Although this is the most precise method it has greater cost, especially when it is desired to estimate biomass in extensive or hard-to-reach zones. Given the operational difficulties, methods have been developed which combine inventories with information acquired by remote sensing [6,7] and environmental variables (i.e. slope, exposition, altitude and ground curvature) [8].

Passive remote sensors use solar energy to obtain spectral information about objects. Among the most used are the medium-resolution multispectral satellites Landsat ETM+ and Aster (30 and 15 meters, respectively) [9]. The information provided by these sensors allows relating reflectance, vegetation indices and/or transformations of different spectral bands to above-ground biomass. One of the advantages of using these images is that some of them may be acquired free or at low cost. The same is true for the Digital Elevation Models (DEM), SRTM ("Shuttle Radar Topography Mission") and Aster GDEM (Aster "Global Digital Elevation Model"), from which it is possible to derive topographic information to be correlated with above-ground biomass [10,11].

LiDAR ("Light Detection and Ranging") is an active remote sensor that obtains a cloud of points with X,Y, and Z coordinates which provide information about the structure of objects on the ground [12]. Measurements are obtained by estimating the distance between the sensor and its objective by the time between the emission and return of a laser pulse [13]. Data of this type have been used by various authors [14-16] to obtain metrics or high-resolution models related to forest heights. One of the models most used is that of Digital Surface Models (DSM), which represent the surface of the Earth including all objects on it, and the Canopy Height Model (CHM), which estimates the height of the trees. One of the main difficulties and disadvantages of this type of data, compared to multi-spectral images, is the high cost of acquiring and processing them.

There are two major types of statistical models, those which focus on the data and those which focus on algorithms [17]. The focus based on the data begins by assuming a stochastic model such as linear or logistic regression, whose fit is used both to predict responses (Y) to future data (X) and to extract information about the nature of the relation between variables X and Y. The approach based on algorithms focuses on prediction and uses estimations of its exactitude to validate the models, without considering the type of relation between the variables. The latter approach is generally called machine-learning models.

In the construction of statistical models for the estimation of above-ground biomass both approaches have been used. Initially techniques

were employed such as multiple regression, k-NN classification and neural networks [18] on data from active and/or passive sensors. However, in recent years the use of machine-learning type algorithms based on classification and regression trees has increased. Among these, the Random Forest regression has been used to estimate forest parameters [10, 19-22], with better results than other methods.

About 22% of the surface area of continental Chile (16.6 million ha) is covered with native forests and forest plantations, concentrated mainly in the central part [23]. This makes the country an interesting place for biomass studies, to estimate the capacity for carbon storage, to implement management measures or to estimate the potential production of biofuel.

The objective of this study is to compare models of above-ground forest biomass estimation, combining satellite information obtained from the Landsat ETM+ sensor, Aster elevation data, information derived from LiDAR point clouds and field measurements.

Material and Methods

Test site

The study area is an experimental forest property of the Universidad de Chile named Pantanillos, located in central Chile (Figure 1), between coordinates UTM (WGS84) 744187 – 746042 meters east and 6074000 – 6070800 meters north.

Pantanillos has a surface area of approximately 400 ha, the majority of which is covered with plantations of *Pinus radiata* D. Don and in lesser quantity of *Eucalyptus globulus* Labill. It also has sectors of second-growth native forest, in which hualo (*Nothofagus glauca* Phil. Krasser) is the predominant species.

Field reference data

Field data were compiled in February, 2011, by establishing a systematic net of conglomerates every 200 m. Each conglomerate was formed by five sampling subunits, one central and four at 30 m distance in each of the cardinal directions. Each subunit consisted of concentric circular plots of 2, 4 and 8 m radius, which included trees with DBH ≥ 5, 10 and 20 cm, respectively (Figure 1). The following state variables were recorded in each plot: DBH (cm), species, total height (m), and other complementary descriptors. A total of 83 conglomerates were measured which encompass 330 plots (some plots were missing due to accessibility issues).

Estimation of the above-ground biomass with field data

The following allometric functions were used to estimate the above-ground biomass of individual trees:

E. globulus [24]:

$$BA = -92 + 2.3 * LN(DBH) \tag{1}$$

P. radiata [25]:

$$BA_F = e^{\left(1.028*LN\left[(DBH+1.9)^2*H\right]-4.892\right)} \tag{2}$$
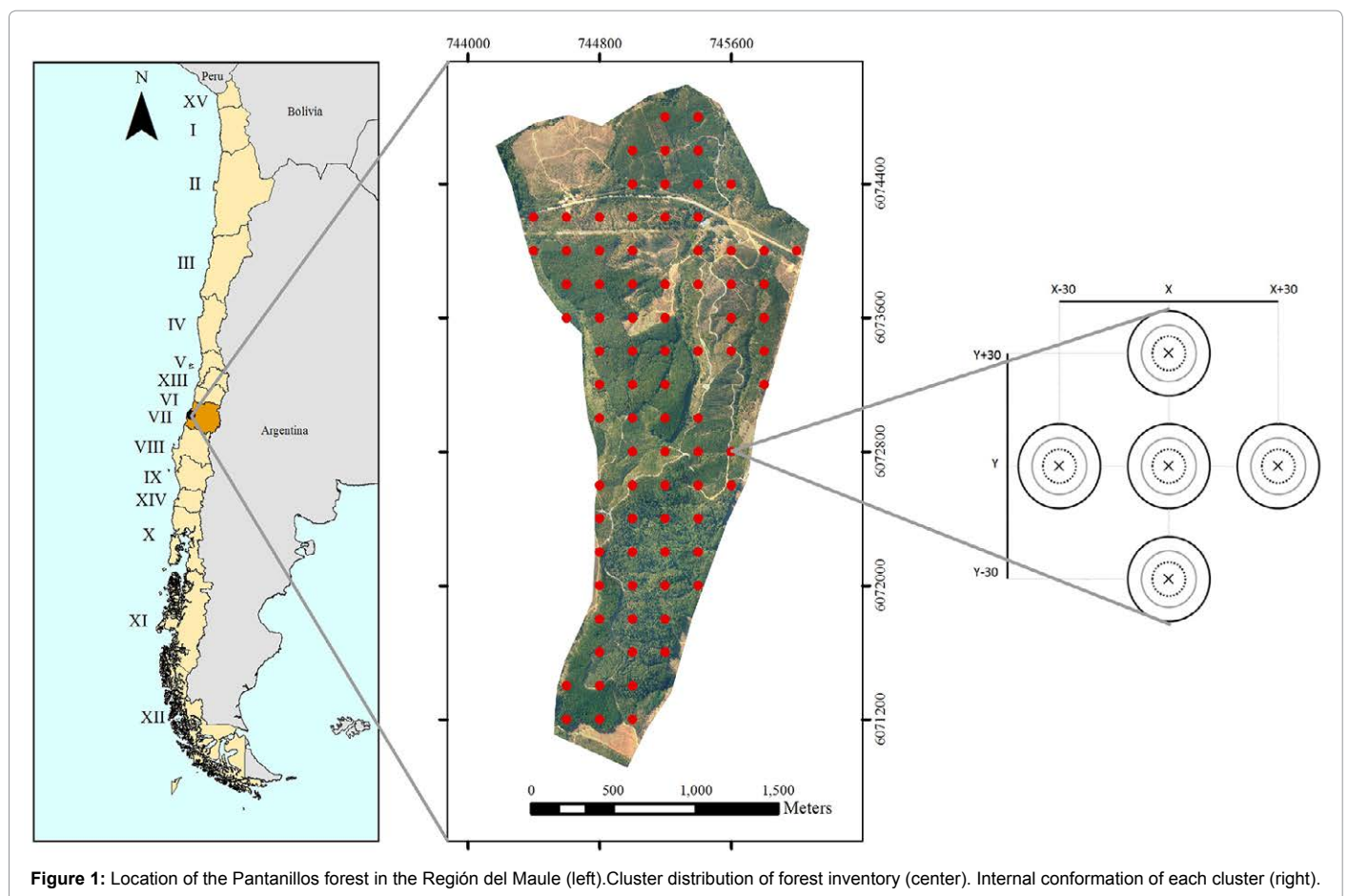
Other species [26]:

Figure 1: Location of the Pantanillos forest in the Región del Maule (left).Cluster distribution of forest inventory (center). Internal conformation of each cluster (right).

$$BA = e^{\left(-2.9149 + 1.9711*LN(DBH) + 0.7442*LN(H)\right)} \qquad (3)$$

BA is the total above-ground biomass (kg), BA_F is stem biomass (kg), DBH (cm) and H is height (m).

For radiata pine BA_F is the above-ground biomass of the stem with bark; thus it was necessary to add the above-ground biomass component of branches and needles (non-stem biomass). This was done using the component equations of Corvalán and Hernández [27] for the percentage of stem above-ground biomass (BF%) and that of the non-stem biomass (BNF%).

The final equation for the stem above-ground biomass (BF%) was:

$$BF_\% = 51.71 * e^{\left(-98/X^{0.78}\right)} + 33 \qquad (4)$$

$$X = E * \sqrt{1000/N} * AB \qquad (5)$$

Where X is the mean distance between trees, E is the age of the stand, AB is the basal area and N is the number of trees. Finally, BNF% was calculated as the complement of the stem biomass area, assuming a constant bark biomass of 10%.

$$BNF_\% = 100 - (BF_\% - 10\%) \qquad (6)$$

Total above-ground biomass per hectare in each plot was calculated by summing the individual stem biomass of each tree multiplied by its respective expansion factor, plus the non-stem biomass expressed as an aggregate function of total stem biomass divided by total biomass.

Processing of satellite images

We obtained two Landsat ETM+ scenes from 8 December 2010 and 9 January 2011 (Path/row 233/85). Both images were rectified by polynomial functions using the UTM projection (WGS84 zone 18 south).

A radiometric correction was performed by conversion of Digital Number (DN) to radiance using the specific gain and offset of the sensor. We then applied an atmospheric correction using the method of Dark Object Subtraction proposed by Chávez [28], and a topographic correction using the method of Civco [29]. Reflectance was obtained by applying a simplified model with the correction parameters of Chuvieco [30]. Since the images were taken after 2003, that is, they contain the SLC-off error, both images were fused using the method of Scaramuzza *et al.* [31] (Figure 2). The reflectance bands were also used to obtain the Tasseled Cap transformation [32] and the vegetation indices shown in Table 1.

Processing of the aster elevation digital model

We obtained the free Aster GDEM Elevation Digital Model from the web site of the same name, with a spatial resolution of approximately 30 meters. From this we obtained topographical variables of aspect (in azimuth degrees), slope (in percentage) and land curvature (concave, convex or flat).

Processing of LiDAR data

LiDAR data were obtained in March 2011 with a mean density of the point cloud of 5 points per m^2, and a true three-band ortho image (visible RGB) with a spatial resolution of 0.5 m.

As a first step to obtain the Digital Terrain Model (DTM) from the LiDAR, we used the multi-scale curvature algorithm to identify the points at ground level [39]. The points not classified as ground were interpolated to obtain a Digital Surface Model (DSM) [40]. We also obtained a Canopy Height Model (CHM) as the difference between DTM and DSM, which shows the tree heights. The three models (DTM, DSM and CHM) were obtained at a spatial resolution of 1 m. As with the Aster GDEM, from the LiDAR DTM we obtained the topographic variables elevation, slope, aspect and land curvature.

Parceling

Parceling consisted in the identification of homogeneous units according to the dominant vegetation present. This was done by photo interpretation of the true ortho image complemented with field data. We identified 5 types of vegetation cover: *P. radiata* plantations, *E. globulus* plantations, second-growth native forests, mixtures of forest covers and a class called other, which included bare areas, buildings and areas outside the forest property.

Estimation of above-ground biomass

We used the method based on classification and regression trees called Random Forest (RF), available in the statistical software R-Project [41], under the name random forest package. This algorithm, proposed by Breiman [42], constructs a tree of classification or
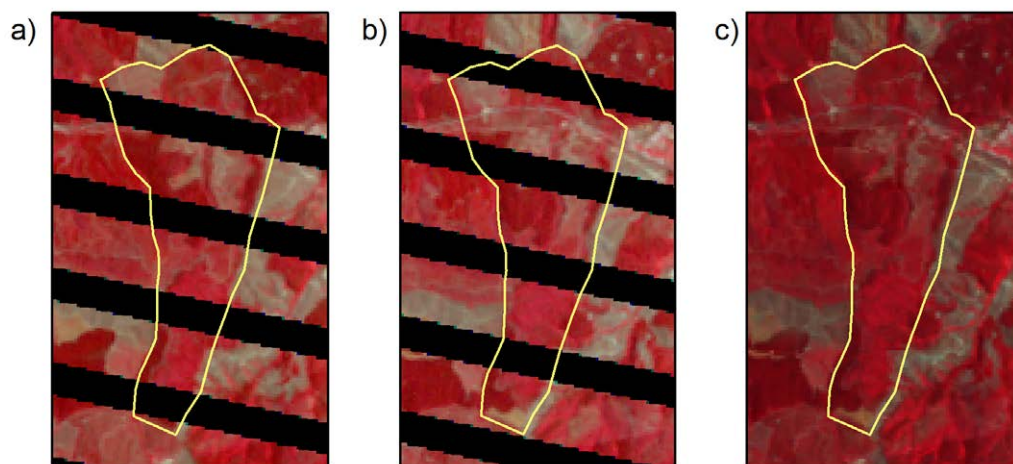

Figure 2: Correction of Landsat ETM+ SLC-off. a) Landsat ETM+SLC-off December/2010. b) Landsat ETM+SLC-off January/2011, c) Landsat ETM+ corrected.

Index	Equation	Reference
Difference Vegetation Index (DVI)	$DVI = \rho_{NIR} - \rho_R$	Roujean and Breon [33]
Simple Ratio (SR)	$SR = \dfrac{\rho_{NIR}}{\rho_R}$	Jordan [34]
Red Green Ratio (RGR)	$RGR = \dfrac{\rho_R}{\rho_G}$	Gamon and Surfus [35]
Red Green Index (RGI)	$RGI = \dfrac{\rho_G - \rho_R}{\rho_G + \rho_R}$	Coops et al. [36]
Normalized Difference Vegetation Index (NDVI)	$NDVI = \dfrac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}$	Rouse et al. [37]
Corrected Normalized Difference Vegetation Index (NDVIc)	$NDVI_c = NDVIv\left[1 - \dfrac{\rho_{MIR} - \rho_{MIR\,min}}{\rho_{MIR\,max} - \rho_{MIR\,min}}\right]$	Nemani et al. [38]

ρ_{NIR} = Near Infrared, ρ_R = Red, ρ_G = Green, ρ_{MIR} = Middle Infrared.

Table 1: Vegetation Indices utilized in the study.

regression using a process called bagging ("Bootstrap + Aggregation") [43,44]. In RF each node is separated using the best division chosen among a subset of randomly chosen predictor variables. This algorithm produces an internal error estimation, thus the results do not require cross-validation or independent validation [45]. The error of estimation is calculated by leaving out of the bootstrap sample about one third of the reference cases not used in the construction of the tree. These data are so-called "out-of-bag" (OOB), and are predicted using the tree created by the bootstrap sample. On the average each datum is an OOB about 1/3 of the time; these predictions are averaged and the mean square error (MSE) is calculated [46].

RF provides a pseudo R^2 similar to the R^2 of a standard linear regression; it is called pseudo because it uses the values predicted by RF to calculate the MSE. Pseudo R^2 may take negative values, which are interpreted as very poor fits [47]. RF also calculates the relevance of the predictor variables, assigning a score that depends on the changes in the error when the values of this variable are permuted (called %IncMSE); if the effect is larger the variable is assigned greater importance [48].

We estimated above-ground biomass for *P. radiata*, *E. globulus* and second-growth native forest separately. Since RF includes an element of randomness, each estimation was repeated 100 times. To evaluate the sources of information, the root mean square error (RMSE) (equation 7) and percentage RMSE (equation) were calculated from the MSE returned by RF.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(yi - \hat{yi}_{OOB}\right)^2} \qquad (7)$$

$$RMSE\% = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n}y_i}*100 \qquad (8)$$

Where the y_i are the observed data.

Variables were extracted based on a point cloud established every

30 m, in which were also included the coordinates of each point and the point above-ground biomass calculated with field data when available. For the variables obtained with LiDAR data the extraction was performed based on an area of influence of 8 m radius based on the net of points, obtaining the mean value for the area. In the extraction of DCM and DSM the standard deviation was also estimated.

To compare the capacity of Landsat, Aster GDEM and LiDAR to estimate above-ground biomass, two sets of predictors which occupy combined sources of information were used (Table 2).

Results

With the Landsat + GDEM set using all the predictors the best fit was obtained in pine plantations, 56.02%. However, for eucalyptus and second-growth native forest the pseudo R^2 were low, with values of -5.73% and 2.5%, respectively. The selection of variables improved the fit and decreased the error (Table 3). The estimations in *P. radiata* and second-growth hualo selected variables related to topography and the spectral response, while in *E. globulus* only three predictors derived from Landsat were selected (Figure 3a).

In estimations using Landsat and LiDAR the fit values increased for *P. radiata* (pseudo R^2=75.72%) and second-growth hualo (pseudo R^2=13.90%) compared to those obtained using Landsat + GDEM, while *E. globulus* continued with negative pseudo R^2 (pseudo R^2=-0.92%).

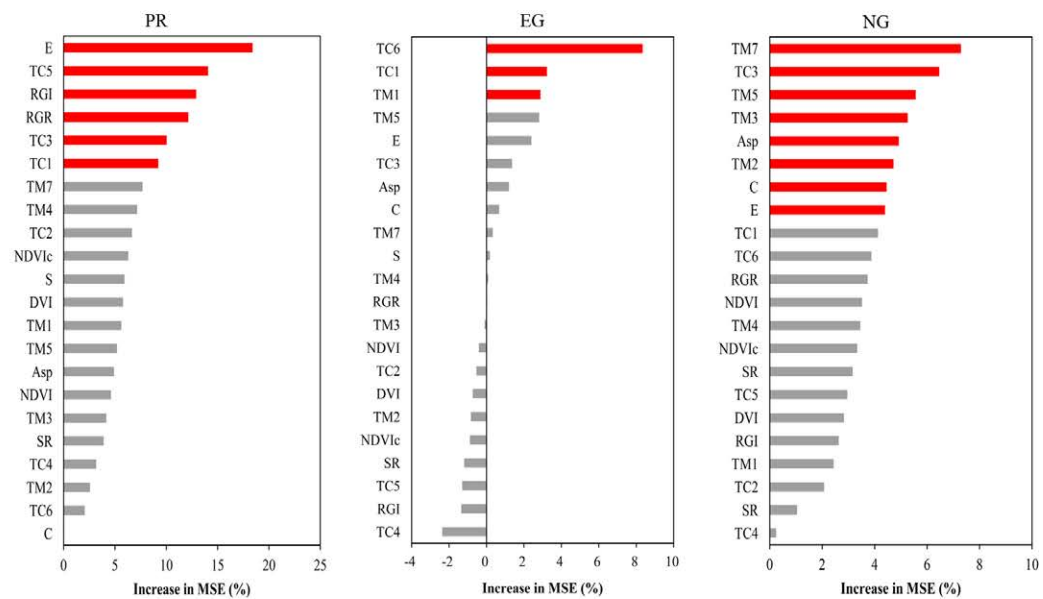| Landsat +GDEM Set	Landsat+LiDAR Set
Landsat: Reflectivity (TM1 - TM5 y TM7) Vegetation Index Tasseled Cap (TC1-TC6)	Landsat: Reflectivity (TM1 - TM5 y TM7) Vegetation Index Tasseled Cap (TC1-TC6)
GDEM: Elevation (E) Slope (S) Aspect (Asp) Curvature (C)	LiDAR: Elevation (Li_E) Slope (Li_S) Aspect (Li_Asp) Curvature (Li_C)
	DSM (Li_DSM) Standard deviation DSM (Li_DSM_sd) CHM (Li_CHM) Standard deviation CHM (Li_CHM_sd)

Table 2: Sets of predictors used in the estimation of aboveground biomass.

		PR	EG	NG
Landsat + GDEM	Pseudo R^2	60.26 (± 0.55)	37.84 (± 2.19)	7.76 (± 1.24)
	RMSE	44.58 (± 0.31)	25.81 (± 0.83)	55.2 (± 0.37)
	RMSE%	51.59 (± 0.32)	60.83 (± 1.95)	75.24 (± 0.51)
Landsat + LiDAR	Pseudo R^2	77.22 (± 0.37)	37.92 (± 1.78)	30.95 (± 1.05)
	RMSE	33.75 (± 0.27)	25.79 (± 0.37)	47.76 (± 0.37)
	RMSE%	35.40 (± 0.28)	60.79 (± 0.87)	65.09 (± 0.50)

PR= Pinus radiata, EG= Eucalyptus globulus and NG= Nothofagus glauca.

Table 3: Fits and mean errors for estimations with selection of variables for the two proposed sets of predictors.
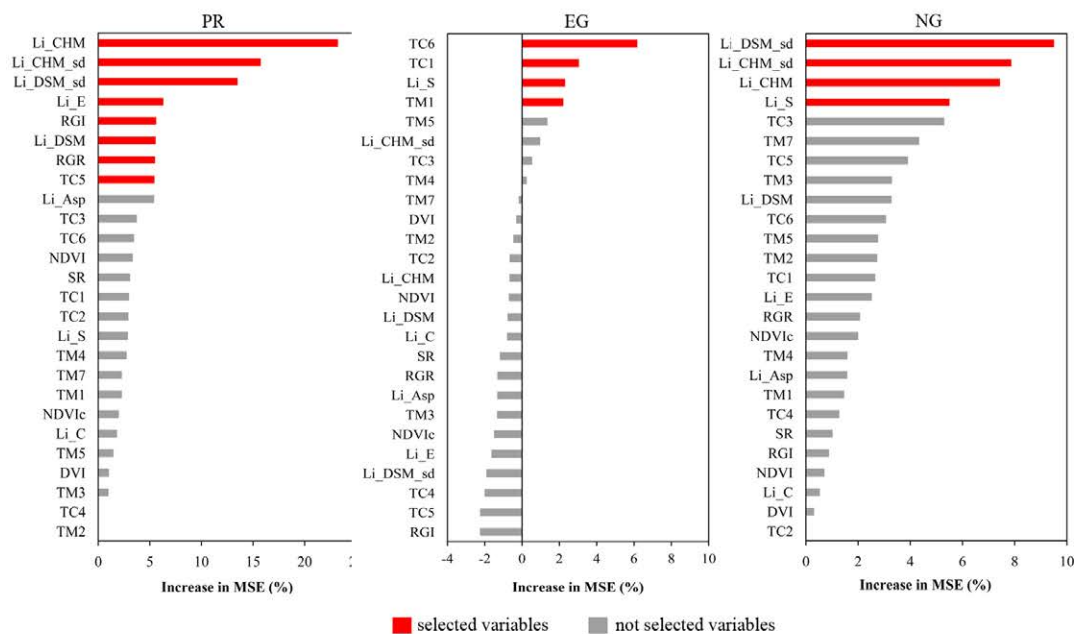

Figure 3: The graphs show the order of importance of the variables obtained in RF for the Landsat + GDEM and Landsat + LiDAR sets according to the dominant species. Red bars indicate the variables selected for the second estimation. PR= Pinus radiata, EG= Eucalyptus globulus and NG= Nothofagus glauca (native forest).

With selection of variables the fit in *P. radiata* was improved using variables of Landsat and LiDAR. For *E. globulus* a fit of only 37.92% was achieved (similar to that of Landsat + GDEM with selection of predictors) in which the variables of canopy structure were not relevant, in contrast to the estimations in second-growth hualo, in which only this class of variables was selected (Figure 3b).

Discussion

The set of predictors which used variables from Landsat and GDEM produced poorer fits than the second set, which incorporated information related to canopy height extracted from the LiDAR point cloud. These results concur with those of Lefsky *et al.* [49], who indicated that the metrics of canopy structure, mainly height and its standard deviation are highly correlated with the above-ground biomass of forests. Latifi *et al.* [19] also found this relation using a genetic algorithm for variable selection, although they also used other predictors extracted directly from the point cloud such as height percentiles.

In the estimations of pine and eucalyptus plantations and a second-growth native forest, both the fits and the type of predictors selected were considerably different for the two sets of predictors. The differences in fit among the three vegetation types may be due both to their horizontal and vertical structures. In pine plantations the horizontal structure is more homogeneous than in areas with native forest, which have not only a more complex and irregular canopy geometry but are also less dense. Clark *et al.* [50] found that plantations had higher correlations with metrics from LiDAR than forests with irregular structure. Our results agree with those of Popescu *et al.* [51], who obtained fits of 0.33 for deciduous forests (RMSE 44 ton/ha) and 0.82 for pine (RMSE 29 ton/ha), similar to our results.

Although eucalyptus plantations are homogeneous horizontally, the vertical structure of individual trees produced a large difference in the results. In pine plantations, the canopy structure allows a greater proportion of points to be intercepted in the canopy, which allows the creation of more precise models of the canopy surface, and thus also more precise DCM. On the contrary, in eucalyptus plantations trees are less leafy and canopies are sparser, thus many more points do not fall in the canopy of the trees but rather in the ground or the undergrowth, thus the DSM and DCM are less precise. LiDAR point clouds with higher densities may be suitable to compensate this problem but computing cost may also rise.

Conclusions

We conclude that the use of data which include information related to the canopy structure, such as the Digital Surface Models and Digital Canopy Models, produced better results (greater correlations) and lower errors. This was true in the estimations of pine and second-growth native forest but not for the estimations in *E. globulus*.

The differences in the fits of the estimations may be due to the precision obtained in the canopy models, given the vertical structure present in the different kinds of cover evaluated. In the estimations for *E. globulus* the utilization of LiDAR data did not improve the fit, which may indicate that the point density (5 points/m^2) is not dense enough to represent correctly the tree canopies, sub-estimating the height of the stand.

This study estimated biomass using digital models related to canopy height, which estimates the mean height of the vegetation in each pixel. However, it is also possible to use other variables extracted directly from the LiDAR cloud points which may reveal the structure of other strata (such as understory); using these variables could improve the fits estimations.

References

1. Gibbs H, Brown S, Niles JO, Foley JA (2007) Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environmental Research Letters 2: 045023.

2. Dong J, Kaufmann RK, Myneni RB, Compton JT, Kauppi PE, et al. (2003) Remote sensing estimates of boreal and temperate forest woody biomass carbon pools, sources and sinks. Remote Sensing of Environment 84: 393-410.

3. Bravo F, Bravo-Oviedo A, Diaz-Balteiro L (2008) Carbon sequestration in Spanish Mediterranean forests under two management alternatives: a modeling approach. European Journal of Forest Research 127: 225-234.

4. Schlegel B, Gayoso J, Guerra J (2000) Medición de la capacidad de captura de carbono en bosques de Chile y promoción en el mercado mundial. Manual de procedimientos: muestreos de biomasa forestal. Chile 24.

5. Fehrmann L, Kleinn C (2006) General considerations about the use of allometric equations for biomass estimation on the example of Norway spruce in central Europe. Forest Ecology and Management 236: 412-421.

6. Labrecque S, Fournier RA, Luther JE, Piercey D (2006) A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. Forest Ecology and Management 226: 129-144.

7. Ji L, Wylie BK, Nossov DR, Peterson B, Waldrop MP, et al. (2012) Estimating aboveground biomass in interior Alaska with Landsat data and field measurements. International Journal of Applied Earth Observation and Geoinformation 18: 451-461.

8. Lu M (2006) The potential and challenge of remote sensing-based biomass estimation. International Journal of Remote Sensing 27: 1297-1328.

9. Wulder MA, White JC, Fournier RA, Luther JE, Magnussen S (2008) Spatially explicit large area biomass estimation: three approaches using forest inventory and remotely sensed imagery in a GIS. Sensors 8: 529-560.

10. Pierce AD, Farris C, Taylor AH (2012) Use of random forests for modeling and mapping forest canopy fuels for fire behavior analysis in Lassen Volcanic National Park, California, USA. Forest Ecology and Management 279: 77-89.

11. Main-Knorn M, Moisen GG, Healey SP, Keeton WS, Freeman EA, et al. (2011) Evaluating the Remote Sensing and Inventory-Based Estimation of Biomass in the Western Carpathians. Remote Sensing 3: 1427-1446.

12. Harding D, Lefsky M, Parker G, Blair J (2001) Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. Remote Sensing of Environment 76: 283-297.

13. Lefsky MA, Cohen WB, Parker G, Harding DJ (2002) Lidar remote sensing ecosystem studies. Bioscencie 52: 19-30.

14. Hyde P, Dubayah R, Peterson B, Blair JB, Hofton M, et al. (2005) Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. Remote Sensing of Environment 96: 427-437.

15. Boudreau J, Nelson RF, Margolis HA, Beaudoin A, Guindon L, et al. (2008) Regional Aboveground Forest Biomass Using Airborne and Spaceborne LiDAR in Québec. Remote Sensing of Environment 112: 3876–3890.

16. Kellndorfer J, Walker W, Lapiont E, Kirsch K, Bishop J, et al. (2010) Statistical fusion of lidar, InSAR, and optical remote sensing data for forest stand height characterization: A regional-scale method based on LVIS, SRTM, Landsat ETM+, and ancillary data sets. Journal of Geophysical Research 115: 10.

17. Breiman L (2001) Statistical Modeling: The Two Cultures. Statistical Sciencie 16: 199-301.

18. Koch B (2010) Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. ISPRS Journal of Photogrammetry and Remote Sensing 65: 581-590.

19. Latifi H, Nothdurft A, Koch B (2010) Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: optimizing variable selection on optical/LiDAR-derived predictors. Forestry 83: 395-407.

20. Gleason CJ, Im J (2012) Forest biomass estimation from airborne LiDAR data using machine learning approaches. Remote Sensing of Environment 125: 80-91.

21. Immitzer M, Atzberger C, Koukal T (2012) Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. Remote Sensing 4: 2661-2693.

22. Leutner BF, Reineking B, Müller J, Bachmann M, Beierkuhnlein C, et al. (2012) Modelling Forest α-Diversity and Floristic Composition — On the Added Value of LiDAR plus Hyperspectral Remote Sensing. Remote Sensing 4: 2818-2845.

23. CONAF (2011) Catastro de los Recursos Vegetacionales Nativos de Chile. Monitoreo de Cambios y Actualizaciones.

24. Bennett LT, Weston CJ, Attiwill PM (1997) Biomass, nutrient content and growth responses to fertilisers of six-year-old Eucalyptus globules plantations at three contrasting sites in Gippsland, Victoria. Australian Journal of Botany 45:103-121.

25. Madgwick HAI (1994) Pinus radiata: Biomass, Form and Growth. Ed. HAI Madgwick, Rotura, New Zealand 428.

26. Gayoso J, Guerra J, Alarcón D (2002) Contenido de carbono y funciones de biomasa en especies nativas y exóticas. Documento Nº1. Proyecto FONDEF Medición de la capacidad de captura de carbono en bosques de Chile y promoción en el mercado mundial. Universidad Austral de Chile Instituto Forestal 53.

27. Corvalán P, Hernández J (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; Santiago, Chile 100.

28. Chávez PS (1988) An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sensing of Environment 24: 459-479.

29. Civco DL (1989) Topographic normalization of Landsat Thematic Mapper Digital Imagery. Photogrammetric Engineering and Remote Sensing 55: 1303-1309.

30. Chuvieco E (2010) Teledetección Ambiental. La observación de la tierra desde el espacio. 3ra Edición. Ariel Ciencia, España, Barcelona 590.

31. Scaramuzza P, Micijevic E, Chander G (2004) SLC gap-filled products, phase one methodology. Landsat Technical Notes.

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

33. Roujean JL, Bréon FM (1995) Estimating PAR absorbed by vegetation from Bidirectional Reflectance Measurements. Remote Sensing of Environment 51: 375-384.

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

35. Gamon JA, Surfus JS (1999) Assessing leaf pigment content and activity with a reflectometer. New Phytologist 143: 105-117.

36. Coops N, Johnson M, Wulder M, White J (2006) Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation. Remote Sensing of Environment 30: 69-76.

37. Rouse J, Haas R, Schell J, Deering D (1973) Monitoring vegetation system in the great plains with ERTS. Third ERST Symposium 1: 309-317.

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

39. Evans JS, Hudak AT (2007) A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. Geoscience and Remote Sensing 45: 1029-1038.

40. Isenburg M, Liu Y, Shewchuk J, Snoeyink J, Thirion T (2006) Generating Raster DEM from Mass Points via TIN Streaming. Geographic Information Science 186-198.

41. R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

42. Breiman L (2001) Random forests. Machine Learning 45: 5-32.

43. Bauer E, Kohavi R (1999) An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. Machine Learning 36: 105-139.

44. Dietterich TG (2000) An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. Machine Learning 40: 139-157.

45. Prasad A, Iverson L, Liaw A (2006) Newer Classification and Regression Tree Techniques: Baggins and Random Forests for Ecological Prediction. Ecosystems 9: 181-199.

46. Grömping U (2009) Variable Importance Assessment in Regression: Linear Regression versus Random Forest. The American Statistician 63: 308-319.

47. Liaw A, Wiener M (2002) Classification and Regression by random Forest. R News 2: 18-22.

48. Reif D, Motsinger A, McKinney B, Crowe J, Moore J (2006) Feature selection using a random forests classifier for the integrated analysis of multiple data types. IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology 1-8.

49. Lefsky MA, Cohen WB, Harding D, Parker G, Acker SA, et al. (2002) Lidar remote sensing of above-ground biomass in three biomes. Global Ecology & Biogeography 2: 393-399.

50. Clark ML, Clark DB, Roberts D (2004) Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. Remote Sensing of Environment 91: 68-89.

51. Popescu SC, Wynne RH, Scrivani JA (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.