Assessing understory development in forest plantations using laser imaging detection and ranging (LiDAR)

Evaluación del desarrollo del sotobosque en plantaciones forestales mediante LiDAR

JAIME HERNÁNDEZ1, M. PAZ ACUÑA1, PATRICIO CORVALÁN1 & JAVIER A. SIMONETTI2,*

1Laboratorio de Geomática y Ecología del Paisaje, Facultad de Ciencias Forestales y Conservación de la Naturaleza, Universidad de Chile, Casilla 9206, Santiago, Chile
2Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile
*Corresponding author: jsimonet@uchile.cl

ABSTRACT

Forestry plantations are expected to be managed in ways to conserve biodiversity while producing goods and services. This goal implies a significant challenge as plantations tend to reduce species richness. The presence of well developed understory enhances the value of plantations as habitat for native fauna. Here, we develop a straightforward method to assess the availability of understory in forestry stands using laser imaging detection and ranging (LiDAR) data and aerial RGB high resolution images. Based on field and airborne acquired data for *Pinus radiata* stands in central Chile, the digital crown model (DCM), derived from the subtraction of the digital terrain model (DTM) from the digital surface model (DSM) is a more reliable predictor of understory height that variables like terrain slope, aspect, plantation age and canopy height in forests and plantations which have not complete closed canopy. The correlation between DCM and understory though decreases while the actual height of the plantation canopy increases, rendering DCM a conservative estimate of understory development. The use of DCM will allow a fast and cost/effective estimate of habitat suitability in forestry plantations.

Key words: forestry plantations, kriging, LiDAR, *Pinus radiata*, understory height.

RESUMEN

Las plantaciones forestales deberían ser manejadas de forma que conserven biodiversidad al tiempo que provean bienes y servicios. Este es un desafío significativo pues las plantaciones tienden a reducir la riqueza de especies nativas. La presencia de un sotobosque desarrollado incrementa el valor de las plantaciones como hábitat para la fauna nativa. En este trabajo desarrollamos un método sencillo para evaluar la disponibilidad de sotobosque en plantaciones forestales empleando imágenes LiDAR y RGB de alta resolución. En base a datos de campo, LiDAR e imágenes aéreas para rodales de *Pinus radiata* en Chile central, el modelo digital de copa (DCM), obtenido de sustraer el modelo digital de terreno (DTM) del modelo digital de superficie (DSM) es un predictor más confiable del desarrollo del sotobosque que variables como la pendiente del terreno, la exposición, la edad de la plantación y la altura del dosel de la plantación en situaciones en las cuales en dosel superior no está completamente cerrado. La correlación entre DCM y el sotobosque sin embargo decrece con la altura del dosel de la plantación, lo que hace de DCM un estimador conservador del desarrollo del sotobosque. El uso de DCM permitirá una evaluación rápida y costo/efectiva de la disponibilidad de hábitat para fauna nativa en plantaciones forestales.

Palabras clave: interpolación, LiDAR, *Pinus radiata*, plantaciones forestales, sotobosque.

INTRODUCTION

Forestry plantations covers over 265 million ha worldwide, growing 2 % yearly (FAO 2011). Landscape transformation brought about by these plantations has impinged upon biological diversity, particularly planted forests based on introduced species (Carnus et al. 2006). Indeed, the Convention of Biological Diversity (CDB) has set as a decennial goal to enhance the role of forestry plantations in the conservation of biodiversity (CBD 2010). Increasing evidence suggests that structurally complex plantations, with a well developed understory could be a way forward to achieve this goal. In fact, species richness and abundance tend to be higher in planted
forests with well developed understory than those devoid of such vegetation, therefore minimizing the reduction of species richness of soil insects, herpetozoans, mammals and birds among other taxa (see Simonetti et al. 2012 for a review).

In this framework, to accurately and cost/effectively assess the presence of understory vegetation in planted forests is mandatory. Precise assessment of understory coverage including its spatial heterogeneity enables quantification and characterization of the amount of suitable habitat that planted forests could provide, contributing to mitigate the impacts of landscape transformation.

Chilean conifer plantations are a case in point. Currently, 1.5 million ha are a monoculture of *Pinus radiata* D. Don. Stands with well developed understory hold more species of insects, mammals and birds than those with scarce or no undergrowth. Among species using plantations, at least temporarily, are endangered species like *Leopardus guigna* (Molina), a fact that contributes to achieve CBD’ goals (Simonetti 2006, Estades et al. 2012). Therefore, assessing how much of the area covered by plantations could be suitable as surrogate habitat for native fauna would provide solid managerial information to properly engage the forestry industry in biological conservation. Here we develop a method to assess understory development using laser imaging detection and ranging (LiDAR) data and aerial RGB high resolution images (see Baltsavias 1999, Wher & Lohr 1999 and Elmqvist et al. 2001 for details regarding the use of these techniques).

LiDAR data, in conjunction with various sources of forest inventories data, can be successfully used to quantify different aspects of forests 3-D structure, such as volume and biomass estimates and canopy height profiles, among other features (Koch & Dees 2008). Hence, small-footprint LiDAR may be used to quantify components of forest structure needed to assess animal-habitat relationships (see review by Vierling et al. 2008). In fact, airborne laser scanning data can be used to predict habitat quality and to map species distributions as a function of habitat structure (Bradbury et al. 2005, Goetz et al. 2007). Regarding the assessment of understory vegetation, LiDAR data can be utilized for to derive both a variety of environmental factors which account for the presence of understory vegetation, including canopy structure, forest successional stage and topography (e.g., Falkowski et al. 2009) and, to characterize height and cover of the understory vegetation either as woody vegetation or suppressed trees integrating only leaf-on and leaf-off LiDAR (Hill & Broughton 2009).

One critical step in using point clouds to derive digital terrain and surface models (DTM and DSM, respectively) is the identification and classification of ground returns, which impinges upon the quality and precision of any LiDAR product. Errors in the estimation of the ground level will affect the subsequent estimation of an object above it, including all types of vegetation. Many algorithms have been developed to generate DTMs using LiDAR data. These algorithms can be classified into two categories based on type of data used: (a) point clouds and (b) raster range image. A DTM can be constructed from the ground points, and a digital surface model (DSM) can be derived from the highest points within a defined grid box (Hollaus et al. 2006). An object height model can be calculated then by subtracting DTM from the DSM. In a forest environment, without human built structures, all objects are assumed vegetation given a digital crown model (DCM). In this study, we deal with the use of DTM and DSM to estimate understory height in forestry plantations. Our aim is to develop a simple step algorithm for processing these types of data in order to advance a simple novel way to estimate understory development, moving from comparisons of observed versus estimated values at plot level to the assessment of observed versus estimated understory heights surfaces, which we obtained applying geo-statistical interpolation on field and LiDAR data, respectively. This procedure will enable the assessment of the availability of suitable habitat for native fauna in forestry plantations, which could be used temporarily used for the native fauna, contributing to the fulfillment of the CBD targets.

**METHODS**

*Study area*

The study area is Pantanillos Forest Research Station of the University of Chile (WGS84 UTM 18s: 358260 S, 728170 W, Fig 1), in the Maule region, coastal range of south-central Chile. The area was originally covered
by *Nothofagus*-dominated forests (Maulino forest), composed mainly by deciduous *Nothofagus glauca* (Phil.) Krasser and *Nothofagus obliqua* (Mirb.) Oerst., and evergreen species such as *Nothofagus dombeyi* (Mirb.) Oerst., *Cryptocarya alba* (Mol.) Looser, *Persea lingue* Nees, and *Aristotelia chilensis* (Mol.) Stuntz, among others (San Martín & Donoso 1996). Currently, this type of forest is highly fragmented, immersed in a landscape dominated by exotic Monterey pine (*P. radiata*) plantations (Grez et al. 1998, Echeverría et al. 2006). At the Forest Research Station there are plantations of different age whose understory is well developed because each plantation receives two to three thinning and prunings during the 15 to 20 years rotation period allowing light penetration and plant development (Estades & Escobar 2005). This understory is composed mainly by shrubs (e.g. *Azara integrifolia* Ruiz et Pavón, *Escallonia pulverulenta* (Ruiz et Pav.) Pers., and small trees such as *N. glauca*, *C. alba*, *Lithraea caustica* Hook. et Arn., and *Luma apiculata* (DC) Burret (Estades & Escobar 2005).

**Field data collection and processing**

Field data was gathered through systematic sampling of a grid of clusters spaced 200 m x 200 m over all types of vegetation. Each cluster contains five subplots, one central and four neighboring ones 30 m apart, placed in the four geographic directions (N, S, W and E; Fig. 1). The coordinates of the plot center were located using a Garmin double frequency GPS Topcon GRS-1 (WGS84 UTM18s). We measured DBH, height, and crown ratio of all trees above 5, 10 and 20 cm of DBH in concentric rings of 2, 4 and 8 m in every plot. Understory was assessed only at the 4 m radii ring, recording species, average height, mean coverage (%) and plant density. A total of 98 clusters, comprising 490 plots were sampled. Using this data, a raster was obtained by applying ordinary kriging, representing understory height (H), as an estimate of its development.

**LiDAR and image data acquisition**

We decided to use LiDAR because the management of pine plantations, including pruning and thinning, determines that the canopy is not completely closed, which in turn allows both enough light to reach the ground understory development and returns of LiDAR pulses.

Discrete waveform laser scan data were acquired in March 31, 2011 by Digimapas Chile Ltd. (http://www.digimapas.cl) using a set Harrier 56/G4 Dual System mounted on a Piper PA-24 Comanche. A digital image acquisition sensor (VIS) and LiDAR scanning were set together and the flight was conducted with a nominal height over the ground of 580 m AGL without GPS errors, with an average speed between 180 to 210 km h\(^{-1}\) and leading to an average point density of 4.64 points per square meter (see Table 1 for flight and system details). Additionally, we acquired an aerial image (VIS), with the three channels (red, green and blue) in the visible range of the electromagnetic spectrum, with 1 m spatial resolution, obtained at the same time of LiDAR data and therefore it was truly orthorectified.

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

Área de estudio en la Región del Maule, centro-sur de Chile. Se indica la posición de los grupos de muestreo. La imagen de la derecha representa la configuración interna de cada unidad de muestreo, y muestra la posición relativa de cada subunidad de muestreo.
LiDAR and image data processing

To process LiDAR point clouds into XYZ Ascii format, we used LasTools (Isenburg & Shewchuk 2010). First, we convert LiDAR data into *.las format and then we used the multiscale curvature algorithm (mcc-lidar) to classify ground points only (Evans & Hudak 2007). We used 1.5 and 0.3 as scale and threshold parameters. The optimal scale parameter depends on, first, the scale of the objects on the ground (i.e. rocks, trees) and second, on the sampling interval (post spacing) of the LiDAR data. LiDAR sensors are capable of collecting high density data (e.g., 8-10 pulses m$^{-2}$) which translate to 0.35 m pulse. This is $1/\sqrt{8}$ pulses m$^{-2}$ = 0.35 m pulse). Sparser LiDAR data (e.g., 0.25 pulses m$^{-2}$) translate to a spacing of 2.0 m pulse ($1/\sqrt{0.25}$ pulses m$^{-2}$ = 2.0 m pulse). Evans & Hudak (2007) recommended 0.3 as a good starting value to try for the curvature threshold in forest landscapes. Once the ground points were classified, we used the interpolation to raster LasTools (las2dem.exe) to obtain a DTM, using only ground points, and a DSM using all points. Further processing, subtracting DTM from DSM allow us to obtain a DCM. Complementary, the VIS image was used to identify and classify all stands in the study area into five stand types: pine plantations (53 %), eucalyptus plantations (6 %), native forest (29 %), mixed stands (6 %), and other types (6 %). A polygon shape file was finally obtained for all former types.

The DCM was sliced into eight height stratum as follows (in meters): 0-0.5, 0.5-1.0, 1.0-2.0, 2.0-4.0, 4.0-8.0, 8.0-16.0 and 16.0-32.0. Each one was then transformed into a raster containing only pixels within the correspondent height class. These pixels were separated into two sets, 60 % for training and performing ordinary kriging and 40 % for validation. The interpolated rasters were named H050, H100, H200, H400, H800, H1600 and H3200 to reflect the height stratum they represented. Slope (SLOPE) and aspect (ASPECT) were directly calculated from the DTM data. Finally, the year of plantation (YEAR) was added as an additional variable to complete a final set of 11 predictors. These variables are the most frequently used to model forest and plantation attributes (see Hernández et al. 2007).

Modeling

To be able to model the relations between all preprocessed predicted variables and the target variable, understory height (H), we sampled all rasters using 2000 plots of 4 m radii, randomly located over the study area. This way we obtained a master table that contains 2000 records and the 11 predictors plus understory height (H). Using this dataset we analyzed correlation among variables and fitted multiple linear models.

RESULTS

Interpolation of field and LiDAR data

The interpolation of heights using LiDAR data available for the seven strata height classes generates not only prediction surfaces but also errors, i.e. uncertainty surfaces, giving an indication of how good the predictions are. The error for all height classes has low values, near to zero, and also constrained standard deviation except for higher stratum where their values increase. The same behavior was observed for the kriging standard error, suggesting estimates are accurate (Table 2).

Correlations between understory height and predictors

The predictor DCM had the strongest correlation to understory height, exhibiting values that decrease as stratum height increase (Table 3). There is also a significant ($\alpha = 0.05$) positive Pearson's product moment correlation coefficient (> 0.3) between the overall stands (All) and predictors DCM, SLOPE, H400, H1600 and H3200. The same tendency was observed for stand of 0-0.5 m dominant height and DCM, and for stand of 8.0-16.0 m and 4.0-8.0 m dominant height.

Linear model

Figure 2, bottom left and right, shows the relation between the residuals of estimates and slope. There is evidence of an additive effect on residues when high slopes and high heights
of plantations are combined that can be seen in the middle zone of the study area. On the contrary, in the bigger native forest stand, in the south-east area, there are also high slopes but the behavior of residues are the opposite, concentrating low values, probably because of the open canopy structure of this vegetation. Aspect was codified into eight classes following a clockwise fashion from 1 (N) to 8 (NW) and therefore, the overall negative correlation is affected by this codification. An interested behavior can be seen for the year of plantation variable (YEAR), which tends to increase its negative correlation with H as the age of the plantation increases. When the dominant canopy develops into denser structure, the understory exhibit less development.

Table 4 illustrates the linear model summary for the general form $H = f$ (predictors). The sampling points were selected for each model according to their dominant height stand, from DCM, as indicated, fitting models for each one. Most predictors showed strong linear relation according to the stand dominant height under consideration. When the actual dominant height is low, i.e. less than 2 m, height stratum predictors have low explanation power. The overall model, including all ten predictors gave the best results. Strongly significant linear relation was also found at 99%, for model 0 -
0.5 in DCM predictor, and for model 8.0 - 16.0 in SLOPE predictor. The associated residues showed a slightly positive bias that make estimates to have a tendency to produce lower values than observed ones (Fig. 3).

Stratum height variables, from H050 to H3200, did not provide strong correlation in any of the stands. This fact could be related to an inadequate selection of height classes, which separate height values in arbitrary classes. Nevertheless, when using them as predictors in linear models, most of them showed significant statistical relation to understory height (Table 4). All models exhibit strong lineal relation for a subset of predictors but, different ones, according to the dominant height of the stand under consideration. When the actual dominant heights are low, i.e. less than 2 m, the height stratum predictors have very little explanation power, which is probably related to the absence of either understory or plantation species development. Both components tend to be

Fig. 2: Interpolated rasters of both observed (top-left) and predicted (top-right) understory average heights. Bottom left: Map of standard residues of overall linear model. Bottom right: map of slope classes.

Cuadriculados interpolados de la altura media del sotoboques observada (arriba-izquierda) y predicha (arriba-derecha). Abajo-izquierda: mapa de los residuos estandarizados de modelo lineal general. Abajo-derecha: mapa de clases de pendientes.
Summary of linear models for the general form $H = f$ (predictors). Every column includes a different set of predictors as height strata varies. Significant correlation at $\alpha = 0.001$ are denoted by ‘***’, $\alpha = 0.01$ by ‘**’, $\alpha = 0.05$ by ‘*’ and $\alpha = 0.1$ by ‘.’; - indicates it was not included in the indicated model.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Statistical significance for predictors used in lineal multiple models.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All 0 - 0.5 0.5 - 1.0 1.0 - 2.0 2.0 - 4.0 4.0 - 8.0 8.0 - 16.0 16.0 - 32.0</td>
</tr>
<tr>
<td>Intercept</td>
<td>. ** . * . ** . **</td>
</tr>
<tr>
<td>H050</td>
<td>* . . * .</td>
</tr>
<tr>
<td>H100</td>
<td>- - - -</td>
</tr>
<tr>
<td>H200</td>
<td>** - - .</td>
</tr>
<tr>
<td>H400</td>
<td>*** - - .</td>
</tr>
<tr>
<td>H800</td>
<td>*** - - -</td>
</tr>
<tr>
<td>H1600</td>
<td>** - - -</td>
</tr>
<tr>
<td>H3200</td>
<td>- - - -</td>
</tr>
<tr>
<td>SLOPE</td>
<td>*** * . * ** . **</td>
</tr>
<tr>
<td>ASPECT</td>
<td>*** * . ** . **</td>
</tr>
<tr>
<td>DCM</td>
<td>*** *** .</td>
</tr>
<tr>
<td>YEAR</td>
<td>** * . . .</td>
</tr>
</tbody>
</table>

Parameter of the models

- Residual std. error: 0.3457 0.2469 0.295 0.3186 0.2427 0.3591 0.4157 0.407
- Degrees of freedom: 886 177 110 136 83 62 196 71
- Multiple R-square: 0.3048 0.1503 0.1074 0.1915 0.1451 0.2535 0.2726 0.3235
- p-value: 2.2E-16 2.2E-05 0.05 0.0001 0.09 0.03 7.41E-10 0.002

Fig. 3: Left: Histogram of residuals from the model that uses all sampling units and all ten predictors. Right: Scatter plot of observed understory heights ($H$) against predicted ones (fitted).

Izquierda: histograma de los residuos del modelo que emplea todas las unidades muestrales y los diez predictors. Derecha: diagrama de dispersión de la altura observada del sotobosque ($H$) frente a los valores predictivos (f).
mixed up and their discrimination is difficult to achieve. The overall model, including all predictors gave the best results, and can be use in understory mapping.

**DISCUSSION**

The Convention on Biological Diversity has established that by 2020, forestry plantations should be managed in ways to ensure the conservation of biodiversity. This goal aims to reduce pressures upon biodiversity, minimizing the effects of forestry practices upon it (CBD2010). To achieve it, forestry plantations should be a suitable surrogate habitat for native species, and the presence of a well developed understory is a key factor. Therefore, forestry plantations holding such undergrowth could provide both goods and services while contributing to conserve a suite of native species (Simonetti et al. 2012). To foster wildlife-friendly practices in otherwise “productive lands” (Fisher et al. 2008), an assessment of the understory development is required to evaluate the potential of forestry stands as surrogate habitat for fauna as a necessary step to integrate these areas devoted to forestry production into wildlife conservation programs as well. The use of LiDAR data in this assessment is feasible, albeit a conservative approach. The height from DCM had the strongest correlation to understory heights, but such correlation decreases while the actual height of the dominant canopy, i.e. plantations, increases. This fact could be accounted for by the number of returns available in each vegetation layer, which gets lower when upper canopy layer gets higher. This effect is stronger as plantation density increases (Lefsky et al. 2002).

The predicted understory height map provided reliable information for ecological applications at a very high spatial resolution and it is comparable to similar studies (Turner et al. 2003, Goetz et al. 2007, Vierling et al. 2008, Bergen et al. 2009). Hotspots of higher values of understory height can be detected and its continuity is captured in the predicted map (Fig. 2 top-right) as can be seen when draw against observed understory height map (Fig. 2 top-left). It is important to consider that estimation were done at one square meter cells and, although simple linear modeling showed low r-square values, the residuals are controlled within the range of 0.24 and 0.41 meters, allowing the use of this type of understory maps in further studies. However, the estimates exhibit some bias between 0.2 and 0.5 m for high understory heights when it has values near or above 2.0 m, tending to underestimate their values. This effect can be related to the original point cloud acquisition and the vegetation architecture (Lefsky et al. 2002). The gaps in the dominant canopy allow for some of the LiDAR pulses to reach the lower vegetation which has also a low probability to return from the top of it, generating lower heights. As these points can mainly be gather in the gaps, no high angles can be used as they the pulses need to be as vertical as possible to actually receive any returns, which also reduce the possibilities of top understory points to be form in the final cloud. In any event, we can hypothesize that the point cloud density, from underneath the main canopy layer in a given small area, should be related to the understory coverage and density, and we can only have good samples of this structure in the gaps. On another hand, the negative relation between canopy density and understory development has been well documented for managed native forest and plantations (Castedo-Dorado et al. 2012, Shatford et al. 2009, Ogden et al. 1997). This relation states that the lower density in trees the bigger the understory development, which means higher heights. Putting everything together, it would be possible to estimate understory density and coverage by using the density of return points in the point cloud or, at least, in DCMs.

One of the advantages of our approach relies on the fact that it requires a minimum of field information, reducing associated time and operational costs. Furthermore, it may be applicable to any type of forests, either native or plantations. This means that it is possible to built understory maps at the landscape level in order to feed the large scales required to properly plan spatially explicit conservation activities (e.g., Lindenmayer et al. 2006). Another advantage is availability of large LiDAR datasets, increasingly more available for such applications, especially in the central coastal range (e.g., Cartus et al. 2012). In this region, land use change and fragmentations processes are very dynamic (Echeverría et al.
and such maps can play an important role in biodiversity conservation planning, potentially incorporation of plantations to ensure both habitat provisioning and landscape-level connectivity for the fauna. The simple model advanced here will then allow the rapid and conservative estimate of how much of the 1.5 million ha currently allocated to Pinus could be incorporated into wildlife conservation in Chile as well as abroad, contributing to render the forest industry a more sustainable one.

ACKNOWLEDGMENTS: This work has been partially supported by Fondecyt 195046, Iniciativa Transversal 3, Programa Domeyko-Biodiversidad, Universidad de Chile. We would also like to express our thanks to Biocomsa Consortium, and therefore INNOVA CORFO CHILE, for funding part of the field work and LiDAR data acquisition. Thanks are due to A. Camaño (Gerencia de Medio Ambiente, Arauco) for critical comments at the onset of this research.

LITERATURE CITED


BRADBURY RB, RA HILL, DC MASON, SA HINSLEY, JD WILSON & H BALTZER (2005) Modeling relationships between birds and vegetation structure using airborne LiDAR data: a review with case studies from agricultural and woodland environments. Ibis 147: 443-452.


LINDENMAYER DB, JP FRANKLIN & J FISHER (2006) General management principles and a checklist of


Editorial responsibility: Lohengrin A. Cavieres
Received January 18, 2013; accepted August 13, 2013