

Análisis temporal de la estructura forestal mediante métricas derivadas de LiDAR *full-waveform*

Pablo Crespo-Peremarch^(1,2), Luis Ángel Ruiz^(1,2), Ángel Balaguer^(1,3), Javier Estornell^(1,2)

- ⁽¹⁾ Grupo de Cartografía GeoAmbiental y Teledetección (CGAT), Universitat Politècnica de València, Camí de Vera s/n 46022-Valencia, España.
⁽²⁾ Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría, Universitat Politècnica de València, Camí de Vera s/n 46022-Valencia, España. Emails: pabcrepe@cgf.upv.es, laruiz@cgf.upv.es, jaescr@cgf.upv.es
⁽³⁾ Departamento de Matemática Aplicada, Universitat Politècnica de València, Camí de Vera s/n 46022-Valencia, España. Email: abalague@mat.upv.es

Resumen: Los sistemas LiDAR *full-waveform* permiten registrar la onda completa proveniente de la interacción entre la vegetación y el pulso de energía enviado por el sistema, obteniendo información sobre las propiedades físicas de la vegetación con un nivel de detalle mayor que la obtenida mediante LiDAR discreto. En este artículo se estudia la evolución de variables de estructura forestal (biomasa, volumen, altura, etc.) a partir de métricas extraídas de LiDAR *full-waveform* de dos fechas diferentes (2010 y 2012). Para ello, se realizó el preprocesado de los datos, incluyendo la generación de nuevas ondas pseudovericales para evitar el error posicional debido al uso de ángulos de escaneo que se encuentran fuera del nadir y que generan ondas no verticales. Ello implicó la voxelización de todos los valores extraídos, asignando el valor máximo de amplitud a cada vóxel. Se calcularon diferentes métricas *full-waveform* directa o indirectamente relacionadas con variables de estructura de la vegetación. A partir de ellas y junto con los datos tomados en campo (2010), se generaron y validaron varios modelos de variables de estructura forestal. Estos modelos obtenidos para 2010 se aplicaron también a los datos de 2012, lo que permitió comparar y cartografiar las estimaciones obtenidas para ambas fechas y analizar la evolución de la masa forestal.

Palabras clave: LiDAR *full-waveform*, voxelización, modelos de regresión, estructura forestal

Time analysis of forest structure using metrics extracted from full-waveform LiDAR

Abstract: Full-waveform LiDAR systems record the full wave coming from the interaction between the canopy and the energy pulse sent by the system, giving more detailed information about the physical properties of the canopy compared to discrete LiDAR systems. This paper studies the evolution of forest structure variables (biomass, volume, height, etc.) using metrics extracted from full-waveform LiDAR on two different dates (2010 and 2012). This required the pre-processing of the data, as well as the creation of additional pseudovertical waves to avoid positional error due to the use of off-nadir scanning angles that produce non-vertical waves. This involved the voxelisation of all the extracted values, with the maximum amplitude value assigned to each voxel. Different full-waveform metrics were calculated that were directly or indirectly related to canopy structure variables. Using these variables and field data from 2010, several models of forest structure variables were generated and validated. The models calculated from 2010 data were then applied to 2012 data. The datasets obtained on the two dates were compared and mapped so that the evolution of forest mass could be analyzed.

Keywords: LiDAR *full-waveform*, voxelisation, regression models, forest structure

1. INTRODUCCIÓN

La teledetección ha demostrado su importancia para la caracterización de la estructura vegetal en bosques poco densos, siendo el mayor desafío aquellos con una densidad media-alta. Disponer de un inventario de variables de estructura como la biomasa es de utilidad para determinar el balance de carbono (Lefsky *et al.*, 1999).

Algunos atributos forestales pueden ser cuantificados directamente a partir de datos LiDAR (*Light Detection and Ranging*), mientras que otros pueden ser estimados mediante modelos empíricos obtenidos a partir de estos datos, como la biomasa o el área basimétrica (Lim *et al.*, 2003).

Los sistemas de LiDAR discreto son los más empleados hasta la fecha, almacenando las coordenadas tridimensionales del primer y último retorno, y en algunas ocasiones de retornos intermedios. Sin embargo,

los sistemas de LiDAR *full-waveform*, de mayor complejidad, almacenan la onda completa permitiendo extraer propiedades físicas de los objetos. La amplitud de la onda en un punto está relacionada con la cantidad de energía recibida, por lo que dicha amplitud está a su vez relacionada con las propiedades del objeto en ese punto (Hermosilla *et al.*, 2014a). Ello permite una estimación más completa de variables de estructura.

En este artículo se analiza el preprocesado de los datos LiDAR *full-waveform* para la obtención de un conjunto de métricas extraídas a partir de la onda. También se generan y evalúan modelos de estimación de variables de estructura y se obtienen resultados comparativos para dos fechas diferentes. Además de analizarse algunas particularidades y problemas inherentes al procesado de LiDAR *full-waveform*, se plantea su aplicabilidad para estudios de caracterización de la estructura forestal.

2. MATERIAL Y MÉTODOS

El área de estudio corresponde a la zona de Panther Creek, en el estado de Oregón, al noroeste de los Estados Unidos. Tiene una superficie de 2258 ha y su altitud varía entre 100 y 700 m. La especie dominante es Douglas-fir (*Pseudotsuga menziesii*) ocupando más de la mitad de la superficie total. En ocasiones esta especie aparece combinada con otros tipos de coníferas. La altura de los árboles en la zona supera en ocasiones los 60 m aunque es muy variable debido a que su uso mayoritario es el aprovechamiento maderero.

Los datos LiDAR, cuyas características de vuelo se observan en la Tabla 1, se distribuyeron en formato LAS 1.3. Además, la empresa Watershed Sciences, Inc. facilitó un modelo digital del terreno (MDT) generado a partir de los últimos retornos de los pulsos. Tras la evaluación del MDT empleando 33 puntos de control GPS, se obtuvo un error medio cuadrático de 0,19 m.

Tabla 1. Características de los datos LiDAR

Característica	Descripción
Fecha de adquisición	15/07/2010 y 10/06/2012
Sensores	Leica ALS60 (2010) y ALS50 (2012)
Avioneta	Cessna Caravan 208B
Frecuencia pulso	105 kHz
Altura media vuelo	900 m
Ángulo escaneo	±15° desde nadir
Almacenamiento	256 bins
Intervalo	2 ns
Tamaño huella	~0,25 m
Densidad	≥ 8 puntos/m ²
Solape	≥ 50% entre pasadas (≥ 100% zona estudio)
Frecuencia registro posición	2 Hz
Frecuencia medición altitud	200 Hz
Frecuencia medición IMU	200 Hz

Para la toma de datos de campo se establecieron 78 teselas circulares de 16 m de radio. El error en la precisión horizontal y vertical en su localización fue inferior a 0,3 m, medido con un receptor Trimble R-8 GNSS y una estación total Leica TPS 800. Se catalogaron todos los árboles con un diámetro a la altura de pecho (DBH) igual o mayor a 2,5 cm que se encontraban en el interior de las teselas.

Tabla 2. Variables de estructura de la vegetación

Abreviatura	Descripción	μ	σ
AGB (t/ha)	Biomasa aérea	310,4	202,0
BA (m ² /ha)	Área basimétrica	46,4	22,8
QMD (m)	Diámetro medio cuadrático	0,332	0,136
SDI	Índice densidad estándar	288,3	115,6
V (m ³)	Volumen	361,9	171,5
CH (m)	Altura vegetación	22,90	8,92

A partir de las mediciones tomadas en campo (altura de la vegetación, DBH, especie, etc.) se aplicaron ecuaciones alométricas por especie (Standish *et al.*, 1985) para la obtención de los valores correspondientes

a las variables de estructura (ver Tabla 2) en cada tesela. Estos valores se utilizaron como referencia en los modelos calculados a partir de las métricas extraídas de los datos LiDAR *full-waveform*.

2.1. Preprocesado de los datos LiDAR

Es preciso realizar un preprocesado de los datos *full-waveform* para eliminar el ruido y convertir las ondas en valores con coordenadas tridimensionales y amplitudes asociadas (Hermosilla *et al.*, 2014a).

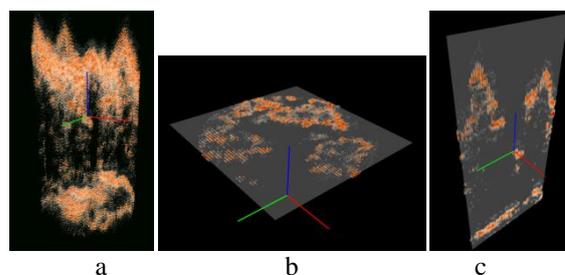


Figura 1. (a) Detalle de visualización 3D de una tesela; (b) sección horizontal; y (c) sección vertical

La mayoría de las trayectorias de los pulsos LiDAR *full-waveform* no coinciden con el nadir. Esto implica que cuanto más se aleje el haz del nadir, más oblicua es la trayectoria y más deformación existe. Este error se ve incrementado cuando se trabaja en zonas forestales ya que el número de puntos obtenidos desde diferentes posiciones aumenta considerablemente. Si no se tuviera este aspecto en cuenta existiría ambigüedad en cuanto a la información altimétrica registrada. Para resolver este error se siguió un proceso de generación de ondas pseudoverticales con los siguientes pasos (Hermosilla *et al.*, 2014b):

- Eliminación del ruido mediante umbral dinámico, cuyo valor es igual a la media más cuatro veces la desviación típica de la onda. Todos los puntos que superan este umbral forman la nueva onda, la cual es suavizada por medio de un filtro Gaussiano.
- Georreferenciación de las *full-waveform* por medio de los metadatos de los archivos LAS, y obtención de la altura normalizada mediante el MDT. De este modo se generaron los puntos con sus coordenadas tridimensionales y su amplitud extraída de la propia *full-waveform* corregida.
- Voxelización de los puntos generados. Cada vóxel con un tamaño de 0,25 x 0,25 x 0,30 m, siendo 0,25 m el tamaño de la huella y 0,30 m la distancia recorrida por la onda en el *temporal sample spacing* (2 ns). A cada vóxel se le asignó el valor de amplitud máxima de los puntos incluidos en él, preservando así el mayor retorno de las diferentes ondas originales (Figura 1).
- Generación de las ondas pseudoverticales extrayendo los valores de cada uno de los vóxeles de una misma columna, formando así una onda vertical.

Tabla 3. Resultados de la evaluación mediante validación cruzada efectuada para cada variable de estructura considerando el estrato global (78 teselas) y el estrato Douglas-fir (47 teselas)

Variables	Datos	R ²		ECM		EAM		ERM	
		μ	σ	μ	σ	μ	σ	μ	σ
AGB	Global	88,04%	9,29%	67,27 t/ha	19,11 t/ha	50,58 t/ha	10,53 t/ha	21,43%	5,11%
	Douglas-fir	90,17%	15,47%	67,97 t/ha	36,32 t/ha	54,78 t/ha	28,11 t/ha	20,86%	8,05%
BA	Global	74,25%	15,26%	11,01 m ² /ha	2,76 m ² /ha	8,38 m ² /ha	1,96 m ² /ha	20,50%	4,20%
	Douglas-fir	79,18%	30,03%	10,08 m ² /ha	4,93 m ² /ha	8,39 m ² /ha	3,83 m ² /ha	23,46%	12,47%
QMD	Global	83,41%	16,30%	0,06 m	0,02 m	0,04 m	0,02 m	14,44%	4,70%
	Douglas-fir	87,71%	14,80%	0,05 m	0,02 m	0,04 m	0,02 m	11,92%	4,28%
SDI	Global	66,61%	21,38%	64,52	18,00	46,96	10,23	18,11%	5,21%
	Douglas-fir	74,24%	30,95%	59,15	32,30	48,74	22,71	21,40%	10,69%
V	Global	69,10%	26,00%	84,25 m ³	25,86 m ³	66,51 m ³	15,46 m ³	24,26%	5,84%
	Douglas-fir	90,68%	9,15%	75,62 m ³	27,74 m ³	59,76 m ³	16,20 m ³	17,55%	6,42%
CH	Global	70,17%	24,64%	4,65 m	1,70 m	3,52 m	1,18 m	18,77%	7,78%
	Douglas-fir	84,14%	17,24%	4,28 m	1,92 m	3,17 m	1,18 m	15,04%	5,26%

2.2. Extracción de las métricas *full-waveform*

Las métricas derivadas de las ondas pseudoverticales se basaron en las propuestas por Duong (2010). *HOME* es la distancia desde la mediana de la onda hasta el terreno; *WD* es la distancia desde el comienzo de la onda hasta el suelo; *NP* es el número de picos; *ROUGH* es la distancia desde el comienzo de la onda hasta el primer pico; *FS* es el ángulo vertical desde el comienzo de la onda hasta el primer pico; *RWE* es la cantidad de energía de la onda; *HTMR* es el cociente entre *HOME* y *WD*; y *VDR* es la diferencia entre *WD* y *HOME* dividida entre *WD*. Para cada una de estas métricas se obtuvo tanto la media como la desviación típica de los valores presentes en cada una de las teselas.

2.3. Generación de los modelos de regresión

Se generaron modelos de estimación de las variables de estructura recogidas en la Tabla 2 empleando los valores obtenidos para las 78 teselas a partir de los datos de campo como referencia y las métricas extraídas de las ondas pseudoverticales como variables independientes. La zona de estudio se dividió en dos estratos: uno en el que la especie dominante es Douglas-fir y otro mixto. Se generaron dos grupos de modelos de estimación, uno considerando todas las teselas (78) y otro únicamente con las teselas de Douglas-fir (47).

La selección de variables independientes se llevó a cabo mediante análisis paso a paso, limitando a tres el número máximo de variables en el modelo. Se ensayaron varios tipos de modelos de regresión: lineal, polinomial, inverso y logarítmico, seleccionando en cada caso aquél cuya R² ajustada fuera mayor. Se realizó una validación cruzada con el fin de evaluar los resultados y poder garantizar una independencia entre los datos de entrenamiento y de evaluación.

Una vez generados los modelos, se aplicaron a toda la zona de estudio. De esta forma, se obtuvieron mapas de variables de estructura para 2010 y 2012. A partir de ellos se obtuvieron mapas de diferencias para analizar la evolución.

3. RESULTADOS

3.1. Validación de los modelos

En la Tabla 3 se recogen los siguientes indicadores estadísticos como resultado de la evaluación mediante validación cruzada para los dos estratos considerados: coeficiente de determinación (R²), error cuadrático medio (ECM), error absoluto medio (EAM) y error relativo medio (ERM). Para todos ellos se calcularon la media y la desviación típica del proceso de validación cruzada.

3.2. Evolución de las variables estructurales

Los modelos de regresión aplicados a los datos *full-waveform* de 2010 y 2012 sirvieron para obtener las estimaciones de las diferentes variables estructurales y su variación. La Figura 2 muestra la AGB y la Figura 3 su variación entre las dos fechas. En la Tabla 4 se muestran los valores estadísticos de las variaciones para las diferentes variables estructurales entre ambas fechas.

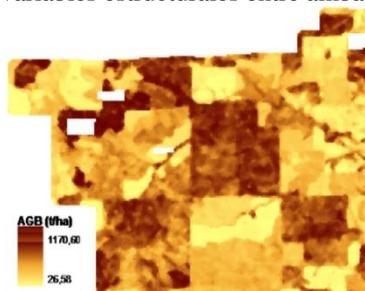


Figura 2. Estimación de la biomasa aérea (AGB) en el año 2010

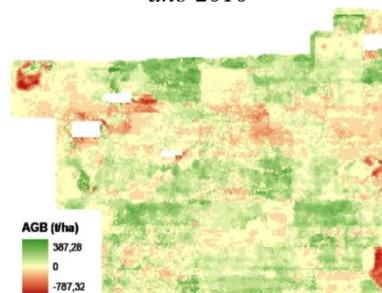


Figura 3. Variación de la biomasa aérea (AGB) entre 2010 y 2012

Tabla 4. Resultados estadísticos de las variaciones entre 2010 y 2012 de las diferentes variables estructurales

Variable	μ	σ	Mín.	Máx.
AGB (t/ha)	10,14	39,51	-787,32	387,28
BA (m ² /ha)	1,79	5,29	-84,37	42,89
QMD (m)	0,008	0,040	-0,470	0,158
SDI	10,09	29,95	-405,64	189,25
V (m ³)	18,45	51,38	-618,59	340,77
CH (m)	1,07	3,35	-29,76	15,40

4. DISCUSIÓN Y CONCLUSIONES

A partir de los resultados de la Tabla 3 se observa que para todo el área de estudio la variable AGB presenta el mejor ajuste ($R^2 = 88,04\%$) y la variable SDI el más bajo ($R^2 = 66,61\%$). Considerando únicamente las teselas de Douglas-fir se obtienen las mejores estimaciones para el V ($R^2 = 90,68\%$) y la AGB ($R^2 = 90,17\%$); en este caso las estimaciones más bajas se encuentran también en el caso de la variable SDI ($R^2 = 74,24\%$).

Los resultados muestran que la estimación mejora cuanto más homogénea es la vegetación. Ello se manifiesta especialmente para la variable V, variando el valor de R^2 69,10% a 90,68% al estratificar la muestra. Ello concuerda con estudios anteriores (Ruiz *et al.*, 2014), mejorando algunas variables como AGB y obteniendo un ajuste más bajo en QMD.

En cuanto a la evolución temporal, hay dos áreas situadas al noroeste y sureste de la zona de estudio donde todas las variables estimadas disminuyen. Esto es debido a que durante esos dos años ha habido una tala de árboles muy localizada. La Tabla 4 muestra cómo todos los promedios de las variables son positivos, lo que indica que a pesar de estas talas con valores negativos muy significativos, existe un balance positivo a nivel general, debido al crecimiento progresivo de la vegetación y a algunos focos de reforestación.

Durante el procesado de los datos se detectó un problema metodológico propio de los datos *full-waveform*, consistente en la aparición de un bandeo en las métricas calculadas. Este bandeo está relacionado con la densidad de los datos ocasionada por solape entre las pasadas de adquisición, y deriva del proceso de voxelización, en el que se asigna el valor máximo de todas las amplitudes que atraviesan cada vóxel.

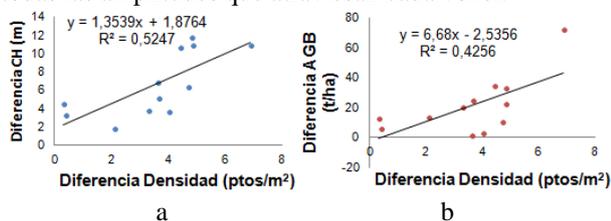


Figura 4. Relación entre las diferencias de densidad de puntos y de las variables estimadas: CH (a), y AGB (b)

Para valorar este hecho, se realizó un estudio de la influencia del bandeo en los valores de las variables de estructura finales, comparando estas variables y la densidad en pares de zonas colindantes. Los resultados preliminares indican la existencia de una tendencia en la que a medida que la diferencia entre densidades aumenta,

también lo hacen las diferencias de CH (Figura 4a) y AGB (Figura 4b). Como conclusiones, los datos *full-waveform* permiten estimar variables de estructura con gran precisión, al extraer más información de los diferentes estratos de la vegetación en alturas, con respecto a la obtenida por los sistemas de LiDAR discreto. Sin embargo, es preciso resolver aspectos metodológicos tales como los errores de bandeo generado en las propias métricas extraídas del *full-waveform*, el cual se traduce en ligeros errores en las estimaciones de las variables. Este error deriva del proceso de voxelización, y deberá resolverse en futuros trabajos.

La optimización del uso de la tecnología LiDAR, en especial *full-waveform*, permitirá aumentar la eficiencia en la estimación de variables de estructura, fundamentales en tareas de inventarios forestales, mantenimiento de los bosques y prevención de incendios forestales.

5. AGRADECIMIENTOS

Este trabajo ha sido financiado gracias al proyecto CGL2013-46387-C2-1-R del Ministerio de Economía y Competitividad.

6. BIBLIOGRAFÍA

- Duong, H.V. 2010. *Processing and application of ICESat large footprint full waveform laser range data*. PhD Thesis, Delft University of Technology.
- Hermosilla, T., Ruiz L.A., Kazakova, A.N., Coops, N.C., & Moskal, L.M. 2014a. *Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data*. International Journal of Wildland Fire, 23(2), 224-233.
- Hermosilla, T., Coops, N.C., Ruiz, L.A., & Moskal, L.M. 2014b. *Deriving pseudo-vertical waveforms from small-footprint full-waveform LiDAR data*. Remote Sensing Letters, 5(4), 332-341.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., & Harding, D. 1999. *Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests*. Remote Sensing of Environment, 70, 339-361.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. 2003. *Lidar remote sensing of forest structure*. Progress in Physical Geography, 27(1), 88-106.
- Ruiz, L.A., Hermosilla, T., Kazakova, A.N., & Moskal, L.M. 2014. *Comparison of forest structure estimates using discrete and full-waveform LiDAR metrics*. ForestSat-2014, 4-7 Nov., Riva del Garda, Italy.
- Standish, J.T., Manning, G.H., Demaershalk, J.P. 1985. *Development of biomass equations for British Columbia tree species*. Canadian Forestry Service, Pacific Forest Research Center, Information Report BC-X-264 (Victoria, BC).