
XVI Congreso Nacional de Tecnologías de la Información Geográfica 25, 26 y 27 de Junio de 2014. Alicante.

Estimación de variables dasométricas a partir de datos LiDAR PNOA en masas regulares de *Pinus halepensis* Mill.

J. Cabrera^a, M.T. Lamelas^{bc*}, A.L. Montealegre^{ac}, J. de la Riva^{ac}

^aDepartamento de Geografía y O.T., Universidad de Zaragoza

^bCentro Universitario de la Defensa, Academia General Militar

^cGrupo GEOFOREST-IUCA, Universidad de Zaragoza

Resumen

El conocimiento de las masas forestales es fundamental para su correcta gestión y ordenación. En ocasiones no basta con un inventario cualitativo del monte, siendo necesaria una valoración cuantitativa, mediante la estimación de variables dasométricas. La tecnología LiDAR aporta una nueva perspectiva a los inventarios forestales al ofrecer de forma directa información tridimensional de toda la superficie. El IGN inició en 2008-2009 la captura de datos LiDAR para gran parte de España, dentro del Plan Nacional de Ortofotografía Área (PNOA). Este trabajo pretende evaluar la adecuación de estos datos para estimar variables dasométricas en masas regulares de *Pinus halepensis* Mill. El área de estudio son los montes “Dehesa de los Enebrales” y “Valdá y Carrilanga” (Daroca, Zaragoza). Se han generado modelos de regresión lineal múltiple entre las variables dasométricas, obtenidas en 61 parcelas de campo, y una colección de variables estadísticas extraídas de la nube de puntos LiDAR. Los coeficientes de determinación corregidos obtenidos son 0,867 para la estimación del volumen, 0,854 para el área basimétrica, 0,858 para la densidad y 0,799 para la altura media. Las variables LiDAR introducidas en los modelos en general incluyen al menos un estadístico referente a altura (m) y otro a la distribución horizontal de la nube de puntos.

Palabras clave: LiDAR; PNOA; inventario forestal; variables dasométricas; regresión lineal múltiple.

* E-mail: tlamelas@unizar.es.

1. Introducción

La gestión y ordenación de montes implica, en muchas ocasiones, la estimación de variables dasométricas que tradicionalmente se han obtenido mediante un muestreo sistemático por parcelas, requiriendo una gran inversión económica e intenso trabajo de campo. A este respecto, la teledetección ofrece diferentes alternativas metodológicas basadas en el uso de imágenes ópticas y más recientemente de los sensores activos como el SAR (*Synthetic Aperture Radar*) y el LiDAR (*Light Detection And Ranging*). La ventaja de estos frente a las primeras es que son capaces de capturar información continua sobre la altura de la vegetación, que es un parámetro imprescindible en la estimación de otras variables dasométricas.

En el caso concreto de los sensores LiDAR, la información de carácter tridimensional que proporcionan ya se ha utilizado con éxito en la estimación de diferentes parámetros forestales, como la altura del dosel, el volumen de madera, el diámetro de las copas, etc., de forma automática y con mayor precisión a la alcanzada con técnicas de inventario por parcelas (Vosselmann y Maas, 2010). Dado que la práctica totalidad del territorio español ha sido sobrevolado recientemente (2008-2012) en el marco del Plan Nacional de Ortofotografía Aérea (PNOA) con sensores LiDAR, surge la necesidad de explotar esta nueva fuente de información topográfica con potencial para el estudio de las masas forestales.

En este contexto, el objetivo principal de este trabajo es evaluar la idoneidad de los datos LiDAR para estimar variables dasométricas en masas regulares de *Pinus halepensis* Mill. y obtener modelos estadísticos que relacionen adecuadamente la información aportada por el LiDAR con la medida directamente en el campo, con el fin último de ayudar a la gestión de este tipo de montes, tan habituales en la Comunidad Autónoma aragonesa.

2. Metodología, materiales, datos y herramientas

2.1. Zona de estudio

El área objeto de estudio corresponde a los montes “Dehesa de los Enebrales” (964 ha) y “Valdá y Carrilanga” (138 ha), ambos situados en el municipio de Daroca, provincia de Zaragoza y enmarcados en el sector centro-oriental de la Cordillera Ibérica (Fig. 1).

La vegetación que puebla ambos montes es principalmente una masa pura de *Pinus halepensis* procedente de repoblación. El monte “Dehesa de los Enebrales” fue repoblado entre los años 1908 y 1920, siendo la edad actual de la masa de unos 100 años, con una densidad muy variable, en función de las condiciones particulares de cada zona (tratamientos de mejora, condicionantes bióticos, etc.). El monte “Valdá y Carrilanga” fue repoblado durante los años 1956 y 1957, conformando una masa de unos 56 años de edad más homogénea.

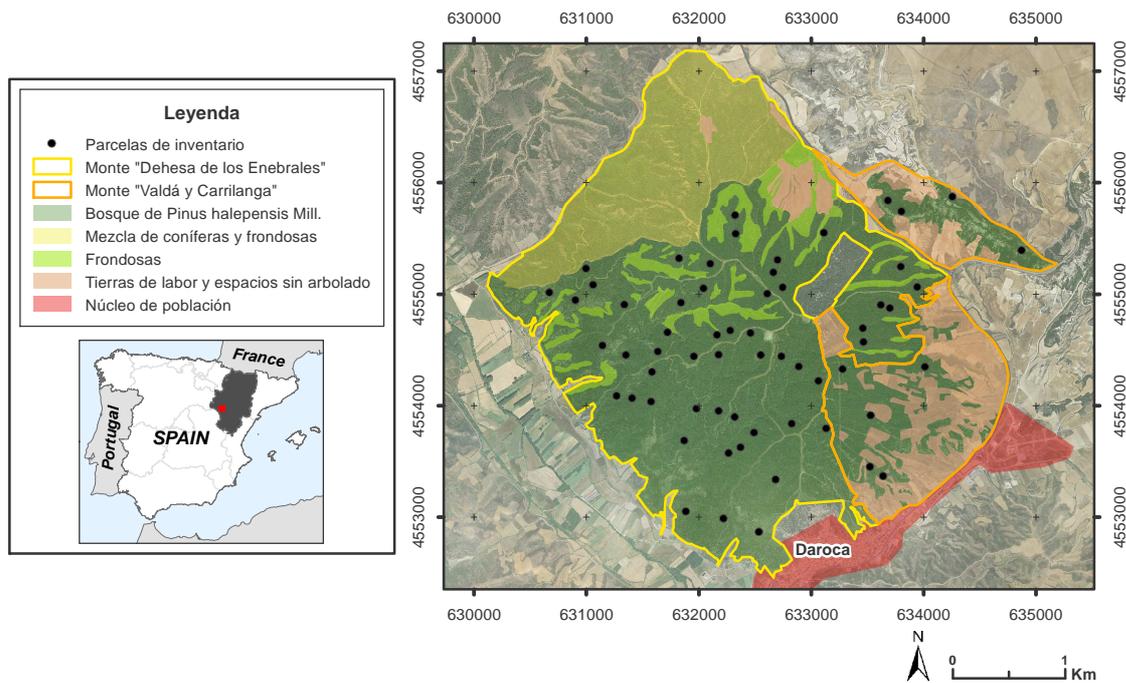


Fig. 1. Localización del área de estudio.

2.2. Datos LiDAR

Los datos LiDAR, proporcionados por el PNOA, fueron capturados durante los meses de agosto y octubre del año 2010 con un sensor Leica ALS60 de pulsos discretos y espejo oscilante. El desfase temporal con la fecha de adquisición de los datos de campo tomados en junio-julio de 2013, no se consideró importante, ya que, en ese intervalo de tiempo, no se habían producido cambios significativos en la estructura forestal.

La información fue suministrada en 9 ficheros en formato binario LAS de 2x2 km sin clasificar, por lo que, tras eliminar los errores debidos a ruido, fue necesario llevar a cabo un proceso de filtrado para separar los retornos del terreno del resto de información topográfica. Con este propósito se empleó el software FUSION LDV v.3.20, desarrollado por el Departamento de Agricultura y el Servicio Forestal de Estados Unidos, que incluye un algoritmo iterativo de clasificación, adaptado de Kraus y Pfeifer (1998). Los puntos clasificados como suelo fueron interpolados para crear un modelo digital de elevaciones (MDE) que permitió normalizar la nube de puntos y pasar a alturas sobre el terreno. Después, se calcularon, para cada parcela, un conjunto de variables derivadas de la nube de puntos LiDAR que tienen relación con variables dasométricas.

2.3. Datos de campo

Los modelos de estimación deben recoger toda la variabilidad que potencialmente puede presentar el área de estudio, por lo que se diseñó una categorización de la zona, en base a la cual se establecieron el número de parcelas de cada clase. Se consideró que algunas variables como la fracción de cabida cubierta (FCC) o la altura dominante, influyen enormemente en las variables dasométricas de una masa, y que variables como la pendiente o la presencia de sotobosque pueden influir en la precisión del modelo digital de elevaciones (MDE), por lo que esta información fue generada a partir de los datos LiDAR y se determinó el número de

parcelas a inventariar en cada categoría en función de su superficie y situándolas de manera aleatoria estratificada por categoría.

Se determinó un único estrato en los montes, ya que son masas con diferente edad y estructura pero de la misma especie principal. Posteriormente, para garantizar la validez estadística de los modelos obtenidos, se seleccionaron un total de 61 parcelas circulares de 15 metros de radio, que equivalen a una fracción de muestreo de 0,58%. En el campo fueron localizadas mediante un GPS-GNSS de precisión submétrica modelo GS15 de Leica.

En relación a las mediciones realizadas en el campo, se obtuvo el diámetro normal de todos los pies de la parcela y para la medición de las alturas se eligieron cuatro pies de forma aleatoria, a ser posible un pie de cada clase diamétrica presente con mayor frecuencia en el monte (22,5-27,4 cm; 27,5-32,4 cm; 32,5-37,4 cm; 37,5-42,4 cm). A partir de estos datos se calcularon para cada parcela los siguientes parámetros dasométricos: volumen con corteza, área basimétrica, densidad y altura media.

Para la estimación del volumen con corteza (m^3/ha), se ha empleado la tarifa de cubicación del Tercer Inventario Forestal Nacional (IFN3) en la provincia de Zaragoza. Aunque es necesario conocer la altura de todos los árboles de la parcela, como se ha comentado anteriormente esta únicamente se midió en cuatro árboles de cada parcela, por lo que para la estimación de la altura de los árboles no medidos en campo, se generó un modelo predictivo generalizado siguiendo una metodología similar a la que se presenta en el apartado 2.4. El área basimétrica se obtiene sumando las secciones normales (a 1,3 metros de altura) de todos los pies de la parcela y dividiendo entre la superficie de la misma, para extrapolarlo a una hectárea. La densidad, expresada en número de pies por hectárea, se calcula refiriendo a la hectárea los pies contabilizados en la parcela. Finalmente, la altura media de cada parcela, se ha obtenido realizando el promedio de las alturas individuales de cada árbol.

2.4. Modelos predictivos para la estimación de variables

Una vez obtenidas las variables dependientes, a partir de los datos de campo, y las variables independientes, en el caso de la estimación de la altura de los árboles mediante trabajo de campo, y en el caso de variables dasométricas mediante el tratamiento de los datos capturados por el sensor LiDAR, se relacionaron mediante un modelo de regresión lineal.

Previamente, se preseleccionaron las variables independientes que mejor se correlacionaban con la variable a estimar. Posteriormente, estas variables se incluyeron en un proceso de selección, en el que finalmente se eligieron el menor número de las que aportasen más información, empleando para ello el método de selección "paso a paso". Una vez seleccionadas las variables independientes se ajustó un modelo de regresión lineal por el método de mínimos cuadrados, comprobándose el cumplimiento de las hipótesis básicas del modelo de regresión lineal y analizando la idoneidad tanto de las variables originales como de nuevas variables generadas mediante transformación logarítmica. Se debe tener en cuenta que la transformación logarítmica inversa introduce un sesgo (subestimación) en el modelo, ya que el error aditivo, generado en el modelo lineal, se convierte en multiplicativo al aplicarla (Baskerville, 1972). Para la eliminación de este sesgo, el resultado final debe ser multiplicado por un factor de corrección (FC), calculado a partir del error estándar de estimación (SEE).

Una vez seleccionado el modelo óptimo, comparando para ello los estadísticos que indican la bondad del ajuste: el coeficiente de determinación ajustado (R^2 corregido) y la media (sesgo) y la raíz de la media cuadrática de los errores (RMSE), se realizó un proceso de validación cruzada dejando un dato fuera (Leave-one-out cross-validation LOOCV), al disponer únicamente de los datos de campo empleados para el ajuste del modelo.

3. Resultados

Para construir el modelo de estimación de la altura de cada árbol, en primer lugar se incluyó únicamente el diámetro normal (dn) medido en campo, obteniéndose bajos coeficientes de correlación, por lo que se probó a introducir en los modelos variables de masa, ya que según la literatura consultada, en general mejoran los resultados del modelo (Diéguez-Aranda et al., 2005). Se comprobó que el modelo se ajustaba mejor a los datos utilizando el área basimétrica de cada parcela (G), que había sido medida en campo (modelos 1 y 2 en tabla 1), y aplicando una transformación mediante logaritmos naturales, ya que el parámetro R^2 corregido pasaba de 0,555 a 0,628 (modelo 3 en tabla 1).

Tabla 1. Resumen de los modelos analizados para la estimación de la altura

Nº	Variable dependiente	Variables independientes	Hipótesis básicas	R^2 corregido	RMSE	Sesgo
1	h	dn, G	Sí	0,555	1,786	0,00
2	Ln(h)	dn, G	No (homocedasticidad)	-	-	-
3	Ln(h)	Ln(dn), Ln(G)	Sí	0,628	1,736*	-0,01*

Finalmente, una vez validado, el modelo es el siguiente:

$$h_i = 0,776 * G^{0,179} * dn_i^{0,660} * CF \quad (1)$$

Donde:

h_i : altura del árbol (m).

dn_i : diámetro normal del árbol (cm).

G: área basimétrica de la parcela (m^2/ha).

CF: coeficiente de corrección del sesgo del modelo, 1,009.

En la tabla 2 se presenta un breve resumen de las principales variables dasométricas obtenidas a partir de las 61 parcelas inventariadas. Se puede observar que los datos cubren un amplio rango de valores reconociéndose en el inventario gran parte de la variabilidad presente en la zona de estudio. En la tabla 3 se presenta el resumen de los modelos de estimación de variables dasométricas seleccionados que son los siguientes:

$$Vcc = -53,300 + 7,857 * Percentil\ 10 + 1,686 * \%retornos_1 \quad (2)$$

$$G = -0,192 + 0,375 * \%retornos_2 \quad (3)$$

$$N = 20,247 * \%retornos_2^{1,470} * Percentil\ 95^{-1,257} * CF_1 \quad (4)$$

$$Hm = 3,065 * Percentil\ 50^{0,596} * CF_2 \quad (5)$$

Donde:

Vcc: volumen con corteza por hectárea (m³/ha).

Percentil 10: percentil 10 de la altura de los puntos LiDAR (m).

%retornos₁: porcentaje de retornos por encima de 3 ó 6 metros (depende del monte) respecto del total de primeros retornos (%).

G: área basimétrica (m²/ha).

%retornos₂: porcentaje de primeros retornos por encima de 3 o 6 metros (en función del monte) (%).

N: densidad (pies/ha).

Percentil 95: percentil 95 de la altura de los puntos LiDAR (m).

CF₁: coeficiente de corrección del sesgo del modelo de densidad, 1,045.

Hm: altura media (m).

Percentil 50: percentil 50 de la altura de los puntos LiDAR (m).

CF₂: coeficiente de corrección del sesgo del modelo de altura, 1,003.

Tabla 2. Resumen de las principales variables dasométricas de los montes “Dehesa de los Enebrales” y “Valdá y Carrilanga”. Dg: diámetro cuadrático medio; Hm: altura media; N: densidad de pies por hectárea; G: Área basimétrica; Vcc: Volumen maderable con corteza por hectárea.

Variables	Dehesa de los Enebrales				Valdá y Carrilanga			
	Media	Máximo	Mínimo	Desviación típica	Media	Máximo	Mínimo	Desviación típica
Dg (cm)	31,4	49,0	23,8	4,9	20,8	26,0	17,7	3,0
Hm (m)	13,0	16,1	9,1	1,2	9,5	11,9	6,0	1,6
N (pies/ha)	336,8	806,4	14,1	162,9	817,4	1471,3	14,1	401,2
G (m ² /ha)	23,5	37,3	2,7	7,2	25,6	36,6	0,8	11,4
Vcc (m ³ /ha)	119,5	195,6	13,0	38,4	113,7	179,4	2,1	54,0

Tabla 3. Resumen de los modelos de estimación de variables dasométricas. Hm: altura media; N: densidad de pies por hectárea; G: Área basimétrica; Vcc: Volumen maderable con corteza por hectárea; Ln Logaritmo neperiano

Nº	Variables dependientes	Variables independientes	Hipótesis básicas	R ² corregido	RMSE	Sesgo
1	Vcc	Percentil 10, Porcentaje de retornos por encima de 3 ó 6 metros respecto del total de primeros retornos.	Sí	0,867	15,42	0,16
2	G	Porcentaje de primeros retornos por encima de 3 o 6 metros.	Sí	0,854	3,05	0,01
4	Ln(N)	Ln(Percentil 95), Ln(Porcentaje de primeros retornos por encima de 3 o 6 metros).	Sí	0,857	0,34	0,00
3	Ln(Hm)	Ln(Percentil 50)	Sí	0,799	0,07	0,00

4. Discusión y conclusiones

Los resultados presentados reafirman la hipótesis de partida del mismo, mostrando la utilidad de los datos LiDAR PNOA para el inventario de masas regulares de pino carrasco. La metodología empleada tanto para el diseño del inventario de campo, como para el establecimiento de los modelos de regresión, ha resultado adecuada, obteniendo precisas estimaciones de variables como el volumen, área basimétrica, densidad y altura media para toda la zona de estudio. En este sentido, cabe destacar la utilidad de los estadísticos LiDAR en la estratificación y planificación del inventario de masas forestales, permitiendo, como se ha demostrado en este trabajo, realizar un muestreo aleatorio estratificado, en cierto modo dirigido, discriminando tipologías de masa a escala de detalle, con un mínimo trabajo de campo.

El buen ajuste de los modelos de estimación generados pone de manifiesto, no solo la adecuación de la metodología empleada, sino también, y aún más importante, la correlación existente entre los estadísticos obtenidos de la nube de puntos LiDAR y las variables dasométricas medidas en campo. Concretamente, han sido diferentes percentiles (mayores o menores, según la variable a estimar) y el porcentaje de retornos LiDAR por encima de una altura determinada, los estadísticos LiDAR que mejor ajuste han proporcionado en los modelos establecidos.

El empleo de una cobertura de datos LiDAR tres años anterior a la toma de datos de campo y una baja densidad de puntos LiDAR (0,5 puntos/m²), no ha afectado a la existencia de altas y significativas correlaciones entre las variables obtenidas en campo y las obtenidas a partir de los datos LiDAR. Por consiguiente, los resultados obtenidos en este trabajo demuestran la utilidad de los datos LiDAR PNOA en el inventario de masas regulares monoespecíficas, no siendo extrapolable dicha afirmación a otras tipologías de masa, debido a la dificultad que plantea la discriminación entre especies y clases diamétricas a partir de datos LiDAR. En este sentido, sería deseable que se estudiara la utilidad de estos datos en otras masas, irregulares, pluriespecíficas, montes bajos, etc., así como, en inventarios de grandes superficies, donde el uso de esta tecnología podría suponer un aumento en la precisión de los resultados y una disminución en los costes. Así mismo, teniendo en cuenta que la cobertura LiDAR PNOA va a ser un producto presumiblemente periódico, sería deseable evaluar la adecuación de los modelos ya generados, empleando las futuras coberturas LiDAR.

Agradecimientos

Los autores agradecen al Centro Nacional de Información Geográfica los datos LiDAR PNOA proporcionados para la realización de este trabajo.

Referencias

- Baskerville, G.L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.*, 2, 49-53.
- Kraus, K. & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53, 193-203.
- Diéguez-Aranda, U., Barrio Anta, B., Castedo Dorado, F. y Álvarez González, J. G. (2005). Relación altura-diámetro generalizada para masas de *Pinus sylvestris* L. procedentes de repoblación en el noroeste de España. *Invest Agrar: Sist Recur For*, 14(2), 229-241.
- Vosselmann, G. & Maas, H. G. (2010). *Airborne and terrestrial laser scanning*. Dunbeath: Whittles Publishing.