

Cartografía de variables forestales a partir de datos LiDAR

Apellidos, nombre	Ruiz Fernández, Luis Ángel (laruiz@cgf.upv.es)
Departamento	Ingeniería Cartográfica, Geodesia y Fotogrametría
Centro	ETSI. Geodésica, Cartográfica y Topográfica



1 Introducción

La tecnología LiDAR (*Light Detection and Ranging*) se ha extendido ampliamente en los últimos años para su utilización en la generación de inventarios forestales y, en general, en la obtención de información sobre la estructura forestal. Aunque su empleo está aún condicionado por la adquisición de datos de campo para calibrar y evaluar los modelos, su aplicación puede ya considerarse como operativa para la estimación y la elaboración de mapas de un buen conjunto de variables forestales.

Se presenta una metodología para la estimación de variables forestales a partir de datos LiDAR aerotransportado, basada en los siguientes pasos:

- (1) el pre-tratamiento de los datos para la obtención del modelo digital del terreno y el modelo de altura de vegetación normalizado;
- (2) la extracción de características o atributos a partir de los datos LiDAR para cada parcela de muestreo;
- (3) la generación de modelos de estimación de variables forestales; y
- (4) su extensión para la generación de mapas de variables forestales en zonas más amplias.

2 Objetivos

Una vez que el estudiante lea con detenimiento este documento, será capaz de:

- Identificar las metodologías para estimar variables forestales a partir de datos LiDAR.
- Interpretar los resultados obtenidos tras modelizar variables forestales a partir de datos LiDAR.

3 Desarrollo

El presente artículo presenta la metodología general para la estimación de variables forestales a nivel de parcela utilizando datos LiDAR. Existen numerosos estudios para la estimación de variables forestales a partir de imágenes y, sobre todo, de datos LiDAR aerotransportado (Andersen et al., 2006; Wulder et al., 2012). En estos estudios se evalúa el potencial de estos sistemas para la estimación de variables de inventario forestal, como son la biomasa, el volumen, la altura dominante, el área basimétrica, etc. No obstante, la mayoría de ellos se centran en áreas de gran homogeneidad forestal, muy diferentes a la diversidad propia de ciertos tipos de bosques, como son los mediterráneos. Otros estudios analizan la densidad de retornos LiDAR y el tamaño de las parcelas de campo necesarios para optimizar los resultados (Ruiz et al., 2014).

A continuación se describe el procedimiento completo para la generación, aplicación y evaluación de modelos de estimación de variables forestales a nivel de parcela empleando datos LiDAR aerotransportado, y su posterior extensión a zonas más amplias, con la finalidad de obtener mapas de variables forestales estimadas a nivel de monte. Se describen los métodos empleados desde la adquisición y el preprocesado de los datos hasta la obtención de los mapas de variables.

La figura 1 muestra el procedimiento general.

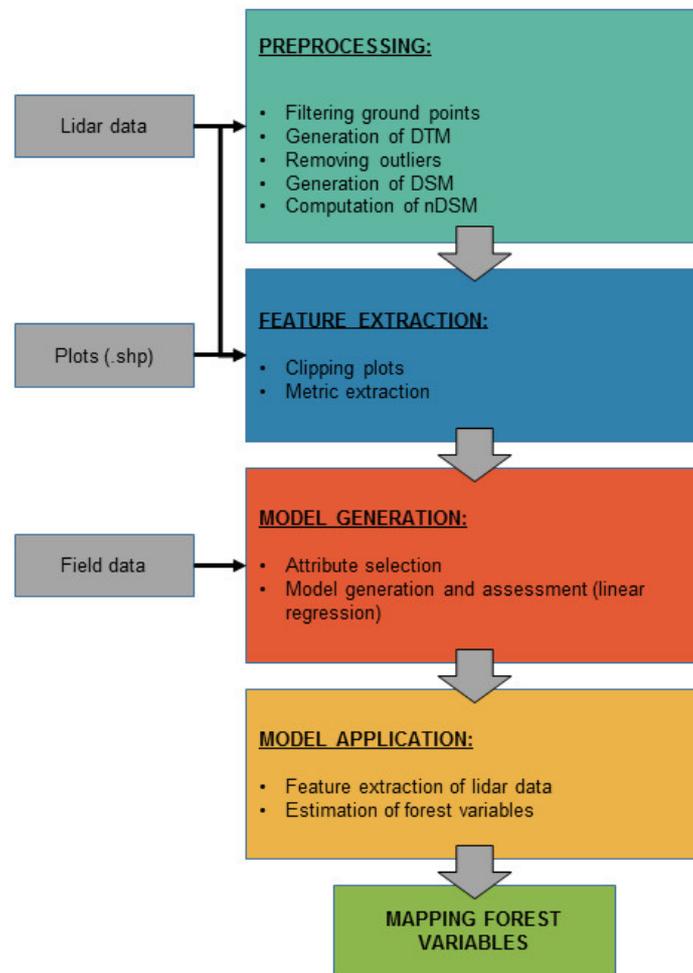


Figura 1.- Procedimiento general para la obtención de cartografía de variables forestales a partir de datos LiDAR.

3.1 Datos necesarios

Los datos básicos necesarios son los siguientes:

- Datos LiDAR discreto de alta densidad, normalmente con una densidad nominal de >4 ptos./m² aunque depende de las variables objeto de estudio, obtenidos mediante vuelo programado o a través de datos ya disponibles.
- Datos de campo de un conjunto de parcelas de muestreo (normalmente circulares de entre 12 y 25 m de radio) con mediciones necesarias para calcular las variables forestales deseadas (biomasa, volumen, área basimétrica, etc.) mediante ecuaciones alométricas estándar para el tipo de bosque

La figura 2 muestra un ejemplo de un conjunto de parcelas de campo circulares distribuidas en una zona de estudio piloto (figura 2a) y de la distribución de archivos que contienen los datos LiDAR (*files* de 1km x 1km) (figura 2b).

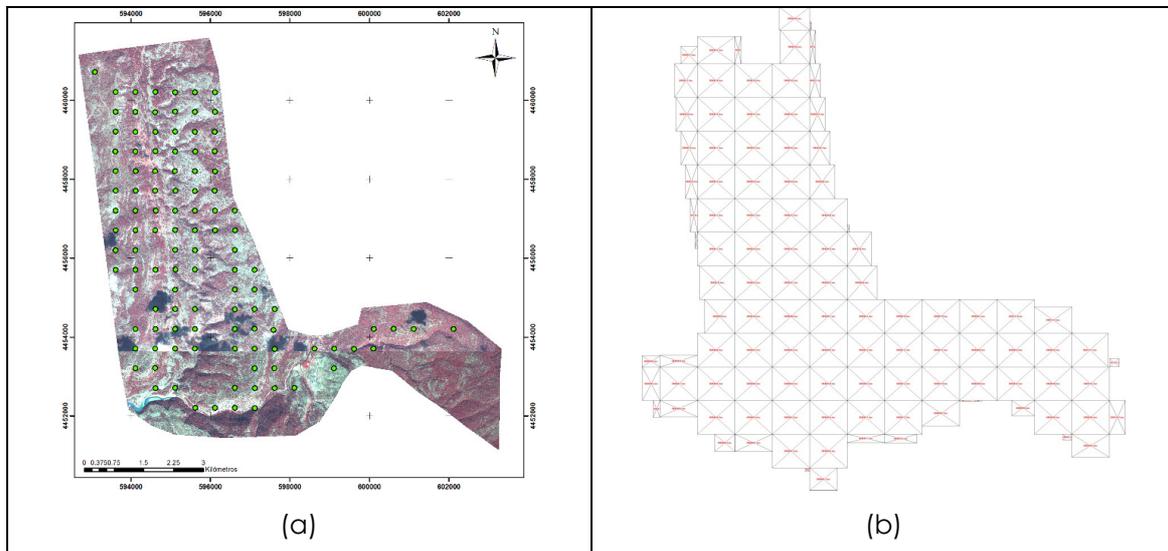


Figura 2.- Ejemplo de datos en una zona piloto: (a) distribución de parcelas de campo circulares (25m radio); y (b) archivos LiDAR de 1km x 1km cada uno.

3.2 Preprocesado de los datos LiDAR

En primer lugar, si fuera necesario se realiza un filtrado de puntos anómalos (figura 3), para ello suelen aplicarse filtros adaptativos que eliminan aquellos puntos LiDAR cuyo valor de cota (z) sea anómalamente alto o bajo con respecto a un vecindario local que se va desplazando en toda la zona de estudio. El tamaño de la ventana de filtrado se elegirá en función de la densidad de puntos.

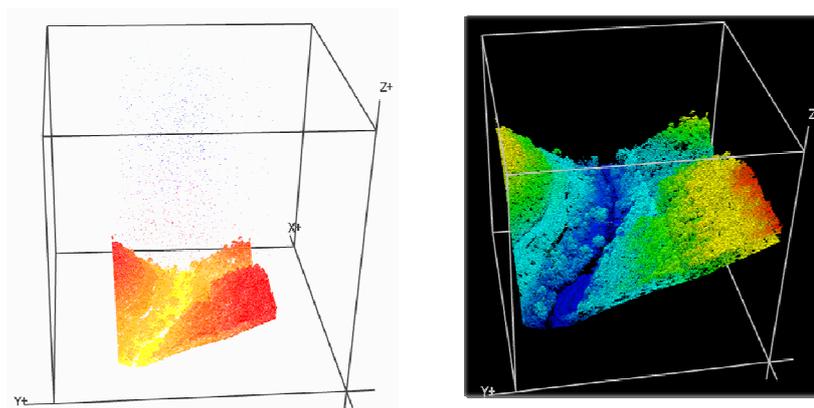


Figura 3.- Filtrado de puntos anómalos (a) antes y (b) después del filtrado.

El siguiente paso consiste en la generación del modelo digital del terreno (MDT). Para ello, se aplica un algoritmo que identifique y clasifique los puntos correspondientes al terreno y a continuación se genera una superficie mediante la interpolación de dichos puntos (figura 4b). Por otro lado, a partir de la nube de puntos original y estableciendo



un tamaño de celda definido en función de la densidad de puntos LiDAR, se obtiene el modelo digital de superficies (MDS) asignando el valor de los puntos con elevación máxima en cada celda (figura 4c). Una vez obtenidos ambos, la diferencia entre los valores de MDS y MDT en cada celda permite generar el modelo digital de superficies normalizado con respecto al terreno (nMDS), también llamado modelo de alturas de vegetación en aplicaciones forestales (figura 4d).

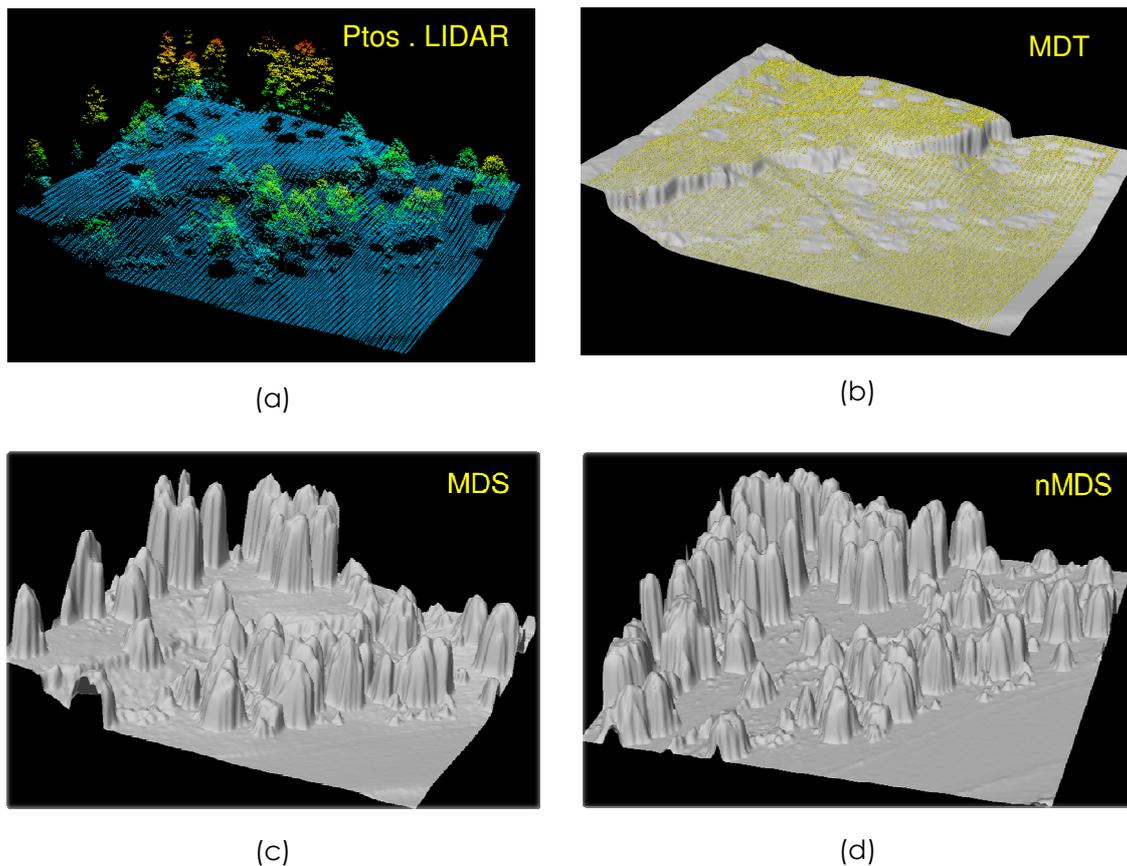


Figura 4.- Detalle de: (a) una nube de puntos LiDAR; (b) superficie del modelo digital del terreno (MDT) con los puntos del terreno extraídos en amarillo; (c) modelo digital de superficies (MDS); y (d) modelo digital de superficies normalizado (nMDS).

3.3 Cálculo de las características descriptivas

El cálculo de características o atributos descriptivos se realiza a nivel de parcela, de forma que pueda compararse con los datos tomados en campo para posteriormente poder generar los modelos. Se pueden distinguir tres grandes grupos de características:

- Las extraídas directamente a partir de la **nube de puntos** de cada parcela, que pueden ser a su vez:
 - variables **estadísticas** de la distribución en alturas de los puntos
 - variables derivadas de los **perfiles de densidad**

- las extraídas a partir del **modelo de superficies** normalizado (nMDS)
- las variables derivadas de los datos de **intensidad** LiDAR

Variables estadísticas: Dentro del primer grupo, las características más comunes son índices estadísticos como la media, desviación típica, máximo, *skewness* y *kurtosis* de cada parcela a partir de la nube de puntos normalizada (datos originales con el valor de altura de cada punto corregido respecto a la cota del MDT obtenido). Estas características se complementan con los valores de los percentiles de altura. El percentil de alturas p indica la altura por debajo de la cual se encuentra un % de puntos dentro de la parcela. Así, por ejemplo, el valor del percentil 90 nos da la altura por debajo de la cual se encuentra el 90% de los puntos. Las variables estadísticas de la distribución en alturas de los puntos LiDAR proporcionan información sobre la estructura interna de la vegetación.

Perfiles de densidad: A partir de la nube de puntos pueden generarse además perfiles de densidad, esto es, histogramas de alturas del conjunto de puntos dentro de cada parcela (figura 5a). A partir de estas curvas se extraen parámetros que las caracterizan, tales como los dos valores de máxima densidad de puntos (PDMAX y PDMAX2), o las alturas correspondientes a las dos densidades de puntos máximas (HDMAX y HDMAX2), como puede verse en el ejemplo de la figura 5c.

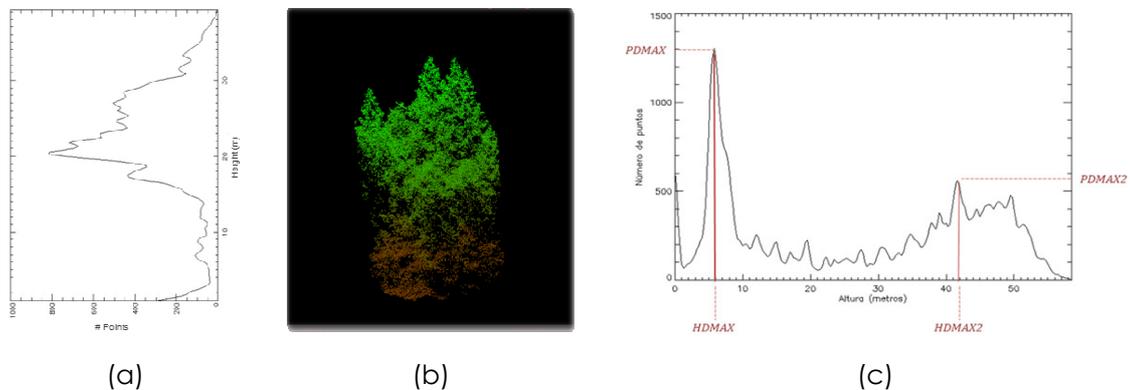


Figura 5.- Ejemplo de perfil de densidad (a) y nube de puntos (b) de una parcela. Algunos parámetros que permiten caracterizar la curva de perfil de densidad (c).

Modelos de superficie (nMDS) (Figura 6): Las características calculadas del nMDS proporcionan información sobre los valores máximos de altura en cada celda o píxel y su distribución espacial dentro de la parcela. Así, además de los valores de la media y desviación típica correspondientes a cada parcela, pueden calcularse variables de textura de primer o segundo orden, como por ejemplo aquellas derivadas de la matriz de coocurrencias de niveles de gris (Haralick et al., 1973) o variables calculadas a partir del semivariograma del nMDS de una parcela determinada (Ruiz et al., 2011).

Intensidad de los puntos LiDAR: Además de las coordenadas x , y , z , cada retorno LiDAR contiene información sobre la intensidad de la radiación reflejada en ese punto, esto es, información espectral correspondiente a la longitud de onda del sistema, generalmente NIR. De forma similar a las variables estadísticas de altura, pueden



obtenerse valores estadísticos de la distribución de las intensidades (media, desviación típica, skewness, kurtosis, percentiles, etc.).

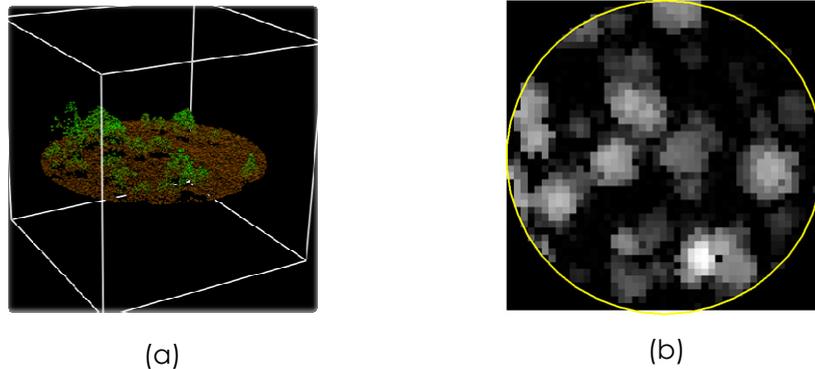


Figura 6.- (a) Representación 3D de la nube de puntos de una parcela circular (en verde la vegetación y en marrón el terreno); (b) modelo digital de superficies normalizado (nMDS), donde cada píxel representa el valor de altura máxima de los puntos incluidos en él.

Existen varios programas que permiten realizar el tratamiento de las nubes de puntos, desde la obtención de los modelos digitales hasta la extracción de características, uno de los programas gratuitos más empleados en aplicaciones forestales es *FUSION* (http://forsys.cfr.washington.edu/FUSION/fusion_overview.html).

3.3 Generación de modelos

Para la creación de los modelos es preciso disponer de datos de campo de un conjunto de parcelas, de forma que sirvan de datos de referencia. Así, los modelos de estimación pueden calcularse mediante técnicas de regresión múltiple, considerando las variables forestales (biomasa, volumen, altura dominante, etc.) una a una como variables dependientes y las métricas derivadas de los datos LiDAR descritas en la sección 3.3 como variables independientes. Sin embargo, para la generación de modelos robustos y extrapolables a toda la zona de estudio o incluso a otras zonas de características similares, es conveniente reducir el número de variables independientes al máximo y evitar problemas de multicolinealidad, producida cuando estas variables están muy correlacionadas entre sí. Para ello, se emplean técnicas de **selección de variables**, como por ejemplo la regresión múltiple paso a paso, en la que se van introduciendo variables en el modelo en función de su relevancia en la regresión, o bien se van eliminando en función de su correlación con otras variables. Así por ejemplo, si se aplica la técnica de regresión lineal múltiple, para cada variable forestal se obtendrá un modelo del tipo

$$V = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

donde V representa la variable forestal (dependiente), x_i las métricas empleadas en el modelo (variables independientes) y a_i los coeficientes obtenidos mediante el método de regresión ordinaria de mínimos cuadrados.



La evaluación de los modelos se realiza relacionando las variables observadas en campo y las predichas a partir de los datos LiDAR, analizando el **coeficiente de determinación R^2** , definido como una medida descriptiva del ajuste global del modelo que representa la proporción de varianza explicada por el mismo. El coeficiente R^2 se calcula como el cuadrado del coeficiente de correlación del momento producto de Pearson:

$$R = \frac{\sum_{i=1}^N (\phi_i - \bar{\phi}_i) \cdot (\phi_{iobs} - \bar{\phi}_{iobs})}{\sqrt{\sum_{i=1}^N (\phi_i - \bar{\phi}_i)^2 \cdot \sum_{i=1}^N (\phi_{iobs} - \bar{\phi}_{iobs})^2}}$$

donde i es el valor pronosticado a partir de los datos LiDAR para la parcela i , $iobs$ es el valor observado en campo para esa misma parcela, y N es el número de parcelas analizadas. Además, también se calcula el **error medio cuadrático (EMC)**, que representa la media de las diferencias promedio entre los valores pronosticados y observados, definido mediante la ecuación:

$$EMC = \sqrt{\sum_{i=1}^N \frac{(\phi_i - \phi_{iobs})^2}{N}}$$

Una práctica habitual cuando no se dispone de un número elevado de parcelas de campo (lo cual es muy común debido al coste económico que ello supone) es la evaluación mediante el procedimiento de **validación cruzada**, consistente en la utilización de $(n-1)$ subconjuntos de parcelas para la generación del modelo y 1 para la validación, repitiendo el proceso n veces de forma que el subconjunto de parcelas de validación es siempre distinto e independiente del resto.

3.4 Mapas de variables forestales

Para la obtención de los mapas de variables forestales es preciso aplicar el modelo obtenido a la zona de estudio completa. Para ello:

- Se divide el territorio en celdas cuadradas con una superficie equivalente a las de las parcelas de muestreo en campo.
- Se calculan para cada celda las métricas seleccionadas en el modelo (x_1, x_2, \dots, x_n) .
- Se aplica la ecuación del modelo para todas y cada una de las celdas, obteniendo así un valor numérico de la variable forestal en cada celda del mapa, con formato de imagen (píxel = celda).
- Edición del mapa en un SIG.

La figura 7 muestra un ejemplo de los mapas de biomasa total y de fracción de cabida cubierta obtenidos en la zona de la figura 2.

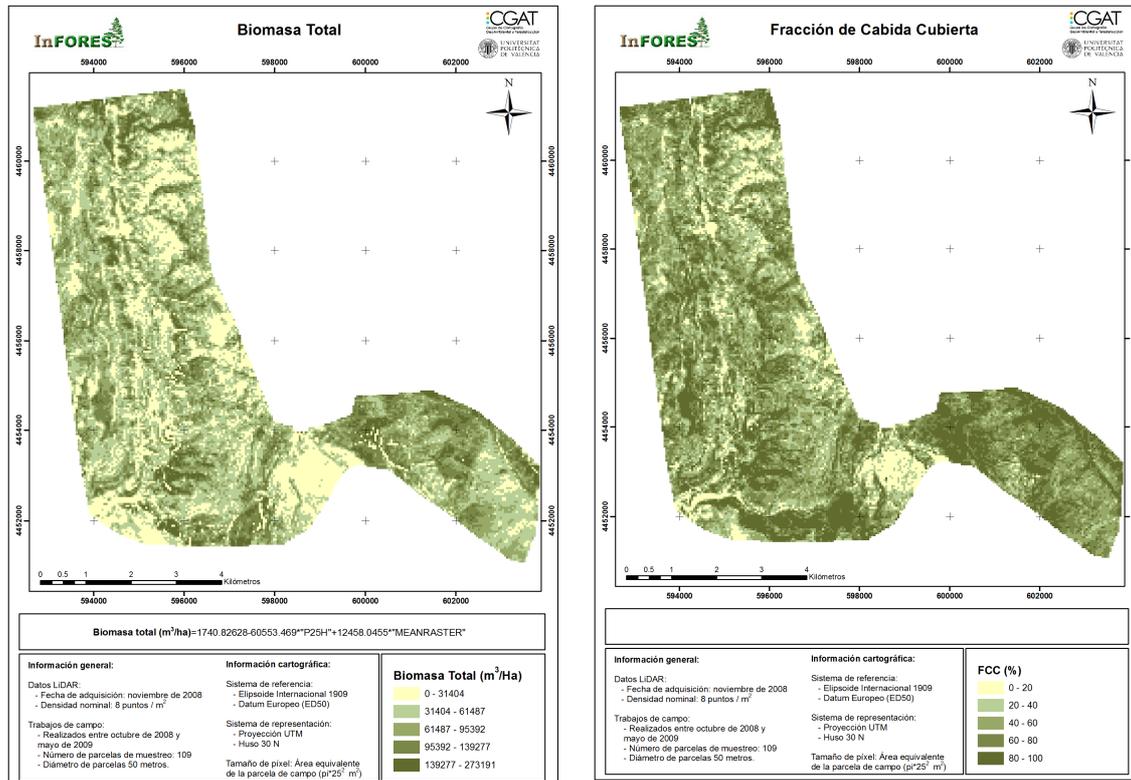


Figura 7.- Ejemplos de dos mapas (biomasa y fracción de cabida cubierta) obtenidos para la zona de la figura 2 aplicando el procesado de datos LiDAR descrito.

4 Cierre

Se ha presentado un procedimiento genérico para la estimación de variables forestales a partir de datos LiDAR aéreos y mediciones de referencia en parcelas de campo. Se ha descrito el procesado de los datos LiDAR que incluye: (i) la generación de los modelos digitales del terreno (MDT), de superficies (MDS) y de superficies normalizado (nMDS); (ii) la extracción de características o métricas; (iii) la generación de los modelos de predicción; y (iv) la extensión del modelo a toda la superficie de trabajo para obtener los mapas de variables forestales finales.



5 Bibliografía

5.1 Artículos:

- Andersen HE., Reutebuch S.E., McGaughey R.J., 2006. Active remote sensing. In: Shao G., Reynolds K.M. (eds) *Computer Applications in Sustainable Forest Management*. Springer, Dordrecht.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Texture features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3 (6) (1973), pp. 610-622.
- Ruiz, L.A., Recio, J.A., Fernández-Sarría, A., Hermosilla, T., 2011. A feature extraction software tool for agricultural object-based image analysis. *Computers and Electronics in Agriculture*, 76(2), pp. 284-296.
- Ruiz, L.A., Hermosilla, T., Mauro, F., Godino, M., 2014. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. *Forests*, 5(5), pp. 936-957.
- Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., ... Gobakken, T., 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, pp. 196-209.

5.2 Referencias de fuentes electrónicas:

- FUSION/LDV LIDAR analysis and visualization software. Pacific Northwest Research Station, USADA Forest Service. Autor: Robert J. McGaughey. Disponible en: http://forsys.cfr.washington.edu/FUSION/fusion_overview.html