

## Haciendo posible la docencia de proyectos de inventario forestal basados en datos procedentes de sensores láser aerotransportados exclusivamente con ArcGIS

Antonio San José

Luis Garmendia

Departamento de Ingeniería del Software e Inteligencia Artificial (DISIA)

Facultad de Informática. Universidad Complutense de Madrid

Francisco Mauro

Victoria Núñez

Antonio García Abril

Departamento de Proyectos y Planificación Rural.

Escuela Técnica Superior de Ingenieros de Montes. Universidad Politécnica de Madrid

### Abstract

During the last decade the use of airborne laser scanner data (LiDAR) for forest inventory purposes has become possible. The use of this data allows a significant reduction of the cost associated to this kind of projects. However, in order to treat LiDAR data, highly specialized software is required and the licenses are not usually at disposal for educational purposes because of their high cost. Furthermore, it happens that this highly specialized software is often also unknown by teachers.

This project shows how it is possible to completely develop medium or small scale LiDAR based inventory projects using only ArcGIS 10. This software is the most widespread and used software for geographic information management in educational institutions. Hence, thanks to this project and the free available laser data provided by the Spanish National Aerial Orthophotography Plan (PNOA), it is possible to implement teaching activities where the students can completely develop LiDAR based inventory projects without using any additional specialized software. This also allows saving money since no additional licenses are required.

**Keywords:** *LiDAR (Light Detection and Ranging); ArcGIS Desktop 10; \*.las files; shapefile; Forest inventory; normalization*

### Resumen

Durante la última década se ha hecho factible el uso de sensores láser aerotransportados (LiDAR) para desarrollar tareas de inventario forestal. El uso de este tipo de sensor ha revolucionado el inventario forestal, pues ha permitido reducir ostensiblemente los costes asociados a los inventarios de campo. Sin embargo, el tratamiento de la información de datos LiDAR requiere generalmente el uso de software altamente especializado cuyas licencias son muy costosas, no suelen estar disponibles para usos docentes y en muchos casos son desconocidas por los propios profesores.

El trabajo realizado muestra que usando el software de tratamiento de información geográfica más extendido y usado en el ámbito universitario, ArcGIS 10, se puedan desarrollar ejemplos a pequeña y mediana escala en los que se detalla exhaustivamente el proceso de inventario basado en datos LiDAR. De esta forma, y gracias a que actualmente la adquisición de datos LiDAR es gratuita en España gracias al Plan Nacional de

Ortofotografía Aérea (PNOA), se pueden realizar experiencias de clase en las que los alumnos desarrollen completamente proyectos de inventario forestal basado en datos LiDAR, sin necesidad de realizar una inversión económica adicional destinada a la obtención de software especializado.

**Palabras clave:** *LiDAR (Light Detection and Ranging); ArcGIS Desktop 10; archivos \*.las; shapefile; Inventario forestal, normalización.*

## 1. Introducción

El uso de sensores LiDAR en tareas de inventario forestal es una actividad que ha crecido fuertemente durante los diez últimos años. Tal ha sido el desarrollo, que en la actualidad, el Inventario forestal basado en datos LiDAR ha pasado de ser un campo al que sólo se dedicaban investigadores a ser una actividad que ya realizan empresas especializadas del sector. Estas empresas mayoritariamente emplean el método de estimación Area Based Approach (ABA), que ha sido probado con éxito en numerosas ocasiones (García, Riaño, Chuvieco, & Danson, 2010; Næsset & Gobakken, 2005; Næsset, Bollandsås, & Gobakken, 2005; Næsset, 2009). Ésta transferencia de conocimientos se ha realizado con cierto éxito y actualmente se requiere que las escuelas forestales proporcionen a sus alumnos competencias suficientes sobre el funcionamiento de dicha tecnología, que actualmente se encuentra en expansión.

Uno de los problemas que presenta el uso de datos LiDAR en tareas docentes es que requiere software altamente especializado, muchas veces incluso desconocido para los potenciales profesores. Además algunos de estos programas requieren el pago de costosas licencias, lo cual supondría un gasto extra que muchas universidades no podrían o no estarían dispuestas a asumir. Las alternativas no comerciales especializadas en el tratamiento de datos LiDAR, tales como FUSION (Mc Gaughey, 2011), también tienen ciertas desventajas a la hora de emplearlas en tareas docentes, como por ejemplo, la carencia de una interfaz gráfica o el ser herramientas muy específicas. Con el fin de solventar el problema de cómo enfocar la docencia de técnicas de Inventario Forestal basado en datos LiDAR, proponemos en este trabajo, un organigrama de una sesión de clase de prácticas en las que el software empleado sería ArcGIS 10 (Environmental Systems Research Institute., 2011) dada la enorme difusión de este software entre la comunidad universitaria y los profesionales del sector forestal.

El planteamiento inicial se centra en la necesidad de aplicar el método ABA a los datos LiDAR empleando el software ArcGIS Desktop 10. El tratamiento de datos LiDAR con este software concreto suscita un problema fundamental: la visibilidad de los campos de información incluida en los datos LiDAR normalmente proporcionados en formato \*.las (ASPRS, 2010). Debido a su codificación binaria, ArcGIS es incapaz de mostrar todos los campos con la información inicial de éstos archivos. Esta información es esencial para su posterior tratamiento y la elaboración de productos derivados de los datos LiDAR y a continuación se muestra una forma de proceder para acceder a toda la Información necesaria de los datos \*.las empleando ArcGIS 10. Si bien ArcGIS 10 presenta el problema de ser un software que requiere el pago de unas licencias, también es cierto que es un Sistema de Información Geográfica (SIG) de propósito general cuyo uso está totalmente implantado en la gran mayoría de las Universidades. Esto hace que no haya que realizar inversiones extra en licencias para el desarrollo de las prácticas.

## 2. Objetivos

Los objetivos del presente trabajo se centran en desarrollar un ejemplo, o caso de estudio, de Inventario Forestal basado en datos LiDAR mediante el uso de ArcGIS 10. El trabajo consistiría en transformar los datos LiDAR en archivos con formato shapefile, mostrando todos los campos de información, lo cual permitiría su uso sin restricciones. El ejemplo que se expone se desarrolla con los datos proporcionados por el Grupo de Investigación, Tecnologías y Métodos para la Gestión Sostenible, y en él se desarrolla todo el proceso de estimación de dos variables de interés forestal como son la biomasa y el volumen

maderable, si bien este mismo esquema se podría aplicar a cualquier otra variable de interés forestal.

### 3. Metodología

La organización del modelo de desarrollo propuesto abarca las siguientes fases:

- Creación de un espacio de trabajo: genera automáticamente el espacio de trabajo donde se almacenarán, tratarán y analizarán los datos LiDAR.
- Carga de datos LiDAR: el procedimiento muestra cómo cargar los datos en el espacio de trabajo, transformando los archivos \*.las en clases de entidad de una geodatabase compatible con el formato shapefile. A continuación implementa el script para extraer la información de la nube de puntos almacenada en formato binario.
- Normalización de los datos: una vez cargados los datos con todos sus atributos, se aplica este modelo que permite referir los datos de elevación al nivel del suelo.
- Extracción de variables predictoras en las parcelas de campo.
- Ajuste de modelos. En esta parte se referiría al alumno a textos generales de estadística básica en los que se aborden temas de regresión simple o múltiple.
- Aplicación de Modelos y Generación de Mapas. Se calcularían las mismas variables predictoras que se calcularon para las parcelas en cada una de las celdas de una malla regular en la que el tamaño de la celda coincidiría con el tamaño de la parcela. Posteriormente se aplicarían los modelos obtenidos en el paso anterior a cada una de estas celdas.

Para ilustrar cómo sería el trabajo práctico a realizar se desarrolla el siguiente caso de estudio.

### 4. Caso de estudio

En este apartado se va a presentar un caso de estudio similar a lo que se desarrollaría en una clase práctica de 4 horas de duración.

Tanto los datos LiDAR como los datos y los cálculos realizados con el trabajo de campo han sido proporcionados por el Grupo de Investigación, Tecnologías y Métodos para la Gestión Sostenible.

El área de estudio seleccionada es una zona del pinar de Valsaín, localizado en el municipio de San Ildefonso o La Granja, en la provincia de Segovia. Se trata de un área forestal de 303,356 Ha (3,03356 km<sup>2</sup>), extensión equiparable a una unidad de gestión de pequeño tamaño, pero suficientemente grande como para ilustrar la metodología a seguir. Sus coordenadas geográficas en el centro del área de estudio son 4° 1' 42,5" Oeste y 40° 48' 56" Norte. La figura 1 muestra dos mapas de localización:

**Figura 1 Zona de estudio (izquierda) y detalle del MTN25 (derecha).**



A continuación se describe la aplicación del procedimiento propuesto en este trabajo y sus resultados correspondientes.

#### 4.1. Carga de datos

Para establecer el espacio de trabajo nos basamos en la aplicación del modelo de creación de geodatabases (ver figura 2). Para este caso se ha optado por crear una geodatabase de archivo (.GDB), que tiene mayor capacidad de almacenamiento (de 1TB a 256TB).

El modelo comprende también la creación de un conjunto de entidades (Feature Dataset). Uno de los parámetros a establecer es el sistema de coordenadas; para este caso se ha elegido el sistema de coordenadas geográfico WGS 1984 UTM Zona 30 Norte.

**Figura 2 Modelo de creación del espacio de trabajo.**



A continuación se procede a cargar los datos en la geodatabase, aplicando el modelo de carga de datos mostrado en la figura 3. Se ha seleccionado la opción de utilizar la herramienta LAS to Multipoint y ejecutar el script de ESRI Support ExplodeLASMultipoint (Crawford, 2009), mostrado en la figura 4:

Figura 3 Modelo de carga de datos.

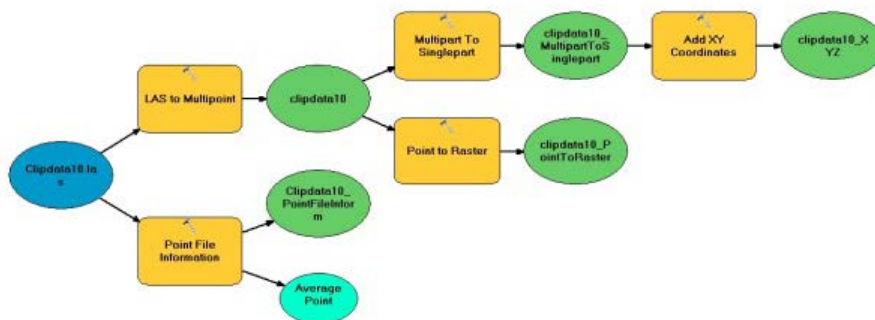
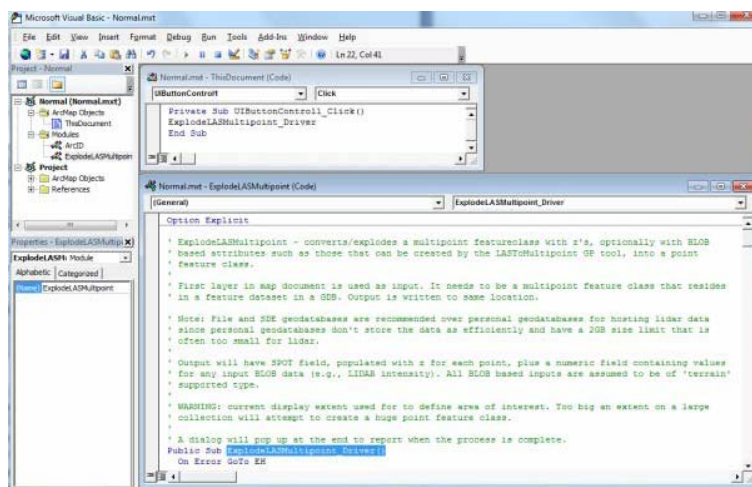


Figura 4 Script ExplodeLASMultipoint en Visual Basic.



Este modelo importa los archivos \*.las, convirtiéndolos en formato shapefile, transformándolos en entidades de puntos y mostrando sus atributos. También calcula las coordenadas de latitud y longitud para cada uno de los puntos. Del mismo modo se ha desarrollado un botón en ArcMap mediante un script importado al editor de Visual Basic. Este botón permite ejecutar el script ExplodeLASMultipoint.

La entidad de multipuntos (ver figura 5) presenta la nube de puntos agrupada en clústers, por lo que el número total de entidades es de 2.111, mientras que el número total de puntos individualizados (ver figura 6) es de 7.127.699. Así mismo los atributos en formato binario de la entidad de multipuntos se extraen para ser representados en la entidad de puntos.

Figura 5 Muestra de la tabla de la entidad de multipuntos.

OID *	Shape *	Intensity	ScanAngle	Return	NumReturns	GpsTime	Class	PointCount
1	Multipoint Z	Blob	Blob	Blob	Blob	Blob	Blob	1698
2	Multipoint Z	Blob	Blob	Blob	Blob	Blob	Blob	1691
3	Multipoint Z	Blob	Blob	Blob	Blob	Blob	Blob	1677
4	Multipoint Z	Blob	Blob	Blob	Blob	Blob	Blob	1614

Figura 6 Muestra de la tabla de la entidad de puntos.

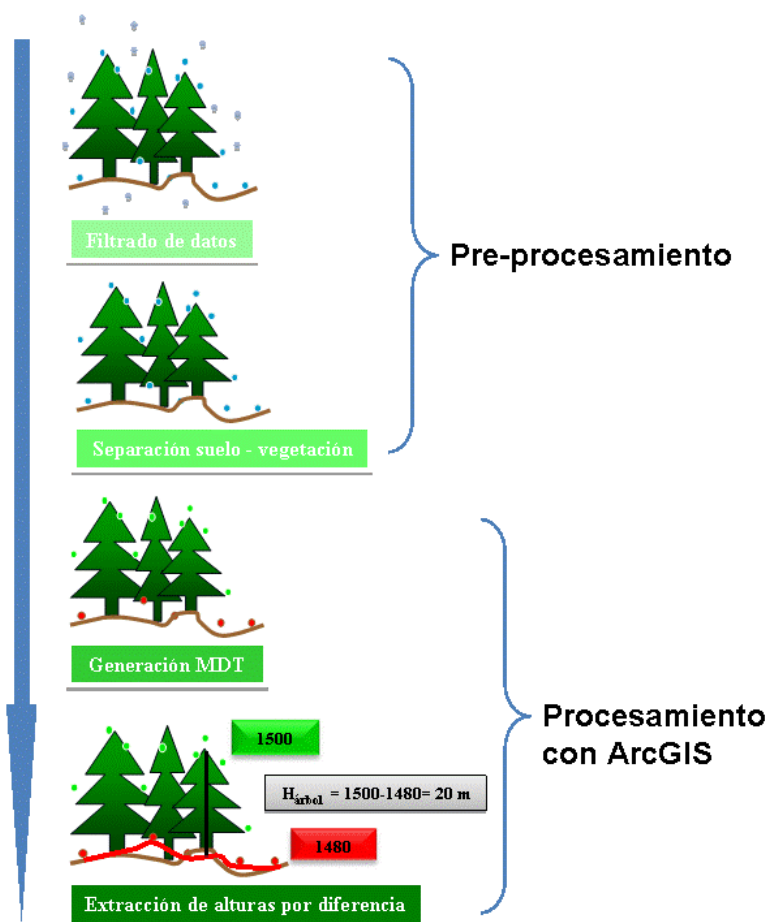
OBJECTID *	SHAPE *	Spot	Intensity	ScanAngle	Return	NumReturns	GpsTime	Class
1	Point Z	1363,0399999999999	0	-14	3	4	227522,280972568	4
2	Point Z	1361,11	52	-14	2	3	227522,280990762	4
3	Point Z	1366,2899999999999	6	-14	1	3	227522,281008913	4
4	Point Z	1366,2	2	-15	1	3	227522,281027107	4

La herramienta Point File Information muestra el rango de valores de elevación entre 1.300,39 y 1.606,14 metros para el conjunto de datos. Así mismo, el valor de distancia media entre puntos es de 64,8429 centímetros.

#### 4.2. Normalización

El proceso de normalización (ver figura 7) de datos permite referir las alturas del conjunto de datos al nivel del suelo mediante la ejecución del modelo.

Figura 7 Modelo de normalización.



Partiendo de todo el conjunto de datos cargados en ArcGIS, el modelo de normalización (ver figura 8) realiza las operaciones necesarias para añadir un nuevo campo a la tabla de la nube de puntos con los datos de elevación normalizados.

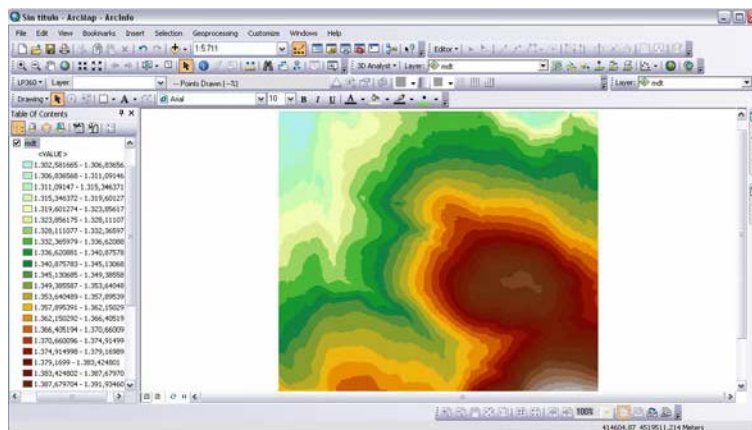
Figura 8 Modelo de normalización.



Para llevar a cabo este proceso, el modelo selecciona los puntos clasificados como suelo, en este caso son 2.522.982 puntos seleccionados como suelo (clase 2 (ASPRS, 2010)).

A continuación crea un modelo digital del terreno (MDT) en formato raster con los valores de elevación sobre el nivel del mar, como se puede observar en la figura 9:

Figura 9 MDT de los puntos clasificados como suelo.



Por último extraen los valores de elevación y se los resta a los valores de la nube de puntos, dando como resultado un campo con los valores de elevación referidos al nivel del suelo (ALTURA), como se muestra en la figura 10:

Figura 10 Muestra de la tabla con los valores de elevación normalizados.

OBJECTID	SHAPE	Spot	Intensity	ScanAngle	Return	IlumReturns	Class	GpsTime	POINT X	POINT Y	POINT Z	RASTERVALU	ALTURA
109403	Point Z	1324,24	78	-10	1	1	2	227707,834417	414229,27	4519597,27	1324,24	1322,470947	1,769053
690138	Point Z	1324,24	78	-10	1	1	2	227707,834417	414229,27	4519597,27	1324,24	1322,470947	1,769053
107961	Point Z	1329,88	164	-11	1	1	2	227707,684463	414252,11	4519588,94	1329,88	1328,472046	1,407954
689738	Point Z	1329,88	164	-11	1	1	2	227707,684463	414252,11	4519588,94	1329,88	1328,472046	1,407954

### 4.3. Análisis forestal

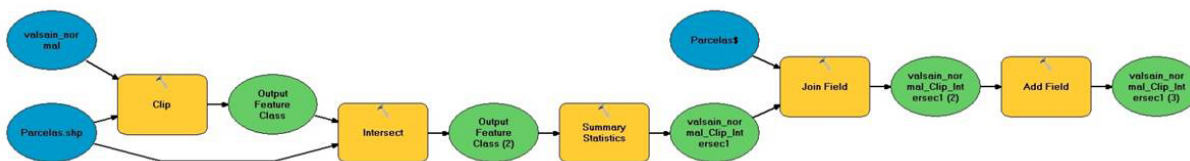
El análisis forestal comprende dos pasos: uno aplicado a las parcelas y otro aplicado a todo el conjunto de datos que conforman el área de estudio.

#### 4.3.1. Extracción de variables predictoras a nivel de Parcela

El análisis forestal aplicado a las parcelas (ver figura 11) realiza los cálculos de ciertas variables estadísticas relativas a la elevación de los puntos en las parcelas, para compararlas con los datos obtenidos en el trabajo de campo.

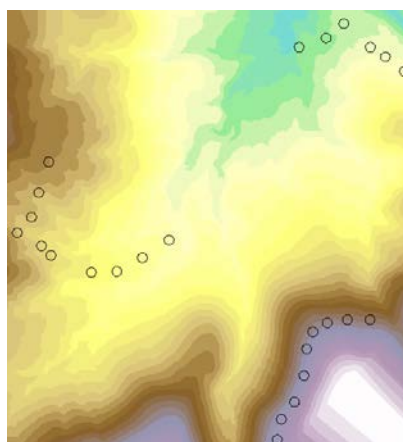


**Figura 11 Modelo de análisis forestal aplicado a las parcelas.**



En primer lugar extrae los datos LiDAR de las parcelas (ver figura 12) y calcula las variables estadísticas del atributo de elevación, que servirán de variables predictoras al elaborar los modelos para la estimación de las variables forestales seleccionadas.

**Figura 12 Muestra de la distribución de las parcelas del área de estudio.**



A continuación carga los datos de biomasa y/o volumen maderable de las parcelas calculados en el trabajo de campo, los une a la tabla de las parcelas y prepara un fichero de datos a utilizar en la elaboración de los modelos para la predicción de las variables de interés.

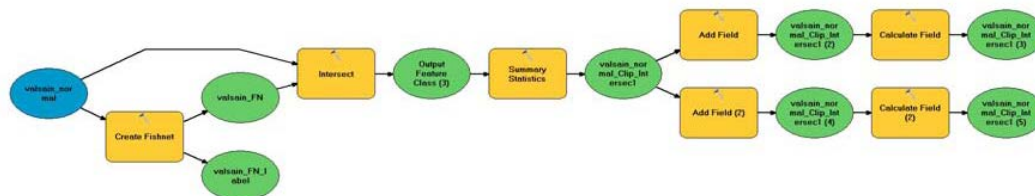
#### **4.3.2. Generación de modelos de regresión para la predicción de variables forestales a partir de variables LiDAR.**

Una vez obtenidas las variables derivadas del LiDAR en las parcelas para las cuales se conocen los valores de las variables forestales a estimar, hay que proceder a ajustar un modelo de regresión en el que éstas últimas quedarían como una función de las primeras. Éste es el único paso para el cual habría que salir de ArcGIS 10, si bien esto no supone ningún problema pues el ajuste del modelo de regresión puede hacerse con cualquier paquete estadístico como R (R Development CoreTeam, 2010), Statgraphics (Statgraphics.net, Statgraphics.net), o similares... Esta parte podría, incluso realizarse con software de ofimática como Microsoft Office Excel u OpenOffice Calc, disponible de forma asequible para cualquier usuario.

#### **4.3.3. Extracción de variables predictoras y aplicación del modelo a toda una zona forestal.**

El análisis forestal del área de estudio (ver figura 13) extrapola los valores obtenidos con el modelo anterior aplicados al área de estudio.

**Figura 13 Modelo de análisis forestal aplicado al área de estudio.**



En primer lugar crea una malla con tamaño de celda igual al tamaño de las parcelas. En este caso se trata de una malla de 2.303 celdas de 35,45 x 35,45 metros (1256,7 m<sup>2</sup>) cada una. Estas celdas se cruzan con la nube de puntos mediante un recorte de la entidad de puntos.

A continuación, para cada celda de la malla, se calculan las mismas variables estadísticas derivadas de la nube de puntos que se calcularon a partir del atributo elevación en las parcelas.

Finalmente se añaden dos nuevos campos en los que se aplican los modelos de predicción, tanto el de volumen forestal como el de biomasa, realizados en el paso anterior.

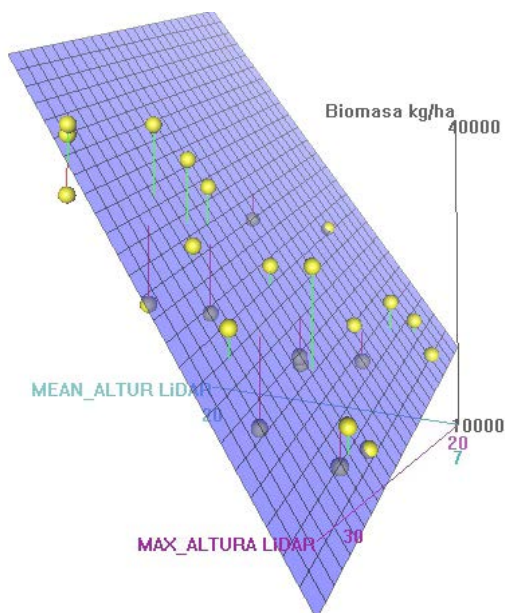
En este caso, los modelos de predicción aplicados son los siguientes:

$$\text{Volumen (m}^3\text{/ha)} = -286.1671564561 + 60.5713592113 * \text{MEAN\_ALTUR} \quad (1)$$

$$\text{Biomasa (Kg/ha)} = 16646,04 + 3014,87 * \text{MEAN\_ALTUR} + -1007,85 * \text{MAX\_ALTURA} \quad (2)$$

Estos modelos se generaron empleando el software Statgraphics. A continuación se muestran las gráficas de regresión de los modelos de predicción, tanto del volumen forestal como de la biomasa forestal (ver figura 14), así como una muestra de la tabla de los datos del volumen forestal y la estimación de biomasa forestal (ver figura 15):

**Figura 14 Gráficos de regresión de los modelos de predicción.**



**Figura 15 Tabla con los datos de volumen forestal y de estimación de biomasa.**

FID	Shape *	OBJECTID *	FID malla valsain	FREQUENCY	MEAN ALTURA	MAX ALTURA	BIOMASA	VOLUMEN
9	Polygon	1	9	2845	7,769441	20,508892	19400,075252	184,438459
10	Polygon	2	10	7662	7,880464	22,914365	17310,447781	191,163287
11	Polygon	3	11	2715	9,232867	22,568896	21735,878002	273,08012
12	Polygon	4	12	2532	9,943399	22,41437	24033,841961	316,118018
13	Polygon	5	13	2402	8,982524	20,140537	23428,605977	257,916545
14	Polygon	6	14	2466	8,509041	19,901694	22241,832804	229,237018
15	Polygon	7	15	2549	7,555011	18,853247	20422,231123	171,450118

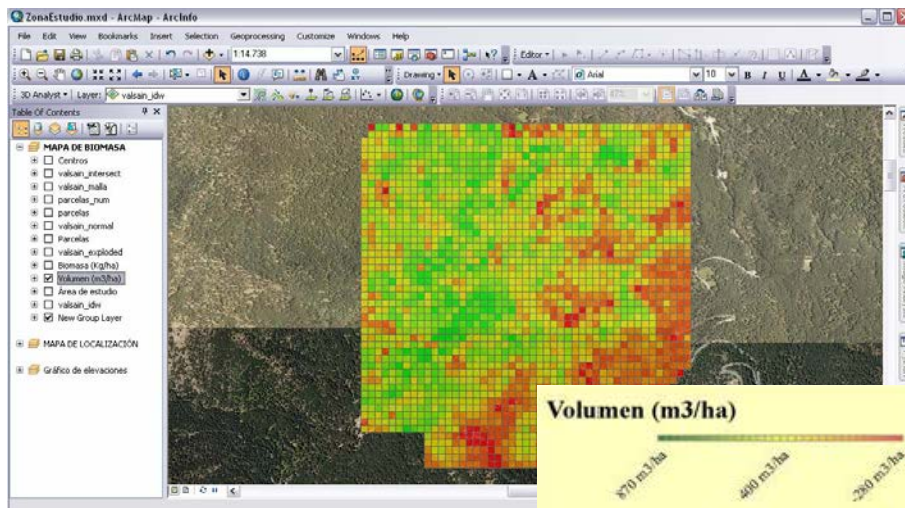
## 5. Resultados

El resultado final de todo este proceso de análisis forestal es el mapa de volumen forestal (ver figura 16) y el mapa de estimación de biomasa (ver figura 17). Estos se representan a través de los atributos volumen y de biomasa calculados con el modelo de predicción en el último análisis.

Como el área de estudio está dividida en celdas, su simbolización se realiza a través del número de celda. En este caso, se ha llevado a cabo una unión de la tabla de atributos de la malla con la tabla que contiene las estadísticas calculadas con la herramienta Summary Statistics.

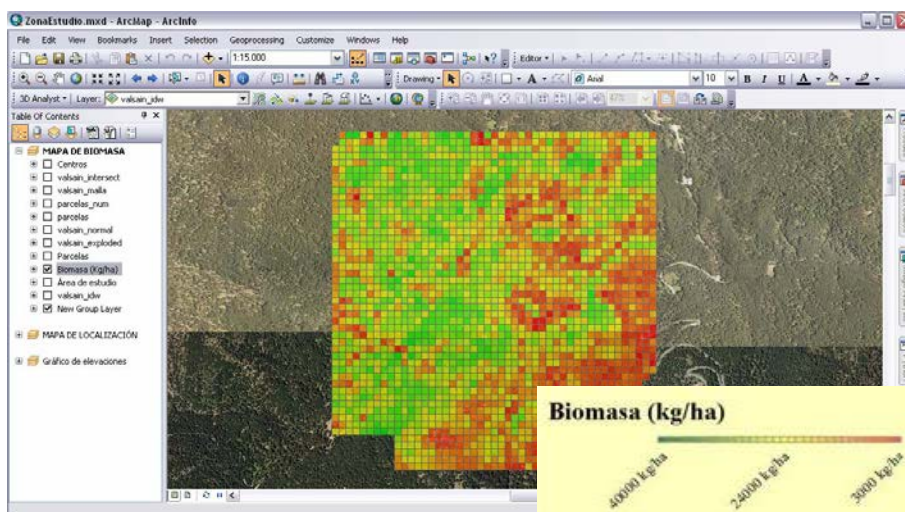
Es importante marcar la opción de mantener solo los registros coincidentes, ya que la malla cubre una zona mayor que la cubierta por los datos LiDAR, por lo que hay celdas de la malla sin información. Una vez unidas las tablas ya se tiene un mapa con estimaciones de las variables forestales. En éste tipo de mapas las celdas suelen quedar coloreadas en función del valor predicho para la variable forestal que se esté analizando.

Figura 16 Volumen forestal en la zona de estudio.



El volumen forestal se calcula en  $\text{m}^3/\text{ha}$ . Los valores más bajos (rojo) son negativos, alcanzando  $-280 \text{ m}^3/\text{ha}$ , mientras que los valores más elevados (verde) alcanzan los  $870 \text{ m}^3/\text{ha}$ .

Figura 17 Estimación de biomasa en la zona de estudio.



La biomasa forestal se calcula en  $\text{kg}/\text{ha}$ . Los valores más bajos (rojo) alcanzan los  $3000 \text{ kg}/\text{ha}$ ; los valores más elevados (verde) alcanzan los  $40000 \text{ kg}/\text{ha}$ .

## 6. Conclusiones

La metodología empleada se basa en el uso de Model Builder para generar procesos automáticos que permitan hacer del tratamiento de los datos LiDAR una tarea ágil y simple.

Además, estos modelos quedan plasmados en diagramas que facilitan al alumno el seguimiento de todo el proceso de datos a realizar durante el ejercicio.

Todos los procesos quedan plasmados en modelos que no dependen de los datos de entrada con lo cual son modelos totalmente aplicables en cualquier zona de estudio.

Todo el tratamiento de la información geográfica se realiza con un único software, el cual es familiar para la gran mayoría de usuarios del sector forestal y cuyo uso está implantado en la mayor parte de las universidades. La ventaja ha sido implementar un modelo que permite realizar el proceso de cálculo de variables independientes (datos LiDAR) desde ArcGIS. El objetivo ha sido sustituir el uso del software FUSION para calcular variables relacionadas con la elevación por un modelo de Model Builder en ArcGIS. Esto está incluido en el modelo de geoprocésamiento forestal.

Entre las aportaciones originales destaca la creación de un botón en ArcMap mediante un script importado al editor de Visual Basic. El botón permite ejecutar el script ExplodeLASMultipoint, que separa las entidades multipuntos en entidades de punto almacenadas en formato shapefile. Estas entidades de punto recogen todos los atributos de los archivos \*.las que se almacenan en formato binario en la entidad de multipuntos. De este modo, el botón se convierte en el paso fundamental del planteamiento de este trabajo: manejar y analizar datos LiDAR con ArcGIS.

## 7. Referencias

ASPRS. (2010). *LAS SPECIFICATION version 1.3-R11*. 5410 Grosvenor Lane, Suite 210

Bethesda, Maryland 20814-2160: The American Society for Photogrammetry & Remote Sensing.

Crawford, C. (2009). ExplodeLASMultipoint.

<http://arcscrips.esri.com/details.asp?dbid=16285>

Environmental Systems Research Institute. (2011). *ArcGIS desktop: Release 10*. Redlands, California USA.

García, M., Riaño, D., Chuvieco, E., & Danson, F. M. (2010). Estimating biomass carbon stocks for a mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, 114(4), 816-830.

Mc Gaughey, R. J. (2011). *FUSION/LDV: Software for LIDAR data analysis and visualization*. Seattle, Washington. USA: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, University of Washington.

- Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 113(1), 148-159.
- Næsset, E., Bollandsås, O. M., & Gobakken, T. (2005). Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sensing of Environment*, 94(4), 541-553.
- Næsset, E., & Gobakken, T. (2005). Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sensing of Environment*, 96(3-4), 453-465.
- R Development CoreTeam. (2010). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Statgraphics.net. (Statgraphics.net). Statgraphics centurion XVI. C/ Bravo Murillo 350, 1º, Madrid, España:

### **Correspondencia:**

Antonio García Abril. ETSI Montes. Departamento de Proyectos y Planificación Rural. Ciudad Universitaria S/N. 28040 Madrid.

Phone: + 34 913366401

(Para más información contacte con):

Antonio San José Albacete

E-mail: [sanjoalbacete@yahoo.es](mailto:sanjoalbacete@yahoo.es)