



---

**Universidad de Valladolid**

Máster

Gestión Forestal basada en Ciencia de Datos/  
Forest Management based on Data Science  
(DATAFOREST)

Herramienta de segmentación  
automática de masas forestales a partir  
de sensores remotos: LiDAR y Sentinel 2

Alumno/a: Fernando Bezares Sanfelip

Tutor/es: Felipe Bravo Oviedo y José Luis Tomé Morán

06/2019



**INDICE**

RESUMEN .....	4
ABSTRACT.....	4
1.- INTRODUCCIÓN .....	5
2.- OBJETIVOS.....	5
3.- MATERIALES Y METODOLOGÍA.....	6
3.1. Área de trabajo .....	6
3.2. Datos de entrada .....	7
3.2.1 LiDAR .....	7
3.2.2. Información LiDAR: Disponibilidad .....	11
3.2.3 Los formatos archivos LiDAR (LAS y LAZ) y su clasificación .....	12
3.2.3 Sentinel 2 .....	14
3.3. Python .....	17
3.4. QGIS .....	17
3.5. METODOLOGÍA. Flujo de trabajo.....	18
3.5.1 Procesado LiDAR.....	19
3.5.2 Procesado de Sentinel .....	21
3.5.3 Procesado Ráster .....	21
3.5.4 Segmentación con Segmentation (Meanshift) OTB.....	24
3.5.5 Suavizado de contornos Chaiken algorithm (GRASS).....	25
4.- RESULTADOS.....	26
5.- CONCLUSIÓN Y DISCUSIÓN .....	31
6.- AGRADECIMIENTOS .....	32
7.- BIBLIOGRAFÍA .....	33
ANEJO 1. RESULTADOS DE LA HERRAMIENTA. SUBPRODUCTOS .....	36
ZONA PILOTO 1: Monte Periurbano en Madrid (Valdelatas).....	36
ZONA PILOTO 2: Sierra de Guadarrama: Peñalara y Cabeza de Hierro.....	42
ANEJO 2. GUÍA DE USUARIO .....	48
A2.1 Presentación .....	48
A2.2 Requerimientos Técnicos .....	48
A2.3 Cómo Acceder A La Herramienta .....	48
A2.4 Estructura.....	52
A2.5 Inputs .....	52
A2.6 Outputs .....	52
A2.7 Recomendaciones.....	53
A2.8 Ejemplo De Utilización De La Herramienta.....	53

## RESUMEN

La delineación de rodales tradicionalmente se ha basado en cartografía y criterio experto de profesionales forestales. La aparición de sensores remotos permite disponer de información que puede ser utilizada para rodalizar de forma semiautomática. Este estudio busca desarrollar una herramienta para automatizar los procesos de segmentación requeridos en la gestión forestal. Para ello, la herramienta se desarrolla en Python 3 y emplea como fuentes de datos, información de LIDAR aéreo e imágenes de Sentinel 2. La metodología empleada, comienza con el procesado de información LIDAR usando el software de FUSION para obtener información de la estructura de la vegetación. A continuación, se combina un índice de vegetación normalizado, generado a partir de la información satelital, con las imágenes de las variables LiDAR. Como resultado, se obtiene un raster multibanda normalizado. Finalmente, el algoritmo de segmentación de Orfeo Tools recibe el raster multibanda, junto con una serie de parámetros de entrada. El resultado del software es un mapa regiones homogéneas del bosque, que no debe ser considerado como una solución definitiva, sino como un apoyo que el gestor puede emplear para determinar unidades de gestión. Para conseguir mejores resultados de segmentación, cada proyecto y tipo de masa requiere una combinación diferente de parámetros, que ha de ser determinada, para cada tipo de masa. Además del archivo de segmentación, las variables raster y otros sus productos se presentan como resultados. Como conclusión, la herramienta que se presenta es una solución moderna, gratuita, rápida, fiable y basada en datos abiertos que combina algoritmos para automatizar las etapas iniciales de un plan de gestión forestal.

**Palabras clave:** LiDAR, QGIS, GIS, Python 3, Sentinel 2, Segmentación de imágenes, programación

## ABSTRACT

Stand typing has traditionally been based on cartography and expert criteria of professional foresters. Remote sensing appearance allows obtaining information suitable for automatizing delineation processes. In light of what is previously exposed, this study aims to develop a tool to automatize stand delineation process required in the management of the woods. To achieve that goal, the tool will be developed in Python 3 using as data sources Airborne LiDAR information and Sentinel 2 spatial images. The methodology used, first, processes the LiDAR information, using the FUSION software algorithm, to obtain rasterized information of the vegetation structure variables. Next, a Sentinel-2 derived NDVI raster file is combined with the LiDAR rasters. As a result, a normalized multiband raster (NMR) is generated. Finally, the Orfeo ToolBox Meanshift algorithm is given the NMR as input along with a set of parameters. The software output, a map of the homogeneous forest regions, should not be considered as an ultimate solution, but as a guideline that the forester could use to finally determine forest stands. In order to get better results, for each project and forest type, a different parameter combination has to be trained for every forest type. In addition to the segmentation shapefile, the raster variable files and other byproducts are presented as results. To conclude, the tool presented consists in a modern, free, open data, reliable and fast solution, that combines, algorithms to automatize the first stage of a forest management plan.

**Key words:** LiDAR, QGIS, GIS, Python 3, Sentinel 2, image segmentation, coding

## 1.- INTRODUCCIÓN

En la actualidad el uso de herramientas informáticas y la automatización de procesos manuales o mecánicos, tanto en la academia como en empresas es una realidad (World Economic Forum, 2018). La creciente preocupación medioambiental y el desarrollo de políticas de gestión sostenible, ha aumentado la demanda de tecnologías de teledetección que sean capaces de proporcionar información y datos para poder cumplir dichas políticas y enfrentar los retos medioambientales. Un ejemplo de este uso, es el empleo de Sistemas de Información Geográfica (SIG) y de sensores remotos para ayudar la gestión forestal (White *et al.*, 2016).

En el ámbito forestal la incorporación del LiDAR (*Light Detection And Ranging*) como fuente de datos, ha supuesto una gran revolución para la monitorización de las masas forestales. El LiDAR aéreo (o *Aerial Laser Scanner*, ALS) ha probado ser efectivo para la adquisición de variables necesarias para el inventario de existencias forestales, principalmente a partir de la altura con errores entre 0.5 a 1m ( Næsset & Økland, 2002, Wulder *et al*, 2008). El ALS también puede resultar de interés para obtener otras variables como: volumen, fracción de cabida cubierta o biomasa (Lim *et al*, 2003).

La aparición del LiDAR no es la única tecnología que ha contribuido a facilitar la gestión de las masas forestales, ya que el desarrollo en la carrera espacial y en la tecnología satelital ha permitido, tanto mejorar las telecomunicaciones, como obtener una mayor información sobre nuestros bosques. Esta información es principalmente proporcionada a través de imágenes satelitales obtenidas en las misiones y programas espaciales desarrollados por las agencias espaciales. La Agencia Europea del Espacio (habitualmente conocida por su acrónimo en inglés: ESA de *European Spatial Agency*) en su programa espacial Copernicus está desarrollando el programa de satélites Sentinel con el objetivo de monitorizar la tierra y vigilar diferentes aspectos ambientales: calidad de las aguas, limpieza atmosférica, deforestación, erosión, tasa de deshielo, variaciones térmica (*European Spatial Agency*, ESA, 2015) .En este estudio nos centraremos en la información proporcionada por el grupo de satélites Sentinel 2 para el cálculo del índice de vegetación normalizado (NDVI) que ofrece información acerca de la actividad fotosintética de una región (Gandhi *et al*, 2015).

Estas dos tecnologías aunadas con el desarrollo de los sistemas de información geográfica han facilitado el análisis de diferentes tipos de información espacial en una misma plataforma, permitiendo el desarrollo de gran cantidad de estudios. Hoy en día el uso de la información generada por estas plataformas en (SIG) es prácticamente indispensable para el desarrollo de proyectos de ingeniería.

## 2.- OBJETIVOS

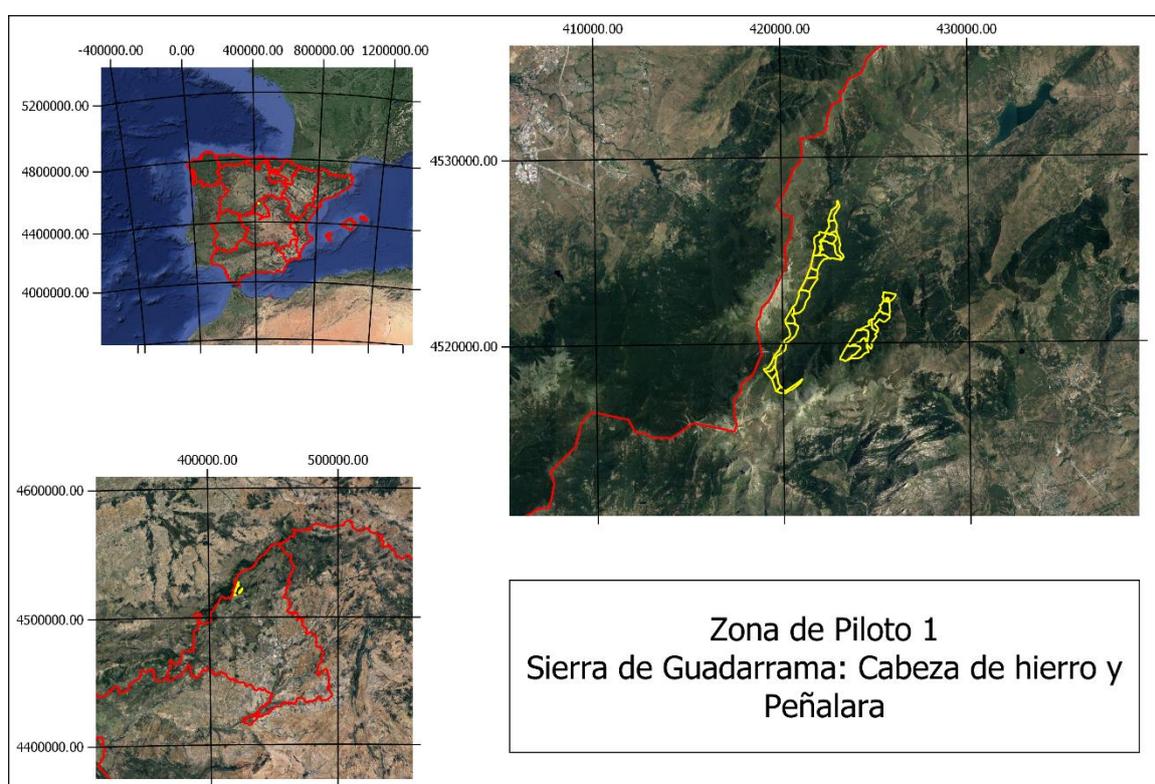
El objetivo de este proyecto es contribuir al estudio de las masas forestales creando una herramienta de segmentación automática de “masas homogéneas” a partir de información en abierto proporcionada por sensores remotos. Esta herramienta buscará las zonas homogéneas en masas forestales, contribuyendo a hacer una clasificación previa de tipos de masa, que facilite la preparación del trabajo de campo para una delineación final de tipos de masa.

### 3.- MATERIALES Y METODOLOGÍA

#### 3.1. Área de trabajo

Con objeto de poner a punto la herramienta se han seleccionado dos zonas piloto con distintas características: una zona eminentemente forestal de la Sierra de Guadarrama (Madrid) y un monte periurbano en el entorno de la ciudad de Madrid.

La primera zona piloto se encuentra enmarcada dentro de la Comunidad de Madrid, en concreto en la zona norte de la Sierra de Guadarrama: Cabecera del valle de Lozoya, montes de Cabeza de Hierro y Peñalara (Figura 1). Se trata de un bosque de coníferas de alta montaña dominada principalmente por *Pinus sylvestris*; L. Las formaciones vegetales de esta zona piloto se caracterizan por ser bosques monoespecíficos cerrados (fracciones de cabida cubierta elevada) en los cuales sus pies alcanzan alturas alrededor de los 20-25 metros. Pudiendo darse un subpiso de *Pteridium aquilinum* con pies aislados de alguna especie acompañante como *Taxus baccata*, *Ilex aquifolium*.



**Figura 1.** Zona piloto 1. Comunidad de Madrid. Sierra de Guadarrama: Peñalara y Cabeza de Hierro.

La zona piloto 2 es un monte mediterráneo mixto de *Pinus halepensis*. L. y *Quercus ilex ballota* emplazado al norte de la ciudad de Madrid (Valdelatas). En esta zona piloto a diferencia de la primera, se trata de un bosque de menores densidades y alturas, con espacios abiertos y dos estratos claramente diferenciados entre el pino (dominante) y la encina.

Se eligieron estas zonas debido a que Agresta S.Coop. ya había trabajado previamente sobre el terreno y poseen unas rodalizaciones realizadas manualmente combinando estas mismas fuentes de datos. Esto nos resulta de gran interés ya que estas rodalizaciones previas sirven como una verdad terreno sobre la que comparar nuestros resultados obtenidos con la segmentación automática. Con esta comparativa se apreció efectivamente, que existe un sentido físico y real entre el resultado de la segmentación y las rodalizaciones previas.

Finalmente, realizar el estudio sobre estas dos zonas nos permitirá observar cómo se comporta la herramienta frente a dos tipologías de masas completamente diferentes, pudiendo así recomendar configuraciones de la herramienta de segmentación para estas dos tipologías de masa.

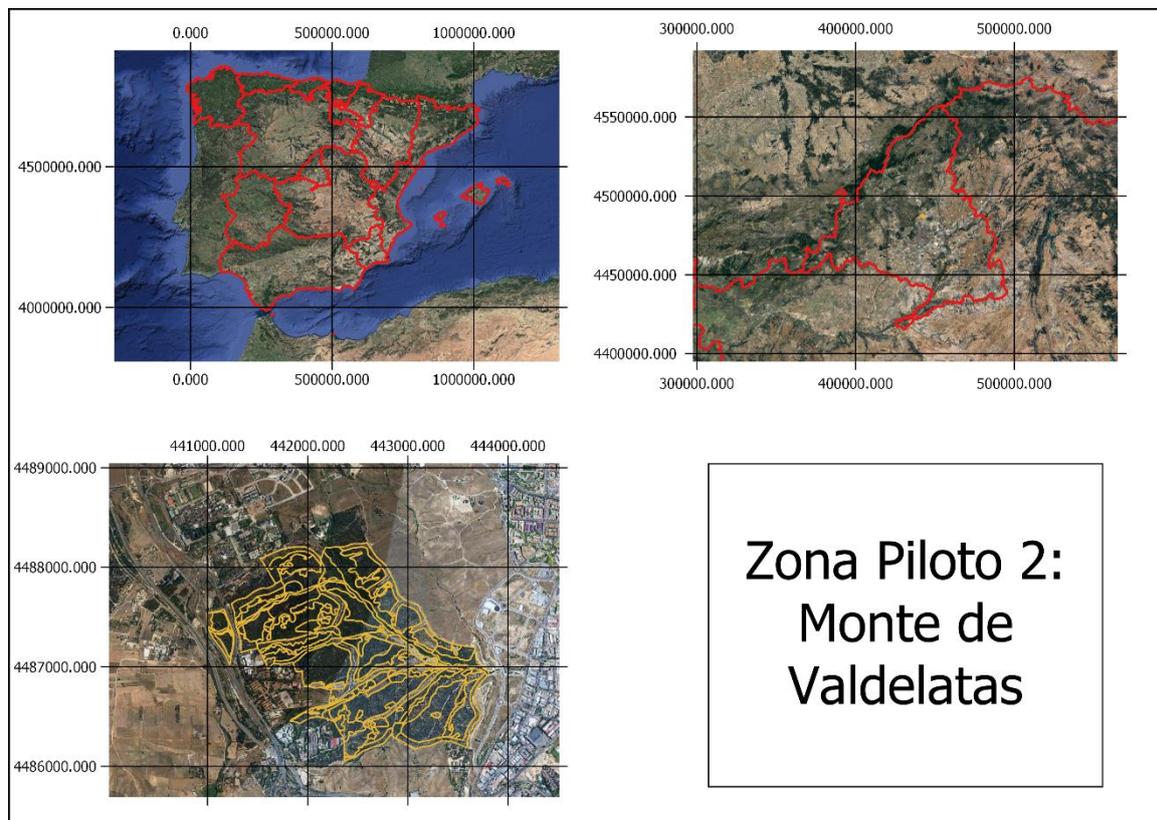


Figura 2. Zona piloto 2. Bosque periurbano: Monte de Valdelatas.

## 3.2. Datos de entrada

### 3.2.1 LiDAR

El LiDAR es una tecnología que utiliza haces de radiación electromagnética, emitida en pulsos, para obtener la información de posición de un objeto o superficie. Originariamente el LiDAR fue desarrollado para tecnología militar en los años 60 del siglo XX (Nelson, 2013) siendo posteriormente utilizado (1970's) en proyectos topográficos para la creación de modelos de superficies. Finalmente, en el mundo forestal esta tecnología fue introducida en la década de los 80's. Esta trayectoria histórica de más de 30 años la presenta ya como una herramienta consolidada e imprescindible en el sector forestal.

El LiDAR es un sensor activo (Figura 2), es decir, emite su propia radiación electromagnética y no hace uso de fuentes de radiación externa (pasivo) como puede ser un sensor óptico. Al ser activo nos ofrece la ventaja de poder operar en lugares no expuestos a radiación solar o incluso de noche.

Para obtener la posición de un punto, esta herramienta calcula el tiempo que tarda en volver el haz de luz pulsada, desde que sale del emisor hasta que se refleja en el objeto de interés y vuelve hacia el emisor que lo recibe (Tomé *et al*, 2013). En cada pulso, al colisionar el haz con diferentes elementos que encuentre, genera un retorno que puede quedar o no registrado en el receptor. Por lo tanto, **un único pulso puede tener varios retornos** que ofrecen información de los objetos que el haz ha encontrado en su camino. Por ejemplo, cuando un haz atraviesa un árbol, su primer retorno será aquel que impacte

contra la copa o la parte superior, posteriormente se espera que el segundo retorno sea aquel de un estrato inferior, una rama el tronco. Así se continuaría hasta llegar al suelo o hasta que el haz pierda su energía.

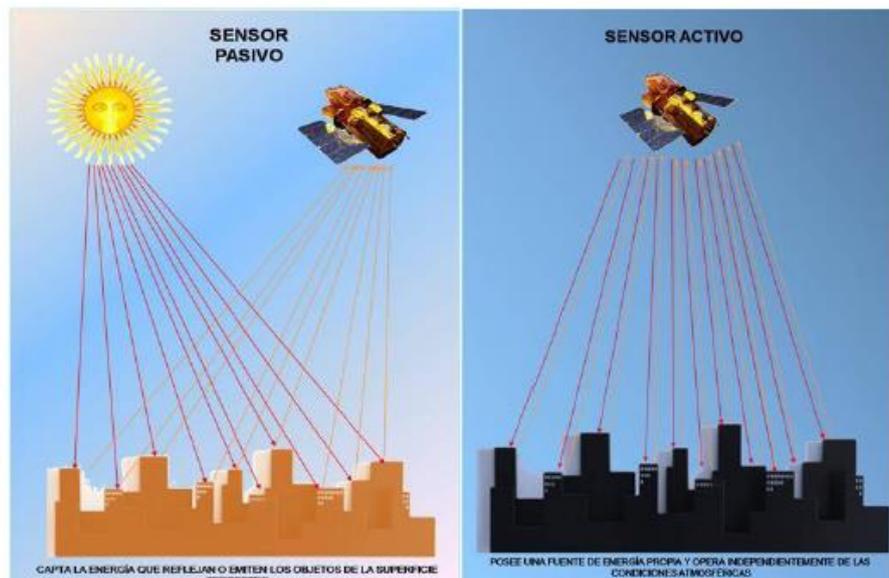
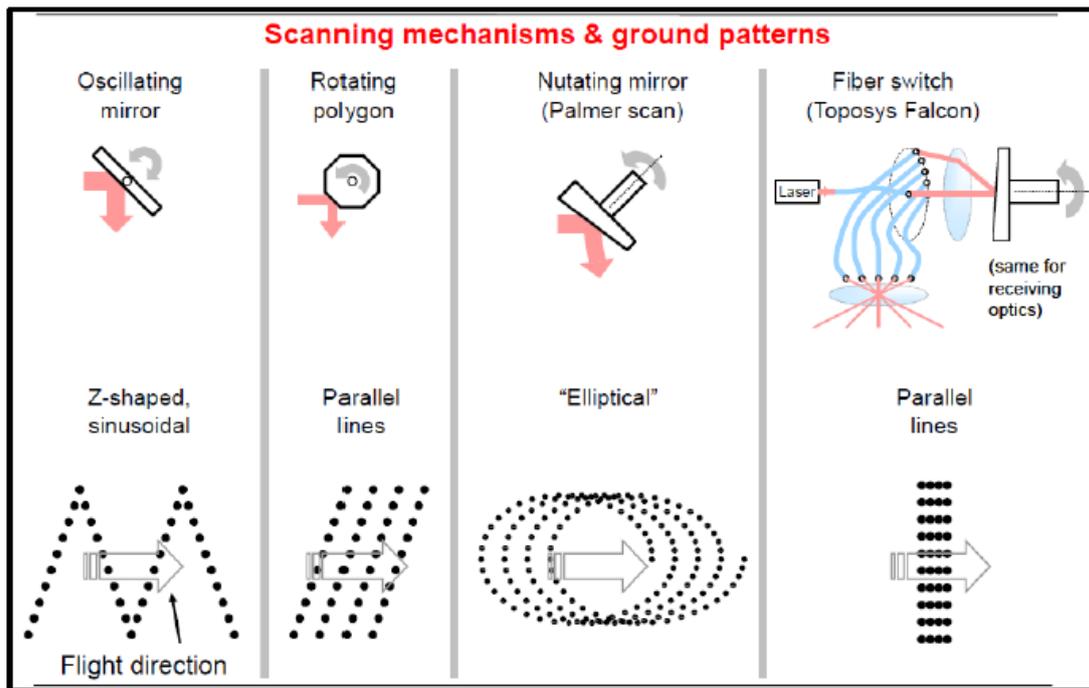


Figura 3. Comparativa entre los sensores activos y pasivos. Fuente (INEGI, 2013)

Una posible clasificación de los sistemas de escaneo LiDAR se basa en las posibles plataformas que portan el sensor, en este sentido el LiDAR aerotransportado o ALS es el más utilizado en los estudios forestales de grandes superficies, por lo que en este estudio nos centraremos en este tipo de LiDAR.

La tecnología ALS está pensada para la recogida masiva de datos en la cual el LiDAR se embarca en una aeronave desde la cual se toman los datos. Está compuesta por 4 elementos principales:

- **El propio emisor/receptor de haces de luz pulsada o sensor láser.** Existen diferentes tipos de emisores laser en función de sus mecanismos internos de reflexión de la radiación y de los patrones de puntos obtenidos a medida que avanza la aeronave. Estos quedan sintetizados la figura 3. Otra cualidad a indicar de sensor laser es que los pulsos emitidos por este, se tratan principalmente de radiación infrarroja cercana (1064nm).



**Figura 4.** Patrones de datos y mecanismos internos de un sensor laser en LiDAR aerotransportado. Fuente Baltsavias, (2008)

- **Un sistema de inercial de navegación (INS):** informa acerca de la trayectoria de navegación y giros del avión para posteriormente ser tenidos en cuenta en la formación y corrección de los de los datos.
- **Un sistema de posicionamiento global (GPS)** que obtiene la posición en todo momento y se la transmite al ordenador de a bordo.
- **Ordenador de abordo:** posee un sistema de cronometraje de alta precisión, con el que calcula la posición (x, y, z) de cada punto intersecado por el haz emitido.

La densidad de puntos obtenidos en un escaneo LiDAR es sensible a diferentes parámetros y depende de la configuración del equipo ALS, pero no solo eso, sino que la configuración y diseño del vuelo también puede ser decisiva en las características de dicha información. Las principales variables que afectan a esta información son:

- **Sistema de escaneo,** presentados anteriormente.
- **Altura y velocidad del vuelo,** inversamente proporcionales a la densidad de puntos. Un incremento en altura o velocidad supondrá una disminución de la densidad de puntos.
- **Frecuencia de pulsos,** directamente proporcional con la densidad de puntos, más frecuencia se traducirá en una mayor densidad de puntos
- **Ángulo de escaneo:** Angulo máximo del láser respecto al NADIR, es decir el ángulo máximo respecto a la perpendicular formada por el aparato y la superficie terrestre. Condiciona el **FOV (Field of View)** o campo de visión a menos ángulo menos superficie entra en la imagen.

Es necesario explicar también el concepto de **precisión** de los sensores LiDAR. Esta precisión se puede descomponer en dos tipos:

- **Vertical:** depende del objeto contra el que colisione el pulso. Este valor puede variar en función de la dureza y homogeneidad de la superficie de contacto. El Plan Nacional de Ortofotografía Aérea (PNOA), a partir del cual extraemos los datos, garantiza 20 cm de precisión vertical:

- $\pm 0.05$  m en superficies duras y homogéneas del terreno abierto.
- $\pm 0.10$  m en superficies blandas/vegetación (terreno plano de rodadura).
- **Horizontal:** condicionada principalmente por el relieve de la zona de trabajo.
  - $\pm 0.10 - 0.20$  m en todos los terrenos para PNOA (Instituto Geográfico Nacional (IGN), 2018). Estos valores pueden modificarse en terrenos muy montañosos, donde además la altura de vuelo y la divergencia del haz también son determinantes.

Otra de las características a destacar del LiDAR es que la radiación electromagnética emitida no tiene el mismo comportamiento con todas las superficies. Esta característica nos ayuda a la vez a obtener más información sobre la superficie de estudios. Caben destacar tres casos que en el ámbito forestal nos son de gran ayuda:

- Las superficies solidas generan un único retorno. Ejemplo zonas desarboladas o suelo desnudo
- El agua no genera un retorno.
- En superficies permeables como cubiertas forestales podemos obtener más de un retorno.

Esta última característica resulta de gran interés para el ámbito forestal. Sin embargo, debido a la pérdida de energía, no siempre todos los pulsos penetran totalmente la vegetación hasta llegar al suelo. En una superficie boscosa los pulsos van penetrando en la masa cada vez con menos energía, por lo cual la proporción de primeros retornos frente al total es siempre muy elevada. Pese a tener una baja proporción con respecto al total de pulsos emitidos, esta capacidad del LiDAR de penetrar bajo el dosel de copas es fundamental para el estudio de la estructura vertical de la vegetación. En la siguiente figura elaborada por MacGaughey y sus colaboradores (2006), podemos observar la pérdida progresiva de energía al penetrar los pulsos LiDAR en la masa, donde el 69 % de los retornos obtenidos son primeros retornos, siendo los cuartos retornos apenas un 1 % del total.



Figura 5. Densidad de pulsos que atraviesan el dosel arbóreo. Fuente: (MacGaughey, Andersen, & Reutebuch, 2006)

Como ya hemos observado el LiDAR es una tecnología de captura masiva de datos. Sin embargo, para poder explotar adecuadamente esta información es necesario hacer un procesamiento de los mismos. Existen tres tipos fundamentales de software utilizables para trabajar con datos LiDAR:

- **Captura:** tras la captura de los datos, un programa, localizado en el ordenador los sensores de captura, genera los datos brutos, los almacena y exporta a herramientas que permiten su visualización.
- **Proceso:** esta fase se realiza sobre los datos brutos generados en la fase anterior. el objetivo es transformar, corregir, exportar o reducir el tamaño de los datos en bruto y prepararlos para la fase de análisis.
- **Análisis:** en esta fase se obtiene información cuantitativa adicional de los a partir de los datos procesados.

### 3.2.2. Información LiDAR: Disponibilidad

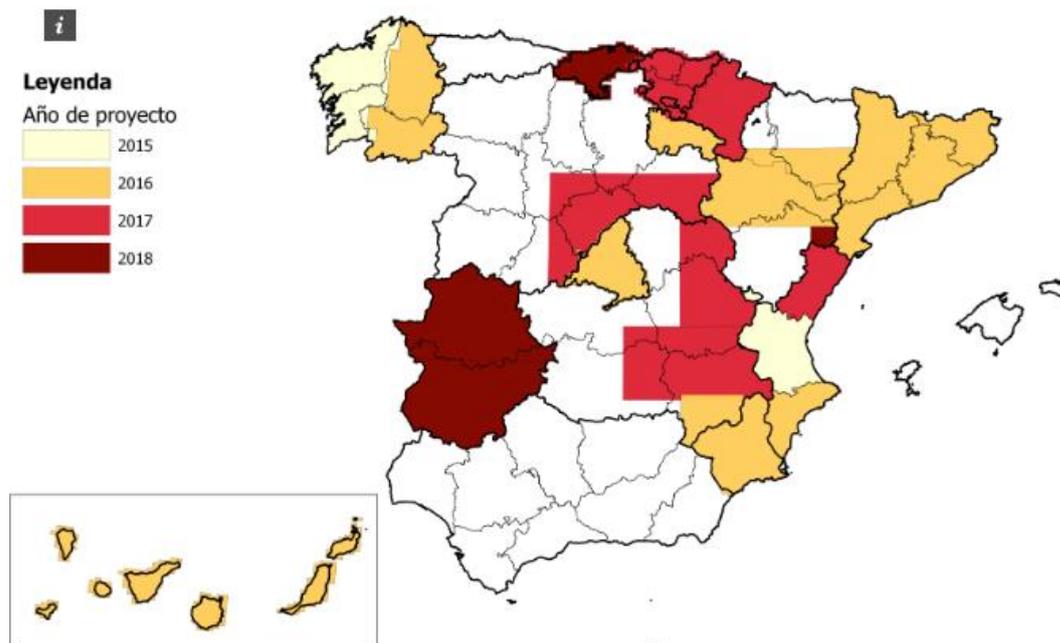
A nivel nacional, hasta hace poco tiempo la disponibilidad de datos LiDAR era escasa y en la gran mayoría de las situaciones no era gratuita. Esta situación cambió drásticamente a partir de 2008 con iniciativas como las del País Vasco en las que se sobrevoló toda la comunidad autónoma proporcionando esta información. Fue ese mismo año cuando el IGN (Instituto Geográfico Nacional) motivado por una demanda europea para elaborar cartografía de zonas inundables, incluyo dentro del Plan Nacional de Ortofotografía Aérea (PNOA) un programa de captura de datos LiDAR para el desarrollo de modelos digitales de elevación de alta precisión. La realización de este proyecto fue posible gracias a la cofinanciación de la Administración General del Estado y las Comunidades Autónomas, bajo la coordinación del IGN.

Los datos empleados LiDAR del PNOA se distribuyen como el resto de productos cartográficos generados por el IGN a través de la web de descarga del CNIG (Centro Nacional De Información Geográfica). El producto LiDAR **proporcionado por esta web se distribuye en ficheros de 2x2 km de extensión en formato “.laz” con un tamaño de pixel de 25 cm. El sistema geodésico de referencia es el ETRS89** en la Península Ibérica (Instituto Geográfico Nacional (IGN), n.d.). En la actualidad se ha finalizado la primera cobertura y estamos en fase de realización de la segunda cobertura del Programa PNOA LiDAR, con diferentes características en sus datos tal como se puede apreciar en la figura 5:

**Tabla 1.** Especificaciones técnicas de los vuelos PNOA LiDAR. Extraído en junio de 2019 de: Fuente: <http://pnoa.ign.es/especificaciones-técnicas>

PNOA- LiDAR	Densidad(p /m2)	Distancia entre puntos (m)	Precisión altimétrica de la nube de puntos	Precisión altimétrica de Modelo Digital del Terreno
1ª Cobertura	0.5	1.4	RMSE z ≤ 0.2 m	RMSE z ≤ 0.5 m
2ª Cobertura	1	1	RMSE z ≤ 0.15 m	RMSE z ≤ 0.5 m

## SEGUNDA COBERTURA



**Figura 6.** Disponibilidad de los Datos LiDAR de la 2 cobertura. Extraído en Marzo 2019 de : <http://pnoa.ign.es/estado-del-proyecto-lidar>

Dado que las zonas de piloto seleccionadas en este estudio se encuentran en la Comunidad de Madrid, se han utilizado datos correspondientes a la segunda cobertura del 2016 (Figura 6).

### 3.2.3 Los formatos archivos LiDAR (LAS y LAZ) y su clasificación

Los archivos LAS (LAZ es la versión comprimida) es un tipo de archivo diseñado para contener información proveniente de LiDAR u otras fuentes que proporcionen datos en forma de nube de puntos. Este formato nace como propuesta de un formato abierto para el tratamiento de información LiDAR. Los archivos combinan la información procedente de un GPS de la unidad inercial de navegación y del láser para generar tripletas (x, y, z) de coordenadas (The American Society for Photogrammetry & Remote Sensing, 2013).

Los archivos LAS contienen datos binarios están formados por un bloque de encabezado, una cantidad de registros de longitud variable, los registros de los datos de tipo punto, y un número cualquiera de registros de longitud variable extendida. Los registros de longitud variable contienen diferentes tipos de datos normalmente, información acerca de la proyección, metadatos, información de usuario (The American Society for Photogrammetry & Remote Sensing, 2013). Por otro lado, los registros de longitud variable extendida, permiten incluir una mayor cantidad de datos y ser añadidos al final de un archivo LAS, pudiendo indicar su proyección por ejemplo sin tener que modificar la capa.

Los registros de los datos de tipo punto poseen diversos formatos registrados de 0 a 10. Los formatos preferidos se encuentran entre el 6 y el 10. Estos formatos han sido mejorados de tal forma que son capaces de soportar 256 clases y e incluyen un slot de bits para puntos solapados. Estos formatos contienen 30 bits (los formatos 0-5 contienen 20 bits), lo cual, permiten capturar mayor número de retornos (hasta 15), mayor número

de clases de puntos, aportan mayor precisión del ángulo de escaneo y, finalmente, es obligatorio indicar el tiempo de captura de los datos. Los puntos están calificados de acuerdo al siguiente estándar definido por la sociedad americana de fotogrametría y teledetección:

**Tabla 2.** Clases estándar ASPRS para las clases de punto para los formatos de datos tipo punto 6 -10. Formato de archivo LAZ 1.4. Fuente:(The American Society for Photogrammetry & Remote Sensing, 2013)

Valor de clasificación	Significado
0	Creado, sin clasificación
1	Sin clasificar
2	Suelo
3	Vegetación baja
4	Vegetación media
5	Vegetación alta
6	Edificios
7	Puntos bajos (ruido)
8	Reservado
9	Agua
10	Railes y líneas ferroviarias
11	Carreteras
12	Reservado
13	Alambre- Conductor (Fase)
14	Alambre- Protector (Escudo)
15	Torre transmisión
16	Conector de estructura de alambre (ejemplo: aislante)
17	Cubierta de puentes
18	Ruido Alto
19-63	Reservado
64-255	Definible por el usuario

El producto que ofrece el CNIG (nuestra fuente de datos) aporta los datos clasificados por defecto en formato LAZ 1.3, La versión anterior a la indicada en la tabla 2. No obstante, una de las principales diferencias entre la versión 1.3 y las 1.4 es el aumento de número de bits del formato, permitiendo un aumento en el número de clases para la clasificación de los puntos. Los archivos estarán clasificados siguiendo la tabla 3, de tal forma que, si se habla de los puntos de clase 2, nos referiremos a los puntos de suelo. Si nos referimos a los puntos de las clases 3-5 serán puntos de la vegetación.

**Tabla 3.** Clases estándar ASPRS para las clases de punto. Formato de archivo LAZ 1.3.(The American Society for Photogrammetry & Remote Sensing, 2009)

Valor de clasificación	Significado
0	Creado, sin clasificación
1	Sin clasificar
2	Suelo
3	Vegetación baja
4	Vegetación media
5	Vegetación alta
6	Edificios
7	Puntos bajos (ruido)
8	Punto clave del Modelo (punto de masa)
9	Agua
10	Reservado para definición de ASPRS
11	Reservado para definición de ASPRS
12	Puntos de solape
13-31	Reservado para definición de ASPRS

### 3.2.3 Sentinel 2

El Sentinel 2 es un satélite perteneciente a la familia Sentinel, desarrollada para cumplir exclusivamente las necesidades técnicas y objetivos requeridos por el programa espacial Copernicus de la Agencia Espacial Europea (ESA). En este programa es una actualización del antiguo programa GMES (*Global Monitoring for Environment and Security*). El proyecto Copernicus busca, al igual que hacía el GMES, la monitorización de la tierra y servir ante situaciones de emergencia y seguridad ciudadana (European Spatial Agency, ESA, 2015). Respecto a la monitorización del planeta, el proyecto se centra específicamente en: gestión del territorio, medio marino, dinámica atmosférica y calidad y cambio climático. Además, cumple otras misiones como por ejemplo proporcionar información para prácticas agrícolas, forestales, gestión de la seguridad alimentaria. Esta información es esencial para la observación de la evolución de la vegetación en el planeta y todas aquellas aplicaciones relativas al estudio de la vegetación, ya sea a nivel de producción, sanidad, paisaje etc...

Hoy en día están en órbita 4 constelaciones de la familia Sentinel: Sentinel1, Sentinel 2 y Sentinel 3 formadas por 2 satélites (A y B) cada una, e independiente a estas constelaciones, un satélite Sentinel 5P(European Spatial Agency (ESA), 2019a). Estas constelaciones y sus objetivos son, de acuerdo con lo propuesto por la (European Spatial Agency (ESA), 2019b):

- Sentinel 1: Su objetivo es proveer de imágenes radar tanto de día como de noche.
- Sentinel 2: Su objetivo es **la distribución de imágenes ópticas de gran resolución destinados a servicios terrestres**. El Sentinel 2-A fue puesto en órbita el 23 de junio de 2015. El S2-B que posteriormente emitido el 7 de marzo de 2017, con el mismo objetivo que su gemelo anterior. Serán los satélites con cuya información operaremos.
- Sentinel 3: destinados al monitoreo del océano y la superficie terrestre.

- Sentinel 4-5: proporcionarán información acerca de la composición atmosférica el Sentinel 4 desde una órbita geoestacionaria y Sentinel 5 desde una polar.
- Sentinel 5P: Esta destinado a reducir los huecos de datos entre el Envisat, tomados por el espectrómetro SCIAMACHY<sup>1</sup>, y la emisión del Sentinel 5.

Puesto que el Sentinel 2 es la constelación elegida para desarrollar este proyecto, vamos a desarrollar de forma más detallada sus objetivos y características. En cuanto a sus objetivos:

- Busca sistematizar la adquisición de imágenes multiespectrales de gran resolución aumentando la frecuencia en el periodo de revisita, es decir el tiempo en el que el satélite vuelve a barrer por el mismo punto.
- Obtener una continuidad con las imágenes multiespectrales proporcionadas por las series de satélites SPOT y USGS LANDSAT.
- Proporcionar datos para futuros productos desarrollables a partir de imágenes multiespectrales.

Respecto a sus características, el Sentinel 2 está diseñado para tener un periodo de vida de 7.25 años, aun así, sus baterías y propulsores están preparados para afrontar 12 años de operaciones. Los dos satélites de esta constelación orbitan de forma helio-sincrónica desfasados 180° uno respecto del otro. La altura media de orbita es de 756 km.

Los satélites de esta constelación vienen equipados con un sensor multiespectral MSI (*MultiSpectral Instrument*). EL MSI funciona de forma pasiva recogiendo la radiación electromagnética a medida que el satélite se desplaza. Una vez recoge el rayo de luz, este es dividido en dos planos focales gracias a un filtro aplicado sobre los receptores. En uno se recoge la franja del espectro visible y el Infrarrojo cercano (NIR), mientras que en el otro se obtiene la radiación infrarroja de onda corta (SWIR). La configuración del telescopio del MSI permite una amplitud de campo (FOV, *Field Of View*) de 290 km.

El MSI recoge la radiación reflejada por la tierra en 13 bandas espectrales desde el visible e infrarrojo cercano (VNIR) hasta el SWIR. Esta información queda recogida en la siguiente tabla:

---

<sup>1</sup> espectrómetro de imágenes cuyo objetivo principal es realizar mediciones globales de gases traza en la troposfera y en la estratosfera.

**Tabla 4.** Bandas espectrales para los MSI de Sentinel 2. Fuente: (European Spatial Agency (ESA), 2019c). Extraída en Febrero 18, 2019, de : <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>.

Bandas	S2A		S2B		
	Longitud de onda central (nm)	Ancho de banda (nm)	Longitud de onda central (nm)	Ancho de banda (nm)	Resolución espacial (m)
1	442.7	21	442.2	21	60
2	492.4	66	492.1	66	10
3	559.8	36	559	36	10
4	664.6	31	664.9	31	10
5	704.1	15	703.8	16	20
6	740.5	15	739.1	15	20
7	782.8	20	779.7	20	20
8	832.8	106	832.9	106	10
8a	864.7	21	864	22	20
9	945.1	20	943.2	21	60
10	1373.5	31	1376.9	30	60
11	1613.7	91	1610.4	94	20
12	2202.4	175	2185.7	185	20

**Tabla 5.** Descripción de las bandas en las que se muestra la correspondencia de cada una de ellas y cuál es su longitud de onda central. Fuente: (Valcarcel, 2016)

Banda	Resolución (m)	Longitud de onda central (nm)	Descripción
B1	60	443	Ultra azul (Costa y Aerosol)
B2	10	490	Azul
B3	10	560	Verde
B4	10	665	Rojo
B5	20	705	Visible e Infrarrojo Cercano (VNIR)
B6	20	740	Visible e Infrarrojo Cercano (VNIR)
B7	20	783	Visible e Infrarrojo Cercano (VNIR)
B8	10	842	Visible e Infrarrojo Cercano (VNIR)
B8a	20	865	Visible e Infrarrojo Cercano (VNIR)
B9	60	940	Onda Corta Infrarroja (SWIR)
B10	60	1375	Onda Corta Infrarroja (SWIR)
B11	20	1610	Onda Corta Infrarroja (SWIR)
B12	20	2190	Onda Corta Infrarroja (SWIR)

Los datos proporcionados por los servicios de Sentinel son una recopilación de “gránulos” o teselas, que **son orto fotografías en UTM/WGS84 de 100x100 km<sup>2</sup> de tamaño** (European Spatial Agency (ESA), 2019d). Al estar proyectados en UTM quiere decir que la tierra está dividida en 60 zonas. Cada zona posee 6° de ancho longitudinal y 8° de ancho latitudinal. Dependiendo de la banda seleccionada poseeremos una resolución

---

espacial (tamaño de pixel diferente), en nuestro estudio, al **emplear la B8 y B4 las resoluciones son de 10 m.**

### 3.3. Python

Python es un lenguaje de programación creado por Guido van Rossum en 1991. En su experiencia con el sistema operativo Amoeba, y basándose en parte de la sintaxis del lenguaje ABC y el manejo de errores de lenguajes como Modula 3, decidió crear un lenguaje que mejorara (entre otras características) la extensibilidad de lenguajes como ABC, con los que trabaja en dicho sistema operativo. Pensó que crear un lenguaje específico para Amoeba no arreglaría el problema, por lo que priorizó la extensibilidad del lenguaje, creando Python de 1989 a 1991. Anecdóticamente, el nombre le viene dado del legendario grupo cómico inglés de los 70's "Monty Python's Flying Circus".

Python es un lenguaje de programación orientada a objetos, interactivo e interpretado. Una de las principales características de Python es su sintaxis clara y su potencia, lo cual lo hace un sistema amigable y de fácil uso con el usuario nuevo en la programación. Python incorpora clases, módulos, excepciones, escritura dinámica, y tipos de datos dinámicos de alto nivel. Python puede extenderse en C o C++ y es un lenguaje apropiado para la programación de interfaces para aplicaciones. Cabe destacar que Python es un lenguaje portable que funciona en diversas variantes de Unix, en Mac y en los sistemas operativos Windows 2000 y superiores. (The Python Software Foundation, 2019)

Python es un lenguaje multipropósito o de propósito general, es decir, no está diseñado para desarrollar y acometer un único tipo de actividades, su basta cantidad de librerías abarcan áreas como el procesado de texto, protocolos de internet, ingeniería de software, operación con interfaces de sistema entre otras. Además, existen muchas extensiones de Python desarrollados por terceros para trabajar sobre otras áreas. Un ejemplo de ello, y relacionado con este proyecto, es la Fundación Geoespacial de Código Libre (OSGeo) que ha desarrollado librerías para el desarrollo de actividades de análisis geoespacial, teledetección, desarrollo y edición de archivos espaciales, por ejemplo.

Para este proyecto se emplearon la versión de Python 3.7. junto a las librerías proporcionadas por QGIS en su instalación y que forman parte también del proyecto OSGeo (Gdal y GRASS).

### 3.4. QGIS

QGIS es el sistema de información geográfica de código abierto y de fuente abierta (FOSS: *Free Open Source Software*). EL proyecto nació en mayo de 2002 y se estableció como tal en junio del mismo año en la plataforma SourceForge (<https://sourceforge.net/>). Gracias a este proyecto se ha conseguido crear un sistema de información geográfica libre y gratuito (ya que este tipo de software solía ser costoso y de tipo propietario) que pueda ser ejecutado en un ordenador personal.

QGIS funciona sobre todas las plataformas Unix: Windows, Mac, Linux... y está desarrollado con la caja de herramientas Qt y C++, lo que le ha permitido que sea un software fácil de comprender y ejecutar por los usuarios nuevos. El proyecto tenía como objetivo inicial ser un visor de gran cantidad de formatos ráster y vectoriales, el cual ha evolucionado hasta convertirse hoy en el GIS libre líder a nivel global.

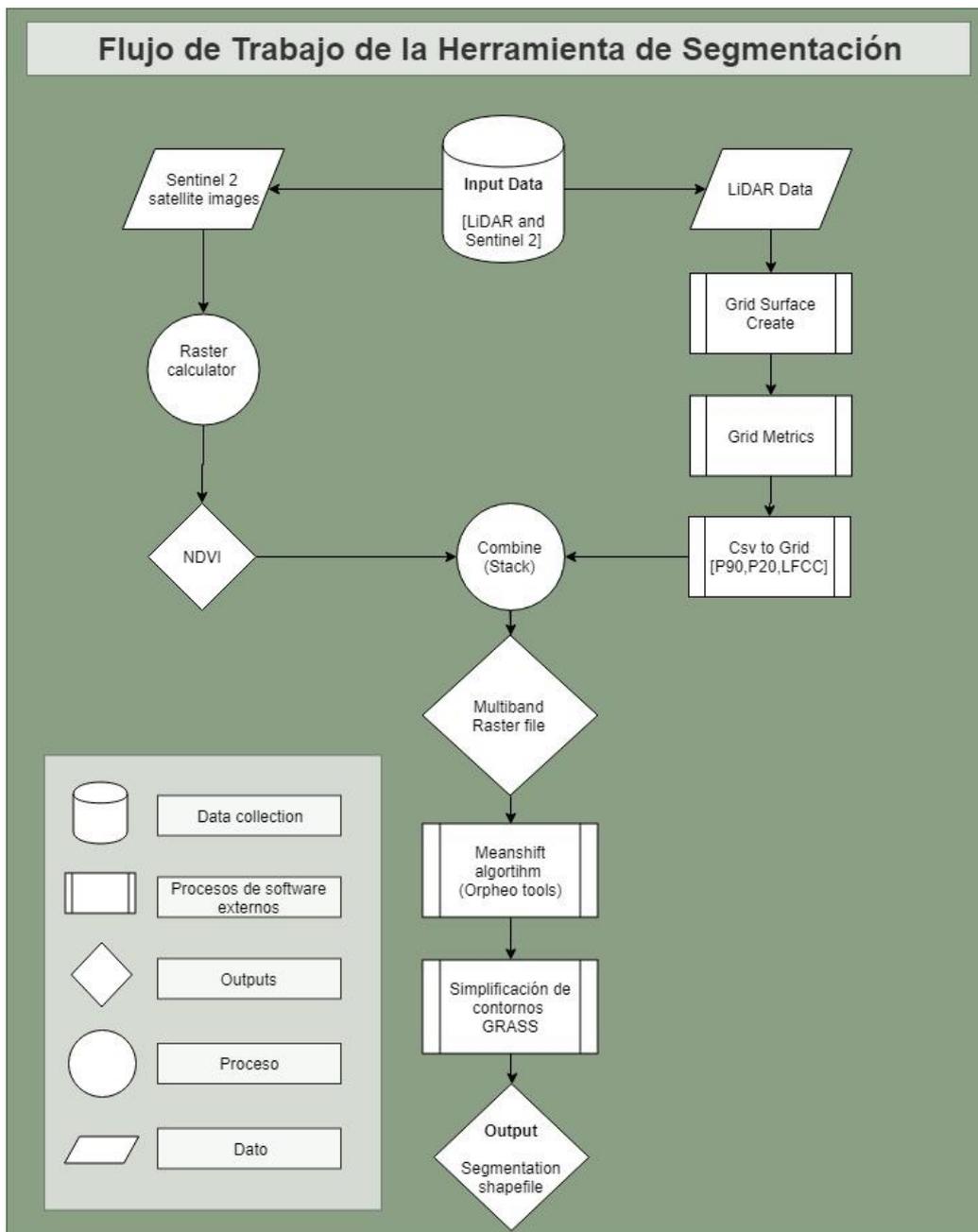
QGIS está licenciado bajo una licencia tipo GNU o *General Public License*, esto significa que el usuario puede modificar y revisar el código fuente y que se garantiza que el usuario siempre podrá acceder a un software GIS modificable de forma gratuita. Es un proyecto oficial de la Open Source Geospatial Foundation (OSGeo). Actualmente se encuentra en la versión 3.4 (estable) y en desarrollo de la 3.6. QGIS ofrece la posibilidad de hacer tus propios desarrollos apoyándose en el lenguaje de programación Python a través de lo

que ellos han denominado PYQGIS, que es la librería de QGIS para ejecutar código Python.(QGIS Development Team, 2018)

Dentro de OSGeo, también encontramos librerías de manejo de datos espaciales como son SAGA (Conrad et al, 2015), GRASS (GRASS Development Team, 2017) y GDAL (GDAL Development Team, 2017) que emplearemos para el procesamiento de datos, debido a su carácter abierto y su conexión con la interfaz y estructura de QGIS.

### 3.5. METODOLOGÍA. Flujo de trabajo

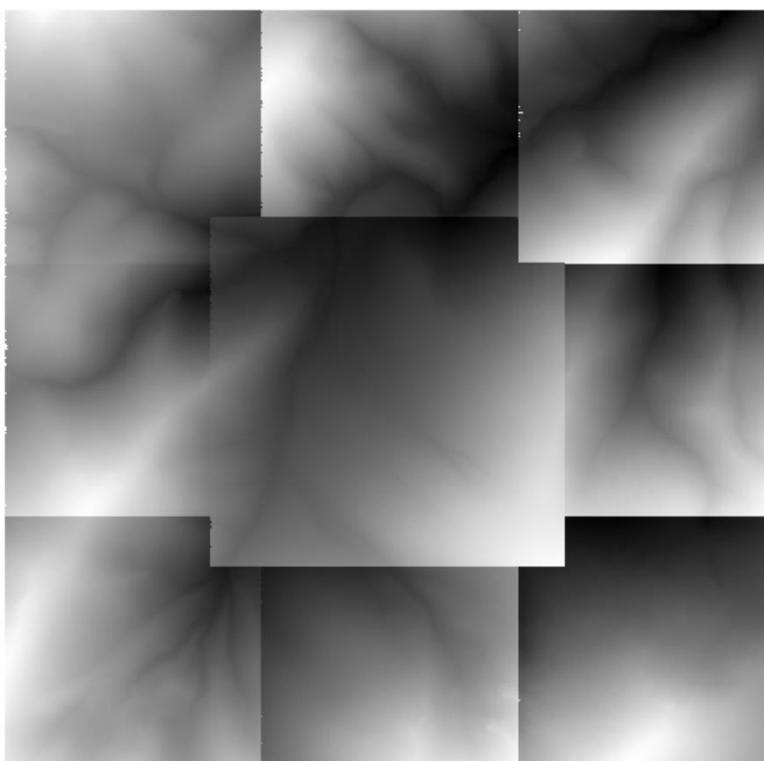
A continuación, se presenta un esquema de flujo de trabajo generalizado, sobre el cual se basará la herramienta de segmentación. Cada uno de los pasos se explicarán en mayor profundidad en sus correspondientes apartados:



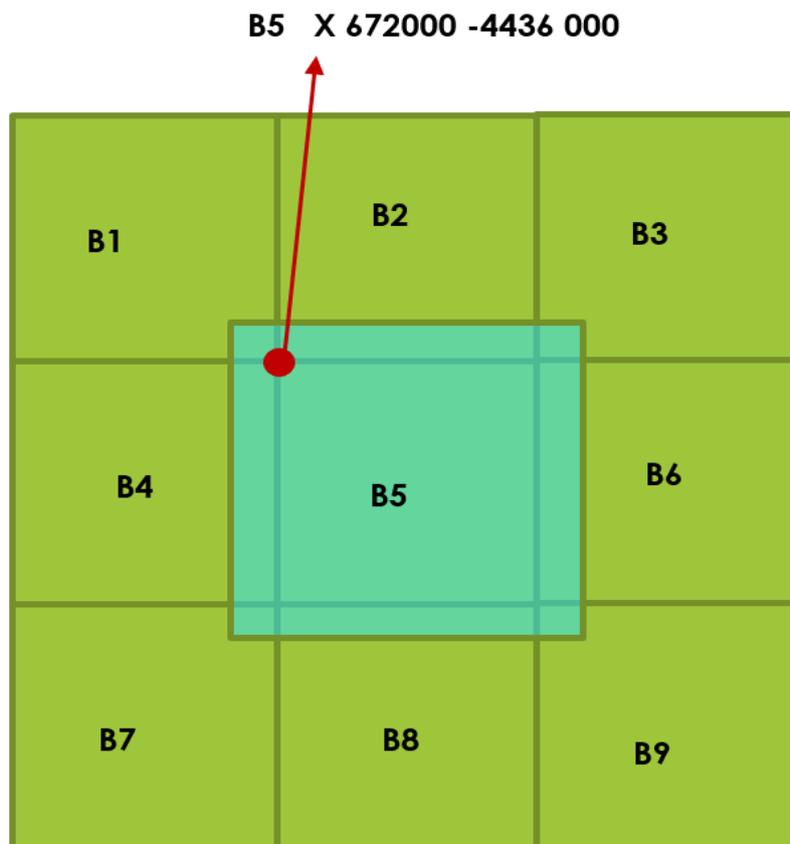
**Figura 7.** Flujo de trabajo general para la herramienta de segmentación. Se aplicarán dos tipos de datos diferentes LiDAR e imágenes aéreas proporcionadas por el satélite Sentinel 2.

### 3.5.1 Procesado LiDAR

La rama derecha del diagrama de flujo aborda el procesado de los datos LiDAR a través de la herramienta FUSION (McGaughey, 2018). FUSION es un software gratuito no abierto distribuido por el Departamento de Agricultura de los Estados Unidos de América. Al ser FUSION un software no abierto, no podemos modificar el código, estando detallados los procesos realizados en la documentación técnica proporcionada con la instalación. La herramienta desarrollada en una primera fase, comienza realizando el procesado de los datos llamando a FUSION. Primero, transforma los archivos LiDAR en formato “.laz” en un modelo digital de elevaciones (MDE) en formato “.dtm” propio de Fusion. Para ello el usuario, especificará las clases que querrá procesar (por defecto opera con la clase 2 para el suelo) y el tamaño de pixel (recomendable como mínimo 2m para PNOA, 0,5 pulsos/m<sup>2</sup>). Una vez generado el MDE hay que convertirlo a un formato raster de intercambio de archivos “.asc” para poder cargarlo en cualquier gis convencional. Al realizar esta transformación es posible que aparezcan en el resultado final (MDE.asc) un borde serrado en el límite izquierdo de la imagen. Para evitar este efecto, en la conversión de los datos se han empleado los LiDAR que rodean a la escena objetivo (denominaremos a estos datos que rodean la hoja objetivo marco). Sobre este marco se ha aplicado un buffer de 200 m que incluya la escena objetivo, de tal forma que el borde aserrado no afecte a la zona objetivo. Finalmente, los bordes sobrantes procesados, que contienen el borde de sierra y que quedan fuera de la zona de interés, se recortan. Se obtiene así, un MDE sin bordes en el procesado y que no dará problemas a la hora de unirlos con los MDE colindantes. Los MDE resultantes se obtienen en formato “.asc” y con una resolución de 2x2 m.



**Figura 8.** Marco de 9 MDE con MDE con el buffer aplicado superpuesto. Observamos que a la izquierda tenemos el borde aserrado. El buffer aplicado en el MDE hace que el borde de la izquierda quede fuera de la zona de interés.



**Figura 9.** Esquema de organización de las capas. El marco corresponde con las capas verdes (B1, B2, B3, B4, B6, B7, B8, B9). La hoja objetivo es la central, la B5. El buffer de 200m es el margen de la imagen azul que entra en el marco.

Una vez obtenidos los MDE, la herramienta calcula las métricas en continuo para cada bloque LiDAR es (GridMetrics). Para ello se le deberán indicar a la herramienta las clases a procesar, que en nuestro caso serán las correspondientes a vegetación y suelo (2,3,4,5). Sobre estas clases se calcularán las métricas, aquellos puntos que no pertenezcan a dichas clases no se procesarán. Además, se podrán definir el tamaño de pixel, la altura mínima de procesamiento de alturas y coberturas. El resultado de esta operación son dos archivos de texto “.csv” uno que recoge las métricas calculadas a los retornos en función de su altura y otro en función a su intensidad. Dentro de las métricas por elevación encontramos una serie de variables, que tienen correlación con estructura de la vegetación, en concreto para este desarrollo hemos elegido las siguientes métricas:

1. Elevación P90: elevación a la que se recoge el 90 por ciento de los retornos de un pulso. Esta variable está muy correlacionada con la altura dominante de la vegetación.
2. Elevación P20: elevación a la que se recoge el 20 por ciento de los retornos de un pulso. Aunque de un modo simplista esta variable tiene una cierta relación con la altura base de copa.
3. Porcentaje de primeros retornos por encima de la altura de corte: Se relaciona con la fracción de cabida cubierta.

Mediante algoritmos de Fusión (CSV2Grid) extraemos estas variables de formato de texto a formato ráster “.asc” y posteriormente el desarrollo une los raster generados por bloques individuales en un único ráster .tif (algoritmo *GDAL:merge* (Warmerdam et al, 2019)) para cada variable que represente la zona de estudio

### 3.5.2 Procesado de Sentinel

En la rama de izquierda del diagrama se muestra el procesado de las escenas obtenidas con Sentinel 2. Empezaremos calculando el índice de vegetación normalizado (NDVI) que se define con la siguiente expresión (a la derecha bandas empleadas en Sentinel 2):

$$NDVI = \frac{(NIR - ROJO)}{(NIR + ROJO)} = \frac{(B8a - B4)}{(B8a + B4)}$$

Este índice muestra la actividad fotosintética de la vegetación y se mueve en un rango de [-1,1] siendo 1 la mayor actividad fotosintética. Este índice aprovecha la alta reflectancia de la vegetación fotosintéticamente activa, para las longitudes de onda comprendidas en el rango del infrarrojo cercano y del rojo. Es decir, la vegetación no fotosintéticamente activa (carente de clorofila) quedará reflejada en este índice con valores bajos.

Para obtener el índice de vegetación normalizado, recurrimos al algoritmo de GRASS *vegetation index* (Kamble & Chemin, 2018). Este algoritmo está integrado en QGIS y por lo tanto no ofrece problemas de compatibilidad. Al igual que el que desarrollamos, este algoritmo emplea la misma fórmula anteriormente expuesta.

El resultado de este procesado es un ráster “.tif” de gran tamaño, que excede el tamaño del área de estudio, por lo tanto este ráster se recorta a través de la extensión de otra capa (ráster de las variables generados anteriormente) empleando *GDAL:cliprasterbyextent*. (QGIS Development Team, 2019). Empleamos GDAL ya que se trata de una librería abierta y gratuita instalada por defecto dentro de QGIS, de esta forma evitamos errores e incompatibilidades.

### 3.5.3 Procesado Ráster

Previamente a la creación del ráster multibanda se deberán normalizar las variables rasterizadas. Esto quiere decir que, para poder ser comparables dichas variables, éstas deben estar dentro del mismo rango de valores. El rango de valores elegido, es de 0 a 100. Los valores nulos o los no data serán reclasificados con el valor -9999.

Al reclasificar los no data o nulos con un valor numérico, conseguimos que el algoritmo de segmentación *meanshift* identifique correctamente estos valores, de lo contrario el algoritmo generará resultados no deseados y aumentará el tiempo de procesado, ya que no reconocería (y, por lo tanto, no agruparía) estos no-data. El resultado sería una segmentación pixel a pixel de los pixeles sin datos (*no data*) (Figura 11).



**Figura 10.** Segmentación defectuosa en la zona piloto 1. Valores no-data no reclasificados.

Dicha reclasificación se ha realizado de diferentes formas para cada variable:

1. Para la variable, fracción de cabida cubierta, no se han realizado modificaciones, puesto que, por defecto, esta capa ya está escalada de 0 a 100.
2. Las capas relacionadas con las variables de alturas de la vegetación (P90 y P20), se calculó:
  - para cada pixel de la capa < percentil 90 la siguiente expresión:

$$valor\ normalizado = \frac{valor\ del\ pixel - 2}{percentil\ 90\ de\ la\ capa - 2} * 100$$

En lugar de emplear el valor máximo de la capa para normalizar se emplea el percentil 90. Esto se justifica porque en las zonas piloto, es posible encontrar, puntos de alturas máximas que no se corresponden con la vegetación, como por ejemplo una catenaria o un edificio. Si observamos que la altura máxima, quedase muy por debajo de las alturas de vegetación, es aconsejable aplicar un percentil más bajo, por ejemplo, P90. Se resta dos porque es la mínima altura de procesado del LiDAR. Así hacemos que las curvas tengan de origen 0.

- Para pixels cuyo valor >p90:

$$valor\ normalizado = 100$$

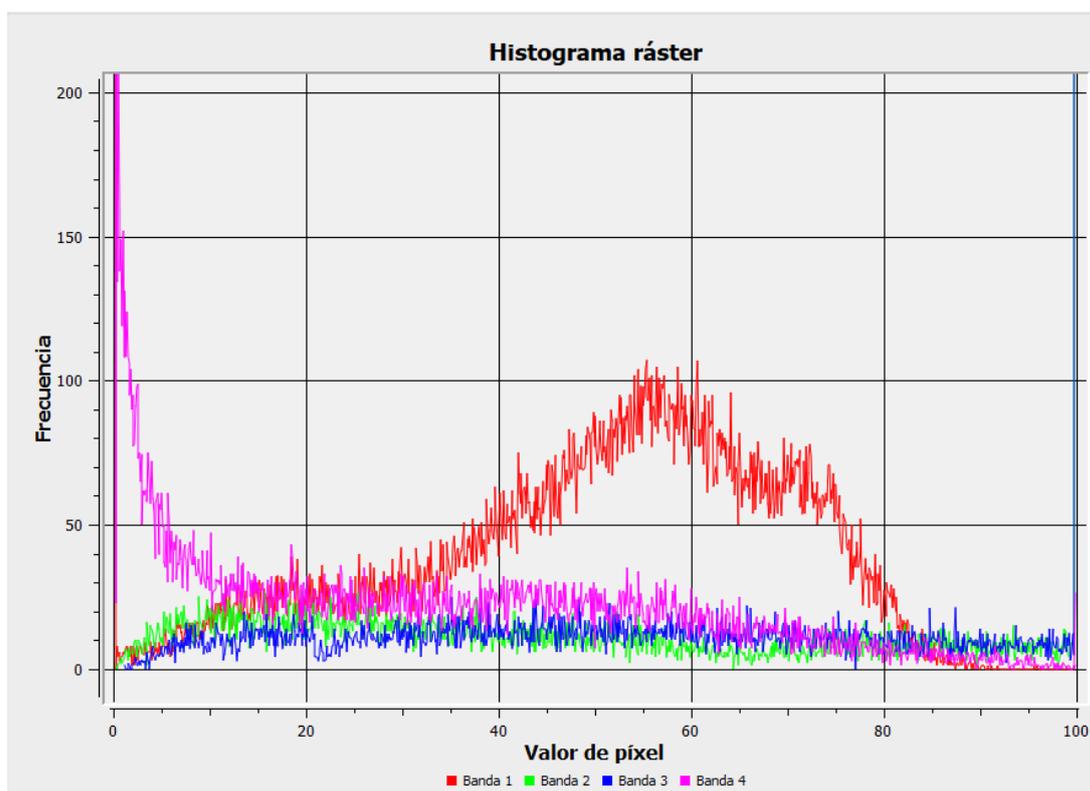
Se considera que aquellos valores por encima del percentil indicado corresponden con elemento que está por encima de la altura de la vegetación y se reclasifica como 100, con el máximo valor de altura.

3. El índice de vegetación ha sido normalizado teniendo en cuenta únicamente los valores entre 0 y 1, los cuales representan a la vegetación fotosintéticamente activa (todo aquel valor menor que cero se considera 0). A estos valores se les ha aplicado la siguiente reclasificación:

$$\text{valor normalizado} = \frac{\text{valor del pixel}}{\text{valor máximo de la capa}} * 100$$

Una vez obtenidos los resultados del procesado LiDAR y de Sentinel 2 y normalizadas las variables, procedemos a crear un ráster multibanda que recopile la información de todas las capas haciendo un *stack* (o pila) de capas. Nos hemos apoyado en el comando *GDAL:merge* (Warmerdam et al, 2019) para este proceso. Al tener todos los ráster de entrada la misma extensión (.tif) y tamaño nos evitamos tener problemas de incompatibilidades. Como resultado final, obtenemos un ráster multibanda .tif, que recoge la información, tanto del LiDAR como del Sentinel en una sola imagen. Es importante destacar, la posición en la que se colocarán las bandas, ya que, si en un futuro necesitamos hacer una operación determinada con su valor, será necesario conocer dicho orden. El orden de las bandas es alfabético, de esta forma, la primera banda tomará el ráster que empiece por la “a” si la hubiere y así sucesivamente.

Finalmente observaremos que las bandas deben quedar representadas aproximadamente de la siguiente forma:



**Figura 11.** Histograma de frecuencias para las bandas del ráster multibanda. Banda 1: NDVI, banda 2: P20, banda 3: P90, banda 4: LFCC.

### 3.5.4 Segmentación con Segmentation (Meanshift) OTB.

Empleamos el ráster multibanda como input para el algoritmo de segmentación meanshift, obtenido en la caja de herramientas Orpheo Tool Box, OTB (OTB-Team, 2018). Este algoritmo funciona principalmente uniendo aquellos píxeles de una imagen multiespectral que se encuentren a un rango de distancia inferior a un rango de radio (y opcionalmente a la distancia espacial por debajo del rango espacial) formando así un clúster (Youssefi, 2018).

Los algoritmos de segmentación disponibles son dos versiones diferentes del algoritmo de segmentación *meanshift* (uno de los cuales es multiproceso), el primero se fundamenta en componentes conectados basados en píxeles simples según un criterio definido por el usuario (generalmente la distancia), y otro es *watershed* para el gradiente de intensidad. Nosotros emplearemos el primero.

El algoritmo de segmentación de OTB funciona con dos modos principales que condicionan el archivo de salida:

- Modo ráster: devolverá un ráster en el que los píxeles vecinos son clasificados con la misma etiqueta, es decir de vuelve una segmentación de etiquetas y colores. Este modo carga toda la imagen a la memoria, por lo tanto, no puede procesar grandes escenas.
- Modo vectorial: En este modo se devolverá una capa *shapefile*. En esta capa se crearán teselas segmentadas de acuerdo a los parámetros y algoritmos establecidos por el usuario. En este modo se hace un uso más eficiente de la memoria, de tal forma que se pueden procesar imágenes de mayor tamaño. Finalmente, al estar activada la opción *stitch* se unirán los segmentos separados accidentalmente por el esquema de teselado. Este modo es el implementado en la herramienta.

Parámetros a modificar por el usuario:

- *Spatial radius*: radio espacial de los vecinos o **el umbral en distancia espacial para considerar píxeles en el mismo segmento**(Youssefi, 2018). Valor por defecto: 5.
- *Range radius*: Rango de radios que definen el radio en el espacio multiespectral expresado en unidades radiométricas. También puede definirse como **el umbral en la distancia euclidiana de la firma espectral (expresado en unidad de radiometría) para considerar los píxeles en el mismo segmento** (Youssefi, 2018). Valor por defecto: 5.
- *Maximum number of iterations*: El algoritmo iterativo terminará las convergencias una vez se haya alcanzado el número máximo de iteraciones. Valor por defecto: 100.
- *Minimum region size*: Tamaño mínimo de una región en la segmentación (en píxeles). Los clústeres más pequeños se fusionarán con los clústeres vecinos adyacentes con la radiometría más cercana. Si se indica en 0 no habrá segmentación. Tamaño mínimo de región es el tamaño del píxel. Valor por defecto: 100.

El resultado será un *shapefile* que, cumpliendo con las condiciones anteriores, haya realizado una segmentación del monte.

Cabe destacar que para cada escena o cada área de estudio las segmentaciones resultantes de los valores por defecto de la herramienta serán diferentes (en cuanto a tamaño de los segmentos, por valores de las bandas de la imagen multiespectral, texturas, tamaño total del área de estudio...), por lo que es aconsejable que el usuario

se identifique con la herramienta un poco y con los parámetros. Normalmente, la segmentación por defecto es útil como referencia a un segundo intento de segmentación, si el resultado con los valores por defecto no ha sido satisfactorio. Con un poco de práctica, el usuario podrá realizar segmentaciones satisfactorias al primer intento.

### 3.5.5 Suavizado de contornos Chaiken algorithm (GRASS)

La segmentación realizada por *meanshift* nos devolverá un archivo con esquinas muy angulosas y de un aspecto artificial, por eso aplicaremos el último paso antes de obtener el producto final. Se aplicará un suavizado de vértices, que funciona de forma inversa a la generalización de contornos (reducción del número de vértices).

El algoritmo de Chaiken suaviza una curva mediante la sustitución iterativa de cada punto por dos nuevos puntos: uno a 1/4 del camino al siguiente punto y otro a 1/4 del camino hasta el punto anterior, así “complica” la geometría aumentando sus elementos y por lo tanto su complejidad. Este algoritmo se caracteriza por generar una nueva línea que inscriba la original, de tal forma que los puntos de esta nueva están a una distancia umbral de la original (Bundala et al., 2007).

Empleamos el algoritmo de Chaiken implementado dentro de GRASS (y por lo tanto de QGIS) bajo el nombre de *grass7:v.generalize* (identificador del algoritmo en QGIS 3). Este módulo acepta de entrada un archivo vectorial a generalizar, es decir, para reducir su número de vértices. En este módulo mayoría de los algoritmos son de generalización. Sin embargo, encontramos en menor medida aquellos que buscan una función contraria, el suavizado. En el suavizado, se intenta devolver un aspecto original o real a una línea, principalmente, mediante el aumento del número de vértices, y por lo tanto, la creación de contornos más complejos (Bundala et al., 2007). Este aumento de vértices puede ser, en ocasiones excesivo, por lo que, es recomendable que emplear un algoritmo de generalización que elimine u exceso de vértices si fuere necesario. En este módulo los algoritmos de generalización/simplificación y de suavizado funcionan línea a línea, es decir que al suavizar una línea no se afecta a otra.

Generalmente, debido a que no buscamos un resultado preciso en el suavizado de contornos, sino que persigue una finalidad estética, el parecido con una rodalización real, en la herramienta usaremos los valores predeterminados para Chaiken (definidos a continuación), de tal manera que el usuario no tiene por qué preocuparse de modificar ningún parámetro de entrada.

De acuerdo con la documentación (Bundala et al., 2007) el algoritmo emplea:

- Un archivo vectorial de entrada, en este caso un shapefile con la segmentación realizada por *meanshift*.
- Un valor umbral sobre el cual realizar el suavizado. Se da en unidades cartográficas (para latitud longitud: en grados decimales): En general, el grado de simplificación incrementa con el aumento el valor umbral. Para Chaiken el valor por defecto es 1.

## 4.- RESULTADOS

Como resultado se ha obtenido una herramienta eficiente, rápida y sólida, la cual ha sido testada en ordenadores de diferentes características y especificaciones comprobando que la herramienta puede adaptarse a todo tipo de máquinas. La herramienta durante su funcionamiento proporciona información al usuario, de esta forma el usuario es capaz de saber en qué momento del procesado se encuentra. En caso de encontrar algún error, cuenta con un manejo de excepciones para ayudar al usuario a hacer funcionar la herramienta correctamente.

**Tabla 6.** Especificaciones de los ordenadores en los que se ha probado la herramienta.

Procesador	Ram instalada	nº archivos. laz (1000 x 1000)	Tiempo de procesado
Intel ® Core™ i7	16 GB	12	10 mins
Intel ® Core™ i5	4GB	12	12 mins

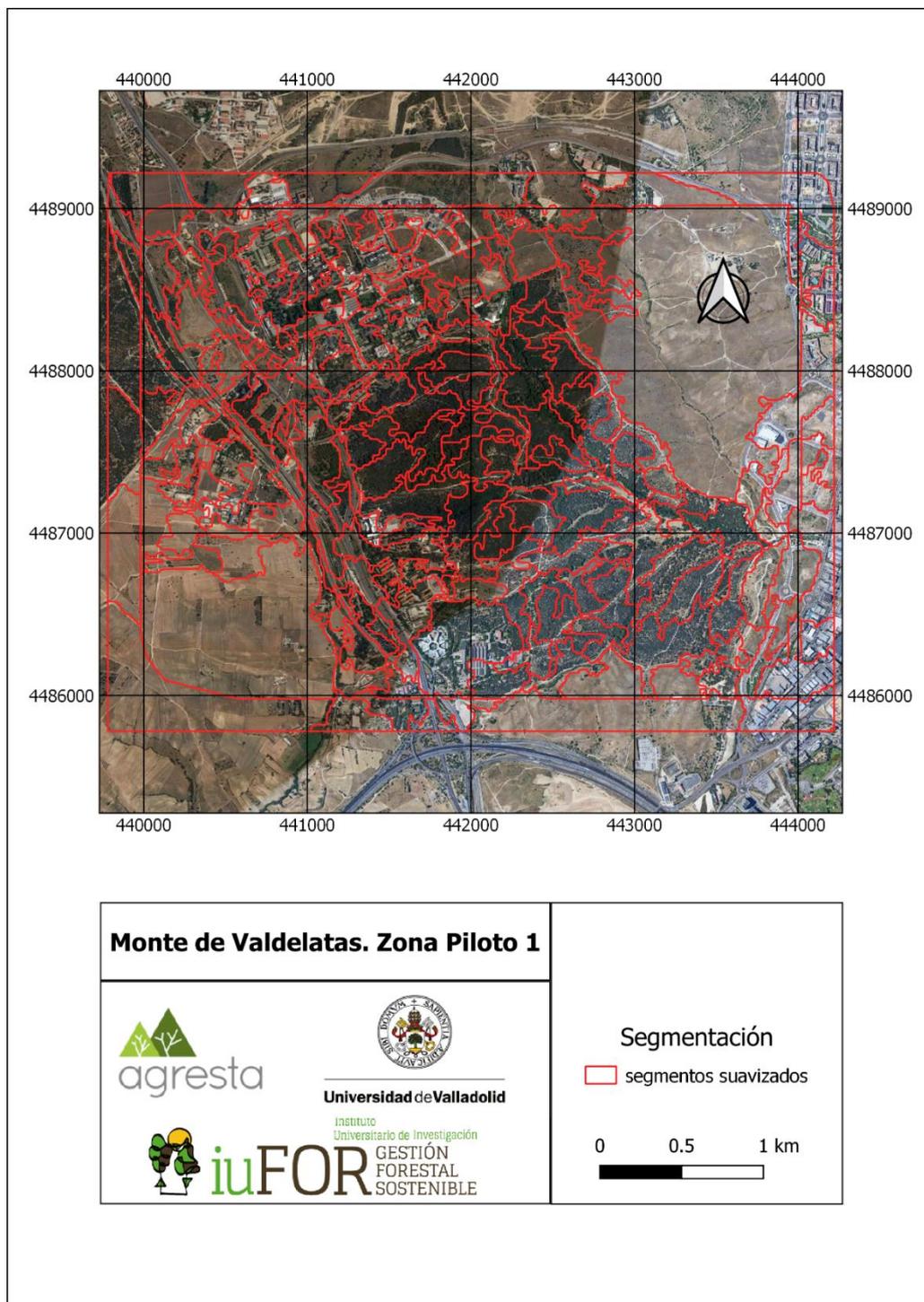
Se pueden emplear 2 tipos de datos: LiDAR y Sentinel. Si se utiliza **LiDAR + Sentinel 2**, todos los **resultados finales** serán almacenados en una carpeta llamada **'Productos finales' + el nombre introducido por el usuario**. Además, se generará una carpeta de **archivos temporales** con el nombre **'temp'+el nombre definido por el usuario**. Dentro de esta carpeta se encontrarán los archivos temporales necesarios para la ejecución de la herramienta. Si sólo se utiliza LiDAR, se generarán dos carpetas, **'Productos\_finales\_LiDAR' + el nombre introducido por el usuario** y **'temp\_LiDAR'**. Sin embargo, con Sentinel, sólo se generará una carpeta **'Productos\_finales\_Sentinel'**. **No borre nada de estas carpetas mientras la herramienta está procesando**. En las carpetas temporales, también se incluirán tres subcarpetas, en las que para cada archivo LAZ se devolverá un subproducto diferente útil en los procesos de planificación, a partir de los cuales se generarán los productos finales. Estas carpetas son:

- MDE\_+ nombre introducido: almacena los MDE en formato .asc y .dtm de cada archivo LAZ.
- Métricas\_+nombre introducido: contiene las métricas individuales de cada archivo.
- Rasters\_+nombre introducido: se encuentran las variables seleccionadas en formato .asc, para cada archivo LiDAR.

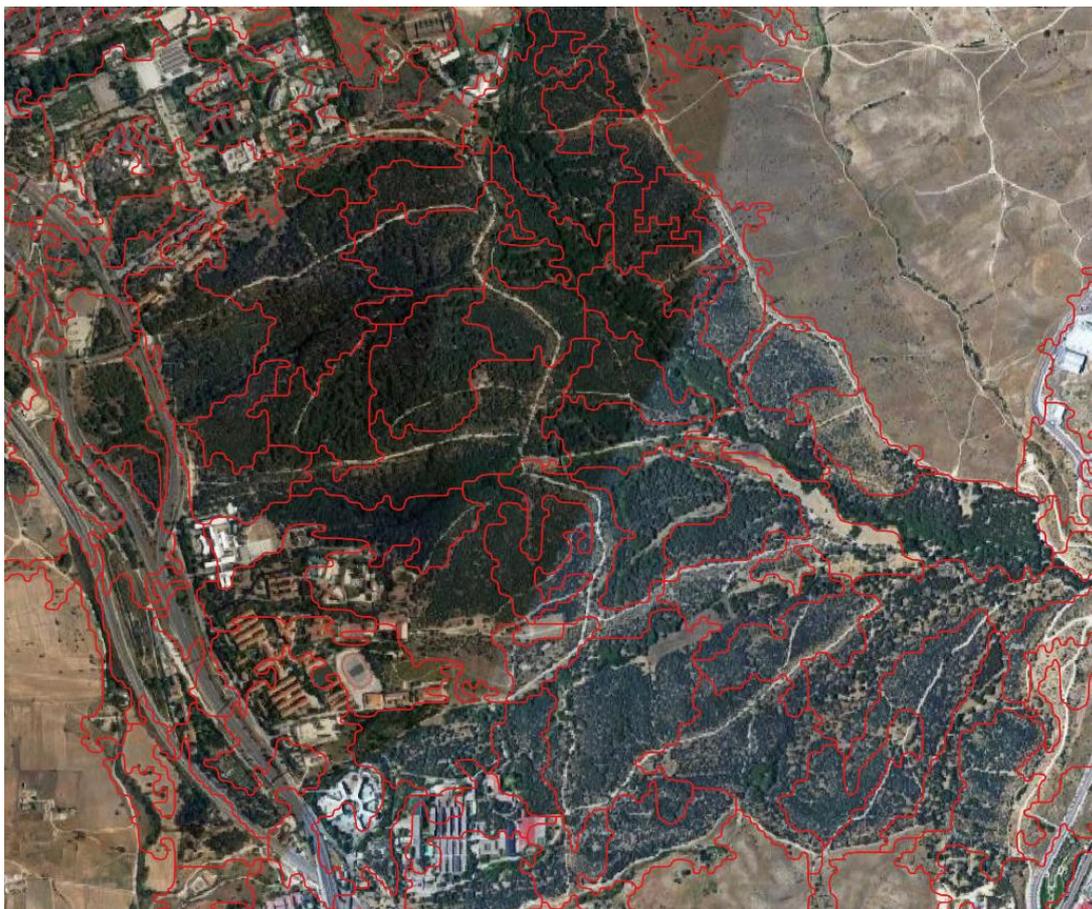
La herramienta ofrece de productos finales al fin de la ejecución, las siguientes capas:

- Un modelo digital de elevaciones del terreno
- Tres rastres que representan diferentes variables forestales, por defecto la herramienta ofrece:
  - La altura dominante
  - La altura del primer verticilo vivo
  - Fracción de cabida cubierta
- Un ráster con el índice de vegetación normalizado para la zona de estudio
- Una segmentación preliminar sin un suavizado de bordes mediante el algoritmo meanshift de OTB
- Por último, a propuesta de los rodales con un suavizado de bordes mediante Chaiken

La herramienta fue probada con diferentes datos (PNOA-LiDAR .laz, e imágenes Sentinel, .jp2) correspondientes a dos zonas piloto. En ambas zonas se busca ver el funcionamiento de la herramienta, su estabilidad, tiempo de procesado y los resultados de la segmentación. A continuación, se muestran los resultados de segmentación ofrecidos por la herramienta en la carpeta de ‘Productos finales’ en cada una de las dos zonas. Los mapas de los subproductos: MDE, alturas, FCC y NDVI se adjuntan el Anexo 1.



**Figura 12.** Segmentación con algoritmo meanshift. Variables Spatial radius: 4 Range Radius: 12, Maximun iteration number: 150, Minimun region size: 80 m.

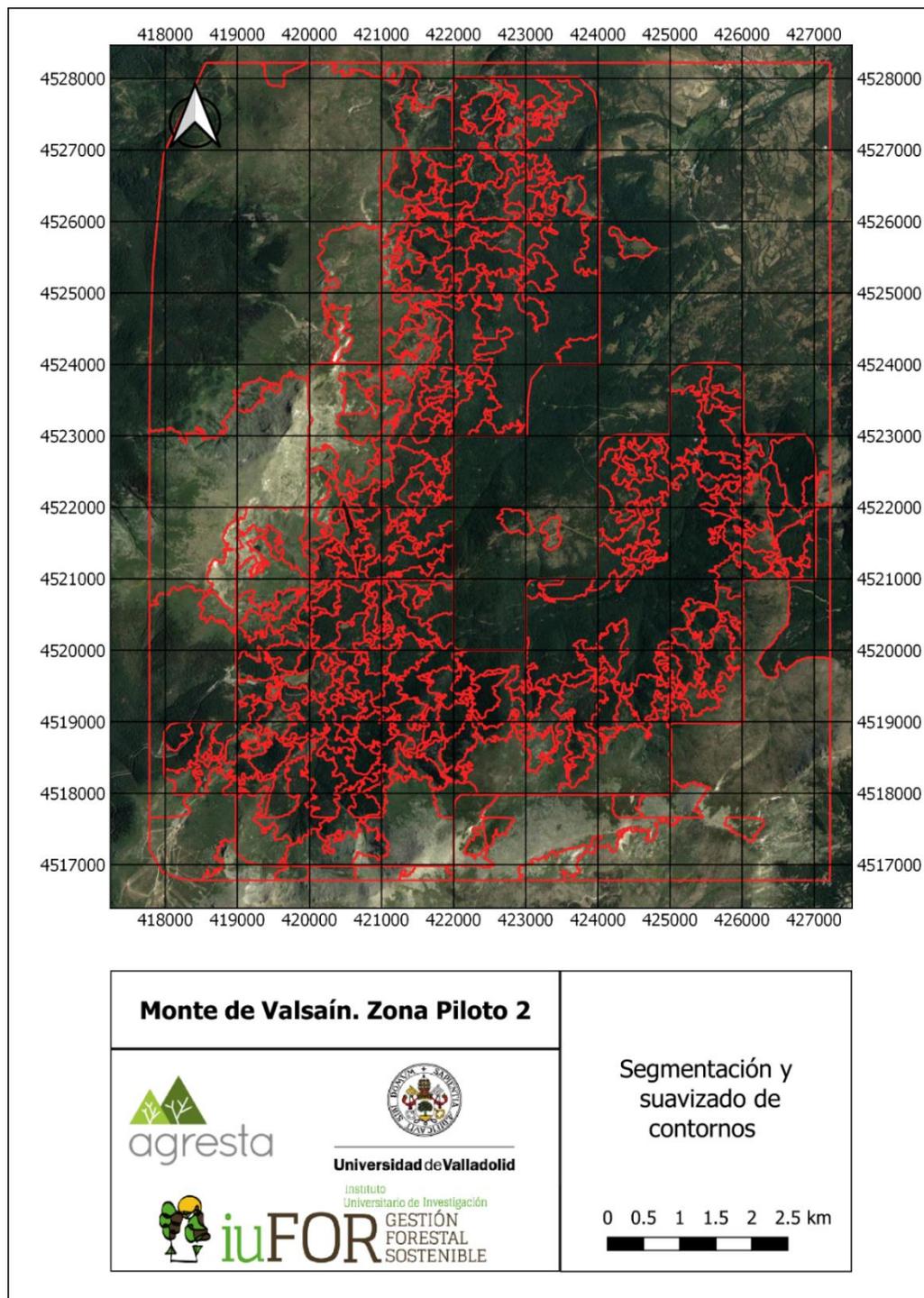


**Figura 13.** Propuesta de rodalización con simplificación de contornos de Valdelatas. La foto está superpuesta a la imagen obtenida por Google Sat (open source en QGIS)

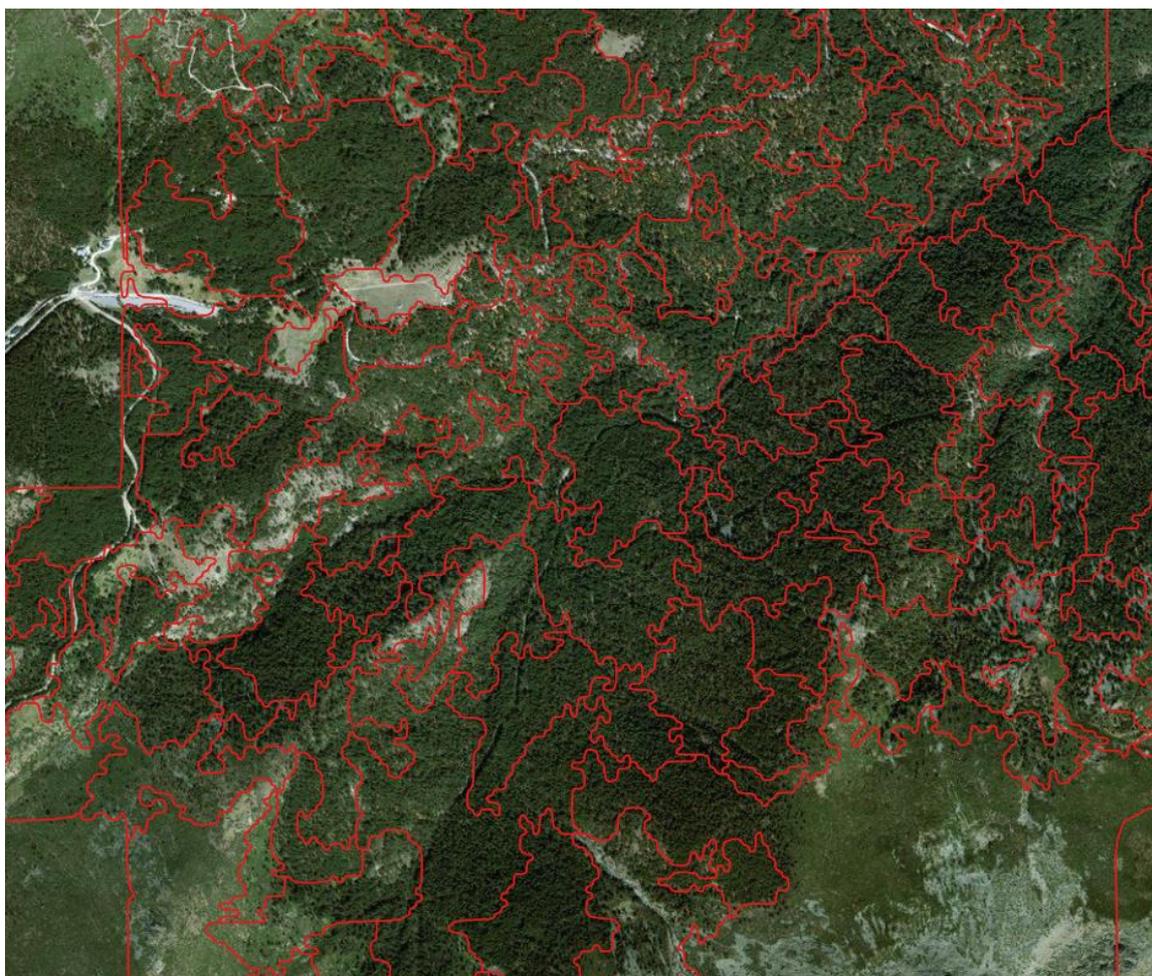
Respecto, a los resultados obtenidos en Valdelatas, observamos que:

- a) El modelo de elevaciones (figura 1, Anexo 1) es un modelo preciso y homogéneo en resolución sin influencia de los bordes de sierra, que podrían haber modificado el resultado final.
- b) Los ráster de variables forestales (Anexo 1) se ajustan a las clases especificadas por el usuario, por lo tanto, las zonas sin valores corresponden en este caso a las zonas sin vegetación, ya que especificamos en el procesado de LiDAR, emplear las clases de vegetación (2,3,4,5). Este es el motivo por el que se puede ver un 'moteado' en la imagen. En lo referente al patrón de colores, tiene sentido que el P20, altura de copa, tenga un color más claro (alturas inferiores) respecto al P90, ya que la altura de copa es menor que la altura dominante del pie.
- c) Las segmentaciones por lo general, se ajustan bien al territorio descrito, aun así, es necesario apuntar, que el resultado final depende de los valores definidos por el usuario al inicio, de los archivos utilizados y de la zona de estudio, lo que puede cambiar el efecto de los valores por defecto de la herramienta. A pesar de esto, podemos observar como la segmentación (figura 13) se ajusta bien a las diferentes tonalidades y texturas de la imagen que corresponden con las diferentes teselas del mosaico de vegetación. Esta segmentación no se ajusta únicamente, a las texturas de la foto, sino que al superponerlo con las características que observamos en la realidad, se ajusta adecuadamente al sentido biológico y de gestión que se requiere. Es cierto, que pueden darse algunas segmentaciones arbitrarias, esto se puede dar bien, a la naturaleza del algoritmo o a un mal ajuste y calibrado del mismo.

El empleo de la herramienta en una mayor área, como la zona de piloto de la Cabecera del Valle de Lozoya nos muestra los siguientes resultados:



**Figura 14.** Propuesta de rodalización con simplificación de contornos sobre Sierra de Guadarrama: Peñalara y Cabeza de Hierro. La foto está superpuesta a la imagen satelital ofrecida por Bing Sat para una mejor interpretación.



**Figura 15.** Detalle de la propuesta de rodalización para la zona piloto de la Cabecera del valle de Lozoya.

Los resultados apreciados en la Cabecera del valle de Lozoya respecto a los ráster, modelos digitales de elevaciones e índices de vegetación normalizado, no varían respecto a las impresiones generales observadas en Valdelatas, los resultados son satisfactorios. Sin embargo, en lo respectivo a la segmentación, a pesar, de haber encontrado una combinación satisfactoria de parámetros, resulta más complicada la configuración del algoritmo de segmentación que en el caso de Valdelatas. Es posible que este efecto se produzca debido a:

- a) La mayor homogeneidad de las masas forestales de esta zona de estudio, influye en una mayor complejidad en la detección de los distintos segmentos por la homogeneidad de los datos y la textura de la capa ráster multibanda. Ello sumado a que, si en el algoritmo se fuerza a tener un tamaño mínimo de tesela reducido, el algoritmo busque uniones entre celdas que biológicamente no guarden gran sentido (Pascual et al, 2019).
- b) Una mala combinación de los parámetros de entrada del algoritmo. Este efecto, en teoría, se debe minimizar con un breve entrenamiento del usuario con la herramienta.

Por lo tanto, se puede concluir que, si bien la herramienta diseñada funciona de un modo eficiente, para obtener uno resultados óptimos el usuario debe desarrollar unas capacidades mínimas orientadas a la configuración eficiente del algoritmo de segmentación utilizado, trascendente a la hora de obtener unos buenos resultados finales.

## 5.- CONCLUSIÓN Y DISCUSIÓN

La metodología desarrollada se inspira en proyectos parecidos de automatización de datos LiDAR, en los cuales metodologías similares se han probado obteniendo reconocimiento entre la comunidad forestal, como en la herramienta SilviLiDAR (Crespo & Díez Rabanos, 2018). Este proyecto, además del procesado utiliza la información generada para proponer medidas de gestión orientadas a las distintas situaciones selvícolas identificadas para cada zona. En la herramienta propuesta en este estudio, se trabaja con segmentación orientada a objeto en lugar de pixel a pixel, incluyéndose además el uso de las imágenes Sentinel como dato de entrada en la segmentación, lo que supone complementar a la información de estructura del LiDAR con una mayor información a la hora de alimentar el algoritmo de segmentación. Sin embargo, no se proponen medidas de actuación silvícola asociadas a los distintos segmentos detectados.

En otros trabajos (Cantero Fauquier, 2016), se plantean metodologías enfocadas a modelos de árbol individual, donde se procesan los datos LiDAR para la delineación de copas a partir de la generación de Modelos de altura de vegetación con y sin el aporte de ortofoto. Sin embargo, se trata de un flujo de trabajo de carácter semiautomático que carece del componente de automatización, que ofrecen tanto SilviLiDAR o en esta herramienta. Por otro lado, su utilización no genera los subproductos (por ejemplo, ráster con variables LiDAR) orientados a la planificación y gestión forestal que complementan el producto desarrollado. Este tipo de trabajos se centran más en observar el resultado de la segmentación sobre determinado tipo de cubiertas forestales, mientras que este proyecto se presenta como una herramienta de gestión.

Una de las ventajas de la herramienta es que está desarrollada a partir de datos y software gratuitos y en su mayoría abiertos (excepto FUSION, pero que también es gratuito). Además, el script está implementado sobre el API (*Application Programming Interface*) de QGIS y utiliza de sus librerías. Esta característica permite, que la herramienta sea fácilmente distribuible y accesible por otros usuarios, ya que la gran mayoría de funciones requeridas por el programa está incluida en QGIS y no requieren de gran cantidad de instalaciones paralelas de programas o librerías.

Los resultados obtenidos en las zonas heterogéneas o en zonas forestales con densidades medias, donde se aprecien cambios o diferencias en la estructura de la vegetación, son satisfactorios. En las áreas donde, la diversidad de masas forestales sea mayor (y preferiblemente el tamaño sea menor), como en Valdelatas, la herramienta podrá dar un resultado más aproximado a la segmentación final que desea el usuario. Sin embargo, en aquellas zonas homogéneas con coberturas altas, se recomienda trabajar en la configuración del algoritmo y emplear zonas de estudio de menor tamaño, o fragmentar la zona en unidades a procesar individualmente. También es aconsejable el empleo de la herramienta en áreas homogéneas, con valores de los parámetros de entrada por defecto o mayores (dependerá de la zona, pudiendo ser ligeramente inferiores a los dados por defecto). De tal forma, que el resultado de la segmentación sea más grosero en cuanto a número de divisiones, pero con un mayor sentido biológico, así la herramienta pudiera servir como una alternativa para generar unidades del paisaje en este tipo de estructuras. Posteriormente el usuario podría detallar manualmente, aquellas regiones de interés que no obtuvieron el resultado esperado.

Como se comentaba previamente al comienzo de este trabajo fin de master, el objetivo de este script es generar unidades homogéneas del monte, a partir de datos Satelitales y LiDAR que, faciliten el trabajo de teselado que requiere cualquier planificación forestal. La herramienta no plantea unos límites inamovibles, de esta forma el usuario tras validar los resultados del teselado previo en campo, puede corregir manualmente determinados polígonos que no crea convenientes o acertados. De esta forma, el gestor ahorra tiempo

de gabinete en el proceso de delimitación del monte ganando precisión al mismo tiempo. El resultado generado, está llamado a ser una herramienta muy potente para acabar de afinar los límites reales de los tipos de masa durante el trabajo de campo, un trabajo de campo que puede ser más intenso merced al tiempo que ahorra la herramienta en la preparación del mismo en gabinete.

El resultado obtenido, depende en gran medida del conocimiento de la parametrización del algoritmo de segmentación en cada caso, es previsible que los futuros usuarios de la herramienta generen un conocimiento experto asociados a distintas formaciones forestales que optimice los resultados obtenidos. Poder manipular los datos de entrada de algoritmo permite que los resultados de la herramienta se ajusten, tanto a proyectos paisajísticos, como de ordenación forestal, aprovechamientos, conectividad. Es interesante, por lo tanto, entrenar la herramienta, para determinar rangos de parámetros que se ajusten a las diferentes tipologías de masa.

En las líneas futuras de trabajo, se debería mejorar la interfaz gráfica de la herramienta proveyendo de menús despegables, botones de dos posiciones (ejemplo para el tamaño de los archivos LiDAR) y no sólo entradas de texto para el valor de los parámetros. Además, sería positivo aumentar el número de algoritmos de segmentación, para que el usuario pudiera tener diferentes calidades de segmentado, en función del tipo de segmentación que desee. Una opción interesante sería implementar la Object Based Image Segmentation, que herramientas como Google Earth Engine, ya está implementando y proveen de mejores segmentaciones (Goerick, 2018). Es preciso, como se comentó previamente, entrenar la herramienta en diferentes tipologías de masa, pudiendo determinar así los rangos ideales de los parámetros del algoritmo de segmentación para cada caso concreto. Finalmente, sería conveniente independizar el código del api de QGIS, ya que de esta manera, la herramienta seguiría funcional incluso si QGIS dejara de actualizar sus versiones y dar soporte a las antiguas.

## 6.- AGRADECIMIENTOS

Quisiera expresar mi agradecimiento por la supervisión, apoyo y confianza de durante estos meses a mis tutores: José Luis Tomé que, a pesar de sus compromisos, siempre ha buscado el momento para guiarme y aconsejarme en el proyecto, y a Felipe Bravo por su supervisión y por ponerme en contacto con Agresta s.coop. gracias al master DataForest ([dataforest.uva.es](http://dataforest.uva.es)).

En Agresta, dar las gracias a Iñigo Escamochero por sus consejos, cuando el mundo del código parecía oscuro e irresoluble y a Nacho por ayudarme a mejorar la herramienta con su revisión de la segmentación.

A César López Leiva por su interés y charlas.

Por estar a mi lado: a mi novia, mi hermano, mis padres y mis amigos (vosotros sabéis quiénes sois).

## 7.- BIBLIOGRAFÍA

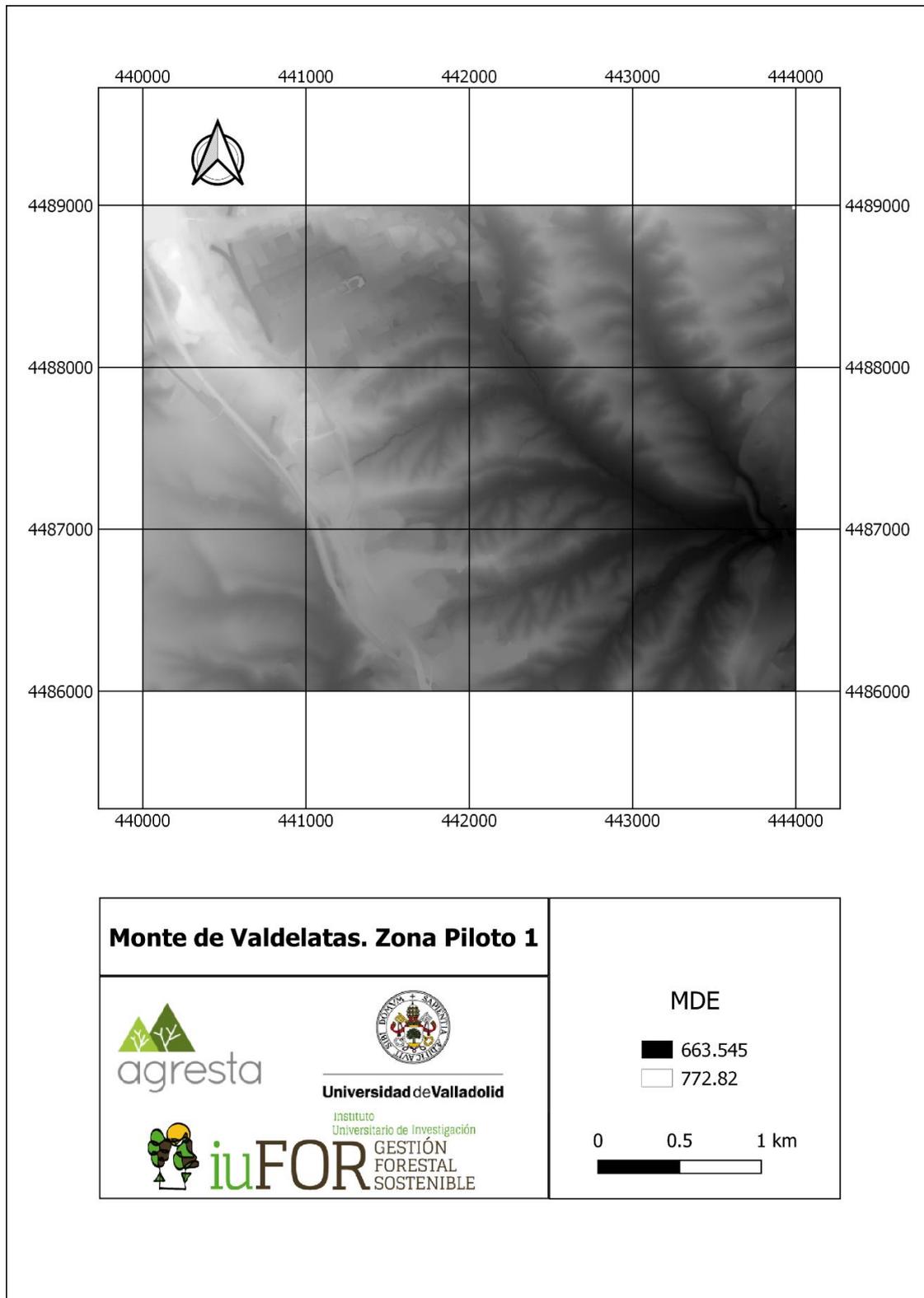
- A Wulder, M., W Bater, C., C Coops, N., Hilker, T., & C White, J. (2008). The role of LiDAR in sustainable forest management. *The Forestry Chronicle*, 84(6), 807–826. Retrieved from <http://pubs.cif-ifc.org/doi/abs/10.5558/tfc84807-6>
- Bundala, D., Bergenheim, W., & Metz, M. (2007). GRASS: v.generalize. Retrieved from <https://grass.osgeo.org/grass77/manuals/v.generalize.html>
- Cantero Fauquier, F. J. (2016). Generación de herramientas de localización de árboles individuales con el módulo de procesamiento de QGIS a partir de información LiDAR y ortofotografía aérea.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., ... Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA). *Geosci. Model Dev.*, 8, 1991-2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Crespo Rodrigo, A., & Díez Rabanos, F. J. (2018). SilviLiDAR: Herramienta de manejo de datos LiDAR para la gestión forestal. *Montes*, 134, 23–26.
- European Spatial Agency (ESA). (2015). Overview / Copernicus / Observing the Earth / Our Activities / ESA. Retrieved February 18, 2019, from [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Overview3](https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview3)
- European Spatial Agency (ESA). (2019a). Copernicus: Overview. Retrieved from [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Overview4](http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4)
- European Spatial Agency (ESA). (2019b). Missions - Sentinel Online. Retrieved February 18, 2019, from <https://sentinel.esa.int/web/sentinel/missions>
- European Spatial Agency (ESA). (2019c). MSI Instrument – Sentinel-2 MSI Technical Guide – Sentinel Online. Retrieved February 18, 2019, from <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>
- European Spatial Agency (ESA). (2019d). Sentinel-2 - Data Products - Sentinel Handbook. Retrieved February 19, 2019, from <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products>
- Gandhi, M., S.Parthiban, Thummalu, N., & Christy, A. (2015). Ndvi: Vegetation Change Detection Using Remote Sensing and Gis – A Case Study of Vellore District. *Procedia - Procedia Computer Science*, 57, 1199–1210. <https://doi.org/10.1016/j.procs.2015.07.415>
- GDAL Development Team. (2017). GDAL - Geospatial Data Abstraction Library. Open Source Geospatial Foundation.
- Goerick, N. (2018). *EEUS 2018- Image Segmentation and object based methods*. Retrieved from <https://www.youtube.com/watch?v=2R0aTaMtYTY&t=219s>
- GRASS Development Team. (2017). Geographic Resources Analysis Support System (GRASS). Open Source Geospatial Foundation.
- INEGI. (2013). *Imágenes de satélite Elementos de percepción remota ¿Qué es percepción remota?* Retrieved from

- [http://www3.inegi.org.mx/contenidos/temas/mapas/imagenes/imgsatelite/metadatos/elem\\_per\\_rem.pdf](http://www3.inegi.org.mx/contenidos/temas/mapas/imagenes/imgsatelite/metadatos/elem_per_rem.pdf)
- Instituto Geográfico Nacional (IGN). (n.d.). Plan Nacional de Ortofotografía Aérea. Presentación. Retrieved February 18, 2019, from <http://pnoa.ign.es/presentacion>
- Instituto Geográfico Nacional (IGN). (2018). *Plan Nacional de Ortofotografía Aérea ESPECIFICACIONES TÉCNICAS*. Retrieved from <https://www.fomento.gob.es/NR/rdonlyres/F1D425E0-2BEC-44B9-9A48-1AEF9CE31B25/148276/ESPECIFICACIONESTECNICAS.pdf>
- Kamle, B., & Chemin, Y. (2018). GRASS GIS manual: i.vi. Retrieved from <https://grass.osgeo.org/grass74/manuals/i.vi.html>
- Lim, K., Treitz, P., Wulder, M. A., St-Onge, B., & Flood, M. (2003). Lidar remote sensing of forest canopy and stand structure. *Progress in Physical Geography*, 27(1), 88–106. <https://doi.org/10.1191/0309133303pp360ra>
- McGaughey, R. J. (2018). FUSION/LDV: Software for LIDAR Data Analysis and Visualization. 214 p Seattle: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, University of Washington.
- McGaughey, R. J., Andersen, H. E., & Reutebuch, S. E. (2006). Considerations for planning, acquiring, and processing LIDAR data for forestry applications. In *Proceedings of the eleventh Biennial USDA Forest Service Remote Sensing Applications Conference*. Salt Lake City, ID, April 24-28, 2006.: American Society of Photogrammetry and Remote Sensing, Bethesda, MD. (in press).
- Næsset, E., & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, 79(1), 105–115.
- Nelson, R. (2013). How did we get here? An early history of forestry LiDAR. *Can J Remote*, 39, 6–17.
- OTB-Team. (2018). Segmentation - Segmentation. Retrieved from [https://www.orfeo-toolbox.org/CookBook/Applications/app\\_Segmentation.html](https://www.orfeo-toolbox.org/CookBook/Applications/app_Segmentation.html)
- Pascual, A., Bravo, F., & Ordoñez, C. (2019). Assessing the robustness of variable selection methods when accounting for co-registration errors in the estimation of forest biophysical and ecological attributes. *Ecological Modelling*. <https://doi.org/https://doi.org/10.1016/j.ecolmodel.2019.04.018>
- QGIS Development Team. (2018). QGIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from <http://qgis.osgeo.org>”.
- QGIS Development Team. (2019). Clip raster by extent. Retrieved from [https://docs.qgis.org/2.8/en/docs/user\\_manual/processing\\_algs/gdalogr/gdal\\_extracton/cliprasterbyextent.html](https://docs.qgis.org/2.8/en/docs/user_manual/processing_algs/gdalogr/gdal_extracton/cliprasterbyextent.html)
- The American Society for Photogrammetry & Remote Sensing. (2009). Las Specification 1.3. *The American Society for Photogrammetry & Remote Sensing*, 17. Retrieved from [http://www.asprs.org/a/society/committees/standards/asprs\\_las\\_spec\\_v13.pdf](http://www.asprs.org/a/society/committees/standards/asprs_las_spec_v13.pdf)

- The American Society for Photogrammetry & Remote Sensing. (2013). *LAS SPECIFICATION VERSION 1.4 – R13 15 July 2013*. American Society for Photogrammetry & Remote Sensing. Retrieved from [http://www.asprs.org/a/society/committees/standards/LAS\\_1\\_4\\_r13.pdf](http://www.asprs.org/a/society/committees/standards/LAS_1_4_r13.pdf)
- The Python Software Foundation. (n.d.). General Python FAQ. Retrieved from <https://docs.python.org/3/faq/general.html#what-is-python>
- Tomé Morán, J. L., Sanjuanbenito García, P., & Fernández Landa, A. (2013). Cartografía de Vegetación en la Comunidad de Madrid utilizando información LiDAR del Plan Nacional de Ortofotografía Aérea (PNOA). Vitoria-Gazteiz: 6º Congreso Nacional Forestal 10/06/2013 - 14/06/2013.
- Valcarcel, J. (2016). Cartografía Digital: Sentinel 2, el satélite de moda. Retrieved February 18, 2019, from <https://digimapas.blogspot.com/2016/04/sentinel-2-el-satelite-de-moda.html>
- Warmerdam, F., Rouault, E., & Others, A. (2019). GDAL: gdal\_merge.py. Retrieved from [https://gdal.org/gdal\\_merge.html](https://gdal.org/gdal_merge.html)
- White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing*, 42(5), 619–641. <https://doi.org/10.1080/07038992.2016.1207484>
- World Economic Forum. (2018). *The Future of Jobs Report. Executive Summary* (Vol. 5). <https://doi.org/10.1177/1946756712473437>
- Youssefi, D. (2018). LSMSSegmentation - Exact Large-Scale Mean-Shift segmentation, step 2. *OTB Cookbook*. Retrieved from [https://www.orfeo-toolbox.org/CookBook/Applications/app\\_LSMSSegmentation.html](https://www.orfeo-toolbox.org/CookBook/Applications/app_LSMSSegmentation.html)

## ANEJO 1. RESULTADOS DE LA HERRAMIENTA. SUBPRODUCTOS

### ZONA PILOTO 1: Monte Periurbano en Madrid (Valdelatas)



**Figura 16.** Modelo Digital de Elevaciones de Valdelatas. Clases empleadas en el proceso: 2 (suelo). Tamaño de píxel 2 metros

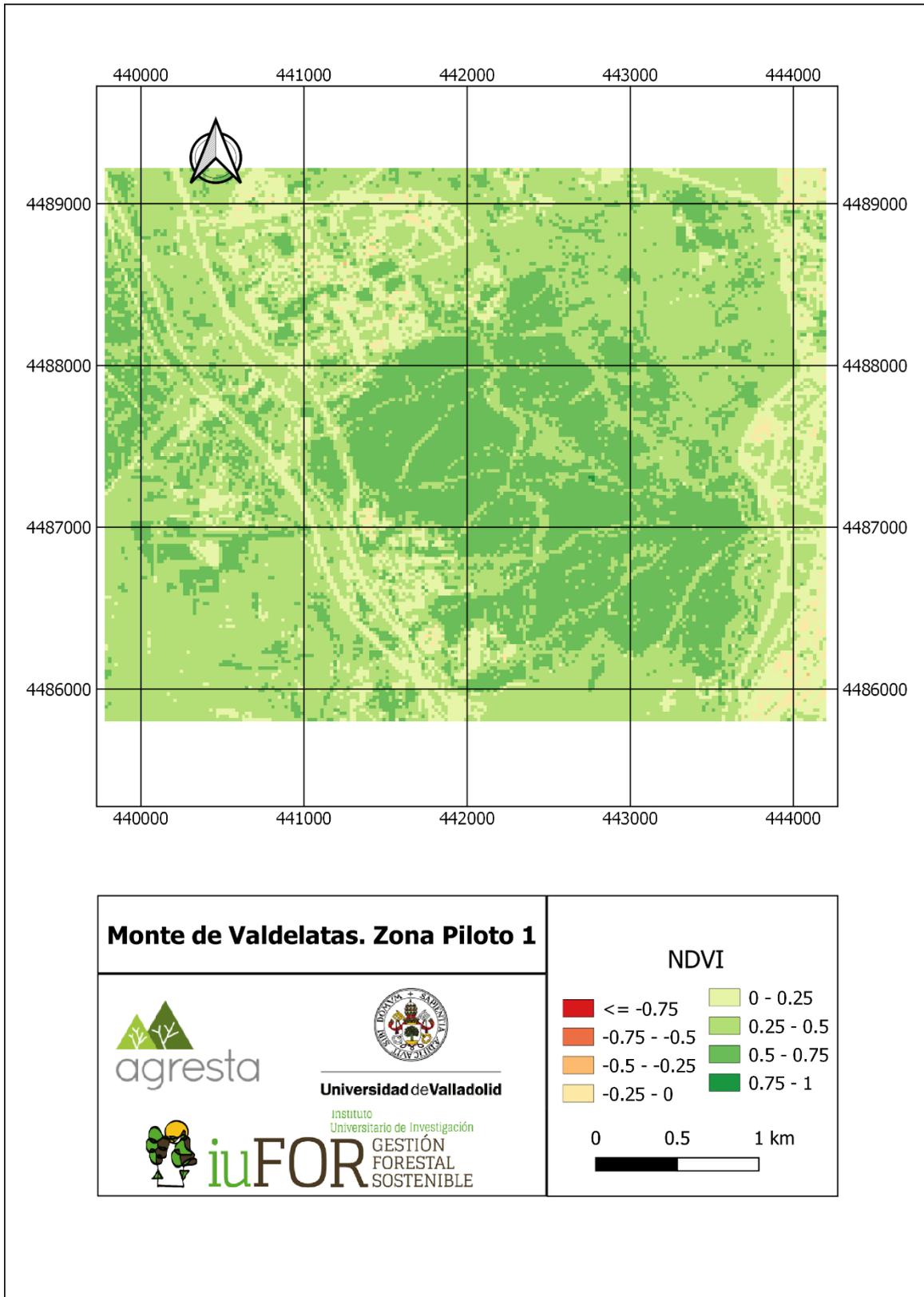
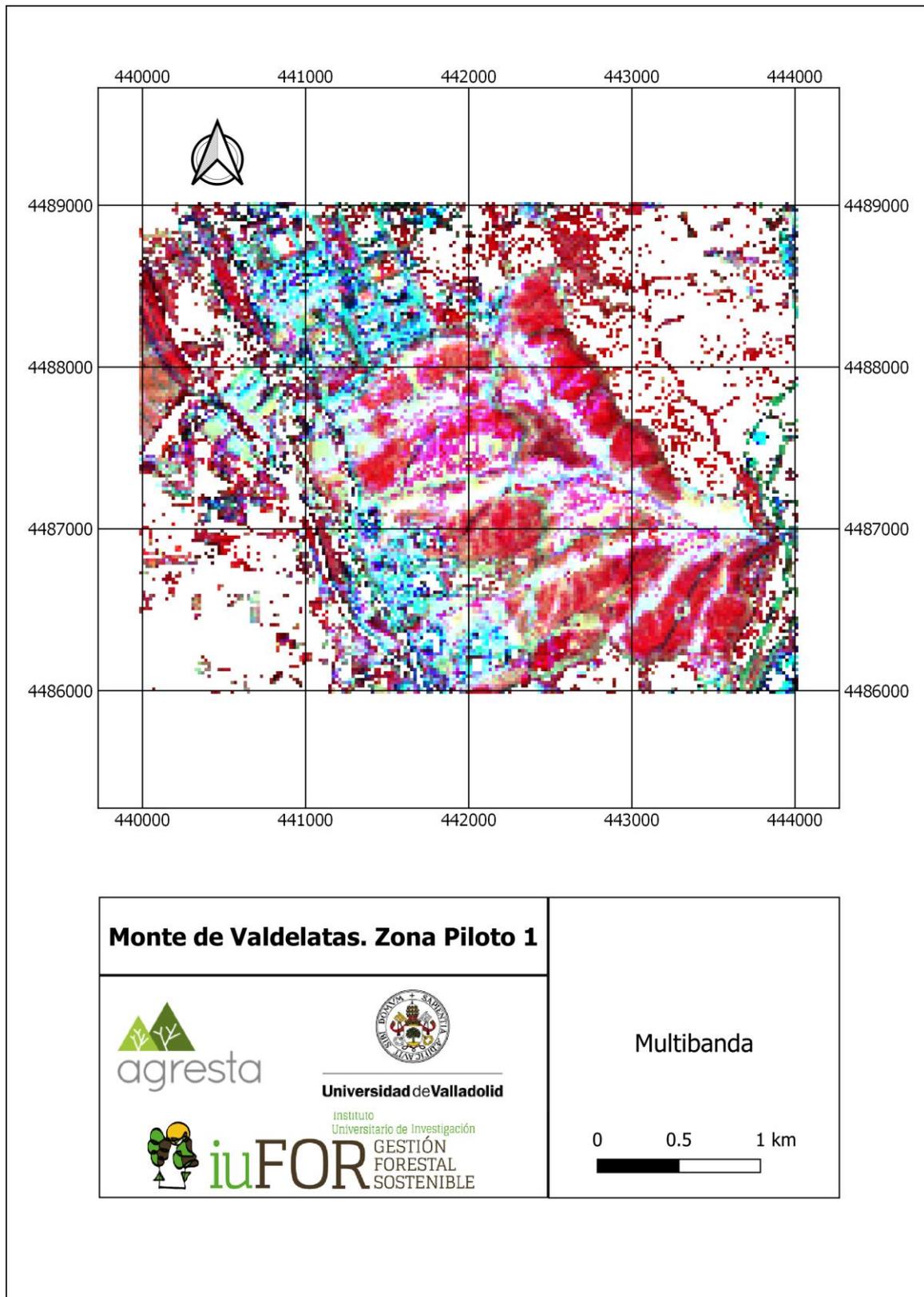
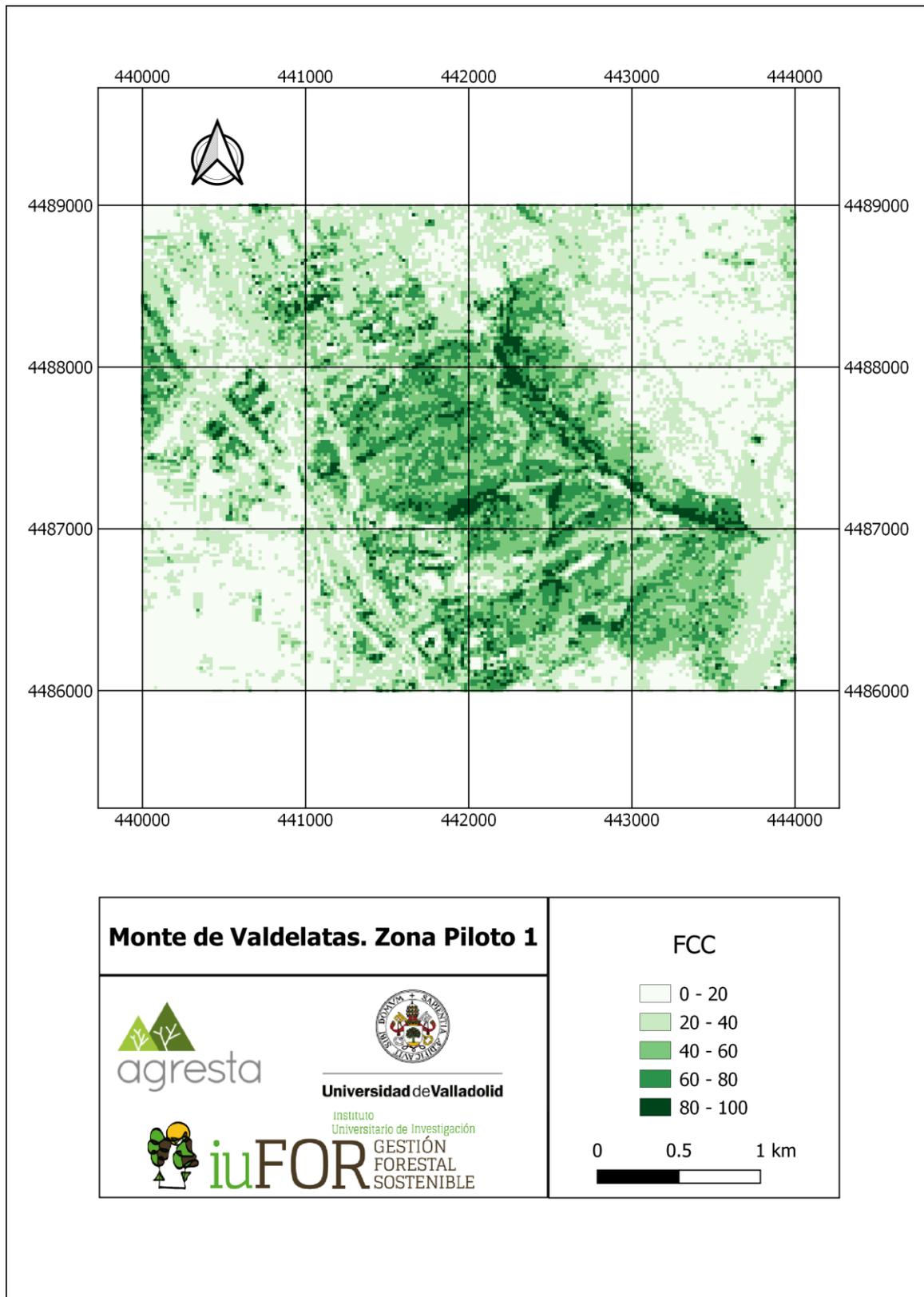


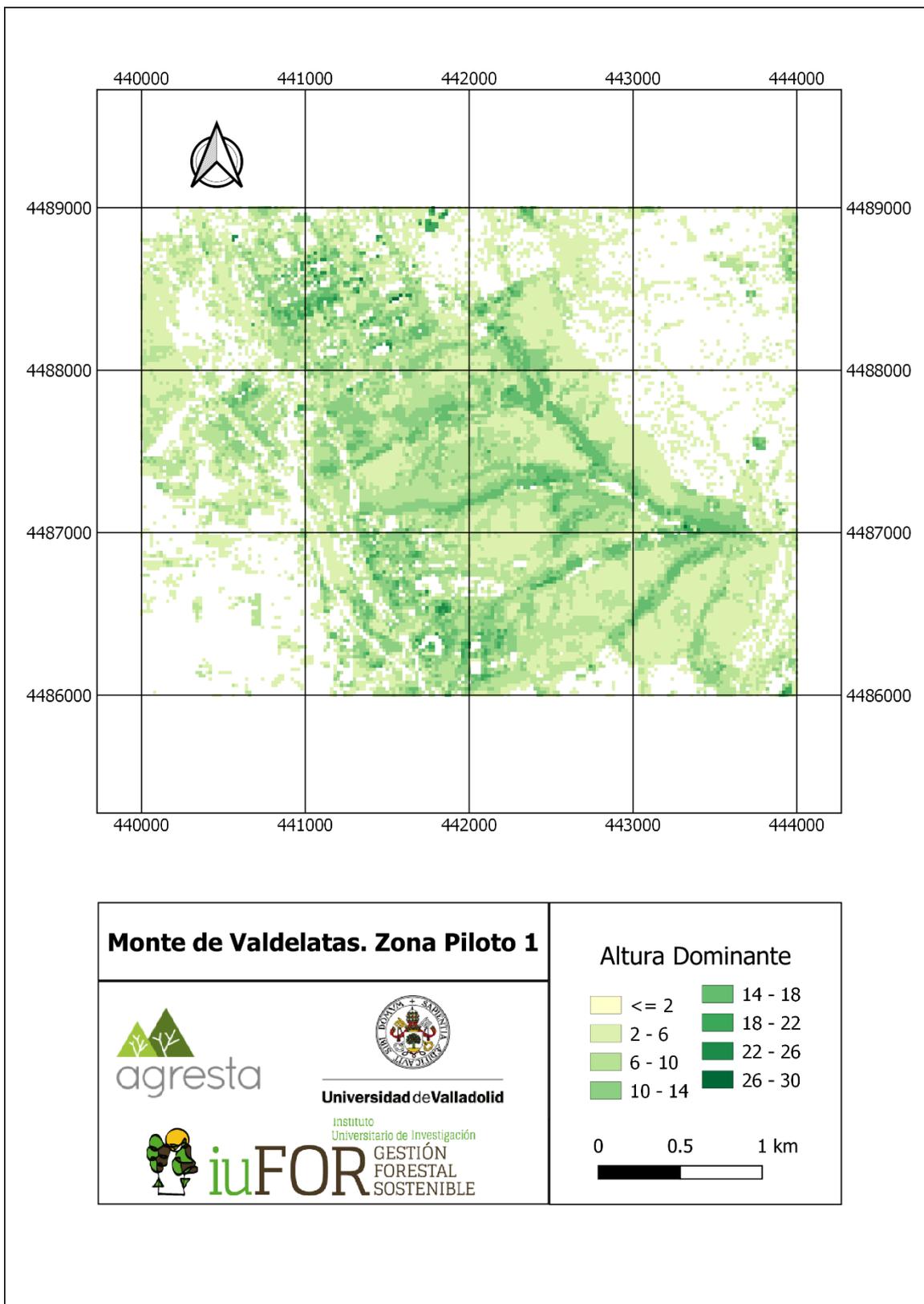
Figura 217. Índice de vegetación normalizado de Valdelatas.



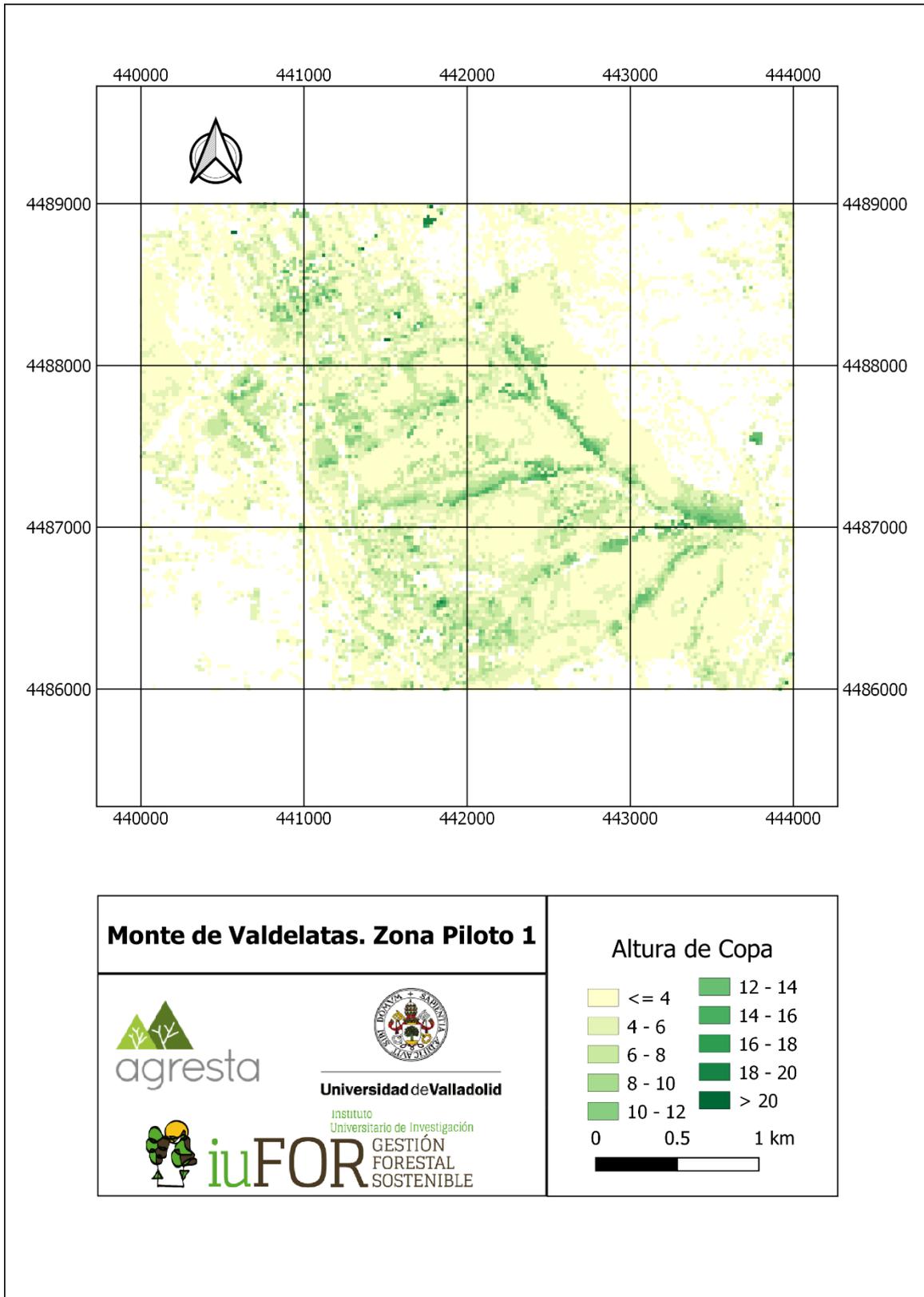
**Figura 3.** Ráster multibanda. Combinación del índice de vegetación normalizado, la altura dominante (P90), altura del primer verticilo vivo (P20) y la fracción de cabida cubierta (FCC)



**Figura 4.** Fracción de cabida cubierta (FCC). Tamaño de píxel 20 m. Altura mínima de procesamiento 2 metros. Clases empleadas para el procesamiento: 2, 3, 4, 5. Equivalentes a las clases de suelo y vegetación.

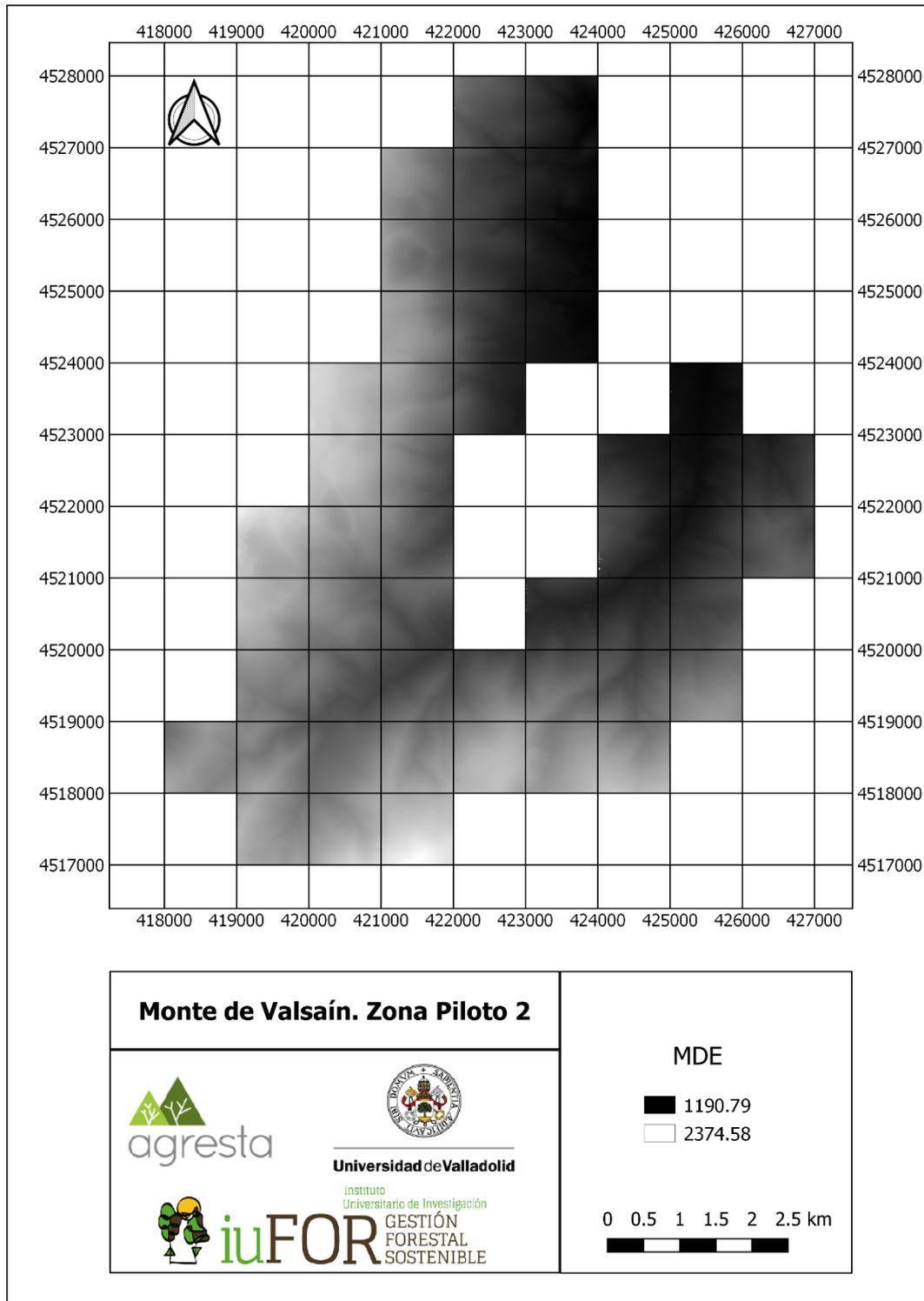


**Figura 5.** Altura dominante de la vegetación (P90). Tamaño de pixel 20 m. Altura mínima de procesado 2 metros. Clases empleadas para el procesado: 2, 3, 4, 5. Equivalentes al as clases de suelo y vegetación.

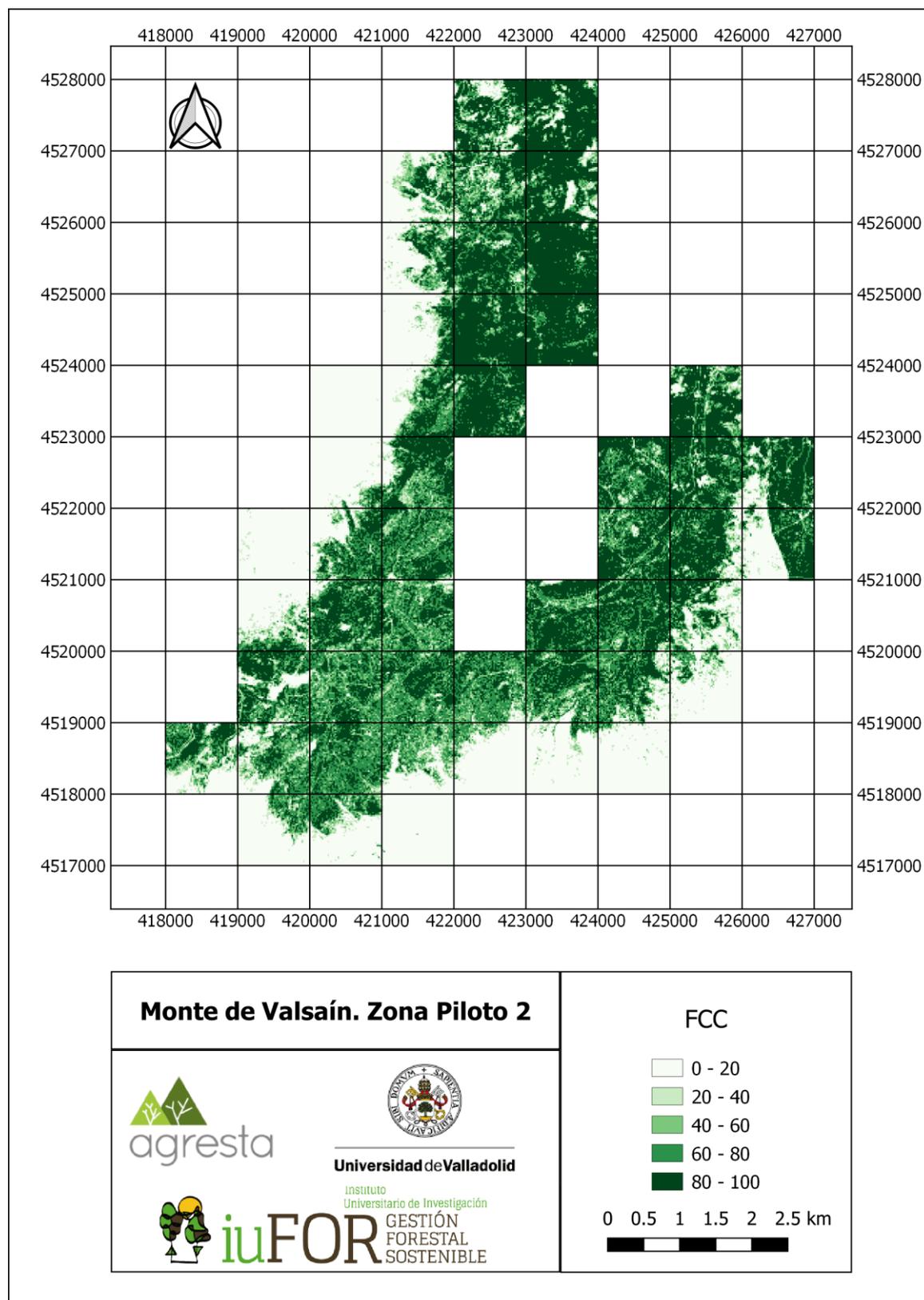


**Figura 6.** Altura del primer verticilo vivo (P20). Tamaño de pixel 20 m. Altura mínima de procesado 2 metros. Clases empleadas para el procesado: 2, 3, 4, 5. Equivalentes al as clases de suelo y vegetación.

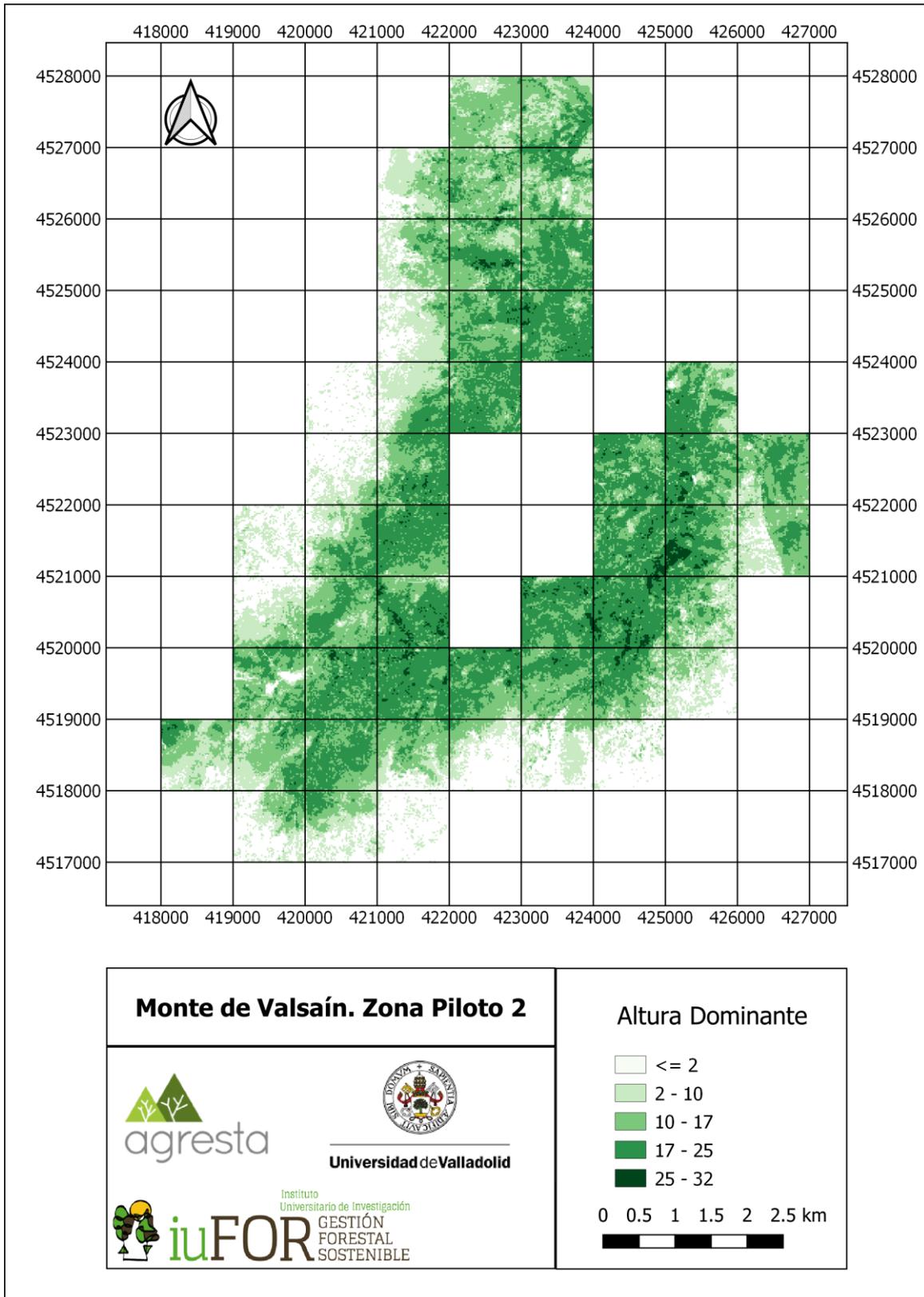
## ZONA PILOTO 2: Sierra de Guadarrama: Peñalara y Cabeza de Hierro



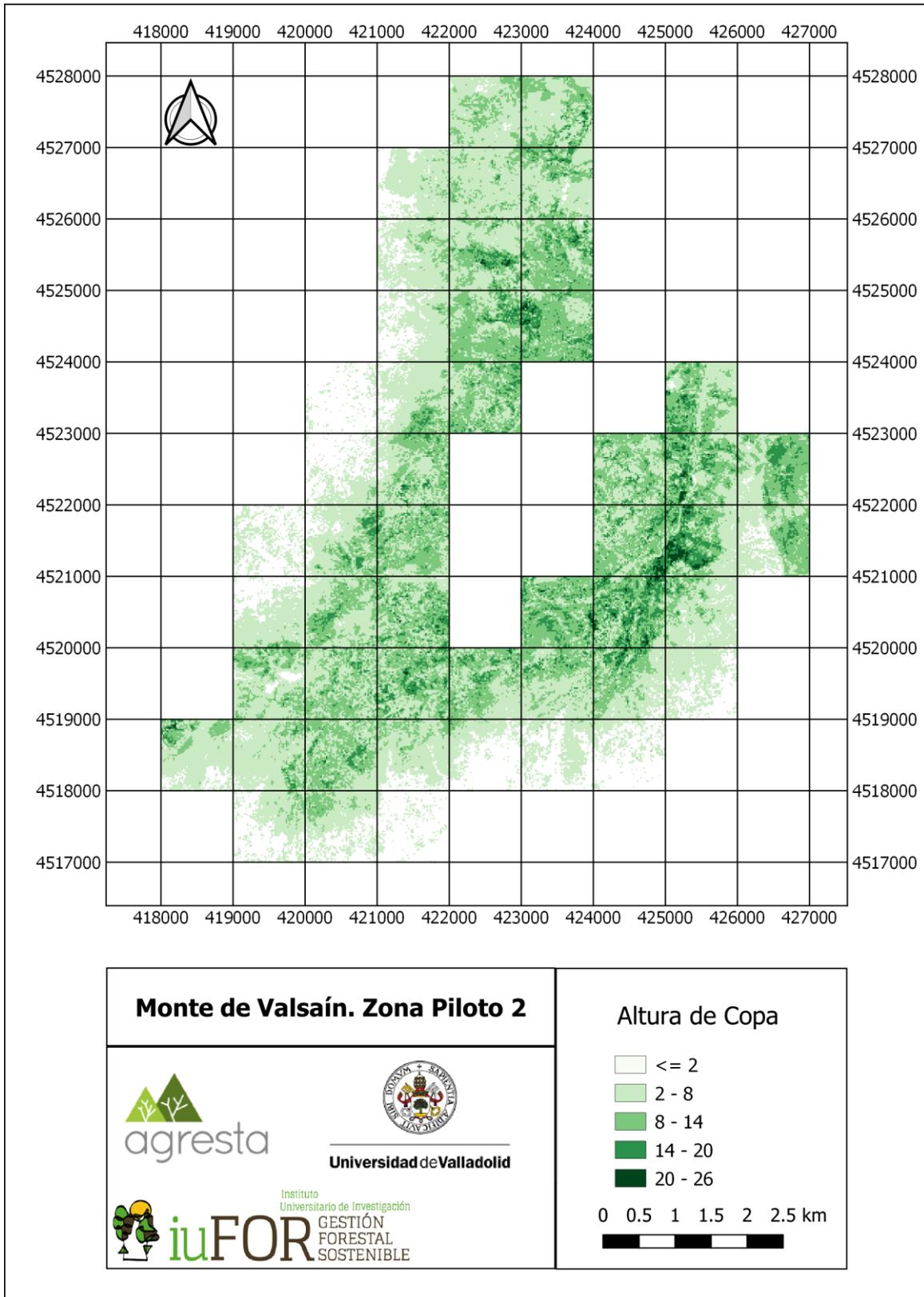
**Figura 7.** Modelo Digital de Elevaciones de la sierra de Guadarrama: Peñalara y Cabeza de Hierro. Clases empleadas en el proceso: 2 (suelo). Tamaño de pixel 2 metros



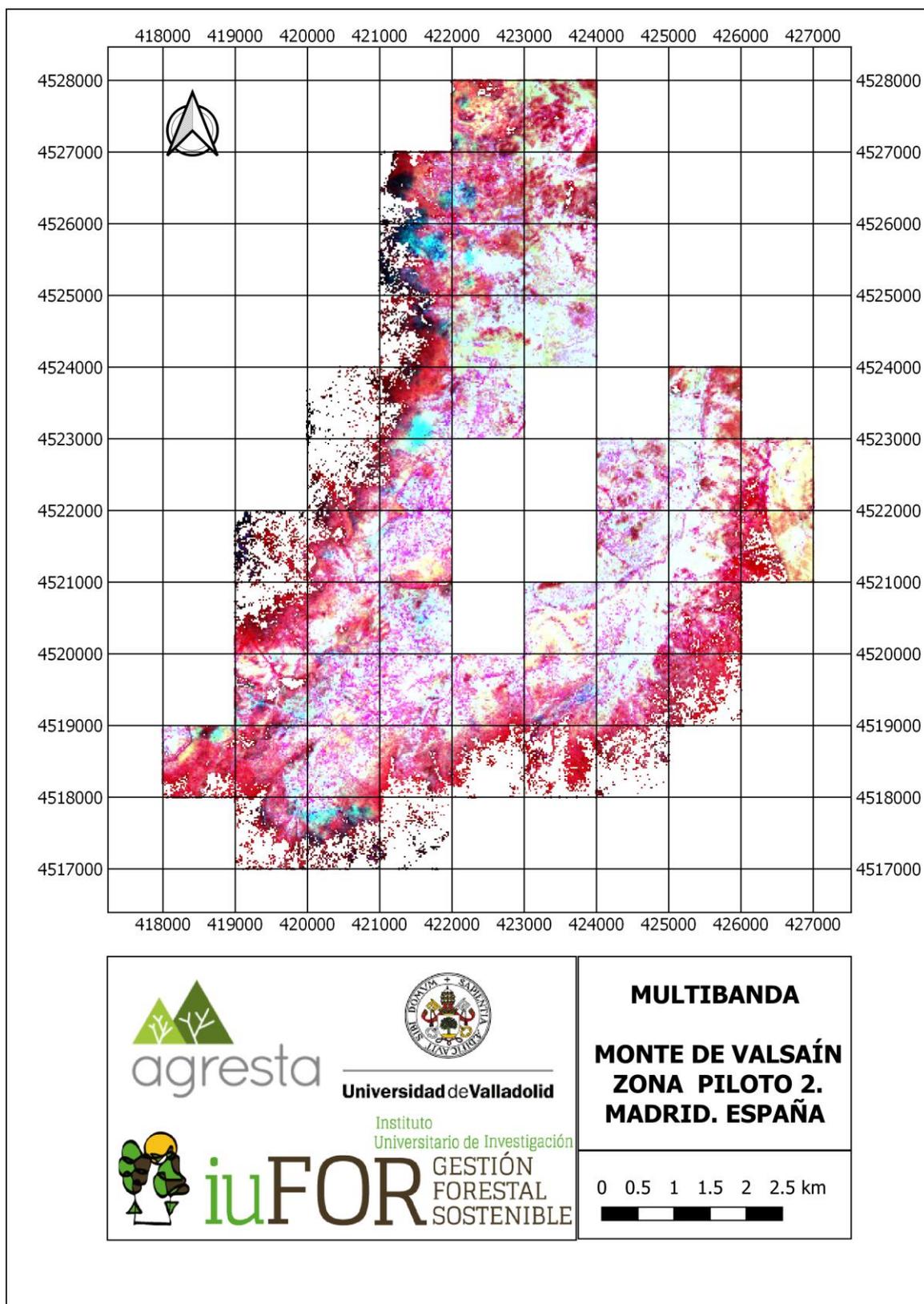
**Figura 8.** Fracción de cabida cubierta (FCC) la sierra de Guadarrama: Peñalara y Cabeza de Hierro. Tamaño de pixel 20 m. Altura mínima de procesado 2 metros. Clases empleadas para el procesado: 2, 3, 4, 5. Equivalentes al as clases de suelo y vegetación.



**Figura 918.** Altura dominante de la vegetación (P90) la sierra de Guadarrama: Peñalara y Cabeza de Hierro. Tamaño de píxel 20 m. Altura mínima de procesamiento 2 metros. Clases empleadas para el procesamiento: 2, 3, 4, 5. Equivalentes a las clases de suelo y vegetación.



**Figura 10.** Altura del primer verticilo vivo (P20) la sierra de Guadarrama: Peñalara y Cabeza de Hierro. Tamaño de pixel 20 m. Altura mínima de procesado 2 metros. Clases empleadas para el procesado: 2, 3, 4, 5. Equivalentes al as clases de suelo y vegetación.



**Figura 11.** Ráster multibanda Sierra de Guadarrama: Peñalara y Cabeza de Hierro. Combinación del índice de vegetación normalizado, la altura dominante (P90), altura del primer verticilo vivo (P20) y la fracción de cabida cubierta (FCC)

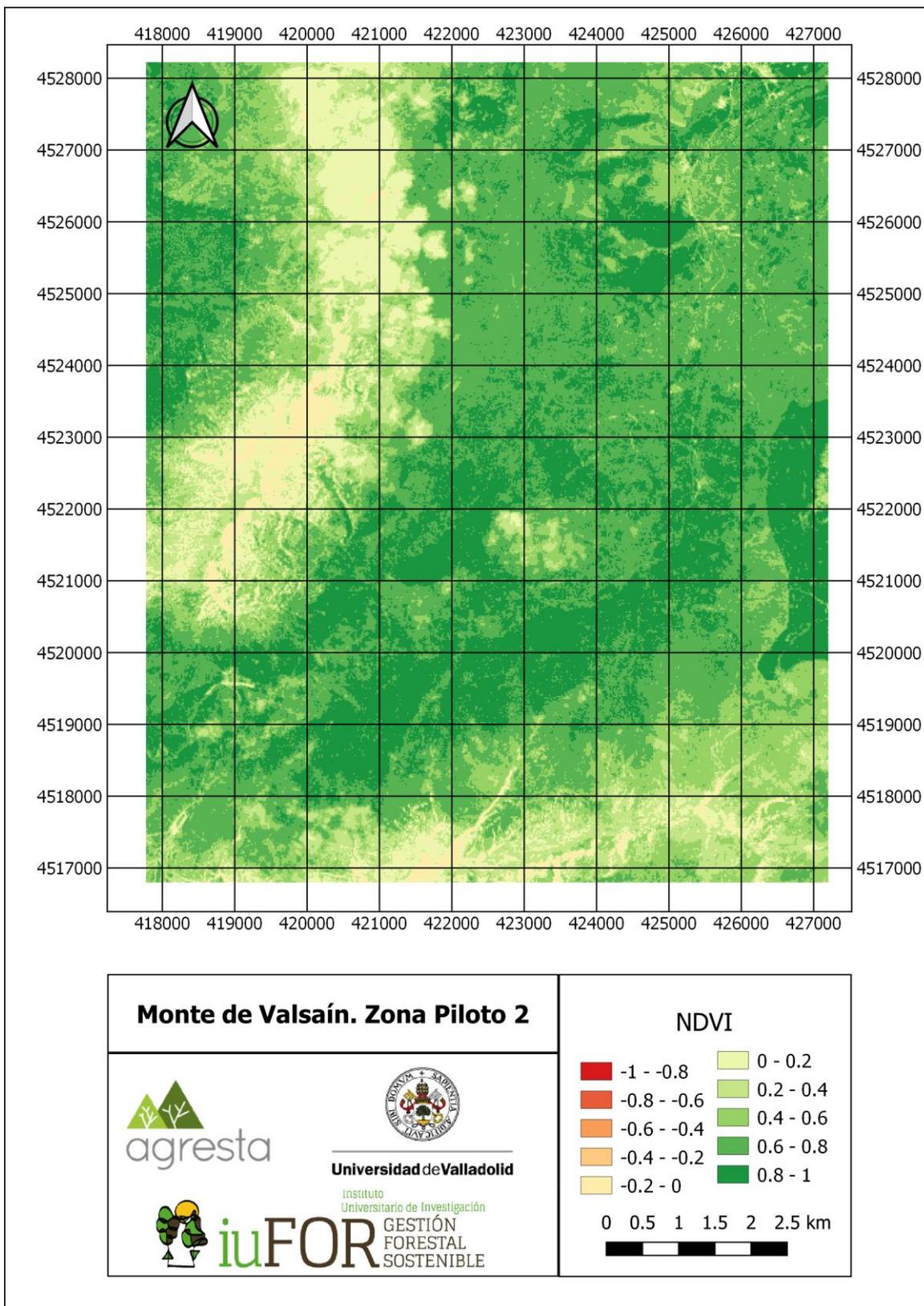


Figura 13. NDVI Cabecera del valle de Lozoya

## ANEJO 2. GUÍA DE USUARIO

### A2.1 Presentación

Esta herramienta pretende facilitar la gestión forestal sugiriendo posibles rodales de un territorio determinado, automatizando el proceso de procesamiento de datos LiDAR e imágenes satelitales. Esta aplicación pretende ser amigable con el usuario, por lo tanto, está implementada sobre la interfaz, que nos proporciona la API de QGIS 3x. Una de las características de este desarrollo es que se apoya sobre softwares gratuitas y abiertas, salvo FUSION (gratuito pero cerrado), con el objetivo de hacer esta aplicación lo más accesible y distribuible posible.

### A2.2 Requerimientos Técnicos

Para el correcto funcionamiento de la herramienta, es necesario descargar los softwares sobre los que se apoya:

**QGIS 3x:** sistema de información geográfica sobre el cual está programada la herramienta haciendo uso de su API. **Es importante que sea una versión 3x las versiones inferiores no podrán emplear la herramienta**, ya que a partir de esa versión QGIS implementa un uso total del lenguaje de programación Python para su api. En esta versión están incluidas librerías de procesamiento de datos espaciales como son GDAL y GRASS a través de las cuales accederemos posteriormente de forma interna con QGIS.

Descarga: <https://qgis.org/es/site/forusers/download.html>

**FUSION:** software de procesamiento de datos LiDAR desarrollado por el departamento de agricultura de los Estados Unidos. Es gratuito y se puede encontrar en: <http://forsys.cfr.washington.edu/fusion/fusionlatest.html>. A pesar de ser, un software cerrado la documentación incluida detallan las características de cada algoritmo. Para el funcionamiento de FUSION es necesario que descargues también LAsTools, software libre de procesamiento de LiDAR. Al descargar el software <https://rapidlasso.com/lasools/> es necesario buscar en la carpeta de archivos binarios (bin) el LASzip.dll y pegarlo en la carpeta de instalación de FUSION. De esta forma la instalación estará completa, aun así, en el enlace de descarga de FUSION, se explica al detalle la instalación.

**Orfeo Toolbox:** Finalmente descargaremos esta librería de algoritmo de tratamiento de imágenes satelitales y datos ráster y vectoriales. Emplearemos el algoritmo de segmentación de imágenes meanshift incluido en este distribuible. Recomendamos incluir la carpeta de este descargable en C:\ o C:\ Archivos de programa, ya que la herramienta buscará por defecto en estas rutas para acceder a los algoritmos, de no se rasí el usuario deberá especificar la ruta de la carpeta.

Descarga: <https://www.orfeo-toolbox.org/download/>

Por lo general todos estos softwares son fáciles de descargar y no debería presentar problemas a la hora de instalarlos.

### A2.3 Cómo Acceder A La Herramienta

Primero se deberá descargar la herramienta que se encuentra en el siguiente repositorio de GitHub: [https://github.com/fernandobezares/Segmentation\\_Tool/tree/master](https://github.com/fernandobezares/Segmentation_Tool/tree/master)

Abrimos QGIS 3x y nos dirigimos a la ventana de procesos y abrimos la caja de herramientas:

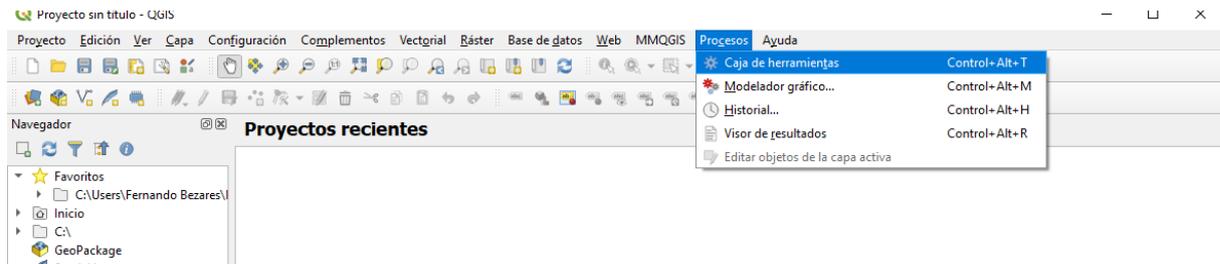


Figura 1. Abrir caja de Herramientas

Una vez abierta, incluimos el script de Python con la herramienta de segmentación. Clicaremos sobre el logo de Python y pincharemos sobre “Abrir modelo de script ...” de esta forma podrás acceder a la edición del script, y podrás correrlo pinchado sobre el triángulo verde (RUN)

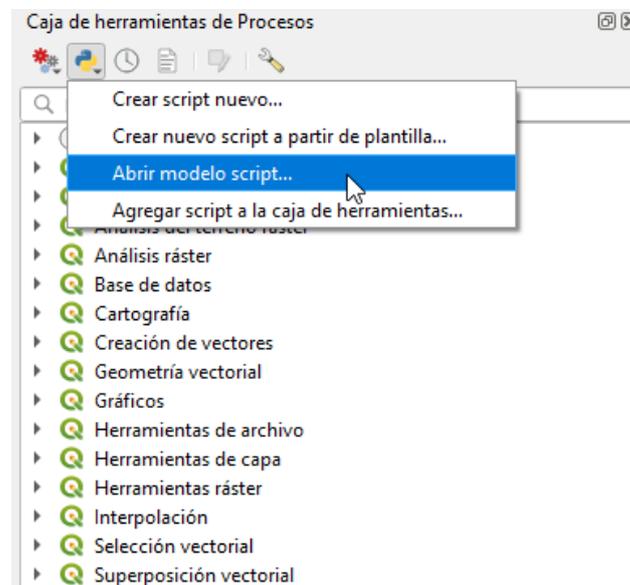


Figura 19. Abrir la edición del script

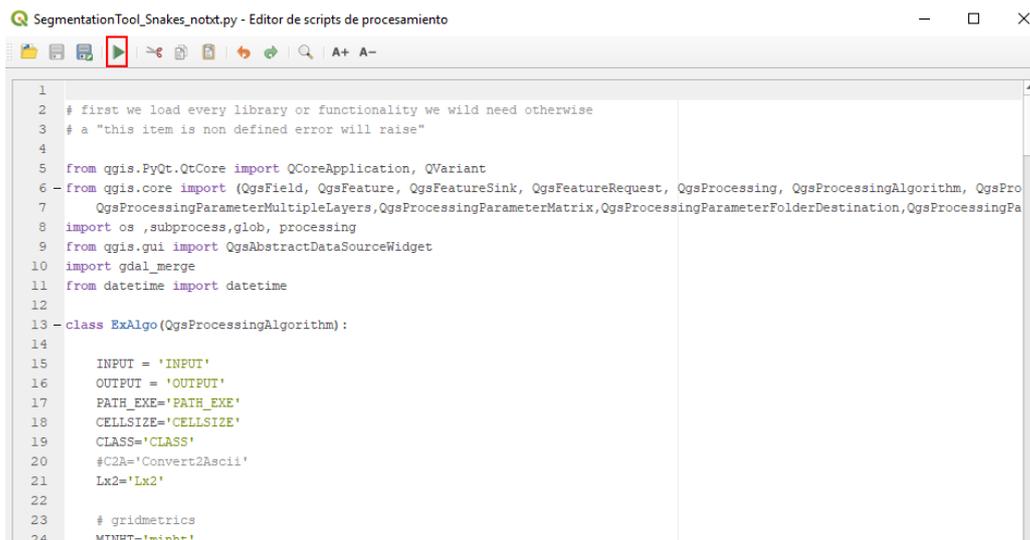


Figura 2. Pantalla de edición del script, en verde el botón para iniciar el script.

Otra opción es incluir el script en la caja de scripts de tal forma que te aparezca al abrir QGIS dentro de la caja de herramientas, como un algoritmo más, y que al pinchar en él no te abra la edición, sino que se ejecute. Veamos cómo se hace:

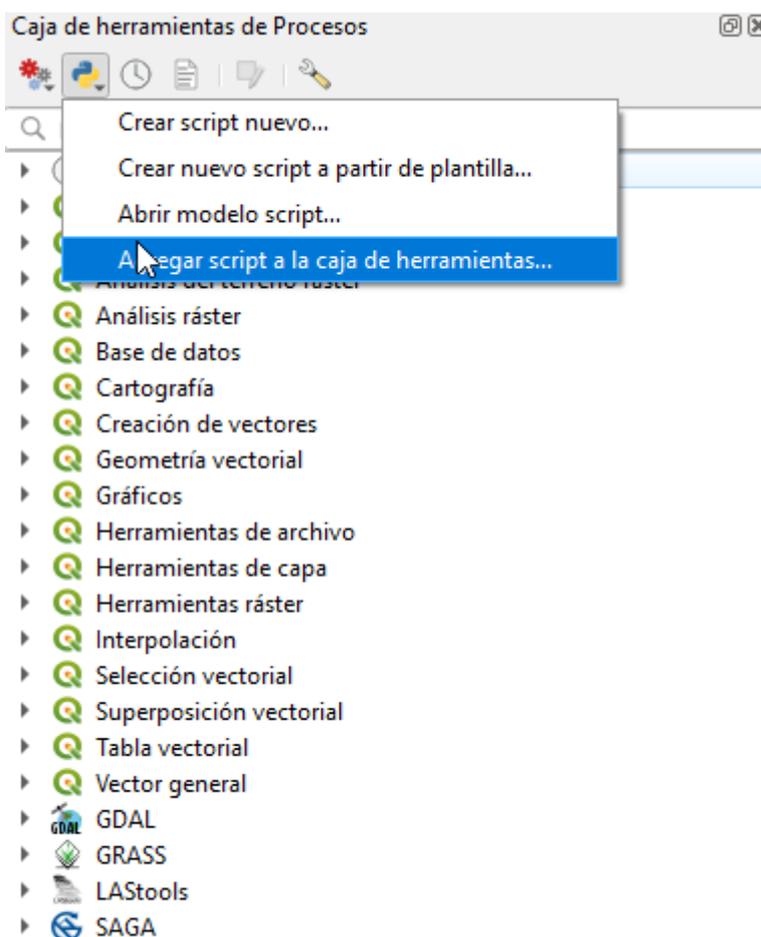


Figura 3. Agregar script a la caja de herramientas.

Una vez indicada la localización del script, aparecerá el logo de Python con nuestro script:

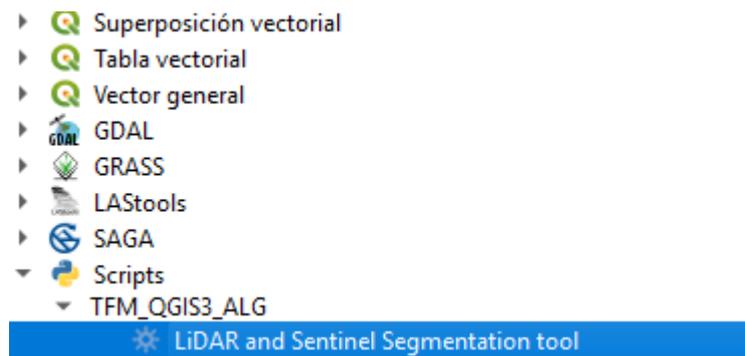


Figura 4. Script añadido

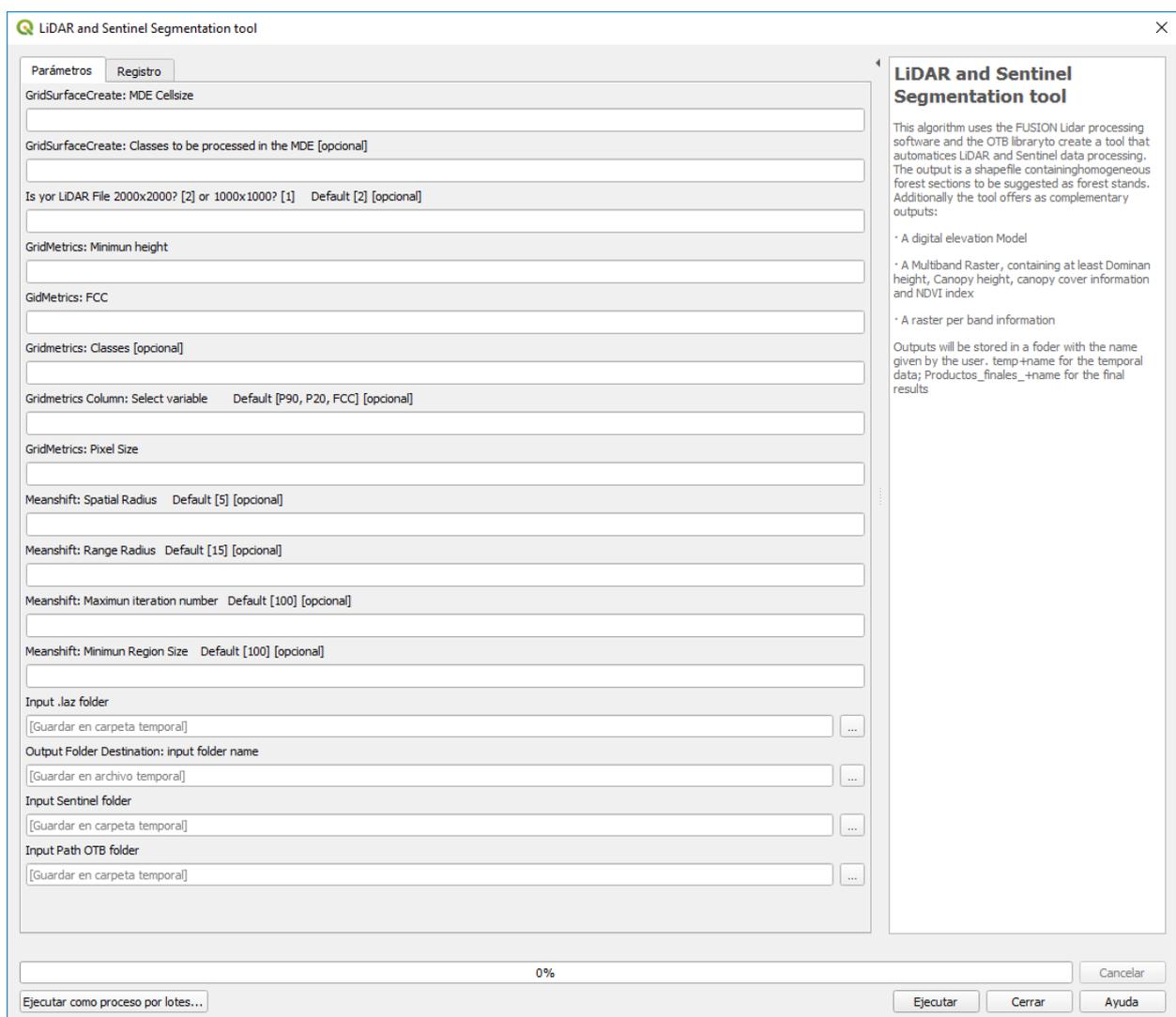


Figura 5. Interfaz de la Herramienta.

## A2.4 Estructura

La herramienta está formada de una ventana o interfaz principal varias entradas, abajo encontraremos los inputs y outputs y en la parte superior los parámetros con los que alimentaremos los algoritmos. En cada entrada indica qué algoritmo es, y si el valor de entrada es opcional o no. A la izquierda encontramos una pequeña descripción de lo que hace la herramienta. Por defecto, el script buscará la carpeta de instalación de FUSION y de OTB en **C:\ o en C:\ Archivos de programa**. Si no los encontrara, el programa pediría la ubicación de la carpeta de FUSION o la de OTB. **Es importante, que, si no lo encuentra, sólo se indique la ruta de la carpeta, no es necesario que indicar subcarpetas**, el programa encontrará lo que necesite desde la carpeta raíz ejemplo: C:\Agesta\TFM\OTB-6.6.1-Win64 o C:\FUSION. El número de casillas variará, en función de si el programa encuentra las carpetas o no.

## A2.5 Inputs

Se pedirá al usuario indicar la carpeta donde se encuentren **los archivos .laz y o la carpeta de Sentinel donde se encuentren las escenas a 20 m de precisión (archivos .jp2)**. **Los archivos .laz deberán ser aquellos descargados de la página del CNIG correspondientes al PNOA**, ya que la herramienta está preparada para trabajar con los nombres sistematizados del PNOA.

## A2.6 Outputs

Es recomendable que, como output, se escriba el nombre de la zona de estudio o algo identificativo la herramienta automáticamente generará carpetas donde almacene los resultados. Se pueden emplear 2 tipos de datos: LiDAR y Sentinel. Si se utiliza **LiDAR + Sentinel 2**, todos **los resultados finales** serán almacenados en una carpeta llamada **'Productos finales' + el nombre introducido por el usuario**. Además, se generará una carpeta de **archivos temporales** con el nombre **'temp'+el nombre definido por el usuario**. Dentro de esta carpeta se encontrarán los archivos temporales necesarios para la ejecución de la herramienta. Si sólo se utiliza LiDAR, se generarán dos carpetas, **'Productos\_finales\_LiDAR' + el nombre introducido por el usuario y 'temp\_LiDAR'**. Sin embargo, con Sentinel, sólo se generará una carpeta **'Productos\_finales\_Sentinel**. **No borre nada de estas carpetas mientras la herramienta está procesando**. En las carpetas temporales, también se incluirán tres subcarpetas:

- 1) MDE: se incluyen los dtm y ascii para cada laz
- 2) Ráster: los ascii para cada variable procesada por cada laz
- 3) Métricas: métricas individuales de cada dtm

La herramienta ofrece como productos finales al terminar la ejecución, las siguientes capas:

- f) Un modelo digital de elevaciones del terreno
- g) Tres rastres que representan diferentes variables forestales, por defecto la herramienta ofrece:
  - a. La altura dominante
  - b. La altura del primer verticilo vivo
  - c. Fracción de cabida cubierta
- h) Un ráster con el índice de vegetación normalizado para la zona de estudio
- i) Una segmentación preliminar sin un suavizado de bordes
- j) Por último, a propuesta de los rodales con un suavizado de bordes (Chaiken y Snakes)

## A2.7 Recomendaciones

No introducir archivos o nombres de outputs con caracteres especiales (ejemplo: Peñalara), pues podrían dar un error en el reconocimiento de los archivos. FUSION **no reconoce los espacios en las rutas**, por lo que es aconsejable que los nombres de las rutas de salida no incluyan espacios.

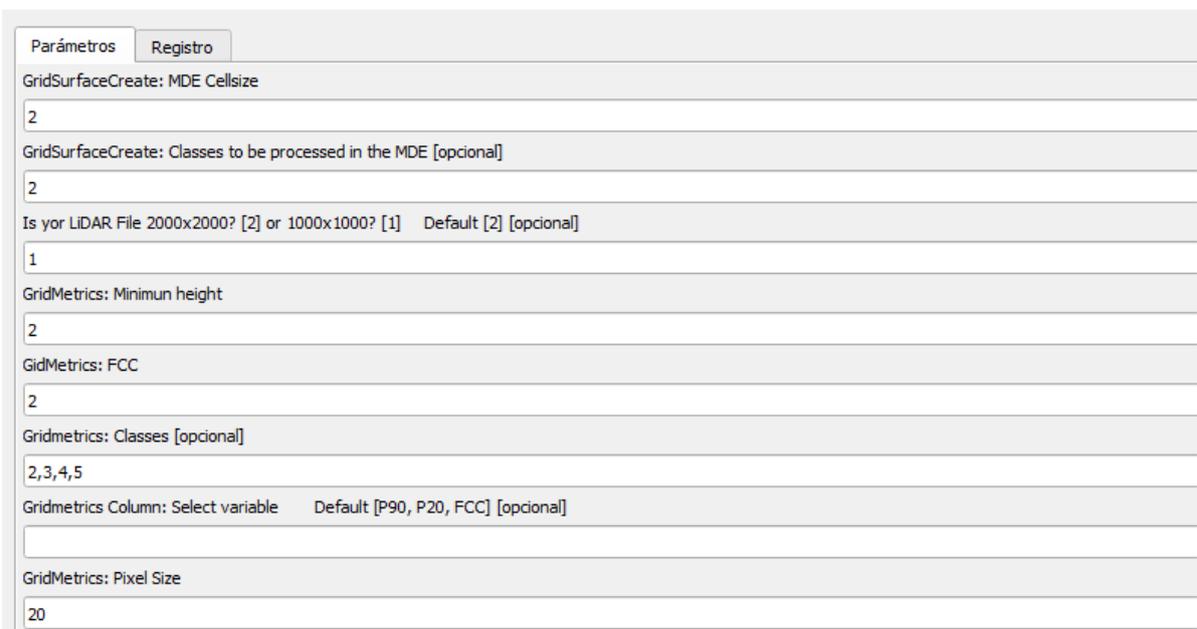
## A2.8 Ejemplo De Utilización De La Herramienta

Abrimos la herramienta como se especificó anteriormente, una vez en la pantalla principal rellenamos los valores opcionales que deseemos. La herramienta, también está preparada para funcionar únicamente con datos LiDAR o datos de Sentinel. Al introducir uno de los dos datos la herramienta reconocerá cuáles son y operará acordeamente. En este ejemplo operaremos con LiDAR y Sentinel.

Los parámetros introducidos, correspondientes al procesado de LiDAR para este ejemplo son:

- Tamaño de celda del MDE: 2
- Clases LiDAR a procesar: en el PNOA 2 es el suelo.
- Tamaño del archivo LiDAR: en el caso de este ejemplo es 1000 x 1000 pero la mayoría de los casos es 2000x2000. El usuario deberá abrir un visor o consultar en los metadatos del archivo para conocer estos detalles. Personalmente recomiendo abrir el visor y medir uno de los lados.
- Altura mínima de procesado: 2 metros, queremos obviar los matorrales
- FCC: 2
- Clases a procesar en las métricas: suelo (2) y vegetación (3,4,5)
- Tamaño de pixel de las métricas: 20 metros

LiDAR and Sentinel Segmentation tool



Parámetros	Registro
GridSurfaceCreate: MDE Cellsize	
2	
GridSurfaceCreate: Classes to be processed in the MDE [opcional]	
2	
Is yor LiDAR File 2000x2000? [2] or 1000x1000? [1] Default [2] [opcional]	
1	
GridMetrics: Minimun height	
2	
GridMetrics: FCC	
2	
Gridmetrics: Classes [opcional]	
2,3,4,5	
Gridmetrics Column: Select variable Default [P90, P20, FCC] [opcional]	
GridMetrics: Pixel Size	
20	

Figura 6. Parámetros LiDAR. Ejemplo.

Parámetros para la segmentación meanshift:

En este caso dejamos por defecto los parámetros, para ver si el resultado obtenidos con esta configuración es adecuados. Por lo general, **mayores valores en los spatial radius y range radius supondrán mayores tamaños de rodal.**

Meanshift: Spatial Radius Default [5] [opcional]

Meanshift: Range Radius Default [15] [opcional]

Meanshift: Maximum iteration number Default [100] [opcional]

Meanshift: Minimum Region Size Default [100] [opcional]

**Figura 7.** Parámetros algoritmo de segmentación. Ejemplo

## Inputs y outputs

Recordamos que **se recomienda no usar caracteres especiales**. Se indica donde están los .laz de la zona de estudio y las imágenes Sentinel (en el ejemplo: C:\Agresta\TFM\Datos\L1C\_T30TVK\_A010559\_20190315T110702\S2B\_MSIL1C\_20190315T105819\_N0207\_R094\_T30TVK\_20190315T144132.SAFE\GRANULE\L1C\_T30TVK\_A010559\_20190315T110702\IMG\_DATA). **En la carpeta de Sentinel deberán encontrarse los archivos .jp2**

Indicamos un nombre para la carpeta de salida, ignore el .dtm, en este caso escribimos la zona de estudio (valdelatas), en el ejemplo sería valdelatas.dtm. Finalmente, indicamos dónde se encuentra la carpeta de instalación de OTB.

Input .laz folder  
C:\Agresta\TFM\Datos\laz\_valdelatas

Output Folder Destination: input folder name  
C:\Agresta\TFM\Datos\results\_seg\_tool\_copia\valdelatas.dtm

Input Sentinel folder  
10702\S2B\_MSIL1C\_20190315T105819\_N0207\_R094\_T30TVK\_20190315T144132.SAFE\GRANULE\L1C\_T30TVK\_A010559\_20190315T110702\IMG\_DATA

Input Path OTB folder  
C:\Agresta\TFM\OTB-6.6.1-Win64

**Figura 8.** Inputs y outputs herramienta. Ejemplo

Una vez estén todos los parámetros indicados, ejecutamos:

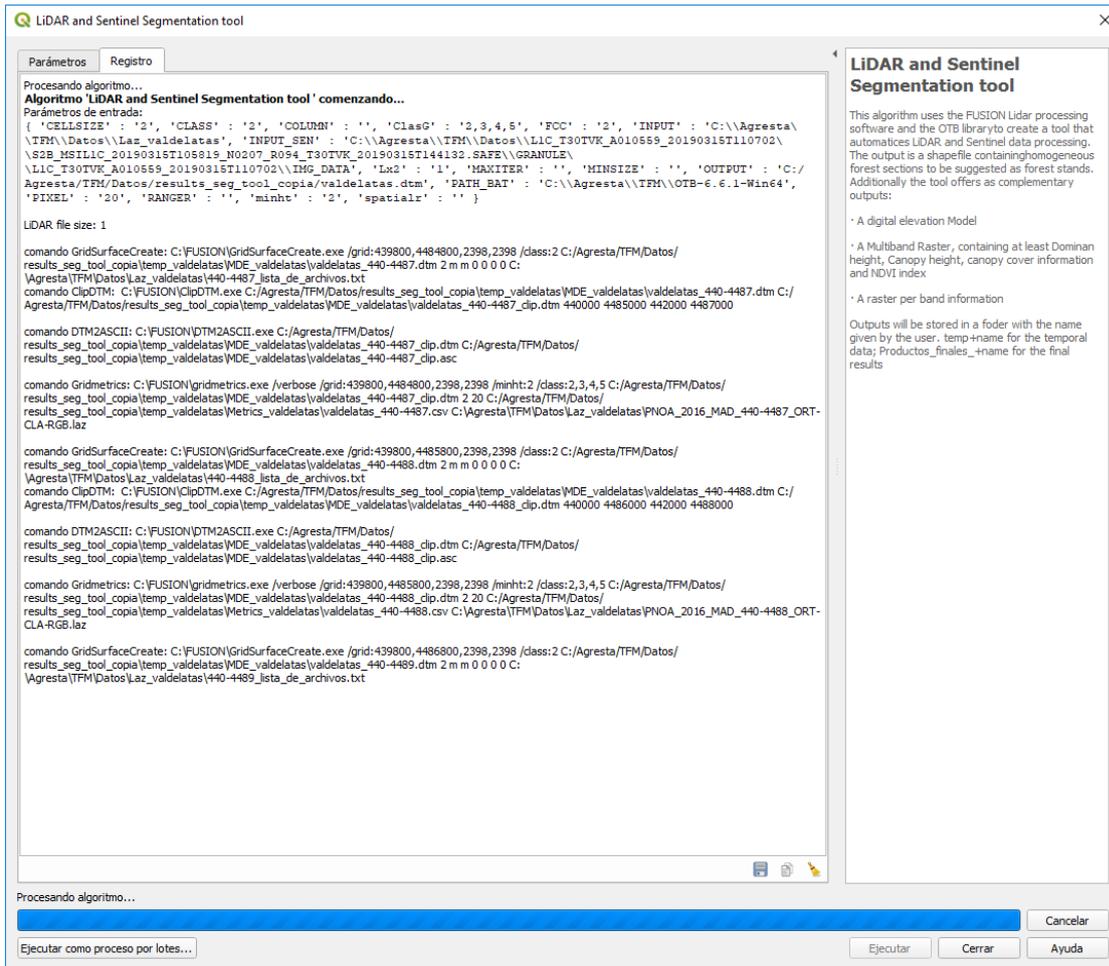


Figura 9. Ejemplo de ejecución de la herramienta.

Las impresiones en pantalla indican los comandos que se están enviando a los ejecutables que se encuentran en la carpeta de FUSION. El proceso finalizará cuando salga el mensaje 'PROCESO FINALIZADO', eso nos asegurará que están nuestros archivos en las carpetas.

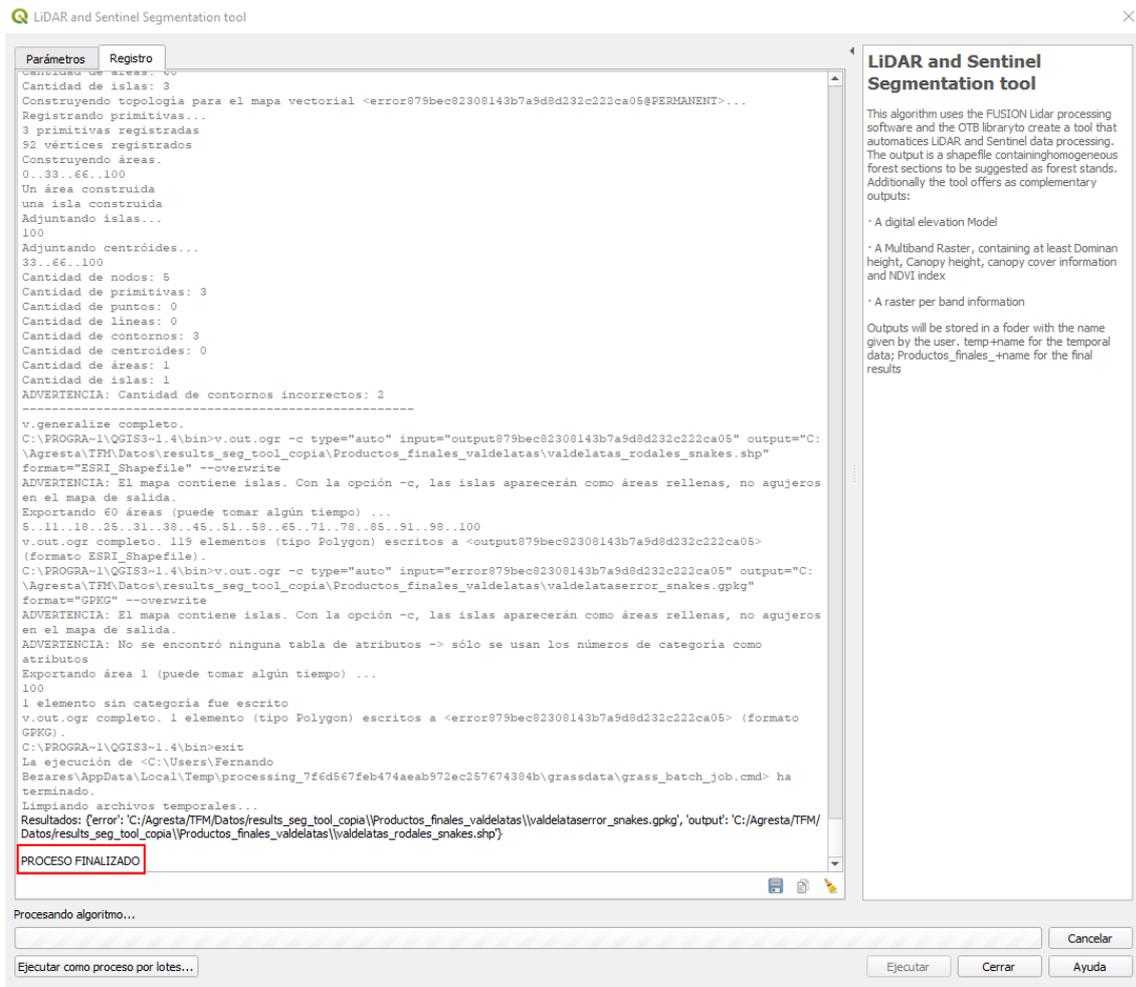


Figura 10 . Final de ejecución.

