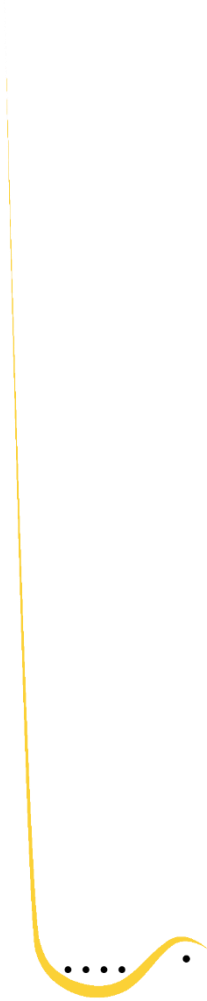


Proyecto dasolidar – fase beta	
Integración del Lidar en la gestión del medio natural en Castilla y León	
	<p>Este documento recoge información variada del proyecto dasolidar, desde una guía rápida de “primeros pasos”, hasta los detalles estadísticos de obtención de los modelos. Por lo tanto, no es un documento para leer “así de seguido”, sino para consultar cosas concretas.</p> <p>En un futuro próximo, las consultas relacionadas con esta iniciativa (dasolidar) se hará de forma más sencilla, en forma de consulta – respuesta, desde el propio proyecto LidarQgis (será una de las utilidades del complemento de Qgis “dasoraster”, en elaboración).</p> <p>Hasta que eso esté disponible, aquí vamos a ir recogiendo contenidos variados, con lo que este documento se irá completando y mejorando de forma progresiva.</p> <p>Documento en elaboración: contenidos pendientes:</p> <ul style="list-style-type: none"><li>• Fiabilidad – intervalos de confianza</li><li>• Como comparar dasolidar con otras fuentes</li><li>• Nueva estructura de directorios en LidarData</li><li>• Terminar de metadatar las capas</li><li>• Incluir en el proyecto las capas de zonificación, de fechas y de modelos</li><li>• Complemento Qgis dasoraster</li><li>• Modelos, rangos de validez y estadísticos</li><li>• Intervalos de confianza en muestreo</li><li>• Control de calidad en IOFs con Lidar</li><li>• Uso de dasolidar en IOFs</li><li>• Etc.</li></ul> <p>Últimos contenidos o mejoras añadidos al documento:</p> <ul style="list-style-type: none"><li>• Guía rápida (1)</li><li>• Fechas de vuelo (4.3)</li><li>• Zonificación (7.2)</li><li>• </li></ul>
	Versión: 08/11/2024
	Manual de consulta

# 1 Guía rápida del proyecto dasoLidar

## 1.1 Primeros pasos

¿Qué es el dasoLidar?

Una iniciativa para poner a disposición de los técnicos productos y herramientas Lidar destinados a la gestión del medio natural.

¿Por dónde empiezo?

Cargando el proyecto LidarQgis.

¿Dónde está ese proyecto?





Se accede mediante un enlace directo  que está en el escritorio<sup>1</sup>

¿Está disponible para todos?

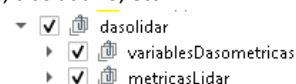
Está disponible para los que están en la lista de usuarios dasoLidar<sup>2</sup>

¿Qué puedo hacer una vez abierto el proyecto?

1. Se puede empezar cargando la nube de puntos Lidar y mostrando un perfil:

- Haz zoom a un sitio concreto
- Hacer click en la capa "cargar\_nubeDePuntos\_LidarPNOA2" para activarla.
- Elegir el botón  y hacer click sobre un punto del mapa ⇒ Eso carga una capa con los puntos Lidar
- Menú Ver -> Perfil de elevación  Trazar una línea con el ratón:   
Terminar la línea pulsando: 
- En el perfil se pueden hacer mediciones, cambiar en ancho del perfil (tolerancia), trasladarlo, etc.

2. Se pueden visualizar en el proyecto capas ráster con información dasométrica como el volumen estimado de madera en pie o la altura dominante Lidar. Esas capas están dentro del grupo dasoLidar.



¿Qué variables dasométricas puedo consultar?

Se puede empezar por la altura dominante Lidar, que es una métrica Lidar, y seguir con el Volumen de madera en pie, que es una variable dasométrica estimada usando Lidar y datos de parcelas.

Quiero planificar unas intervenciones selvícolas (cortas, claras, clareos, etc.), ¿me puedo apoyar en esta información?

Estas capas se pueden consultar visualmente y para eso no hace falta saber mucho más que las unidades en que están expresadas (altura: metros; volumen: m3/ha). Si se van a hacer cálculos de existencias, crecimientos, etc. hay que hacerlo teniendo en cuenta la interpretación y precisión de cada métrica o variable. En todo caso, siempre se puede contrastar las estimaciones dasolidar con datos de parcelas con objeto de validar si el modelo usado funciona bien en el monte elegido.

¿Qué es eso de la ubicación de red; dónde está toda esta información?

<sup>1</sup> Ese enlace directo se crea al iniciar Qgis (después de que el usuario ha sido dado de alta). El proyecto y las capas que usa están en una ubicación de red compartida por todos los SSTT

<sup>2</sup> Se ha diseñado una implantación progresiva, con objeto de poder atender las incidencias del proyecto en su fase de pruebas (beta). En todo caso, cualquier técnico que quiera acceder a esta información y participar en esta fase beta, puede mandar un correo electrónico a [benmario@jcyl.es](mailto:benmario@jcyl.es) indicando:

Asunto: Usuario DasoLidar [cod332]

Sustituir [cod332] por el código 332 personal de acceso a la red

Texto del mensaje: Nombre y apellidos

Teléfono (incluido código de acceso directo)

Servicio territorial y unidad/sección en la que trabaja

Funciones principales en el puesto de trabajo

Si tiene previsto o desea prospectar algún uso concreto de los productos y herramientas Lidar.

Es un espacio de almacenamiento que nos ha habilitado informática para este proyecto, igual que cualquier otra unidad de red, pero a la que, por el momento, no asignamos letra de unidad.

¿Hay más documentación además de esta guía rápida?

Este ManualDasolidar.pdf está en la ubicación de red y se puede cargar desde Qgis.

## 1.2 ¿Qué más se puede hacer con dasolidar?

**Pendiente:** en este apartado incluiremos, sin entrar en detalles, algunos ejemplos de uso de los productos y herramientas dasolidar. El objetivo es sólo dar ideas de las posibilidades que tiene trabajar con productos derivados del Lidar.

### 1.2.1 Ejemplo relacionado con IOFs

...

### 1.2.2 Ejemplo relacionado con lotes de corta

...

### 1.2.3 Ejemplo relacionado con ...

...

### 1.2.4 Explorar que datos hay en la ubicación de red

En principio no será necesario bucear en la ubicación de red (LidarData), porque la idea es trabajarlo todo dentro Qgis. No obstante, se puede bucear en este directorio y ver donde están:

- Los ficheros con las nubes de puntos Lidar (lazFiles),
- Las capas dasolidar (ráster con métricas Lidar y variables dasométricas),
- Los proyectos de Qgis (hay más)
- Las mallas con enlaces directos a los lazFiles
- La documentación
- La cocina y ficheros auxiliares de gestión del proyecto

## 1.3 ¿Qué puedo encontrar en este manual de consulta?

Cuando uno toma contacto con este proyecto es probable que surjan preguntas como:

- ¿Por qué la información dasométrica está en capas ráster y no en vectoriales?
- ¿En qué unidades están expresadas las distintas variables?
- ¿Me puedo fiar de los valores que me dan estas capas?
- ¿De dónde sale esta información? ¿Es pública? ¿Cómo y quién ha generado estas capas?
- ¿Qué es eso del Lidar-PNOA2? ¿Hay PNOA1 o PNOA3? ¿Qué significa PNOA?
- ¿Qué otras capas se pueden consultar?
- ¿Puedo hacer cálculos con estas capas para un rodal o lote concreto, en vez de mirar píxel a píxel?
- ¿Son datos actuales?
- Si el Lidar son nubes de puntos ¿esos píxeles de las capas ráster son los puntos Lidar?
- ¿Puedo visualizar las nubes de puntos Lidar? ¿En 3D?
- ¿Alguna de estas capas me informa de las especies arbóreas presentes?
- ¿Cómo comparo mis datos de campo con los de estas capas?
- ¿Hay una guía rápida que pueda leer para hacerme una idea general de todo esto?

Estas preguntas tocan aspectos muy distintos del proyecto dasolidar y este documento pretende dar respuesta a todos ellos y a otras cuestiones más técnicas, que no son necesariamente del interés de los

usuarios del proyecto dasolidar, pero que es información relevante que debe estar a mano. No es necesario leer ni entender todo lo que aquí se cuenta.

La respuesta a esta última pregunta es: **si, hay una guía rápida y una “chuleta”, que están, como todo lo que aquí se cuenta, en el proyecto LidarQgis.**

## INDICE

1	Guía rápida del proyecto dasoLidar.....	2
1.1	Primeros pasos.....	2
1.2	¿Qué más se puede hacer con dasoLidar?.....	3
1.2.1	Ejemplo relacionado con IOFs .....	3
1.2.2	Ejemplo relacionado con lotes de corta .....	3
1.2.3	Ejemplo relacionado con .....	3
1.2.4	Explorar que datos hay en la ubicación de red.....	3
1.3	¿Qué puedo encontrar en este manual de consulta?.....	3
2	Proyecto dasoLidar: introducción.....	8
2.1	Planteamiento.....	8
2.2	Hoja de ruta .....	8
2.3	¿Dónde está todo esto? .....	10
2.3.1	Puerta de entrada:.....	10
2.3.2	Pero ¿dónde está el proyecto, las capas, las herramientas y demás? .....	10
2.3.3	¿Por qué usamos una ruta tan rara y no una unidad de red? .....	10
2.3.4	El acceso directo LidarQgis .....	10
2.3.5	Lista de usuarios dasoLidar .....	11
2.3.6	Versiones alfa, beta y realease .....	11
3	LidarQgis.....	11
3.1	Presentación .....	11
3.2	Las capas del proyecto LidarQgis .....	12
3.2.1	Malla 2k para cargar nubes de puntos .....	12
3.2.2	Trabajar con las nubes de puntos Lidar en Qgis: visualización.....	13
3.2.3	Capas ráster con información dasométrica .....	14
4	LidarData .....	15
4.1	Directorio dasoLidar.....	16
4.1.1	Contenidos.....	16
4.1.2	Métricas Lidar (ML) y variables dasométricas (VD) .....	16
4.1.2.1	Aclaraciones previas .....	16
4.1.2.2	Listado de de métricas Lidar y variables dasométricas.....	17
4.2	Directorio PNOA2.....	18
4.2.1	Cuadrantes de vuelo.....	18
4.3	Fechas de vuelo.....	19
4.3.1	Los ficheros Lidar (lazFiles) .....	21
4.4	¿Qué más hay en LidarData? .....	22
4.4.1	.aux .....	22

5	Origen de los datos brutos: PNOA, IFN, MFE .....	22
5.1	El Plan Nacional de Ortofotografía Aérea (PNOA) .....	22
5.2	El IFN4 .....	22
5.3	El MFE25 .....	22
6	Procesado de lazFiles y obtención de las métricas Lidar .....	22
6.1	Métricas Lidar .....	22
6.2	Generación de capas ráster .....	23
7	Estimación de variables dasométricas a partir del Lidar .....	23
7.1	Referencia espacial .....	23
7.2	Zonificación Lidar (estratificación) .....	23
7.2.1	Introducción .....	23
7.2.2	Las especies arbóreas en la estratificación .....	24
7.2.3	El ámbito geográfico en la estratificación .....	25
7.2.4	La fecha de vuelo en la estratificación .....	26
7.3	Procesado del IFN4 .....	28
7.3.1	Variables dasométricas .....	28
7.3.2	Ajustes Lidar .....	29
7.3.2.1	Cálculo de las variables dasométricas en las parcelas IFN4 .....	29
7.3.2.2	Cálculo de las métricas Lidar en las parcelas IFN4 .....	29
8	Regresiones Lidar .....	29
8.1	El modelo lineal .....	29
8.1.1	Planteamiento estadístico .....	29
8.1.2	Hipótesis de partida del modelo lineal .....	31
8.1.2.1	Normalidad y homocedasticidad .....	31
8.1.2.2	Colinealidad .....	31
8.1.2.3	Independencia de la muestra .....	31
8.1.2.4	Representatividad .....	31
8.1.3	Inferencia (predicción, estimación, etc.) .....	31
8.1.4	Validación del modelo .....	32
8.1.4.1	Partición de la muestra train/test .....	32
8.1.4.2	Outliers .....	32
8.1.5	Ámbito de aplicación – espacio muestral .....	32
8.1.6	Selección de variables explicativas .....	32
8.1.7	Significación de los estimadores de los parámetros .....	33
8.1.7.1	Cuantía .....	33
8.1.7.2	Significación estadística .....	33
8.1.8	Explicaciones para alumnos .....	33
8.1.8.1	Planteamientos .....	33

8.1.8.2	Regresión versus muestreo .....	34
8.1.8.3	El espacio muestral: población y muestra .....	34
8.1.8.4	Muestreo en ámbito territorial.....	35
8.1.8.5	Elección de la muestra en regresión y en muestreo.....	35
8.1.9	Hay que definir bien la variable que estimamos .....	35
8.1.10	Escala (resolución espacial) y representatividad del “caso” individual .....	35
8.1.11	Como hacemos una regresión .....	35
8.1.12	Relación lineal entre variables .....	35
8.1.12.1	Transformación de las variables explicativas .....	35
8.1.12.2	Transformación de las variables explicada .....	35
8.1.12.3	Linealización del modelo .....	36
8.2	Resultados.....	36
8.2.1	Método y calidad del ajuste .....	36
8.2.2	Verificación de los requisitos del modelo lineal .....	36
8.2.3	Intervalos de confianza.....	36
8.2.4	El talón de Aquiles del método y la cura en salud .....	37
8.2.5	Incorporación de la autocorrelación espacial a la inferencia .....	38
8.3	Correlación entre variables.....	39
8.4	Relación entre variables explicadas y explicativas.....	39
8.4.1	Por especie y provincia .....	39
8.4.2	Por estratoLidar .....	39
8.5	Ecuaciones dasoLidar .....	39
8.6	Control de calidad .....	39
9	Documentación de capas Lidar .....	39
Apéndice 1.	Gráficos de dispersión de las variables explicativas <i>versus</i> explicadas.....	39
Apéndice 2.	Intervalos de confianza en la predicción .....	40
10	Notas personales y bitácora .....	40

## 2 Proyecto dasoLidar: introducción

### 2.1 Planteamiento

Esta iniciativa tiene como objetivo proporcionar a los técnicos de la Consejería capas y herramientas que les permitan **aprovechar el potencial dasométrico y selvícola del Lidar**. Inicialmente, está implementada para el Lidar-PNOA2, con la vista puesta en el Lidar-PNOA3, que se volará en Castilla y León en 2025.

Un principio fundamental del proyecto es que los productos y herramientas deben ser lo más accesibles posible para los usuarios. La **accesibilidad** se entiende tanto en términos de disponibilidad, es decir, que la información esté fácilmente localizable, como de facilidad de uso, de modo que la interacción con esta información no requiera procedimientos excesivamente largos o complejos.

Cuando hablamos de **productos y herramientas**, nos referimos tanto a las nubes de puntos Lidar propiamente dichas (archivos .laz o lazFiles) como a capas derivadas y herramientas que utilizan esta información y/o facilitan su uso.

**QGIS** va a ser la vía principal de interacción para las aplicaciones dasométricas y selvícolas del Lidar. Esto permitirá integrar el Lidar de manera continua en la forma de trabajar de los técnicos, que, en su gran mayoría, utilizan QGIS.

### 2.2 Hoja de ruta

Esta iniciativa llega después de 10 años de colaboración con el IGN en el Lidar PNOA1 y PNOA2 participando en las distintas fases de la generación de los lazFiles y productos derivados.

El primer hito de este proyecto fue una jornada celebrada en Íscar el 6 de julio de 2023, con técnicos de las nueve provincias, en la que se analizó el potencial del Lidar en la gestión del medio natural y las necesidades que, a priori, podrían relacionarse con esta fuente de información, evaluando en qué medida el Lidar puede satisfacer dichas demandas.

El segundo hito es la celebración de un curso ECLAP (5 y 6 de junio de 2024), para el cual se han generado una serie de capas provisionales con métricas Lidar y variables dasométricas, y se ha habilitado una ubicación en la red que facilita el acceso tanto a los archivos Lidar con las nubes de puntos (.laz) como a las mencionadas capas. En el curso se presentó esta iniciativa, se explicaron las capas y algunos usos posibles, y se expusieron los fundamentos y el funcionamiento de Silvilar, un complemento de QGIS destinado a facilitar aplicaciones selvícolas del Lidar, como la priorización de intervenciones o la localización de zonas con determinadas características selvícolas.

Este curso también sirvió para iniciar una fase alfa de pruebas de los productos, su disponibilidad y su integración en la forma de trabajar de los técnicos. Se constituyó un grupo de alfa-testers formado por los 16 alumnos del curso, los técnicos Sigmena de los SSTT y otros técnicos que solicitaron incorporarse a esta fase debido a su relación con este tipo de herramientas o por tener alguna necesidad concreta que podría ser atendida con estos productos o herramientas.

En septiembre de 2024 se ha cerrado esta fase alfa con una ronda de contactos, se han depurado los productos y herramientas que se ponen a disposición de los usuarios de la fase beta y se ha diseñado la iniciativa de difusión de noviembre de 2024. La fase beta de pruebas, terminará en enero de 2024, de forma que en marzo de 2025 se inicie la fase operativa de este proyecto.

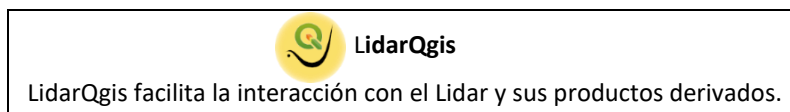


Plan de trabajo / Hoja de ruta		
2014-2024	Colaboración IGN – JCyL	Convenios PNOA, procesado, etc.
2023 (julio)	Reunión de Íscar	Necesidades ⇔ aplicaciones Lidar
2024 (Jul-dic)	Generación de productos	Proyecto lidarQgis, capas dasoLidar, silvilidar, dasoraster, documentación, etc. *
2024 (jul)	Curso Lidar 1 ⇒ Fase alfa	20 usuarios ⇒ feedback en septiembre
2024 (nov)	Curso Lidar 2 ⇒ Fase beta	+16 usuarios ⇒ feedback en enero
2024 (nov)	Seminario Lidar 1	Difusión SSTT (gestión forestal)
2025 (ene)	Fin de fase beta	Depuración de productos y herramientas
2025 (feb)	Seminario Lidar 2	Difusión SSTT
2025 (marzo-)	Fase operativa – dasoLidar	Asistencia y apoyo a usuarios
2025	Desarrollo de herramientas	Medios propios y asistencias técnicas
2025	IFN4 – unidades de gestión	MFE25, SIGPAC, montes, rodales.
2025-2026	PNOA3	Disponibilidad y eficiencia

\* i/ inventario Lidar en IOF

## 2.3 ¿Dónde está todo esto?

### 2.3.1 Puerta de entrada:



Es un proyecto Qgis que Incluye:

- Utilidades para cargar y visualizar los ficheros Lidar con las nubes de puntos (lazFiles)
- Capas derivadas del Lidar
- Herramientas para sacar partido a toda esta información
- Ayudas

Existen diversas formas de trabajar con la información Lidar y diversas aplicaciones para visualizar o generar salidas a partir de las nubes de puntos<sup>3</sup>. Sin embargo, en esta iniciativa, hemos decidido canalizar todo esto a través de Qgis. Esta aplicación no solo ofrece la posibilidad de visualizar y operar con nubes de puntos en nuestros proyectos, sino que también admite complementos (propios o de terceros) que añaden las funcionalidades que deseemos, con el único límite de nuestra imaginación, capacidad y, en algunos casos, de las restricciones propias de la seguridad digital.

### 2.3.2 Pero ¿dónde está el proyecto, las capas, las herramientas y demás?

El proyecto LidarQgis está aquí:

[\\repoarchivohm.jcyl.red\MADGMNSVPI\\_SCAYLEVueloLIDAR\\$\\qgis\LidarQgis.qgz](\\repoarchivohm.jcyl.red\MADGMNSVPI_SCAYLEVueloLIDAR$\\qgis\LidarQgis.qgz)

Esta ruta apunta a una ubicación de red en la que hemos puesto el proyecto LidarQgis y todas las capas y herramientas de las que tira.

Está en una ubicación provisional en la que vamos poniendo toda la información que ponemos a disposición de los usuarios del proyecto dasolidar. A todo eso, lo llamamos LidarData e incluye:

- Los ficheros con las nubes de puntos (lazFiles)
- Las capas ráster con información dasométrica y de diverso tipo
- El resto de la información, documentación y herramientas de esta iniciativa

Las capas ráster las hemos generado a partir del Lidar, parcelas de campo del IFN4, cartografía forestal (MFE25), etc.

### 2.3.3 ¿Por qué usamos una ruta tan rara y no una unidad de red?

La ruta de LidarData es una ubicación de red imposible de memorizar 😞:

[\\repoarchivohm.jcyl.red\MADGMNSVPI\\_SCAYLEVueloLIDAR\\$](\\repoarchivohm.jcyl.red\MADGMNSVPI_SCAYLEVueloLIDAR$)

Esta ubicación de red es de solo lectura para todos los usuarios Lidar.

Este proyecto se ha construido usando los recursos disponibles y, en lo que se refiere a recursos informáticos, tener de esta ubicación de red común para todos los servicios territoriales ha sido un paso muy importante y la mejor de las alternativas que hemos explorado. Confiamos en que en el futuro podamos hacer que todos estos recursos estén disponibles con una interacción más amigable.

### 2.3.4 El acceso directo LidarQgis

Hemos creado un acceso directo que llamamos LidarQgis (😊), que se copia en el escritorio de los usuarios Lidar (más adelante se explica quiénes son los usuarios Lidar). Esto se realiza mediante un script que se ejecuta al iniciar QGIS, por lo que los accesos directos no aparecen si no se ha iniciado esta aplicación

---

<sup>3</sup> Aplicaciones de escritorio con interfaz de usuario, utilidades en línea de comandos y aplicaciones web.

después del 1 de noviembre de 2024 (y después de haber sido dado de alta en la lista de usuarios dasoLidar).

### 2.3.5 Lista de usuarios dasoLidar

El acceso a la ubicación de red requiere la habilitación expresa por parte del Servicio de Informática, para lo cual se ha elaborado un listado de usuarios Lidar que incluye a técnicos de servicios territoriales y centrales que han solicitado su inclusión en el proyecto con motivo del seminario dasoLidar del 21 de noviembre de 2024, así como a los que han participado en las dos ediciones del curso Lidar-ECLAP de 2024.

Nos hemos centrado en potenciales usuarios relacionados con gestión forestal, restauración, promoción, defensa, informes y espacios. Este listado es un punto de partida y, dado que no es fácil de elaborar, es seguro que está incompleto.

Todos aquellos técnicos que comprueben que no tienen acceso a esta información, pero desean utilizarla pueden solicitarlo a Jose Bengoa ([benmarjo@jcyl.es](mailto:benmarjo@jcyl.es)) con el siguiente contenido:

**Asunto:** Usuario DasoLidar – [cod332]<sup>4</sup>  
**Texto del mensaje:** Nombre y apellidos  
Teléfono (incluido código de acceso directo)  
Servicio territorial y unidad/sección en la que trabaja (indicar provincia)  
Funciones principales en el puesto de trabajo  
Si tiene previsto o desea prospectar algún uso concreto de los productos y herramientas Lidar.

Este último dato es opcional y está destinado a conocer mejor las necesidades y demandas de los usuarios. Inicialmente, la orientación de los productos y herramientas es netamente dasométrica y selvícola, pero no se trata de un planteamiento restrictivo, sino de un ámbito inicial de trabajo.

### 2.3.6 Versiones alfa, beta y realease

Se ha diseñado una implantación progresiva, con objeto de poder atender las incidencias del proyecto en sus fases de pruebas (alfa y beta: julio 2024 - enero 2025). Cuando finalice la fase de pruebas se lanzarán a partir de marzo de 2025 las versiones *release* o explotación, que deben tener la funcionalidad deseada del proyecto.


El campo de los productos y aplicaciones Lidar está en pleno desarrollo y en los próximos meses y años iremos teniendo productos y funcionalidades nuevas, empezando por un nuevo vuelo PNOA-Lidar (el tercero), nuevas herramientas que incluiremos en Qgis y una interacción más sencilla con todo ello asistida por lenguaje natural.

## 3 LidarQgis

### 3.1 Presentación

El proyecto LidarQgis.qgz (antiguo LidarPNOA2.qgz) está planteado con un número reducido de capas para facilitar la toma de contacto y poner el foco en las que se usan con más frecuencia.

Para una primera aproximación vale con abrir el proyecto y activar alguna de las capas que incluye el proyecto. Las capas se pueden usar/consultar de varias formas:

- Visualmente: se han usado leyendas de colores que destacan las diferencias dasométricas entre unas y otras masas forestales.
- Valores puntuales: p. ej. se puede consultar la “altura dominante Lidar” o el “volumen estimado” en un punto concreto (un píxel o celda de la correspondiente capa ráster) con el botón .
- Se puede consultar el valor correspondiente a una parcela (circular o cuadrada, del radio que se desee) o a un rodal, lote o polígono que se desee. Eso se puede hacer con las herramientas propias

---

<sup>4</sup> El cod332 es el código personal para iniciar el ordenador en el puesto de trabajo.

de Qgis o, de forma más sencilla, con los correspondientes botones del complemento dasoraster:

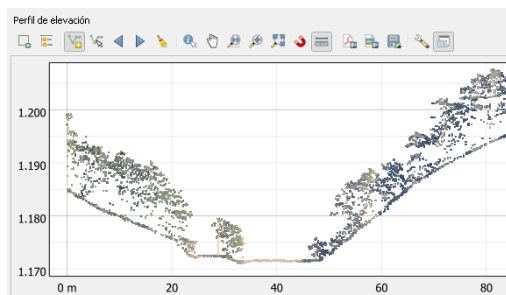


Otra cosa que se puede hacer es cargar la nube de puntos Lidar correspondiente a una zona concreta y visualizar un perfil. La carga se hace con el botón del mencionado complemento dasoraster:

- Para cargar el fichero con la nube de puntos lidar en el proyecto. hacer click sobre un punto del mapa (después de activar el puntero de carga con ). Eso crea un grupo llamado *lidarDescargado* y coloca el fichero laz descargado dentro.
- Para mostrar un perfil: Menú Ver -> Perfil de elevación.



Trazar una línea con el ratón ( ) y terminar la línea pulsando el botón derecho. En el perfil se pueden hacer mediciones, cambiar en ancho del perfil (tolerancia), trasladarlo, etc.



Como se ha indicado, el proyecto incluye pocas capas, no solo para facilitar la toma de contacto, sino también porque puede ser copiado (preferentemente en O:\Sigmena\usuarios\...) y adaptado a las necesidades de cada uno. Por supuesto, también se puede trabajar directamente con este proyecto o copiar sus capas y pegarlas en los propios proyectos base de cada uno.

## 3.2 Las capas del proyecto LidarQgis

Al abrir el proyecto se muestran los cuadrantes del vuelo Lidar PNOA2 en Castilla y León (ver apartado 5.1. El Plan Nacional de Ortofotografía Aérea (PNOA)). La tabla de contenidos incluye estos tres elementos:

- Un grupo de capas dasoLidar (aquí está la información dasométrica en capas ráster)
- Un grupo con unas pocas capas de referencia (malla Lidar con celdas de 2x2 km y cuadrantes de vuelo del Lidar-PNOA2 , división administrativa, MFE25, montes de UP y poco más)
- Capas wms con ortofotos y mapa IGN 25k.

### 3.2.1 Malla 2k

Aviso: este apartado se va a eliminar porque a partir del 21/11/2024 se usará el complemento dasoraster para este cometido

Es una capa vectorial denominada “cargar\_nubeDePuntos\_LidarPNOA2” consistente en una malla de 2 km de lado, con hiperenlaces que apuntan a los ficheros Lidar correspondientes (lazFiles).<sup>5</sup>

Los ficheros Lidar del PNOA2 están organizados en bloques de 2x2 km y, en este proyecto, hemos mantenido esta organización de los ficheros.

Para cargar la nube de puntos hay que:

1. Ubicarse en la localización deseada,
2. Activar la capa “cargar\_nubeDePuntos\_LidarPNOA2”
3. Hacer click en el botón (habrá pasado de gris a color al seleccionar la capa)
4. Pinchar en el mapa con esa herramienta

La capa cargar\_nubeDePuntos\_LidarPNOA2 tiene los hipervínculos que apuntan a la ubicación en la que están los ficheros Lidar (más adelante se explica donde están estos ficheros y algunos detalles sobre ellos).

Esta forma de funcionar se parece a como gábamos antiguamente las ortofotos. Supuso un avance respecto a la carga manual que se hacía previamente, pero hoy en día ya no trabajamos así con las

<sup>5</sup> Además, hay otra capa “cargar\_nubeDePuntos\_LidarPNOA2\_SWext” que está destinada a la carga de ficheros del SW que perteneciendo al huso 30 están en huso 29 extendido. Es algo transitorio, la idea es que en la versión operativa (enero de 2025) esté solucionado este inconveniente.

### 3.2.2 Trabajar con las nubes de puntos Lidar en Qgis: visualización

Ya se ha explicado más arriba cómo cargar una nube de puntos en un proyecto usando el botón 🗲. No es objeto de este manual explicar las utilidades de QGIS, pero sí queremos mostrar los primeros pasos que, en algunos casos, son suficientes y, en todo caso, muestran el potencial de trabajar con las nubes de puntos dentro de QGIS.

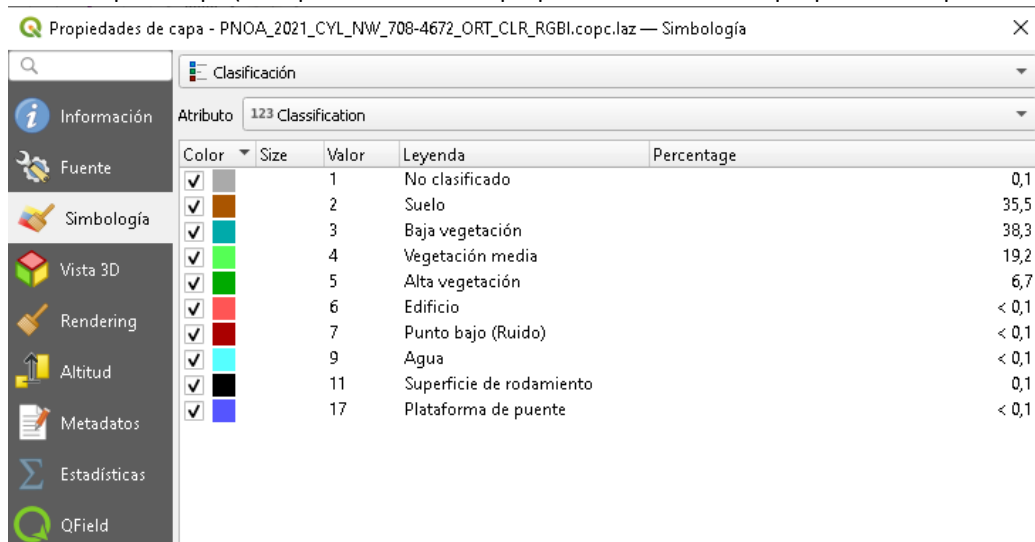
Una vez cargado un fichero Lidar, este aparece en el panel izquierdo (capas) dentro de un grupo llamado "lidarDescargado". Si nos acercamos mucho a esa capa, podremos ver que no se trata de una capa ráster con píxeles, sino de puntos individuales que, al estar tan juntos, parecen un ráster.

Con esta capa podemos hacer dos cosas que dan mucho de sí:

#### 1. Cambiar la **visualización**:

Propiedades de la capa ⇒ Simbología ⇒ Elegir en el desplegable de arriba:

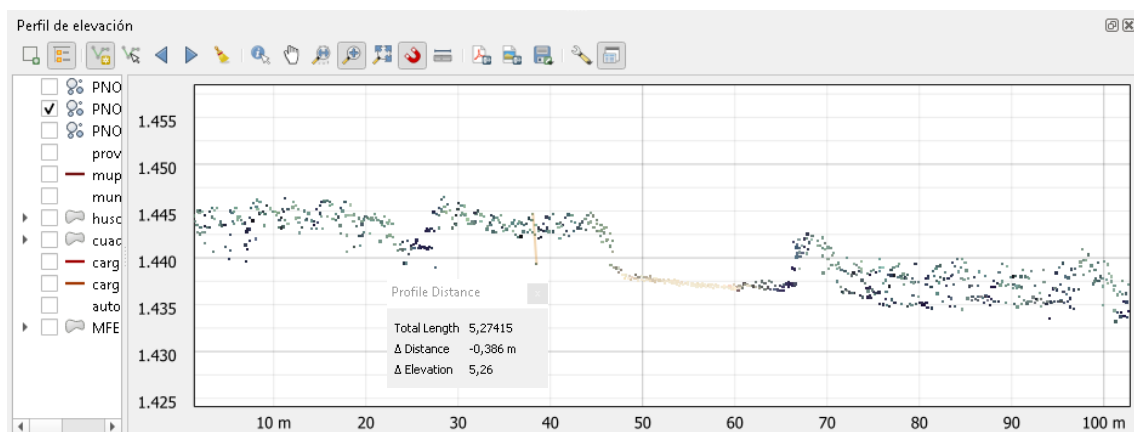
- RGB (es lo que aparece por defecto si los puntos tienen color asignado)
- Clasificación (categorías ASPRS; ver abajo)
- Atributo por rampa (cada punto tiene varias propiedades o atributos que podemos representar).




Ya que tenemos abierto el panel de Propiedades de la capa, podemos consultar otras pestañas como "Información" y "Estadísticas" que ayudan a conocer el contenido de estos ficheros.

#### 2. Mostrar un **perfil**:

Menú Ver ⇒ Perfil de elevación ⇒ Trazar una línea con el botón 📏



En el panel "Perfil de elevación" podemos usar el mencionado botón para trazar la línea en la que queremos mostrar el perfil, o usar una línea que ya tengamos en shape, por ejemplo, con la traza de una

posible pista forestal; podemos desplazar el perfil dentro del mapa, podemos medir distancias/alturas, podemos cambiar la anchura del perfil (  ), etc.

Trasteando la visualización y el perfil se saca mucho partido al Lidar.

---

### ¿Qué es la ASPRS?

---

La Sociedad Americana de Fotogrametría y Teledetección (ASPRS) es una organización que trabaja en el campo de la fotogrametría, la teledetección y los sistemas de información geográfica. ASPRS propuso en su momento el formato “las”, que ha sido adoptado por la mayoría de las aplicaciones. Este formato se ha mejorado con su versión comprimida (lastools) e indexada (pdal), sin perder el cumplimiento de las especificaciones originales de ASPRS.

Existen varias versiones del formato “las”, y cada una incluye en sus especificaciones las propiedades que se pueden asignar a cada punto y las categorías en la clasificación de los puntos. En Castilla y León, trabajamos con el formato “las” 1.4, usando el formato de punto 8, que permite incluir 4 canales o bandas de color (RGBA). La clasificación de los puntos tiene algunas clases comunes a todos los formatos, como suelo, vegetación (baja, media, alta), edificio, ruido, etc. ASPRS asignó y definió orientativamente las principales clases y dejó margen para crear otras clases (códigos 64 en adelante).

Clase	Significado
0	Nunca clasificado
1	No asignado
2	Terreno
3	Vegetación baja
4	Vegetación media
5	Vegetación alta
6	Edificio

Clase	Significado
7	Punto bajo
8	Reservado
9	Agua
10	Ferrocarril
11	Carretera o camino
12	Reservado
13	Protector-cable (señal)

Clase	Significado
14	Conductor-cable (fase)
15	Torre de transmisión
16	Conector-cables (aisla)
17	Plataforma del puente
18	Ruido alto
19-63	Reservado
64-255	Definido por el usuario

El IGN intentará consensuar una definición más precisa de las clases y asignar algunas adicionales, por lo que próximamente tendremos una versión más completa de esta clasificación.

La mayor parte de las cosas que buscamos al cargar un fichero Lidar se resuelven con la visualización y los perfiles. En Qgis se pueden hacer más cosas con esta información porque ha ido incluyendo y mejorando de forma progresiva diversas herramientas para trabajar con las nubes de puntos y seguirá haciéndolo en las próximas versiones. No es objeto de este manual explorar estas cuestiones. Los usuarios interesados pueden consultar las herramientas disponibles en la “Caja de herramientas de procesos” poniendo “Lidar” en el buscador. Ahí hay herramientas para recortar y unir ficheros, filtrar y adelgazar nubes de puntos, exportar a capa de puntos, calcular la densidad de puntos por celda, consultar metadatos, generar diversos ráster (entre ellas las Lastools), reproyectar, generar COPC, etc. Las Lastools integradas en Qgis permiten ampliar su funcionalidad con esta batería de herramientas, empezando por lasinfo y todas las de conversión de formatos y de preprocesado y terminando por las que trabajan con la clasificación de los puntos y las que permiten generar modelos digitales del terreno (MDT) o de superficies (MDS).

### 3.2.3 Capas ráster con información dasométrica

Las capas han cambiado de nombre, actualizar este apartado

El grupo dasoLidar incluye los ráster con métricas Lidar y variables dasométricas.

Son capas ráster derivadas (no nubes de puntos)

Diferencia entre métricas y variables

Se puede trabajar con estas capas para obtener valores para unos determinados polígonos (rodales, lotes, montes, provincias, etc.): **Estadísticas de zona (Análisis ráster)**

Tiene las capas con sus estilos

Todos los detalles sobre el contenido de cada capa se pueden consultar en el apartado 4.1 Directorio dasoLidar.

Datos concretos para las provincias y para CyL

Son capas generadas dentro de este proyecto: metodología en el apartado ...

Este grupo incluye, a su vez, tres grupos de capas:

- Grupo PNOA2\_2017-2021
  - VCC\_\_\_\_IFNxPNOA2
  - Alt95\_m\_PNOA2
  - Grupo CoberturaArboladaLidar
    - Cob5m\_PRT\_PNOA2
    - Cob3m\_PRT\_PNOA2
  - CoberturaEstratosRptoHD
    - CobEstr\_MidHD\_TopHD\_TLR\_PNOA2
    - CobEstr\_200cm\_MidHD\_TLR\_PNOA2
    - CobEstr\_050cm\_200cm\_TLR\_PNOA2
  - CoberturaEstratosAbsolutos
    - CobEstrDe0025a0150cm\_TLR\_PNOA2
    - CobEstrDe0150a0250cm\_TLR\_PNOA2
    - CobEstrDe0250a0500cm\_TLR\_PNOA2
  - DasometriaEstimada\_IFNxLidar
    - IAVC\_\_\_\_IFNxPNOA2
    - Abas\_\_\_\_IFNxPNOA2
  - AltMax\_r02m
    - AltMaxSobreMdk\_CE\_cm
    - AltMaxSobreMdk\_NE\_cm
    - AltMaxSobreMdk\_NW\_cm
- Grupo PNOA1\_2010-2014
  - Alt95\_m\_PNOA1
  - Cob3m\_PRT\_PNOA1
- Grupo PNOA1\_vs\_PNOA2
  - Alt95\_cm\_PNOA1\_PNOA2
  - Cob3m\_PNOA2-PNOA2
- Grupo cartoBasica (capas vectoriales de referencia)
  - auton\_ign\_e25\_etr89, prov\_ign\_e25\_etr89, mun\_ign\_e25\_etr89
  - mup\_ex\_etr89
  - MFE25CyL — mfe25cyl
  - Grupo Cuadrantes&Husos: cuadrantes y husos\_h29h30
- Grupo wms
  - Ortofotos ITACYL correspondientes a fechas relevantes para el Lidar
  - OrtofotoIGN (por si la del ITACYL no está disponible)
  - Mapas raster del IGN

## 4 LidarData



Esta ubicación está disponible para todos los técnicos incluidos en el listado de usuarios Lidar.

Se trata de un espacio limitado que nos ha habilitado el Servicio de Informática. En cualquier caso, es suficiente para poner en marcha este proyecto y confiamos en que en un futuro cercano contemos con más recursos, lo que nos permitirá poner más material a disposición de todos. Por indicación de informática, no asignamos unidad de red a esta ubicación.

LidarData incluye información diversa, que incluye los ficheros con las nubes de puntos Lidar (lazFiles), capas dasoLidar (ráster con métricas Lidar y variables dasométricas), los proyectos de Qgis con botones y

complementos ad-hoc (como Silvilidar), mallas para facilitar el acceso a los ficheros, documentación y otros ficheros y herramientas para sacar partido de esta información.

LidarData está organizado en tres directorios:

- **dasoLidar**: capas con métricas Lidar y variables dasométricas, documentación y cocina.
- **PNOA2**: ficheros laz de los 5 cuadrantes (lazcopc) y carpeta auxiliar (.aux)
- **qgis**: incluye el proyecto principal (LidarQgis.qgz) y otros auxiliares

## 4.1 Directorio dasoLidar

06/11/2024: Las capas han cambiado de nombre y la estructura de directorios, ídem. Pendiente actualizar

### 4.1.1 Contenidos

Las capas dasoLidar son capas ráster de 10 m de píxel que cubren toda Castilla y León y representan métricas Lidar y variables dasométricas. Son archivos TIFF de una sola banda, por lo que hay uno por cada métrica o variable.

Cada métrica o variable tiene una definición que está vinculada a un determinado tamaño de píxel o celda<sup>6</sup>, por lo que esta referencia espacial, los 10 x 10 m, forma parte de su definición.<sup>7</sup>

dasoLidar está organizado en cinco directorios:

- doc
- PNOA1\_2010-2014
- PNOA2\_2017-2021
- PNOA2\_vs\_PNOA1
- varios

Como puede verse, los productos derivados del PNOA2, los del PNOA1 y los obtenidos comparando métricas de ambos se han separado en directorios distintos. Además, se ha incluido un directorio de documentación (doc) y otro (varios) que incluirá diversa información auxiliar.

La comparativa PNOA2\_vs\_PNOA1 no debe interpretarse necesariamente como un cambio en la cubierta vegetal, ya es una primera aproximación consistente en restar los valores PNOA2 – PNOA1. Las diferencias se deberán en parte a cambios en la cubierta vegetal y en parte a diferencias en las características de los vuelos (sensor, configuración del sensor, altura de vuelo y, sobre todo, estado fenológico de la vegetación). El sensor, su configuración, y la altura y velocidad del vuelo condicionan la densidad, el tamaño de la huella, la penetrabilidad, la sensibilidad para identificar retornos, entre otros factores. Por lo tanto, las capas del directorio **PNOA2\_vs\_PNOA1** son solo una pista sobre lo que puede haber pasado entre las fechas de uno y otro vuelo, hasta tanto elaboremos un trabajo específico de comparación.

### 4.1.2 Métricas Lidar (ML) y variables dasométricas (VD)

#### 4.1.2.1 Aclaraciones previas

Cuando hablamos de capas con métricas Lidar, nos referimos a capas ráster con valores generados exclusivamente a partir de las nubes de puntos Lidar. Por otro lado, las capas con variables dasométricas recogen variables medidas en campo o estimadas usando mediciones de campo. Estas variables pueden ser de árbol individual (diámetro normal, altura total, altura de fuste, altura hasta la primera rama viva,

---

<sup>6</sup> Usamos estos dos términos indistintamente.

<sup>7</sup> Hay una excepción que es la altMax\_r02m, que registra la altura del retorno más alto sobre el suelo en cada celda de 2x2 m. Como esta resolución da lugar a ficheros grandes, esta métrica se ha repartido en cinco ficheros, correspondientes a los cinco cuadrantes de vuelo de Castilla y León (ver más adelante información sobre los cuadrantes de vuelo). Para reducir el tamaño de los ficheros esta variable está almacenada en centímetros (int 16 bit en lugar de float 32 bit; cada fichero ronda los 4 GB). Por el momento (septiembre 2024) solo está disponible para CE, NE y NW. En el futuro se generarán para los otros dos cuadrantes.



volumen de fuste, etc.) o de masa (altura dominante, densidad, área basimétrica, fracción de cabida cubierta, volumen en pie con corteza, volumen de leñas, biomasa aérea, etc.). No obstante, las variables dasométricas que se recogen en estos directorios son siempre variables de masa (estimadas utilizando parcelas de campo y métricas Lidar). Las métricas Lidar no requieren parcelas, mientras que las variables dasométricas sí.

Algunas métricas Lidar pueden ser fácilmente interpretables al asociarse directamente con variables dasométricas bien conocidas en el ámbito forestal. Tal es el caso de las dos más habitualmente usadas, que hemos denominado altura dominante Lidar y cobertura arbolada Lidar.

- **Altura dominante lidar (alt95):** percentil 95 de las alturas de los retornos sobre el suelo (primeros retornos).
- **Cobertura Lidar 3m (cob3m):** porcentaje de retornos que están a más de 3 m sobre el suelo (primeros retornos).

La primera métrica se asocia con el concepto forestal de “altura dominante”. De hecho, si el vuelo Lidar se ha realizado cuando el arbolado tenía hojas, podemos considerarla como una medida muy precisa de la denominada “altura dominante de parcela” (referida a una parcela de 10x10 m).

La segunda métrica representa el porcentaje de vegetación que está a más de 3 metros de altura, es decir, corresponde principalmente al arbolado joven y adulto. Está relacionada con la fracción de cabida cubierta (FCC), pero difiere en que los huecos dentro de las copas no se contabilizan como cobertura, mientras que, en la acepción habitualmente generosa de la FCC, sí. En consecuencia, los valores de la Cob3m son siempre inferiores a los de la FCC. No obstante, ambas están correlacionadas de forma bastante estrecha.

Para caracterizar la cobertura arbolada también puede usarse como referencia los 5 m (en coherencia con algunas definiciones internacionales), pero eso deja fuera la cubierta de especies arbóreas que no han alcanzado esa talla y no por ello deja de ser cubierta arbolada (las definiciones de arbolado hablan siempre de cobertura de especies que pueden alcanzar los 5 metros de altura in situ).

#### 4.1.2.2 Listado de de métricas Lidar y variables dasométricas

Se han generado métricas lidar para ambos vuelos PNOA, pero sólo se han estimado variables dasométricas para el PNOA2 utilizando los datos de las parcelas del IFN4 y la cartografía MFE25.

- PNOA2\_2017-2021
  - **metricasLidar** (mss significa metros sobre el suelo)
    - **Alt95\_m\_PNOA2.tif:** Percentil 95 de altura sobre el suelo (primeros retornos)
    - **Cob3m\_PRT\_PNOA2.tif:** % de primeros retornos que están a 3 o más mss
    - **Cob5m\_PRT\_PNOA2.tif:** % de primeros retornos que están a 5 o más mss
    - **CobEstrDe0025a0150cm\_TLR\_PNOA2.tif:** % de retornos que están entre 0,25 y 1,5 mss
    - **CobEstrDe0150a0250cm\_TLR\_PNOA2.tif:** % de retornos que están entre 1,5 y 2,5 mss
    - **CobEstrDe0250a0500cm\_TLR\_PNOA2.tif:** % de retornos que están entre 2,5 y 5,0 mss
    - **CobEstr\_050cm\_200cm\_TLR\_PNOA2.tif:** % de retornos entre 0,5 y 2,0 mss
    - **CobEstr\_200cm\_MidHD\_TLR\_PNOA2.tif:** % de retornos entre 0,5 mss y la mitad de la Alt95
    - **CobEstr\_MidHD\_TopHD\_TLR\_PNOA2.tif:** % de retornos que están encima de la mitad de la Alt95
    - **altMax\_r02m:**
      - AltMaxSobreMdk\_NW\_cm.tif: Altura en cm del retorno más alto en cada celda de 2x2 m.
      - AltMaxSobreMdk\_NE\_cm.tif: Altura en cm del retorno más alto en cada celda de 2x2 m.
      - AltMaxSobreMdk\_CE\_cm.tif: Altura en cm del retorno más alto en cada celda de 2x2 m.
    - **OtrasMetricas**
      - Alt100\_cm\_PNOA2.tif: Altura máxima sobre el suelo
      - Alt80\_cm\_PNOA2.tif: Percentil 80 de altura sobre el suelo (primeros retornos)
      - Alt65\_cm\_PNOA2.tif: Percentil 65 de altura sobre el suelo (primeros retornos)
      - Alt35\_cm\_PNOA2.tif: Percentil 35 de altura sobre el suelo (primeros retornos)
      - Alt20\_cm\_PNOA2.tif: Percentil 20 de altura sobre el suelo (primeros retornos)
      - CobEstrDe0025a0150cm\_PRT\_PNOA2.tif: % de primeros retornos entre 0,25 y 1,5 mss
      - CobEstrDe0150a0250cm\_PRT\_PNOA2.tif: % de primeros retornos entre 1,5 y 2,5 mss
      - CobEstrDe0250a0500cm\_PRT\_PNOA2.tif: % de primeros retornos entre 2,5 y 5,0 mss
      - CobEstrDe0500a9999cm\_PRT\_PNOA2.tif: % de primeros retornos por encima de 5,0 mss

- Cob0m\_PRT\_PNOA2.tif: % de primeros retornos rpto al total de retornos (cf)  
 CobEstrDe0500a9999cm\_TLR\_PNOA2.tif: % de retornos por encima de 5 mss
- variablesDasometricas
    - **VCC\_\_\_IFNxPNOA2.tif:** Volumen en pie con corteza en m3/ha
    - **IAVC\_\_\_IFNxPNOA2.tif:** Incremento anual en volumen en pie con corteza en m3/ha.año
    - **Abas\_\_\_IFNxPNOA2.tif:** Área basimétrica en m<sup>2</sup>/ha

Nota: cada capa tiene su correspondiente archivo de estilos (.qml). Recomendamos usar esta representación de las variables, que mantendremos en las diferentes capas y versiones.
  - estimacionesIFNxLidar
 

Es un duplicado temporal de variablesDasometricas, que se eliminará cuando todos los usuarios hayan pasado del proyecto LidarPNOA2.qgz de la fase alfa al LidarQgis.qgz de la fase beta.
  - PNOA1\_2010-2014
    - **Alt95\_m\_PNOA1.tif** Percentil 95 de altura sobre el suelo (primeros retornos)
    - **Cob3m\_PRT\_PNOA1.tif** % de primeros retornos que están a 3 o más mss
  - PNOA2\_vs\_PNOA1
    - **Alt95\_cm\_PNOA1\_PNOA2.tif** Resta de Alt95 del PNOA2 menos PNOA1 (en cm)
    - **Cob3m\_PNOA2-PNOA2.tif** Resta de Cob3m del PNOA2 menos PNOA1
- \* Aclaraciones:
- mss: metros sobre el suelo
- Los % de primeros retornos correspondientes a determinados intervalos de alturas se refieren al total de primeros retornos (no es el caso de Cob0m\_PRT\_PNOA2.tif).
- Los % de retornos correspondientes a determinados intervalos de alturas se refieren al total de retornos (todos los retornos, no solo primeros).
- (cf) indica que tengo pendiente confrontar si es correcto.

## 4.2 Directorio PNOA2

### 4.2.1 Cuadrantes de vuelo

En este directorio se encuentran los archivos Lidar propiamente dichos, con las nubes de puntos. Están organizados en cuadrantes de vuelo y, dentro de cada cuadrante, en bloques de 2x2 km. Cada bloque se identifica por las coordenadas UTM de su esquina superior izquierda (en miles de km).

Estos ficheros se han obtenido mediante vuelos realizados entre 2017 y 2021 (Lidar-PNOA2), dentro del Plan Nacional de Ortofotografía Aérea (<https://pnoa.ign.es/>). En Castilla y León participan CNIG, JCYL e ITACYL:

CNIG -> <https://www.cnig.es> -> Centro Nacional de Información Geográfica

JCYL -> <https://www.jcyl.es> -> Junta de Castilla y León

ITACYL -> <https://www.itacyl.es> -> Instituto Tecnológico Agrario de Castilla y León

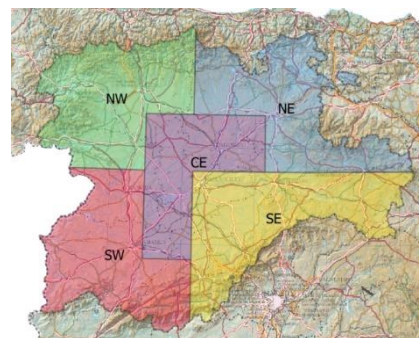
Los vuelos se han licitado en cinco "cuadrantes". Las empresas adjudicatarias de los vuelos han sido:

SE, CE y NW -> SERVICIOS POLITECNICOS AÉREOS, S.A. (SPASA)

NE y SW -> UTE TOPCAD INGENIERIA - PRIMUL MERIDIAN

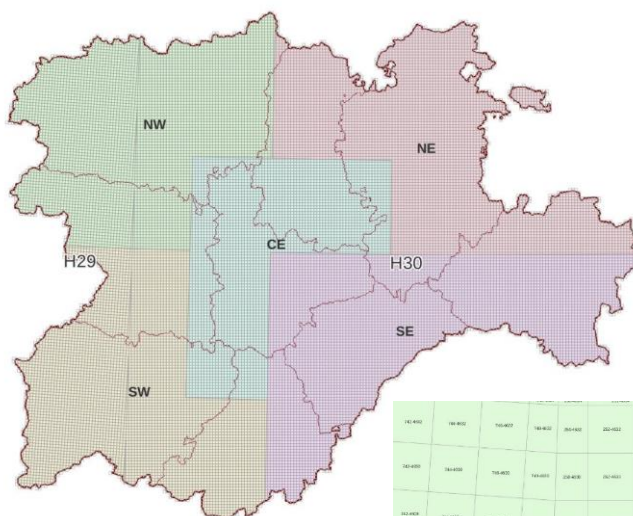
Ver las fechas de vuelo en el apartado 4.3 Fechas de vuelo.

El detalle de los cuadrantes puede consultarse en el proyecto lidarQgis.qgz (antiguo LidarPNOA2.qgz). Los lazFiles están organizados por cuadrantes de vuelo, de forma que cada directorio incluye los lazFiles correspondientes a un vuelo. Esos ficheros cubren una superficie que excede ligeramente la delimitación de cada cuadrante, por lo que hay cierto solape, necesario para garantizarla cobertura total licitada. En los bordes de las zonas reales de vuelo los lazFiles son parcialmente incompletos (no hay puntos en todo el bloque de 2x2 km), pero esto no es un inconveniente porque eso ocurre siempre fuera del ámbito estricto del cuadrante (el licitado).

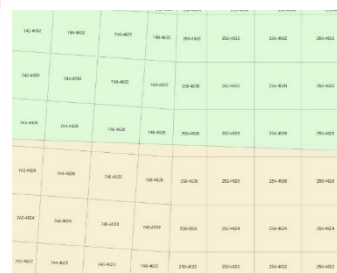


En consecuencia, en las franjas cercanas a los límites de los cuadrantes de vuelo hay bloques 2x2 km en dos versiones, voladas en fechas distintas y posiblemente por empresas distintas y con sensores distintos.

Los cuadrantes NW y SW incluyen zonas en huso 29 (además del 30 que es el del resto de Castilla y León). En esas zonas del oeste de León, Zamora y Salamanca, las coordenadas correspondientes a la proyección en el huso 29 es diferente a las que se obtienen proyectando en el 30 extendido. Los lazFiles que están en el ámbito del huso 29 están proyectados en dicho uso. No obstante, cuando se cargan en Qgis, en un proyecto creado en el huso 30, se reprojecan al vuelo y aparecen reubicadas en las coordenadas correspondientes al huso 30. Se ha incluido información del sistema de proyección en los lazFiles para que todo esto ocurra de forma automática al cargarlos en Qgis, sin intervención del usuario.



Sin embargo, todas las capas dasoLidar generadas están proyectadas en huso 30 para su visualización en una sola capa, como esta adjunta de las mallas 2x2 km, que incluye tanto la zona del huso 30, como la del 29 proyectada en el 30 extendido.



El nombre de los ficheros lidar (lazFiles) incluye las coordenadas de su esquina superior izquierda (en miles de km); dichas coordenadas están en el huso original del bloque (29 o 30). En todo caso, se puede deducir el huso de un bloque a la vista de las coordenadas ya que la coordenada X de los bloques 2x2 de Castilla y León correspondientes al huso 30 son siempre inferiores a 610.000 (códigos de bloque 240-YYYY a 600-YYYY) y las que están en el huso 29 son siempre superiores a 650.000 (códigos de bloque 656-YYYY a 754-YYYY).

El número de cuadrados de 2x2 km que cubren todo Castilla y León es de 24.481 bloques. Debido a los solapes entre las zonas voladas, el número total de bloques 2x2 km es de 26.589.

En la siguiente tabla se indican los sensores utilizados, así como la precisión y las densidades de pulsos licitadas:

Cuadrante	Inicio vuelo	Fin vuelo	Pulsos/m2	RMSE xy	RMSE z	Sensor principal
CYL SE	nov-17	dic-18	1	0,2	0,15	LEICA ALS80
CyL-NW	jun-21	oct-21	2	0,3	0,15	OPTECH GALAXY T2000
CyL SW	sep-19	jul-20	4	0,3	0,15	RIEGL LMS-Q1560
CyL NE	oct-19	oct-21	4	0,3	0,15	RIEGL LMS-Q1560
CyL C	sep-19	oct-19	1	0,3	0,15	LEICA ALS80

Se pueden consultar más detalles, incluidas las especificaciones de cada licitación en: <https://pnoa.ign.es/web/portal/pnoa-lidar/segunda-cobertura>. Aviso: si se consulta esta página desde un PC del trabajo es posible que no proporcione información actualizada debido a que el navegador de internet muestra una versión cacheada de esta web, obtenida hace bastante tiempo, antes de que el IGN volcara toda la información de la segunda cobertura (Lidar-PNOA2). Se recomienda acceder a esta ruta desde el móvil o desde un PC fuera de la red de la Junta.

### 4.3 Fechas de vuelo

Los vuelos Lidar se planifican normalmente mediante un barrido en pasadas este-oeste con cierto solape entre ellas. Los puntos de cada pasada se registran en un intervalo pequeño de tiempo, por ejemplo, entre 10 y 25 minutos (si la pasada de vuelo es de 50 o 100 km y el avión vuela entre 250 y 300 km/h).

Al componer los bloques de 2x2 km se incluyen todos los puntos volados en ese cuadrado, que normalmente incluye más de una pasada. Las pasadas pueden ser del mismo día, de días distintos o incluso de años distintos, con lo que no se puede asignar una única fecha a cada bloque 2x2, sino un intervalo de fechas que va del primer retorno al último.

El formato ASPRS incluye en su cabecera un campo para registrar el año de generación del fichero “.las” pero no un campo para almacenar el rango de fechas de vuelo. Para cada punto, la propiedad “time” si contiene esta información, pero no está sintetizada en la cabecera.

Para solventar este inconveniente se ha utilizado el campo System Identifier (sysid) para anotar esta información. Las especificaciones ASPRS crearon este campo para recoger el dispositivo de captura (modelo de sensor Lidar, normalmente un Leica, Optech o Riegl), pensando en ficheros “.las” de salida de los dispositivos de captura. Pero como esos ficheros se postprocesan antes de ofrecerlos al público, la ASPRS optó por recoger también en este campo el tipo de procesado (“MERGE”, “MODIFICATION”, “EXTRACTION”, “TRANSFORMATION” y “OTHER”).

La realidad es que los ficheros son objeto de varios procesos y no tiene mucho sentido mencionar únicamente el último de ellos, razón por la cual, se ha obviado esta especificación ASPRS y se ha utilizado este campo para recoger el intervalo de fechas de vuelo con el siguiente formato:

FechasDeVuelo\_AAAAMMDD\_AAAAMMDD

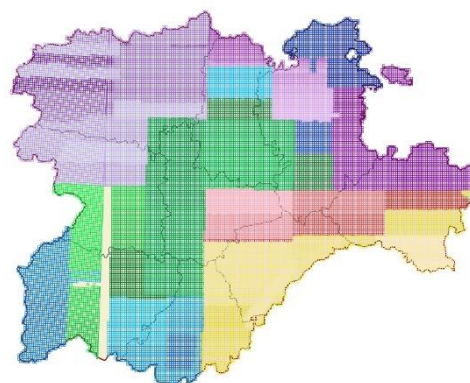
El primer AAAMMDD se refiere a año, mes día del primer retorno y el segundo al último.

La distribución de fechas de vuelo es diversa (los distintos colores reflejan distintos intervalos de fechas y la intensidad de color el día concreto; para ver datos concretos, consultar el proyecto de trabajo LidarQgis.qgz (antiguo LidarPNOA2plus.qgz):

Para tener una información más detallada de las fechas de vuelo se podría hacer un histograma de fechas de cada fichero 2x2 km. Esto permitiría un mejor criterio de uso de esta información. Sin embargo, por el momento no se ha realizado. Si se desea incluir esta información en el formato “.las”, sería necesario utilizar un VLR

A modo de resumen, estas son las fechas de vuelo de los distintos cuadrantes:

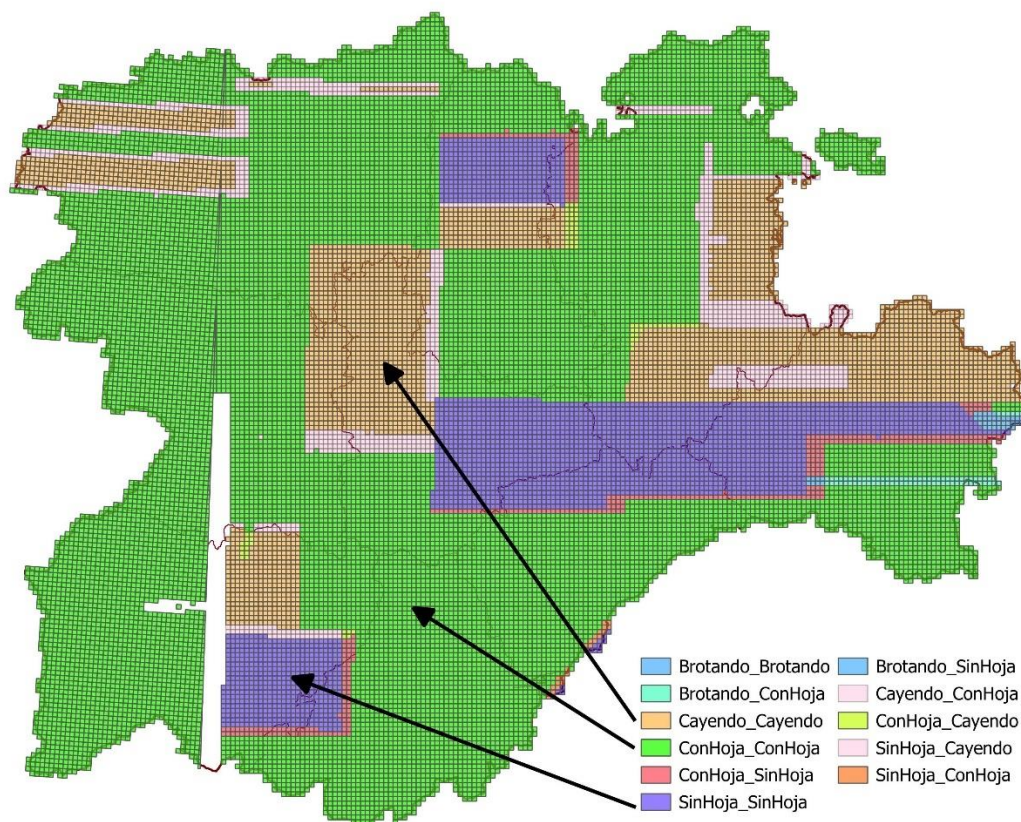
- Cuadrante CE: volado todo en 2019 (3810 bloques)
- Cuadrante NE: volado principalmente en 2020 y 2021 (595 bloques terminados en 2019, 1540 bloques en 2020 y 3665 bloques en 2021)
- Cuadrante NW: volado todo en 2021 (5715 bloques)
- Cuadrante SE: volado principalmente en 2017 y 2018 (2105 bloques terminados en 2017, 3418 bloques en 2018 y 123 bloques en 2019; supongo que en 2019 los que solapan con el cuadrante CE)
- Cuadrante SW: volado en 2019 y 2020 (1993 bloques terminados en 2019 y 3617 terminados en 2020)



Numero de bloques terminados de volar por cuadrante y año						
Cuadrante	2017	2018	2019	2020	2021	Total
CE			3.810			3.810
NE			595	1.540	3.665	5.800
NW					5.715	5.715
SE	2.105	3.418	123			5.646
SW			1.993	3.617		5.610
Total	2.105	3.418	6.521	5.157	9.380	26.581

Basándonos en las fechas inicial y final de cada bloque podemos deducir si el vuelo se ha hecho con hoja o sin hoja, con el siguiente resultado:





### La propiedad time de los puntos lidar

La propiedad time es el numero de segundos desde las 0 horas del 6 de enero de 1980 (GPS time) menos 1E9. Para obtener la fecha correspondiente al valor time de un punto concreto, se puede convertir a timestamp (referida a 1 de enero de 1970) sumandole 1315964800 segundos y usar un conversor on-line (o con código) de timesatmp a hora UTC u hora local (p. ej. <https://timestamp.online/>). En realidad, a día de hoy, el resultado tiene un desfase de 18 segundos respecto a la UTC (por los segundos intercalares con que se ha ido corrigiendo la UTC). También se puede convertir a GPS time (sumando 1E9) y usar un conversor de GPS time a UTC on-line (p. ej. <https://gwosc.org/gps/>) y, si el conversor tiene en cuenta los segundos intercalares, (como es el caso) el resultado es UTC exacto. Para obtener la hora local de Europa central (y Madrid) hay que sumar una o dos horas, según si la fecha cae dentro del horario de invierno o verano.

#### 4.3.1 Los ficheros Lidar (lazFiles)

Los archivos Lidar están en formato laz (lazFiles), que es la versión comprimida de los lasFiles (ASPRS: versión las-1.4, formato de punto 8). Están indexados en formato copc con pdal para una carga rápida en QGIS.

El nombre de los ficheros incluye esta información:

- AAAA (año de vuelo correspondiente al último punto volado en el bloque): esta información no se ha incluido en el nombre en el cuadrante SW (la versión de los lazFiles del cuadrante SW es provisional y el año que figura en el nombre es, provisionalmente, 2023, hasta tanto generemos un dataset completo definitivo). Ver más detalles en el apartado 4.3 Fechas de vuelo.
- VV (valores CE/NE/NW/SE/SW): cuadrante de vuelo
- XXX, YYYY: coordenadas UTM de la esquina superior izquierda del bloque 2x2 km.
- ORT: alturas ortométricas
- CLR: clasificación de puntos revisada

- RGBI: coloreado RGBI (incluye 4 bandas: rojo, verde, azul e infrarrojo)

Directorios que contienen los cuadrantes y el formato de nombre de fichero<sup>8</sup>:

- ce -> PNOA\_AAAA\_CYL\_CE\_XXX-YYY\_ORT\_CLR\_RGBI.copc.laz (3810 files, 140 GB)
- ne -> PNOA\_AAAA\_CYL\_NE\_XXX-YYY\_ORT\_CLR\_RGBI.copc.laz (5800 files, 1,64 TB)
- nw -> PNOA\_AAAA\_CYL\_NW\_XXX-YYY\_ORT\_CLR\_RGBI.copc.laz (5715 files 741 GB)
- se -> PNOA\_AAAA\_CYL\_SE\_XXX-YYY\_ORT\_CLR\_RGBI.copc.laz (5646 files 277 GB)
- sw -> Pendiente

Como a fecha 1 de octubre de 2024 no disponemos de una versión definitiva del cuadrante SW, hemos incluido una versión provisional:

- sw\_semiDepurado\_20231009 -> PNOA-2023-CYL-SW-XXX-YYY-ORT-000-RGBI\_LF14PF8.copc.laz (5618 files 1,32 TB)

## 4.4 ¿Qué más hay en LidarData?

### 4.4.1 .aux

Este directorio incluye algunos ficheros auxiliares y varios directorios.

- Ficheros auxiliares:
  - lidar\_copc.gpkg: mallas 2x2 para carga de lazFiles
  - Incidencias.txt: notas sobre incidencias, sugerencias, tareas pendientes, etc.

En este directorio hay otros ficheros temporales que se eliminarán en el futuro, como la malla lidar 2x2 del sw H29 extendido, para descargar ficheros que debían estar en coordenadas H30, pero los tenemos en H29 (lidar\_copc\_sw\_H29extendido.gpkg).

- Directorios:
  - cartoBasica: cartografía de referencia, por el momento división administrativa, montes de UP y mfe25.
  - usuarios: lista de usuarios incluidos en esta iniciativa
  - accesosDirectos: accesos directos que se copian al escritorio de los usuarios incluidos en esta iniciativa.
  - icon: iconos para los botones y accesos directos

## 5 Origen de los datos brutos: PNOA, IFN, MFE

### 5.1 El Plan Nacional de Ortofotografía Aérea (PNOA)

### 5.2 El IFN4

### 5.3 El MFE25

## 6 Procesado de lazFiles y obtención de las métricas Lidar

### 6.1 Métricas Lidar

La generación de métricas lidar se ha realizado con desarrollo propio (cartolidar). A continuación se recogen los principales criterios de generación de las métricas.

---

<sup>8</sup> Como se puede comprobar, no se ha unificado el criterio de uso de guión medio y bajo en el nombre de los ficheros. Este asunto está pendiente.

Pte...

## 6.2 Generación de capas ráster

Pendiente

## 7 Estimación de variables dasométricas a partir del Lidar

La estimación de variables dasométricas en toda la superficie arbolada de Castilla y León se basa en modelos de regresión lineal en los que la variable explicada es una variable dasométrica y las explicativas son métricas Lidar. A partir de estos modelos y las capas con métricas Lidar se generan las capas con variables dasométricas estimadas.

### 7.1 Referencia espacial

Tanto las capas de métricas Lidar como las de variables dasométricas estimadas tienen un tamaño de píxel de 10 m.

Por su parte, las regresiones no se obtienen con datos obtenidos para cada píxel, sino para cada parcela del IFN4, que es la fuente de realidad-terreno. Esto tiene el inconveniente de que estas parcelas tienen radio variable:

- En el círculo de 5 metros se miden todos los pies mayores (diámetro medio normal  $\geq 75$  mm)
- En el anillo de 5 - 10 metros se miden los pies de diámetro medio normal  $\geq 125$  mm
- En el anillo de 10 - 15 metros se miden los pies de diámetro medio normal  $\geq 225$  mm
- En el anillo de 15 - 25 metros se miden los pies de diámetro medio normal  $\geq 425$  mm

Como consecuencia, el círculo exterior de la parcela, de radio 25 m, no cuenta con información dasométrica completa. Se podría asumir que la dasometría del círculo de 25 m es homogénea en toda su superficie, pero se considera que es una asunción excesiva para los objetivos que se plantean, con lo que se ha optado por adoptar el círculo de 15 m como referencia espacial para las variables que intervienen en la regresión.<sup>9</sup>

Como las métricas lidar están generadas en una malla de 10x10 m, es necesario integrar varias celdas de esta malla para obtener una métrica a nivel de parcela. En concreto se seleccionan las celdas cuyo centro está dentro del círculo de 15 m de la parcela IFN y se genera una métrica igual a la correspondiente media aritmética.

La generación de capas con variables dasométricas estimadas por regresión se lleva a cabo aplicando las regresiones a cada píxel de 10 m. Esta dualidad en la referencia espacial (círculos de 15 m para el ajuste y celdas de 10x10 m para la inferencia) no conlleva sesgos en los valores inferidos, pero debe tenerse en cuenta en todo lo relacionado con la fiabilidad y los intervalos de confianza. Es esta cuestión también interviene de forma muy determinante la autocorrelación espacial, cosa que se analiza en el apartado 8.2.5 Incorporación de la autocorrelación espacial a la inferencia.

### 7.2 Zonificación Lidar (estratificación)

#### 7.2.1 Introducción

Tanto para el ajuste como para la inferencia se han definido espacios muestrales, de forma que las regresiones se han construido y aplicado por separado para cada uno de ellos. Con frecuencia se llama estratos a los espacios muestrales, por lo que en este texto usaremos indistintamente ambos términos.

Los estratos o espacios muestrales se han definido en función de:

---

<sup>9</sup> De hecho, se ha propuesto al Ministerio que el IFN5 incluya una medición más exhaustiva en el interior del círculo de 15 m y dejar la componente variable de las parcelas sólo para el anillo de 15-25 m.

- Las especies arbóreas
- La fecha de vuelo
- El ámbito geográfico

Como las regresiones se construyen y se aplican dentro de cada estrato, es necesario disponer de un número suficiente de parcelas dentro de cada uno y este es uno de los condicionantes en su delimitación. Por otro lado, hay más factores que influyen en las regresiones, pero no vale con tener información acerca de ellos en las parcelas con las que se construyen las regresiones, sino que también debe estar disponible en todo el territorio en el que se aplican (en nuestro caso, Castilla y León), razón por la cual no podemos usar todos los criterios de estratificación que nos gustaría.

Con mucha frecuencia la definición de estratos Lidar omite un “elefante en la habitación” y es que la historia de cada masa forestal que puede influir de forma importante en la relación que hay entre métricas Lidar y variables dasométricas. El itinerario selvícola, es decir, la secuencia de intervenciones pasadas (podas, clareos y claras) determinan la densidad y esbeltez de la masa principal por un lado y la distribución vertical de la nube de puntos Lidar por otro. Y, lo que es realmente importante, condiciona la relación entre esos dos grupos de variables. Esto invita a usar la información disponible sobre intervenciones selvícolas como factor de estratificación; sin embargo, esta cuestión no se ha incluido en el procesado actual, posponiéndola para cuando se disponga de información suficiente sistematizada al respecto.

Otro factor como el origen de la masa (repoblación o regeneración natural) también condicionan la densidad, tanto más cuanto más joven es la masa y, por lo tanto, influye en el mismo sentido que los demás aspectos del itinerario selvícola. Sin embargo, la distribución geográfica de este factor es más consistente, con lo que la segmentación territorial permite minorar sus efectos.

En definitiva, se ha intentado disponer de estratos suficientemente amplios para contar con un número suficiente de parcelas y suficientemente acotados como para que la relación entre las variables explicativas (métricas Lidar) y las explicadas (variables dasométricas) esté bien definida, sea lo más homogénea posible y no de lugar a residuos elevados.

Como resultado, los estratos más importantes cuentan con 200-500 parcelas, suficientes para estimar con precisión la relación entre las métricas Lidar y las variables dasométricas. Téngase en cuenta que se usa un número limitado de variables explicativas, por lo que no hay una merma significativa de grados de libertad por esa causa. Este número de parcelas permite, además, caracterizar bien el error residual y, con ello, el margen de error en las estimaciones.

Para algunos estratos, de menor entidad, se cuenta con un número bajo de parcelas (60-100) que, no obstante, permiten tener una idea de su fiabilidad.

### 7.2.2 Las especies arbóreas en la estratificación

Para construir las regresiones, se ha usado la información de las especies presentes en las parcelas IFN mientras que, en la aplicación, se ha usado la información del MFE25.

Eso significa que la construcción de las regresiones se ha realizado con información precisa y cierta, mientras que su aplicación a todo el territorio de Castilla y León se ha hecho con información menos precisa y, puntualmente, menos fiable. Con carácter general, el MFE25 tiene una buena identificación de las especies principales de cada tesela, pero no dispone de información de cómo se distribuyen éstas dentro de la tesela, por lo que en una determinada celda 10x10 m, puede haber una o varias de las especies presentes en la tesela. La escala 1:25.000 ha permitido mejorar la discriminación de las especies arbóreas, reduciendo en buena medida las teselas con mosaicos de especies principales respecto a anteriores ediciones. No obstante, la existencia de teselas de mosaico es un hecho y debe contarse con ello. Por su parte, las mezclas pie a pie son, por naturaleza, inseparables en la cartografía, con lo que también hay que trabajar con esta realidad.

La segmentación de las teselas de mosaico para tener por separado la superficie cubierta por cada especie principal es un objetivo asequible. Cuando se trabaja localmente, como en una ordenación, se puede hacer mediante fotointerpretación manual y cuando se trabaja en extensiones más amplias mediante teledetección.



Por su parte, las mezclas de especies, bien con coexistencia en el estrato superior o bien como dominancia de una y subpiso de otras, requeriría la definición de estratos concretos de mezclas de especies. No obstante, este paso debe darse después de la segmentación de los mosaicos y, en todo caso, de forma bastante restrictiva para no mermar en exceso los estratos. En definitiva, queda pospuesto para la siguiente edición del proyecto.

El conclusión: para cada parcela del IFN se ha asignado una especie principal, que es la que más área basimétrica ocupa y los estratos construidos con el MFE25 se han definido también en función de la especie principal.

Las especies principales seleccionadas, ordenadas por el número de parcelas IFN, son:

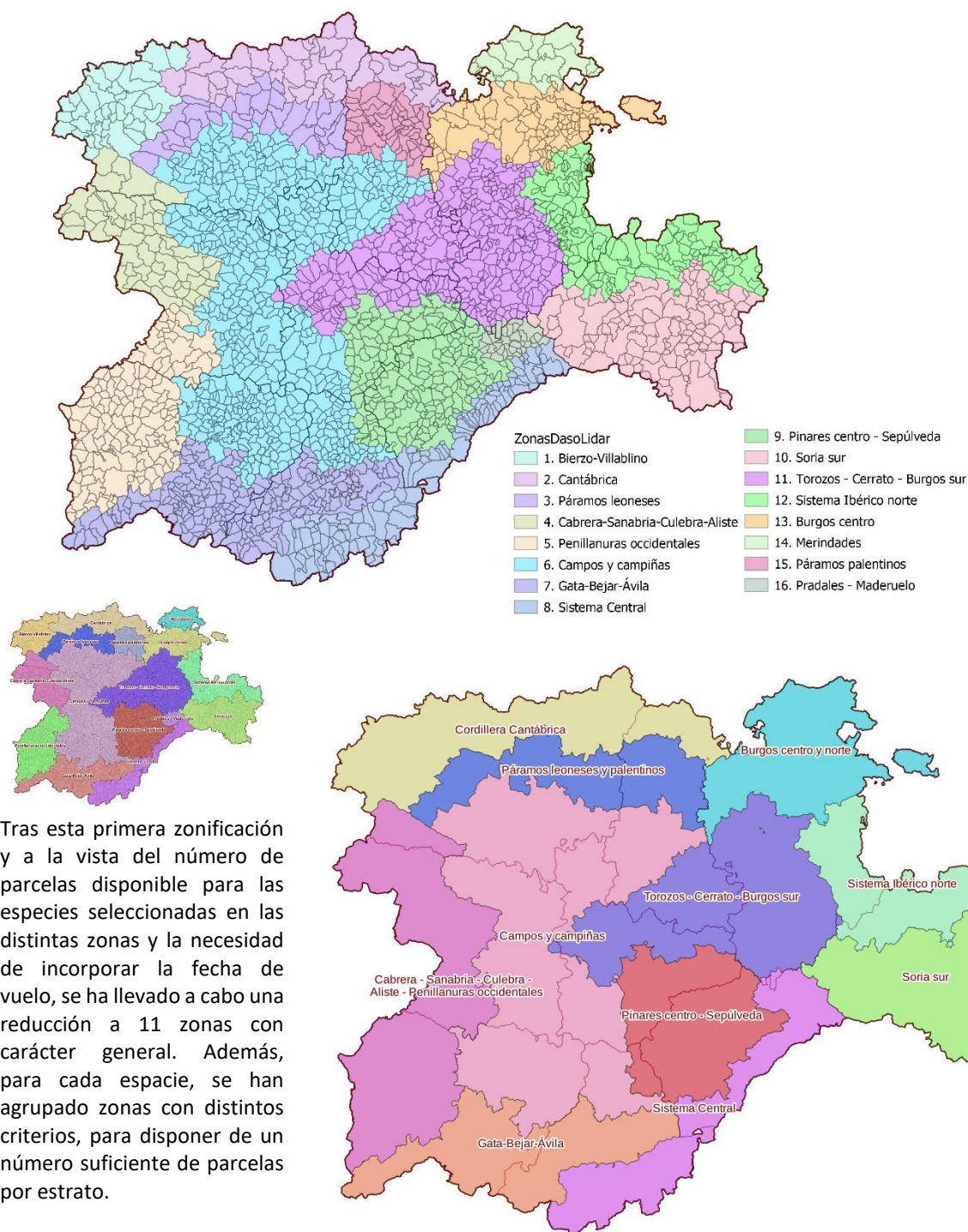
Especie principal	Parcelas	CodIFN	Cod2L	Especie principal	Parcelas	CodIFN	Cod2L
<i>Pinus pinaster</i>	1.844	26	Pt	<i>Fagus sylvatica</i>	268	71	Fs
<i>Quercus ilex</i>	1.716	45	Qi	<i>Populus x canadensis</i>	235	258	Lq
<i>Pinus sylvestris</i>	1.556	21	Ps	<i>Quercus petraea</i>	138	42	Qt
<i>Quercus pyrenaica</i>	1.519	43	Qp	<i>Populus nigra</i>	133	58	Lg
<i>Pinus pinea</i>	545	23	Pp	<i>Castanea sativa</i>	106	72	Cs
<i>Quercus faginea</i>	519	44	Qf	<i>Pinus halepensis</i>	76	24	Ph
<i>Pinus nigra</i>	472	25	Pn	<i>Pinus radiata</i>	62	28	Pr
<i>Juniperus thurifera</i>	428	38	Jt	Resto	500	-	-

Quedan fuera especies de cierta importancia en Castilla y León como *Quercus robur*, *Quercus suber*, *Betula alba*, *Pinus uncinata*, especies ribereñas, etc.

### 7.2.3 El ámbito geográfico en la estratificación

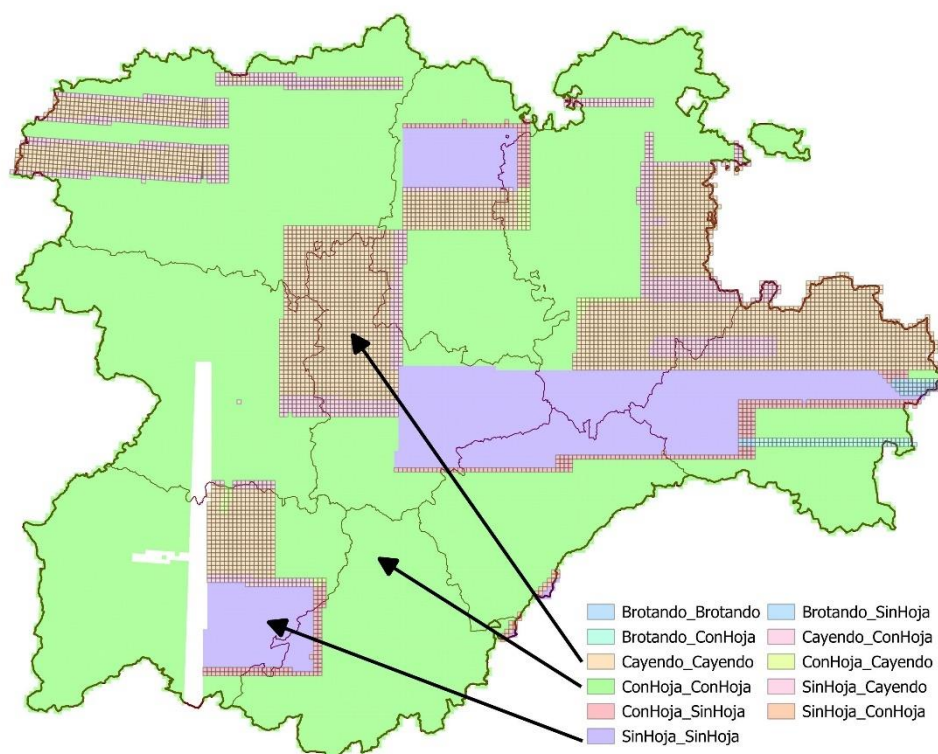
Para cada especie principal se han definido ámbitos geográficos destinados a reducir la variabilidad en la relación entre variables explicativas y explicadas. Esta variabilidad se debe tanto a factores naturales como antrópicos; de hecho, se considera que estos últimos (en concreto, el régimen selvícola) son los más determinantes de dichas relaciones, especialmente en lo que se refiere a variables como el diámetro cuadrático medio y el número de pies por hectárea. La partición del territorio en ámbitos geográficos reducidos tiene la limitación del número de parcelas disponible en cada uno de ellos. Además, no se cuenta con información suficientemente elaborada de grado de intervención por área geográfica, con lo que la delimitación de estos ámbitos geográficos tiene un carácter prospectivo. Un análisis de los resultados y de fuentes relativas a la gestión permitirá redefinir de estos ámbitos.

Se ha hecho una primera delimitación general de ámbitos geográficos, por municipios completos (para facilitar la tarea), con la mirada puesta en que sirva para todas las especies, integrando zonas, si es necesario y con el objetivo de definir 10-15 zonas. Se han tenido en cuenta aspectos orográficos o geomorfológicos, de la naturaleza del sustrato, biogeográficos (naturalidad o no de los pinares), administrativos (por posibles diferencias de “tradición selvícola”), y de cuadrante de vuelo (por usar diferentes sensores y densidades). El resultado no satisface plenamente las diferencias existentes, de acuerdo a todos estos factores, pero adopta criterios de todos ellos, obteniéndose, en primera instancia el siguiente resultado:



#### 7.2.4 La fecha de vuelo en la estratificación

El siguiente mapa sintetiza el estado fenológico del arbolado en el momento del vuelo de cada bloque 2x2 km. Se han asumido tres situaciones principales: “con hoja” (verde), “sin hoja” (lila) y “resto” (beige y otros colores, malla 2x2 visible). Pueden consultarse algunos detalles sobre este mapa en el apartado 4.3. Fechas de vuelo.



Como se puede comprobar, la mayor parte del territorio ha sido volada con hoja. Hay tres zonas voladas sin hoja (Pa, Va-So y Sa-Av) y diversas zonas voladas en época de transición fenológica.

Para tener en cuenta estas cuestiones, sin distorsionar en exceso el enfoque sintético de la zonificación y teniendo en cuenta la distribución regional de las especies caducifolias, se han hecho las siguientes consideraciones:

- Integrar los páramos palentinos y leoneses en una sola zona para poder dividirla por fecha de vuelo en dos categorías: (1) con hoja y (2) sin hoja. Esto afecta fundamentalmente a rebollo y quejigo. Para el rebollo<sup>10</sup> se ha segregado un estrato sin hoja con 39 de las 245 parcelas de rebollo en esta zona. Las 24 parcelas voladas en época de transición, aun cuando se podrían usar por corresponder a primeros de septiembre<sup>11</sup> se han excluido de los ajustes por contar con suficientes parcelas en el estrato. El estrato sin hoja sirve para comparación de modelos con el estrato con hoja.
- Para el quejigo se han descartado las parcelas voladas sin hoja en todos los estratos considerados. En uno de ellos (Soria, compuesto por las zonas de Ibérico Norte y Soria sur) hay 31 parcelas voladas sin hoja que se destinan al mismo uso que para el rebollo (en ese estrato hay 107 parcelas voladas con hoja).
- A la vista de que hay una zona volada en febrero de 2020 en Salamanca y parte de Ávila (sin hoja) se plantea dividir la zona denominada “Gata - Béjar - Ávila” en varios estratos para las especies caducifolias. De ellas, la más abundante es el rebollo, que cuenta con 274 parcelas en esta zona: 192 han sido voladas con hoja<sup>12</sup>, 30 sin hoja y 52 incluyen puntos de ambos periodos. Las 82 parcelas de rebollo total o parcialmente voladas sin hoja se han excluido de los ajustes. Las 30 parcelas voladas sin hoja se usan para comparación con los ajustes sin hoja. Las 52 con mezcla de puntos se descartan de los ajustes.

<sup>10</sup> Para el rebollo, los páramos leoneses y palentinos se han agrupado con las tres grandes zonas del interior de la cuenca del Duero, en las que el rebollo no tiene entidad suficiente para constituir un estrato.

<sup>11</sup> Se trata principalmente de la primera semana de octubre de 2019 (Palencia) y 2021 (León).

<sup>12</sup> La parte incluida en el cuadrante sur-este en septiembre de 2018 y la incluida en el sur-oeste en agosto de 2019 y julio de 2020.

- Otras especies caducifolias presentes en la mencionada zona volada sin hoja en Gata – Béjar – Ávila son el fresno y el chopo, pero la primera no se ha incluido en este trabajo y la segunda tiene una representación escasa en esta zona.
- No se utiliza la información sobre el estado fenológico en las tres grandes zonas, mayoritariamente perennifolias, que componen el centro de la cuenca del Duero ( “Campos y campiñas”, “Torozos – Cerrato - Burgos Sur” y “Pinares centro - Sepúlveda”).
- Las tres franjas que figuran “en transición” en el lado oeste y norte de la provincia de León fueron voladas el 7 de octubre de 2021. Hemos tomado el 1 de octubre como fecha de referencia para inicio de la fase de caída de la hoja, con lo que el 7 de octubre la mayor parte de las especies tienen la mayor parte de su hoja. En conclusión, no excluimos las parcelas de esas franjas y las tratamos junto con las voladas con hoja.
- La zona denominada “Sistema Ibérico norte” fue volada mayoritariamente en un intervalo pequeño de tiempo (primera quincena de octubre de 2021), con lo que compone un bloque bastante homogéneo en lo que se refiere a la fase fenológica.
- La zona denominada “Soria-sur” es la que tiene mayor disparidad de estados fenológicos, siendo las dos especies caducifolias más afectadas el quejigo y el chopo. No se trata de una zona importante para el chopo, pero sí para el quejigo por lo que se define un estrato principal de quejigo con hoja en esta zona (que para esta especie se integra con el Ibérico norte), con 107 parcelas y uno accesorio sin hoja, para comparación, con 31 parcelas.
- El chopo de producción no se ha estratificado por zonas geográficas, pero sí se han separado las 39 parcelas voladas sin hoja de las 196 voladas con hoja. Las primeras se usan para comparación, como se ha comentado para otras especies.
- Con carácter general se han descartado de los ajustes las parcelas voladas sin hoja o en transición. Aparte de los casos ya comentados, esto afecta sólo a algunas parcelas de quejigo en zonas no mencionadas previamente, al chopo del país, no incluido en este trabajo y a dos parcelas de roble albar. No hay parcelas voladas sin hoja para haya ni castaño.

## 7.3 Procesado del IFN4

Generación de capas regionales con métricas Lidar

Cálculo de variables dasométricas en las parcelas del IFN4

### 7.3.1 Variables dasométricas

Se va a trabajar con las siguientes variables dasométricas (variables explicadas en el modelo lineal del apartado 8. Regresiones Lidar):

- VCC: Volumen con corteza en m<sup>3</sup>/ha
- AB: Área basimétrica en m<sup>2</sup>/ha
- NP: Numero de pies mayores por hectárea (mayores: Dn > 7,5 cm)
- DG: Diámetro cuadrático medio en cm ( $AB = NP \cdot \pi \cdot DG^2 / 40000$ )
- IAVC: Incremento anual en volumen con corteza en m<sup>3</sup>/ha·año
- BA: Biomasa aérea del arbolado en t/ha
- VLE: Volumen de leñas gruesas en m<sup>3</sup>/ha

Hay otras dos variables dasométricas de masa fundamentales que son la altura dominante (HD) y la fracción de cabida cubierta (FCC) que tienen unas particularidades propias:

- HO: La altura dominante tiene diversas definiciones y diversas formas de medirla o estimarla. Hay cierta unanimidad en el concepto aproximativo de esta variable dasométrica: la altura de los árboles más altos de la masa, pero, cuando analicé esta cuestión (hace algún tiempo, Bengoa, 1999), llegué a la conclusión de que las definiciones canónicas adoptadas en España no eran razonables y, en la práctica no se aplicaban, por lo que proponía la adopción del concepto de altura dominante como media de una serie de alturas máximas de parcela de 10x10m. Trascurridos 25 años desde aquella propuesta, podemos aparcar la medición de campo y sustituirla por una métrica Lidar conceptual y procedimentalmente muy parecida, que es la de la media de los percentiles 95 de alturas de primeros retornos lidar con hoja de una muestra de celdas de 10x10m. Considero que esta métrica Lidar es más objetiva y precisa que la mayor parte de los procedimientos de cálculo más o menos reglados.
- FCC: La Fracción de Cabida Cubierta es una variable que se estima siempre de forma aproxima y no cuenta con una definición formal objetivable. Se refiere a la fracción del suelo cubierta por las copas de los árboles (en %) y se estima normalmente *de visu*. Lo habitual es considerar las copas de los árboles 'opacas' y descontar del 100% los huecos entre las copas, de forma que se asigna el 100% cuando hay tangencia de copas y éstas sólo dejan pequeños huecos entre ellas. En realidad, muchas parcelas a las que se asigna una FCC del 100% tienen pequeños huecos intra- e inter- copas por los que llegaría luz directa al suelo si ésta incidiera verticalmente.

Las métricas Lidar que expresan la intercepción de pulsos de luz por las copas antes de llegar al suelo se refieren a este mismo concepto de cobertura del suelo por la vegetación, pero entendida a una escala más detallada, por lo que los valores de Cob3m y Cob5m (ver apartado 6.1. Métricas Lidar) lo que da lugar a valores inferiores a la FCC. En consecuencia, para expresar la cobertura de las copas se pueden expresar tanto las métricas Lidar Cob3m y Cob5m como la variable dasométrica FCC, pero siempre teniendo en cuenta que son variables distintas. La relación entre la FCC y la Cob3m o la Cob5m puede expresarse mediante un modelo lineal, con la particularidad de que, en este caso, la variable explicada, FCC, tiene una componente subjetiva mayor (menor precisión) que las variables explicativas, que están más objetivadas (para un determinado sensor, altura de vuelo y estado fenológico de la vegetación), con lo que el fundamento del modelo lineal está, en cierta forma, invertido

### 7.3.2 Ajustes Lidar

#### 7.3.2.1 Cálculo de las variables dasométricas en las parcelas IFN4

#### 7.3.2.2 Cálculo de las métricas Lidar en las parcelas IFN4

Pte...

## 8 Regresiones Lidar

### 8.1 El modelo lineal

#### 8.1.1 Planteamiento estadístico

Algunas variables dasométricas de masa se pueden estimar a partir de métricas lidar, de acuerdo a un modelo lineal:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \epsilon_i$$



Que se aplica en forma matricial:

$$y = X\beta + \epsilon \quad \mathbf{y} = \mathbf{X}^T \beta + \epsilon$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ 1 & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_n \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix}$$

Donde:

- La variable explicada ( $y_i, y$ ) es una variable dasométrica medida en una parcela de campo que, en este caso, es una parcela del IFN4 (ver 7.3.1. Variables dasométricas)
- Las variables explicativas ( $x_{i1} \dots x_{ip}, X$ ) son métricas Lidar (originales o transformadas) obtenidas para esas mismas parcelas de campo (ver apartado 6.1. Métricas Lidar).
- Los parámetros ( $\beta_0 \dots \beta_p, \beta$ ) los que estimamos en el ajuste.
- ( $\epsilon_i, \epsilon$ ) es el error o componente aleatoria no explicada por el modelo.

Los estimadores de los parámetros  $\beta$  son los que minimizan los residuos:

$$\hat{\beta} = \arg \min_{\beta} (\mathbf{y} - \mathbf{X}^T \beta)^2$$

El estimador de la  $y$  se obtiene con esos estimadores de los parámetros:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}$$

Estimador de su varianza:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{n - p} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p}$$

Error medio residual:

$$RSE = \sqrt{\frac{1}{n - p - 1} RSS}$$

Porcentaje de varianza explicada por el modelo:

$$R^2 = \frac{\text{Suma de cuadrados totales} - \text{Suma de cuadrados residuales}}{\text{Suma de cuadrados totales}}$$

$$1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \bar{y})^2}$$

$$R^2_{ajustado} = 1 - \frac{SSE}{SST} \times \frac{n-1}{n-p-1} = R^2 - (1 - R^2) \frac{n-1}{n-p-1} = 1 - \frac{SSE/df_e}{SST/df_t}$$

F test:

$$F = \frac{(TSS - RSS)/(p-1)}{RSS/(n-p)}$$

## 8.1.2 Hipótesis de partida del modelo lineal

### 8.1.2.1 Normalidad y homocedasticidad

Normalidad en la distribución de los residuos:

- Media cero en todo el rango -> visualmente, test
- Varianza constante en todo el rango (homocedasticidad)

Tests de permutación, bootstrapping

### 8.1.2.2 Colinealidad

Importante si tenemos la tentación de incluir variables de más y si queremos valorar la significación de las variables.

### 8.1.2.3 Independencia de la muestra

Hay que evitar determinados sesgos en la muestra con la que construimos el modelo

Cada elemento de la muestra debe elegirse de forma independiente a todos los demás.

Problema en las parcelas: desplazamientos (ya que estoy aquí hago otras parcelas cerca: conglomerados de parcelas)-> cuidado en la inferencia, especialmente en los intervalos de confianza.

### 8.1.2.4 Representatividad

Fundamental en lo que se refiere a la relación entre las variables explicativas y la explicada. Y en muestreo : FUNDAMENTAL.

## 8.1.3 Inferencia (predicción, estimación, etc.)

Valor estimado para cada X:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}$$

Diferenciar entre la varianza (dispersión) en la estimación de:

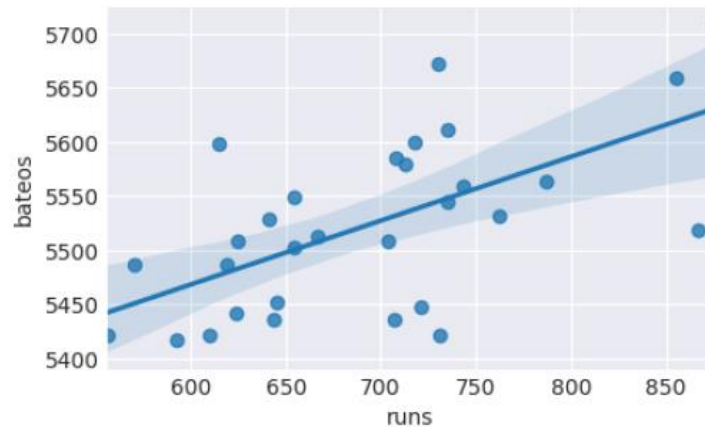
- Intervalo de confianza para las predicciones puntuales:

$$\hat{y} \pm t_{\alpha/2, n-2} \sqrt{MSE \left( 1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)}$$

- Intervalo de confianza para estimación de la media poblacional

$$\hat{y} \pm t_{\alpha/2, n-2} \sqrt{MSE \left( \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)}$$

Los intervalos se amplian conforme nos alejamos de las medias muestrales de X:



#### 8.1.4 Validación del modelo

##### 8.1.4.1 Partición de la muestra train/test

Partimos la muestra en dos grupos: uno para construir el modelo y otro para validarlo. Normalmente 80%/20%.

Puede ayudar a detectar sesgos o cuestiones que se nos han escapado porque no tenemos buena base estadística o no verificamos algunas hipótesis en las que se basa la predicción.

Como interpretamos la validación? Significación estadística. Puede ayudar a volver atrás para revisar errores o posibles malas prácticas.

##### 8.1.4.2 Outliers

Hay que analizarlos, uno por uno y ver si son errores o no, y en este caso pueden aportar información relevante para interpretar el modelo

#### 8.1.5 Ámbito de aplicación – espacio muestral

Se adopta como hipótesis que esta relación es válida para un determinado ámbito de aplicación, que queda definido por:

- La especie principal de la parcela
- Un ámbito territorial y de tipo de gestión dentro de cada especie
- El rango de valores de validez de las variables explicativas

Por lo tanto, es necesario definir los distintos ámbitos de aplicación, con el objeto de llevar a cabo, en cada uno de ellos, los correspondientes ajustes de regresión de forma independiente: cada ámbito de aplicación constituye la población en la cual se obtiene una muestra con la que se construye el modelo lineal.

#### 8.1.6 Selección de variables explicativas

Mientras que las variables explicadas son las mismas para todos los casos (ver 7.3.1. Variables dasométricas), las variables explicativas seleccionadas para el modelo lineal cambian de uno a otro.

Aunque solo sea por azar, cualquier variable que incluyamos en el modelo mejora el porcentaje de varianza explicada, pero eso no significa que sea adecuado incluirla.



Hay que evitar el vicio de incluir variables de más.

La selección de variables explicativas (entre las disponibles, que se enumeran en el apartado 6.1. Métricas Lidar) se lleva a cabo por el método forward stepwise usando el criterio de Akaike (AIC: Akaike Information Criterion)<sup>13</sup> y restringiendo el número máximo de variables explicativas a 5 (número máximo de parámetros estimados: 6).

Métodos de selección de variables:

- Selección paso a paso
  - Forward
  - Backward
  - Mixto

Criterios de selección de variables:

- Cp
- AIC
- BIC
- R2ajustado.

Yo uso El método Akaike (AIC), que tiende a ser más restrictivo e introducir menos predictores que el R2-ajustado. En todo caso pongo tope de antemano. Tengo pendiente el tratamiento adecuado de los valores nulos y noData de algunas variables de explicativa.

## 8.1.7 Significación de los estimadores de los parámetros

### 8.1.7.1 Cuantía

Que sea pequeño (0.000041) no indica que sea poco significativo xq depende de la magnitud empleada para las variables explicativas

### 8.1.7.2 Significación estadística

Indica si es significativamente distinto de cero, es decir, si la correspondiente variable aporta algo (influye, o explica parte de la variación de y)

-> t de Student -> significación (p value) -> Intervalos de confianza para los parámetros

$$\hat{\beta}_j \pm t_{df}^{\alpha/2} SE(\hat{\beta}_1)$$

## 8.1.8 Explicaciones para alumnos

### 8.1.8.1 Planteamientos

Objetivo de la regresión:

- Ver cómo influyen un conjunto de variables explicativas (predictoras, regresoras, independientes, de entrada, features, X) en una variable explicada (respuesta, predicha, dependiente, de salida, y): si las variaciones de la variable respuesta se pueden explicar o relacionar con las explicativas. Calcular en qué medida afecta cada x a la y si esa afección, relación o influencia es significativa.
- Calcular el valor de Y a partir de las X.
  - Inferir (estimar) valores concretos

---

<sup>13</sup> Ver ejemplo de uso y referencias en <http://verso.mat.uam.es/~amparo.baillo/MatEstII/RegMultVarSel.html>

- Inferir (estimar) medias

En toda regresión hay una parte no explicada por el modelo: componente que a veces llamamos aleatoria (no lo es en la práctica, pero sí en el modelo teórico). Es el residual y lo caracterizamos con el ECMR (mse). El error estándar (su raíz cuadrada) es una magnitud fundamental y se expresa en las unidades de la variable explicada (p. ej. m<sup>3</sup>/ha)

Cuando decimos que medimos el volumen de un árbol, no es cierto, lo estimamos y eso tiene un margen de error: error de regresión que expresamos con el ECMR.

Atención porque en los cálculos de volumen de un árbol hay otra fuente de error que es muy importante y que normalmente no tenemos bien acotada: la relación altura diámetro -> Ver conclusiones de los cálculos de Zamora.

Otra forma de expresar cómo de buena es la regresión es la  $R^2$ : porcentaje de varianza explicada por el modelo. El resto de la varianza se debe a otras causas no incluidas en el modelo (no es realmente aleatoria). Como el  $R^2$  es un porcentaje parece más fácil de interpretar que el ECMR, pero lo que más nos interesa es el error estándar.

Las variables X que nos ayudan a explicar las variaciones de y (estimar el valor de y para un caso concreto) pueden ser varias -> regresión múltiple X: x<sub>1</sub>, x<sub>2</sub>, x<sub>3</sub>, etc.

Si tenemos muchas candidatas a variable explicativa corremos el riesgo de pensar que cuantas más, mejor. Pero no es así: el exceso puede ser más peligroso que el defecto por efectos indeseados.

En regresión hay una fase que es construir el modelo (elegirlo y calcular los estimadores de los parámetros y otra que es la inferencia (es más automática, pero cuidado con las variables explicativas que tienen margen de error).

La regresión lineal simple es fácil de representar e inspeccionar visualmente. La múltiple lo es un poco menos, pero también se puede inspeccionar visualmente, especialmente los residuos frente a cada una de las variables explicativas.

No se debe esconder una mala gestión estadística detrás de los números, coeficientes y parámetros estadísticos. Una vestimenta estadística elegante puede esconder una mala base conceptual y práctica.

#### 8.1.8.2 Regresión versus muestreo

Dos campos de la estadística, que tienen sus planteamientos, sus bases estadísticas, sus procedimientos, etc. Pero reflexionando sobre esta cuestión en el caso de las IOFs me ha llevado a integrar los dos campos, porque hay veces que jugamos a hacer modelos de regresión y en realidad estamos flirteando con el muestreo:

*regresiones débiles se apañan si la muestra con la que se construye la regresión es representativa*

En una regresión hay una componente que es el intercept y que es muy importante en estos casos, y que adquiere una interpretación concreta si trabajamos con variables explicativas normalizadas (o calculamos la y para los valores medios de las X): podemos vestir un muestreo de regresión y eso no es bueno, al menos si no se hace explícitamente.

Gráfica de nube de puntos poco relacionada con la X.

La interpretación puede ser válida si la aplico a todo el ámbito con el que he construido la regresión, pero no a un rodal concreto.

#### 8.1.8.3 El espacio muestral: población y muestra

Más importante que el número es la representatividad.

Media poblacional y media muestral (estimadora de aquella)

Varianza poblacional y varianza muestral (estimadora de aquella)

#### 8.1.8.4 Muestreo en ámbito territorial

Uno de los requisitos del muestreo es que las muestras sean aleatorias o por lo menos no estén relacionadas entre sí: en el territorio las muestras son más parecidas cuanto más cerca estén y eso es difícil de cuantificar y de incorporar al modelo.

Una muestra muy grande pero concentrada en una zona o en un tipo de lugares tiene menos valor que una más pequeña pero sin esos sesgos.

#### 8.1.8.5 Elección de la muestra en regresión y en muestreo

Elección - aleatoriedad – representatividad

- En regresión es importante que la variable explicada no condicione la muestra porque eso introduce sesgos, sin embargo, no lo es tanto para las variables explicativas (salvo que juguemos a hacer regresión cuando en realidad estamos haciendo muestreo).
- En muestreo, es muy importante la representatividad, que se consigue con aleatoriedad

Importante es espacio muestral: tiene que estar bien definido y debemos muestrearlo en todo su ámbito (territorial o dasométrico) **para ambos casos**.

#### 8.1.9 Hay que definir bien la variable que estimamos

La variable no queda definida sólo indicando el parámetro, sino tb u resolución espacial y demás detalles de cómo se obtiene.

#### 8.1.10 Escala (resolución espacial) y representatividad del “caso” individual

Cuando hablamos del valor de una variable dasométrica en un punto es una falacia: siempre es en una porción de terreno.

No es lo mismo que la variable explicada sea el volumen por hectárea en una parcela de 10 m de radio que en una de 25. La esperanza matemática (la media poblacional o su estimador que es la muestral) pueden ser iguales pero sus varianzas son muy diferentes y si queremos acotar el error de estimación hay que concretar esta cuestión y ser consecuentes con ello.

En un raster dasométrico, un pixel no es un punto sino una celda: ¿Cuánta heterogeneidad espacial queremos integrar? Eso influye en el tamaño del pixel que más nos interesa.

Visualizarlo en Qgis

#### 8.1.11 Como hacemos una regresión

Se puede hacer con una aplicación específica de estadística (de pago: spss, sas, statgraphics, etc.; libres: jasp, jamovi, etc.), con código (paquetes para python, R, Matlab) o, simplemente, con Excel.

No es mejor una regresión si se hace con un software más potente: el modelo ajustado debe ser el mismo en todos: cambia el uso, la interfaz, la información que aporta, etc.

#### 8.1.12 Relación lineal entre variables

La linealidad se refiere a los parámetros

En regresión múltiple podemos chequear la linealidad visualizando los residuos frente a las variables o con determinados estadísticos.

##### 8.1.12.1 Transformación de las variables explicativas

Sin problema en regresión

##### 8.1.12.2 Transformación de las variables explicada

Cuidado: factor de corrección para evitar sesgos

### 8.1.12.3 Linealización del modelo

$$y = A \cdot x^B \rightarrow \log(y) = \log(A) + B \log(x) \rightarrow YY = AA + B \cdot XX$$

Cuidado con los cálculos cuando se deshace la linealización  $\rightarrow$  sesgos

## 8.2 Resultados

### 8.2.1 Método y calidad del ajuste

Los modelos se han ajustado por mínimos cuadrados (OLS) usando el paquete statsmodels (<https://www.statsmodels.org>). La bondad de los ajustes se ha evaluado principalmente a través del error cuadrático medio residual (expresado como desviación estándar, para disponer de valores expresados en las mismas unidades que las variables explicadas) y, secundariamente, mediante la R<sup>2</sup> ajustada con los grados de libertad.

También se han estimado las varianzas de los estimadores de los parámetros para conocer hasta qué punto las distintas variables explicativas contribuyen a explicar las variables explicadas (t-test, p-value). No se han estandarizado las variables explicativas. Se han calculado los rangos y varianzas de las variables explicativas y se han acotado sus ámbitos de aplicación en el intervalo definido por la  $\mu \pm 1.5 \cdot \sigma$  ( $\mu$ : media muestral;  $\sigma$ : desviación estándar muestral).

Además, se ha evaluado el cumplimiento de los requisitos del modelo lineal, por su influencia en la fiabilidad de los intervalos de confianza en la inferencia.

### 8.2.2 Verificación de los requisitos del modelo lineal

Hola

### 8.2.3 Intervalos de confianza

Se han obtenido estos dos indicadores para el valor medio de todas las variables explicativas, como forma de expresar hasta qué punto el modelo es capaz de estimar con precisión la variable dasométrica, en función de la representatividad de la muestra con la que se hace la inferencia:

- Intervalo de confianza para las predicciones puntuales:

$$\hat{y} \pm t_{\alpha/2, n-2} \sqrt{MSE \left( 1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)}$$

- Intervalo de confianza para estimación de la media poblacional

$$\hat{y} \pm t_{\alpha/2, n-2} \sqrt{MSE \left( \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)}$$

El intervalo de confianza para la media poblacional, correspondiente a cada vector de valores de las variables explicativas, es tanto más estrecho cuanto mayor sea el número de muestras con el que se ajusta el modelo. Para caracterizar la calidad de la inferencia y, en su caso, identificar las regresiones que no cumplen unos mínimos, se ha establecido un valor máximo del 10% para la semiamplitud del intervalo del 95% de confianza (dejando 2.5% a cada lado), expresado en porcentaje del valor de la variable explicada para el tramo central de la regresión (el correspondiente al valor medio de las variables explicativas).

Eso significa que las predicciones hechas con los modelos que cumplen este requisito permiten estimar medias poblacionales con un error inferior al 10% en el 95% de las ocasiones, siempre y cuando el subrodal o lote en el que se aplica no ha tenido un régimen selvícola muy diferente del habitual en el ámbito de aplicación del modelo.

En el Apéndice 2 se recogen los estadísticos que caracterizan estos intervalos de confianza en la inferencia.

La predicción de valores concretos (parcelas concretas) tiene mayor dispersión y sólo puede reducirse aumentando el tamaño muestral con el que se construye el modelo hasta cierto punto, siendo irreductible la dispersión inherente a factores no contemplados en el modelo. En la inferencia, esta dispersión se reduce incrementando la superficie de aplicación del modelo, siendo esta reducción tanto más eficaz cuanto más representativa acabe siendo la superficie de aplicación.

#### 8.2.4 El talón de Aquiles del método y la cura en salud

Una de las claves de esta metodología, que puede considerarse una debilidad, reside no tanto en la construcción de los modelos y la evaluación de su calidad, como en su aplicación en rodales o lotes concretos.

El modelo lineal se construye para una determinada población (ámbito de aplicación) con una muestra más o menos representativa de la misma. Dicha representatividad se consigue muestreando ejemplares de la población sin sesgos motivados por la relación de la variable explicada con las explicativas y cubriendo adecuadamente el rango de validez para cada variable explicativa.

Este planteamiento (junto con el resto de los requisitos del modelo lineal) permite obtener modelos consistentes, que funcionan bien en la inferencia cuando se aplican en muestras representativas del ámbito de aplicación definido, pero no necesariamente en situaciones particulares que no lo son.

El modelo asume cierta heterogeneidad en la población (de ahí la inclusión de un error residual  $\epsilon$ , que integra los factores que afectan a la variable explicada y no se han incluido en el modelo) y la clave reside en que el error residual esté adecuadamente distribuido y compensado no solo cuando se ajusta el modelo como cuando se aplica en inferencia en un lugar concreto.

Si el modelo se utiliza para hacer inferencia en un subrodal o lote concretos, es probable que dicho subrodal o lote esté dentro del ámbito de aplicación del modelo pero constituye un caso particular no representativo en el sentido de que la relación entre las variables explicativas y explicada difiera de la media poblacional. Esto puede ser debido a causas naturales (suelo, rocosidad, altitud, pendiente, etc.) o antrópicas (diferente régimen selvícola, podas, etc.). En ese caso, aún cuando sea elevado el número de celdas para el que se hace la inferencia, los errores residuales no se compensan y se obtiene cierto sesgo en la estimación, que es difícil evaluar.

Para poner de manifiesto y acotar esta posible deficiencia en la aplicación de los modelos, en los gráficos de dispersión que muestran la relación entre la variable explicada y las principales variables explicativas de los modelos se han representado las parcelas con nueve colores distintos en función de su posición geográfica (malla de 3 x 3 en el correspondiente rango de latitud y longitud de cada ámbito de aplicación), de forma que una distribución localizada de alguno de los colores dentro del eje de ordenadas de la nube de puntos pondría en evidencia una subpoblación con una relación entre variables no representativa de la población y, por lo tanto, induciría a excluir o particularizar esa subpoblación del ámbito de aplicación general del modelo.

Los gráficos de dispersión obtenidos, que pueden consultarse en el Apéndice 1, reflejan que no hay patrones geográficos claros que obliguen a subdividir los ámbitos geográficos definidos para cada especie, dentro del margen que permite el número de parcelas disponible- Se ha establecido un número mínimo de 20 parcelas para valorar la creación de un ámbito geográfico.

Esta conclusión apoya a la hipótesis de que no hay problemas importantes de heterogeneidad dentro de los ámbitos geográficos definidos y da cierta confianza en la aplicación del modelo en zonas geográficas concretas dentro del ámbito de aplicación.

Es habitual, aunque no siempre, que las diferencias importantes entre regímenes selvícolas estén ligados a tres factores:

- El origen de la masa (regeneración natural o plantación).
- La densidad de plantación (incluye sobre todo en los primeros estadíos de la masa).
- La unidad administrativa (provincia, sección territorial o comarca).

El análisis realizado para afianzar la coherencia geográfica de los modelos apunta a una cierta coherencia en lo que se refiere a las diferencias de régimen selvícola ligado a ámbitos geográficos distintos.

Por su parte, la propia definición de ámbitos geográficos también contribuye parcialmente a reducir las diferencias por los distintos orígenes de masa, en la medida en la que algunos de ellos se caracterizan por una prevalencia mayoritaria de uno u otro origen (plantación y regeneración natural).

Debe asumirse que no hay un pleno control de este talón de Aquiles y, por lo tanto, en la aplicación de los modelos debe valorarse si el régimen selvícola del subrodal o lote en cuestión difiere mucho del mayoritario en el ámbito geográfico de aplicación para establecer las cautelas que correspondan a una incertidumbre en la inferencia superior a las que se indica en este documento.

### 8.2.5 Incorporación de la autocorrelación espacial a la inferencia

Referencias:

Giraldo, R. 2002. Introducción a la geoestadística. Teoría y aplicación.

Díaz Viera, M.A. 2002. Geoestadística aplicada.

<https://scholar.google.es/citations?user=fhPkQQ8AAAAJ&hl=es>

Para conocer la autocorrelación espacial (varianza ligada a la distancia entre puntos) se lleva a cabo un análisis estructural de los datos. Si tuviéramos datos de las variables explicadas medidos en parcelas a distintas distancias (desde parcelas contiguas -con una distancia entre sus centros de 10 m- con elevada autocorrelación hasta parcelas separadas 1 km, en que la autocorrelación es pequeña) podríamos llevar el análisis estructural con estas variables (VCC, AB, H0, DG, etc.). Como no los tenemos, llevamos a cabo este análisis con determinadas variables explicativas, asumiendo que el fenómeno de la autocorrelación se produce en la misma medida, de forma que usamos el semivariograma generado con estas como estimador del de aquellas. Para ello usamos las variables explicativas mejor relacionadas con las explicadas.

Para este análisis se asume que tanto las variables explicativas como las explicadas son variables aleatorias regionalizadas (llamamos  $Z(\underline{x})$  a la variable aleatoria en el punto  $\underline{x}$ ).

Si asumimos la hipótesis de estacionariedad (primer y segundo orden), entonces  $m = E[Z(\underline{x})]$  y  $\sigma^2 = \text{Var}[Z(\underline{x})]$  son invariantes para todo  $\underline{x}$ .

Además, esa variable  $Z(\underline{x})$  presenta una estructura espacial de correlación que depende de  $h$ , de forma que para cualquier par de variables aleatorias  $Z(\underline{x})$  y  $Z(\underline{x+h})$  la función de covarianza es:

$$C(h) \equiv C(\underline{x+h}, \underline{x}) = E[Z(\underline{x+h}) \cdot Z(\underline{x})] - m^2$$

$$C(h) = \text{Cov}[Z(\underline{x+h}), Z(\underline{x})] \equiv E[\{Z(\underline{x}) - m\} \cdot \{Z(\underline{x+h}) - m\}]$$

La varianza de la muestra o varianza a priori es  $C(0) = \sigma^2$  (¿se puede llamar autocorrelación?)

El semivariograma,  $\gamma(h)$ , representa la variación de la diferencia de variables aleatorias obtenida usando pares de puntos:

$$2\gamma(\underline{x}_i, \underline{x}_j) = \text{Var}[Z(\underline{x}_i) - Z(\underline{x}_j)] \Rightarrow \gamma(\underline{x}_i, \underline{x}_j) = \frac{1}{2} E\left[\{Z(\underline{x}_i) - Z(\underline{x}_j)\}^2\right]$$

Y, en consecuencia,  $\gamma(h) = \sigma^2 - C(h)$

Díaz Viera expone esta relación como  $\gamma(h) = C(0) - C(h)$ , pero no me convence porque si  $C(0) = E[Z(\underline{x+h}) \cdot Z(\underline{x})] - m^2$ , y  $E[Z(\underline{x+h}) \cdot Z(\underline{x})] = m^2$ , entonces  $C(0) = 0$ .

El semivariograma solo depende de la distancia entre los puntos ( $h$ )

El elefante en la habitación:

Para

### 8.3 Correlación entre variables

### 8.4 Relación entre variables explicadas y explicativas

#### 8.4.1 Por especie y provincia

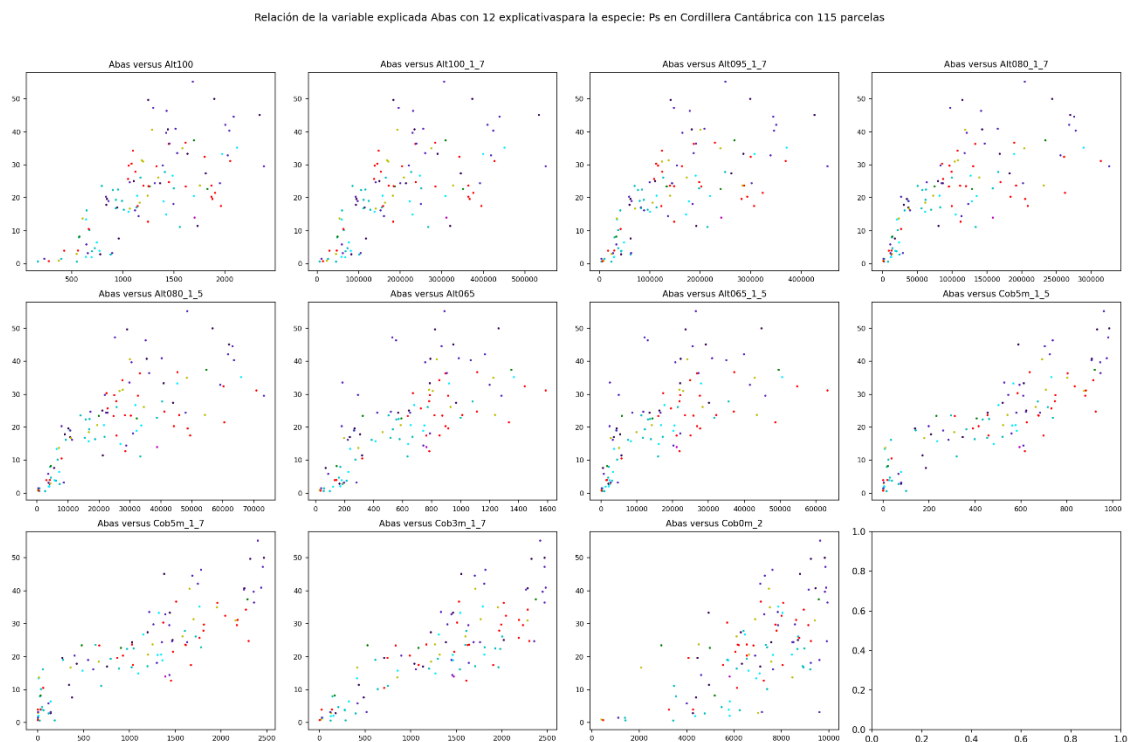
#### 8.4.2 Por estratoLidar

### 8.5 Ecuaciones dasoLidar

### 8.6 Control de calidad

## 9 Documentación de capas Lidar

## Apéndice 1. Gráficos de dispersión de las variables explicativas *versus* explicadas



## Apéndice 2. Intervalos de confianza en la predicción

### 10 Notas personales y bitácora

Ver información asociada en O:\Sigmena\usuarios\COMUNES\Bengoa\Lidar\\_planLidar:

- Plan de Trabajo
- Cursos ECLAP

Bitácora:

- 202405. Generar una versión preliminar (alpha) de capas Lidar (métricas Lidar y variables dasométricas) para usarlas en el curso 2024MAV005.
- 202406. Curso 2024MAV005
- 202406-202408. Triple tarea:
  - Interacción con alpha-testers para depurar los productos
  - Generación de la versión beta de las capas Lidar
  - Documentar el procesado y los productos
- 202409a. Publicación de la versión beta de las capas y productos Lidar. Documentación, manual y factsheet (chuleta).
- 202409b. Ronda de contactos o videoconferencia con Alpha-testers
- 202410. Iniciativas de difusión de los productos dasométricos derivados del Lidar.
- 202411. Curso ECLAP 2024MAV006



