



---

**Universidad de Valladolid**

TESIS DOCTORAL

**Aplicación de imágenes de satélites y datos LiDAR en la  
modelización de recursos forestales**

**Forest resource modelling combining satellite imagery and  
LiDAR data**

Autora

Ángela Blázquez Casado

Directores

Dr. José Miguel Olano Mendoza

Dr. José Ramón González Olabarría

Dr. Francisco Rodríguez Puerta

Enero 2019



Instituto  
Universitario de Investigación  
GESTIÓN  
FORESTAL  
SOSTENIBLE





---

**Universidad de Valladolid**

**PROGRAMA DE DOCTORADO EN CONSERVACIÓN Y USO  
SOSTENIBLE DE SISTEMAS FORESTALES**

TESIS DOCTORAL

**Aplicación de imágenes de satélites y datos LiDAR en la  
modelización de recursos forestales**

**Forest resource modelling combining satellite imagery and  
LiDAR data**

Presentada por Ángela Blázquez Casado para optar al grado  
de Doctora con Mención Doctorado Industrial por la  
Universidad de Valladolid

Dirigida por:  
Dr. José Miguel Olano Mendoza  
Dr. José Ramón González Olabarría  
Dr. Francisco Rodríguez Puerta



*“La educación es el arma más poderosa  
que puedes usar para cambiar el mundo”*

Nelson Mandela

*A mi familia*

## INDICE

---

<b>Nota a los lectores</b> .....	9
<b>Resumen</b> .....	11
<b>Abstract</b> .....	13
<b>Capítulo 1. Introducción</b> .....	18
1.1 La importancia de los recursos forestales.....	18
1.2 Uso de imágenes de satélite en la evaluación de recursos forestales.....	20
1.3 La tecnología LiDAR en la evaluación de recursos forestales.....	24
1.4 Integración de imágenes de satélite y datos LiDAR como herramientas en la evaluación de recursos forestales.....	27
1.5 La modelización forestal integrando información procedente de sensores remotos.....	31
1.6 Aproximación a distintas escalas.....	32
1.7 Objetivos y alcance.....	34
1.8 Marco de estudio.....	34
<b>Chapter 2. Combining low density LiDAR and satellite images to discriminate species in mixed Mediterranean forest</b> .....	42
Abstract.....	42
2.1 Introduction.....	43
2.2 Material and methods.....	46
2.2.1 Study area.....	46
2.2.2 Field data.....	48
2.2.3 Remote sensing data.....	49
2.2.4 Remotely sensed data processing.....	49
2.2.4.1 LiDAR data processing.....	49
2.2.4.2 Tree crown selection.....	53
2.2.4.3 Plèiades data processing.....	53
2.2.5 Classification technique.....	54
2.3 Results.....	55
2.3.1 Training model selection.....	55
2.3.2 Validation model.....	57
2.4 Discussion.....	58
2.5 Conclusion.....	62
2.6 References.....	63

<b>Chapter 3. Stand types discrimination comparing machine learning algorithms in monteverde, Canary Islands</b> .....	72
Abstract.....	72
3.1 Introduction.....	73
3.2 Material and methods.....	74
3.2.1 Study area.....	74
3.2.2 Field data.....	74
3.2.3 LiDAR data.....	75
3.2.4 Multispectral imagery data source.....	75
3.2.5 Segmentation process.....	76
3.2.6 Data analysis.....	76
3.3 Results and discussion.....	77
3.4 References.....	81
<b>Chapter 4. Fire and burn severity assessment: calibration of Relative Differenced Normalized Burn Ratio (RdNBR) with field data</b> .....	86
Abstract.....	86
4.1 Introduction.....	87
4.2 Material and Methods.....	89
4.2.1 Data Sources.....	89
4.2.1.1 Studied fires.....	89
4.2.1.2 Field data.....	90
4.2.2 Satellite image collation and processing and RdNBR calculation.....	93
4.2.3 Methodological approaches.....	93
4.3 Results.....	95
4.4 Discussion.....	100
4.5 References.....	104
<b>Capítulo 5. Discusión general, Futuras líneas de trabajo y Conclusiones</b> .....	112
5.1 Discusión general.....	112
5.2 Futuras líneas de trabajo.....	117
5.3 Conclusiones.....	119
<b>Agradecimientos</b> .....	124
<b>Referencias</b> .....	128









## **Nota a los lectores**

---

Esta tesis doctoral está basada en cinco capítulos originales. El cuerpo de la tesis está representado en los capítulos dos, tres y cuatro, los cuales, han sido o serán publicados en revistas científicas de impacto. Así, cada uno de ellos contiene su propia bibliografía. El resto de las citas incluidas en la tesis pueden ser consultadas en el epígrafe final de referencias.

### **Capítulo 1** Introducción

**Chapter 2** Combining low density LiDAR and satellite images to discriminate species in mixed Mediterranean forest

Ángela Blázquez-Casado, Rafael Calama, Manuel Valbuena, Marta Vergarechea, and Francisco Rodríguez. *Under reviewer to Annals of Forest Science*.

**Chapter 3** Stand types discrimination comparing machine-learning algorithms in monteverde, Canary Islands

Miguel García-Hidalgo, Ángela Blázquez-Casado, Beatriz Águeda and Francisco Rodríguez. *Forest Systems*, 27 (3), eSC03, 5 pages (2018). <http://doi.org/10.5424/fs/2018273-13686>.

**Chapter 4** Fire and burn severity assessment: calibration of Relative Differenced Normalized Burn Ratio (RdNBR) with field data

Adrián Cardil, Blas Mola-Yudego, Ángela Blázquez-Casado and José Ramón Gonzalez-Olabarria. *Journal of Environmental Management* 235: 342-349. doi: 10.1016/j.jenvman.2019.01.077

**Capítulo 5** Discusión general, Futuras líneas de trabajo y Conclusiones



## Resumen

---

El sector forestal tiene un papel relevante en la transición hacia una economía innovadora y eficiente en el consumo de sus recursos. Conocer la disponibilidad espacial de los recursos forestales y su evolución temporal es crítico en la gestión forestal, tanto de los recursos maderables como de los no maderables. El uso de información procedente de sensores remotos se está convirtiendo en una opción cada vez más rigurosa y asequible para el desarrollo de esta tarea. Así, el conocimiento que estas herramientas proporcionan sobre el estado de desarrollo de las masas forestales y la disponibilidad de sus recursos permite hacer frente a los diferentes escenarios futuros que plantea el actual contexto de cambio global.

Esta tesis caracteriza y evalúa diferentes recursos forestales mediante la combinación de información continua procedente de imágenes de satélite y datos LiDAR, con diferentes niveles de resolución espacial y espectral. Estos datos, apoyados en trabajo de campo, han sido calibrados y validados, demostrando un gran potencial. Discriminar diferentes especies y tipos de masa, tanto a nivel de árbol individual como de objeto, son objetivos alcanzables mediante el uso adecuado de estas herramientas, disminuyendo la dependencia histórica del trabajo de campo e integrando el cambio de escala en los inventarios tradicionales. Esta tesis desarrolla herramientas robustas capaces de evaluar recursos forestales a gran escala mediante modelos mixtos lineales y técnicas de modelización basadas en aprendizaje automático.

El capítulo 2 combina datos LiDAR de baja densidad e imágenes Plèiades para discriminar entre *Pinus pinaster* Ait. y *Pinus pinea* L. a nivel de árbol individual en bosques mixtos Mediterráneos. La predicción de los productos forestales no maderables derivados de estas dos especies es particularmente importante debido a la considerable diferencia en cuanto a la gestión y producción de cada uno de ellos. El modelo combinado con ambas fuentes de información aporta la mayor precisión, del 83.3% y 63% en las masas puras y mixtas, respectivamente. Esta aproximación es fácilmente aplicable a grandes áreas, aportando valor añadido a los productos forestales y ofreciendo una herramienta útil para los gestores.

El capítulo 3 compara cuatro metodologías de modelización de aprendizaje automático diferentes: RF, SVML, SVMR y ANN (*Random Forest, Linear Support Vector Machine, Radial Support Vector Machine y Artificial Neural Networks*) para la clasificación de las distintas tipologías de masas forestales presentes en los bosques de cobertura completa de monteverde Canario mediante información multitemporal procedente de sensores remotos. Los modelos comparados presentan una precisión entorno al 90%, con resultados muy similares, lo que hace que ninguno de ellos se postule como claramente superior a los demás, a pesar de que pequeñas diferencias entre ellos pueden ser importantes cuando se aplican a grandes superficies.

El capítulo 4 evalúa la severidad post-incendio mediante la calibración del índice espectral RdNBR (*Relative Differenced Normalized Burn Ratio*) con variables medidas en 28 grandes incendios repartidos por tres países del sur de Europa. El índice proporciona la capacidad de evaluar, a través de variables que reflejan la intensidad (altura de quemado en la corteza) y la severidad (pérdida de arbolado y matorral), el impacto inducido por el fuego sobre los sistemas forestales. Este puede ser traducido en estimaciones de pérdida de biomasa cuando los modelos predictivos, con un 67% de fidelidad, son aplicados sobre superficies grandes y heterogéneas. La variación en la respuesta de la vegetación tras el incendio se refleja en la importancia del tiempo transcurrido entre el incendio y su evaluación en la formulación de los modelos. Los modelos mixtos reflejan la existencia de variaciones importantes entre los incendios debido a contextos específicos que modifican estas relaciones, y, el uso de BRT (*Boosted Regression Trees*) muestra que las relaciones entre las medidas de severidad y la respuesta espectral no tienen que ser necesariamente lineales.

## Abstract

---

Forestry sector plays an important role in the transition towards a new economy, driven by efficient resource consumption. Understanding the spatial distribution of forest resources and its temporal evolution is critical in forest management, both for timber and non-timber resources. Remote sensing information is becoming an increasingly precise and affordable option for the accomplishment of this task. The knowledge provided by these tools regarding stand development and availability of resources enables predicting future global change scenarios.

This Doctoral Thesis assesses different forest resources combining continuous information derived from satellite images and LiDAR data at different spatial and spectral resolution levels. This information, supported by field work, has been calibrated and validated, showing a great potential. Species and stand types discrimination, both at individual tree and object levels, can be accomplished with these tools, decreasing the historical dependence of field work and integrating the scale change in traditional inventories. This PhD work aims to develop robust tools able to evaluate large-scale forest resources, by means of linear mixed models and machine learning.

In chapter 2, low density LiDAR data and Plèiades images were combined to discriminate between *Pinus pinaster* Ait. and *Pinus pinea* L. at individual tree level in mixed Mediterranean forests. Non-timber forest products prediction derived from these two species is particularly important due to the considerable difference in management and production of each of them. The model generated with both information sources provides the highest precision in the pure and mixed stands, 83.3% and 63%, respectively. This approach is easily applicable over large areas, providing added value to forest products and a useful tool for managers.

In chapter 3, four different machine learning modelling methodologies were compared, RF, SVML, SVMR and ANN (Random Forest, Linear Support Vector Machine, Radial Support Vector Machine and Artificial Neural Networks), in order to classify different forest typologies in full-coverage forests of Canarian monteverde, using remote sensing multitemporal information. Compared

models show an accuracy around 90%, and are very similar, so none of them is postulated as clearly superior to the others, although small differences between them might reveal relevant when applied over large surfaces.

In chapter 4, post-fire severity was assessed by calibrating the spectral index RdNBR (Relative Differenced Normalized Burn Ratio) with variables measured in 28 large fires across three southern European countries. The index provides the ability to assess, through variables that reflect intensity (mean bark char height) and severity (loss of trees and shrub cover), the impact induced by fire on forest ecosystems. The results can be incorporated into biomass loss estimations when predictive models, with 67% of accuracy, are applied across large and heterogeneous regions. The importance of the time between the fire and its assessment is reflected in the model formulations, due to vegetation response variations after the fire. Mixed models highlight significant variations between fires due to specific contexts that modify these relationships. The use of BRT (Boosted Regression Trees) shows that relationships between severity measures and spectral information may not be linear.







# CAPÍTULO 1

---

## Introducción

*“Caminante, son tus huellas  
el camino y nada más;  
Caminante, no hay camino,  
se hace camino al andar”*

Antonio Machado

## Capítulo 1. Introducción

---

### 1.1 La importancia de los recursos forestales

A lo largo de la historia del ser humano, los bosques han tenido una enorme importancia. Tanto la madera, como muchos otros productos forestales han supuesto la base de la economía y el bienestar de nuestros antepasados. Actualmente, la actividad forestal es fuente de alimentación, medicinas y combustible para más de un billón de personas, contribuyendo al desarrollo económico en las zonas rurales, especialmente en las zonas más desfavorecidas (FAO 2018). En este sentido, los bosques y el sector forestal juegan cada vez un papel más importante en la transición hacia una economía innovadora y eficiente en el consumo de recursos (Wong and Prokofieva 2014). Esto es, una economía basada en el conocimiento y la utilización de recursos ecológicos, procesos y métodos biológicos para proporcionar bienes y servicios de forma sostenible en todos los sectores económicos (*bioeconomy*) (Carus 2017).

El concepto de servicios ecosistémicos se refiere a los bienes y servicios que los ecosistemas proveen a los seres humanos. Estos servicios son múltiples. Por ejemplo, los servicios de apoyo como el reciclado de los nutrientes; los servicios de aprovisionamiento referidos a la cantidad de bienes o materias primas que un ecosistema ofrece como la madera, el agua o los alimentos; los servicios de regulación, como la protección frente a inundaciones, la depuración de las aguas, la polinización de los cultivos y los servicios culturales que proporcionan beneficios recreativos, estéticos y espirituales (Hassan et al. 2005). Tradicionalmente son los servicios de aprovisionamiento los que se han cuantificado a través de un beneficio económico directo, por ejemplo, la madera u otros recursos no maderables como el piñón o la resina (Calama et al. 2010). Sin embargo, en los últimos tiempos ha aumentado la conciencia de otros servicios como los servicios culturales y de soporte como la biodiversidad (Rego et al. 2014). La persistencia de estos servicios está amenazada por los cambios de usos, el cambio climático y, en definitiva, los diferentes componentes del cambio global (Thom and Seidl 2016).

De acuerdo con el último informe del estado de los bosques en Europa (FAO 2015), el valor de mercado total de los productos forestales no maderables se estima en 2277 millones de euros para toda la región europea, de los cuales el 73.0% fueron generados por productos derivados directamente de las plantas, aunque los datos disponibles son difícilmente comparables debido a la falta de homogeneidad en la recolección de estos por los propios países y a la cantidad variable de países que los recogen estos datos anualmente. En el anterior informe publicado en 2011 (FAO 2011) participaron cinco países más en la recolección de dichos datos y se estimó que el valor de mercado total de los productos forestales no maderables fue durante el anterior periodo de 2763 millones de euros, un 17.6% menos que en el informe actual. Sin embargo, el porcentaje de productos derivados directamente de plantas fue del 76.6%, solamente un 3.6% más que en el mercado actual. Por tanto, puede considerarse que el mercado de los productos forestales no maderables ha sido estable en el tiempo a pesar de la crisis sufrida en los mercados europeos.

En zonas donde el crecimiento de los árboles está más limitado por factores climáticos como, por ejemplo, en la Cuenca del Mediterráneo los aprovechamientos de los productos forestales no maderables (corcho, setas, trufas, piñón, etc.) tienen un valor superior al de la madera (Vázquez-Piqué and Pereira 2008; Palahí et al. 2009; Calama et al. 2010). De hecho, en Europa el sector de los productos forestales alimenticios (setas, castaña, piñón, etc.) está liderado por Italia, España y Turquía (44600, 42100 y 22700 toneladas respectivamente). Si bien, en términos económicos el patrón es ligeramente diferente, encabezando los datos de valor económico España, Italia y Portugal con 196, 88 y 55 millones de euros respectivamente (FAO 2015).

En los ecosistemas mediterráneos, los incendios son una de las principales amenazas que se ciernen sobre sus servicios ecosistémicos (Seidl et al. 2015). La superficie forestal media afectada por incendios en los países miembros de la Unión Europea fue de 410776 ha en los últimos 5 años (2013–2017), un 0.23% de la superficie forestal total. El impacto de los incendios se concentra de modo desproporcionado en los países del Suroeste de Europa (España, Italia y Portugal), que con un 23% de la superficie forestal, tuvieron el 83% de la superficie quemada (JRC-EFFIS 2017). Los datos de superficie quemada pueden ser engañosos, ya que la cantidad de superficie afectada por un incendio no indica, por sí sola, ni la gravedad del daño, ni las pérdidas económicas asociadas. En

este sentido, la inclusión de la severidad de un incendio permite estimar de un modo más certero la pérdida de recursos. Esto es muy relevante, ya que la cuantificación de los daños reales es clave para evaluar adecuadamente las medidas políticas y prácticas de manejo forestal sostenible relacionadas con la prevención y lucha contra el fuego.

La gestión de los montes o, en su caso, la ausencia de gestión tiene un impacto crítico sobre la disponibilidad de los servicios ecosistémicos (Asta et al. 2002; Arozena et al. 2015). Sin embargo, asegurar la persistencia de la masa es siempre el objetivo principal en la gestión forestal, ejecutando aprovechamientos que serán la base de la financiación del coste de las mejoras necesarias para mantener la multifuncionalidad de nuestros bosques. Así, la gestión forestal actual y futura debe de ir encaminada hacia una optimización sostenible y multifuncional de los recursos y servicios ecosistémicos forestales de los bosques (Wong and Prokofieva 2014).

### **1.2 Uso de imágenes de satélite en la evaluación de recursos forestales**

La teledetección se define como una técnica de adquisición de datos de la superficie terrestre desde la distancia usando la señal reflejada desde la superficie terrestre en una o más regiones del espectro electromagnético (Campbell and Wynne 2011). Los inicios de esta disciplina se remontan al año 1858, cuando se tomaron las primeras imágenes desde el aire mediante globo aerostático en Francia (Campbell and Wynne 2011). Ya en 1909 se tomaron las primeras imágenes desde avión, para poco después convertirse en una herramienta de uso habitual para la observación de la superficie durante la Primera Guerra Mundial (Fig. 1).



**Fig. 1:** Fotograma tomado sobre la ciudad de Soria el 22/08/1946. Escala del vuelo: 40,890. Fuente: Fototeca Digital CNIG – Vuelo AMS-46/47. Ministerio de Defensa (CECAF) procedente de los archivos del Ejército del Aire. Nombre del fichero: H0350\_294\_050.ecw.

Los satélites son un paso más allá en la observación de la superficie terrestre. El primer satélite, Sputnik I, fue lanzado por la Unión Soviética en 1957 y fue el detonante de una carrera tecnológica de naturaleza política, militar, tecnológica y científica, marcando el inicio de la era espacial (Garber 2007). Esta tecnología, que en su origen era de naturaleza militar y de uso muy restringido, ha ido popularizándose para usos civiles. De hecho, en la actualidad la teledetección se usa de forma habitual en muy diversas disciplinas, siendo un instrumento indispensable para la observación y el estudio de los ecosistemas y sus recursos, así como los cambios que éstos están sufriendo en el nuevo contexto de cambio global (Tielidze 2015; Urban et al. 2018).

A lo largo de las últimas décadas la disponibilidad de imágenes satelitales ha aumentado notablemente. Las imágenes procedentes de los satélites de las diversas misiones Landsat han sido las más utilizadas, a pesar de su resolución espacial media (30 metros en el multiespectral), por la disponibilidad de su amplio repositorio (lanzado en 1972), buena resolución espectral y su alta resolución temporal (16 días de revista). Un ejemplo de ello es la creación de mapas de cobertura forestal en Canadá (Wulder et al. 2008) o la evaluación de

la dinámica que experimentan los bosques europeos durante los más de 40 años que lleva este satélite registrando imágenes (Potapov et al. 2015).

Otras misiones posteriores, como SPOT o Plèiades fueron lanzados siendo uno de sus objetivos la disminución de la resolución espacial, llegando hasta 2.5 y 0.5 metros respectivamente (Fig. 2). Un ejemplo de su uso lo encontramos en el trabajo de Meng et al. (2016) en el que se estudia la diversidad estructural de los bosques de la región autónoma de Guangxi Zhuang en China mediante información procedente del satélite SPOT 5. Sus resultados indicaron que variables como el área basimétrica pueden predecirse de manera fiable utilizando variables espectrales o de textura procedentes de estas imágenes. Otro ejemplo es el trabajo de Akbari and Kalbi (2017), donde se ha examinado la capacidad real de las imágenes de muy alta resolución Plèiades para determinar diversidad de especies en bosques mixtos del norte de Irán. Si bien es indudable que estos satélites con muy alta resolución espacial aportan gran precisión, su elevado coste y su baja resolución espectral limitan su utilización.



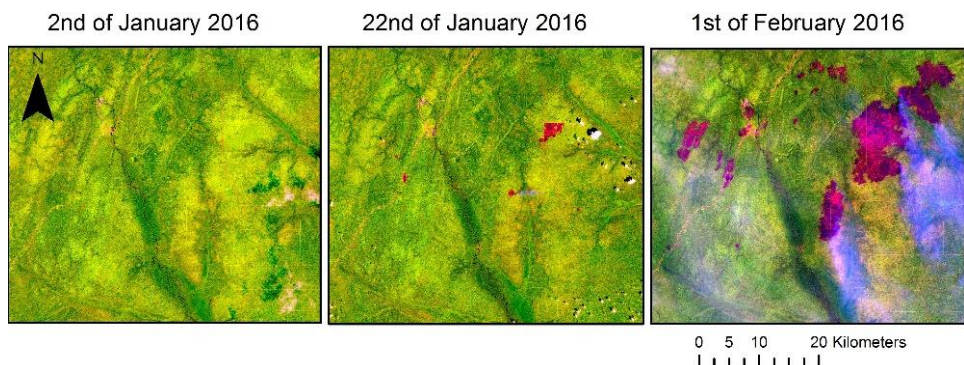
**Fig. 2:** Imagen multispectral de muy alta resolución tomada por el satélite Plèiades.

En este contexto, los satélites Sentinel 2A y 2B, en órbita desde junio de 2015 y marzo de 2017, respectivamente, han sido diseñados expresamente para monitorizar la superficie de la tierra y sus cambios. Estos satélites presentan alta resolución espectral (13 bandas) y resolución espacial media (hasta 10 metros).



Sin embargo, el potencial de esta misión se basa en su trabajo en conjunto para proveer de una cobertura global de la superficie terrestre cada 5 días. Actualmente, se está evaluando la potencialidad de Sentinel para el monitoreo de la cobertura terrestre. Por ejemplo, en el ámbito agrícola, Immitzer et al. (2016) generaron una clasificación para el mapeo de seis tipos de cultivos de verano a partir de imágenes de Sentinel. Li et al. (2018) construyeron índices de vegetación basados en las bandas del visible, borde del rojo e infrarrojo cercano de Sentinel para evaluar el contenido de clorofila en las copas de los árboles de cultivos agroforestales con el objetivo de detectar la capacidad fotosintética y estadios de estrés en el desarrollo de las plantas.

Otro de los usos potenciales en el ámbito forestal es el estudio de perturbaciones (Hirschmugl et al. 2017). Concretamente, los daños causados por incendios (Fig. 3) y por tormentas de viento y nieve son algunos de los más habituales en los bosques templados de Europa (FAO 2015). Mejorar el conocimiento del estado actual de los bosques y, en concreto, de estas masas donde existe riesgo o donde ya han sufrido una perturbación pudiéndose cuantificar la pérdida de recursos forestales tanto maderables como no maderables, es objeto de evaluación.



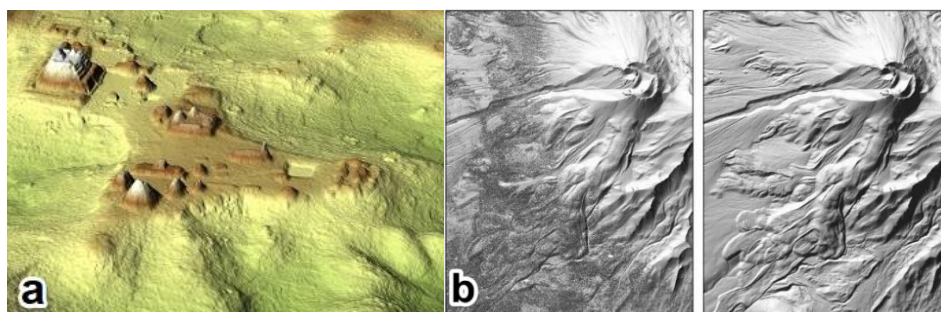
**Fig. 3:** Progresión en áreas forestales quemadas en la República del Congo con imágenes de Sentinel (Verhegghen et al. 2016).

De cara a la gestión forestal, el volumen de existencias es una de las variables biofísicas más comunes en las que se basa la planificación de una masa forestal. La exploración de las posibilidades que ofrecen, en este sentido, las imágenes multiespectrales procedentes de satélites también han sido

estudiadas por diversos autores (Chrysafis et al. 2017; Mura et al. 2018) y, en concreto, la estimación de recursos forestales no maderables, especialmente la biomasa o el stock de carbono (Castillo et al. 2017; Muhati et al. 2018). La integración de las imágenes de satélite en la planificación forestal es de gran utilidad en la toma de decisiones de cara a definir la planificación de una masa forestal.

### 1.3 La tecnología LiDAR en la evaluación de recursos forestales

La tecnología LiDAR (*Light Detection And Ranging*) aerotransportada se basa en un escáner laser incorporado a un medio aerotransportado. LiDAR utiliza pulsos de láser para medir distancias siendo capaz de generar nubes de puntos tridimensionales que miden tanto la cobertura como la superficie de la tierra (Glennie et al. 2013). El concepto de LiDAR fue desarrollado entre los años 60 y 70 como herramienta para el estudio del relieve y mediciones batimétricas (Lohani and Ghosh 2017). En las últimas décadas, multitud de disciplinas han incorporado el LiDAR, encontrando soluciones innovadoras a problemas de mapeo mediante esta herramienta altamente precisa. Evaluaciones arqueológicas de ciudades hoy desaparecidas (Fisher et al. 2017), seguimiento de derrumbes en la superficie (Daehne and Corsini 2013) o monitoreo de volcanes activos (Behncke et al. 2016) son algunos ejemplos donde se ha demostrado el potencial de los datos LiDAR (Fig. 4).



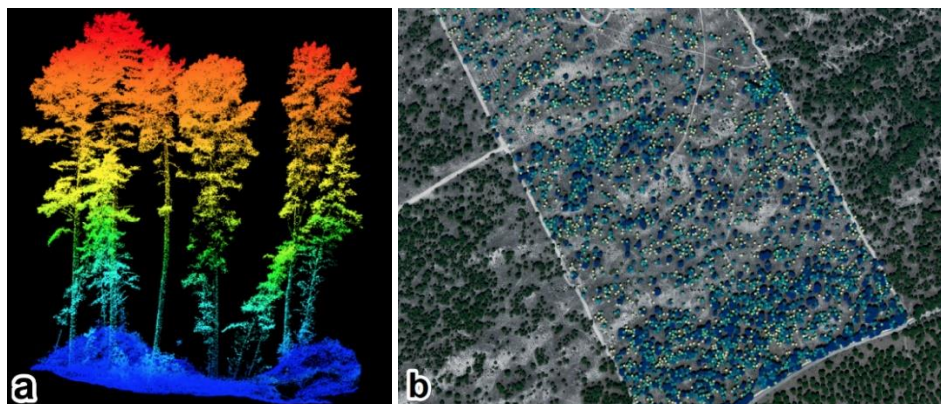
**Fig. 4:** (a) Datos LiDAR para trabajos arqueológicos de la ciudad maya de Caracol (Belice). <https://bit.ly/2BnkUBW>. (b) Datos LiDAR par observación de flujos de lava y otras características superficiales del cono Shastina (California). A la derecha con la vegetación presente y a la izquierda con la vegetación eliminada, USGS. <https://on.doi.gov/2Ew0P08>.

Hoy en día, muchos países en el mundo ya han barrido al completo su superficie con sensores LiDAR aerotransportados. Suiza, Polonia o Países Bajos son algunos ejemplos. En España, el Instituto Geográfico Nacional (IGN) a través del Plan Nacional de Ortofotografía Aérea (PNOA) ha completado recientemente un primer vuelo para toda la superficie nacional a una densidad de 0.5 puntos/m<sup>2</sup>. En los últimos años se ha empezado a realizar un segundo vuelo aumentando la densidad a 1 punto/m<sup>2</sup> y manteniendo su precisión altimétrica a menos de 0.20 m. Habitualmente, estos datos se están adquiriendo con el objetivo de obtener modelos detallados de la elevación del terreno, pero también sirven como fuente de datos para el mapeo de la vegetación y, en concreto, para los inventarios forestales.

La aplicación de LiDAR en el sector forestal ha supuesto una verdadera revolución. Esta herramienta experimentó su primer impulso a principios de los años 90 como herramienta de inventario forestal. Su uso permitía proporcionar información sobre la cobertura vegetal a gran escala y con muy alta precisión (Maltamo et al. 2014). La fracción de cabida cubierta y la altura son variables estructurales medidas directamente con LiDAR y que se usan ya rutinariamente para el estudio y caracterización de las masas forestales (Lim et al. 2003; Næsset et al. 2004; Guerra-Hernández et al. 2016). La gran fortaleza de utilizar LiDAR para aplicaciones forestales es la capacidad de generar con precisión la estructura tridimensional del dosel (Maltamo et al. 2014).

El inventario de los datos LiDAR puede ser realizado mediante dos metodologías distintas: por el método de masa o bien por el método de árbol individual. Los datos LiDAR procesados por el método de masa han sido tradicionalmente usados para la generación de los inventarios forestales a mediana y gran escala, donde uno de los principales objetivos es estimar el volumen de una masa forestal. Los primeros estudios encaminados en este sentido buscaron relaciones entre los atributos del bosque a nivel de parcela basados en mediciones en campo y las distribuciones de altura proporcionados por los datos LiDAR (Maltamo et al. 2014) (Fig. 5a). En estos casos, el conjunto de la masa es el objetivo, como ocurre en los inventarios forestales a nivel nacional o regional (Magnussen et al. 2018). Posteriormente, se desarrolló el procesado de los datos LiDAR por el método de árbol individual. Este método presenta un enfoque dirigido hacia trabajos donde el individuo es la parte importante (Wu et al. 2016). Esta metodología generalmente se basa en la búsqueda de máximos

dentro de la masa forestal para identificar copas individuales. Por ello, requiere de un espacio mínimo entre pies para considerar con mayor exactitud copas individuales y, por tanto, identificar los pies dentro de la masa (Fig. 5b).



**Fig. 5:** (a) Imagen de una parcela de inventario de *Pinus radiata* D. Don. con LiDAR de muy alta resolución en Australia. Fuente: <https://www.forestry.org.au/>. (b) Imagen de la estimación de altura en un inventario de *Pinus pinea* L. con LiDAR a nivel de árbol individual en España. Fuente: <https://fora.es/>.

Existen numerosos estudios que demuestran el potencial de la aplicación del LiDAR para la evaluación de recursos forestales, tanto maderables como no maderables. La cuantificación en un inventario forestal del volumen de madera y su relación con la estimación del recurso de biomasa en base a variables obtenidas directamente de datos LiDAR ya es un hecho (Popescu et al. 2003; Latifi et al. 2010). Recientemente, también se ha estudiado la relación directa de variables LiDAR con el stock de carbono en nuestros bosques (García et al. 2010), con la ocurrencia de hongos (Peura et al. 2016; Thers et al. 2017) o con la deforestación (Knapp et al. 2018).

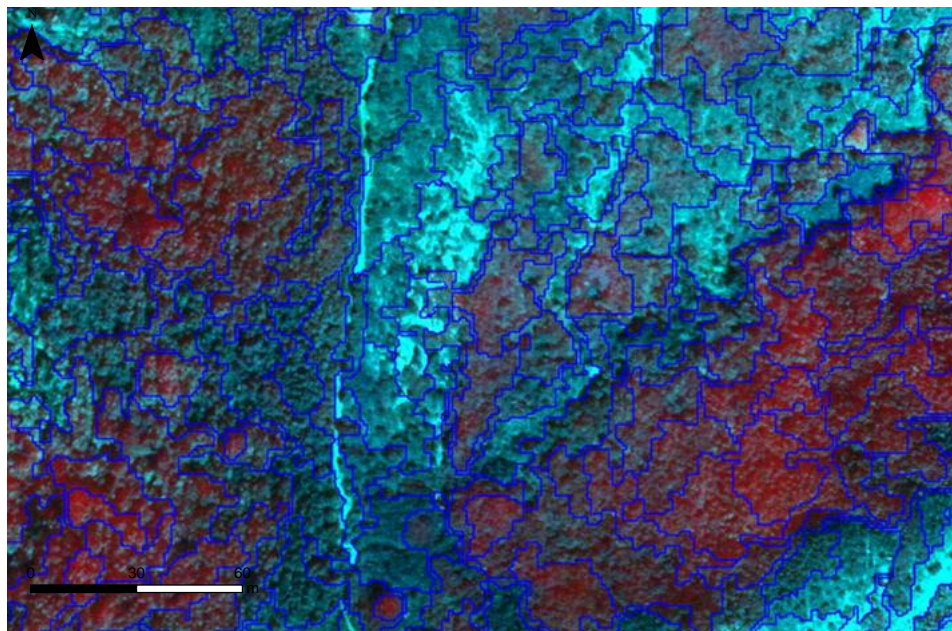
En este sentido, la aparición de series temporales de vuelos LiDAR abre un gran abanico de posibilidades en el estudio de la estructura y evolución de nuestros bosques (Marinelli et al. 2018). Mediante la caracterización con datos LiDAR multitemporales, se puede, por ejemplo, caracterizar las dinámicas de claros en las masas forestales tras el paso una perturbación, como puede ser un incendio (Vepakomma et al. 2008). Así, el monitoreo del crecimiento de los árboles, la dinámica de la biomasa o del secuestro de carbono son objetivos alcanzables

con esta tecnología (Zhao et al. 2018) y especialmente útiles en la toma de decisiones en gestión forestal.

#### **1.4 Integración de imágenes de satélite y datos LiDAR como herramientas en la evaluación de recursos forestales**

El uso combinado de los datos procedentes de sensores remotos, como son los datos LiDAR y las imágenes de satélite, se ha convertido en un medio asequible y riguroso para la caracterización espacial y temporal de los ecosistemas forestales, dado que permiten la provisión de información espacial y temporalmente continua (Blázquez-Casado et al. 2015). La adquisición de estos datos procedentes de sensores remotos no sustituye la necesidad de obtener datos de campo sólidos, pero disminuye notablemente su volumen (Stereńczak et al. 2018). Por tanto, la aplicación de estas tecnologías incide en la reducción en los costes de toma de datos *in situ* generando importantes ahorros (Rodríguez and Lizarralde 2015).

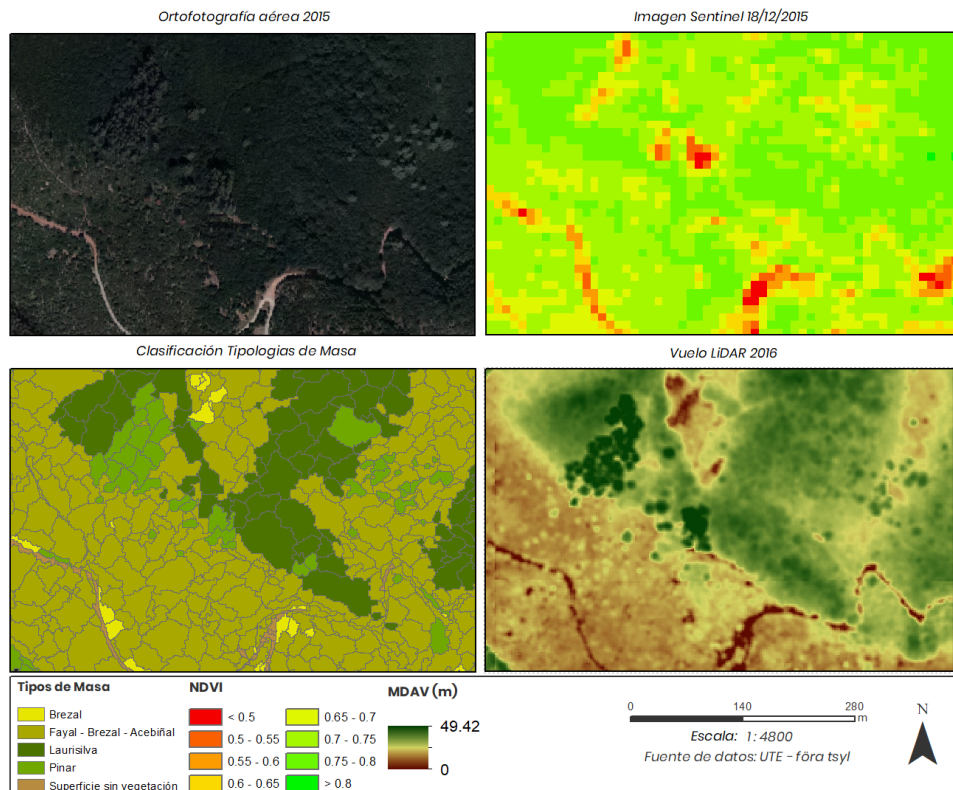
La combinación de estos datos procedentes de dos fuentes de información diferente (satélite y LiDAR), proporcionan mejores resultados que de forma individual cada uno de ellos (Koch 2011). Algunos ejemplos son la evaluación de la integración de los datos para delimitar los diferentes tipos de cobertura presentes sobre la superficie terrestre (Sasaki et al. 2012; Wang and Glennie 2015) o una vez delimitada una superficie forestal, delinear zonas homogéneas y generar mapeo de hábitats con diversos objetivos (Hatten 2014; Onojeghuo and Onojeghuo 2017), o incluso delimitar y clasificar objetos que representan directamente el individuo (Sasaki et al. 2012) (Fig. 6).



**Fig. 6:** Ejemplo de delimitación de zonas homogéneas en una superficie forestal combinando datos LiDAR e imagen multiespectral en el Parque Rural de Anaga, Tenerife, España. Fuente: (Blázquez-Casado 2017).

Identificar no sólo la localización de las masas forestales, sino identificar con precisión las especies es un paso más allá, si bien se trata de un objetivo realmente complejo. Numerosos autores han estudiado diferentes metodologías para recoger la diversidad específica mediante la combinación de datos LiDAR e imágenes de satélite en diversos ecosistemas forestales como son los bosques tropicales (Dupuya et al. 2013), bosques de sabana (Naidoo et al. 2012; Sarrazin et al. 2012), bosques boreales (Nordkvist et al. 2012; Dalponte et al. 2014) o bosques templados (Heinzel and Koch 2012; Shi et al. 2018). Los resultados de las clasificaciones llegan a superar el 80% de precisión, dependiendo de la resolución de los sensores utilizados, del número de especies a clasificar y de si la clasificación se realiza sobre árbol individual o sobre objeto. Así, la combinación de datos procedentes de sensores remotos es una herramienta clave para la localización de las masas forestales, pero también para su caracterización (Fig. 7). Entre los aspectos más estudiados están los parámetros biofísicos como la altura, el área basimétrica o la densidad, que permiten cuantificar el volumen de madera (Hudak et al. 2006; Tonolli et al. 2011). Esto sirve a posteriori para la cuantificación de otros productos forestales como biomasa,

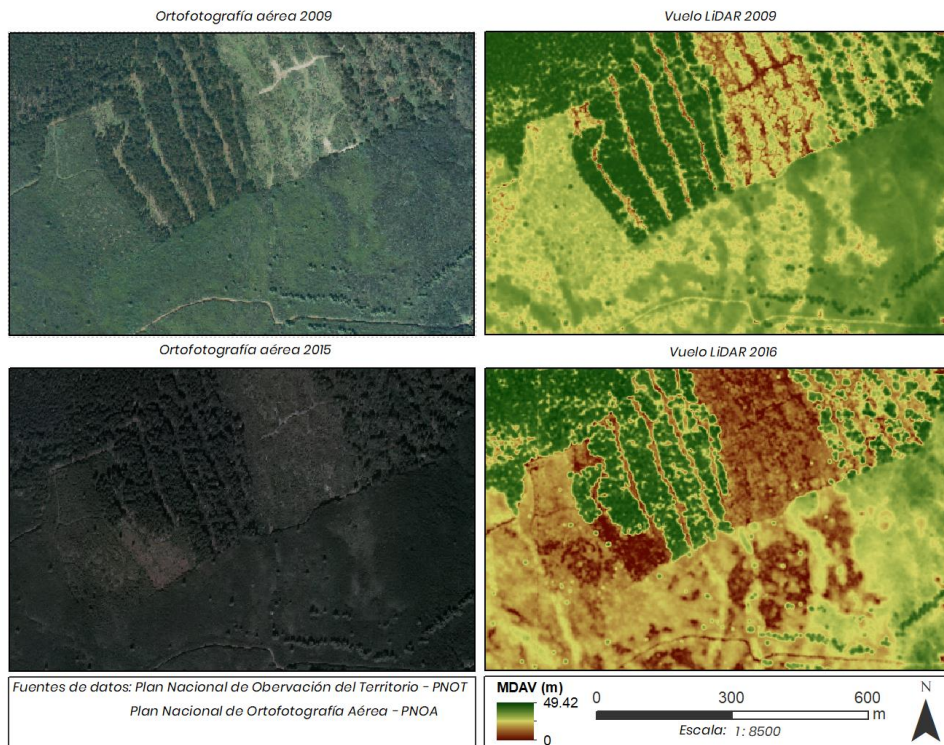
stock de carbono u otros servicios ecosistémicos (Ediriweera et al. 2014; Karna et al. 2015; Vauhkonen 2018).



**Fig. 7:** Ejemplo de clasificación de tipologías de masa con imagen de Sentinel y datos LiDAR. Fuentes de datos: Plan Nacional de Observación del Territorio (PNOT), Plan Nacional de Ortofotografía Aérea (PNOA).

Otro enfoque es el de los estudios multitemporales (Fig. 8). La disponibilidad de series temporales de datos procedentes de sensores remotos hace que sea posible evaluar la dinámica de las masas forestales y sus productos (Matasci et al. 2018; Zhao et al. 2018), aunque en el caso de datos LiDAR la disponibilidad de series temporales actualmente es complicada. Así, es habitual encontrar trabajos multitemporales donde se busca estudiar la potencialidad de la integración de las dos herramientas solamente con un vuelo LiDAR y múltiples imágenes de satélite. Por ejemplo, Badreldin and Sanchez-Azofeifa (2015), evaluaron la dinámica del recurso de biomasa para monitorear el equilibrio del

ecosistema ante el efecto del cambio climático y de las actividades antropogénicas mediante el uso de imágenes Landsat multitemporales y datos LiDAR. Sin embargo, también existen algunos trabajos donde se emplean datos multitemporales tanto LiDAR como satelitales para evaluar dinámicas o cambios sufridos en un periodo de tiempo determinado. Por ejemplo, Wulder et al. (2009) demostraron que la estructura forestal después de un incendio y los cambios absolutos y relativos sufridos en la estructura del bosque están fuertemente relacionados, siendo los cambios absolutos y relativos en la estructura de la vegetación y en la cobertura los más relacionados con los distintos índices espectrales de severidad de incendio. Así, conociendo las condiciones de desarrollo de las masas forestales, es posible predecir diferentes escenarios futuros, lo que podría servir como base de las directrices de gestión.



**Fig. 8:** Ejemplo de cambio en una masa forestal con información LiDAR multitemporal. Fuentes de datos: Plano Nacional de Observación del Territorio (PNOT), Plan Nacional de Ortofotografía Aérea (PNOA).



## 1.5 La modelización forestal integrando información procedente de sensores remotos

La modelización es la abstracción o simplificación de fenómenos que se dan en la naturaleza con el objetivo de predecir y/o explicar dicho fenómeno (Vanclay 1994). La clave de la correcta gestión sobre las masas forestales radica en el conocimiento expreso del desarrollo de las especies forestales y de su comportamiento frente a posibles cambios (Diéguez-Aranda et al. 2009). En este sentido, la generación de modelos de crecimiento y producción de los sistemas forestales ha sido una de las ramas más desarrolladas en investigación forestal desde finales del siglo XVIII, aunque como destaca Tesch (1980), el estudio del crecimiento forestal se remonta a la época de Aristóteles. Este tipo de modelos han evolucionado desde tablas que predicen variables tradicionales como la altura (Assman 1970), a modelos cuyo objetivo es la predicción de la fijación del carbono, tanto a nivel de árbol como de masa. Sin embargo, la creciente necesidad de caracterizar cambios en el espacio a gran escala ha requerido un cambio en el paradigma de la modelización, siendo posible el análisis a gran escala de estos procesos utilizando modelos controlados por sensores remotos.

Existen numerosos métodos de modelización capaces de estimar variables forestales de masa como la altura, el área basimétrica, el volumen y la biomasa en pie a partir de datos obtenidos por sensores remotos (García et al. 2018; Pearse et al. 2018). Hasta ahora, los modelos de regresión paramétricos han dominado la modelización forestal. Los modelos de regresión lineal múltiple, no lineales, lineales generalizados o mixtos son algunos de los más utilizados (Hudak et al. 2006; Popescu 2007; Sullivan et al. 2017). Sin embargo, el hecho de que las relaciones entre las variables predictoras y variable respuesta no sean siempre lineales, ni sigan una distribución normal, y la toma de datos masiva que actualmente se da mediante los sensores remotos, está haciendo que las técnicas de modelización basados en aprendizaje automático se presenten como la alternativa más eficiente en la modelización actual.

El objetivo de las técnicas de modelización no paramétricas es imitar, mediante una estructura matemática, el funcionamiento de un cerebro biológico. Son estructuras capaces de aprender a resolver problemas a partir del conocimiento extraído de una muestra de entrenamiento. Los modelos no paramétricos más usados son los vecinos más cercanos ( $k$ -NN:  $K$ -Nearest

*Neighbour*), los árboles aleatorios (RF: *Random Forest*) y la máquina de vector soporte (SVM: *Support Vector Machine*) (Xu et al. 2018). Además, existen otros modelos como las redes neuronales (ANN - *Artificial Neural Networks*) o los árboles de regresión iterados (BRT - *Boosted Regression Trees*), también aplicados en trabajos relacionados con la modelización en el sector medioambiental (Nitze et al. 2012; Youssef et al. 2016).

Los modelos paramétricos siempre parten del supuesto de normalidad, aunque muchas de las variables implicadas en la modelización forestal no los cumplan. Pero la gran ventaja de los modelos paramétricos es que la relación entre los predictores y la variable respuesta se fija a través de los coeficientes, por lo que se asume robustez en la predicción cuando son validados. Esto implica que los modelos resultantes pueden ser aplicados sobre otros contextos posteriores cuando se cumplan características similares. En cambio, los modelos no paramétricos no están sujetos al supuesto de normalidad, ni tampoco se ven afectados por problemas de colinealidad. Sin embargo, el enfoque de caja negra sobre las relaciones entre las variables hace que este tipo de modelos sean difíciles de interpretar. Por tanto, ninguna técnica se revela como claramente superior a las demás para predecir atributos de los inventarios forestales y la mejor aproximación depende de los objetivos (Brososfske et al. 2014).

## 1.6 Aproximación a distintas escalas

La recogida de información sobre la estructura forestal ha dependido históricamente del trabajo de campo, de alcance limitado cuando se trata de estudiar zonas extensas y/o remotas (Bergen and Dronova 2007). La integración del trabajo en campo con la información procedente de sensores remotos implica el cambio de escala, desde escala de monte o región a global. Para hacer efectivo este cambio de escala, existen varios sensores que aportan información espacial completa de toda la superficie global, pero con muy diversas resoluciones espaciales.

La unidad mínima de trabajo es un factor a tener en cuenta. Estudios a nivel de árbol individual o a nivel de parcela o segmento, requieren información con distinta resolución espacial. Xu et al. (2017) haciendo una estimación de la

biomasa a nivel de árbol, encontraron que la varianza residual de las ecuaciones alométricas fueron las responsables de la mayor parte de la incertidumbre, mientras que la varianza de la altura LiDAR tuvo poca contribución en el error. Posteriormente, al integrar el error a nivel de árbol individual en los análisis a escala de parcela, se encontró que los errores se reducían aún más al incrementar la escala, estando concentrados en la estimación de biomasa.

La relación a nivel de píxel o de parcela es conceptualmente similar a la relación entre árboles. La principal diferencia es que la relación se establece entre parámetros biofísicos y la variable objetivo de todos los árboles dentro del área de interés (Rodríguez-Veiga et al. 2017). Clark et al. (2005) observaron una disminución de la precisión en sus clasificaciones cuando aumentaba el tamaño del píxel de sus datos. En cambio, Ghosh et al. (2014) encontraron que sus clasificaciones globales a una resolución intermedia presentan mayor precisión que a una resolución menor o mayor, y que este efecto podría deberse a que el tamaño del píxel más pequeño recoge demasiada variabilidad espectral intraespecífica haciendo que no exista apenas separabilidad espectral entre especies. Y, por el contrario, en el caso del tamaño del píxel más grande, recoge escasa variabilidad haciendo, de igual manera, que no exista separabilidad espectral entre especies.

Es importante que exista equilibrio entre la resolución espacial de la información utilizada y la unidad mínima de trabajo. También existe relación entre la precisión de los modelos y la escala de trabajo. Kimberley et al. (2017) estudiaron la caracterización del error de predicción en función de la escala en la productividad del *Pinus radiata* D. Don. en Nueva Zelanda y concluyeron que la precisión mejora gradualmente a medida que aumenta la escala. La modelización en los trabajos a pequeña escala generalmente aporta buena precisión. Del mismo modo, trabajos a gran escala, como la regional, aportan resultados de gran calidad, ya que sus errores se ven compensados.

## 1.7 Objetivos y alcance

### Objetivo global

En la presente tesis se trata de aportar nuevos conocimientos de la potencialidad de la teledetección combinando información continua procedente de imágenes satelitales y de datos LiDAR, para desarrollar herramientas capaces de evaluar distintos recursos forestales a gran escala.

### Objetivos específicos

- i. Combinación de datos LiDAR de baja densidad e imágenes de satélite de alta resolución Plèiades para discriminar especies pie a pie en masas mixtas de bosques mediterráneos.
- ii. Comparación de diferentes metodologías de modelización de aprendizaje automático RF, SVML, SVMR y ANN (*Random Forest, Linear Support Vector Machine, Radial Support Vector Machine y Artificial Neural Networks*) para la clasificación de las distintas tipologías de masas forestales presentes en los bosques de monteverde de la isla de Tenerife en Islas Canarias (España).
- iii. Evaluación de la severidad post-incendio a través del calibrado del índice espectral RdNBR (*Relative Differenced Normalized Burn Ratio*) con datos de campo con el objeto de ser capaces de cuantificar la pérdida del recurso tras una perturbación.

## 1.8 Marco de estudio

El cuerpo de esta tesis se estructura en tres capítulos bien diferenciados (Fig. 9), en los que se aborda la evaluación de recursos forestales o, en su caso, la pérdida de estos a distintas escalas: (i) Clasificación de especies a nivel de árbol individual para la cuantificación de la producción de piñón y resina en masas mixtas de *Pinus pinea* L. y *Pinus pinaster* Ait.; (ii) Clasificación de las diferentes tipologías de masa a escala regional en bosques multifuncionales de monteverde; y (iii) Calibración de la severidad post-incendio en diferentes sucesos repartidos en el sur de Europa. En todos los casos, la información de

partida procede de diferentes sensores remotos, tanto a nivel estático como multitemporal. Además, también son utilizados diferentes métodos estadísticos para la evaluación de los recursos en cada uno de los trabajos.

En el capítulo primero, se realizó un estudio para caracterizar con precisión, a nivel de árbol individual, el número de pies de cada especie presentes en bosques mixtos de *Pinus pinea* L. y *Pinus pinaster* Ait. de la meseta norte española. Los modelos fueron entrenados con masas puras, con el objetivo de cuantificar, con la mayor precisión posible, la producción de piña y resina en este tipo de masas mixtas. Para ello, se hizo una clasificación basada en datos LiDAR, a nivel de árbol individual, procedentes del vuelo PNOA 2010 de baja densidad e imágenes Plèiades de alta resolución espacial tomadas en el año 2014.

En el segundo capítulo, se identificaron las diferentes tipologías de masa presentes en los bosques de monteverde con cobertura completa, elevada densidad y continuidad vertical de la isla de Tenerife. El objetivo fue caracterizar la diversidad de estos mediante una comparación de cuatro métodos de clasificación para identificar las diferentes tipologías de masa. Para ello, se tomó como base de datos de partida la información multitemporal, tanto LiDAR a nivel de masa (dos vuelos en 2009 y 2016) como espectral mediante imágenes de Sentinel (dos imágenes en los años 2015 y 2017) con resolución espacial media.

En el tercer capítulo, se evaluó la capacidad de determinar la severidad tras el paso de un incendio. El estudio ha sido desarrollado a partir de datos tomados en 28 grandes incendios repartidos por tres países del sur de Europa con el objetivo de considerar la pérdida de biomasa. Para ello, se calibró el índice espectral RdNBR a través de datos de afectación tomados en campo mediante dos métodos estadísticos distintos, tomando como base de datos el repositorio multitemporal de imágenes Landsat con baja resolución espacial.

Además, en esta tesis se han utilizado distintos métodos estadísticos, tanto modelos clásicos paramétricos como no paramétricos, para estimar las relaciones existentes entre los datos tomados en campo y las diferentes variables procedentes de los sensores remotos. En el capítulo primero fue

utilizado *Random Forest* para realizar la clasificación de las dos especies objeto de estudio. En el segundo capítulo, se enfrentaron cuatro modelos no paramétricos como son *Random Forest*, *Linear Support Vector Machine*, *Radial Support Vector Machine* y *Artificial Neural Networks* para la clasificación de las diferentes tipologías de masa presentes en los bosques de monteverde. Por último, en el capítulo tercero se evaluó la capacidad de predecir la severidad de un incendio y, por tanto, el nivel de afectación en el recurso de biomasa mediante dos métodos estadísticos, uno lineal (modelos mixtos) y otro no lineal (*Boosted Regression Trees*, BRT).

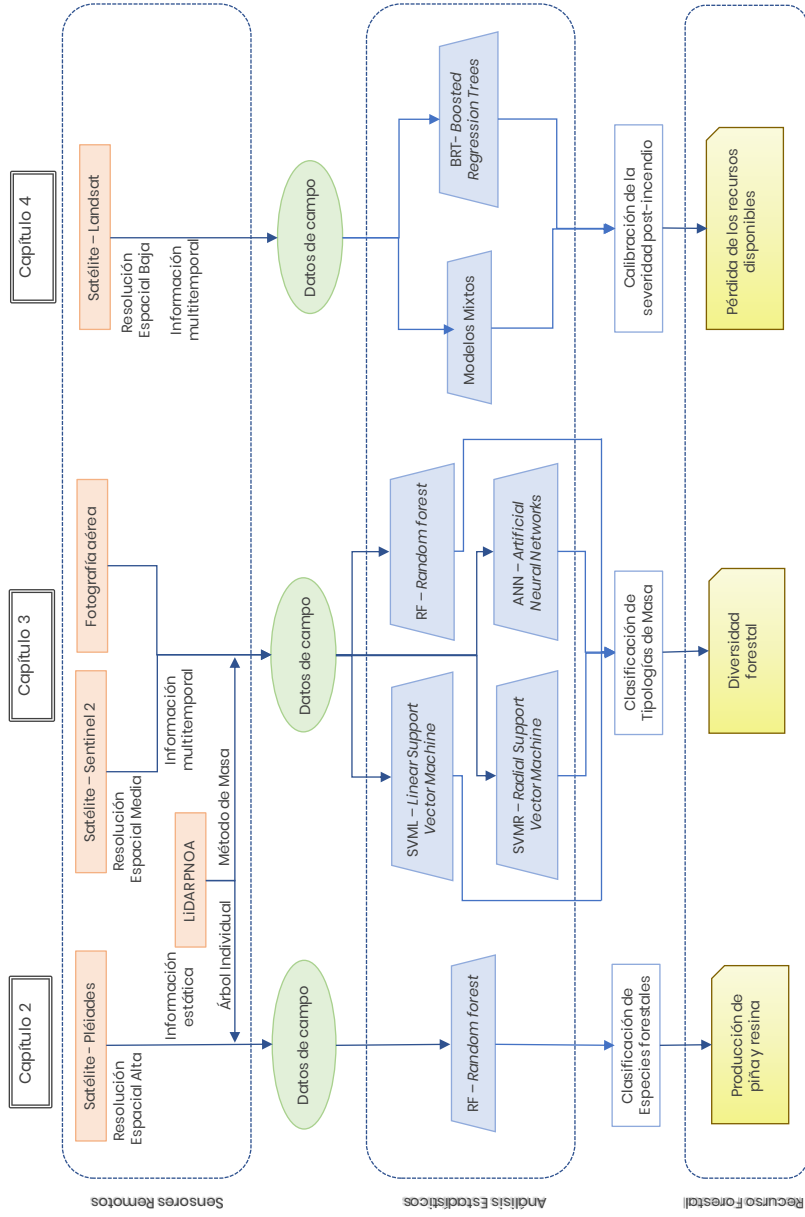


Fig. 9: Esquema general del cuerpo de la tesis.









## CHAPTER 2

---

Combining low density LiDAR and satellite images to discriminate species in mixed Mediterranean forest

*“En la naturaleza, nada y todo es perfecto.  
Los arboles pueden ser contorsionados,  
doblados de maneras raras,  
y aun así son hermosos”*

Alice Walker

## **Chapter 2. Combining low density LiDAR and satellite images to discriminate species in mixed Mediterranean forest**

---

### **Abstract**

The discrimination of tree species at individual level in mixed Mediterranean forest based on remote sensing is a field which has gained greater importance. In those stands, the capacity to predict the quality and quantity of non-wood forest products is particularly important due to the very different goods the two species produce. To assess the potential of using low density airborne LiDAR data combined with high resolution Pleiades images to discriminate two different pine species in mixed Mediterranean forest (*Pinus pinea* L. and *Pinus pinaster* Ait.) at individual tree level.

In a first stage, a Random Forest model was trained using plots from the pure stand dataset, determining which LiDAR and satellite variables allow us to obtain better discrimination between groups. The model constructed was then validated by classifying individuals in an independent set of pure and mixed stands. The model combining LiDAR and Pleiades data provided greater accuracy (83.3% and 63% in pure and mixed stand validation, respectively) than the models which only use one type of covariables. The automatic crown delineation tool developed allows two very similar species in mixed Mediterranean conifer forest to be discriminated using continuous spatial information on surface: Pleiades images and open source LiDAR data. This approach is easily applicable over large areas, enhancing the economic value of non-wood forest products and aiding forest managers to accurately predict this production.

## 2.1 Introduction

Forest management is currently undergoing important changes and challenges with the up-to-date development and use of new tools built up based on remote sensing technology. In general, field data have always been gathered manually, but nowadays, the application of new technology in forest inventories has revolutionized this field and has brought down the costs associated with forest inventories (Barrett et al. 2016). These new technologies have the greatest potential for evaluating forest characteristics, given that canopy vegetation height, one of the key data on forest resources evaluation, is a function of species composition, site quality, and age; hence this data can be used for land cover classification, habitat mapping, and forest management (Dubayah and Drake 2000). Additionally, forest parameters related to tree height and stem volume at both, plot and individual tree level, can be estimated using the information provided by remote sensors (Popescu et al. 2003; Hollaus et al. 2009; Tonolli et al. 2011; Ruiz et al. 2016; Castaño-Díaz et al. 2017; Arias-Rodil et al. 2018).

An increasing amount of evidences suggest that mixed-species forests, in comparison with monospecific ones, provide a greater number of ecological and socio-economic goods and services, even in Mediterranean ecosystems (Barba et al. 2013; Riofrío et al. 2017). Consequently, the management of these forests should be based on a comprehensive assessment of their wide diversity and multifunctionality, as well as on accurate species classification. Furthermore, when the management of these stands is aimed at individual tree level rather than stand level, species discrimination is essential in order to provide an inventory for the different resources. The use of remote sensing techniques offers the advantage that such questions can be addressed at large scales in mixed forests, including at landscape level.

The high economic and social importance of non-wood forest products (NWFP) is one of the main factors that distinguishes Mediterranean forests from other temperate forests (Calama et al. 2008). The production of Stone pine nuts from *Pinus pinea* L. and resin from *Pinus pinaster* Ait. in the Mediterranean forests of the Northern Plateau of the Iberian Peninsula has made an important contribution to rural activity since ancient times (Lopez-García 1980; Nanos et al.

2000). In many Stone pine forests, pine nuts are more profitable than timber, with prices reaching over 100 €/kg, in turn favouring the NWFP transformation industry. In the case of resin, tapping is an activity which is currently undergoing resurgence. It was very important until the end of the 1970s, when the increase in the labour costs and international competitiveness of Chinese natural resins tapped from *Pinus massoniana* Lamb. rendered this traditional activity no longer profitable (Soliño et al. 2018). In the recent years, within the context of the global financial crisis resulting in higher rates of rural unemployment in Spain, together with the drop of Chinese resin exportations, and an increasing demand for high-quality secondary products for cosmetics, pharmacy and healthcare, resin tapping has become again a profitable activity, and many of the abandoned stands are being tapped again (Rodríguez-García et al. 2015; Soliño et al. 2018). *Pinus pinea* L. and *Pinus pinaster* Ait. share autoecological conditions (Gordo et al. 2012), so pine nuts and resin are both harvested in Mediterranean mixed pine-forests. Hence, the accurate assessment of nut and resin yields is particularly important, not only for the forest management decision-making process but also as regards to the associated industrial process and evaluation of market prices.

Classification of tree species assisted by remote sensing data is motivated by a wide variety of applications for the sustainable forest management of wood resources and ecosystem services (European Environmental Agency 2006; Fassnacht et al. 2016). Thus, to know the tree species amount in mixed stands has the highest importance, due to the fact that different species produce very different NWFP and, therefore, determine the forest management. Several studies have used remote sensing data, Light Detection and Ranging (LiDAR) and multispectral images to predict stand classification (Korpela et al. 2010; Ørka et al. 2012; Ballanti et al. 2016), for both species groups and individual tree species (Suratno et al. 2009a; Deng et al. 2016). Distinguishing between two species of the same genus is especially difficult, since the differences between them use to be subtle. Species that show real-world differences in physical and/or spectral properties should be clearly recorded by the sensors (Vauhkonen et al. 2014; Fassnacht et al. 2016). However, when the studied species are similar, slight differences in crown structure (e.g. umbrella and apical crowns in *P. pinea* and *P. pinaster* respectively) and stem slenderness could be key to discriminate among them, while spectral information concerning chlorophyll contents should help in this work.

The combined use of LiDAR and satellite images results in a much more complicated process than if LiDAR or satellite images are separately used. Besides, when the forest studied contains several species, each of them with their own differential structural characteristics, the challenge is yet bigger (Vauhkonen et al. 2014). In such stands, where species must be correctly classified, additional complexity is involved in the modelling process (crown individualization) if low-point density (<1 point per m<sup>2</sup>) LiDAR data is used (Suratno et al. 2009b; Dalponte et al. 2012). Remote sensing images have been widely used for tree species classification in order to obtain spatially detailed species information (Xie et al. 2008; Dalponte et al. 2012). New satellite images with high spatial resolution are available today, such as the 50 cm images provided by the Pleiades constellation, which have led to improved analyses. Moreover, Pleiades images have recently been used in studies focusing on forest biomass modelling and forest structure mapping through image texture analysis (Beguet et al. 2014; Rougier and Anne 2014; Maack et al. 2015).

Current estimation and forecasting of cone and resin production is based on individual tree-level models (Nanos et al. 2001; Moreno-Fernández et al. 2013; Calama et al. 2016), thus sound estimates should be based on previous correct assessment of tree level features as species and size. Individual tree delineation from LiDAR data can be used to define individual structure for many wood species (Popescu et al. 2003; Valbuena et al. 2016a). The individual crown delineation allows us to obtain tree-level variables from LiDAR variables which are enclosed in the algorithms for species classification. Therefore, these tree-level LiDAR variables might also be applied to predict individual tree NWFP production by means of the currently existing models, which depends on the tree as an individual entity (Popescu 2007). Furthermore, species discrimination is often determined satisfactorily using satellite data. Hence, the operational process in pine stands should consist on first carrying a process of crown delineation using LiDAR data, then applying classification models to identify species and extract tree-level attributes, and subsequently implementing pre-existing production models for each NWFP. Thus, the integration of the methods should permit assessing NWFP in large areas (Lee et al. 2016).

The aim of our study is to analyse the potential of using low density LiDAR data combined with high resolution Pleiades images to discriminate different Mediterranean pine species in pure and mixed *Pinus pinaster* Ait. and *Pinus pinea*

L. forests, based on an automatic algorithm for crown delineation. The practical motivation is that by distinguishing these two species at a tree scale will permit us to apply the pre-existing models for different non-wood forest products (resin or pine nuts) which act at a tree level scale (e.g. Calama et al. 2016; Nanos et al. 2004). For this purpose, (i) in a first phase, a Random Forest model was trained with a dataset collected in pure stands of both species, to determine which LiDAR and satellite variables allow us to obtain the best discrimination between groups; and (ii) the Random Forest constructed was then validated by classifying individuals into an independent set of mixed and pure stands. Our hypotheses are: (i) combining LiDAR and Pleiades data to discriminate species in mixed Mediterranean conifer forests should be more efficient than the use of each technique separately; (ii) this combination should provide an efficient tool to discriminate pine species in mixed Mediterranean forests at individual tree level, and (iii) the results for *Pinus pinea* L. should be more accurate than those for *Pinus pinaster* Ait. due to the fact that the former usually grows in open stands.

## 2.2 Material and methods

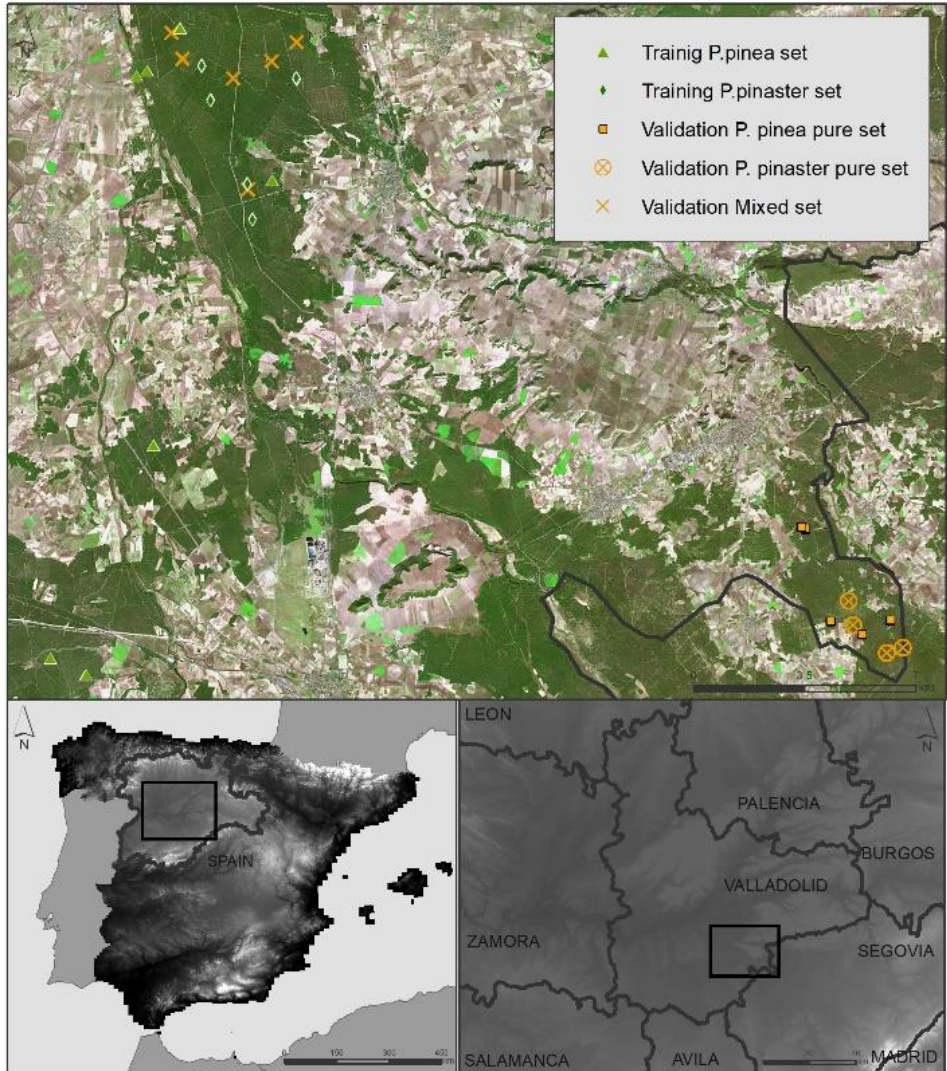
### 2.2.1 Study area

Forested land accounts for almost a third of the total land area of Spain (18.27 million ha). Pure forests of *Pinus pinea* L. cover 401,701 ha while *Pinus pinaster* Ait. forests make up 1,131,901 ha (MAGRAMA 2012). The study area is situated in the Northern Plateau of the Iberian Peninsula (Fig.1), a region with 65,275 ha of pure *Pinus pinea* L. forests, 271,000 ha of pure *Pinus pinaster* Ait. forests and 53,000 ha of mixed *Pinus pinea* L. and *Pinus pinaster* Ait. forests (López et al. 2009). The Duero river basin is on the Northern Plateau, 700–800 m a.s.l., with a homogenous Mediterranean continental climate, characterized by low average annual rainfall (450 mm), strong summer drought (<50 mm) and low average annual temperatures (12 °C). Soils mainly present a very high sand content (> 90%) and very low water holding capacity (WHC < 100 mm), except in the upper areas where limestone soils show a percentage of clays and limes over 40–50%, reaching WHC values > 250 mm.

The plots within the study area were clustered in three different forest types: pure stands of *Pinus pinea* L., pure stands of *Pinus pinaster* Ait., and mixed stands, where the least represented species should account for at least 35% of the total



number of trees. All considered stands were at mature development stage, where regeneration is quite poor.



**Fig.1** Study area. Locations of training and validation sets.

### 2.2.2 Field data

In this study we used a subset from the network of permanent plots installed by the Forest Research Centre of the Spanish National Institute for Agricultural Research (INIA-CIFOR) in the province of Valladolid. Mixed and pure *Pinus pinaster* Ait. plots were installed between summer of 2015 and spring of 2016. Pure even aged *Pinus pinea* L. plots were installed in 1996 in cooperation with the regional forest service and have been regularly monitored since then (the most recent inventory was carried out in 2016). Plots were selected in order to cover the whole range of site conditions, stand stocking and age identified within the region. The plots were all located in public mature forests, and at installation, the stand should have not been altered (thinnings, final cuttings or pruning) during at least the previous five years.

The plots were circular in shape, with variable radius. Pure plots included at least 20 trees and mixed plots included at least 10 trees of the least represented species. Plot centre coordinates were recorded using a high precision submeter positioning GPS, with an accuracy below 1 m. Polar coordinates (x,y) of each tree within the plot were computed after measuring their distance and heading from the centre of the plot using a VERTEX IV distance meter and a compass. Trees were then numbered and measured. Tree measurements included species, diameter at breast height (1.3 m) (DBH, cm), total height (Height, m), crown base height (Live\_Branch, m) and crown radius (Crown\_Radio, m). Afterwards, each tree measured during the field work was plotted on a map. From the whole available network of plots, the less represented conditions of composition were both pure *Pinus pinaster* Ait. plots and mixed P. pinaster – P. pinea ones. Thus, all the available plots from pure *Pinus pinaster* Ait. stands and mixed stands were included in the sample. In a second phase, we selected all the available pure *Pinus pinea* L. plots located in the vicinity of the pure P. pinaster and mixed plots. At the end, twenty-seven plots were selected for the present study (Fig.1): nine pure *Pinus pinaster* Ait. plots, six mixed plots and twelve pure *Pinus pinea* L. plots. Of these, seven plots in pure *Pinus pinea* L. stands and five plots in pure *Pinus pinaster* Ait. stands were randomly selected and used for training the model. The remaining fifteen plots – five in pure *Pinus pinea* L. stands, four in pure *Pinus pinaster* Ait. stands and six in mixed stands– were therefore used to validate the model. All individual trees measured inside each selected plot were enclosed in the sample.

### 2.2.3 Remote sensing data

Low-density LiDAR data for the study area was provided by the Spanish Geographic Institute (PNOA 2018). Point clouds were captured in 2010 using a laser scanner, Leica ALS60 sensor, with a mean density of 0.5 points per m<sup>2</sup> and a root mean square error of 20 cm in Z altimetric accuracy. Digital files, classified and coloured, were downloaded in .laz format (2x2 km).

Pleiades is a constellation of the National Centre for Space Studies (CNES, French Government) which provide very-high-resolution images. It was designed by the Optical and Radar Federated Earth Observation program (ORFEO 2018) and was launched in October 2003. Pleiades 1A and Pleiades 1B operate as a constellation in the same orbit, phased 180° apart. The identical twin satellites capture images which are transferred, orthorectified and geometrically corrected by the Spanish Remote Sensing National Plan (PNT 2018). Images were captured in June–July 2014 and each one covers an area of 400 km<sup>2</sup>. The Pleiades images comprise four multispectral bands with a spatial resolution of 2 m. The spectral ranges of the four bands are 430–550 nm (Blue), 490–610 nm (Green), 600–720 nm (Red) and 750–950 nm (Infrared). Additionally, there is a panchromatic band (480–830 nm) with a spatial resolution of 0.5 m.

The time delay in the field calibration of the remote sensing data do not increase the level of uncertainty since in these mature stands large changes in such short lapse of time are not expected, particularly if there is not any disturbance. Low productivity in these inner Mediterranean stands (mean annual volume increment about 1 m<sup>3</sup> ha<sup>-1</sup>) confirms this issue. In addition, as field data came from a net of permanent plots, we can confirm little change on time in the structure (number of stems per ha, stand composition) in the last years.

### 2.2.4 Remotely sensed data processing

#### 2.2.4.1 LiDAR data processing

LiDAR point files (in .laz format), which include the field plots, were compiled and normalized, subtracting their corresponding ground class points. Points below 3 m were eliminated as it was considered that this was the lowest level at which a distinction between bushes and trees could be drawn. Points over 50 m were considered 'mistakes', as there are no trees above this height in the study area,

hence, they were also removed. The algorithm for tree delineation is the one by Valbuena et al. (2016) written as a function in a SQL language for use in PostgreSQL 9.2 with the POSTGIS 2.0 add-on, which gave the database manager geometrical and other GIS functions. It was applied to each normalised point cloud in .shp format. The algorithm is based on cloud point LiDAR analysis, searching directly for relative maximums. Once a relative maximum has been located, the next lower LiDAR return is found, and the distance between them is calculated. The algorithm determines whether this distance implies that the point is a new apex or whether it forms part of the first crown found, thus allowing individual tree crowns to be delineated. To reach each crown, the algorithm divides the point cloud into layers half a metre deep and analyses the distance from each point to the crown polygons already delineated in each layer, in a downward direction (Valbuena et al. 2016a). This process delineates the horizontal projection of each crown as a polygon that takes in all the points which impact on each tree, identifying individual tree crowns drawing them as vector format (Fig.2). The result takes the form of a .shp layer that contains each crown polygon for each tree in graphic form and an associated table with the following data for each crown: maximum height of tree crown (Max\_Height, m), minimum height of tree crown (Min\_Height, m), crown area of individual tree (Crown\_Area, m<sup>2</sup>), crown height of individual tree (Crown\_Height, m), understood as the difference between Max\_Height and Min\_Height. Density (Density, trees per ha) was calculated from the centroid of the polygon using the 6-tree sampling method (Prodan 1968).



**Fig.2** Crown delineation in a mixed stand using the Valbuena et al. (2016) algorithm.

In order to characterize the structural differences between *Pinus pinea* L. and *Pinus pinaster* Ait., three relations were derived from the previous variables: Heights relation (HR), calculated by the relationship between maximum height and minimum height ( $((\text{Max\_Height} - \text{Min\_Height}) / \text{Max\_Height})$ ); Area and height relation (AHR), estimated as the quotient between crown area and maximum tree height ( $\text{Crown\_Area} / \text{Max\_Height}$ ), it reveals the relationship between *Pinus pinea* L. and *Pinus pinaster* Ait. outline crowns, something each species is usually characterized by; and Crown and Height relation (CHR), calculated as the relationship between maximum height and crown base height ( $((\text{Max\_Height} - \text{Crown\_height}) / \text{Max\_Height})$ ), it expresses the relationship between height tree and natural pruning height, which is generally higher in *Pinus pinea* L. than in *Pinus pinaster* Ait. Thus, six LiDAR variables got into modelling: Crown\_Area, Crown\_Height, Density, HR, AHR, and CHR (Table 1 and Table 2).

**Table 1.** Descriptive statistics of training dataset. Pp\_p: *Pinus pinea* L. individual trees in pure stand plots, Pt\_p: *Pinus pinaster* Ait. individual trees in pure stand plots.

	Minimum		Maximum		Mean		Standard Deviation	
	Pp_p	Pt_p	Pp_p	Pt_p	Pp_p	Pt_p	Pp_p	Pt_p
<b>FIELD DATA</b>								
DBH (cm)	17.7	28.3	63.7	48.0	35.9	36.6	10.7	5.2
Height (m)	6.7	11.7	15.4	18.8	10.7	15.0	2.5	1.4
Live_Branch (m)	2.0	0.0	9.0	12.7	4.9	7.8	2.0	2.5
Crown_Radius (m)	1.7	0.9	5.8	3.9	3.3	2.5	0.9	0.6
<b>LIDAR DATA</b>								
Max_Height (m)	6.4	9.3	14.4	16.1	9.9	13.5	2.4	1.5
Min_Height (m)	3.0	3.0	11.5	12.0	6.3	8.0	2.1	2.2
Crown_Area (m <sup>2</sup> )	16.9	19.5	80.1	100.8	34.9	30.0	17.7	12.7
Crown_Height (m)	0.8	0.9	10.2	11.2	3.7	5.5	1.7	2.1
Density (trees/ha)	55.0	63.0	237.0	216.0	114.3	107.5	42.7	24.1
HR	0.12	0.08	0.77	0.78	0.37	0.41	0.13	0.15
AHR	1.69	1.34	7.20	7.04	3.46	2.23	1.34	0.88
CHR	-0.86	-0.86	-0.41	-0.26	-0.67	-0.51	0.10	0.13
<b>SATELLITE DATA</b>								
Blue	96.4	93.2	123.2	129.4	107.6	107.6	7.0	6.9
Green	120.6	116.2	146.0	141.0	130.4	127.2	5.2	5.0
Red	135.4	133.6	155.1	156.7	143.6	143.2	4.0	4.3
Infrared	246.6	220.3	306.4	278.1	269.2	249.1	13.3	13.7
NDVI	0.26	0.21	0.35	0.31	0.30	0.27	0.02	0.02
EVI	0.42	0.34	0.59	0.53	0.51	0.45	0.04	0.05

**Table 2.** Descriptive statistics of validation dataset. Pp\_p: *Pinus pinea* L. individual trees in pure stand plots, Pt\_p: *Pinus pinaster* Ait. individual trees in pure stand plots, Pp\_m: *Pinus pinea* L. individual trees in mixed stand plots, Pt\_m: *Pinus pinaster* Ait. individual trees in mixed stand plots.

	Minimum				Maximum				Mean				Standard Deviation			
	Pp_p	Pt_p	Pp_m	Pt_m	Pp_p	Pt_p	Pp_m	Pt_m	Pp_p	Pt_p	Pp_m	Pt_m	Pp_p	Pt_p	Pp_m	Pt_m
<b>FIELD DATA</b>																
DBH (cm)	34.6	27.2	26.9	22.5	71.0	58.6	57.2	52.6	44.6	41.0	40.0	32.7	7.1	6.9	7.5	6.2
Height (m)	12.4	11.5	9.6	11.0	19.9	21.6	16.7	16.9	15.0	16.2	12.6	13.7	1.7	2.8	1.6	1.6
Live_Branch (m)	3.8	2.6	2.4	2.7	9.8	14.2	9.1	11.3	6.4	8.6	5.4	7.9	1.7	3.6	1.5	1.9
Crown_Radius (m)	2.6	1.8	2.1	1.4	6.0	3.7	5.9	3.7	3.9	2.7	3.8	2.5	0.7	0.5	0.8	0.6
<b>LIDAR DATA</b>																
Max_Height (m)	11.1	7.7	6.8	10.0	18.4	18.0	16.0	15.6	14.0	13.7	12.0	12.7	1.7	2.8	1.6	1.3
Min_Height (m)	2.1	2.2	3.1	4.2	13.3	13.8	12.2	10.8	8.4	7.8	7.4	7.4	2.6	2.5	1.7	1.7
Crown_Area (m <sup>2</sup> )	23.5	12.5	19.5	19.5	87.7	75.9	105.2	107.0	42.7	26.0	38.6	36.2	16.9	14.9	18.5	17.7
Crown_Height (m)	2.3	1.2	0.5	1.6	11.9	14.9	11.2	10.3	5.5	5.9	4.6	5.3	2.1	3.1	1.7	1.8
Density (trees/ha)	41.0	60.0	48.0	68.0	227.0	179.0	206.0	197.0	106.6	113.8	108.5	122.7	44.6	28.9	36.4	42.0
HR	0.2	0.1	0.08	0.15	0.8	0.9	0.78	0.70	0.4	0.4	0.38	0.41	0.2	0.2	0.13	0.13
AHR	1.7	0.9	1.53	1.53	6.1	5.8	9.01	9.62	3.0	1.9	3.23	2.87	1.0	1.0	1.53	1.51
CHR	-5.1	-4.8	0.22	0.30	-0.7	0.1	0.92	0.85	-2.0	-0.9	0.62	0.59	1.0	1.0	0.13	0.13
<b>SATELLITE DATA</b>																
Blue	65.6	54.5	90.6	92.0	87.9	99.4	120.5	113.6	74.7	72.3	102.2	103.0	5.1	11.6	18.7	6.7
Green	86.3	74.4	113.5	117.0	107.3	115.2	133.5	131.5	96.8	93.4	123.7	124.8	4.7	10.5	17.4	4.1
Red	102.1	96.7	132.1	132.4	121.5	135.5	148.4	147.1	111.2	109.7	139.8	140.3	4.0	9.8	17.0	3.8
Infrared	199.3	136.9	201.1	228.9	270.4	269.9	287.5	282.7	238.1	208.1	253.0	252.5	17.2	34.7	29.6	15.7
NDVI	0.27	0.17	0.20	0.22	0.43	0.38	0.36	0.34	0.36	0.30	0.29	0.28	0.03	0.05	0.04	0.03
EVI	0.45	0.27	0.32	0.36	0.76	0.65	0.61	0.59	0.63	0.51	0.48	0.47	0.06	0.09	0.08	0.06

### 2.2.4.2 Tree crown selection

Following the process of individual tree crown delineation, a subset of trees with clearly identified crowns and which had been matched with trees in the field inventory of the plot were selected for the analysis. Selected training data in pure stands is more reliable in relation to mixed ones, due to the extreme importance of tree position's high precision at individual tree level. Once the delineation algorithm generated the crowns, both shapes were matched one to one automatically though QGIS intersect tool (QGIS Development Team 2017), throwing away crowns when more than one measured tree coincided on one delineated crown to guarantee that each crown has been issued to coincide with one measured tree. 95.5% and 81.0% of trees measured has respectively overlapped with crown delineation data in *P. pinea* and *P. pinaster* pure stands. On the other hand, overlapping between field data measured and delineated crowns in mixed stands was 126.0% due to an overestimation of delineated crowns. Thus, the algorithm is properly applied to this kind of stands. This resulted in 42 *Pinus pinea* L. trees and 76 *Pinus pinaster* Ait. trees from the pure forest plots to be used for training the model (Table 1), while 82 *Pinus pinea* L. (36 from pure plots and 46 from mixed plots) and 71 *Pinus pinaster* Ait. (36 from pure plots and 35 from mixed stands) were selected for the validation process (Table 2).

### 2.2.4.3 Plèiades data processing

Firstly, 0.5 m high resolution multispectral images were created from the pansharpening algorithm of high-resolution panchromatic and lower resolution multispectral imagery. Based on the previous individual crown delineation, variables derived from Pleiades images were estimated for each object. Pleiades reflectance information was assigned to crown delineation with the mean values of all pixels within each crown, estimating Blue band value, Green band value, Red band value, and Infrared band value. Data from each band was assigned to the generated objects from the crown delineation using QGIS Zonal Statistics Plugin.

In addition, different image bands were used to calculate two spectral indices: NDVI Normalized Difference Vegetation Index (Rouse et al. 1973) and EVI Enhanced Vegetation Index (Liu and Huete 1995). NDVI was selected as it is the most commonly used vegetation indicator, capable of assessing the quantity, quality,

and developmental stage of vegetation. It was calculated using Pleiades bands following the equation:  $[(\text{Infrared} - \text{Red}) / (\text{Infrared} + \text{Red})]$ , where infrared and red refers to the recorded Pleiades values in the infrared and red bands, respectively. EVI is an optimized vegetation index designed to enhance the vegetation signal with improved sensitivity which reduces the adverse effects of environmental factors such as atmospheric conditions and soil background. This was calculated using the equation:  $[2.5 * (\text{Infrared} - \text{Red}) / (\text{Infrared} + 2.4 * \text{Red} + 1)]$ . Statistics for the training and validation dataset are also shown in [Table 1](#) and [Table 2](#). The six variables from satellite data that were finally available for modelling were: Blue band, Green band, Red band, Infrared band, NDVI index and EVI index.

### 2.2.5 Classification technique

Random Forest classifier ([Breiman et al. 2001](#)) is a machine learning methodology based on independent decision trees providing diverse ways to explore numerically and graphically complex relationships, improving the accuracy of the model prediction ([Valbuena et al. 2016b](#); [Vega Isuhuaylas et al. 2018](#)). It averages several decision trees, so, there is a significantly lower risk of overfitting and it also reduces the chance of stumbling across a classifier that does not perform well. A large number of classification trees are produced from a random subset of training data, with permutations introduced at each node, selecting the most common classification result. In this study, Random Forest was used to classify species based on LiDAR and satellite data at individual tree level.

R software ([R Core Team 2016](#)) was used to generate the model, applying Random Forest library as is common practice in forest science. Three parameters are needed to optimize the model: (i) *mtry*, number of variables randomly sampled as candidates at each split; (ii) *ntree*, number of bootstrap replicates; and (iii) *nodesize*, minimum size of terminal nodes. To identify those parameter values leading to the best species discrimination, the model was optimized based on out-of-bag error estimates (OOB), defined as the rate of classification error estimated for the different subsampling sets from the training set used to train the model classifier. In addition, random forest algorithm can estimate importance of the variable in every model, showing the mean decrease of accuracy, which is defined as the loss of accuracy measured by the OOB-error



when leaving out a variable (Breiman et al. 2001). Higher values of those statistics indicate higher importance of each variable for the classification.

Three training models were developed. First, the 'LiDAR model' was constructed only using LiDAR variables; then the 'Spectral model' was built including only satellite variables; and finally, all available variables were incorporated into the so-called 'Complete model'. Each of them was designed using the same approach, attempting to identify the best combination of the three Random Forest parameters and they were evaluated based on out-of-bag error estimation. The training model which provided the greatest predictive capacity was then validated in terms of classification accuracy on the three independent validation datasets: (i) individual trees in pure *Pinus pinea* L. stand plots, (ii) individual trees in pure *Pinus pinaster* Ait. stand plots, and (iii) individual trees in *Pinus pinea* L. and *Pinus pinaster* Ait. mixed stand plots. Validation was done by: (i) building confusion matrices which show the relationship between the false positive fraction and true negative fraction (or vice versa), (ii) estimating overall accuracy (OA) as  $(\text{Number of correctly classified } Pinus\ pinea\ L. + \text{ number of correctly classified } Pinus\ pinaster\ Ait.) / \text{ total number of trees used in the validation}$ , and (iii) calculating receiver operating characteristic (ROC) area under the curve (AUC) in order to quantify the uncertainty in models' prediction (Zipkin et al. 2012). ROC curve is a plot of the sensitivity of a diagnostic test against one minus its specificity, as the cut-off criterion for indicating a positive test is varied (Everitt and Skrondal 2010). As discussed in Yesilnacar (2005), the quantitative-qualitative relationship between the AUC and prediction accuracy can be classified as follows: 50-60% (poor), 60-70% (average), 70-80% (good), 80-90% (very good), and 90-100% (excellent).

## 2.3 Results

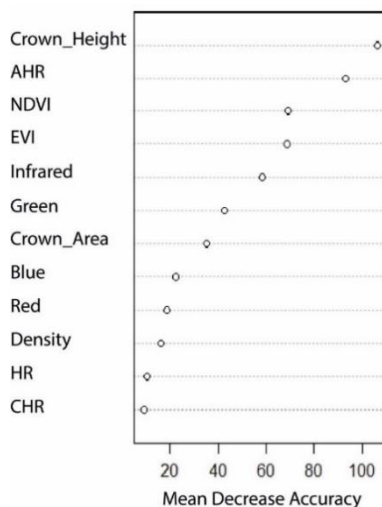
### 2.3.1 Training model selection

To achieve the best discrimination between the two pine species, *Pinus pinea* L. and *Pinus pinaster* Ait., three training models were developed based on data from individual trees in pure stand plots. The 'Spectral model' is less accurate than the 'LiDAR model' (Table 3), so species discrimination based on the latter is better. However, the 'Complete model', which combines LiDAR and satellite variables provided the lowest out-of-bag error estimates.

**Table 3.** Accuracy of the three training models in classifying *Pinus pinea* L. and *Pinus pinaster* Ait. in pure stands.

Training Models	N° of variables	OOB %
<b>LiDAR model</b>	6	13.56
<b>Spectral model</b>	6	20.34
<b>Complete model</b>	12	6.78

Thus, the model selected was the one combining LiDAR and Pleiades data since it provides the greatest accuracy. The ‘Complete model’ possesses a high degree of predictive and explanatory power, correctly classifying 93.22% of cases (i.e. OOB estimate of error rate was 6.78%). In other words, a high proportion of *Pinus pinea* L. and *Pinus pinaster* Ait. trees were correctly identified. Fig.3 shows the relative importance of each variable in the selected model, as provided by Random Forest, revealing which variables from LiDAR and satellite data are the most significant in discriminating between species. The most important variables defining the characterisation of the composition are: Crown\_Height, AHR (area and height relation), NDVI index, EVI index, Infrared band, Green band, Crown\_Area, Blue band, Red band, Density, HR (Heights relation), and CHR (Crown and height relation). The first three variables are highly related to differences in crown and stem shape between the two considered species.



**Fig.3** Mean estimated variable importance in the Random Forest ‘Complete model’. The maximum value corresponds to the main discriminating variable and the rest are presented in relation to this score.

### 2.3.2 Validation model

The validation process was carried out on an independent sample of pure and mixed plots. Table 4 shows the confusion matrix obtained for the selected 'Complete model' for the validation on pure datasets. In the case of *Pinus pinea* L. trees, the model correctly classified 72.2% of individuals. In contrast, the level of efficiency obtained in the case of *Pinus pinaster* Ait. was 94.4%. The overall accuracy of the model when used for plots in pure stands was 83.3%, a mid-precision value between the results observed in the training model and those from the validation performed on mixed datasets. The results indicate that *Pinus pinea* L. is well-discriminated although sub-estimated, because *Pinus pinea* L. are sometimes classified as *Pinus pinaster* Ait., whereas the opposite does not occur. Hence, the 'Complete model' achieves a high-confidence classification of both *Pinus pinea* L. and *Pinus pinaster* Ait. trees. Also, the area under the curve (AUC) value also indicates that the model classifier is discriminating between both species with an average rate of 89.6% as shown in Fig.4a. Thus, the model applied in pure stands showed reasonably very good accuracy in its prediction.

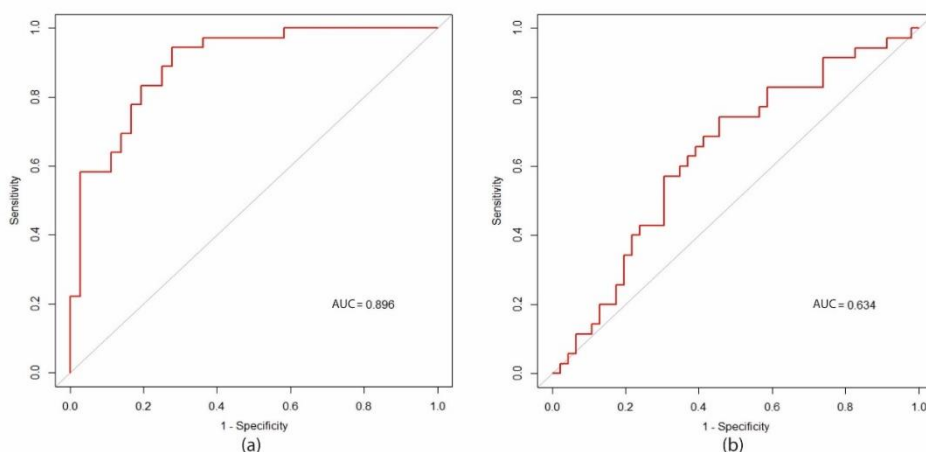
Accuracy results for the selected model on the mixed dataset validation are presented in Table 5. In this case, regarding the efficiency of the model in classifying the two species when compared with field data, 63.0% of individual trees were correctly classified. The model can be classified as average accuracy. In the case of *Pinus pinaster* Ait. 74.1% were successfully classified while in *Pinus pinea* L. the percentage was 54.3%. In the same way, AUC value for mixing validation was 63.4% (Fig.4b). The model applied in mixed stands had lower prediction power than in pure stands. However, the model prediction accuracy should be classified as average.

**Table 4.** Confusion matrix obtained from the validation on pure datasets. P. pinea: *Pinus pinea* L., P. pinaster: *Pinus pinaster* Ait.

		Predicted			User's accuracy (%)
		P. pinea	P. pinaster	Total	
Observed	P. pinea	26	10	36	72.2
	P. pinaster	2	34	36	94.4
	Total	28	44	72	-
Customer's accuracy (%)		92.9	77.3	-	83.3

**Table 5.** Confusion matrix obtained from the validation of the mixed stand dataset. P. pinea: *Pinus pinea* L., P. pinaster: *Pinus pinaster* Ait.

		Predicted			User's accuracy (%)
		P. pinea	P. pinaster	Total	
Observed	P. pinea	25	21	46	54.3
	P. pinaster	9	26	35	74.3
	Total	34	47	81	-
Customer's accuracy (%)		73.5	55.3	-	63.0



**Fig.4** ROC curve of pure and mixed dataset. a: pure stand dataset; b: mixed stand dataset. AUC: area under the curve.

### 2.4 Discussion

LiDAR data provide a widely accepted tool for characterizing forest structure and attributes such as total biomass or volume (Lim et al. 2003; Næsset 2004; Miura and Jones 2010; Nguyen et al. 2016). In addition, LiDAR information has been extensively used to identify individual tree species (Holmgren and Persson 2004; Zhang et al. 2016; Dechesne et al. 2017). Similarly, satellite images have also been used for tree crown delineation (Wang 2010; Lin et al. 2011) and species classification (Leckie et al. 2005; Ballanti et al. 2016). Finally, the combined use of multispectral and LiDAR data has been evaluated as a tool for classifying individual trees (Holmgren et al. 2008; Suratno et al. 2009a).

Statistical analyses have been carried out applying the random forest algorithm, but it could have been done applying other statistical methods, such as linear discriminant analysis. In fact, linear discriminant analysis was carried out, but its results were similar to the ones obtained through random forest. In this way, it was decided to work with random forest algorithm since in using this method several decision trees are averaged, resulting in a more robust analysis with lower overfitting risk.

As hypothesized, the random forest model which only includes LiDAR variables has a greater discriminative power than the model with only spectral variables because the structural differences between the two pine species are more evident than their spectral singularities. However, the model combining information from both sources is even more precise. This finding is in accordance with those of Dalponte et al. (2012), who obtained higher accuracy using models with mixed spectral and LiDAR information than with models based on separate information sources. These authors performed the crown delineation manually and used spectral images and low as well as high density LiDAR data to classify tree species in the Alps mountain area. They discriminated between six species, including evergreen species (*Abies alba* Mill., *Picea abies* (L.) H. Karst., *Pinus mugo* Turra, and *Pinus sylvestris* L.) and deciduous species (*Fagus sylvatica* L. and *Larix decidua* Mill.), and found that evergreen species displayed similar structural and spectral responses, with the exception of *Pinus mugo* Turra, the structure of which is clearly different. The response of deciduous species, however, is well discriminated, taking into account the absence of leaves during the winter season and the spectral intensity at the beginning of the vegetative development. The mean accuracy of the classification between evergreen species was 58%, slightly lower than the results for mixed stands obtained in this study due to the fact that the differences in structural variables between *Pinus pinea* L. and *Pinus pinaster* Ait. are greater than those of their spectral variables. Hence, structural variables have a greater discriminative power than spectral variables in the classifier model.

Including all variables in the complete model maybe sub-optimal due to some of them seeming likely to be useless. In addition, machine learning algorithms like random forest are able to manage correlations between variables without any normalization process and the relationship between them need not be lineal.

Then, the importance of structural variables in discriminating between species is related with the wide differences in individual allometry observed between them. *Pinus pinea* L. usually shows a flattened, short and wide umbrella-like crown, with a long clean stem. On the contrary, *Pinus pinaster* Ait. in pure stands usually presents an apical crown-shape, with lower Live\_Branch, closing the visible ground. Differences in Crown\_Height reflect the importance of this structural variable that the model considers as the most important in these mature stands. This Crown\_height (Max\_Height – Min\_Height) is usually greater in *Pinus pinaster* Ait. than in *Pinus pinea* L. The second variable in importance is AHR which reflects the weight of the structural characteristics involved in its formulation (Crown\_Area and Crown\_Height) for each species. Regarding the relationship between the Crown\_Area and AHR, high values are indicative of *Pinus pinea* L. structure. The third LiDAR variable in importance is Crown\_Area, although it is moderately important. It is noteworthy to mention that this variable is not density-dependent because this type of stands shows very low canopy cover with very little crown overlapping. In addition, the spectral index variables – NDVI, EVI and Infrared band - selected for the model give higher values in *Pinus pinea* L. than *Pinus pinaster* Ait. stands. This may be linked with both the greater plasticity of *Pinus pinea* L., often displacing *Pinus pinaster* Ait. (Bravo-Oviedo et al. 2010), and the generalized decay (associated with decolouration and defoliation) observed in *Pinus pinaster* Ait. forests within the region (Prieto-Recio et al. 2015).

The existence of mixed stands of *Pinus pinaster* Ait. and *Pinus pinea* L. trees indicates that both species have a similar ecological strategy and species mixture may be desirable in terms of increasing and diversifying productivity. However, the conditions currently present in the stands are likely to lead to future dominance of *Pinus pinea* L. over *Pinus pinaster* Ait. given the greater competition tolerance of the former (Ledo et al. 2014). Both species' spectral response in mixed stands could become more uniform than in pure ones. Thus, this uniformity might reflect the fact that living together, both pines become alike, indicating their ecological competition. Furthermore, tree allometry depends on whether a tree grows in a pure or a mixed stand (Forrester et al. 2017).

As expected, the training model selected in this study achieves greater accuracy in the pure stand validation set due to the fact that it has only been trained with pure stands, as shown in Table 1 and Table 2. Therefore, the average accuracy

when validating on pure stands is higher than on mixed stands. In any case, the degree of accuracy is greater for *Pinus pinaster* Ait. classification. Nevertheless, in pure stands, when the model classifies a tree as *Pinus pinea* L. it is commonly correct (only 5.6% of *Pinus pinaster* Ait. trees are classified as *Pinus pinea* L.). In the case of *Pinus pinaster* Ait., the failure rate is higher, with 27.8% being misclassified. However, in mixed stands, *Pinus pinea* L. trees are underestimated while *Pinus pinaster* Ait. are overestimated by around 34%. This may be due to the wider range of variability in the allometric variables of *Pinus pinea* L. individuals growing in mixed stands while *Pinus pinaster* Ait. individuals tend to maintain their pyramidal crown structure. A decrease in accuracy between training and validation models seems logical, since validation sample is an independent data set from the one used to train the model, and this validation set includes plots from mixed stands, where the reduction in accuracy is expected to be higher, since tree allometry attributes are different in these types of stands. Those divergences may be improved in future research by using hyperspectral images and higher resolution LiDAR data (Holmgren et al. 2008; Dalponte et al. 2012). However, this data is usually hardly available and aerial photographs could be useful, but they are not always taken in the infrared channel as some areas in Spain, and they are also extremely dependent on the orography. From NWFP management planning point of view, the maximum interest is to know the percentage of each species correctly classified at stand level. Thus, in mixed stands, classification errors for one species could be compensated with reversed classification errors for the other species.

The main strength of this study is the application of an automatic crown delineation tool together with satellite-based continuous spatial information and open source low density LiDAR data for distinguishing two very similar species at tree level in a mixed conifer forest. The low density of the point cloud LiDAR data is not a limiting issue in these open stands involving wide tree crowns, thus it is highly likely that LiDAR sensor hit them a few times. Also, considering that training data was collected in pure stands where no classification mistake is expected and then validated in pure and mixed stands, the results are still remarkable. With all these, the tool developed could avoid doing stem by stem traditional inventory. The main limitation may be the Pleiades images, which currently cannot easily be replaced by open source multi-spectral satellite images, given the lower spatial resolution of the latter in comparison to Pleiades images. Working with a high level of detail, such as individual tree level, requires a great degree of spatial assessment between the different sources of

information, in order to prevent combining information from inside and outside the crown delineation.

## **2.5 Conclusion**

The species classification tool developed allows, jointly used with an algorithm for automatic crown delineation, to discriminate two very similar species in mixed Mediterranean conifer forests using continuous spatial information on surface: Pleiades images and open source LiDAR data. The method presented here should be valid for regional or landscape-scale assessment of resources computed at tree-level (such as NWFP), enhancing the economic value of the forest production and providing support for forest management decision making.



## 2.6 References

- Arias-Rodil M, Diéguez-Aranda U, Álvarez-González JG, et al (2018) Modeling diameter distributions in radiata pine plantations in Spain with existing countrywide LiDAR data. *Ann For Sci* 75:36. doi: 10.1007/s13595-018-0712-z
- Ballanti L, Blesius L, Hines E, Kruse B (2016) Tree species classification using hyperspectral imagery: A comparison of two classifiers. *Remote Sens* 8:1–18. doi:10.3390/rs8060445
- Barba J, Curiel Yuste J, Martínez-Vilalta J, Lloret F (2013) Drought-induced tree species replacement is reflected in the spatial variability of soil respiration in a mixed Mediterranean forest. *For Ecol Manage* 306:79–87. doi: 10.1016/j.foreco.2013.06.025
- Barrett F, McRoberts RE, Tomppo E, et al (2016) A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. *Remote Sens Environ* 174:279–289. doi: 10.1016/j.rse.2015.08.029
- Beguet B, Chehata N, Boukir S, et al (2014) Classification of forest structure using very high resolution Pleiades image texture. *Geosci Remote Sens Symp (IGARSS), 2014 IEEE Int* 2324–2327. doi: 10.1109/IGARSS.2014.6946936
- Bravo-Oviedo A, Gallardo-Andrés C, del Río M, Montero G (2010) Regional changes of Pinus pinaster site index in Spain using a climate-based dominant height model. *Can J For Res* 40:2036–2048. doi: 10.1139/X10-143
- Breiman L, Friedman J, Olshen R, Stone C (2001) Classification and regression trees. Chapman Hall, New York
- Calama R, Gordo FJ, Mutke S, Montero G (2008) An empirical ecological-type model for predicting stone pine (*Pinus pinea* L.) cone production in the Northern Plateau (Spain). *For Ecol Manage* 255:660–673. doi: 10.1016/j.foreco.2007.09.079
- Calama R, Gordo J, Madrigal G, et al (2016) Enhanced tools for predicting annual stone pine (*Pinus pinea* L.) cone production at tree and forest scale in Inner Spain. *For Syst* 25. doi: 10.5424/fs/2016253-09671
- Castaño-Díaz M, Álvarez-Álvarez P, Tobin B, et al (2017) Evaluation of the use of low-density LiDAR data to estimate structural attributes and biomass yield in a short-rotation willow coppice: an example in a field trial. *Ann For Sci* 74:69. doi: 10.1007/s13595-017-0665-7
- Dalponte M, Bruzzone L, Gianelle D (2012) Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens Environ* 123:258–270. doi: 10.1016/j.rse.2012.03.013
- Dechesne C, Mallet C, Le Bris A, Gouet-Brunet V (2017) Semantic segmentation of forest stands of pure species combining airborne lidar data and very high resolution multispectral imagery. *ISPRS J Photogramm Remote Sens* 126:129–145. doi: 10.1016/j.isprsjprs.2017.02.011
- Deng S, Katoh M, Yu X, et al (2016) Comparison of tree species classifications at the individual tree level by combining ALS data and RGB images using different algorithms. *Remote Sens* 8. doi: 10.3390/rs8121034

- Dubayah RO, Drake JB (2000) Lidar Remote Sensing for Forestry Applications. *J For* 98:44–46
- European Environmental Agency (2006) European forest types: Categories and types for sustainable forest management reporting policy. 114
- Everitt BS, Skrondal A (2010) *The Cambridge Dictionary of Statistics*
- Fassnacht FE, Latifi H, Stereńczak K, et al (2016) Review of studies on tree species classification from remotely sensed data. *Remote Sens Environ* 186:64–87. doi: 10.1016/j.rse.2016.08.013
- Forrester DI, Benneter A, Bouriaud O, Bauhus J (2017) Diversity and competition influence tree allometric relationships – developing functions for mixed-species forests. *J Ecol* 105:761–774. doi: 10.1111/1365-2745.12704
- Gordo FJ, Montero G, Gil L (2012) La problemática de la regeneración natural de los pinares en los arenales de la Meseta Castellana
- Hollaus M, Dorigo W, Wagner W, et al (2009) Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. *Int J Remote Sens* 30:5159–5175. doi: 10.1080/01431160903022894
- Holmgren J, Persson Å (2004) Identifying species of individual trees using airborne laser scanner. *Remote Sens Environ* 90:415–423. doi: 10.1016/S0034-4257(03)00140-8
- Holmgren J, Persson Å, Söderman U (2008) Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int J Remote Sens* 29:1537–1552. doi: 10.1080/01431160701736471
- Korpela I, Ole Ørka H, Maltamo M, et al (2010) Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fenn* 44:319–339. doi: 10.14214/sf.156
- Leckie DG, Tinis S, Nelson T, et al (2005) Issues in species classification of trees in old growth conifer stands. *Can J Remote Sens* 31:175–190. doi: 10.5589/m05-004
- Ledo A, Cañellas I, Barbeito I, et al (2014) Species coexistence in a mixed Mediterranean pine forest: Spatio-temporal variability in trade-offs between facilitation and competition. *For Ecol Manage* 322:89–97. doi: 10.1016/j.foreco.2014.02.038
- Lee J, Cai X, Lellmann J, et al (2016) Individual tree species classification from airborne multi-sensor imagery using robust PCA. *IEEE J* 1–15
- Lim K, Treitz P, Wulder M, et al (2003) LiDAR remote sensing of forest structure. *Prog Phys Geogr* 27:88–106. doi: 10.1191/0309133303pp360ra
- Lin C, Lo CS, Thomson G (2011) A textural modification of the MMAC algorithm for individual tree delineation in forest stand using aerial bitmap images. *Proc - 4th Int Congr Image Signal Process CISP 2011* 3:1604–1608. doi: 10.1109/CISP.2011.6100501
- Liu HQ, Huete A (1995) A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans Geosci Remote Sens* 33:457–465

- Lopez-García P (1980) Estudio de semillas prehistóricas en algunos yacimientos españoles. *Trab Prehist* 37:419–432
- López C, Espinosa J, Bengoa J (2009) Mapa de Vegetación de Castilla y León
- Maack J, Kattenborn T, Fassnacht FE, et al (2015) Modeling forest biomass using very-high-resolution data - combining textural, spectral and photogrammetric predictors derived from spaceborne stereo images. *Eur J Remote Sens* 48:245–261. doi: 10.5721/EuJRS20154814
- MAGRAMA (2012) Criteria and indicators for sustainable forest Management, in Spanish forests 2012
- Miura N, Jones SD (2010) Characterizing forest ecological structure using pulse types and heights of airborne laser scanning. *Remote Sens Environ* 114:1069–1076. doi: 10.1016/j.rse.2009.12.017
- Moreno-Fernández D, Cañellas I, Calama R, et al (2013) Thinning increases cone production of stone pine (*Pinus pinea* L.) stands in the Northern Plateau (Spain). *Ann For Sci* 70:761–768. doi: 10.1007/s13595-013-0319-3
- Næsset E (2004) Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sens Environ* 91:243–255. doi: 10.1016/j.rse.2004.03.009
- Nanos N, Tadesse W, Montero G, et al (2000) Modelling resin production distributions for *Pinus Pinaster* Ait using two probability functions. *Ann For Sci* 57:379–377. doi: 10.1051/forest:2000128
- Nanos N, Tadesse W, Montero G, et al (2001) Spatial stochastic modeling of resin yield from pine stands. *Can J For Res* 31:1140–1147. doi: 10.1139/cjfr-31-7-1140
- Nanos N, Calama R, Montero G, Gil Luis (2004) Geostatistical prediction of height/diameter models. *For Ecol Manage* 195: 221–235. doi: 10.1016/j.foreco.2004.02.031
- Nguyen HT, Hutryra LR, Hardiman BS, Raciti SM (2016) Characterizing forest structure variations across an intact tropical peat dome using field samplings and airborne LiDAR. *Ecol Appl* 26:587–601. doi: 10.1890/15-0017
- ORFEO (2018) ORFEO
- Ørka HO, Gobakken T, Næsset E, et al (2012) Simultaneously acquired airborne laser scanning and multispectral imagery for individual tree species identification. *Can J Remote Sens* 38:125–138. doi: 10.5589/m12-021
- PNOA (2018) Plan Nacional de Ortofotografía Aérea
- PNT (2018) Plan Nacional de Teledetección
- Popescu SC (2007) Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy* 31:646–655. doi: 10.1016/j.biombioe.2007.06.022

- Popescu SC, Wynne RH, Nelson RF (2003) Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Can J Remote Sens* 29:564–577. doi: 10.5589/m03-027
- Prieto-Recio C, Martín-García J, Bravo F, Diez JJ (2015) Unravelling the associations between climate, soil properties and forest management in *Pinus pinaster* decline in the Iberian Peninsula. *For Ecol Manage* 356:74–83. doi: 10.1016/j.foreco.2015.07.033
- Prodan M (1968) *Forest bimetrics*, Oxford, UK
- QGIS Development Team (2017) QGIS Geographic Information System. Open Source Geospatial Foundation Project
- R Development Core Team (2016) R: A Language and Environment for Statistical Computing
- Riofrío J, del Río M, Pretzsch H, Bravo F (2017) Changes in structural heterogeneity and stand productivity by mixing Scot pine and Maritime pine. *For Ecol Manage InProcess*: doi: 10.1016/j.ibmb.2009.12.007
- Rodríguez-García A, Martín JA, López R, et al (2015) Influence of climate variables on resin yield and secretory structures in tapped *Pinus pinaster* Ait. in central Spain. *Agric For Meteorol* 202:83–93. doi: 10.1016/j.agrformet.2014.11.023
- Rodríguez F, Lizarralde I (2015) Inventario forestal con tecnología LiDAR, ¡Más barato y más preciso! *Forestalis* 24:33–34
- Rougier S, Anne P (2014) Improvements of Urban vegetation segmentation and classification using multi - temporal Pleiades images. *South-Eastern Eur J Earth Obs Geomatics* 3:409–414
- Rouse JW, Hass RH, Schell JA, Deering DW (1973) Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resour Technol Satell Symp* 1:309–317. doi: citeulike-article-id:12009708
- Ruiz LÁ, Recio JA, Crespo-Peremarch P, Sapena M (2016) An object-based approach for mapping forest structural types based on low-density LiDAR and multispectral imagery. *Geocarto Int* 6049:1–15. doi: 10.1080/10106049.2016.1265595
- Soliño M, Yu T, Alía R, et al (2018) Resin-tapped pine forests in Spain: Ecological diversity and economic valuation. *Sci Total Environ* 625:1146–1155. doi: 10.1016/j.scitotenv.2018.01.027
- Suratno A, Seielstad C, Queen L (2009a) Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS J Photogramm Remote Sens* 64:683–693. doi: 10.1016/j.isprsjprs.2009.07.001
- Suratno A, Seielstad C, Queen L (2009b) Mapping Tree Species Using LIDAR in Mixed-Coniferous Forests. *Proc Silvilaser 2009* 1–10
- Tonolli S, Dalponte M, Neteler M, et al (2011) Fusion of airborne LiDAR and satellite multispectral data for the estimation of timber volume in the Southern Alps. *Remote Sens Environ* 115:2486–2498. doi: 10.1016/j.rse.2011.05.009

- Valbuena M, Santamaría J, Sanz F (2016a) Estimation of diameter and height of individual trees for *Pinus sylvestris* L. based on the individualising of crowns using airborne LiDAR and the National forest inventory data. *For Syst* 25:1–11. doi: 10.5424/fs/2016251-05790
- Valbuena R, Maltamo M, Packalen P (2016b) Classification of forest development stages from national low-density lidar datasets: a comparison of machine learning methods. *Rev Teledetección* 15–25. doi:10.4995/raet.2016.4029
- Vauhkonen J, Ørka HO, Holmgren J, et al (2014) Tree species recognition based on Airborne Laser Scanning and complementary data source. 27. doi:10.1007/978-94-017-8663-8
- Vega Ispahuaylas LA, Hirata Y, Santos LCV, Torobeo NS (2018) Natural forest mapping in the Andes (Peru): A comparison of the performance of machine-learning algorithms. *Remote Sens* 10. doi: 10.3390/rs10050782
- Wang L (2010) A Multi-scale Approach for Delineating Individual Tree Crowns with Very High Resolution Imagery. *Photogramm Eng Remote Sens* 76:371–378. doi:10.14358/pers.76.4.371
- Xie Y, Sha Z, Yu M (2008) Remote sensing imagery in vegetation mapping: a review. *J Plant Ecol* 1:9–23. doi: 10.1093/jpe/rtm005
- Yesilnacar EK (2005) The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey (Ph.D Thesis) Department of Geomatics the University of Melbourne, p. 423.
- Zhang Z, Kazakova A, Moskal LM, Styers DM (2016) Object-based tree species classification in urban ecosystems using LiDAR and hyperspectral data. *Forests* 7:1–16. doi: 10.3390/f7060122
- Zipkin EF, Grant EHC, Fagan WF (2012) Evaluating the predictive abilities of community occupancy models using AUC while accounting for imperfect detection. *Ecol Appl* 22, 1962–1972.









## CHAPTER 3

---

Stand types discrimination comparing machine learning algorithms in monteverde, Canary Islands

*“Mira profundamente en la naturaleza  
y entonces comprenderás todo mejor”*

Albert Einstein

### **Chapter 3. Stand types discrimination comparing machine learning algorithms in monteverde, Canary Islands**

---

#### **Abstract**

The main objective of this study is to determine the best machine-learning algorithm to classify the stand types of Monteverde forests combining LiDAR, orthophotography, and Sentinel data, thus providing an easy and cheap method to classify Monteverde stand types. The area of study covers 1500 ha forest in Monteverde, North Tenerife, Canary Islands. RF, SVML, SVMR and ANN algorithms are used to classify the three Monteverde stand types. Before training the model, feature selection of LiDAR, orthophotography, and Sentinel data through VSURF was carried out. Comparison of its accuracy was performed.

Five LiDAR variables were found to be the most efficient for classifying each object, while only one Sentinel index and one Sentinel band was valuable. Additionally, standard deviation and mean of the Red orthophotography colour band, and ratio between Red and Green bands were also found to be suitable. SVML is confirmed as the most accurate algorithm (0.904, 0.041 SD) while ANN showed the lowest value of 0.891 (0.073 SD). SVMR and RF obtain 0.902 (0.060 SD) and 0.904 (0.056 SD) respectively. SVML was found to be the best method given its low standard deviation. The similar high accuracy values among models confirm the importance of taking into account diverse machine-learning methods for stand types classification purposes and different explanatory variables. Although differences between errors may not seem relevant at a first glance, due to the limited size of the study area with only three plus two categories, such differences could be highly important when working at large scales with more stand types.

### 3.1 Introduction

During the second half of the 20th century traditional forestry practices used in the monteverde forest in Canary Islands changed dramatically. Moreover, the beginning of this century is marked by a shift in the preferences of society as regards the ecosystem services provided by the monteverde forest from traditional forest resources to conservation orientated services, especially as part of one of the current areas of Macaronesia and the importance of its relict flora (Arozena & Panareda, 2013). Nowadays, monteverde conservation is explained by its uniqueness and the need for monitoring Laurisilva dynamics which lead to spot new and innovative classification tools. Recent remote sensing technologies can help to improve its management, providing easy and cheap classification of its stand types reducing cost and time consumption from traditional forest management procedures based on expensive field studies.

Data derived from active and passive remote sensors are of great interest in forestry. In particular, the combination of LiDAR information with Sentinel multispectral images provide a powerful tool for classifying forests with high densities and stocking rates, thus reducing the cost of the estimation process (Zhu et al., 2017). In addition, the volume of data we are dealing with is constantly growing, including the aim at retrieving a wide variety of geographic and ecological characteristics. Consequently, the analyses can only be tackled using computational methods.

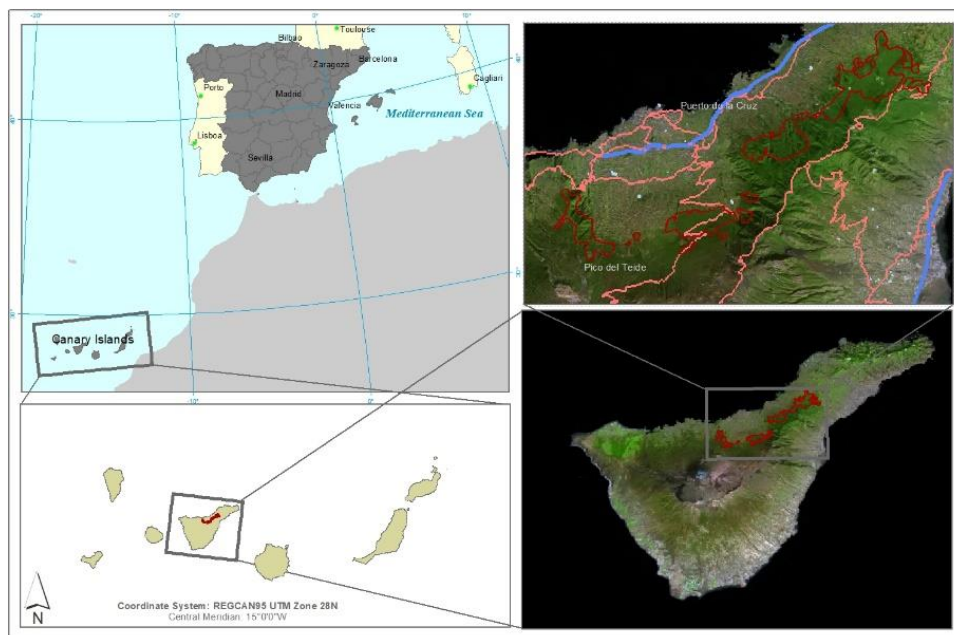
From the wide range of algorithms used to find the rules for object classification in forest sciences, the random forest algorithm (RF) has shown high rates of accuracy. Nevertheless, the Linear and Radial Support Vector Machine (SVML, SVMR) and Artificial Neural Networks (ANN) algorithms are increasingly being taken into consideration in this area (Nitze et al., 2012; Valbuena et al., 2016; Vega Isuhuaylas et al., 2018; Xu et al., 2018). Although its accuracy is not always taken into account, its effectiveness could be easily increased only keeping in mind the most suitable one.

The main objective of this work is to determine the best machine-learning algorithm to classify the three monteverde stand types, Canary Islands, combining LiDAR, orthophotography and Sentinel data.

## 3.2 Material and methods

### 3.2.1 Study area

The study area is located in a 1500 ha evergreen forest in the North of the island of Tenerife in the Canary Islands, between 200–1300 m.a.s.l (Fig. 1). The ecosystem is highly valuable in economic terms but also as regards ecosystem services, characterised by a wide variety of species.



**Fig. 1.** Geographical location of the Canary Islands (upper left), Tenerife (bottom left), study area location in Tenerife (bottom right) and study area in detail (upper right).

### 3.2.2 Field data

A total number of 259 objects were measured during May of 2017. The objects were irregular in shape, collecting homogenous remote information. Each of ones included visual information about stand type (main species and GPS coordinates). Objects were selected with the aim of picking up the highest spectral variability of stand types in monteverde, three forest categories were defined according to its relevance in management (Arozena & Panareda, 2013): (i) Brezal (BR), composed by *Erica arborea* L. shrubby stands with variable cover

and occasional presence of scrub; (ii) Faya-Brezal-Acebiñal (FBA), composed by shrub or tree stands with a high density and average diameter of saplings 5-10 cm; the proportion of *E. arborea* L. in the specific composition of the stand varies and stands may be dominated by other species such as *Morella faya* Aiton., *Laurus novocanariensis* Rivas Mart., Lousã, Fern.Prieto, E.Díaz, J.C.Costa & C.Aguiar., or *Viburnum rigidum* L. and other companion species at different stand types of development; (iii) Laurisilva (LA), mixed stands with a significant presence of *L. novocanariensis* Rivas Mart., Lousã, Fern.Prieto, E.Díaz, J.C.Costa & C.Aguiar. and *V. rigidum* L. with average diameter greater than 10 cm and the rare presence of *E. arborea* L. In addition, two more stand types were included in the database: (iv) Non-Monteverde ground (NMG), defined as bare ground or scrub less than 2 m high; and (v) Non-Monteverde forest (NMF) defined as stand cover composed by other species. The number of samples set out in each category were 32 for BR, 143 for FB, 27 for LA, 6 for NMG, and 51 for NMF, respectively.

### 3.2.3 LiDAR data

The island of Tenerife was scanned using a LiDAR sensor in 2009 and 2016, with an average nominal point density of 0.5 pulses m<sup>2</sup> (PNOA project, Spanish Government). Data from the study area were provided in digital files of 2x2 km extension. Point clouds were automatically classified and coloured, taking RGB orthophotos as a reference. LiDAR data was processed using FUSION software (McGaughey, 2007) and several raster variables (5 m resolution) were generated: digital tree height model (DTHM), digital terrain model (DTM), canopy cover (FCC), height maximum (HEIGHT), height percentile 95 (P95), height percentile 25 (P25), height model growth between 2016 and 2009 (HMGR), canopy cover growth between 2016 and 2009 (FCCGR), height percentile 95 growth between 2016 and 2009 (HGR), height percentile 25 growth between 2016 and 2009 (H25GR) along with standard deviations for all these variables.

### 3.2.4 Multispectral imagery data source

European Space Agency Sentinel satellite images (10 m and 20 m resolution) of the study area captured in December 2015 and January 2017 were employed in order to avoid clouds and deciduous tree reflectance. These orthorectified and atmospherically corrected images were downloaded from Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). Medium value of each of its bands

were calculated. Besides, several vegetation indexes were calculated based on imagery data: NDVI, RNDVI, GNDVI, SAVI, LAI-SAVI, SR, and EVI (10 m resolution), and NDVI705, NDWI, RNDWI, NDII, NDI45, NBR, MSI (20 m resolution) (Henrich et al., 2012). In addition, aerial orthophotograph images (2015) were provided by CNIG-PNOA (Spanish Government), with a resolution of 25 cm in the official reference geodetic system, REGCAN95 - UTM zone 28N projection. Medium value and standard deviation of each of its bands, and the ratio between Red and Green bands, were calculated. All these spectral indices were included as potential predictors in the classification model.

### 3.2.5 Segmentation process

To define object segmentation in the study area, we executed an Object Based Image Analysis which created an image-object through the aggregation of pixels by image segmentation from the Orfeo Tool Box (OTB Development Team, 2017). The two variables we worked with are: (i) the spatial resolution and (ii) range domains, which is the allowable spectral range within each segment for each band at a minimum scale of 40 m<sup>2</sup>.

Data features from each source of information were assigned to the generated objects from segmentation using QGIS Zonal Statistics Plugin (QGIS Development Team, 2017).

### 3.2.6 Data analysis

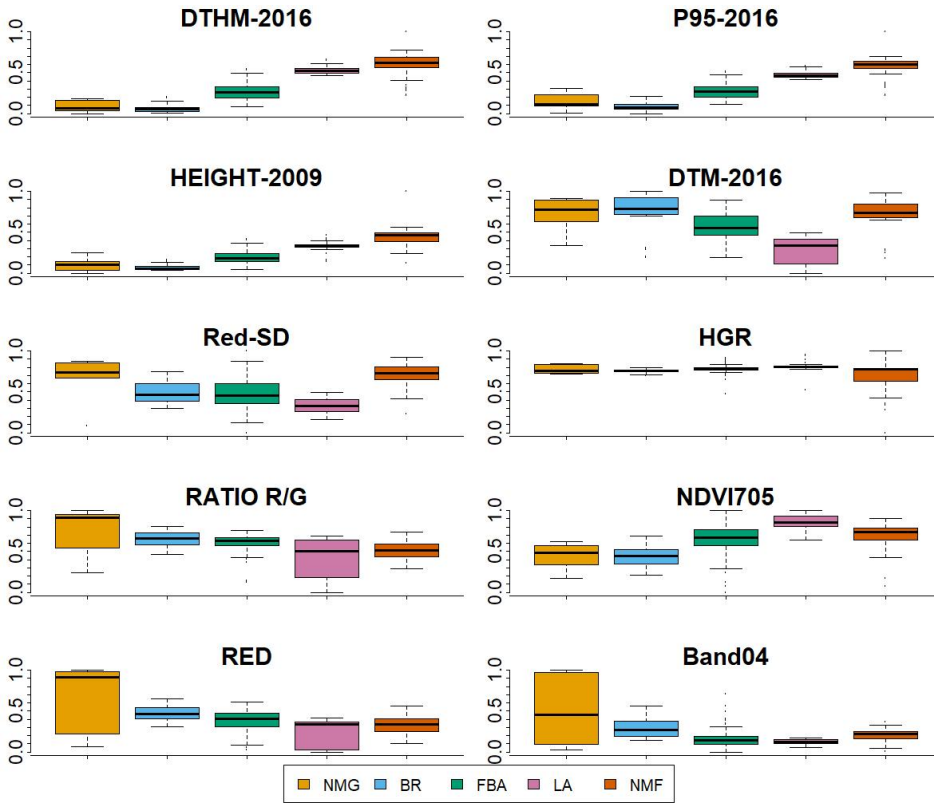
Prior to model training, feature selection using the Variable Selection Using Random Forest (VSURF) was performed (Genuer et al., 2015). In order to conduct a useful comparison between RF, SVM, SVMR, and ANN, caret package in R Software was run using the rf, svmLinear, svmRadial and nnet methods with default parameters (Kuhn et al., 2018). Furthermore, the 'overfitting' problem was reduced by Cross Validation using 10 folds with three repetitions.

### 3.3 Results and discussion

The application of VSURF procedure selected ten features (Fig. 2). Five LiDAR variables were found as the most efficient for classifying each object, while only one Sentinel index and one Sentinel band was valuable. Additionally, standard deviation and mean of the Red orthophotography colour band, and ratio between Red and Green bands were also found to be suitable.

When the data dispersion was analysed, selected LiDAR features showed differences according to the classification factor variable. The feature DTHM-2016 reveals importance at its clear boundaries between the different forest typologies and it shows the difference between forest typologies stage, together with P95-2016 and HGR variables (Fig. 2). The rest of the LiDAR data support distinction and split soil from the other types. The selection of Sentinel NDVI705 reflects red edge radiation and its usefulness with very high spectral resolution reflectance data (Sinergise, 2018). In contrast, the average and the standard deviation of Red orthophotography colour band showed a crucial disjunction among NMF and monteverde.

From the Cross-Validation results, most of the models reached 0.90 mean accuracy (Table 1). SVMl was confirmed as the most accurate method while ANN presented the lowest accuracy value. SVMR and RF obtained the intermediated accuracy values of 0.902 and 0.904 respectively. Simultaneously, Cohen Kappa values did not vary from the achieved accuracy values. Given its low standard deviation, SVMl was found as the best method thanks to the variable influence shown in Table 2 for monteverde stand classification types. Usefulness of LiDAR variables in the classification of this case study is demonstrated by the fact that four (DTMH-2016, P95-2016, HEIGHT-2009, and DTM-2016) out of ten selected variables have the highest accuracy, being the one LiDAR variable remaining (HGR) the sixth one (Table 2).



**Fig. 2.** Distribution of selected features as a result of variable selection applying random forest (VSURF) according to monteverde typologies. Values under normalization from minimum '0' to maximum value '1'. DTHM-2016, digital tree height model based on 2016 LiDAR data; P95-2016, height percentile 95 based on 2016 LiDAR data; HEIGHT-2009, maximum height based on 2009 LiDAR data; DTM-2016, digital terrain model based on 2016 LiDAR data; Red-SD, standard deviation of the Red band orthophotograph data; HGR, height growth based on 2009 and 2016 LiDAR data; RATIO R/G, ratio between Red and Green bands orthophotograph data; NDVI705, Normalized Differenced Vegetation Index 705 2015 image data; RED, Red band orthophotograph data; Band04, Sentinel band 04 2015 image data.



**Table 1.** Accuracy, Accuracy Standard Deviation, Kappa Value and Kappa Value Standard Deviation resulted from each model. ANN, Artificial Neural networks algorithm; SVML, Linear support vector machine algorithm; SVMR, radial support vector machine algorithm; RF, random forest algorithm.

	<b>Accuracy</b>	<b>Accuracy SD</b>	<b>Kappa</b>	<b>Kappa SD</b>
<b>ANN</b>	0.891	0.073	0.817	0.129
<b>SVML</b>	0.904	0.041	0.842	0.070
<b>SVMR</b>	0.902	0.060	0.836	0.106
<b>RF</b>	0.904	0.056	0.841	0.092

**Table 2.** Importance of variables in the linear support vector machine algorithm (SVML) final model for the monteverde typology classification. Values state the loss of accuracy and/or standard deviation according to the lack of each feature from SVML calculated. DTHM-2016, digital tree height model based on 2016 LiDAR data; P95-2016, height percentile 95 based on 2016 LiDAR data; HEIGHT-2009, maximum height based on 2009 LiDAR data; DTM-2016, digital terrain model based on 2016 LiDAR data; Red-SD, standard deviation of the Red band orthophotograph data; HGR, height growth based on 2009 and 2016 LiDAR data; RATIO R/G, ratio between Red and Green bands orthophotograph data; NDVI705, Normalized Differenced Vegetation Index 705; RED, Red band orthophotograph data; Band04, Sentinel band 04 image data.

	<b>Source</b>	<b>Mean</b>	<b>Standard Deviation</b>
<b>SVML</b>		0.904	0.041
<b>DTHM-2016</b>	LiDAR	0.892	0.055
<b>P95-2016</b>	LiDAR	0.903	0.052
<b>HEIGHT-2009</b>	LiDAR	0.909	0.045
<b>DTM-2016</b>	LiDAR	0.880	0.049
<b>Red-SD</b>	Orthophotography	0.908	0.043
<b>HGR</b>	LiDAR	0.886	0.057
<b>RATIO R/G</b>	Orthophotography	0.907	0.046
<b>NDVI705</b>	Sentinel	0.900	0.042
<b>RED</b>	Orthophotography	0.915	0.044
<b>Band04</b>	Sentinel	0.915	0.049

Obtained results differ with other previous ones in the literature. For instance, Valbuena et al. (2016) found that the best machine learning algorithm to determine Mediterranean forest development stages are RF and ANN, and Vega

Isuhuaylas et al. (2018) met with SVM and RF to classify Andes mountain forests and shrubland land cover classes.

The high accuracy values confirm the importance of taking into account diverse machine-learning methods for stand classification purposes and different explanatory aspects. Notwithstanding the small deviation between accuracy values, our work proves that SVML is the best algorithm for the monteverde forest classification due to the minimal results' scattering. Over 90% of cases monteverde stand type (BR, FBA, LA, NMG, or NMF) determined from remote sensing data are correct, though small size of the study area and the only three, plus two, stand types considered here should not be forgotten. Our results confirm machine learning classification is a suitable tool to optimize classification in monteverde forest and, thus, its management.

Boost of machine learning algorithms applied to classify forest is broadly enough demonstrated. Although differences between errors in accuracy and scattering may not seem significant at a first glance, such differences are highly important when working at large scales. Errors may be higher when classifying larger areas with more stand types. So, comparisons between algorithms should be considered when stand classification analyses are performed owing to different behaviour of algorithms relying on stand types, features and size of the study area.

### 3.4 References

- Arozena ME, Panareda, JM, 2013. Forest transition and biogeographic meaning of the current laurel forest landscape in Canary Islands, Spain. *Physical Geography* 34: 211-235.
- Genuer R, Poggi JM, Tuleau-Malot C. 2015. VSURF: An R Package for Variable Selection Using Random Forests. <https://journal.r-project.org/archive/2015/RJ-2015-018/RJ-2015-018.pdf> [8 June 2018].
- Henrich V, Krauss G, Götze C, Sandow C, 2012. IDB - [www.indexdatabase.de](http://www.indexdatabase.de), Entwicklung einer Datenbank für Fernerkundungsindizes. *AK Fernerkundung, Bochum* 45: 10.
- Kuhn M, Jed Wing A, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, et al., 2018. caret: Classification and Regression TrainingR package version 6.0-78. <https://cran.r-project.org/package=caret> [8 June 2018].
- McGaughey R, 2007. Fusion/LDV: Software for Lidar Data Analysis and Visualization; USDA Forest Service, Pacific Northwest Research Station: Portland, OR, USA.
- Nitze I, Schulthess U, Asche H, 2012. Comparison of machine learning algorithms Random Forest, Artificial Neural Network and Support Vector Machine to maximum likelihood for supervised crop type classification. *Proc GEOBIA, Rio de Janeiro (Brazil)*, May 7-9, pp: 35.
- OTB Development Team, 2017. The ORFEO Tool Box Software Guide. Updated for OTB-5.10.0
- QGIS Development Team, 2017. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- Sinergise, 2018. Sentinel 2 EO products. [https://www.sentinel-hub.com/develop/documentation/eo\\_products/Sentinel2EOproducts](https://www.sentinel-hub.com/develop/documentation/eo_products/Sentinel2EOproducts). [24 May 2018].
- Valbuena R, Maltamo M, Packalen P, 2016. Classification of forest development stages from national low-density LiDAR datasets: a comparison of machine learning methods. *Revista de Teledetección* 45: 15-25.
- Vega Isuhuaylas LA, Hirata Y, Ventura Santos LC, Serrudo Torobeo N, 2018. Natural forest mapping in the Andes (Peru): a comparison of the performance of machine-learning algorithms. *Remote Sens* 10: 782.
- Xu C, Manley B, Morgenroth J, 2018. Evaluation of modelling approaches in predicting forest volume and stand age for small-scale plantation forests in New Zealand with RapidEye and LiDAR. *Int J Appl Earth Obs Geoinformation* 73: 386-396.
- Zhu XX, Tuia D, Mou L, Xia GS, Zhang L, Xu F, Fraundorfer F, 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci Remote Sens Mag* 5: 8-36.







## CHAPTER 4

---

Fire and burn severity assessment: calibration of Relative Differenced Normalized Burn Ratio (RdNBR) with field data

*“Salvaje no es quien vive en la naturaleza,  
salvaje es quien la destruye”*

Anónimo

## **Chapter 4. Fire and burn severity assessment: calibration of Relative Differenced Normalized Burn Ratio (RdNBR) with field data**

---

### **Abstract**

The assessment of burn severity is highly important in order to describe and measure the effects of fire on vegetation, wildlife habitat and soils. The estimation of burn severity based on remote sensing is a powerful tool that, to be useful, needs to be related and validated with field data. The present paper explores the relationships between field accessible variables and Relative Differenced Normalized Burn Ratio (RdNBR) index by using linear mixed-effects models and boosted regression trees, based on data from 28 large fires and 668 field measurements across three countries in southern Europe.

The RdNBR clearly reflected the mean height of charred stem and loss of ligneous, living shrub and tree cover during the fire. The paper confirms that remote sensing indices provide an acceptable assessment of fire induced impact on forest vegetation but also highlights there are important between-fire variations due to specific contexts that modify these relationships. These variations can be effectively assessed and should be taken into account in future predictive efforts.



## 4.1 Introduction

Fire is a key ecological process in many forested areas, and its variation in size and severity strongly influences forest structure and wildlife habitats (Herrando and Brotons 2002; Herrando et al., 2003). Large amounts of data are required to define post-fire management plans, either at strategic or tactical level. In this context, remote sensing is a necessary and reliable tool for providing spatially continuous information about the state of forest ecosystems, associated resources and for assessing their evolution overtime (Franklin, 2001; Jones and Vaughan, 2010; Lefsky et al., 2002; Xie et al., 2008).

Remote sensing-based assessments have gained relevance in natural resources management as they are much less time-consuming than assessments entirely based on field inventories, while reducing uncertainties associated with inference and estimation from sample surveys in areas showing fine-scale heterogeneity (Franklin, 2001). In addition, these assessments have also proved their validity to improve the retrieval of representative field data through stratification methods (Adam et al., 2010; Czaplewski and Patterson, 2003; Grafström and Ringvall, 2013). Currently, a variety of remotely sensed images are available for providing information about forest ecosystems by a range of airborne and satellite sensors, from multi-spectral to hyperspectral sensors, with different spatial and temporal resolutions.

The use of multispectral images has also supposed an important step ahead in the impact assessment of the impact of forest disturbances; specifically, to assess wildfire impact, as it has become an essential source of data to map burned area (Chuvienco, 2012; Fernández et al., 1997; Kushlat and Ripple, 1998; Lentile et al., 2006; Miller and Yool, 2002; Pereira, 1999; Roy et al., 2002). In this context, burn severity has become a key variable in order to assess the impact of fire on forest ecosystems and can be defined as the loss of soil and aboveground organic matter in a burned area, being used to analyse non-immediate fire effects (Keeley, 2009; Morgan et al., 2014). Burn severity has been mapped through indices such as the Normalized Burn Ratio (NBR; Key and Benson, 1999) considering the absolute difference change detection protocol (dNBR), or as a relative difference (RdNBR) (De Santis, 2008; French et al., 2008; Keeley, 2009). These indices combined with other metrics derived from available satellite multi-spectral images have been used for mapping burn severity in

several studies related to large fires (Harris and Taylor, 2015; Harris et al., 2011; Keeley, 2009; Malone et al., 2011; Picotte et al., 2016; Quintano et al., 2018; Warner et al., 2017).

Researchers and managers use these remotely sensed assessments and field data to describe and measure the fire effects on vegetation, wildlife habitat and soils (Keeley, 2009), and improve the understanding of fire regimes, the successional pathways and hydro-geomorphological effects (e.g. Harris et al., 2011; Martín-Alcón and Coll 2016). In addition, on regions where fire is a common threat, as on the northern part of the Mediterranean basin (Cardil et al., 2015; Kalabokidis et al., 2015), forest managers require a fast and accurate estimation of burn severity to map and assess fire damages across different landscapes. However, burn severity indexes coming from remote sensing data, while accessible, need to be translated into forest variables to be properly understood by forest managers. For this purpose, it is required to compare the burn severity indices coming from remote sensing data with field measures of burn severity (Cocke et al., 2005; Epting et al., 2005; French A et al., 2008; Miller and Quayle, 2015; Parks et al., 2014) to identify reliable relations. These relationships have been reported as either linear; by using Composite Burn Index (CBI) (Key and Benson, 2005) or its modified version GeoCBI (De Santis, 2008; Santis and Chuvieco, 2009; Veraverbeke et al., 2011, 2010); or as non-linear: as in Parks et al., (2014); Soverel et al., (2010) and Van Wagtenonk et al., (2004).

CBI is an operational methodology developed in the USA for burn severity assessment which provides a good field assessment on fire severity (Key and Benson, 2005). It has been demonstrated, however, that may require modifications to better adapt to specific ecosystems (Santis and Chuvieco, 2009). It is clear that the impact of fire on forested areas depends on fire intensity and time of residence, as well as the tree capacity to protect sensitive tissues (Butler and Dickinson, 2010; Michaletz and Johnson, 2007), which is related to the species and size of each single tree (Catry et al., 2010). This aspect is especially relevant when delayed responses to fire need to be accounted (Soverel et al., 2010; Valor et al., 2017), as each tree responds differently to the same fire intensity depending on its size, species-dependent physiological traits and time from the fire event. These result in a different spectral response and, thus, a different burn severity value that needs to be adjusted (Miller and Quayle, 2015). Although CBI and modified versions of the severity index, have shown their utility, further

studies are required to understand the relation between remote sensing derived indices and ground data, in order to be applicable to large and heterogeneous landscapes consisting on a wide variety of vegetation types. A first step to test this can be the evaluation of the relationships between a remote sensing derived index and field-based measurements of severity; to identify how each individual severity indicator is related to the satellite index, and afterwards to test whether combined also produce a clear satellite signature without major correlation problems. A limited number of studies have investigated the empirical relationships between remote-sensing indices of burn severity and field data in the context of the Mediterranean Basin, even though it is a region heavily affected by forest fires (Cardil et al., 2015, 2014; Kalabokidis et al., 2015; Salis et al., 2013), and characterized by a high variability of vegetation and fuels across its forested landscapes.

The present paper seeks to evaluate the performance of the satellite-derived indices as continuous metrics of burn severity. For this purpose, we tested their correspondence to variables measured in the field, from several fires and we built simple and robust explanatory models to consider linear (linear mixed-effects models) and non-linear effects (Boosted Regression Trees (BRT)), based on data from twenty-eight fires across three countries from Southern Europe, aiming at: (1) assessing correlations among field explanatory variables and RdNBR; and (2) analysing the burn severity variability among the studied fires and the random effect of each fire event.

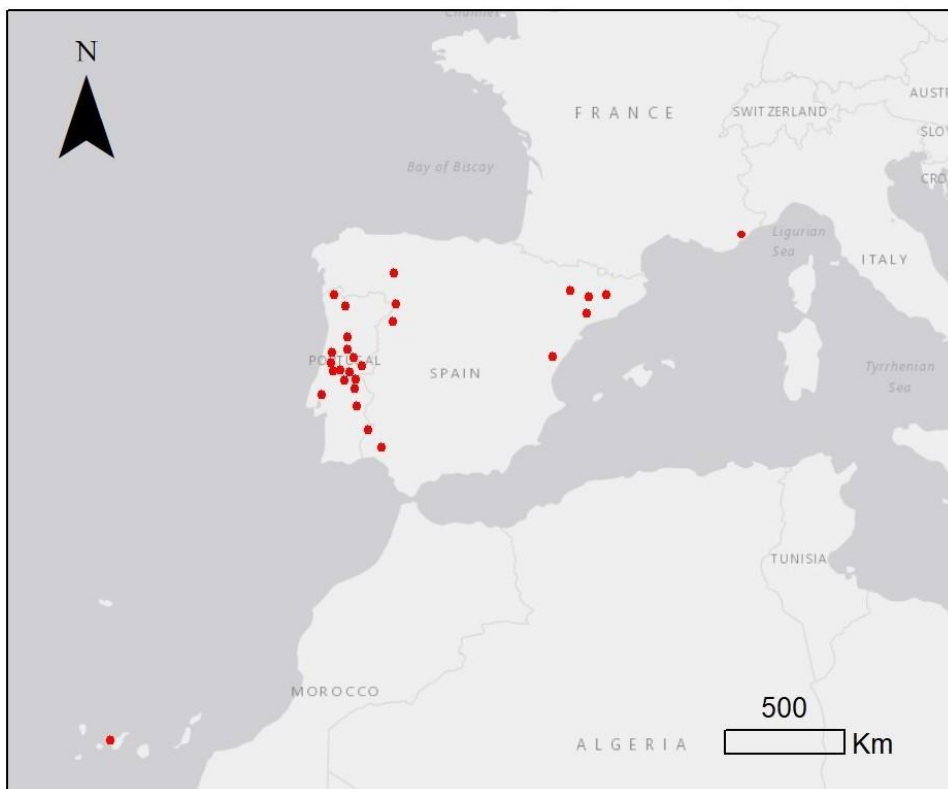
## 4.2 Material and Methods

### 4.2.1 Data Sources

#### 4.2.1.1 Studied fires

In this study, we analysed the area burned by 28 forest fires that occurred along the period 2006–2008 in south-western Europe, entailing the south of France, Portugal and both continental Spain and the Canary Islands (Fig. 1). Overall, fires burned 24,869 ha including agricultural fields, grasslands, shrublands, and forests dominated by conifers, broadleaves, or mixed forests. The fire size ranged from 42.7 ha (Montauroux, France, 2008) to 14,649 ha (Tenerife, Canary Islands, Spain; Table 1). The areas covered a high range of forest types in relation to their tree canopy cover and species composition. The main conifer species in the burned areas were *Pinus pinaster* Ait., *P. halepensis* Mill. and *P. canariensis* Chr.

Sm. Ex DC in Buch, while the most common broadleaves were *Eucalyptus* sp (*Eucalyptus globulus* Labill. and *E. nitens* Deane & Maiden), *Quercus ilex* L. and *Q. suber* L. The area is characterized by annual mean temperatures ranging from 11.9 to 16.9 °C, annual precipitation from 514 to 962 mm and elevation from 61 m a.s.l. to 1521 m, reflecting high variability.



**Fig. 1.** Location of the studied fires in Southern Europe.

#### 4.2.1.2 Field data

Information about the extent of wildfires occurred in the European Union is recorded in a standardized database. In order to add data about burn severity of fires larger than 50 ha, a manual field assessment was created to be used by the European Member States named Field Assessment of Forest Fire Severity (ForFireS). Ground-based burn severity data, measured as the loss of ground and aboveground biomass after the fire, was collected in twenty-eight fires by

surveying a total number of 668 inventory plots following a systematic squared grid. The number of plots varied according to the area burned by each fire, with a mean of 23.2 plots by fire (standard deviation = 5.5). The plots were circular with an area of 500 m<sup>2</sup>.

We created a dichotomic variable to characterize the land use in the selected plots (Forest = 1, forested areas; Forest = 0, agricultural or other areas). For each forested plot, the following variables were measured indicating the state of the plots prior to the fire (measured when possible and reconstructed from stumps and nearby unburned plots when fully combusted) and after the fire (measured) as potential candidates for the statistical analysis (Table 1): mean diameter at breast height (Dm); mean tree height (Hm); maximum, minimum and mean bark char height (CHAR); herbaceous cover after the fire (Herbs); basal area (G); shrub cover (SHRC), ligneous cover (LIG) and tree cover (TREEC) before and after the fire and the corresponding difference (loss) between both periods due to the fire effects; mean slope, aspect and elevation of the plots; dominant tree species and vegetation type (conifer or broadleaved). Note that one of the detailed variables (LIGloss) had discontinued records and was not measured in all plots (N=245).

We also created the dichotomic variable Firetd in order to consider the time between the burn severity assessment and the field work. Firetd is equal to 0 for Initial Assessment plots (IA, plots measured between two or three months after the fire occurrence), and 1 for Extended Assessment plots (EA, plots measured during the next year's campaign).

**Table 1.** Descriptive statistics of the candidate variables considered for the models.

Variable	Unit	Description	Mean	Std. Dev.
Dm	cm	Mean diameter at breast height	12.57	10.67
Hm	dm	Mean tree height	82.77	70.00
CHARmax	dm	Maximum bark char height	57.05	58.03
CHARmin	dm	Minimum bark char height	20.20	24.37
CHARHmean	dm	Mean bark char height	41.58	38.90
CHARH	%	Mean bark char height (% total height)	39.70	40.79
Gpre	m <sup>2</sup> ha <sup>-1</sup>	Basal area before the fire	6.50	9.73
Gpost	m <sup>2</sup> ha <sup>-1</sup>	Basal area after the fire	5.61	9.58
Gloss	%	Loss of basal area during the fire	24.44	39.74
Herbs	%	Herbaceous cover after the fire	12.90	21.06
SHRCpre	%	Total living shrub cover before the fire	45.63	34.51
SHRCpost	%	Total living shrub cover after the fire	7.88	13.43
SHRCloss	%	Loss of shrub cover during the fire	34.53	32.02
LIGpre	%	Total living ligneous cover before the fire	58.47	35.01
LIGpost	%	Total living ligneous cover after the fire	18.58	18.70
LIGloss	%	Loss of ligneous cover during the fire	38.87	31.69
TREECpre	%	Total living tree cover before the fire	28.10	26.26
TREECpost	%	Total living tree cover after the fire	12.02	15.64
TREECloss	%	Loss of tree cover during the fire	15.98	22.10
Slope	degree	Slope	13.72	10.26
Elevation	m	Elevation	460.74	394.36
Aspect	degree	Aspect	177.00	123.29
Sp	-	Dominant Species	-	-
Vegetation type	-	Conifer or Broadleaved plot	-	-
Forest	-	Dichotomic variable (Forest = 0 represents plots located in non-forest areas, and 1 represents plots measured in forest areas)	-	-
Firetd	-	IA plots (initial assessment) Were measured between two or three months after the fire and ET plots (extended assessment) during the next year field campaign	-	-

#### 4.2.2 Satellite image collation and processing and RdNBR calculation

The remote sensing index for assessing burn severity was based on Landsat 5 y 7 satellite imagery, with observations after and before the fires. We collected the closest available images to the fire occurrence and to the field work. Based on these, we computed the Normalized Burn Ratio (NBR) at 30 m resolution using the Landsat NIR and SWIR bands (4 and 7, respectively) and the Relative Differenced Normalized Burn Ratio (RdNBR) (Miller et al., 2009) to assess burn severity in the burned areas. The RdNBR is a variant of the Differenced Normalized Burn Ratio (dNBR) (Key and Benson, 2005) that considers the relative amount of pre- to post-fire change by dividing dNBR by pre-fire NBR value. This index was proposed to remove the bias due to the pre-fire vegetation type and density (Miller et al., 2009). The RdNBR showed a normal distribution although with long tails in both sides; therefore, the values were restricted to -500 to 1500, excluding 18 observations, to ensure a more stable analysis of the relations between the different variables.

#### 4.2.3 Methodological approaches

A preliminary analysis was performed on all the considered variables, in order to identify strong correlations between RdNBR and ground data. The selection of variables included the following criteria: the variables must be available and easy to retrieve, must have a good predictive power, have a physical or biological meaning, and in case they are combined, they should present a limited multi-correlation.

Two modelling approaches were taken to establish the relationships, aiming at different applications: on the one hand, based on linear regression in order to produce models that could be easily applied and understood and, on the other one, based on Boosted Regression Trees (BRT), in order to produce the highest predictive power with the variables available.

The linear regression models were constructed by a mixed model approach, since the variables were grouped according to fire event, reflecting a hierarchical structure. To simplify the modelling methods, the resulting random

factor was applied to the intercept. In addition to the criteria used for variable selection, in this case the variables should have no obvious biases or dependencies in the predictions, had to be significant at the 0.05 level, and the resulting models had to be parsimonious and avoid excessive complexity, since the ultimate objective was to observe the relationship of the variables, and not a full predictive model.

The resulting models under this approach were evaluated quantitatively by examining the magnitude and graphical distribution of the residuals for all the variables included in the model, aiming at finding obvious dependencies that indicate systematic discrepancies. The accuracy of the model predictions was determined by calculating the bias and precision, based on the absolute and relative estimates of bias and RMSE, as well as the coefficient of determination.

With all candidate variables used in the mixed models, an alternative analysis was performed based on BRT. This approach combines statistical and machine learning techniques aiming at the improvement of the performance of a single model by fitting many models and combining them for prediction (Schapire, 2003). In this case, the modelling approach allows the use of incomplete data, does not take any assumption about the shape of the relationship of the variables and the predicted, permitting multiple interactions.

In this case, besides the selection of variables to be included in the models, the models are defined by additional parameters: number of trees, learning rate (or shrinkage, related to the reduction of the impact of any additional tree), bag (random fraction of the residuals is selected to build the tree, per iteration) and number of interactions between variables. The number of trees was estimated through optimization (Ridgeway, 2006) and for the other parameters, several models were tested sequentially, being alternative learning rates fixed ranging from 0.8 to 0.001, and the number of interactions from 1 to 5.

The models and statistical analysis were developed in R version 3.2.4 (R core development team, 2017). The BRT models were based on the *dismo* (Elith et al., 2008) extension of the *gbm* package (Ridgeway, 2006). The mixed models were estimated using the *nlme* package (Pinheiro et al., 2018).



### 4.3 Results

The preliminary analysis showed a significant correlation between the variables considered and the RdNBR. The strongest correlations were for the variable LIGpre and LIGloss ( $R=0.59$  and  $0.57$ , respectively) and CHARH, CHARHmean, TREECloss, SHRCloss ( $R=0.44, 0.40, 0.37, 0.12$ , respectively), with the rest presenting lower correlations. For the dummy variables, both Firetd and Forest showed a high explanatory power (Fig. 2).

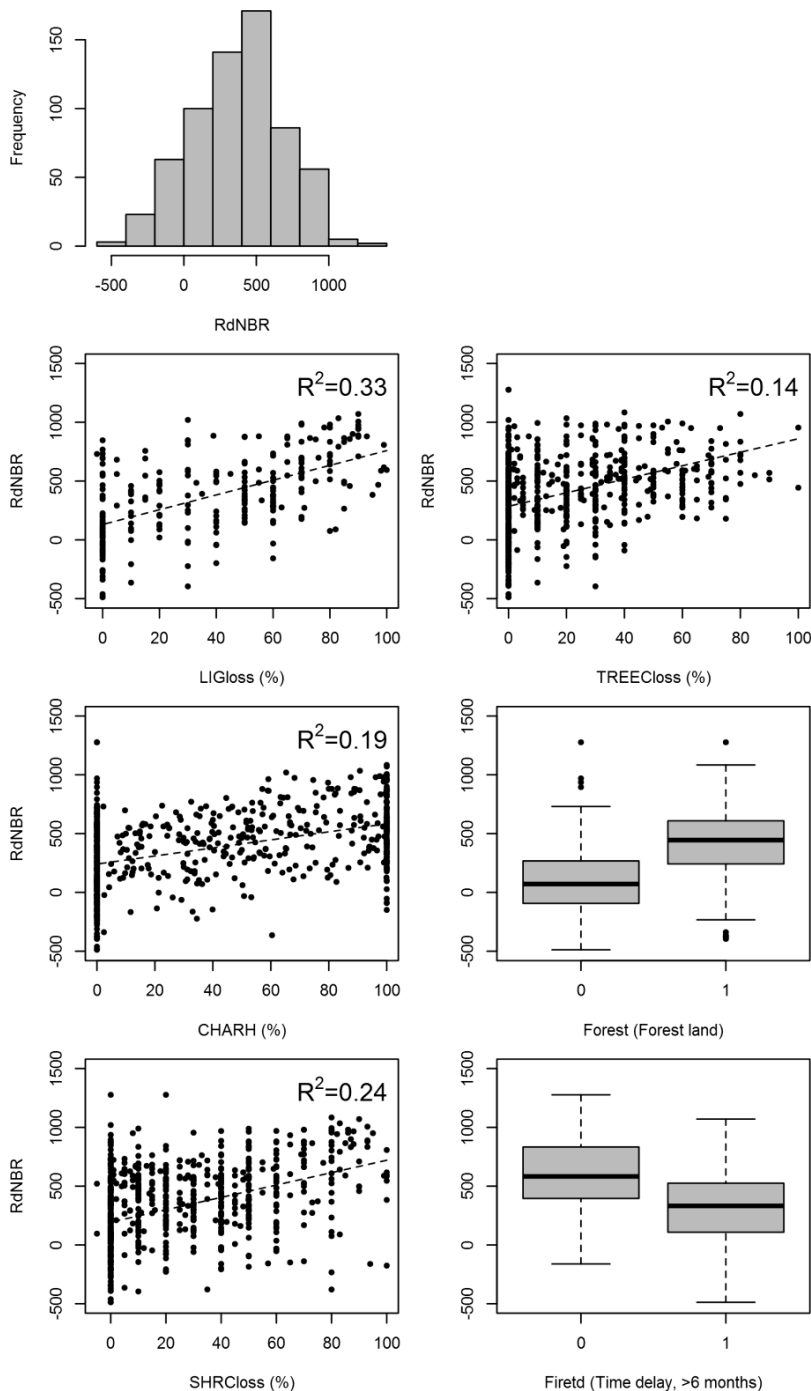
As expected, correlations between the independent variables were found (LIGpre and LIGloss; CHARH and CHARHmean) and, based on this, a subset of variables was selected by considering those variables presenting better results for the statistical models. One of the variables that showed the highest relationships in the preliminary analysis (LIGloss) had discontinued records ( $N=241$ ), which limited the use of the rest of the data since mixed models require complete cases to be fitted. Therefore, two alternative models were fitted, with LIGloss being substituted with surrogate variables ( $N=650$ ).

The final models considered under this approach were:

$$\text{RdNBR}_{ij} = \beta_0 + \beta_1 \text{LIGloss}_{ij} + \beta_2 \text{CHARH}_{ij} + \beta_3 \text{Forest}_{ij} + \beta_4 \text{Firetd}_{ij} + \mu_j + \varepsilon_{ij} \quad [1]$$

$$\text{RdNBR}_{ij} = \beta_0 + \beta_1 \text{SHRCloss}_{ij} + \beta_2 \text{TREECloss}_{ij} + \beta_3 \text{CHARH}_{ij} + \beta_4 \text{Forest}_{ij} + \beta_5 \text{Firetd}_{ij} + \mu_j + \varepsilon_{ij} \quad [2]$$

where RdNBR is the response variable related to LIGloss, TREECloss, SHRCloss, CHARH, Forest, Firetd all as defined in Table 1. Subscripts  $i$  and  $j$  refer to measurement  $i$  in fire  $j$ ,  $\beta_1$ – $\beta_n$  are the parameters to be estimated,  $\mu_j$  and  $\varepsilon_{ij}$  are independent and identically distributed random between-fire and between-measurement factors, with mean equal to 0 and constant variances  $\sigma_{\text{fire}}$  and  $\sigma_{\text{obs}}$ , respectively.

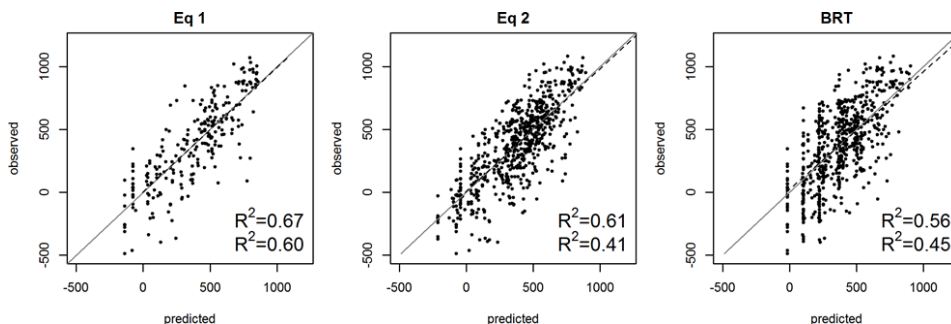


**Fig. 2.** Histogram of the distribution of the frequencies of the RdNBR (Relativized differenced of Normalized Burn Ratio) values, (between -500 and 1500, used as thresholds in the study). Relationships between RdNBR and CHARH (%), TREECloss (%), LIGloss (%), SHRCloss (%), Forest and Firetd.

All variables included in the linear mixed-effects models (Eq1 and Eq2) were statistically significant at  $p\text{-value} < 0.05$ . The linear regression mixed models presented a coefficient of determination for the random and fixed parts of 0.67 and 0.61, for Eq1 and Eq 2, respectively. When only considering the fixed part, these values decreased to 0.60 and 0.41, respectively, as a large percentage of the variability was explained through the random factor (Table 2), reflecting the differences due to specific conditions related to the fire event. Concerning the BRT approach, the best model used a learning rate of 0.01, 5 interactions and 550 trees. In this case, the coefficient of determination was 0.56, but it was reduced to 0.45 when it was estimated using a cross-validation approach (Fig. 3).

**Table 2.** Estimates, standard error (S.E.) and significance level of the parameters and variance components of the mixed models (Eq 1 and Eq 2) to predict the relativized difference of normalized burn ratio (RdNBR). BRT: predictions based on boosted regression trees. The weight for each variable is the average for models with 1, 3 and 5 interactions (tree complexity,  $N=3$ ), for a learning rate 0.01.

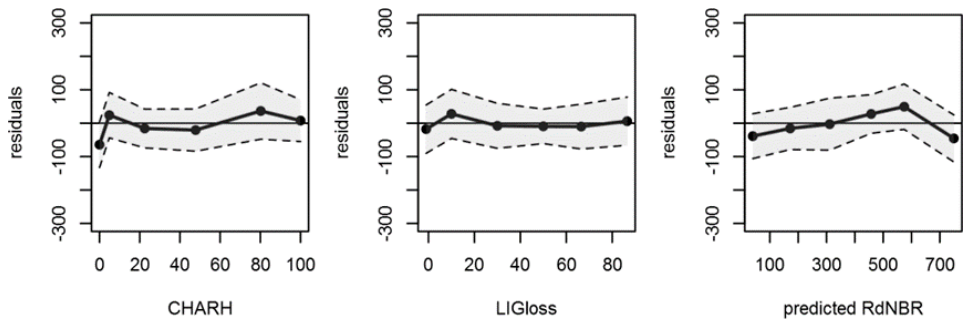
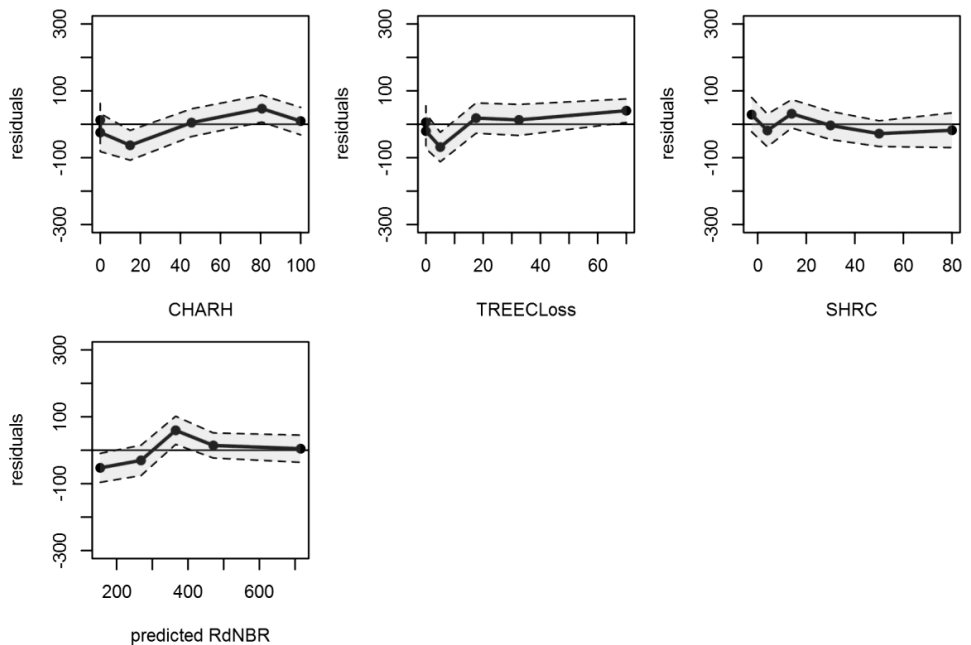
Variable	Value	Std. Error	DF	t-value	p-value
<b>Eq 1</b>					
$\beta_0$	321	62.53	229	5.1415	< 0.001
CHARCH	3.055	0.479	229	6.380	< 0.001
LIGloss	2.568	0.612	229	4.198	< 0.001
Forest	-109.854	45.722	229	-2.403	0.017
Firetd	-252.589	72.841	7	-3.468	0.010
$\sigma_{\text{fire}}$	100.68				
$\sigma_{\text{measurement}}$	201.28				
<b>Eq 2</b>					
$\beta_0$	429.1	56.25	618	7.627	< 0.001
CHARCH	1.042	0.276	618	3.768	< 0.001
TREECloss	3.060	0.490	618	6.246	< 0.001
SHRCloss	1.663	0.404	618	4.110	< 0.001
Forest	-82.910	28.57	618	-2.901	< 0.004
Firetd	252.406	62.09	26	4.064	< 0.001
$\sigma_{\text{fire}}$	139.83				
$\sigma_{\text{measurement}}$	203.08				
<b>BRT</b>					
	<b>Weight (mean)</b>	<b>Std. Error</b>			
CHARCH	28.9 %	1.5			
Firetd	28.4 %	2.2			
TREECloss	19.1 %	1.0			
LIGloss	11.4 %	0.6			
Forest	7.6 %	1.4			
SHRCloss	4.6 %	1.7			



**Fig. 3** Observed versus predicted estimates for the two mixed models: Eq 1 ( $N = 241$ ) and Eq 2 ( $N = 650$ ) and estimates based on Boosted Regressions Trees (BRT).

Concerning model assessment, there was not observed strong systematic biases in the predicted RdNBR and the values of the variables. (Fig. 4), although in Eq2, the use of TREECloss tend, to some extent, underestimate the RdNBR. The bias was  $-2.749$  ( $-0.72\%$ ) and  $-1.471$  ( $-0.38\%$ ) for Eq1 and Eq2, respectively, and  $-0.045$  ( $-0.012\%$ ) for the BRT model (RMSE= $218, 245$  and  $211$ , for Eq1, Eq2 and BRT, respectively). In the case of the BRT approach, it failed systematically to produce RdNBR estimates below  $-67$ .

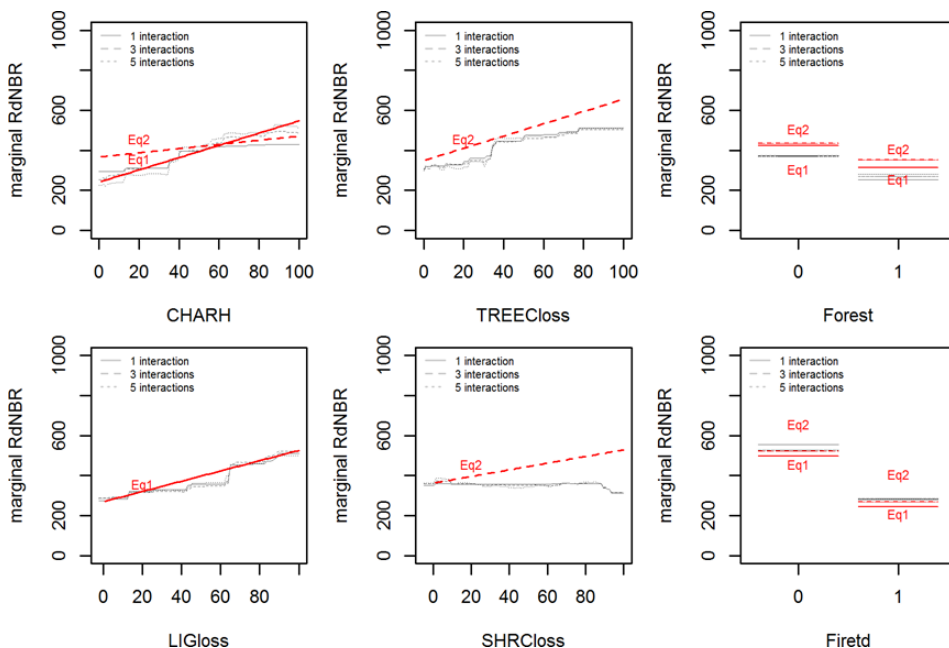
The marginal effect on RdNBR of each single variable included in the statistical analysis was determined from partial dependency plots, which showed the effect of each variable on the response after accounting for the average effects of all other variables for the linear mixed models and with one, three and five interactions for BRT. All relationships were largely consistent along the modelling approaches (Fig. 5). The effects on RdNBR increased with increasing values of CHARH and LIGloss, following a linear trend. A similar effect was found considering TREECloss, although in this case, the values of Eq2 slightly resulted in higher RdNBR than in the BRT approach. For SHRCloss there were divergences between the mixed models, which accounted for a positive relationship, and the BRT estimates, which identified a negative threshold after 80%. In the BRT estimates, alternative values in the number of interactions, learning rates or number of trees used in the calibration, did not alter the nature of the relationships.

**Eq 1****Eq 2**

**Fig. 4** Mean residuals as a function of the variables included in the fixed part of Eq 1 and Eq 2. The residuals have been grouped in 6 tiles of equal number of observations. The figure represents the mean residual value as well as two time the standards errors (discontinuous lines). Predicted: relativized difference of normalized burn ratio (RdNBR).

Regarding the dichotomous variables Forest and Firetd, the effects on the independent variable were higher in plots measured in forest areas between two or three months after the fire occurrence. The estimated contributions of these

variables (Table 2) were highest for CHARCH, TREECloss and LIGloss (average over 20%) and the lowest for SHRCloss (average less than 5%), suggesting the explanatory effects of the latter could be explained by combinations of the other variables. It must be taken into account that the weights slightly changed with the number of interactions, which was particularly affecting the weights of Forest (decreasing till 6% for 5 interactions) and SHRCloss (increasing till 7.3% for 5 interactions).



**Fig. 5** Partial dependence plots of the relativized difference of normalized burn ratio (RdNBR) in the different models tested. The light grey lines result from boosted regression trees (BRT) constructed with 1, 3 and 5 interactions, whereas the straight lines result from two linear mixed models (Eq 1 and Eq2). LIGloss, included in Eq 1 and in the BRT models, presents lower observations (N=241) than the other variables (N=650).

#### 4.4 Discussion

The present study addresses the empirical relationships between a burn severity index derived from remote sensing and field data based on 28 fire events in several countries. Due to the large areas entailed by fires, the use of Landsat based indices fits the purposes of the analysis, as their spectral and spatial resolutions make them particularly suitable for assessing burn severity (Epting et

al., 2005; Quintano et al., 2018; Soverel et al., 2010; Van Wagtendonk et al., 2004). Particularly, the RdNBR has been used to a large success (De Santis, 2008; French et al., 2008; Keeley, 2009). But for their use, it is crucial to calibrate it with field variables, in order to analyze the performance of remote-sensing outputs (French A et al., 2008; Miller et al., 2009) and to assess their accuracy. Most of studies combining field data with remotely sensed assessments of burn severity used the CBI, or adaptations like the GeoCBI, as these indices were designed to define burn severity ecologically, and to measure ground effects visually at each vegetative strata on large and heterogeneous plots. The results of these relationships were statically satisfactory in most of these studies.

However, the integration of multiple variables in a combined index makes the specific assessment of these relationships difficult. In this study, we aimed to analyse the relations of the different specific field variables (based on the ForFireS assessment) in order to better understand the information that these indices are reflecting, therefore using RdNBR as the response variable subject to study. As expected, the variables reflecting fire intensity (as mean bark char height, CHARH), or burn severity, (as loss of tree, ligneous and shrub cover, TREEClass, LIGloss and SHRCross, respectively), were positively related with RdNBR, indicating a loss of above ground biomass and reduced photosynthetic reflectance by living biomass, similar to other studies comparing field indicators and satellite indexes on burn severity (Miller et al., 2009; Miller and Quayle, 2015). Topographical factors, that were not significant in this study, have clear influence on fire behaviour although the association between topography and burn severity cannot be generalized for all ecosystems and regions (Dragozi et al., 2016; Gonzalez-Olabarria et al., 2006; González et al., 2007) due to the prevalence of determined local conditions on the studied regions (for instance, variation on humidity between different elevations and aspects), or how the slope influences fire behaviour (higher slopes increase the rate of spread and intensity, but usually reduce the residence time of fire). Burn severity is expected to be more severe in coniferous forests than in broadleaved forests (Fernandes et al., 2010) but in our study it did not appear as significant variable. The pre-fire tree species composition could play an important role in determining post-fire reflectance and recovery and, subsequently, burn severity (Dragozi et al., 2016). It is clear that, depending on the fire intensity and the specific response of the tree species to fire, different ecosystem responses such as resprouting and regeneration can be expected or be accentuated as time goes on (Keeley et al., 2008). Such variation on vegetation characteristics not only defines an immediate response to fire, but also ecological responses overtime (Keeley et al.,

2008; Liu, 2016). Even more, the season of the year when each fire takes place is expected to influence initial fire severity, and delayed vegetation response (Knapp and Keeley, 2006; Valor et al., 2017).

The results confirm that remote sensing indexes provide an overall acceptable assessment of fire included impact on forest vegetation, which can be translated into estimates of biomass lost even when predictive models are applied across large and heterogeneous regions. Yet it must be taken into account that the interactions between fire seasons, vegetation types and ecological responses, among others, would require of adjustment factors. This specific adjustment was addressed by the use of random factor, which included specific conditions of each fire in a mixed-effects approach. In fact, the statistical analysis showed the important differences between fire events on the RdNBR response, suggesting that even if general relations can be inferred for any type of forest, specific variations on such relations are to be expected for each fire event as it is characterized by a specific behaviour, due to topography, weather, and fuel arrangement (McRae et al., 2005; Estes et al., 2017). The vegetation composition and development stage prior to fires is also expected to vary between fires, especially when fires are allocated across a large and highly variable region such as the South Europe plus the Canary Islands (Valavanidis and Vlachogianni, 2011).

In addition to the general differences of the RdNBR response depending on field measured burn severity, the use of BRT showed that the relationships are not necessary linear and that there are several interactions between the measurements of severity and their reflection in the RdNBR values. BRT is a useful statistical technique for analysis of ecological data (De'Ath, 2007) because non-linear effects and interactions are mechanically considered due to trees are invariant to transformations of the predictors and interactions are automatically included. This is very suitable for complex analysis as it allows addressing the relationship among the explanatory variables and the predicted variables with nearly no assumptions. Although the results of the used of BRT were largely consistent with the mixed models approach, this analysis revealed that the relationship with loss of shrub cover during the fire could be over-estimated in the models. It must notice that the use of BRT does not account for between-fire differences in the same way that is handled by a mixed model.



Finally, a factor influencing the relation between ground measurements and RdNBR were the time lapse between the fire event and the remotely sensed burn severity assessment. The impact of the elapsed time could be derived from the RdNBR sensitivity to ash cover which declines with time since fire (Miller and Quayle, 2015) and the regeneration and response of vegetation after the fire. In that sense, as the time since fire increases, the usage of fire severity measures taken on the field to generate burn severity indexes and relate them to satellite derived images may arise problems, due to the concept of burn severity itself and the increased impact of ecological responses (Keeley, 2009). In order to reduce this time, Quintano et al. (2018) combined the use of Landsat 8 OLI and Sentinel-2A MSI data with the dNBR index and its relative version because Sentinel-2A MSI provides higher spatial and temporal resolutions and can be used in case of unavailability of a pre-fire or a post-fire Landsat image. Although the authors showed that this approach can be adequate for postfire management intervention, the accuracy of burn severity maps was lower than the maps exclusively created with Landsat imagery (Quintano et al., 2018).

The results of this study provide an assessment of relevant variables that should be prioritized when designing field inventories to assess burn severity, especially when a strong relation is found between satellite derived data and the real impacts of fire on forest. Note that fire may also damage soils, a variable that was not measured and correlated in this study. The study has confirmed that the time lapse between the fire event and the severity assessment influences not only the correlation but the interpretation of the severity concept, and therefore, a careful adjustment of the time lapse must be performed according to the objectives of the analysis. On late assessments, the growth of herbs and bushes may play down estimation on tree mortality, whereas early assessments can miss relevant processes, as delayed mortality, plant recovery through resprouting, or recruitment of smaller plants protecting soil against erosion. A satellite assessment on fire severity should be therefore planned according to the expected use of the obtained information, and always try to reduce potential misinterpretations by evaluating the post-fire response traits of the affected species, either trees, bushes or herbs.

## 4.5 References

- Adam, E., Mutanga, O., Rugege, D., 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. *Wetl. Ecol. Manag.* 18, 281–296. doi:10.1007/s11273-009-9169-z
- Butler, B.W., Dickinson, M.B., 2010. Tree injury and mortality in fires: Developing process-based models. *Fire Ecol.* 6, 55–79. doi:10.4996/fireecology.0601055
- Cardil, A., Eastaugh, C.S., Molina, D.M., 2015. Extreme temperature conditions and wildland fires in Spain. *Theor. Appl. Climatol.* 122, 219–228. doi:10.1007/s00704-014-1295-8
- Cardil, A., Salis, M., Spano, D., Delogu, G., Molina Terrén, D., 2014. Large wildland fires and extreme temperatures in Sardinia (Italy). *iForest - Biogeosciences For.* 7, 162–169. doi:10.3832/for1090-007
- Catry, F.X., Rego, F., Moreira, F., Fernandes, P.M., Pausas, J.G., 2010. Post-fire tree mortality in mixed forests of central Portugal. *For. Ecol. Manage.* 260, 1184–1192. doi:10.1016/j.foreco.2010.07.010
- Chuvieco, E., 2012. Remote sensing of large wildfires: in the European Mediterranean Basin. Springer Science & Business Media.
- Cocke, A.E., Fulé, P.Z., Crouse, J.E., 2005. Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data. *Int. J. Wildl. Fire* 14, 189. doi:10.1071/WF04010
- Czaplewski, R.L., Patterson, P.L., 2003. Classification Accuracy for Stratification with Remotely Sensed Data. *For. Sci.* 49, 402–408. doi:10.1093/forestscience/49.3.402
- De'Ath, G., 2007. Boosted trees for ecological modeling and prediction. *Ecology* 88, 243–251
- De Santis, A., 2008. Burn severity estimation from remotely sensed data using simulation models. *Geography* 138.
- De Santis, A., Chuvieco, E., 2009. GeoCBI: A modified version of the Composite Burn Index for the initial assessment of the short-term burn severity from remotely sensed data. *Remote Sens. Environ.* 113, 554–562. doi: <https://doi.org/10.1016/j.rse.2008.10.011>
- Dragozi, E., Gitas, I.Z., Bajocco, S., Stavrakoudis, D.G., 2016. Exploring the relationship between burn severity field data and very high resolution GeoEye images: The Case of the 2011 Evros wildfire in Greece. *Remote Sens.* 8. doi:10.3390/rs8070566
- Elith, J., Leathwick, J., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813.
- Epting, J., Verbyla, D., Sorbel, B., 2005. Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. *Remote Sens. Environ.* 96, 328–339.
- Estes, B.L., Knapp, E.E., Skinner, C.N., Miller, J.D., Preisler, H.K., 2017. Factors influencing fire severity under moderate burning conditions in the Klamath Mountains, northern California, USA. *Ecosphere* 8. doi:10.1002/ecs2.1794

- Fernandes, P.M., Luz, A., Loureiro, C., 2010. Changes in wildfire severity from maritime pine woodland to contiguous forest types in the mountains of northwestern Portugal. *For. Ecol. Manage.* 260, 883–892. doi:<https://doi.org/10.1016/j.foreco.2010.06.008>
- Fernández, A., Illera, P., Casanova, J.L., 1997. Automatic mapping of surfaces affected by Remote Sens. *Environ.* 60, 153–162. doi:[https://doi.org/10.1016/S0034-4257\(96\)00178-2](https://doi.org/10.1016/S0034-4257(96)00178-2)
- Franklin, S., 2001. *Remote Sensing for Sustainable Forest Management*. CRC Press; 1 edition.
- French A, N.H.F., Kasischke, E.S., Halle, R.J., Murphy, K.A., Verbyla, D.L., Hoy, E.E., Allen, J.L., 2008. Using landsat data to assess fire and burn severity in the North American boreal forest region: an overview and summary of results. *Int. J. Wildl. Fire* 17, 443–462.
- Gonzalez-Olabarria, J.R., Palahí, M., Trasobares, A., Pukkala, T., 2006. A fire probability model for forest stands in Catalonia (north-east Spain). *Ann. For. Sci.* 63, 169–176. doi:[10.1051/forest](https://doi.org/10.1051/forest)
- González, J.R., Trasobares, A., Palahí, M., Pukkala, T., 2007. Predicting stand damage and tree survival in burned forests in Catalonia (North-East Spain). *Ann. For. Sci.* 64, 733–742. doi:[10.1051/forest:2007053](https://doi.org/10.1051/forest:2007053)
- Grafström, A., Ringvall, A.H., 2013. Improving forest field inventories by using remote doi:[10.1139/cjfr-2013-0123](https://doi.org/10.1139/cjfr-2013-0123)
- Harris, L., Taylor, A.H., 2015. Topography, Fuels, and Fire Exclusion Drive Fire Severity of the Rim Fire in an Old-Growth Mixed-Conifer Forest, Yosemite National Park, USA. *Ecosystems* 18, 1192–1208. doi:[10.1007/s10021-015-9890-9](https://doi.org/10.1007/s10021-015-9890-9)
- Harris, S., Veraverbeke, S., Hook, S., 2011. Evaluating spectral indices for assessing fire severity in chaparral ecosystems (Southern California) using modis/aster (MASTER) airborne simulator data. *Remote Sens.* 3, 2403–2419. doi:[10.3390/rs3112403](https://doi.org/10.3390/rs3112403)
- Herrando, S., Brotons, L., 2002. Mediterranean areas affected by wildfires: a multi-scale approach. *Ecography.* 25(2), 161–172. Doi:[10.1034/j.1600-0587.2002.250204.x](https://doi.org/10.1034/j.1600-0587.2002.250204.x)
- Herrando, S., Brotons, L., & Llacuna, S. (2003). Does fire increase the spatial heterogeneity of bird communities in Mediterranean landscapes?. *Ibis*, 145(2), 307–317. doi:[10.1046/j.1474-919X.2003.00155.x](https://doi.org/10.1046/j.1474-919X.2003.00155.x)
- Jones, H., Vaughan, R., 2010. *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press.
- Kalabokidis, K., Palaiologou, P., Gerasopoulos, E., Giannakopoulos, C., Kostopoulou, E., Zerefos, C., 2015. Effect of climate change projections on forest fire behavior and values-at-risk in southwestern Greece. *Forests* 6, 2214–2240. doi:[10.3390/f6062214](https://doi.org/10.3390/f6062214)
- Keeley, J.E., 2009. Fire intensity, fire severity and burn severity: A brief review and suggested usage. *Int. J. Wildl. Fire* 18, 116–126. doi:[10.1071/WF07049](https://doi.org/10.1071/WF07049)
- Keeley, J.E., Brennan, T.J., Pfaff, a H., 2008. Fire severity and ecosystem responses following crown fires in California shrublands. *Ecol. Appl.* 18, 1530–1546. doi:[10.1890/07-0836.1](https://doi.org/10.1890/07-0836.1)

- Key, C., Benson, N., 2005. Landscape assessment: remote sensing of severity, the normalized burn ratio and ground measure of severity, the composite burn index. FIREMON: Fire effects monitoring and inventory system. USDA Forest Service, Rocky Mountain Res. Station 164.
- Knapp, E.E., Keeley, J.E., 2006. Heterogeneity in fire severity within early season and late season prescribed burns in a mixed-conifer forest. *Int. J. Wildl. Fire* 15, 37–45. doi:10.1071/WF04068
- Kushlat, J.D., Ripple, W.J., 1998. Assessing wildfire effects with Landsat thematic mapper data Space-borne sensors have been used for the past two decades to observe the effects of forest fires. *Landsat Thematic Mapper (TM) imagery was used to evaluate the effects of the Yellowstone fi* 19, 2493–2507.
- Lefsky, M., Cohen, W., Parker, G., Harding, D., 2002. Lidar Remote Sensing for Ecosystem Studies. *Bioscience* 52, 1.
- Lentile, L.B., Holden, Z.A., Smith, A.M.S., Falkowski, M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E., Benson, N.C., 2006. Remote sensing techniques to assess active fire characteristics and post-fire effects. *Int. J. Wildl. Fire* 15, 319–345. doi:10.1071/WF05097
- Liu, Z., 2016. Effects of climate and fire on short-term vegetation recovery in the boreal larch forests of Northeastern China. *Sci. Rep.* 6, 1–14. doi:10.1038/srep37572
- Malone, S.L., Kobziar, L.N., Staudhammer, C.L., Abd-Elrahman, A., 2011. Modeling relationships among 217 fires using remote sensing of burn severity in southern pine forests. *Remote Sens.* 3, 2005–2028. doi:10.3390/rs3092005
- Michaletz, S.T., Johnson, E.A., 2007. How forest fires kill trees: A review of the fundamental biophysical processes. *Scand. J. For. Res.* 22, 500–515. doi:10.1080/02827580701803544.
- Martín-Alcón, S., Coll, L., 2016. Unraveling the relative importance of factors driving post fire regeneration trajectories in non-serotinous *Pinus nigra* forests. *For. Ecol. Manage.*, 361, 13–22. doi.org/10.1016/j.foreco.2015.11.006
- McRae, D.J., Jin, J.Z., Conard, S.G., Sukhinin, A.I., Ivanova, 534 G.A., Blake, T.W., 2005. Infrared characterization of fine-scale variability in behavior of boreal forest fires. *Canadian Journal of Forest Research* 35.9 2194–2206. doi: 10.1139/x05-096
- Miller, J.D., Knapp, E.E., Key, C.H., Skinner, C.N., Isbell, C.J., Creasy, R.M., Sherlock, J.W., 2009. Remote Sensing of Environment Calibration and validation of the relative differenced Normalized Burn Ratio (RdNBR) to three measures of fire severity in the Sierra Nevada and Klamath Mountains. *Remote Sens. Environ.* 113, 645–656. doi:10.1016/j.rse.2008.11.009
- Miller, J.D., Quayle, B., 2015. Calibration and validation of immediate post-fire satellite-derived data to three severity metrics. *Fire Ecol.* 11, 12–30.
- Miller, J.D., Yool, S.R., 2002. Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. *Remote Sens. Environ.* 82, 481–496. doi:Pii S0034-4257(02)00071-8

- Morgan, P., Keane, R.E., Dillon, G.K., Jain, T.B., Hudak, A.T., Karau, E.C., Sikkink, P.G., Holden, Z.A., Strand, E.K., 2014. Challenges of assessing fire and burn severity using field measures, remote sensing and modelling. *Int. J. Wildl. Fire* 23, 1045–1060.
- Parks, S.A., Dillon, G.K., Miller, C., 2014. A new metric for quantifying burn severity: The relativized burn ratio. *Remote Sens.* 6, 1827–1844. doi:10.3390/rs6031827
- Pereira, J.M.C., 1999. A comparative evaluation of NOAA/AVHRR vegetation indexes for burned surface detection and mapping. *IEEE Trans. Geosci. Remote Sens.* 37, 217–226. doi:10.1109/36.739156
- Picotte, J.J., Peterson, B., Meier, G., Howard, S.M., 2016. 1984–2010 trends in fire burn severity and area for the conterminous US. *Int. J. Wildl. Fire* 25, 413–420.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., 2018. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–137.
- Quintano, C., Fernández-Manso, A., Fernández-Manso, O., 2018. Combination of Landsat and Sentinel-2 MSI data for initial assessing of burn severity. *Int. J. Appl. Earth Obs. Geoinf.* 64, 221–225. doi:https://doi.org/10.1016/j.jag.2017.09.014
- R core development team, 2017. R 3.2.4.
- Ridgeway, G., 2006. Generalized boosted regression models. Documentation on the R Package 'gbm', version 1.5–7.
- Roy, D.P., Lewis, P.E., Justice, C.O., 2002. Burned area mapping using multi-temporal moderate spatial resolution data—a bi-directional reflectance model-based expectation approach. *Remote Sens. Environ.* 83, 263–286. doi:10.1016/S0034-4257(02)00077-9
- Salis, M., Ager, A.A., Arca, B., Finney, M.A., Bacciu, V., Duce, P., Spano, D., 2013. Assessing exposure of human and ecological values to wildfire in Sardinia, Italy. *Int. J. Wildl. Fire* 22, 549–565. doi:10.1071/WF11060
- Schapire, R., 2003. The boosting approach to machine learning – an overview., in: Denison, D., Hansen, M., Holmes, C., Mallick, B., Yu, B. (Eds.), *MSRI Workshop on Nonlinear Estimation and Classification*. Springer, New York.
- Soverel, N.O., Perrakis, D.D.B., Coops, N.C., 2010. Estimating burn severity from Landsat dNBR and RdNBR indices across western Canada. *Remote Sens. Environ.* 114, 1896–1909. doi:10.1016/j.rse.2010.03.013
- Valavanidis, A., Vlachogianni, T., 2011. Ecosystems and biodiversity hotspots in the Mediterranean basin threats and conservation efforts. *Sci Adv Environ Toxicol Ecotoxicol.* 10, 1–24.
- Valor, T., Olabarria, J., Pique, M., Casals, P., 2017. The effects of burning season and severity on the mortality over time of *Pinus nigra* spp. *salzmannii* (Dunal) Franco and *P. sylvestris* L. *For. Ecol. Manage.* 406, 172–183.

- Van Wagtendonk, J.W., Root, R.R., Key, C.H., 2004. Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sens. Environ.* 92, 397–408. doi:10.1016/j.rse.2003.12.015
- Veraverbeke, S., Lhermitte, S., Verstraeten, W.W., Goossens, R., 2011. Evaluation of pre/post-fire differenced spectral indices for assessing burn severity in a Mediterranean environment with Landsat Thematic Mapper. *Int. J. Remote Sens.* 32, 3521–3537. doi:10.1080/01431161003752430
- Veraverbeke, S., Lhermitte, S., Verstraeten, W.W., Goossens, R., 2010. The temporal dimension of differenced Normalized Burn Ratio (dNBR) fire/burn severity studies: The case of the large 2007 Peloponnese wildfires in Greece. *Remote Sens. Environ.*
- Warner, T.A., Skowronski, N.S., Gallagher, M.R., 2017. High spatial resolution burn severity mapping of the New Jersey Pine Barrens with WorldView-3 near-infrared and shortwave infrared imagery. *Int. J. Remote Sens.* 38, 598–616. doi:10.1080/01431161.2016.1268739
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. *J. Plant Ecol.* 1, 9–23. doi:10.1093/jpe/rtm005







## Capítulo 5

---

Discusión general, Futuras líneas de trabajo y Conclusiones

*“El tiempo invertido en los arboles  
nunca es una pérdida de tiempo”*

Katrina Mayer

## Capítulo 5. Discusión general, Futuras líneas de trabajo y Conclusiones

---

Este capítulo presenta una valoración general de las aportaciones realizadas sobre la aplicación de diversas fuentes de información remota en la modelización de los recursos forestales a diferentes escalas en la presente tesis doctoral. Los trabajos realizados han generado herramientas eficientes para la evaluación de estos recursos que permiten ser discutidas de forma general y exponer una serie de conclusiones como síntesis de las que ya han sido mostradas en cada uno de los capítulos de forma independiente, y de las cuales también se derivan posibles líneas de trabajo futuras. Hay que recalcar que esta tesis se enmarca en la empresa föra forest technologies a través del Programa de Investigación de Doctorados Industriales del Ministerio de Economía y Competitividad del Gobierno de España. Así, todas las aportaciones de los trabajos han sido integradas dentro del objetivo principal del programa, es decir, transferir el conocimiento generado en la propia empresa.

### 5.1 Discusión general

La combinación de datos procedentes de distintas fuentes de información, con diferentes niveles de resolución espacial y espectral, ha demostrado un gran potencial para la evaluación de recursos forestales. Las herramientas generadas han sido efectivas para la resolución de diferentes problemas: la identificación de especies a nivel de árbol individual; la clasificación de diferentes tipos de masa en bosques de cobertura completa y; la evaluación de la capacidad de determinar la severidad tras el paso de un incendio.

La integración de diferentes fuentes de información incrementa el poder de los modelos generados (Koch 2011). En el caso del capítulo 2, la información LiDAR de baja resolución ha permitido la delineación automática de las copas de los árboles individuales, pero la resolución espacial de la información combinada debe estar acorde con la escala y la unidad mínima de trabajo (Clark et al., 2005). A la escala de árbol individual, la resolución espacial de la información espectral es crítica. Una resolución muy baja generaría ruido por incluir gran cantidad de información externa al objeto, si bien un exceso de resolución espacial podría generar demasiada variabilidad dentro del objeto (Ghosh et al., 2014).

La recurrencia temporal de la información procedente de sensores remotos permite abordar cambios en los sistemas forestales, tales como la evolución de la cobertura tras la ocurrencia de una perturbación (Chuvioco, 2012; Lentile et al., 2006) o el estado de los ecosistemas y sus recursos (Keeley, 2009). El hecho de estudiar efectos alejados en el tiempo implica recurrir a repositorios (como el de Landsat) que combinen buena resolución espectral y largas series temporales, si bien esto implica cierto sacrificio en cuanto a resolución espacial. En fenómenos ante los que el sistema reacciona con rapidez, el tiempo transcurrido entre el evento a estudiar y la toma de datos puede ser clave en la interpretación del efecto del fenómeno (Harris et al., 2011). La disponibilidad de información periódica es crítica si el objetivo principal es evaluar cambios en la cobertura (capítulo 4) o discriminar coberturas que cambian drásticamente según la estación del año en la que nos encontremos. Por ello, es importante recoger el efecto del paso del tiempo sobre la cobertura, por ejemplo, incluyendo variables que representen esa temporalidad (capítulo 4) o, incluyendo series temporales de imágenes y datos LiDAR en varios momentos en el tiempo con el objetivo de recoger la respuesta de la vegetación a la estacionalidad o el crecimiento (capítulo 3). Si bien es cierto que la recurrencia de la toma de datos, por ejemplo, de LiDAR aerotransportado, tiene un coste económico elevado, ha de ser valorada la posibilidad de realizar vuelos con mayor recurrencia temporal a costa de sacrificar superficie a volar.

La resolución espectral es otro aspecto clave a tener en consideración. Si bien grandes clases, como el agua o la vegetación se pueden identificar con facilidad utilizando amplios rangos de longitud de onda, cuando se persigue un análisis más detallado, como pueden ser diferentes tipos de masa o de especies, su discriminación es más compleja, requiriendo de rangos de longitud de onda más estrechos. En estos casos, la disponibilidad de información con alta resolución espectral, como es el caso de las imágenes hiperespectrales, puede ser un limitante (Dalponte et al., 2014). Así, el hecho de ampliar los rangos del espectro puede ayudar a detectar con mayor precisión diferencias en la sensibilidad espectral de los componentes de la vegetación, permitiendo identificar aspectos adicionales como la variación en la densidad del dosel, procesos de senescencia o fenómenos de estrés (Li et al., 2018; Warner et al., 2017). En este sentido, es interesante el potencial del índice NDVI-705, que incluye la respuesta en la banda del red-edge, recogiendo la reflectancia de la vegetación en el rango en el que la clorofila absorbe gran parte de la luz, aportando

diferencias en la separabilidad de las clases (capítulo 3). Del mismo modo, es interesante la estrecha relación demostrada, mediante modelos mixtos, entre el índice RdNBR y ciertas variables fácilmente medibles, entre las que destacan la altura de quemado en la corteza y el lapso de tiempo entre el incendio y su evaluación. Dichos modelos recogen con mayor eficiencia, la importancia de la variabilidad entre incendios y el impacto inducido por el fuego en la vegetación (capítulo 4).

El análisis de grandes cantidades de datos requiere de algoritmos cada vez más sofisticados. En este sentido, se ha desarrollado una amplia gama de algoritmos de aprendizaje automático (Nitze et al., 2012; Vega Isuhuaylas et al., 2018; Xu et al., 2018), tanto para resolver problemas de clasificación, como de regresión. La precisión de este tipo de modelos suele ser alta. Sin embargo, la capacidad de explicación de la función modelada por estos algoritmos “caja negra” puede verse limitada por su complejidad. Frente a esto, cabe la alternativa de simplificar este tipo de modelos o buscar otros, más sencillos y fácilmente entendibles por los gestores. Así, se puede llevar a cabo una selección previa de variables con el objetivo de evitar sobreajuste, y reducir la dimensionalidad, lo que puede favorecer la comprensión de estos modelos. Esta simplificación puede realizarse de forma automática (Domingo et al., 2018), a través de herramientas como VSURF (Genuer et al., 2015), o mediante criterio experto, basado en el estudio de las matrices de correlación entre variables (Domingo et al., 2017). En este sentido, los resultados de esta tesis parecen indicar que las variables descriptivas de la estructura o del crecimiento suelen tener mayor importancia en los modelos y, por tanto, aportan mayor poder discriminatorio que las variables espectrales, aunque la información espectral multitemporal también aporta valor, siempre que exista cierta separabilidad espectral entre los grupos.

Un aspecto adicional que considerar es la elección entre modelos paramétricos y no paramétricos. A pesar de que los modelos paramétricos son más rápidos y requieren de menos datos, su poder predictivo puede verse limitado, ya que diversos fenómenos ecológicos no responden linealmente (Park et al., 2015; Wu et al., 2018). Por contra, los modelos basados en aprendizaje automático son más complejos y lentos, sobre todo, debido al proceso de entrenamiento al que deben de ser sometidos, pero permiten obtener predicciones más robustas (De'Ath 2017; Zhu et al., 2017). En algunos casos el uso de métodos de regresión

clásicos puede permitir la obtención de mejores ajustes frente a los modelos no paramétricos, como es el caso del capítulo 4, donde se enfrentan modelos mixtos y modelos de aprendizaje automático (BRT) y donde se observa claramente el efecto de la variabilidad entre incendios. En todo caso, también debe tenerse en cuenta que el efecto del tamaño de la muestra puede ser significativo en cuanto al poder predictivo de modelos paramétricos y no paramétricos. La elección de la técnica más adecuada depende del problema a abordar. De forma general, ninguna técnica se revela como claramente superior a las demás en el estudio de los sistemas ecológicos y la mejor aproximación suele depender de los objetivos (Brososfske et al. 2014). En todo caso, hay que considerar que pequeñas diferencias en los ajustes de los modelos pueden ser importantes en su aplicación en grandes áreas y, por tanto, debe haber un compromiso entre permitir errores en el modelo y que el modelo sea generalista.

La disponibilidad de una gran cantidad de datos recogidos por sensores remotos permite comprender procesos que ocurren en grandes áreas. Su uso requiere de un análisis previo de las herramientas a utilizar, considerando la resolución espacial, espectral y temporal, así como el coste económico de la obtención y gestión de dichos datos. Existe información de libre acceso que cubre la superficie global con resoluciones medias y bajas. En cambio, cuando requerimos información a muy alta resolución espacial, la disponibilidad en grandes áreas es menor y normalmente a un coste muy elevado, tanto económico, como computacional. Así mismo, la tendencia parece ir hacia la implementación de modelos cada vez más complejos incluyendo aprendizaje reforzado y actualizaciones periódicas de la información de forma que los ajustes mejoren con el paso del tiempo, aunque esto suponga un aumento de la complejidad de los modelos y del coste computacional. También es importante remarcar que este último aspecto cada vez resulta más fácil de solventar, dado que el almacenamiento y la gestión de los datos es cada vez más eficiente a través de servicios de soportes en la nube para Big Data. La aplicación de estas herramientas en el sector forestal está suponiendo un fenómeno disruptivo, con cambios radicales en prácticas establecidas durante décadas, como son los inventarios forestales, obteniendo modelos con un nivel de detalle muy superior y aportando herramientas clave para la toma de decisiones por el gestor.



## 5.2 Futuras líneas de trabajo

De los trabajos realizados durante esta tesis se pueden derivar diferentes propuestas de futuro:

Aplicar la metodología empleada para discriminar entre las dos especies de pinos, sobre masas mixtas con mayor diversidad de especies.

Realizar la clasificación de especies a nivel de árbol individual utilizando información espectral o SAR de libre acceso, como Sentinel, lo que permitiría valorar si la posible disminución de la capacidad predictiva del modelo puede ser asumida por la disminución del coste económico.

Refinar la metodología de individualización de copas para la monitorización de árboles de forma automática, utilizando para ello la segunda cobertura de vuelos LiDAR-PNOA, la cual está siendo llevada a cabo y puesta a disposición de los usuarios actualmente.

Desagregar con mayor detalle la discriminación de las tipologías de masa presentes en los bosques de cobertura completa, incluso llegando al nivel de especie.

Evaluar el compromiso entre la recurrencia de los vuelos LiDAR-PNOA y la superficie volada mediante vuelos de RPAS con cámaras multiespectrales y LiDAR de cara a realizar monitorización de sistemas forestales.





### 5.3 Conclusiones

A continuación, se presentan una serie de conclusiones específicas que se derivan de los objetivos que se han desarrollado en cada uno de los trabajos realizados.

1.- La combinación de datos LiDAR e imágenes de satélite Plèiades para clasificación de especies es más eficiente que el uso de cada una de las fuentes de datos individualmente. Esta combinación ha permitido discriminar a nivel de árbol individual entre dos especies similares en bosques mixtos de coníferas en la cuenca Mediterránea a partir de datos de entrenamiento en masas puras.

2.- La clasificación de las dos especies en masas mixtas puede considerarse satisfactoria pese a la semejanza de ambas especies, especialmente en cuanto a su respuesta espectral. La compensación de los errores a nivel de rodal muestra que la fortaleza del modelo recae en que ha sido entrenado sobre masas puras, donde es inequívoca la caracterización de la especie, y aplicado posteriormente sobre una masa mixta.

3.- La combinación de datos LiDAR e imágenes Sentinel multitemporales ha permitido clasificar las diferentes tipologías de masa en bosques de cobertura completa con alta densidad y continuidad vertical como son los bosques de monteverde.

4.- Las metodologías empleadas son válidas en su aplicación a grandes áreas, tanto a nivel regional como de paisaje, para la evaluación de los recursos. Los trabajos desarrollados aportan herramientas automáticas que mejoran la estimación del valor económico de la producción de los recursos forestales y dan apoyo al gestor en la toma de decisiones.

5.- Las herramientas utilizadas de *machine learning*: *Random Forest* (RF); *Linear Support Vector Machine* (SVML); *Radial Support Vector Machine* (SVMR) y *Artificial Neural Networks* (ANN), han sido capaces de clasificar con elevada precisión las diferentes tipologías de monteverde en Tenerife.

6.- Pequeñas diferencias entre modelos pueden ser importantes cuando se aplican sobre grandes áreas. Por lo tanto, las comparaciones entre algoritmos deben considerarse sobre los mismos objetos, características y tamaños de área de estudio.

7.- El índice RdNBR puede ser fácilmente entendible a través de variables de campo sencillas de medir como son la altura del quemado en la corteza y la pérdida de cobertura de matorral y arbolado. Además, se confirma que el lapso de tiempo entre el evento y la evaluación influye, no sólo en la correlación, sino también en la interpretación de la severidad. Esto ha supuesto la generación de una herramienta eficiente para la evaluación del impacto inducido por el fuego en los ecosistemas mediterráneos y, por tanto, sobre la pérdida de la biomasa en regiones grandes y heterogéneas.

8.- El análisis de las correlaciones entre variables hace posible la generación de modelos sencillos, comprensibles y robustos. El análisis estadístico muestra la importancia de la variabilidad de la severidad del fuego entre incendios y su influencia en la respuesta del RdNBR, debido a factores físicos como la topografía, el clima y la composición y estado de desarrollo de la vegetación antes del incendio.

9.- A pesar de que, en el estudio sobre la severidad de los incendios, los modelos mixtos arrojan mejores resultados en sus diagnósticos, los análisis del BRT muestran que las relaciones entre la variable dependiente RdNBR y las variables independientes incluidas en el modelo no son necesariamente lineales.





## Agradecimientos

---

*“No hay barrera, cerradura ni cerrojo  
que puedas imponer a la libertad de mi mente”*

Virginia Woolf

## Agradecimientos

---

Este trabajo ha sido financiado a través de las ayudas para contratos para la formación de investigadores en empresas, Doctorados Industriales del Ministerio de Economía y Competitividad del Gobierno de España (DI-14-06953), en la empresa föra forest technologies.

En primer lugar, muchísimas gracias a mis directores de tesis Dr. José Miguel Olano, Dr. Francisco Rodríguez y Dr. José Ramón González por sus aportaciones y su apoyo, especialmente durante estos últimos meses. Quiero hacer especial agradecimiento a Paco e Iñigo por haber apostado por este proyecto conjunto e integrarlo dentro de su empresa cuando prácticamente acababa de arrancar. Ahora ya, es una gran empresa. También quiero agradecer su apoyo al resto de los compañeros de la empresa Rafa, Bea, Saray, Fernando y Javier, cada uno de vosotros habéis aportado vuestro granito.

Agradecer también a Blas Mola y Rafa Calama por su acogida y apoyo durante mi estancia en sus respectivos grupos. Cortas, pero fructíferas y satisfactorias, tanto para esta tesis como para mí personalmente. He aprendido mucho con vosotros. Quiero agradecer el apoyo de la *COST Action FPI301 EuroCoppice* que me permitió visitar en marzo de 2017 la *University of Eastern Finland – Joensuu*. Del mismo modo, agradecer la acogida de los miembros del Departamento de Selvicultura y Gestión de Sistemas Forestales del INIA-CIFOR en Madrid durante mi visita en septiembre de 2017.

Quiero expresar especial agradecimiento a Borque por su comprensión y apoyo incondicional durante el largo camino que supone hacer una tesis, por soportar mis ausencias, por los buenos y malos momentos que hemos compartido... juntos podemos con todo. Gracias por compartir tu vida conmigo.

También quiero dar las gracias a mi familia, mis padres, Andrés y Toñi, por vuestro cariño, esfuerzo y dedicación durante toda mi vida, siempre estáis ahí. Y a mis hermanos que, aunque ahora estáis un poco más lejos, nunca dejaremos de encender ese horno de leña y compartir largas y divertidas sobremesas.

Gracias a la Isabel y al Patín por vuestro apoyo y al Pepe, sobre todo por el soporte logístico.

Mil gracias a la otra familia, a la que eliges, a tod@s mis amig@s, a los que estáis más lejos y los que estáis más cerca por todas las risas y buenos momentos que hemos vivido y que viviremos. Todos habéis sido conscientes de la montaña rusa en la que he viajado durante los últimos años. Gracias al Medra, a la Jules y el Juan, el Jaime y el Tito, la Ali y el Barrabás y un largo etc... por esos conciertos y risas reconstituyentes. Gracias también a todo el ecosistema BRIF, sois todos unos amores.

Gracias también a toda la gente de Zárágózá, Gabi, Rox, Isamar, Aldo y Jon Qmás te Da, ¡cómo me iba a olvidar de vosotros!

Y por supuesto, gracias a las niñas, Yoli y Araceli, ¡vosotras sí que valéis millón!

Y a la Shira, sobre todo por esos largos paseos bajo cero que despejan cuerpo y mente.





## Referencias

---

*“Uno no es lo que es por lo que escribe  
sino por lo que ha leído”*

Jose Luis Borges

## Referencias

---

- Akbari H, Kalbi S (2017) Determining Pleiades satellite data capability for tree diversity modeling. *IForest* 10:348–352. doi: 10.1016/j.jsv.2018.08.051
- Arozena ME, Panareda JM, Rivero B (2015) La herencia del carboneo en el paisaje de la laurisilva canaria. 1535–1542
- Assman E (1970) *The Principles of Forest Yield Study*. Pergamon Press Oxford, New York, NY 506 pp
- Asta J, Erhardt W, Ferretti M, Fornasier F (2002) Mapping lichen diversity as an indicator of environmental quality. *Environment* 7:273–279. doi: Article
- Badreldin N, Sanchez-Azofeifa A (2015) Estimating forest biomass dynamics by integrating multi-temporal Landsat satellite images with ground and airborne LiDAR data in the Coal Valley Mine, Alberta, Canada. *Remote Sens* 7:2832–2849. doi: 10.3390/rs70302832
- Behncke B, Fornaciai A, Neri M, et al (2016) Lidar surveys reveal eruptive volumes and rates at Etna, 2007–2010. *Geophys Res Lett* 43:4270–4278. doi: 10.1002/2016GL068495
- Bergen KM, Dronova I (2007) Observing succession on aspen-dominated landscapes using a remote sensing-ecosystem approach. *Landsc Ecol* 22:1395–1410. doi: 10.1007/s10980-007-9119-1
- Blázquez-Casado Á (2017) Modelling methodologies focussed on different machine learning using Rstudio for stand classification and productivity
- Blázquez-Casado Á, González-Olabarria JR, Martín-Alcón S, et al (2015) Assessing Post-Storm Forest Dynamics in the Pyrenees Using High-Resolution LIDAR Data and Aerial Photographs. 12:841–853
- Broszofske KD, Froese RE, Falkowski MJ, Banskota A (2014) A review of methods for mapping and prediction of inventory attributes for operational forest management. *For Sci* 60:733–756. doi: 10.5849/forsci.12-134
- Calama R, Tome M, Sanchez-Gonzalez M, et al (2010) Modelling non-wood forest products in Europe: a review. *For Syst* 19:69–85. doi: 10.5424/fs/201019S-9324
- Campbell JB, Wynne RH (2011) *Introduction to remote sensing*. Guilford Press
- Carus M (2017) Biobased Economy and Climate Change—Important Links, Pitfalls, and Opportunities. *Ind Biotechnol* 13:41–51. doi: 10.1089/ind.2017.29073.mca
- Castillo JAA, Apan AA, Maraseni TN, Salmo SG (2017) Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery. *ISPRS J Photogramm Remote Sens* 134:70–85. doi: 10.1016/j.isprsjprs.2017.10.016

- Chrysafis I, Mallinis G, Siachalou S, Patias P (2017) Assessing the relationships between growing stock volume and Sentinel-2 imagery in a Mediterranean forest ecosystem. *Remote Sens Lett* 8:508–517. doi:10.1080/2150704X.2017.1295479
- Chuvieco, E., (2012). *Remote sensing of large wildfires: in the European Mediterranean Basin*. Springer Science & Business Media.
- Clark ML, Roberts DA, Clark DB (2005) Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens Environ* 96:375–398. doi:10.1016/j.rse.2005.03.009
- Daehne A, Corsini A (2013) Kinematics of active earthflows revealed by digital image correlation and DEM subtraction techniques applied to multi-temporal LiDAR data. *Earth Surf Process Landforms* 38:640–654. doi:10.1002/esp.3351
- Dalponte M, Ørka HO, Ene LT, et al (2014) Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sens Environ* 140:306–317. doi:10.1016/j.rse.2013.09.006
- De'Ath, G., 2007. Boosted trees for ecological modeling and prediction. *Ecology* 88, 243–251
- Diéguez-Aranda U, Rojo A. A, Castedo-Dorado F, et al (2009) Herramientas selvícolas para la gestión forestal sostenible en Galicia
- Domingo D, Lamelas-García M. T, Montealegre-García A., L, De la Riva-Fernández J (2017) Comparison of regression models to estimate biomass losses and CO<sub>2</sub> emissions using low-density airborne laser scanning data in a burnt Aleppo pine forest. *Eur J of Remote Sens*, 50:1, 384–396.
- Domingo D, Lamelas-García M. T, Montealegre-García A., L, García-Martín A, De la Riva-Fernández J (2018) Estimation fo total biomass in Aleppo pine forest stands applying parametric and nonparametric methods to low-density airborne laser scanning data. *Forests*, 9, 158; doi:10.3390/f9030158.
- Dupuya S, Lainé G, Tassin J, Sarrailh JM (2013) Characterization of the horizontal structure of the tropical forest canopy using object-based LiDAR and multispectral image analysis. *Int J Appl Earth Obs Geoinf* 25:76–86. doi:10.1016/j.jag.2013.04.001
- Ediriweera S, Pathirana S, Danaher T, Nichols D (2014) Estimating above-ground biomass by fusion of LiDAR and multispectral data in subtropical woody plant communities in topographically complex terrain in North-eastern Australia. *J For Res* 25:761–771. doi:10.1007/s11676-014-0485-7
- FAO (2018) *the World ' S Forests*. Rome
- FAO (2015) *FOREST EUROPE 2015: State of Europe's Forest*.
- FAO (2011) *FOREST EUROPE, UNECE and FAO 2011: State of Europe's Forests 2011. Status and Trends in Sustainable Forest Management in Europe*

- Fisher CT, Cohen AS, Fernández-Díaz JC, Leisz SJ (2017) The application of airborne mapping LiDAR for the documentation of ancient cities and regions in tropical regions. *Quat Int* 448:129–138. doi: 10.1016/j.quaint.2016.08.050
- Garber S (2007) NASA-history. <https://history.nasa.gov/sputnik/index.html>. Accessed 8 Nov 2018
- García M, Riaño D, Chuvieco E, Danson FM (2010) Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sens Environ* 114:816–830. doi: 10.1016/j.rse.2009.11.021
- García M, Saatchi S, Ustin S, Balzter H (2018) Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery. *Int J Appl Earth Obs Geoinf* 66:159–173. doi: 10.1016/j.jag.2017.11.017
- Genuer R, Poggi JM, Tuleau-Malot C, (2015). VSURF: An R Package for Variable Selection Using Random Forests. <https://journal.r-project.org/archive/2015/RJ-2015-018/RJ-2015-018.pdf> [8 June 2018].
- Ghosh A, Fassnacht FE, Joshi PK, Kochb B (2014) A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int J Appl Earth Obs Geoinf* 26:49–63. doi: 10.1016/j.jag.2013.05.017
- Glennie CL, Carter WE, Shrestha RL, Dietrich WE (2013) Geodetic imaging with airborne LiDAR: The Earth's surface revealed. *Reports Prog Phys* 76:. doi: 10.1088/0034-4885/76/8/086801
- Guerra-Hernández J, Tomé M, González-Ferreiro E (2016) Using low density LiDAR data to map Mediterranean forest characteristics by means of an area-based approach and height threshold analysis. *Rev Teledetección* 0:103–117. doi: 10.4995/raet.2016.3980
- Harris, S., Veraverbeke, S., Hook, S., (2011). Evaluating spectral indices for assessing fire severity in chaparral ecosystems (Southern California) using modis/aster (MASTER) airborne simulator data. *Remote Sens.* 3, 2403–2419. doi:10.3390/rs3112403
- Hassan R, Scholes R, Ash N (2005) ECOSYSTEMS AND HUMAN WELL-BEING: Current State and Trends, Volume 1
- Hatten JR (2014) Mapping and monitoring Mount Graham red squirrel habitat with Lidar and Landsat imagery. *Ecol Modell* 289:106–123. doi: 10.1016/j.ecolmodel.2014.07.004
- Heinzel J, Koch B (2012) Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *Int J Appl Earth Obs Geoinf* 18:101–110. doi: 10.1016/j.jag.2012.01.025
- Hirschmugl M, Gallaun H, Dees M, et al (2017) Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review. *Curr For Reports* 3:32–45. doi: 10.1007/s40725-017-0047-2
- Hudak AT, Crookston NL, Evans JS, et al (2006) Regression modeling and mapping of coniferous forest basal area and tree density from discrete- return LiDAR and multispectral satellite data. *Can J Remote Sens* 32:126–138. doi: 10.5589/m06-007

- Immitzer M, Vuolo F, Atzberger C (2016) First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens* 8. doi: 10.3390/rs8030166
- JRC-EFFIS (2017) *Forest Fires in Europe, Middle East and North Africa 2017*
- Karna YK, Hussin YA, Gilani H, et al (2015) Integration of WorldView-2 and airborne LiDAR data for tree species level carbon stock mapping in Kayar Khola watershed, Nepal. *Int J Appl Earth Obs Geoinf* 38:280–291. doi: 10.1016/j.jag.2015.01.011
- Kimberley MO, Watt MS, Harrison D (2017) Characterising prediction error as a function of scale in spatial surfaces of tree productivity. *New Zeal J For Sci* 47. doi: 10.1186/s40490-017-0100-8
- Keeley, J.E., (2009). Fire intensity, fire severity and burn severity: A brief review and suggested usage. *Int. J. Wildl. Fire* 18, 116–126. doi:10.1071/WF07049
- Knapp N, Fischer R, Huth A (2018) Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. *Remote Sens Environ* 205:199–209. doi: 10.1016/j.rse.2017.11.018
- Koch B (2011) *La teledetección como apoyo a los inventarios forestales nacionales EFN*
- Latifi H, Nothdurft A, Koch B (2010) Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: Application of multiple optical/LiDAR-derived predictors. *Forestry* 83:395–407. doi: 10.1093/forestry/cpq022
- Lehtile, L.B., Holden, Z.A., Smith, A.M.S., Falkowski, M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E., Benson, N.C., (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *Int. J. Wildl. Fire* 15, 319–345. doi:10.1071/WF05097
- Li C, Zhu X, Wei Y, et al (2018) Estimating apple tree canopy chlorophyll content based on Sentinel-2A remote sensing imaging. *Sci Rep* 8:1–10. doi: 10.1038/s41598-018-21963-0
- Lim K, Treitz P, Wulder M, et al (2003) LiDAR remote sensing of forest structure. *Prog Phys Geogr* 27:88–106. doi: 10.1191/0309133303pp360ra
- Lohani B, Ghosh S (2017) Airborne LiDAR Technology: A Review of Data Collection and Processing Systems. *Proc Natl Acad Sci India Sect A - Phys Sci* 87:567–579. doi: 10.1007/s40010-017-0435-9
- Magnussen S, Nord-Larsen T, Riis-Nielsen T (2018) Lidar supported estimators of wood volume and aboveground biomass from the Danish national forest inventory (2012–2016). *Remote Sens Environ* 211:146–153. doi: 10.1016/j.rse.2018.04.015
- Maltamo M, Næsset E, Vauhkonen J (2014) *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*
- Marinelli D, Paris C, Bruzzone L (2018) A Novel Approach to 3-D Change Detection in Multitemporal LiDAR Data Acquired in Forest Areas. *IEEE Trans Geosci Remote Sens* 56:3030–3046. doi: 10.1109/TGRS.2018.2789660

- Matasci G, Hermosilla T, Wulder MA, et al (2018) Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. *Remote Sens Environ* 216:697–714. doi: 10.1016/j.rse.2018.07.024
- Meng J, Li S, Wang W, et al (2016) Estimation of forest structural diversity using the spectral and textural information derived from SPOT-5 satellite images. *Remote Sens* 8: doi: 10.3390/rs8020125
- Muhati GL, Olago D, Olaka L (2018) Quantification of carbon stocks in Mount Marsabit Forest Reserve, a sub-humid montane forest in northern Kenya under anthropogenic disturbance. *Glob Ecol Conserv* 14:e00383. doi: 10.1016/j.gecco.2018.e00383
- Mura M, Bottalico F, Giannetti F, et al (2018) Exploiting the capabilities of the Sentinel-2 multi spectral instrument for predicting growing stock volume in forest ecosystems. *Int J Appl Earth Obs Geoinf* 66:126–134. doi: 10.1016/j.jag.2017.11.013
- Næsset E, Gobakken T, Holmgren J, et al (2004) Laser scanning of forest resources: The nordic experience. *Scand J For Res* 19:482–499. doi: 10.1080/02827580410019553
- Naidoo L, Cho MA, Mathieu R, Asner G (2012) Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J Photogramm Remote Sens* 69:167–179. doi: 10.1016/j.isprsjprs.2012.03.005
- Nitze I, Schulthess U, Asche H (2012) Comparison of machine learning algorithms random forest, artificial neuronal network and support vector machine to maximum likelihood for supervised crop type classification. *Proc tje 4th GEOBIA, May 7–9 – Rio Janeiro – Brazil* 35–40. doi: 10.2355/isijinternational.49.425
- Nordkvist K, Granholm A-H, Holmgren J, et al (2012) Combining optical satellite data and airborne laser scanner data for vegetation classification. *Remote Sens Lett* 3:393–401. doi: 10.1080/01431161.2011.606240
- Onojeghuo AO, Onojeghuo AR (2017) Object-based habitat mapping using very high spatial resolution multispectral and hyperspectral imagery with LiDAR data. *Int J Appl Earth Obs Geoinf* 59:79–91. doi: 10.1016/j.jag.2017.03.007
- Palahí M, Pukkala T, Bonet A, et al (2009) Management of Even-Aged Pine Stands of Catalonia
- Par H, Jeong S J, Ho C H, Kim J, Brown M E, Schaepman M E (2015). Nonlinear response of vegetation green-up to local temperature variations in temperate and boreal forests in the Northern Hemisphere. *Remote Sensing of Environment*, 165, 100–108.
- Pearse GD, Dash JP, Persson HJ, Watt MS (2018) Comparison of high-density LiDAR and satellite photogrammetry for forest inventory. *ISPRS J Photogramm Remote Sens* 142:257–267. doi: 10.1016/j.isprsjprs.2018.06.006
- Peura M, Silveyra Gonzalez R, Müller J, et al (2016) Mapping a 'cryptic kingdom': Performance of lidar derived environmental variables in modelling the occurrence of forest fungi. *Remote Sens Environ* 186:428–438. doi: 10.1016/j.rse.2016.09.003

- Popescu SC (2007) Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy* 31:646–655. doi: 10.1016/j.biombioe.2007.06.022
- Popescu SC, Wynne RH, Nelson RF (2003) Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Can J Remote Sens* 29:564–577. doi: 10.5589/m03-027
- Potapov P V., Turubanova SA, Tyukavina A, et al (2015) Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sens Environ* 159:28–43. doi: 10.1016/j.rse.2014.11.027
- Rego C, Boieiro M, Gonçalves Y, et al (2014) The drosophilids (Diptera: Drosophilidae) from a Laurisilva patch in Madeira with two new records for this island. *Bocagiana* 3141:1–8
- Rodriguez-Veiga P, Saatchi S, Wheeler J, et al (2017) Methodology for Regional to Global Mapping of Aboveground Forest Biomass: Integrating Forest Allometry, Ground Plots, and Satellite Observations. *Earth Obs L Emerg Monit* 7
- Rodríguez F, Lizarralde I (2015) Inventario forestal con tecnología LiDAR, ¡Más barato y más preciso! *Forestalis* 24:33–34
- Sarrazin MJD, Aardt JANV, Asner GP, et al (2012) Fusing small-footprint waveform LiDAR and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. *Can J Remote Sens* 37:653–665. doi: 10.5589/m12-007
- Sasaki T, Imanishi J, Ioki K, et al (2012) Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. *Landsc Ecol Eng* 8:157–171. doi: 10.1007/s11355-011-0158-z
- Seidl R, Schelhaas M, Rammer W, Verkerk PJ (2015) Increasing forest disturbances in Europe and their impact carbon storage. 4:806–810. doi: 10.1038/nclimate2318.Increasing
- Shi Y, Skidmore AK, Wang T, et al (2018) Int J Appl Earth Obs Geoinformation Tree species classification using plant functional traits from LiDAR and hyperspectral data. *Int J Appl Earth Obs Geoinf* 73:207–219. doi: 10.1016/j.jag.2018.06.018
- Stereńczak K, Lisańczuk M, Erfanifard Y (2018) Delineation of homogeneous forest patches using combination of field measurements and LiDAR point clouds as a reliable reference for evaluation of low resolution global satellite data. *For. Ecosyst.* 5:1
- Sullivan FB, Ducey MJ, Orwig DA, et al (2017) Comparison of lidar- and allometry-derived canopy height models in an eastern deciduous forest. *For Ecol Manage* 406:83–94. doi: 10.1016/j.foreco.2017.10.005
- Tesch SD (1980) The evolution of forest yield determination and site classification. *For Ecol Manage* 3:169–182. doi: 10.1016/0378-1127(80)90014-6
- Thers H, Brunbjerg AK, Læssøe T, et al (2017) Lidar-derived variables as a proxy for fungal species richness and composition in temperate Northern Europe. *Remote Sens Environ* 200:102–113. doi: 10.1016/j.rse.2017.08.011

- Thom D, Seidl R (2016) Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. *Biol Rev Camb Philos Soc* 91:760–781. doi: 10.1111/brv.12193
- Tielidze LG (2015) Glaciers change over the last century, Caucasus Mountains, Georgia, observed by the old topographical maps, Landsat and ASTER satellite imagery. *Cryosph Discuss* 9:3777–3806. doi: 10.5194/tcd-9-3777-2015
- Tonolli S, Dalponte M, Neteler M, et al (2011) Fusion of airborne LiDAR and satellite multispectral data for the estimation of timber volume in the Southern Alps. *Remote Sens Environ* 115:2486–2498. doi: 10.1016/j.rse.2011.05.009
- Urban M, Berger C, Mudau TE, et al (2018) Surface moisture and vegetation cover analysis for drought monitoring in the southern Kruger National Park using Sentinel-1, Sentinel-2, and Landsat-8. *Remote Sens* 10. doi: 10.3390/rs10091482
- Vanclay J.K. (1994) Modelling forest growth and yield: applications to mixed tropical forest. *School of Environmental Science and Management Papers*, 537.
- Vauhkonen J (2018) Predicting the provisioning potential of forest ecosystem services using airborne laser scanning data and forest resource maps. *For Ecosyst* 5:1–19. doi: 10.1186/s40663-018-0143-1
- Vázquez-Piqué J, Pereira H (2008) What to take into account to develop cork weight models?: Review and statistical considerations. *Inst Nac Investig y Tecnol Agrar y Aliment* 17:199–215
- Vega Isuhuaylas LA, Hirata Y, Ventura Santos LC, Serrudo Torobeo N, (2018). Natural forest mapping in the Andes (Peru): a comparison of the performance of machine-learning algorithms. *Remote Sens* 10: 782.
- Vepakomma U, St-Onge B, Kneeshaw D (2008) Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sens Environ* 112:2326–2340. doi: 10.1016/j.rse.2007.10.001
- Verhegghen A, Eva H, Ceccherini G, et al (2016) The potential of sentinel satellites for burnt area mapping and monitoring in the Congo Basin forests. *Remote Sens* 8:1–22. doi: 10.3390/rs8120986
- Wang H, Glennie C (2015) Fusion of waveform LiDAR data and hyperspectral imagery for land cover classification. *ISPRS J Photogramm Remote Sens* 108:1–11. doi: 10.1016/j.isprsjprs.2015.05.012
- Warner, T.A., Skowronski, N.S., Gallagher, M.R., (2017). High spatial resolution burn severity mapping of the New Jersey Pine Barrens with WorldView-3 near-infrared and shortwave infrared imagery. *Int. J. Remote Sens.* 38, 598–616. doi:10.1080/01431161.2016.1268739
- Wong J, Prokofieva I (2014) Report presenting synthesis of regional sectorial reviews to describe the “State of the European NWF” StarTree deliverable D1.3. 96 pp



- Wu B, Yu B, Wu Q, et al (2016) Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests. *Int J Appl Earth Obs Geoinf* 52:82–94. doi: 10.1016/j.jag.2016.06.003
- Wu C, Tao H, Zhai M, Lin Y, Wang K, Deng J, Yang H (2018). Using nonparametric modeling approaches and remote sensing imagery to estimate ecological welfare forest biomass. *Journal of Forestry Research*, 29(1), 151-161.
- Wulder MA, White JC, Alvarez F, et al (2009) Characterizing boreal forest wildfire with multi-temporal Landsat and LIDAR data. *Remote Sens Environ* 113:1540–1555. doi: 10.1016/j.rse.2009.03.004
- Wulder MA, White JC, Cranny M, et al (2008) Monitoring Canada's forests. Part I: Completion of the EOSD land cover project. *Can J Remote Sens* 34:549–562. doi: 10.5589/m08-066
- Xu C, Manley B, Morgenroth J (2018) Evaluation of modelling approaches in predicting forest volume and stand age for small-scale plantation forests in New Zealand with RapidEye and LiDAR. *Int J Appl Earth Obs Geoinf* 73:386–396. doi: 10.1016/j.jag.2018.06.021
- Xu Q, Man A, Frederickson MM, et al (2017) Quantification of the uncertainty in the aboveground biomass estimates derived from small-footprint airborne LiDAR. *Prep* 216:514–528. doi: 10.1016/j.rse.2018.07.022
- Youssef AM, Pourghasemi HR, El-Haddad BA, Dhahry BK (2016) Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir Region, Saudi Arabia. *Bull Eng Geol Environ* 75:63–87. doi: 10.1007/s10064-015-0734-9
- Zhao K, Suarez JC, Garcia M, et al (2018) Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. *Remote Sens Environ* 204:883–897. doi: 10.1016/j.rse.2017.09.007
- Zhu XX, Tuia D, Mou L, Xia GS, Zhang L, Xu F, Fraundorfer F, 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci Remote Sens Mag* 5: 8-36

