



**2025** | **16-20**  
**GIJÓN** | **JUNIO**

**9º CONGRESO FORESTAL ESPAÑOL**

**9CFE-1615**

Actas del Noveno Congreso Forestal Español  
Edita: **Sociedad Española de Ciencias Forestales. 2025.**  
ISBN: **978-84-941695-7-1**

Organiza





## Integración de Teledetección y Machine Learning para proporcionar Listas de Árboles Espacialmente Explícitas para *Pinus pinaster*

GONZÁLEZ MESQUIDA, J.B. (1), MAURO-GUTIÉRREZ, F. (1)

(1) iuFOR, EiFAB, Universidad de Valladolid, 42004 Soria, España

Keywords: Forest Management, Forest Structure Descriptors, Growth Models, Spain

### Resumen

La estimación precisa de los descriptores de la estructura forestal es crucial para comprender la dinámica de los ecosistemas, gestionar los recursos forestales y predecir las respuestas al cambio climático. En las últimas décadas, se han desarrollado múltiples modelos de crecimiento y simuladores de incendios forestales que permiten proyectar el comportamiento de los bosques y evaluar cómo las alternativas de gestión afectan a su resiliencia. Estos modelos requieren principalmente dos entradas: i) una descripción de las condiciones climáticas para el crecimiento y ii) listas de árboles espacialmente explícitas, que detallan la composición y estructura de una determinada área boscosa, permitiendo proyectar su evolución bajo distintos escenarios de gestión. Cada registro en estas listas representa un árbol prototipo, que incluye descriptores básicos y un factor de expansión, es decir, la frecuencia esperada de aparición de árboles similares al prototipo por unidad de superficie. Sin embargo, a pesar de su gran potencial, la aplicación de estos modelos se ve limitada por la falta de datos espaciales explícitos, como los mapas de listas de árboles. Este estudio aborda esta limitación proporcionando modelos para obtener descriptores de la estructura de *Pinus pinaster* Aiton que pueden integrarse en simuladores forestales. Para ello, se combinan datos de teledetección y del Inventario Forestal Nacional bajo un esquema de aprendizaje automático, con el objetivo de mejorar las estrategias de gestión forestal en el norte de España bajo diferentes escenarios climáticos. Se consideraron cuatro variables dasométricas (volumen con corteza, área basimétrica, número de individuos por hectárea y biomasa aérea total), y se ajustaron modelos de imputación tanto univariantes como multivariantes, con entre 1 y 3 vecinos, y hasta 5 variables por modelo. Se probaron dos métodos por cada número de vecinos: MSN y Random Forests. La métrica de performance utilizada fue la distancia de Mahalanobis ajustada, penalizando los modelos más complejos. La investigación empleó técnicas avanzadas, como el algoritmo genético y métodos de imputación (k-NN y Random Forest), para generar mapas espacialmente explícitos de estos descriptores a una resolución de 30 metros. Los resultados obtenidos mostraron una alta precisión en la predicción de descriptores forestales, lo que permitirá generar mapas detallados y espacialmente explícitos que mejoren las simulaciones de resiliencia forestal y las estrategias de gestión bajo diferentes escenarios climáticos.

### Palabras clave

LiDAR, imputación, kNN, modelos de crecimiento de árbol individual, lista de árboles.

### 1. Introducción



Los descriptores de la estructura forestal, como el volumen de madera, la densidad de rodal y las dimensiones de los árboles, son fundamentales para comprender la dinámica de los ecosistemas forestales, optimizar la gestión de los recursos y prever los impactos del cambio climático (PEARSE et al., 2025). En las últimas décadas, se han desarrollado múltiples modelos de crecimiento y simuladores de incendios forestales que permiten realizar proyecciones y evaluar cómo las distintas alternativas de gestión contribuyen a lograr estructuras forestales más resilientes (BRAVO et al., 2012; DE CÁCERES et al., 2022). Muchos de estos modelos de crecimiento están disponibles para los principales tipos de bosques de la Península Ibérica, y generalmente requieren dos entradas: i) una descripción de las condiciones climáticas en las que se producirá el crecimiento; y ii) listas de árboles espacialmente explícitas, que son tablas que detallan la composición y estructura forestal de una determinada área boscosa, permitiendo su proyección bajo diferentes escenarios de gestión. Una lista de árboles, o *Tree List* (TL), es una tabla dinámica que detalla la abundancia, tamaño, especies y estado de salud de los árboles dentro de un área forestal dada (TEMESGEN et al., 2003). Cada registro en una TL representa un árbol prototipo, y además de los diferentes descriptores forestales, debe incluir un factor de expansión que expresa la frecuencia esperada de árboles similares por unidad de área.

El advenimiento de diferentes tecnologías de teledetección ha revolucionado la ciencia forestal al proporcionar una visión sinóptica de los paisajes forestales, y se ha posicionado como una herramienta prácticamente omnipresente en el ámbito forestal (FASSNACHT et al., 2024). Especialmente relevante es el caso del LiDAR, cuya elevada correlación con los descriptores estructurales lo ha convertido en una herramienta fundamental a nivel operativo desde principios de los años 2000 (NÆSSET, 2002). En España, la abundancia de datos de teledetección, combinada con extensos datos de campo del Inventario Forestal Nacional (IFN) y los inventarios regionales gestionados por cada Comunidad Autónoma, ofrece la oportunidad de generar descriptores forestales espacialmente explícitos virtualmente para toda la península.

El elevado número de predictores provenientes de datos de teledetección, sin embargo, requiere complejos métodos de selección de las variables. Para encontrar soluciones óptimas entre un amplio conjunto de predictores, es necesario entonces recurrir a métodos metaheurísticos avanzados como *simulated annealing* (PUKKALA, 2024), o los algoritmos genéticos, que simulan procesos de selección natural para lograr soluciones casi óptimas mediante una optimización iterativa.

La integración de estas técnicas avanzadas de selección de variables con nuevos métodos de aprendizaje automático, permiten la creación de relaciones robustas entre las variables de teledetección y las mediciones de campo, y en última instancia, una estimación precisa de los descriptores de la estructura forestal (HUDAK et al., 2020). No obstante, es necesario seleccionar una metodología apropiada de modelado. La imputación aparece aquí como una técnica fundamental, ya que entre sus principales virtudes permite predecir eficazmente múltiples variables de respuesta simultáneamente (HUDAK et al., 2008). La imputación se define como "el reemplazo de mediciones faltantes o no muestreadas para cualquier unidad en la población con mediciones de otra unidad con características similares" (VAN DEUSEN, 1997). Además, existen diferentes métodos de imputación, entre los que se incluyen el vecino más similar, el k-vecino más cercano, la estimación geoestadística y los métodos de imputación tabular (TEMESGEN et al., 2003).



Las técnicas de imputación para generar datos de descriptores forestales en diversos niveles de agregación están bien documentadas (TEMESGEN et al., 2003; HUDAK et al., 2008; ESKELSON et al., 2019; MAURO et al., 2019). Sin embargo, su aplicación en España ha sido limitada, con desarrollos notables principalmente en Cataluña (PUKKALA et al., 2024). Debido a esta escasez de descriptores forestales espacialmente explícitos, lo que obstaculiza su uso en simulaciones de gestión forestal bajo escenarios de cambio global, es esencial expandir estos mapas de descriptores espacialmente explícitos a otras regiones de la Península Ibérica. Como primer paso fundamental entonces, se vuelve necesario desarrollar modelos para diversas especies forestales en toda la península.

## 2. Objetivos

Este estudio busca obtener modelos de predicción para descriptores de estructura forestal de *Pinus pinaster* Aiton combinando datos de teledetección y del Inventario Forestal Nacional bajo un esquema de aprendizaje automático. Estos modelos proporcionarán información que permitirán mejorar las estrategias de gestión forestal bajo diferentes escenarios de cambio climático.

## 3. Metodología

### Área de estudio

La investigación se centró en la región oriental de Castilla y León, abarcando las provincias de Soria, Burgos y Segovia, en donde el pino resinero o negral, *Pinus pinaster* Aiton, tiene una amplia representación. Esta especie fue elegida tanto por su importancia económica como por su distribución regional, cubriendo un amplio abanico de condiciones edafoclimáticas en el área de estudio. El área de estudio, que abarca unas 177.349 hectáreas, fue obtenida seleccionando todos los polígonos del Mapa Forestal de España (MFE) en los que el pino resinero tiene una presencia de más del 50% del área de cada polígono como especie primaria. El conjunto resultante está compuesto por 6292 polígonos que abarcan una superficie media de 28.1 hectáreas, y evidenciando una enorme variabilidad en su extensión, con unos valores mínimo y máximo de 0.2 y 572 hectáreas, respectivamente. Las elevaciones dentro de las teselas oscilan entre los 522 y 1.879 metros sobre el nivel del mar, y con un clima mediterráneo continentalizado, la precipitación anual media varía entre 366 y 941 mm.

Los datos terrestres correspondieron a aquellas parcelas del Inventario Forestal Nacional en su cuarto ciclo (IFN4) que fueron incluidas dentro del subconjunto de teselas del MFE. En total, se consideraron un total de 787 parcelas con coordenadas submétricas. La Tabla 1 evidencia los valores calculados por parcela del volumen total con corteza (VCC), el área basimétrica (G), la densidad de árboles (N) y la biomasa aérea (B), pudiéndose apreciar como las parcelas cubren un amplio espectro de estructuras dasométricas.

**Tabla 1.** Resumen de las variables dasométricas de las 747 parcelas incluidas en el estudio de *Pinus pinaster*. VCC: Volumen con corteza por hectárea, G: área basimétrica. N: número de pies por hectárea. B biomasa aérea total.

Variable	Mínimo	Media	Máximo
VCC (m <sup>3</sup> /ha)	1.07	140.34	589.33
G (m <sup>2</sup> /ha)	0.42	22.51	75.47
N(pies/ha)	5.09	409.45	2656.12

B (Mg/ha)	0.77	79.82	310.26
-----------	------	-------	--------

Los datos auxiliares de teledetección fueron agrupados en cuatro conjuntos de predictores: LiDAR, ópticos, climáticos y topográficos. Los flujos de trabajo de preprocesamiento se basaron principalmente en los métodos implementados en el sistema de monitoreo de carbono de la NASA (CMS), según lo descrito por HUDAK et al. (2020).

Los datos espectrales se obtuvieron mediante la creación de series temporales libres de nubes a partir de imágenes de Landsat 5, 7 y 8, correspondientes a los meses de verano en el hemisferio norte para el período 1984-2023. Se utilizó la colección de datos de reflectancia de superficie de Collection 2 Level 2 Tier 1 en Google Earth Engine (GORELICK et al., 2017), y posteriormente, los valores del Índice de Vegetación de Diferencia Normalizada (NDVI) de los compuestos anuales fueron segmentados temporalmente utilizando el algoritmo LandTrendr (KENNEDY et al., 2018, 2010), con el fin de suavizar las tendencias y resaltar los cambios más abruptos.

Las métricas topográficas fueron generadas con una resolución de 30 metros, e incluyeron elevación, pendiente, exposición, latitud, longitud y métricas de rugosidad del terreno. Esta batería de predictores fue derivada a partir de los Modelos Digitales del Terreno (MDT) generados a partir de la última adquisición ALS del segundo ciclo del Plan Nacional de Ortofotografía Aérea (PNOA-LiDAR2).

Los 19 predictores bioclimáticos de WorldClim2 (FICK Y HIJMANS, 2017) fueron remuestreados a una resolución de 30 metros utilizando interpolación bilineal y representados como normales climáticas para el periodo 1970-2000.

Los datos LiDAR fueron obtenidos a partir del PNOA-LiDAR2, y fueron normalizados utilizando la clasificación del suelo proporcionada por la misma base de datos, mediante el paquete lidR en R (ROUSSEL et al., 2020). Una vez normalizados, las métricas provenientes del escaneo láser aerotransportado para cada parcela se derivaron recortando las nubes de puntos normalizadas con un círculo de 16.93 m de radio centrado en cada parcela, asegurando consistencia con el área de 900 m<sup>2</sup> utilizada para las métricas en cuadrícula. Finalmente, las variables espectrales, de cambio espectral, topográficas, geográficas y climáticas se unieron espacialmente a los centros de las parcelas siguiendo los requisitos del método ABA.

### Modelos de imputación y selección de variables

Los modelos de imputación se basaron en el método de k-Vecinos más próximos (k-NN). Se ajustaron una serie de modelos univariantes y multivariantes probando hasta 3 vecinos, limitando el número de predictores a cinco, y utilizando dos métodos de imputación: Análisis de Correlación Canónica (también llamado Vecino Más Similar o MSN, por sus siglas en inglés) y *Random Forests*, implementadas en el paquete *yaImpute* (CROOKSTON Y FINLEY, 2007). Además, se desarrollaron modelos minimizando el RMSE para cada variable de respuesta individualmente.

Los modelos de imputación k-NN luego fueron validados tanto a nivel de variable como global, comparando métricas como RMSE, BIAS y R<sup>2</sup>. La selección de variables predictoras se realizó mediante el método de eliminación sucesiva del paquete *yaImpute*, utilizando la distancia de bosque aleatorio para k-NN. Tanto para los conjuntos ponderados como para los modelos individuales, la métrica de error minimizada fue la distancia media de Mahalanobis, ajustada para penalizar la complejidad del modelo, similar al R<sup>2</sup> ajustado.



La selección de predictores se llevó a cabo mediante un algoritmo genético generado *ad-hoc*, en el que a partir de un *pool* inicial de 400 individuos (i.e., soluciones binarias posibles) generados al azar, y a partir de sucesivos cruces entre individuos, se van seleccionando los mejores. Los modelos se clasificaron en función de una medida de desempeño, por ejemplo el Error Cuadrático Medio (RMSE), seleccionando el modelo con el menor RMSE en cada generación.

El modelo de imputación multivariante incluyó todas las variables dasométricas como variable de respuesta (i.e., VCC, G, N y B en simultáneo), y los modelos univariantes se realizaron por separado, para evaluar el desempeño de cada variable en ambas situaciones.

Tanto para el conjunto de todas las variables ponderadas como para los modelos de kNN para respuestas individuales la métrica de error que se minimizó fue la distancia de Mahalanobis media entre observaciones y predicciones, corregida por un factor de corrección, que penalizó la complejidad del modelo, de un modo similar al R ajustado.

### **Predicción de descriptores espacialmente explícitos**

Una vez generados y clasificados los modelos de imputación univariantes y el mejor modelo multivariante, se realizó la predicción para obtener mapas espacialmente explícitos de descriptores forestales a 30 metros de resolución para posteriormente ser agregados al nivel de la unidad de manejo.

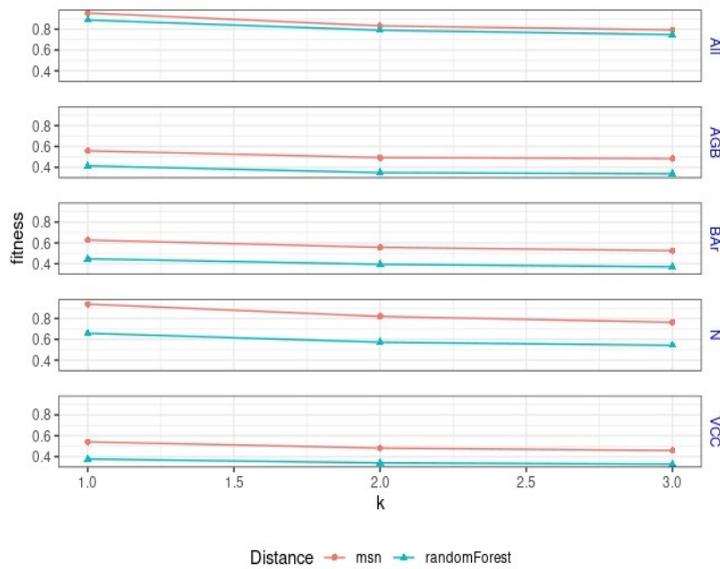
## **5. Resultados**

### **Información auxiliar**

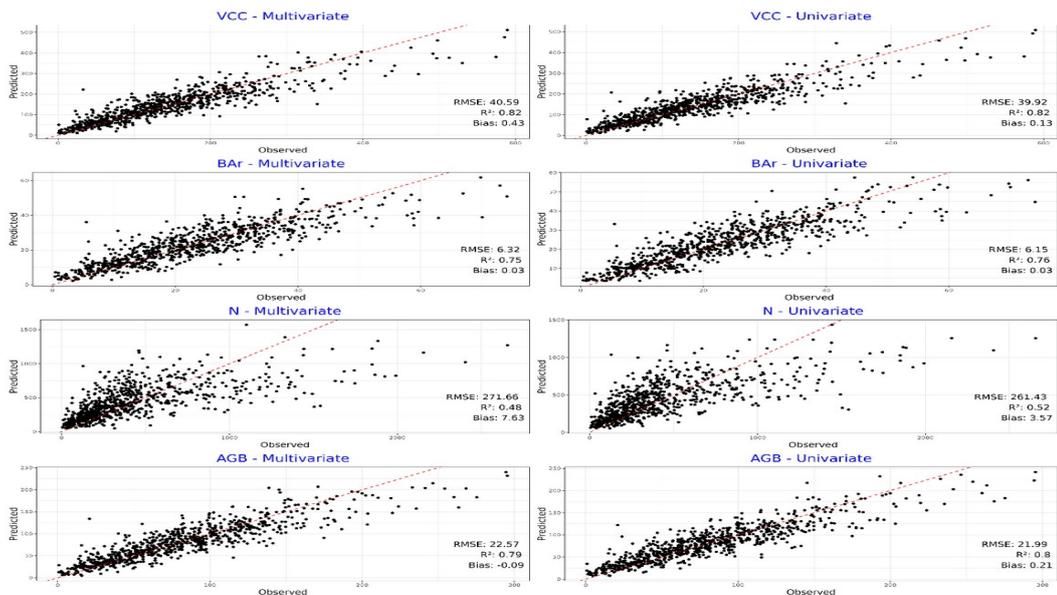
Las diferentes variables auxiliares fueron procesadas obteniendo un total de 206 predictores finales para el área de estudio. Estos predictores fueron agrupados en las siguientes categorías: métricas LiDAR (90), índices topográficos (24), imágenes Landsat (73), y variables climáticas (19).

En la Figura 1 se puede ver que las variables modeladas por separado siempre presentaron una menor distancia de Mahalanobis, lo que implica un mejor ajuste de los modelos en comparación con el modelo multivariante. Adicionalmente, para todos los modelos el método elegido siempre fue el *random forest*, y el número de vecinos que presentó la distancia mínima siempre fue el mayor número de vecinos, es decir  $k = 3$ . Adicionalmente, la complejidad de los modelos siempre se mantuvo en 5 predictores, y en prácticamente todos los casos fueron elegidos 3 predictores lidar, y uno de topografía y uno de clima.

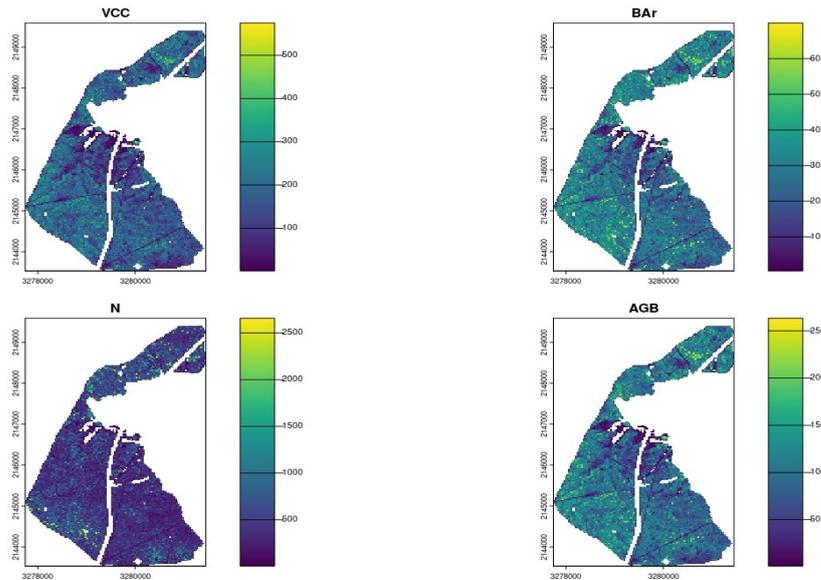
Por otro lado, si bien en términos absolutos el modelo multivariante presentó una distancia de Mahalanobis ligeramente superior a los modelos de variables individuales, cuando analizamos los diagramas de dispersión de los valores observados versus los predichos para las variables modeladas por separado y de forma multivariante (Figura 2), podemos ver que los valores de las métricas de desempeño excelentes para el modelo multivariante. Esto se desprende del hecho de que los estadísticos de bondad de ajuste (RMSE,  $R^2$  y sesgo) fueron, para la misma variable, prácticamente iguales en el modelo multivariante que en el modelo por separado.



**Figura 1.** Desempeño de los modelos para los diferentes métodos: msn y random forest, para el modelo multivariante (ALL) vs los 4 descriptores forestales (VCC, Bar, N y VCC) modeladas por separado.



**Figura 2.** Diagramas de dispersión de los valores observados vs los predichos, para los 4 descriptores forestales en su desempeño en el modelo multivariante y en el modelo individual. Tres indicadores del desempeño fueron considerados: RMSE, R<sup>2</sup> y el bias (sesgo).



**Figura 3.** Mapa de descriptores forestales para un monte de *Pinus pinaster*. Se puede ver los valores predichos a nivel de pixel (30 metros) por el modelo multivariante del volumen con corteza, el Área basimétrica (BA<sub>r</sub>), Número de individuos por hectárea (N) y Biomasa Aérea (AGB).

## 5. Discusión

El presente estudio destaca las ventajas de emplear algoritmos genéticos para la selección de variables en la generación de listas de árboles espacialmente explícitas, una necesidad crítica en la gestión forestal moderna. Los algoritmos genéticos simulan procesos de selección natural, permitiendo una exploración eficiente de grandes espacios de predictores y asegurando la identificación de combinaciones óptimas que mejoran significativamente la precisión de los modelos de imputación. A diferencia de los métodos tradicionales de selección por eliminación sucesiva, que limitan la exploración a un número reducido de combinaciones posibles, los algoritmos genéticos pueden manejar la complejidad y diversidad de los datos de teledetección, particularmente cuando se trabaja con una batería de predictores de teledetección que fácilmente pueden llegar a las centenas de variables. El uso de estos algoritmos facilita la creación de modelos de imputación más robustos, capaces de generar descriptores forestales espacialmente explícitos con alta precisión. Esto es particularmente relevante para regiones como la Península Ibérica, donde la disponibilidad de tales descriptores es relativamente limitada. A pesar del avance significativo en Cataluña (PUKKALA, 2024), la escasez de mapas detallados en otras regiones ha restringido el uso efectivo de simulaciones de gestión forestal, obstaculizando la capacidad de los gestores para proyectar y evaluar la resiliencia de los bosques ante diferentes escenarios de cambio climático.

En este contexto la imputación emerge como una técnica clave, permitiendo la predicción simultánea de múltiples variables de respuesta, lo que es esencial para proporcionar una visión integral de las estructuras forestales. Los métodos como *k*-vecinos más cercanos (*k*-NN) y *Random Forest*, integrados con algoritmos genéticos, ofrecen un enfoque avanzado para el modelado de datos forestales. Estos métodos no solo permiten una mejor representación de las condiciones actuales del bosque, sino que también facilitan la generación de listas de árboles que pueden ser utilizadas en simulaciones dinámicas para evaluar el impacto de



diferentes estrategias de gestión.

La creación de mapas de listas de árboles en el este de Castilla y León representa un avance significativo, proporcionando una base sólida para futuras investigaciones y aplicaciones prácticas en la región. Este estudio subraya la necesidad de seguir desarrollando y refinando estos métodos para garantizar que las herramientas de simulación forestal puedan ser utilizadas de manera más efectiva en toda la Península Ibérica. El desarrollo de descriptores forestales espacialmente explícitos contribuirá no solo a mejorar la gestión local de los recursos forestales, sino también a proporcionar datos esenciales para la planificación y mitigación de los impactos del cambio climático a nivel regional.

## 6. Conclusiones

En resumen, este estudio mejora la gestión forestal en la Península Ibérica mediante el uso de tecnologías avanzadas de teledetección y aprendizaje automático. Se desarrollaron modelos precisos para predecir descriptores de la estructura de *Pinus pinaster*, lo que permitirá generar mapas espaciales explícitos de estos descriptores. Estos avances contribuirán a mejorar las simulaciones forestales y optimizar las estrategias de gestión bajo escenarios climáticos cambiantes.

## 7. Agradecimientos

José Bernardo González Mesquida ha sido financiado dentro de la convocatoria de contratos predoctorales 2022 de la Universidad de Valladolid, cofinanciada por el Banco Santander.

## 8. Bibliografía

ADRAH, E.; MOHD JAAFAR, W. S. W.; BAJAJ, S.; OMAR, H.; LEITE, R. V.; SILVA, C. A.; CARDIL, A.; MOHAN, M.; 2021. Analyzing canopy height variations in secondary tropical forests of Malaysia using NASA GEDI. IOP Conf. Ser.: Earth Environ. Sci. 880 012031.

BRAVO, F.; RODRIGUEZ, F.; ORDÓÑEZ, C.; 2012. A web-based application to simulate alternatives for sustainable forest management: SIMANFOR. Forest Systems, 21(1), 4. <https://doi.org/10.5424/fs/2112211-01953>

CONSEJO ECONÓMICO Y SOCIAL DE CASTILLA Y LEÓN; 2020. El sector forestal en Castilla y León.

CROOKSTON, N. L.; FINLEY, A. O.; 2007. yaImpute: An R Package for kNN Imputation. Journal of Statistical Software, 23(10). <http://www.jstatsoft.org/v23/i10>

DE CÁCERES, M.; MOLOWNY-HORAS, R.; CABON, A.; MARTÍNEZ-VILALTA, J.; MENCUCCINI, M.; GARCÍA-VALDÉS, R.; NADAL-SALA, D.; SABATÉ, S.; MARTIN-STPAUL, N.; MORIN, X.; BATLLORI, E.; AMÉZTEGUI, A.; 2022. MEDFATE 2.8.1: A trait-enabled model to simulate Mediterranean forest function and dynamics at regional scales. Geosci. Model Dev. Discuss. 2022 1–52. <https://doi.org/10.5194/gmd-2022-243>

ESKELSON, B.; TEMESGEN, T.; LEMAY, V., BARRET, M; CROOKSTON, N.; HUDAK, A; 2009. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases, Scandinavian Journal of Forest Research, 24:3, 235-246, DOI: 10.1080/02827580902870490

FASSNACHT, F. E.; WHITE, J. C.; WULDER, M. A.; NÆSSET, E.; 2024. Remote sensing in forestry: current challenges, considerations and directions. For.: Int. J. For. Res. 97 11–37.

FICK, S; Hijmans, R.; 2017. WorldClim 2: new 1-km spatial resolution climate



- surfaces for global land areas. *International journal of climatology*, 37(12), 4302-4315.
- HUDAK, A., CROOKSTON, N., EVANS, J., HALL, D., FALKOWSKI, M.; 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data, *Remote Sensing of Environment*, Volume 112, Issue 5, Pages 2232-2245, ISSN 0034-4257, <https://doi.org/10.1016/j.rse.2007.10.009>.
- HUDAK, A. T.; et al.; 2020. A carbon monitoring system for mapping regional, annual aboveground biomass across the northwestern USA. *Environ. Res. Lett.* 15 095003.
- HUDAK, A. T.; FEKETY, P. A.; KANE, V. R.; KENNEDY, R. E.; FILIPPELLI, S. K.; FALKOWSKI, M. J.; TINKHAM, W. T.; SMITH, A. M. S.; CROOKSTON, N. L.; DOMKE, G. M.; CORRAO, M. V.; BRIGHT, B. C.; CHURCHILL, D. J.; GOULD, P. J.; MCGAUGHEY, R. J.; KANE, J. T.; DONG, J.; 2020. A carbon monitoring system for mapping regional, annual aboveground biomass across the northwestern USA. *Environ. Res. Lett.* 15(9) 095003. <https://doi.org/10.1088/1748-9326/ab93f9>
- MALTAMO, M.; KANGAS, A.; 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. *Can. J. For. Res.* 28(8) 1107–1115.
- MAURO, F.; FRANK, B.; MONLEON, V. J.; TEMESGEN, H.; FORD, K. R.; 2019. Prediction of diameter distributions and tree-lists in southwestern Oregon using LiDAR and stand-level auxiliary information. *Can. J. For. Res.* 49(7) 775–787.
- MEDDENS, A. J.; STEEN-ADAMS, M. M.; HUDAK, A. T.; MAURO, F.; BYASSE, P. M.; STRUNK, J.; 2022. Specifying geospatial data product characteristics for forest and fuel management applications. *Environ. Res. Lett.* 17(4) 045025.
- NÆSSET, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data, *Remote Sensing of Environment*, Volume 80, Issue 1, Pages 88-99, ISSN 0034-4257, [https://doi.org/10.1016/S0034-4257\(01\)00290-5](https://doi.org/10.1016/S0034-4257(01)00290-5).
- PACKALÉN, P.; MALTAMO, M.; 2007. The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sens. Environ.* 109(3) 328–341.
- PEARSE, G. D.; JAYATHUNGA, S.; CAMARRETTA, N.; PALMER, M. E.; STEER, B. S. C.; WATT, M. S.; WATT, P.; HOLDAWAY, A.; 2025. Developing a forest description from remote sensing: Insights from New Zealand. *Sci. Remote Sens.* 11 100183.
- PUKKALA, T.; AQUILUÉ, N.; JUST, A.; CORBERA, J.; TRASOBARES, A.; 2024. Developing kNN forest data imputation for Catalonia. *J. For. Res.* 35(1) 80. <https://doi.org/10.1007/s11676-024-01735-5>
- ROUSSEL, J.-R.; COOPS, N. C.; TOMPALSKI, P.; GOODBODY, T. R. H.; MEADOR, A. S.; BOURDON, J.-F.; BOISSIEU, F. de; ACHIM, A.; 2020. lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sens. Environ.* 251 112061. <https://doi.org/10.1016/j.rse.2020.112061>
- TEMESGEN, B. H.; LEMAY, V. M.; FROESE, K. L.; MARSHALL, P. L.; 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *For. Ecol. Manag.* 177 277–285.
- TEMESGEN, B. H.; LEMAY, V. M.; FROESE, K. L.; MARSHALL, P. L.; 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *For. Ecol. Manag.* 177(1–3) 277–285. [https://doi.org/10.1016/S0378-1127\(02\)00321-3](https://doi.org/10.1016/S0378-1127(02)00321-3)
- TOMPPA, E.; KATILA, M.; 1991. Satellite image-based national forest inventory of Finland. *Int. Arch. Photogramm. Remote Sens.* 28(7–1) 419–424.