




ESTIMACIÓN DE VARIABLES FORESTALES A PARTIR DE SENSORES LIDAR Y ÓPTICOS E INTELIGENCIA ARTIFICIAL

MIHAI TANASE^{1*}, JUAN PABLO MARTINI², PABLO MIRANDA²,
DANIEL GARCIA GARCÍA², VICTORIA WILKE²,
JAIME DIEZ³, SERGIO NATAL³, DANIEL SAN MARTIN³

¹*Instituto de Ciencias Forestales ICIFOR (INIA-CSIC), Madrid, España.*

²*Dpto. Geología, Geografía y Medio Ambiente, Universidad de Alcalá, Alcalá de Henares, España.*

³*Predictia Intelligent Data Solutions S.L., Santander, España.*

RESUMEN. La caracterización precisa y continua de las propiedades de los combustibles forestales es esencial para evaluar el riesgo de incendio y predecir el comportamiento del fuego. Este estudio aborda los desafíos asociados a la escasa disponibilidad temporal de datos lidar, integrando imágenes ópticas (Landsat) con adquisiciones lidar puntuales para estimar variables estructurales del bosque, como la fracción de cubierta (FCC) y la altura del dosel (H), en dos regiones de la España peninsular: Madrid y el País Vasco. Estas variables constituyen indicadores indirectos (proxies) para la caracterización de las propiedades del combustible, ya que reflejan tanto la cantidad como la continuidad espacial de los combustibles. Se compararon modelos de aprendizaje automático (Random Forest y Extreme Gradient Boosting) con arquitecturas de aprendizaje profundo, incluyendo transformers con mecanismos de auto-atención (NeNeT).

Los resultados muestran que los modelos de aprendizaje profundo mejoran significativamente la precisión, con una reducción promedio del error cuadrático medio (RMSE) del 30% respecto a los métodos tradicionales. En particular, NeNeT destacó por su capacidad para capturar relaciones espaciales complejas, mejorando las estimaciones de altura en bosques densos del País Vasco (RMSE reducido de 7,0 a 4,0 m). En contraste, en los bosques mediterráneos más abiertos de Madrid, las diferencias fueron menores, lo que sugiere que métodos menos costosos computacionalmente, como XGB, pueden ser adecuados en ciertos contextos. A pesar de sus ventajas, los modelos de aprendizaje profundo presentan limitaciones operativas, especialmente por sus altas demandas computacionales. Por ejemplo, la generación de mapas históricos en la España peninsular requeriría hasta cuatro años de procesamiento con NeNeT frente a siete meses con XGB, si no se dispone de paralelización. Además, los modelos DL tienden a aprender patrones espurios relacionados con la adquisición de imágenes (por ejemplo, número de observaciones o fechas), lo que puede introducir sesgos si no se controlan adecuadamente.

En conclusión, la combinación de sensores lidar y ópticos, junto con modelos avanzados de inteligencia artificial, permite estimar con alta precisión variables útiles en la gestión de incendios forestales. No obstante, la elección del modelo debe equilibrar la precisión obtenida con los recursos computacionales disponibles, adaptándose al tipo de ecosistema y a las necesidades de la aplicación.

Estimating Forest Variables from Lidar and Optical Sensors Using Artificial Intelligence

ABSTRACT. Accurate and continuous characterization of forest fuel properties is essential for assessing wildfire risk and predicting fire behavior. This study addresses the limitations posed by the scarce temporal availability of lidar data by integrating temporally consistent optical imagery (Landsat) with occasional lidar acquisitions to estimate structural forest variables such as canopy cover fraction (FCC) and canopy height (H) in two regions of peninsular Spain: Madrid and the Basque Country. These variables are widely recognized as proxies for fuel properties characterization, reflecting both the amount and spatial continuity of forest fuels. Machine Learning (ML) models (Random Forest and Extreme Gradient Boosting) were compared with Deep Learning architectures, (DL) including transformer-based models with self-attention mechanisms (NeNeT).

The results show that DL models significantly improve accuracy, with an average reduction in root mean square error (RMSE) of 30% compared to traditional methods. NeNeT, in particular, demonstrated strong performance in capturing complex spatial relationships, improving height estimates in dense forests of the Basque Country (RMSE was reduced from 7.0 to 4.0 m). In contrast, differences were smaller in the more open Mediterranean forests of Madrid, suggesting that less computationally demanding methods like XGB may be suitable in certain contexts. Despite their advantages, DL models present operational limitations, particularly due to high computational demands. For instance, producing historical maps for the peninsular Spain would require up to four years of processing with NeNeT versus seven months with XGB, assuming no parallelization. Moreover, DL models tend to learn spurious patterns related to image acquisition (e.g., number of observations or dates), which can introduce biases if not properly controlled.

In conclusion, combining lidar and optical sensors with advanced artificial intelligence models enables highly accurate estimation of key variables for wildfire management. However, model choice should balance achieved precision with available computational resources, taking into account the ecosystem type and specific application needs.

Palabras clave: variables forestales, teledetección, series temporales, inteligencia artificial.

Keywords: forest variables, remote sensing, time series, artificial intelligence.

Recibido: 22 Abril 2025

Aceptado: 30 Junio 2025

***Correspondencia:** M. Tanase, Instituto de Ciencias Forestales ICIFOR (INIA-CSIC), Madrid, España. E-mail: mihai@tma.ro

1. Introducción

Los incendios afectan aproximadamente 4 millones de km² cada año, siendo uno de los agentes perturbadores más importantes a nivel global (Giglio *et al.*, 2010) y alterando los ecosistemas de diferentes maneras, desde impactos directos como la combustión de la vegetación y la degradación del suelo, hasta la pérdida de biodiversidad y hábitats, la reducción del valor recreativo o la degradación de las cuencas hidrográficas (Lucas-Borja *et al.*, 2021). Los incendios también contribuyen al aumento de gases de efecto invernadero en la atmósfera mediante emisiones directas (combustión) e indirectas (por ejemplo, deforestación y degradación forestal), siendo un factor importante en el cambio en tipos de cobertura terrestre y, por lo tanto, en los flujos de energía y agua.

La predicción y extinción de incendios, componentes clave en la gestión del fuego, requieren información espacial y temporalmente explícita sobre las condiciones meteorológicas relacionadas con incendios y las propiedades del combustible, que se obtienen cada vez más de datos de teledetección adquiridos por sensores aéreos o satélite (Gale *et al.*, 2021). La planificación en gestión de incendios requiere información detallada sobre las propiedades de la estructura forestal (i.e., combustible), incluyendo la disposición horizontal y vertical en el paisaje (continuidad), así como la cantidad de combustible (carga) presente en diferentes estratos (por ejemplo, dosel, sotobosque), para estimar el riesgo y entender el comportamiento del fuego (Miller and Urban, 2000; Quan *et al.*, 2021). La carga de combustible ejerce efectos diferenciales sobre la propagación e intensidad del fuego, con tasas que aumentan o disminuyen según el tamaño, clase, densidad y tipo del combustible. La variación espacial y temporal del combustible afecta la ocurrencia, intensidad y propagación del fuego, habiéndose demostrado relaciones inversas entre la probabilidad de ignición y las propiedades del combustible vivo (Plucinski *et al.*, 2010). La continuidad del combustible es relevante para muchos aspectos del comportamiento del fuego, ya que la combustión depende en gran medida de la estructura del dosel, la

cual afecta la tasa de propagación e intensidad del incendio. Además, la información sobre la carga del combustible trasciende las necesidades inmediatas de gestión del fuego, como estimación de riesgos o modelado del comportamiento del fuego, ya que su distribución espacial se emplea frecuentemente para evaluar las emisiones generadas por incendios y sus efectos sobre la salud de la población (Cascio, 2018). Las variables de estructura forestal como la fracción de cabida cubierta (FCC) y altura constituyen indicadores indirectos (proxies) para la caracterización de los combustibles ya que están relacionadas con su cantidad y continuidad espacial.

Las técnicas de teledetección, utilizadas desde mediados de la década de 1980 para prevención, evaluación y monitoreo de incendios forestales, abordan múltiples aspectos relacionados con la gestión de incendios, incluyendo la caracterización y monitoreo de las propiedades de combustibles (Aragoneses *et al.*, 2023; Gale *et al.*, 2021). Sin embargo, la caracterización de las propiedades de combustibles está limitada por la forma en que actualmente se emplean los sensores, tanto pasivos como activos. De hecho, las adquisiciones lidar satelitales carecen de cobertura espacial continua y son afectadas por la nubosidad mientras que las señales radar están influenciadas por factores no relacionados con la vegetación (precipitación, topografía), dificultando la estimación de las variables forestales (Tanase *et al.*, 2019). Además, el seguimiento de la recuperación post-incendio de la vegetación (combustible) con sensores ópticos es adecuado solo durante las primeras décadas debido a la saturación de la señal (Tanase *et al.*, 2011). Estas limitaciones resaltan la necesidad de un marco integrado que, mediante sinergias y complementariedades, aproveche las fortalezas de diferentes tipos de sensores para mejorar la información obtenida de cada uno individualmente. La revolución digital, acompañada por una revolución en sensores, proporciona cantidades sin precedentes de datos (Wilson, 2013) que permiten comprender y monitorizar mejor fenómenos dinámicos como los cambios en la estructura forestal. Los avances tecnológicos y la mayor capacidad computacional permiten hoy en día enfoques más complejos, aprovechar sinergias entre diferentes tipos de sensores y utilizar modelos avanzados de inteligencia artificial.

La combinación de adquisiciones ópticas temporalmente consistentes (Landsat) con adquisiciones lidar puntuales o temporalmente dispersas ha permitido mejorar el seguimiento forestal (Caughlin *et al.*, 2021; Hudak *et al.*, 2020), facilitando vistas sinópticas a escala de paisaje durante periodos largos de tiempo (Matasci *et al.*, 2018). La dinámica de la estructura forestal puede ser monitorizada combinando sensores ópticos y sensores lidar (Bolton *et al.*, 2020). Los sensores ópticos permiten identificar fácilmente cambios en la cobertura del dosel forestal (continuidad de combustible), mientras que los sensores lidar proporcionan información precisa sobre la estructura vertical, incluyendo la carga y altura del combustible (Sexton *et al.*, 2009). Por lo tanto, se están desarrollando modelos integrados que permiten aprovechar las fortalezas de distintas tecnologías de sensores (Hudak *et al.*, 2020; Zhang *et al.*, 2019). En este contexto, Tanase *et al.* (2024) han empleado una combinación de datos lidar y ópticos para estimar y posteriormente monitorizar variables estructurales del bosque como biomasa aérea, fracción cabida cubierta (FCC) y la altura del dosel (H), que se pueden asimilar de manera directa con las propiedades del combustible o pueden servir como proxies para su caracterización. Este estudio complementa el trabajo anterior (Tanase *et al.*, 2024) comparando algoritmos ampliamente utilizados en aprendizaje automático (random forests y extreme gradient boosting) con algoritmos recientes de aprendizaje profundo que incorporan mecanismos de auto-atención (transformers). Los transformers han mostrado precisión superior respecto a redes neuronales convolucionales tradicionales, aunque con mayores requerimientos computacionales. Por tanto, el estudio tiene como objetivo comparar algoritmos de aprendizaje automático y profundo para caracterizar variables estructurales que se pueden servir para estimar las propiedades de los combustibles y evaluar la precisión lograda y los recursos computacionales necesarios.

2. Área de estudio

Las áreas de estudio se encuentran en la España peninsular y comprenden dos comunidades autónomas (de aquí en adelante denominadas regiones), Madrid y País Vasco, que cubren 8.028 km² y

7.234 km², respectivamente (Fig. 1). La vegetación, bosques y matorrales, ocupan el 55% de la superficie en la región de Madrid y el 56% en el País Vasco. En la región de Madrid (Fig. 1a) la elevación se extiende entre 400 m sobre el nivel del mar en el sur y 2.400 m en el norte. Según la Agencia Estatal de Meteorología (AEMET), el clima es mediterráneo con veranos cálidos y precipitación media anual de 450 mm, que ocurre principalmente en primavera y otoño, presentando un importante gradiente desde las elevaciones bajas del sur (<500 mm año⁻¹) hacia las montañas del norte (1.500 mm año⁻¹). La temperatura media mensual oscila entre 1°C (invierno) y 32°C (verano) con un gradiente desde sur (más cálido) hacia norte (más frío). Las principales especies arbóreas son las encinas (*Quercus ilex* y *Q. pyrenaica*) y los pinos (*Pinus halepensis*, *P. sylvestris*, *P. pinea* y *P. pinaster*), que representan respectivamente el 26% y el 11,5% de las áreas cubiertas por vegetación natural. Los bosques mixtos ocupan un 12% del territorio, quedando el área restante cubierta mayoritariamente por pastizales (16,5%) y matorrales (33%).

El País Vasco se extiende sobre un terreno principalmente montañoso con elevaciones desde el nivel del mar hasta más de 1551 s.n.m. La región presenta un clima oceánico templado en el norte que cambia gradualmente hacia un clima mediterráneo con características continentales en el sur. Las precipitaciones medias superan los 1.000 mm año⁻¹ en el norte, pero disminuyen hasta 450 mm año⁻¹ en el sur. La temperatura media mensual varía entre 5°C (invierno) y 20°C (verano). El terreno montañoso modula la temperatura, siendo comunes los veranos más cálidos en el sur. Las principales especies arbóreas son las coníferas (43%), mientras que los bosques de haya (*Fagus sylvatica*) y robles (*Q. robur*, *Q. petraea*, *Q. pyrenaica*, *Q. faginea* y *Q. ilex*) ocupan respectivamente el 14% y el 22% de la superficie forestal (Fig. 1b). Las plantaciones de eucalipto ocupan el 6%, mientras que los bosques mixtos atlánticos representan aproximadamente el 10% del área. Según el cuarto Inventario Forestal Nacional español (IFN4), estos bosques se caracterizan por una mayor altura promedio (18,6 m frente a 9,7 m) y una mayor fracción de cabida cubierta del dosel arbóreo (77% frente a 55%) en comparación con los bosques de Madrid.

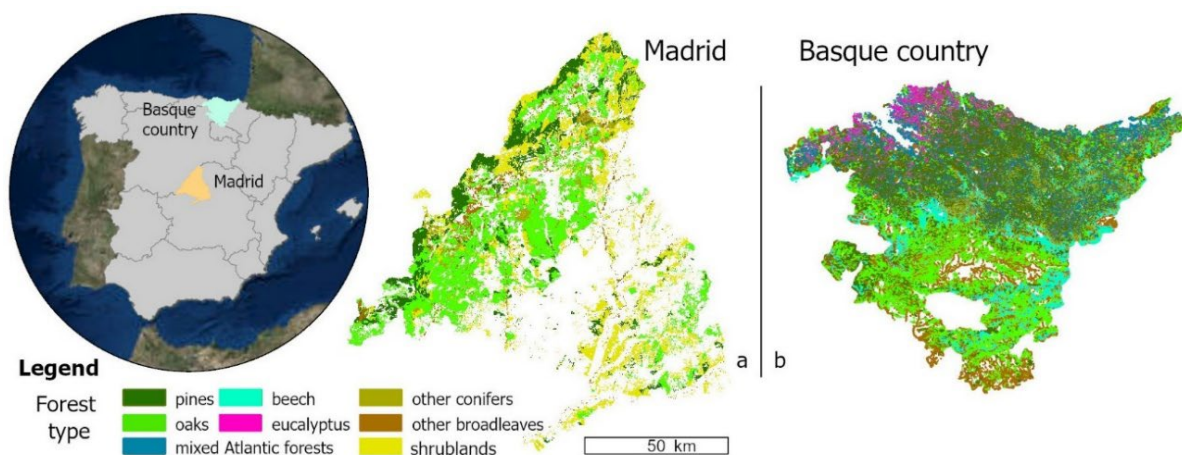


Figura 1 Áreas de estudio, enmascaradas con el mapa forestal, y las principales especies arbóreas.

3. Materiales y Métodos

Para obtener información sobre estructura forestal hemos utilizado modelos de inteligencia artificial entrenados con variables forestales de referencia derivadas de lidar (i.e., altura y fracción cabida cubierta). Además, información topográfica derivada de modelos digitales de elevación ha sido añadida al conjunto de variables predictoras. Este enfoque permite utilizar series temporales para estimar las variables estructurales de interés a largo plazo a partir de modelos entrenados con datos de referencia muy precisos.

3.1. Datos

Los atributos estructurales de interés (FCC y H) se derivaron, a nivel regional, a partir del segundo vuelo lidar realizado en 2016 en Madrid y 2017 en el País Vasco por el Instituto Geográfico Nacional (IGN) dentro del Plan Nacional de Orto fotografía Aérea (PNOA). La altura del bosque (percentil 75 de la elevación normalizada, en metros) y la fracción de cabida cubierta (primeros retornos por encima de dos metros divididos por el número total de primeros retornos dentro de la celda de resolución, en %) se han estimado directamente de las nubes de puntos lidar. Se ha utilizado el percentil 75 ya que muestra la mayor correlación con la altura promedia estimada en el campo en IFN4. Las nubes de puntos clasificadas, en formato comprimido (LAZ), se descargaron del Centro Nacional de Información Geográfica (CNIG) y fueron examinadas para comprobar su consistencia, es decir, densidad de puntos, áreas superpuestas o vacíos, precisión de la clasificación. La altura normalizada (altura sobre el suelo), calculada a partir de modelos digitales de elevación (MDE) generados con los retornos clasificados como terreno, se han utilizado para estimar la H y FCC en píxeles de 30 m, la resolución espacial de las imágenes Landsat. Una altura mínima de 2 m se ha utilizado para limitar los retornos procedentes de la vegetación del sotobosque.

El procesamiento de imágenes Landsat de la Colección 2 Nivel 2, realizado en Google Earth Engine, ha incluido el filtrado temporal para coincidir con el año del vuelo lidar en cada comunidad. Además, el procesamiento ha incluido el mantenimiento de las imágenes con cubierta nubosa inferior al 90%, la estandarización de los nombres de las bandas entre los diferentes sensores (ETM+, OLI), y el filtrado de píxeles afectados por nubes, sombras de nubes, nieve, o con saturación radiométrica a partir de las bandas de calidad disponibles para cada adquisición. También se han aplicado correcciones topográficas (Soenen *et al.*, 2005) para limitar las variaciones de reflectancia de la superficie debido a efectos de iluminación provocados por el relieve (Young *et al.*, 2017). Además, dado que las bandas de calidad proporcionadas con las imágenes Landsat están afectadas por errores (Chen *et al.*, 2021; Souza *et al.*, 2013), se ha utilizado una técnica de desmezclado espectral (Souza *et al.*, 2005) para eliminar nubes residuales y sombras.

Para cada región, las imágenes Landsat procesadas se han utilizado para estimar estadísticas anuales (percentiles 10, 25, 50, 75 y 90, y desviación estándar) para bandas individuales (verde, rojo, infrarrojo cercano, infrarrojo de onda corta) e índices de vegetación (Normalizad Burn Ratio – NBR, Enhanced Vegetation Index – EVI, Normalized Difference Fraction Index – NDFI, y los componentes de humedad y verdor de la transformación Tasseled Cap). El NDFI es un indicador sensible a los cambios en el dosel forestal (Souza *et al.*, 2013). En total, se han calculado 60 variables predictoras a partir de la reflectividad de la superficie Landsat, mientras que otras cuatro variables predictoras se han calculado a partir del número válido de adquisiciones para cada píxel (conteo) y su distribución a lo largo del año (fecha mínima, mediana y máxima). Las últimas cuatro variables predictoras se han calculado utilizando el modelo digital de elevación (MDE) derivado de los datos lidar: elevación, pendiente, orientación y rugosidad. El mapa forestal más reciente, descargado del repositorio en línea del ‘Ministerio para la Transición Ecológica y el Reto Demográfico’ (MITECO), en cada región se ha utilizado para enmascarar las áreas no forestales.

3.2. Modelos

Se han utilizado dos tipos de modelos de inteligencia artificial (IA) para estimar FCC y H a partir de las métricas derivadas de Landsat como variables predictoras: modelos de aprendizaje automático y de aprendizaje profundo. Entre los algoritmos de aprendizaje automático (ML, del inglés Machine Learning), se ha seleccionado el Gradient Boosting (GB), que ha demostrado mejorar la precisión frente a otros tipos de modelos ML (Luo *et al.*, 2019). GB utiliza un algoritmo de aprendizaje basado en árboles combinando el uso de árboles de decisión poco profundos (WL, del inglés Weak Learners), para generar modelos predictivos progresivamente más robustos. Los WL tienen una

capacidad predictiva baja, aunque superior a la predicción al azar (Merghadi *et al.*, 2020), y se benefician de un proceso de potenciación al incorporar árboles adicionales que permiten minimizar la función de pérdida (Freund and Schapire, 1997). El proceso de aprendizaje sigue un enfoque aditivo por etapas, lo que permite al algoritmo mejorar la precisión general al reajustar los errores cuando se añaden nuevos árboles (Jerome, 2001). La versión mejorada, Extreme Gradient Boosting (XGB), incorpora un proceso avanzado de regularización para mejorar la capacidad de generalización en conjuntos de entrenamiento grandes. También se ha utilizado un modelo basado en bosques aleatorios (RF, del inglés random forests) (Breiman, 2001) para establecer una línea base y facilitar comparaciones con estudios previos, ya que se usa ampliamente con datos de teledetección.

Para evaluar la eficacia de los modelos de aprendizaje profundo (DL, del inglés Deep Learning) tipo transformer en teledetección, se compararon diversas arquitecturas de redes neuronales, desde modelos convolucionales simples hasta enfoques híbridos más complejos. El modelo más sencillo (Nested Mean Model, NMM) utiliza dos capas convolucionales con filtros 3x3 y activación lineal rectificada (ReLU, por sus siglas en inglés) para extraer características espaciales, e incorpora una capa de abandono (dropout) para evitar el sobreajuste. Neural Neighbor Embeddings (NeNe), una versión más profunda basada en NMM, añade convoluciones parciales y regulares que permiten procesar datos enmascarados o de baja calidad, conservando la información espacial relevante. U-Net, ampliamente utilizado en tareas de segmentación, combina operaciones de codificación y decodificación en una estructura en “U” que captura tanto características locales como globales; emplea convoluciones parciales, regulares y transpuestas, junto con técnicas de regularización y normalización por lotes. La arquitectura más avanzada, NeNe Transformer (NeNeT), integra convoluciones parciales con capas tipo transformer para capturar relaciones de largo alcance y gestionar eficazmente datos multiescala y multimodales. Esta combinación permite tratar información incompleta y adaptarse a distintos tamaños y resoluciones de entrada. La arquitectura demuestra una gran capacidad de generalización y precisión, siendo especialmente adecuada para tareas complejas en teledetección forestal.

Para los modelos de aprendizaje automático, el conjunto de datos fue dividido en muestras para el entrenamiento (80%) y para la validación (20%) tras extraer unas 10 000 muestras aleatorias utilizando un enfoque de muestreo equilibrado. El equilibrio en el muestreo se ha logrado extrayendo un número constante de muestras para cada intervalo, independientemente del histograma de valores. Para FCC, los intervalos se han definido cada 10 unidades desde 0 hasta 100, extrayéndose 1000 muestras en cada intervalo. Como la altura máxima varía entre regiones, los intervalos de muestreo se han establecido de forma dinámica utilizando percentiles (0, 0,05, 0,25, 0,5, 0,75, 0,95, 0,99, 1) con 1.400 muestras extraídas en cada uno de los siete intervalos. Para reducir la dimensionalidad del conjunto de datos, se ha seleccionado un subconjunto con las variables predictoras más relevantes desde las 68 variables calculadas (sección 3.1). La importancia, calculada utilizando RF, se ha utilizado para ordenar las variables de mayor a menor relevancia, conservándose las 20 primeras (para más detalles véase Tanase *et al.*, 2024).

Los hiperparámetros del modelo XGB se han ajustado mediante una búsqueda en malla (grid search) sobre 384 combinaciones posibles entre el número de iteraciones (150, 200, 250, 300, 350, 400), la profundidad máxima del árbol (4, 5), la tasa de aprendizaje (0,025, 0,05), la intensidad de poda (0,4, 0,8), el porcentaje de muestreo de predictores (0,5, 0,8), el peso mínimo de los nodos (1, 3) y el porcentaje de muestreo de datos de entrenamiento (0,6, 0,8). Los modelos RF fueron entrenados y validados utilizando el mismo marco de trabajo y conjuntos de datos. Sus hiperparámetros se ajustaron mediante una búsqueda en malla sobre 34 combinaciones entre el número de variables en cada árbol (4, 8, 12), el tipo de selección de las variables en cada árbol (varianza, extra trees y maxstat), y el tamaño mínimo de los nodos (3, 5, 10, 15). Posteriormente, utilizando las variables predictoras seleccionadas y los hiperparámetros ajustados, se han entrenado los modelos finales que se han utilizado para predecir las variables forestales de interés sobre el conjunto de validación. La selección de variables predictoras, el ajuste de hiperparámetros y el entrenamiento del modelo se han llevado a cabo para cada variable de interés (FCC, H) en cada región.

Para los modelos DL se ha utilizado la extensión de la variable objetivo para derivar una cuadrícula de referencia sobre la que se han interpolado tanto las variables objetivo como las predictoras. Para mejorar la eficiencia computacional, la cuadrícula de referencia se ha almacenado en sub-cuadrículas de 10x10 km, identificadas por coordenadas (i, j). Posteriormente, las variables objetivo y predictoras se han re-muestrado por las sub-cuadrículas y se han repartido de manera aleatoria en conjuntos de entrenamiento (70%), validación (20%) y prueba (10%), equilibrando la asignación en base al valor del tercer cuartil y los valores extremos de las sub-cuadrículas objetivo (Ec. 1). Para mejorar la asignación de sub-cuadrículas, se ha tenido en cuenta la distribución espacial de las variables forestales de interés, ya que los valores extremos se caracterizan por baja frecuencia y concentración en áreas específicas. Para la normalización de los datos (z-score), se ha utilizado la media y desviación estándar en cada sub-cuadrícula. La parametrización de cada modelo se ha realizado a través de un archivo de configuración que contiene tanto los datos de entrenamiento como los parámetros de construcción del modelo. Los parámetros del modelo incluyen el tipo de arquitectura, así como la tasa de aprendizaje, la función de pérdida, el número de épocas para el entrenamiento, el número de lotes y el tamaño del lote utilizado durante el entrenamiento y la validación. Se han utilizado las 64 variables predictoras derivada de las imágenes Landsat y del MDE. La precisión se ha evaluado mediante el error absoluto medio (MAE, Ec. 2), el error cuadrático medio (RMSE, Ec. 3) y el R^2 (observado vs. predicho). Para XGB, las métricas de error se han calculado sobre las muestras de validación mientras que para el resto de modelos los errores se han calculado sobre las sub-cuadrículas del conjunto de prueba.

$$\text{Upper outlier fence} = Q3 + 1.5RIC = Q3 + 1.5 * |Q3 - Q1| \quad (\text{Ec. 1})$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\text{Predicted}_i - \text{Observed}_i| \quad (\text{Ec. 2})$$

$$\text{RMSE} = \sum_{i=1}^n \frac{1}{n} (\text{Predicted}_i - \text{Observed}_i)^2 \quad (\text{Ec. 3})$$

donde: Q1, Q3 – cuartiles uno y tres, RIC – rango intercuartílico.

4. Resultados y Discusión

En general, las diferencias entre los modelos de aprendizaje automático son marginales, con diferencias entre los valores de R^2 y RMSE dentro del 2–3%, siendo XGB ligeramente más preciso en comparación con RF (Fig. 2). Por variable forestal, la mayor diferencia entre modelos se observa para la altura en el País Vasco, donde el R^2 aumenta de 0,46 a 0,50 y el RMSE disminuye ligeramente (de 7,2 a 7,0 m) al usar el modelo XGB. En los bosques mediterráneos, la precisión fue mayor, con valores de R^2 que varían entre 0,77 y 0,82 para ambos modelos y variables, mientras que el RMSE se mantiene prácticamente constante, en torno a 3 m y 13,5%, independientemente del modelo. Dichos valores representan un 25% (FCC) y un 30% (altura) respecto al promedio de los datos in situ del IFN. Para los bosques atlánticos, los errores RMSE (22,2% y 7,1 m) aumentaron en relación con los promedios in situ, alcanzando el 28% para FCC y el 38% para la altura. Esta sensibilidad diferenciada entre biomas podría explicarse por: i) la menor sensibilidad de los sensores ópticos a la estructura forestal con mayor FCC (saturación de la señal), y ii) la concentración de datos libres de nubes durante el verano, lo cual afecta el cálculo de los percentiles anuales, es decir, menos representativos de los cambios fenológicos anuales y por tanto sensibilidad reducida para variables estructurales.

El análisis de las arquitecturas de modelos muestra diferencias notables en R^2 y en los errores de estimación (MAE y RMSE), siendo las arquitecturas de aprendizaje profundo superiores a los modelos de aprendizaje automático en ambas regiones (Tabla 1 y Fig. 3). En general, NeNeT supera XGB en aproximadamente un 26% para R^2 y un 30% para RMSE, promediando entre regiones y métricas estructurales. En particular, el uso de modelos DL reduce el RMSE en los bosques del País Vasco, de 7 m (XGB) a 4 m (NeNeT), o en términos relativos, de 37,0% a 21,5% con respecto al valor medio in situ de 18,6 m. Para el FCC, el RMSE se reduce de 22,2% a 14,2% con el modelo NeNeT. En los bosques mediterráneos (Madrid), las diferencias fueron menores, especialmente para FCC, que

disminuye de 13,4% a 12,2%, lo que sugiere que el uso de modelos computacionalmente menos costosos puede considerarse una opción viable, al menos cuando se necesita información sobre la continuidad del combustible. Sin embargo, para la altura, el uso de modelos DL proporciona una reducción sustancial del RMSE, de 2,9 a 1,8 m, o en términos relativos, del 30% al 20% respecto al valor promedio in situ de 9,7 m. Las diferencias observadas entre variables pueden estar relacionadas con el contenido informativo de los datos ópticos, que se relacionan más con la capa superior del dosel que con la estructura tridimensional del bosque. Los resultados sugieren que la capacidad de generalización de los modelos ML se reduce considerablemente para variables estructurales forestales que no están directamente relacionadas con la información proporcionada por los sensores ópticos, que se vinculan principalmente con la cobertura del bosque. En cambio, los modelos DL parecen capaces de identificar relaciones adicionales entre las variables objetivo y predictoras, como lo demuestra la gran reducción del RMSE para la altura del bosque, una variable estructural no directamente relacionada con la reflectancia de la superficie, independientemente de la región.

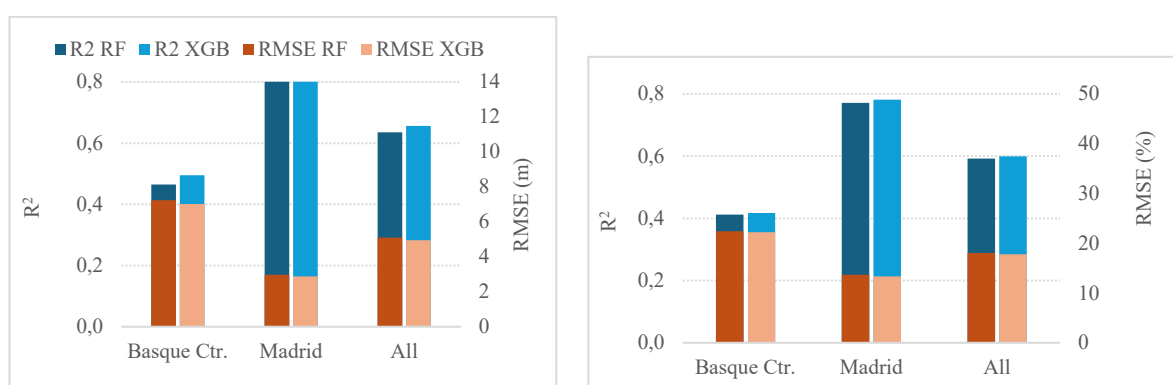


Figura 2 Precisión de la estimación de altura (izquierda) y fracción cabida cubierta (derecha) a partir de modelos de aprendizaje automático: Random Forests (RF) y eXtreme Gradient Boosting (XGB).

Tabla 1 Precisión de diferentes arquitecturas de modelos para la altura del bosque (H) y la fracción cabida cubierta (FCC). Los datos de referencia se han estimado a partir del segundo vuelo lidar. XGB-extreme gradient boost, NMM-nested mean model, NeNe-Neural Neighbor Embeddings, NeNeT- Neural Neighbor Embeddings with transformers.

Región Variable	Madrid (2016)						Basque country (2017)					
	H			FCC			H			FCC		
Model	R²*	MAE	RMSE	R²	MAE	RMSE	R²	MAE	RMSE	R²	MAE	RMSE
XGB	0,82	2,0	2,9	0,78	10,1	13,4	0,50	5,5	7,0	0,42	18,1	22,2
NMM	0,49	1,3	2,9	0,79	9,1	12,2	0,64	3,4	4,4	0,71	10,8	14,5
U-Net	0,69	1,6	2,4	0,78	9,2	12,5	0,61	3,5	4,6	0,62	11,9	16,7
NeNe	0,67	1,7	2,4	0,79	9,2	12,3	0,64	3,3	4,4	0,70	10,7	14,8
NeNeT	0,78	1,41	1,9	0,81	9,1	12,2	0,64	3,1	4,0	0,73	10,3	14,2

* R² – coeficiente de determinación (observado vs. predicho), MAE – error medio absoluto, RMSE – error medio cuadrático

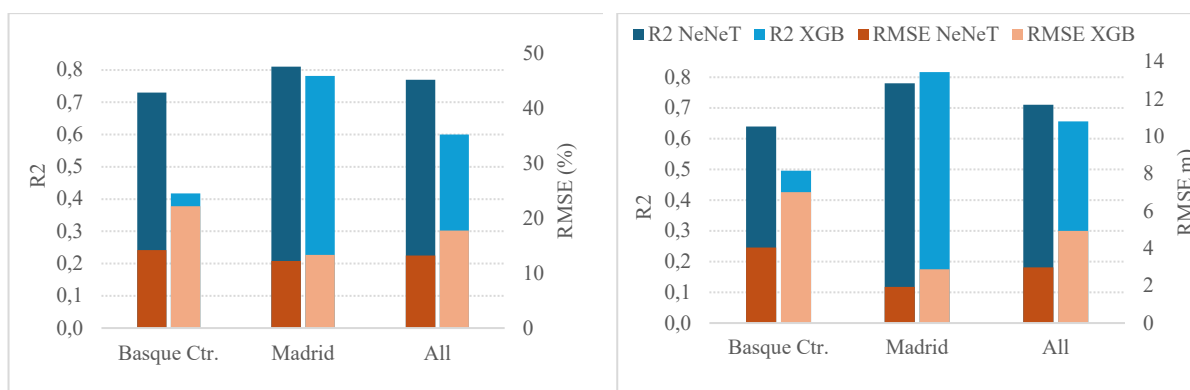


Figura 3 Precisión de la estimación de altura (izquierda) y fracción cabida cubierta (derecha) a partir de modelos de aprendizaje automático (eXtreme Gradient Boosting - XGB) y aprendizaje profundo (NeNe Transformer- NeNeT).

Aunque los modelos de aprendizaje profundo proporcionan estimaciones más precisas, su principal inconveniente para un uso operativo es la mayor demanda computacional, lo que se traduce en un incremento del tiempo necesario tanto para el entrenamiento como, especialmente, para predecir largos periodos. El tiempo medio necesario para entrenar cada modelo DL en la región de Madrid fue de 35 horas en un pequeño clúster (usando 8 CPUs y una GPU), con el número de épocas fijado en 30. En comparación, el tiempo de entrenamiento para XGB fue de aproximadamente 4,5 horas. Más importante aún, el tiempo requerido para predecir una variable para un año en Madrid fue de aproximadamente una hora con el modelo XGB, mientras que con el modelo NeNeT llegó a nueve horas. Si consideramos el tiempo de predicción por kilómetro cuadrado, se necesitarían unos siete meses para generar la cartografía de los últimos 40 años (FCC y H) en la España peninsular usando el modelo XGB, y unos 48 meses con el modelo NeNeT si las regiones y/o variables no se procesan en paralelo.

Otro inconveniente de los modelos DL es su "capacidad" de aprender patrones espaciales no necesariamente relacionados con la variable objetivo. Las variables predictoras como el número de imágenes y la fecha del año afectan los patrones espaciales. La utilidad de estas variables radica en proporcionar información sobre la temporalidad de las adquisiciones, que puede variar particularmente en las primeras décadas (1985-2000) cuando solo un satélite Landsat estaba operativo. Estos patrones están relacionados con las propiedades de adquisición de imágenes (cobertura nubosa, fallos del sensor), y no con patrones espaciales relacionados con la variable objetivo. El uso de convoluciones espaciales hace que los modelos DL aprendan dichos patrones, lo que se traduce en predicciones erróneas. Estos problemas pueden abordarse parcialmente eliminando las variables relacionadas con el conteo y la fecha de adquisición del conjunto de predictores, aunque su importancia fue alta durante el entrenamiento de los modelos ML. Sin embargo, los patrones espaciales causados por el fallo del Scan Line Corrector (Landsat 7) son más difíciles de gestionar, especialmente cuando este satélite fue el único activo (2012-2013). Eliminar por completo Landsat 7 del proceso de composición temporal podría resolver este problema, pero a costa de reducir el contenido informativo, ya que los percentiles calculados son menos fiables durante los primeros años (antes de 2014) y en las regiones del norte, caracterizadas por una alta frecuencia de nubes. En dichas regiones, la falta de suficientes píxeles válidos (no afectados por nubes) sesga los percentiles hacia los periodos del año con menor cobertura nubosa (verano), en detrimento de la primavera, el otoño y el invierno, estaciones con importantes cambios fenológicos y, por tanto, con un mayor contenido informativo.

5. Conclusiones

Este estudio demuestra que la combinación de imágenes ópticas temporalmente consistentes con observaciones lidar esporádicas mejora la capacidad de monitorizar los bosques, lo cual es esencial

para mejorar la evaluación de la dinámica de forestal y así del riesgo de incendios forestales y de la modelización de su comportamiento. Nuestro análisis comparativo pone de manifiesto las fortalezas y limitaciones de distintos enfoques. Los modelos de aprendizaje profundo, que emplean arquitecturas tipo transformer con mecanismos de autoatención (NeNeT), mejoran notablemente la precisión de las estimaciones, logrando una reducción promedia del RMSE de 30% en comparación con los métodos basados en aprendizaje automático, según los resultados en dos regiones ubicadas en la España peninsular. NeNeT captura eficazmente las complejas dependencias espaciales y espectrales inherentes a los datos de teledetección, mostrando notables mejoras en la estimación de atributos estructurales forestales como la altura del dosel. Sin embargo, a pesar de su mayor precisión, los modelos DL presentan desafíos prácticos relacionados con las demandas de recursos computacionales. El tiempo para su entrenamiento es sustancialmente más elevado, requiere hasta ocho veces más tiempo que los métodos basados en aprendizaje automático. Por lo tanto, predecir las variables forestales en periodos temporales prolongados y extensiones espaciales amplias, como el conjunto de datos de la España peninsular, resulta mucho más lento y, por lo tanto, hay limitaciones en su uso operacional sin una considerable infraestructura computacional.

Los resultados indican que los modelos DL avanzados ofrecen ganancias sustanciales de precisión en entornos caracterizados por bosques altos y densos donde los enfoques ML tradicionales muestran limitaciones debido a su dependencia de datos ópticos que capturan principalmente la cobertura del dosel. Por el contrario, en regiones de bosque mediterráneo abierto, los métodos ML ofrecen un rendimiento equivalente en la estimación de la fracción de cabida cubierta, constituyendo una alternativa viable considerando la capacidad computacional necesaria.

Agradecimientos

El trabajo ha sido financiado por el Ministerio de Ciencia e Innovación a través de los proyectos PID 2020-114062RA-I00 y CNS 2022-135251.

Referencias

- Aragoneses, E., García, M., Salis, M., Ribeiro, L.M., Chuvieco, E. (2023). Classification and mapping of European fuels using a hierarchical, multipurpose fuel classification system. *Earth System Science Data*, 15, 1287-1315. <https://doi.org/10.5194/essd-15-1287-2023>
- Bolton, D.K., Tompalski, P., Coops, N.C., White, J.C., Wulder, M.A., Hermosilla, T., Queinnec, M., Luther, J.E., van Lier, O.R., Fournier, R.A., Woods, M., Treitz, P.M., van Ewijk, K.Y., Graham, G., Quist, L. (2020). Optimizing Landsat time series length for regional mapping of lidar-derived forest structure. *Remote Sensing of Environment*, 239, 111645. <https://doi.org/10.1016/j.rse.2020.111645>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- Cascio, W.E. (2018). Wildland fire smoke and human health. *Science of the Total Environment*, 624, 586-595. <https://doi.org/10.1016/j.scitotenv.2017.12.086>
- Caughlin, T.T., Barber, C., Asner, G.P., Glenn, N.F., Bohlman, S.A., Wilson, C.H. (2021). Monitoring tropical forest succession at landscape scales despite uncertainty in Landsat time series. *Ecological Applications*, 31, e02208. <https://doi.org/10.1002/eap.2208>
- Chen, S., Woodcock, C.E., Bullock, E.L., Arévalo, P., Torchinava, P., Peng, S., Olofsson, P. (2021). Monitoring temperate forest degradation on Google Earth Engine using Landsat time series analysis. *Remote Sensing of Environment*, 265, 112648. <https://doi.org/10.1016/j.rse.2021.112648>
- Freund, Y., Schapire, R.E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55, 119-139. <https://doi.org/10.1006/jcss.1997.1504>

- Gale, M.G., Cary, G.J., Van Dijk, A.I.J.M., Yebra, M. (2021). Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour. *Remote Sensing of Environment*, 255, 112282. <https://doi.org/10.1016/j.rse.2020.112282>
- Giglio, L., Randerson, J., Van der Werf, G., Kasibhatla, P., Collatz, G., Morton, D., DeFries, R. (2010). Assessing variability and long-term trends in burned area by merging multiple satellite fire products. *Biogeosciences*, 7, 1171-1186. <https://doi.org/10.5194/bg-7-1171-2010>
- 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. *Environmental Research Letters*, 15, 095003. <http://doi.org/10.1088/1748-9326/ab93f9>
- Jerome, H.F. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29, 1189-1232. <http://doi.org/10.1214/aos/1013203451>
- Lucas-Borja, M.E., Delgado-Baquerizo, M., Muñoz-Rojas, M., Plaza-Álvarez, P.A., Gómez-Sánchez, M.E., González-Romero, J., Peña-Molina, E., Moya, D., de las Heras, J. (2021). Changes in ecosystem properties after post-fire management strategies in wildfire-affected Mediterranean forests. *Journal of Applied Ecology*, 58, 836-846. <https://doi.org/10.1111/1365-2664.13819>
- Luo, X., Lin, F., Chen, Y., Zhu, S., Xu, Z., Huo, Z., Yu, M., Peng, J. (2019). Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. *Scientific Reports*, 9, 15369. <https://doi.org/10.1038/s41598-019-51941-z>
- Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Bolton, D.K., Tompalski, P., Bader, C.W. (2018). Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. *Remote Sensing of Environment*, 216, 697-714. <https://doi.org/10.1016/j.rse.2018.07.024>
- Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D.T., Avtar, R., Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207, 103225. <https://doi.org/10.1016/j.earscirev.2020.103225>
- Miller, C., Urban, D.L. (2000). Connectivity of forest fuels and surface fire regimes. *Landscape Ecology*, 15, 145-154. <https://doi.org/10.1023/A:1008181313360>
- Plucinski, M.P., Anderson, W.R., Bradstock, R.A., Gill, A.M. (2010). The initiation of fire spread in shrubland fuels recreated in the laboratory. *International Journal of Wildland Fire*, 19, 512-520. <http://doi.org/10.1071/WF09038>
- Quan, X., Li, Y., He, B., Cary, G.J., Lai, G. (2021). Application of Landsat ETM+ and OLI Data for Foliage Fuel Load Monitoring Using Radiative Transfer Model and Machine Learning Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 5100-5110. <https://doi.org/10.1109/JSTARS.2021.3062073>
- Sexton, J.O., Bax, T., Siqueira, P., Swenson, J.J., Hensley, S. (2009). A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America. *Forest Ecology and Management*, 257, 1136-1147. <http://doi.org/10.1016/j.foreco.2008.11.022>
- Soenen, S.A., Peddle, D.R., Coburn, C.A. (2005). SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing*, 43(9), 2148-2159. <https://doi.org/10.1109/TGRS.2005.852480>
- Souza, C.M., Roberts, D.A., Cochrane, M.A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment*, 98, 329-343. <https://doi.org/10.1016/j.rse.2005.07.013>
- Souza, C.M., Siqueira, J.V., Sales, M.H., Fonseca, A.V., Ribeiro, J.G., Numata, I., Cochrane, M.A., Barber, C.P., Roberts, D.A., Barlow, J. (2013). Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon. *Remote Sensing*, 5(11), 5493-5513. <https://doi.org/10.3390/rs5115493>

- Tanase, M.A., de la Riva, J., Santoro, M., Pérez-Cabello, F., Kasischke, E. (2011). Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. *Remote Sensing of Environment*, 115, 2075-2085. <https://doi.org/10.1016/j.rse.2011.04.009>
- Tanase, M.A., Mihai, M.C., Miguel, S., Cantero, A., Tijerin, J., Ruiz-Benito, P., Domingo, D., Garcia-Martin, A., Aponte, C., Lamelas, M.T. (2024). Long-term annual estimation of forest above ground biomass, canopy cover, and height from airborne and spaceborne sensors synergies in the Iberian Peninsula. *Environmental Research*, 259, 119432. <https://doi.org/10.1016/j.envres.2024.119432>
- Tanase, M.A., Villard, L., Pitar, D., Apostol, B., Petrila, M., Chivulescu, S., Leca, S., Borlaf-Mena, I., Pascu, I.-S., Dobre, A.-C., Pitar, D., Guiman, G., Lorent, A., Anghelus, C., Ciceu, A., Nedea, G., Stanculeanu, R., Popescu, F., Aponte, C., Badea, O. (2019). Synthetic aperture radar sensitivity to forest changes: A simulations-based study for the Romanian forests. *Science of the Total Environment*, 689, 1104-1114. <https://doi.org/10.1016/j.scitotenv.2019.06.494>
- Wilson, J. (2013). Remote Sensing Technologies and Global Markets. In, *Market Reserch. Instrumentation And Sensors* BBC Research LLC.
- Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R., Evangelista, P.H. (2017). A survival guide to Landsat preprocessing. *Ecology*, 98, 920-932. <https://doi.org/10.1002/ecy.1730>
- Zhang, L., Shao, Z., Liu, J., Cheng, Q. (2019). Deep Learning Based Retrieval of Forest Aboveground Biomass from Combined LiDAR and Landsat 8 Data. *Remote Sensing*, 11, 1459. <https://doi.org/10.3390/rs11121459>