



2025 | 16-20
GIJÓN | JUNIO

9º CONGRESO FORESTAL ESPAÑOL

9CFE-1454

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

Organiza





Modelos predictivos de variables forestales en masas de *Eucalyptus globulus*, *Pinus pinaster* y *Pinus radiata* en el Norte de España a partir de imágenes de satélite ópticas y radar.

TEIJIDO-MURIAS, I., LÓPEZ-SÁNCHEZ, C.A. y BARRIO-ANTA, M.

Grupo de Investigación SMartForest. Departamento de Biología de Organismos y Sistemas, Escuela Politécnica de Mieres, Universidad de Oviedo, Campus Universitario de Mieres, C/Gonzalo Gutiérrez Quirós S/N, 33600 Mieres (España).

Resumen

En este trabajo se desarrollaron modelos automáticos para estimar cinco variables forestales de masa, tales como el número de pies por hectárea, el área basimétrica, la altura dominante, el volumen total con corteza y la biomasa aérea, basándonos en datos de teledetección y variables auxiliares en plantaciones de *Pinus pinaster*, *Pinus radiata* y *Eucalyptus globulus*. El estudio se llevó a cabo en la región forestal más productiva de España, Galicia, Asturias, Cantabria y País Vasco. Se utilizaron como datos de entrenamiento las parcelas del Inventario Forestal Nacional de España (IFN) para obtener la verdad-terreno, y dos fuentes de datos de teledetección: las imágenes ópticas de la constelación Sentinel-2 y las imágenes radar de la constelación Sentinel-1, a partir de las cuales se calcularon diferentes índices y variables de textura. Además, se incluyó información auxiliar sobre climatología y terreno. Se desarrollaron dos modelos alternativos para cada especie y variable forestal: un modelo reducido, que excluye la información radar, y un modelo completo, que considera todas las variables explicativas disponibles. La generación de estos modelos permitirá mejorar la gestión y planificación forestal en la región, ofreciendo estimaciones precisas de variables clave para la evaluación y monitoreo de los recursos forestales.

Palabras clave

Sentinel-2, sentinel-1 , random forest , IFN , google earth engine

1. Introducción

Las plantaciones forestales de pino pinaster (*P. pinaster*), pino radiata (*P. radiata*) y eucalipto blanco (*E. globulus*) ocupan un 34% de la superficie forestal arbolada de la España Atlántica (Galicia, Asturias, Cantabria y País Vasco) (MAPAMA, 2019a) y proporcionan el 88% del volumen de madera que se corta en la región (período 2005-2021). Este volumen de corta representó, en el mismo período, el 62% de volumen total cortado en toda España; lo que convierte a estas cuatro Comunidades en la región con mayor producción forestal del país. Debido a su gran relevancia socioeconómica, el poder contar con una fuente de información adecuada y precisa sobre los recursos que pueden proporcionar estas plantaciones es de gran importancia para los diferentes grupos de interés (gobiernos regionales, propietarios de bosques, empresas madereras, etc.). Hasta hace pocos años, el inventario forestal tradicional era la única herramienta para obtener esta información. Sin embargo, en los últimos años la investigación sobre el uso de datos de teledetección ha demostrado su capacidad para la generación de modelos ráster de información forestal espacialmente continua a diferentes escalas espaciales y temporales.

En este sentido, los modelos de alta resolución espacial para estimar variables forestales a escala regional o suprarregional a partir de datos públicos de LiDAR aerotransportado proporcionan las estimaciones más precisas. Sin embargo, el alto



costo de esos vuelos determina su baja resolución temporal (6 años o más), y reduce la utilidad práctica de estos modelos para obtener estimaciones de máxima actualidad del recurso. En estas circunstancias, el uso de estimaciones basadas en datos de satélites con alta resolución temporal y espacial constituye una solución adecuada para completar la información sobre el recurso forestal entre dos adquisiciones públicas de LiDAR aerotransportado.

El programa Copernicus de la Agencia Espacial Europea pone a disposición de la sociedad datos de sus satélites (constelaciones, Sentinel-1 y Sentinel-2) y constituye una muy buena opción para obtener información forestal de alta resolución espacial y temporal para el monitoreo forestal. La constelación de satélites Sentinel-1 (S1) consiste en dos satélites, Sentinel-1A y Sentinel-1B (lanzados al espacio en 2014 y 2016, respectivamente), equipados con radar de apertura sintética (SAR). Estos satélites operan en una frecuencia de radar en la banda C (5.405 GHz) y proporcionan cobertura global cada 12 días con un satélite y cada 6 días con ambos satélites (actualmente, Sentinel-1B está fuera de servicio), con una resolución espacial en el modo de barrido interferométrico ancho (IW) de 5 x 20 metros. Este sistema admite polarización simple (HH o VV) y polarización dual (HH+HV o VV+VH). En 2015, la Agencia Espacial Europea (ESA) inició la misión Sentinel-2 (S2) lanzando Sentinel-2A (S2A), más tarde fue complementada con el lanzamiento de Sentinel-2B (S2B) en 2017. Los satélites Sentinel-2 utilizan 13 bandas espectrales a través del espectro electromagnético, ofreciendo diversas perspectivas sobre la superficie terrestre. La resolución espacial varía entre bandas, con un rango de 10 a 60 metros por píxel, y la resolución temporal es de 5 días.

La combinación de los datos de S1 y S2 puede contribuir a mejorar la precisión de modelos al aportar los datos de S1 información sobre la estructura del bosque imposible de obtener a partir de datos S2 (FANG et al., 2023). Incluso, en casos especiales, la información de radar S1 se puede utilizar para completar o rellenar la usencia de información óptica debido a la existencia de coberturas de nubes persistentes en las áreas de estudio (ARANGUREN ERICE, 2021).

Por otro lado, la información de campo es siempre necesaria para obtener los valores reales de las variables forestales de interés (verdad-terreno), que posteriormente se convertirán en las variables dependientes de los modelos a desarrollar en función de métricas obtenidas de sensores satelitales. En este sentido, las parcelas del Inventario Forestal Nacional (IFN) proporcionan variables forestales que se actualizan cada 5 años para las especies más productivas de la zona norte de España (MAPAMA, 2019b), por lo que constituye la fuente de datos de campo más valiosa y fiable para el desarrollo a gran escala de este tipo de modelos.

Por tanto, la disponibilidad gratuita y en abierto de datos de satélites de observación de la tierra y de la información de campo proporcionada por el IFN, permite actualmente desarrollar modelos de alta resolución para estimar variables forestales a escala regional o suprarregional. Además, el uso de modernas herramientas de computación en la nube (*cloud-computing*), resultan hoy en día indispensables para obtener una mayor eficiencia en los procesos de cálculo necesarios debido a que se manejan ingentes cantidades de datos. Así, en el ámbito de datos geoespaciales, herramientas como *Google Earth Engine* (GEE) facilitan el acceso a la información de teledetección, acelera el procesamiento de datos y permite compartir metodologías y procedimientos estandarizados para análisis de



los datos. Todo ello está permitiendo la obtención de información precisa y actualizable sobre los ecosistemas forestales con bajo coste y alta resolución espacial y temporal abriendo un nuevo paradigma en la gestión de los recursos forestales a escala global.

2. Objetivos

El presente estudio ha tenido dos objetivos principales: i) desarrollar quince modelos para estimar el número de pies por hectárea (N), área basimétrica (G), altura dominante (H0), volumen total sobre corteza (TV) y biomasa aérea (AGB) para cada una de las tres especies, *Pinus pinaster*, *Pinus radiata* y *Eucalyptus globulus*, en el norte de España utilizando datos de S1, S2 y variables auxiliares en la plataforma *Google Earth Engine* y ii) analizar la contribución de los datos S1 al poder explicativo de los modelos.

3. Materiales y Métodos

3.1. Flujo de trabajo

La descripción detallada de todo el flujo de trabajo realizada se presenta en la figura 1 y la desarrollo de cada uno de los bloques se puede ver en los aparatos siguientes.



3.1. Datos utilizados

Se utilizaron cuatro tipos diferentes de datos para desarrollar los modelos forestales basados en teledetección: i) datos de campo; ii) datos de teledetección; iii) datos de terreno; y iv) datos climáticos.

3.1.1. Datos de campo

Los datos de campo utilizados en este estudio se obtuvieron de la actualización en 2018 del IFN4 para las especies productivas más importantes del norte de España (IFN 4.5) (MAPAMA, 2019a). Se seleccionaron las parcelas para las tres especies de interés en masas puras (80% del área basal total dentro de la parcela perteneciente a la especie de interés). Siguiendo este criterio, se dispuso de un total de 1.471 parcelas dentro del área de estudio, perteneciendo 589 parcelas a *E. globulus*, 474 a *P. pinaster* y 408 a *P. radiata*.

3.1.2. Datos de teledetección

Los datos de teledetección utilizados se obtuvieron a través de la plataforma *Google Earth Engine*, concretamente, accediendo los repositorios de Sentinel-1 y Sentinel-2. Los datos S1 usados correspondieron a imágenes de polarización dual (HH+HV) obtenidas para el año 2018 y albergadas en el repositorio Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected (COPERNICUS/S1_GRD). Por otro lado, las imágenes S2 con un nivel de procesamiento 1C fueron extraídas del repositorio Harmonized Sentinel-2 MSI: MultiSpectral Instrument (COPERNICUS/ S2_HARMONIZED) para el período de estudio (2018).

3.1.3. Variables auxiliares

Se utilizaron dos tipos de variables auxiliares: variables climáticas y variables del terreno. Los datos climáticos se extrajeron del Atlas Climático de la Península Ibérica (NINYEROLA et al., 2005) e incluyen valores de precipitación, temperatura mínima, media y máxima y radiación. Por otra parte, los datos del terreno se obtuvieron a partir del modelo digital de elevaciones (MDE) con 5 m de resolución espacial del Centro Nacional de Información Geográfica (CNIG).

3.2. Preprocesamiento de datos

Los datos de las parcelas de campo del IFN 4.5 seleccionadas se procesaron para obtener las siguientes variables de estado de la masa: dos variables de densidad (número de pies por hectárea y área basimétrica), una variable relacionada con el tamaño (altura dominante) y dos variables de producción (volumen total con corteza y biomasa aérea total). La Tabla 1 incluye los valores de los estadísticos descriptivos de las variables anteriores para las tres especies forestales consideradas en el área de estudio.

Tabla 1. Estadísticos descriptivos de las variables dependientes analizadas (número de pies por hectárea, N; área basimétrica, G; altura dominante, H0; volumen total con corteza, TV; y biomasa aérea, AGB) y extraídas de las parcelas del IFN 4.5 donde G de la especie dominante 80% G total de la parcela.

Especie/	Nº parcelas	Estad.	Variable Forestal						
N	(pies×ha-1)	G	(m2×ha-1)	H	0 (m)	TV	(m3×ha-1)	AGB	(Mg×ha-1)

<i>E. globulus</i>	(589 parcelas)	Mean	833,83	18,30	21,43	148,42	99,44
Min.	10,19	0,44	6,70	0,68		0,98	
Max.	2,695,02	52,25	43,55	522,67		371,55	
Std.	499,93	0,44	7,26	118,14		81,68	
<i>P. pinaster</i>	(474 parcelas)	Mean	574,60	22,60	16,67	164,05	92,26
Min.	10,19	0,42	3,40	0,88		0,80	
Max.	3,176,03	55,73	31,78	460,72		298,64	
Std.	439,15	13,70	6,34	119,37		68,25	
<i>P. radiata</i>	(408 parcelas)	Mean	453,66	27,82	22,55	246,23	127,43
Min.	25,46	0,67	5,70	2,25		1,59	
Max.	1,773,48	66,62	39,55	699,31		356,93	
Std.	294,07	13,54	6,28	147,64		75,38	

Los conjuntos de datos de teledetección disponibles en *Google Earth Engine* tienen el preprocesamiento mínimo para preservar máximo de información posible. En función de las necesidades específicas de cada aplicación, el usuario puede realizar diferentes preprocesamientos para las colecciones de imágenes. En este caso, se utilizaron dos formas diferentes de optimizar los datos de teledetección en GEE, una para los datos de Sentinel-1 y otra para los datos de Sentinel-2.

Respecto a los datos de Sentinel-1, se seleccionaron las mejores imágenes para el año 2018 basadas en la cobertura, dirección y polarización. El proceso seguido para corregir los datos de Sentinel-1 se pueden consultar en MULLISSA et al. (2021) e incluyó corrección de ruido de borde, filtrado de *speckle* y normalización radiométrica del terreno. Después de aplicar las correcciones, se obtuvieron dos índices SAR: Índice de Vegetación Radar (RVIVV) (TRUDEL et al., 2012) e Índice de Doble Depolarización Vertical (VDDPI) (PERIASAMY, 2018) y cuatro operaciones básicas entre bandas (suma, resta, multiplicación y división) (tabla 2).

En cuanto a los datos de Sentinel-2, se seleccionaron las mejores imágenes para el año 2018 basadas en el ángulo solar y el porcentaje de cobertura de nubes. El proceso para corregir las imágenes de Sentinel-2 ha sido el seguido en NOVO-FERNÁNDEZ et al. (2024) e incluyó correcciones atmosféricas, topográficas y BRDF. Después de aplicar las correcciones, se obtuvieron 21 índices espectrales y 18 variables de textura basadas en el Índice de Diferencia Normalizada de Vegetación (NDVI) propuestas por HARALICK et al. (1973) (ver tabla 2).

Tabla 2. Grupos de variables y variables consideradas como candidatas para utilizar como predictoras en los modelos de producción.

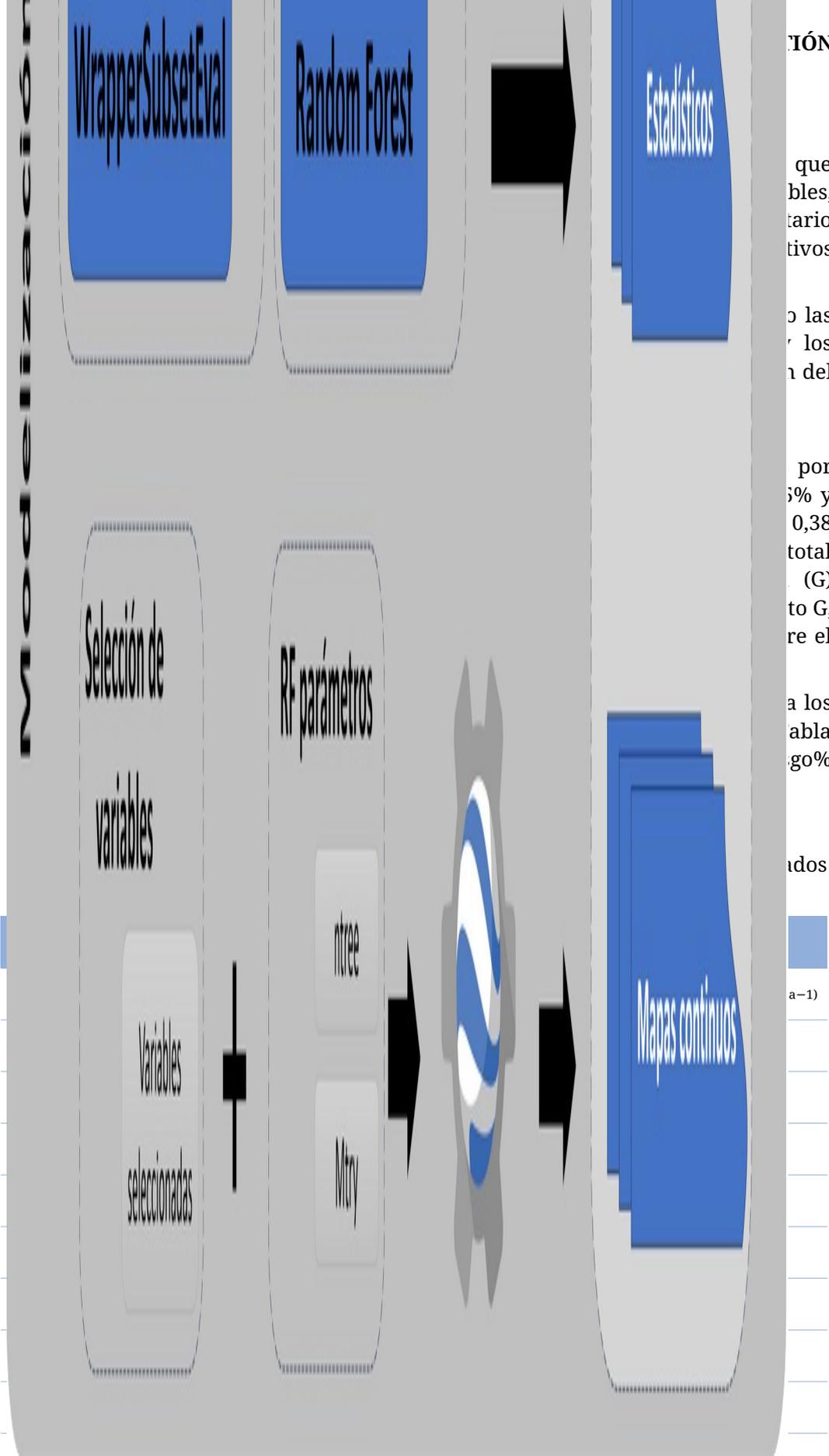
Grupo	Nombre de la variable
Bandas espectrales	Banda 2 - Azul (B2), Banda 3 - Verde (B3), Banda 4 - Rojo (B4), Banda 5 - Vegetación Red-Edge-1 (B5), Banda 6 - Vegetación Red-Edge-2 (B6), Banda 7 - Vegetación Red-Edge-3 (B7), Banda 8 - NIR (B8), Banda 8A - NIR estrecho (B8A), Banda 11 - SWIR-1 (B11), Banda 12 - SWIR-2 (B12).
Índices espectrales	Índice de Reflectancia de Antocianinas (ARI), Índice de Red-Edge de Clorofila (CRE), Índice de Vegetación Mejorado (EVI), Índice de Vegetación Mejorado 2 (EVI2), Índice de Diferencia Normalizada de Vegetación Verde (GNDVI), Índice de Reflectancia de Antocianinas Modificado (MARI), Índice Modificado de Absorción de Clorofila en Reflectancia (MCARI), Índice Modificado de Vegetación Ajustado al Suelo (MSAVI), Índice Modificado de Vegetación Ajustado al Suelo (MSI), Índice Normalizado de Quema (NBR), Índice Normalizado de Quema 2 (NBR2), Índice de Humedad de Diferencia Normalizada (NDMI), Índice de Diferencia Normalizada de Vegetación (NDVI), Índice Simple Específico de Pigmentos (PSSR), Índice de Vegetación Ajustado al Suelo (SAVI), Ángulo de la Capacitancia de Tasseled (TCA), Brillo de Tasseled (TCB), Verdes de Tasseled (TCG), Humedad de Tasseled (TCW).
VARIABLES DE TEXTURA	Momento Angular Segundo (SEC), Contraste (CON), Correlación (COR), Disimilitud (DIS), Energía (ENE), Entropía (ENT), Homogeneidad (HOM), Máximo (MAX), Promedio (MEN), Desviación Estándar (STD).
VARIABLES DE TERRENO	Aspecto (ASP), Relación Aspecto/Espesor (ASR), Curvatura (CU), Elevación (ELV), Índice de Carga Térmica (HLI), Curvatura de Plan (PLC), Curvatura de Perfil (PFC), Pendiente (SLP), Índice de Forma del Terreno (TSD), Índice de Humedad (WI).
VARIABLES DE CLIMA	Temperatura promedio (TM), Temperatura máxima (TMAX), Temperatura mínima (TMIN), Precipitación (PT), Radiación (RA)
BANDAS RADAR	Polarización Vertical-Vertical (VV), Polarización Vertical-Horizontal (VH)
ÍNDICES RADAR	Índice de Vegetación de Radar (RVIVV), Índice de Doble Depolarización Vertical (VDDPI), Diferencia VVVH, suma, producto y división (VVVH_div, VVVH_sum, VVVH_prod, VVVH_ratio).

Finalmente, los datos obtenidos a partir de las parcelas del IFN4.5 y todo el conjunto de variables candidatas a predictoras de los modelos (Tabla 2) subieron al almacenamiento GEE.

3.3. Análisis de datos

Para cada especie y variable del forestal se elaboraron dos tipos de modelos predictivos: i) un modelo completo considerando todas las variables explicativas disponibles descritas en la tabla 2 y ii) un modelo reducido sin las variables radar de S1. El objetivo de este procedimiento fue analizar el poder predictivo de las variables de radar S1 en comparación con el uso solo de variables ópticas S2 en la explicación de la variabilidad de las variables forestales analizadas.

Se realizó una reducción de dimensionalidad de las variables para minimizar la colinealidad de los datos en el software *Weka* utilizando métodos *Wrapper* (ZHIWEI & XINGHUA, 2010) y a continuación se llevó a cabo una optimización de los parámetros asociados (*mtry* y *Nvar*) para el modelo *Random forest* (BREIMAN, 2001) en el software R. Para la evaluación de los modelos, se utilizó el procedimiento *10 repeat k-folds*, $k=10$. El proceso se llevó a cabo mediante el software estadístico R utilizando la biblioteca “*caret*”. Los estadísticos de bondad del ajuste se calcularon como el promedio de los estadísticos obtenidos en el proceso de validación cruzada repetida, es decir, 100 ejecuciones del modelo (validación cruzada de 10 carpetas repetida 10 veces utilizando los datos de entrenamiento). La capacidad de ajuste del modelo se evaluó a través de varios estadísticos, incluyendo el pseudocoefficiente de determinación (R^2), el sesgo (Sesgo), la raíz del error medio cuadrático (REMC) y los valores relativos de estos (Sesgo% y REMC%). La medida de importancia de las variables (VIM) se utilizó



CIÓN

que
bles,
tario
tivos

o las
r los
n del

por
5% y
0,38
total
(G)
to G,
re el

a los
abla
go%

idos

a-1)

Nvar	13	18	13	12	17
R2	0,436	0,378	0,192	0,415	0,394
Sesgo	-0,124	-0,018	-8,992	-1,746	-0,945

Sesgo%	-0,005	-0,001	-0,016	-0,011		-0,010
REMC	8,519	4,115	291,763	74,899		43,545
REMC%	46,00	30,40	68,57	56,32		58,15
<i>P. radiata</i>	mtry	8	6	13	9	2
Nvar	11	11	21	9		2
R2	0,393	0,330	0,120	0,398		0,37
Sesgo	0,051	-0,005	-7,791	-0,209		0,092
Sesgo%	0,002	0,000	-0,017	-0,001		0,001
REMC	8,392	4,156	220,169	92,603		48,658
REMC%	38,13	22,92	61,11	46,74		47,24

Nota: mtry = número de variables muestreadas aleatoriamente como candidatas en cada división del árbol de regresión; N var= número de variables independientes en los modelos finales

Según los valores de VIMR obtenidos (no presentados), el grupo de variables con mayor contribución en los modelos fueron las bandas espectrales (39,11% del VIMR), seguido por los índices espectrales (25,69% del VIMR), las variables de textura (11,69% del VIMR), las variables de terreno (10,63% del VIMR), las variables climáticas (7,62% del VIMR) y por último la información radar de S1 (índices de radar, con el 3,55% del VIMR y las bandas de radar, con el 1,71% del VIMR.

Tabla 4. tablas resumen de VIMR : i) una por grupos de variables.

Bandas radar	Índices radar	Bandas espectrales	Índices espectrales	Variables de Textura	Variables de Terreno	Variables de Clima	
Densidad	1,15%	4,06%	36,04%	29,85%	12,43%	8,70%	7,77%
G	2,31%	2,20%	47,86%	21,10%	11,66%	8,59%	6,29%
Nha	0,00%	5,92%	24,22%	38,60%	13,20%	8,81%	9,24%
Tamaño	1,58%	4,36%	42,21%	14,44%	13,72%	16,80%	6,88%
H0	1,58%	4,36%	42,21%	14,44%	13,72%	16,80%	6,88%
Producción	2,32%	2,64%	40,63%	27,15%	9,92%	9,48%	7,85%
TVha	1,65%	3,69%	36,74%	31,50%	8,51%	11,05%	6,86%
WTha	3,00%	1,59%	44,53%	22,81%	11,33%	7,90%	8,85%
Total	1,71%	3,55%	39,11%	25,69%	11,69%	10,63%	7,62%

Tabla 5. tablas resumen de VIMR : otra por grupo de variable y especie.

Bandas radar	Índices radar	Bandas espectrales	Índices espectrales	VARIABLES de Textura	VARIABLES de Terreno	VARIABLES de Clima
			Pinus pinaster			
			Pinus radiata			
			Eucalyptus globulus			

Analizando los resultados por especie, el grupo de variables con mayor contribución en los modelos de *E. globulus* y *P. radiata* fue el de las bandas espectrales (promedio acumulado de VIMR de 46,19% y 43,45%, respectivamente), no así para *P. pinaster*, donde el grupo más influyente fue el de los índices espectrales, seguido por las bandas espectrales (promedio acumulado de VIMR de 34,04% y 27,70%, respectivamente). La segunda posición para *E. globulus* corresponde a los índices espectrales (promedio acumulado de VIMR de 27,50%) y para *P. radiata* a las variables de textura (promedio acumulado de VIMR de 16,97%). Sin embargo, en las tres especies, únicamente las variables obtenidas de las imágenes de Sentinel-2 presentan una importancia relativa acumulada de entre 58,96% y 73,70% (bandas + índices). En el caso de la contribución de los diferentes conjuntos de datos con cada variable dependiente, se pudo observar que la mayor contribución fue proporcionada por las bandas espectrales en todos los casos, excepto en el número de pies por hectárea.

Otras variables tienen una contribución menor, por debajo del 16% del VIMR acumulado; así, las variables de textura y terreno contribuyen con 11,21% y 9,17% en *P. pinaster*; 17,20% y 10,54% en *P. radiata*; y 6,88% y 12,18% en *E. globulus*. Las variables climáticas contribuyeron solo con 10,81% en *P. pinaster*, 8,99% en *P. radiata* y 3,08% en *E. globulus*. Finalmente, la contribución de las bandas de radar y los índices derivados fue de solo 5,26% (VIMR acumulado), con bandas de radar y índices de radar contribuyendo 0,95% y 6,12% en *P. pinaster*; 0,0% y 4,54% en *P. radiata*; y 4,17% y 0,0% en *E. globulus*, respectivamente.

La contribución de las variables de radar tuvo más importancia en los modelos desarrollados para *P. pinaster*. Para *P. radiata*, las variables contribuyen en los modelos en todos los casos, excepto en AGB y G. En *E. globulus*, las variables de radar sólo ayudaron a explicar la variabilidad en los modelos de producción (AGB y TV) y en los modelos de densidad (G y N), pero no en el modelo de tamaño (H0). La mayor contribución se observó en el modelo AGB para *E. globulus*, donde la banda de radar contribuyó con un 9% del VIMR.

5. Discusión

5.1. Precisión de los modelos

Las imágenes de reflectancia en superficie usadas en este estudio y los niveles de corrección fueron los mismos que utilizaron NOVO-FERNÁNDEZ et al. (2024) la misma región y especies. Al analizar la capacidad de ajuste de los modelos, se observa que los resultados obtenidos en R^2 no difieren de forma significativa de los obtenidos en otros estudios (ASTOLA et al., 2019, DOS REIS et al., 2018). Los



modelos basados en el número de pies por hectárea (N) fueron los menos precisos (R^2 promedio de 0,12 y 0,27). La altura dominante presenta un mejor ajuste con R^2 entre 0,33 y 0,38. Los modelos de densidad para área basimétrica (G) y de variables de producción (TV y AGB) produjeron los mejores resultados, con R^2 de 0,45, 0,45 y 0,44, respectivamente. Estos resultados son habituales, ya que los sensores remotos ópticos predicen la producción (volumen o biomasa) o las variables de densidad (ocupación del sitio) mejor que el número de pies por hectárea o las alturas del rodal (ASTOLA et al., 2019; RAHIMZADEH-BAJGIRAN et al., 2020).

5.2. Contribución de la información de radar

Las bandas ópticas obtuvieron el VIMR más alto en todos los casos, siendo la contribución de la información radar menos importante o nula en algunos modelos. El radar de apertura sintética (SAR) tiene una relevancia significativa en la estimación de diferentes parámetros forestales en grandes áreas, generalmente en el estudio de la biomasa forestal debido a su capacidad para penetrar el dosel forestal hasta cierta profundidad, su sensibilidad al contenido de agua en la vegetación y su independencia del clima (LE TOAN et al., 1992; KASISCHKE et al., 2011).

Numerosos estudios han proporcionado información sobre la predicción de variables forestales utilizando datos satelitales de radar como predictores (ROY et al., 2021; FANG et al., 2023); sin embargo, la mayoría de esos estudios se han llevado a cabo a escalas locales, sobre especies únicas (o sin discriminar especies) o han utilizado un número reducido de parcelas de verdad de campo, siendo difícil encontrar estudios que abarquen áreas grandes o diferentes especies (PERSSON et al., 2021).

Nuestros resultados indican que la inclusión de las bandas de radar de Sentinel-1 no contribuyen a mejorar los modelos forestales desarrollados. Estos resultados son similares a los obtenidos por FANG et al. (2023) pero discrepan de los obtenidos por otros autores como (ROY et al., 2021) que si obtuvieron mejoras significativas al estudiar un área reducida, con 40 parcelas rectangulares de 900 metros cuadrados en el bosque caducifolio tropical, donde *Shorea robusta* era la especie principal. Las discrepancias pueden estar relacionadas con las diferencias en las condiciones del estudio y el tamaño del área de estudio (local vs regional) y la base de datos utilizada (pequeña vs grande).

5.3. Señal de radar, tipo de sensor y topografía compleja

Tras analizar los resultados con detalle, es preciso mencionar ciertas particularidades encontradas en el uso de imágenes radar en el área de estudio que justifican los resultados obtenidos.

En primer lugar, la retrodispersión de las diferentes superficies de la cobertura terrestre está muy influenciada por variables que dependen del sensor, como la longitud de onda (por ejemplo, X, C, L, P) o la polarización (por ejemplo, HH, VV, HV, VH), y por otras que dependen de la superficie, como la cobertura terrestre y las propiedades del terreno (por ejemplo, rugosidad, constante dieléctrica o humedad), y por ambas, como el ángulo de incidencia. En este sentido, en las imágenes de la constelación Sentinel-1, la banda C con la que opera no es la más recomendable para estudios forestales debido a su incapacidad para captar características de la biomasa forestal (LE TOAN et al., 1992; LU, 2006). En segundo lugar, la interferencia constructiva y destructiva de las ondas radar reflejadas en una superficie genera un moteado en la imagen conocido como *speckle*. Dentro de un pixel, la suma de los vectores (amplitud + fase) correspondientes a los



elementos contenidos en él produce un valor. El valor del pixel vecino puede ser diferente por este efecto de la interferencia entre las ondas reflejadas.

Las variaciones anteriormente mencionadas dentro de una imagen y entre imágenes provocan que los valores de retrodispersión disponibles para una parcela varíen significativamente incluso cuando la parcela permanece constante (por ejemplo, el mismo volumen de madera en un mes), lo que complica el uso de este tipo de información para análisis a gran escala. Por tanto, el efecto del *speckle* juega un papel importante en la estimación de las variables forestales objetivo. El tipo de dato de campo recogido es crucial en este tipo de estudios. Así, mientras que parcelas rectangulares, de alrededor de una hectárea, pueden ayudar a combatir este efecto al promediar varios píxeles; las parcelas circulares habituales (las del IFN, por ejemplo) se encuadran dentro de un pixel, o en el mejor de los casos promedian 4 píxeles. Esto juega en contra de la utilidad de este tipo de datos, generando posiblemente más ruido que información que ayude a explicar la variabilidad de las variables objetivo.

En cuanto al tipo de imagen utilizado hay que destacar que Sentinel-1 proporciona dos tipos de datos: i) imágenes de *Single Look Complex* (SLC), que preservan la información de fase, y ii) *Ground Range Detection* (GRD) donde se pierde la información de fase, siendo este último conjunto de datos el único disponible en el repositorio de GEE. Debido a la incompatibilidad de la plataforma GEE para usar imágenes de SLC, no se pueden realizar análisis interferométricos, polarimétricos o tomográficos en GEE. Este tipo de análisis son los más comunes para estimar variables forestales, como la altura del arbolado, que posteriormente se correlaciona con otras variables forestales importantes como el volumen o la biomasa (SOLBERG et al., 2024, RAMACHANDRAN et al., 2023, XING et al., 2023, PAPATHANASSIOU et al., 2021) por lo que este tipo de análisis se debe realizar con otro tipo de herramientas diferentes a GEE.

En cuanto al efecto de la topografía, surgen dos problemas importantes cuando usamos imágenes de radar en áreas montañosas: i) el acortamiento y ii) el deslizamiento; ambos ocasionan la sombra de radar. La sombra de radar sucede cuando el haz de radar no puede iluminar la superficie del suelo; las regiones sombreadas aparecerán oscuras en una imagen, ya que no hay energía disponible que capturar. A medida que el ángulo de incidencia aumenta también aumentarán los efectos de sombra, ya que el haz de radar observa más oblicuamente la superficie. Es posible combinar distintas imágenes de pasadas ascendentes y descendentes, para aumentar la información sobre dos pendientes. Aun así, existen áreas sombreadas para las cuales no se obtiene información. La falta de información en estas áreas hace que las imágenes radar tengan limitaciones en cuanto al mapeado continuo de una variable forestal si lo comparamos con las imágenes ópticas.

5.4. Automatización, limitaciones y futuros desarrollos

Uno de los puntos fuertes de este estudio es la automatización del preprocesamiento de las imágenes, lo que proporciona modelos y herramientas útiles para obtener información replicable en los próximos años. Por el contrario, su principal limitación es la dificultad de disponer de imágenes sin nubes a escala regional o suprarregional, sobre todo si se desea realizar un análisis con escalas temporales inferiores a un año.

Debe tenerse en cuenta que la mayoría de los modelos de estimación de parámetros forestales son solo adecuados para las áreas de estudio específicas



donde se desarrollan; y no son transferibles debido al desconocimiento del preprocesamiento de la imagen. Los novedosos modelos de aprendizaje profundo (*Deep learning*) están permitiendo superar esta limitación, y, aunque requieren un cierto número de parcelas para recalibrar los modelos, este número es muy inferior al necesario para entrenar el modelo desde cero (GE et al., 2023). Este tipo de enfoques puede proporcionar, en los próximos años, una colección de modelos de variables forestales, transferibles, de forma que sea posible su aplicación en diferentes áreas geográficas.

Los futuros desarrollos pasaran por ajustar estos modelos utilizando una serie temporal o probar algoritmos de aprendizaje profundo para minimizar la dependencia de imágenes sin nubes. Además, el desarrollo de modelos para ciertas áreas y especies forestales utilizando otras técnicas como InSAR, PolSAR, PolInSAR o TomoSAR podría ser una solución para obtener una estimación de la producción forestal en estas áreas, aunque deberán realizarse fuera de la plataforma GEE debido a que, como se explicó anteriormente, se perdió la información de fase en el almacenamiento. La utilización de otras imágenes radar como ALOS-2 PALSAR-2 o TanDEM-X también se deberá estudiar para poder determinar el aporte de esta información en la capacidad predictiva de los modelos.

Otras posibles líneas futuras de mejora serán la inclusión de los datos de sensores radar que incorporen los nuevos satélites cuando estos estén disponibles (NISAR) o la transferencia de modelos, o la creación de modelos *Deep learning* cuando exista una correlación temporal entre los datos de campo y los vuelos LiDAR.

6. Conclusiones

La utilización de plataformas como GEE permite obtener los mismos resultados que con un procesamiento de las imágenes en local, pero permitiendo actualizar las imágenes para el área de estudio con un script simple, que no requiere descargar imágenes y además realiza el preprocesamiento de las imágenes en el servidor. Esto permite llevar a cabo análisis sobre área de estudio grandes de manera ágil y rápida. Hemos observado que las variables de radar no proporcionan información adicional a las variables ópticas para predecir variables estado de masa en plantaciones forestales en el norte de España. Por lo tanto, la búsqueda de una contribución significativa de la información de radar pasara por probar diferentes enfoques del preprocesamiento y/o tipo de datos (bandas, satélites, etc.) fuera de la plataforma GEE.

7. Agradecimientos

Este trabajo es un resultado del proyecto de investigación PID2020-112839RB-I00 financiado por la Agencia Estatal de Investigación (AEI) del Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033). Mientras realizaba el presente estudio, el primer autor disfrutó de un contrato Severo Ochoa del Gobierno del Principado de Asturias (código BP21-125).

8. Bibliografía

ARANGUREN ERICE, I.; 2021. Relleno de huecos en series temporales de índices Sentinel-2 en cubiertas agrícolas con datos procedentes de Sentinel-1. Trabajo Fin de Máster, ETSIA, Universidad Pública de Navarra. Disponible en: <https://academica-e.unavarra.es/handle/2454/40687>. Acceso el 8 julio 2024

ASTOLA, H.; HÄME, T.; SIRRO, L.; MOLINIER, M.; KILPI, J.; 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.* 223: 257–273.



- BREIMAN, L.; 2001. Random forests. *Mach. Learn.* 45(1): 5–32.
- DOS REIS, A.A.; CARVALHO, M.C.; DE MELLO, J.M.; GOMIDE, L.R.; FILHO, A.C.F.; JUNIOR, F.W.A.; 2018. Spatial prediction of basal area and volume in *Eucalyptus* stands using Landsat TM data: an assessment of prediction methods. *N. Z. J. For. Sci.* 48(1): 2018.
- FANG, G.; XU, H.; YANG, S.-I.; LOU, X.; FANG, L.; 2023. Synergistic use of Sentinel-1, Sentinel-2, and Landsat 8 in predicting forest variables. *Ecol. Indic.* 151: 110296.
- GE, S.; ANTROPOV, O.; HÄME, T.; McROBERTS, R.E.; MIETTINEN, J.; 2023. Deep learning model transfer in forest mapping using multi-source satellite SAR and optical images. *Remote Sens.* 15(21): 5152.
- HARALICK, R.M.; DINSTEIN, I.; SHANMUGAM, K.; 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, v. SMC-3, n. 6, p. 610–621.
- KASISCHKE, E.S.; TANASE, M.A.; BOURGEOU-CHAVEZ, L.L.; BORR, M.; 2011. Soil moisture limitations on monitoring boreal forest regrowth using spaceborne L-band SAR data. *Remote Sens. Environ.* 115 (1): 227–232.
- LE TOAN, T.; BEAUDOIN, A.; RIOM, J.; GUYON, D.; 1992. Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 30, n. 2, p. 403–411.
- LU, D.; 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7): 1297–1328.
- MAPAMA; 2019. Actualización del Cuarto Inventario Forestal español para el norte de España. Ministerio de Agricultura, Pesca y Alimentación. Gobierno de España.
- MULLISSA, A.; VOLLRATH, A.; ODONGO-BRAUN, C.; SLAGTER, B.; BALLING, J.; GOU, Y.; GORELICK, N.; REICHE, J.; 2021. Sentinel-1 SAR backscatter analysis ready data preparation in Google Earth Engine. *Remote Sens.* 13(10): 1954.
- NINYEROLA, M.; PONS, X.; ROURE, J.M.; 2005. Atlas climático digital de la Península Ibérica. Universitat Autònoma de Barcelona, 45 pp.
- NOVO-FERNÁNDEZ, A.; LÓPEZ-SÁNCHEZ, C.A.; CÁMARA-OBREGÓN, A.; BARRIO-ANTA, M.; TEIJIDO-MURIAS, I.; 2024. Estimating forest variables for major commercial timber plantations in Northern Spain using Sentinel-2 and ancillary data. *Forests.* 15(1): 99.
- PAPATHANASSIOU, K.P.; CLOUDE, S.R.; PARDINI, M.; QUIÑONES, M.J.; HOEKMAN, D.; FERRO-FAMIL, L.; GOODENOUGH, D.; CHEN, H.; TEBALDINI, S.; NEUMANN, M.; ULANDER, L.M.H.; SOJA, M.J.; 2021. Forest applications. En: *Polarimetric Synthetic Aperture Radar. Book series in Remote Sensing and Digital Image Processing.* 25: 59–117. Springer Nature.
- PERIASAMY, S.; 2018. Significance of dual Polarimetric Synthetic Aperture Radar in biomass retrieval: an attempt on Sentinel-1. *Remote Sens. Environ.* 217: 537–549.
- PERSSON, H.J.; JONZEN, J.; NILSSON, M.; 2021. Combining TanDEM-X and Sentinel-2 for large-area species-wise prediction of forest biomass and volume. *Int. J. Appl. Earth Obs. Geoinf.* 96: 102275.
- RAHIMZADEH-BAJGIRAN, P.; HENNIGAR, C.; WEISKITTEL, A.; LAMB, S.; 2020. Forest potential productivity mapping by linking remote-sensing-derived metrics to site variables. *Remote Sens.* 12: 2056.



RAMACHANDRAN, N.; SAATCHI, S.; TEBALDINI, S.; D'ALESSANDRO M.M.; DIKSHIT, O.; 2023. Mapping tropical forest aboveground biomass using airborne SAR tomography. *Sci. Rep.* 13(1): 6233.

ROY, S.; MUDI, S.; DAS, P.; GHOST, S.; SHIT, P.K.; 2021. Estimating above ground biomass (AGB) and tree density using Sentinel-1 data. *Environ. Eng. Sci.*, p. 259–280.

SOLBERG, S.; BOLLANDSÅS, O.M.; GOBAKKEN, T.; NÆSSET, E.; BASAK, P.; DUNCANSON, L.I.; 2024. Biomass change estimated by TanDEM-X interferometry and GEDI in a Tanzanian forest. *Remote Sens.* 16(5): 861.

TRUDEL, M.; CHARBONNEAU, F.; LECONTE, R. 2012. Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. *Can. J. Remote Sens.* 38(4): 514-527.

XING, C.; WANG, H.; ZHANG, Z.; YIN, J.; YANG, J.; 2023. A Review of forest height inversion by PolInSAR: theory, advances, and perspectives. *Remote Sens.* 15(15): 3781.

ZHIWEI, X.; XINGHUA, W.; 2010. Research for Information Extraction Based on Wrapper Model Algorithm. 2010 Second International Conference on Computer Research and Development. IEEE Xplore.