

Martín-Molina, D.; Pugni-Stanek, T.; Merino-de-Miguel, S. Aplicación de la teledetección y los modelos de aprendizaje automático para la detección y seguimiento de daños causados por la banda marrón (*Lecanosticta acicola*) en masas de *Pinus radiata* en Euskadi

Aplicación de la teledetección y los modelos de aprendizaje automático para la detección y seguimiento de daños causados por la banda marrón (*Lecanosticta acicola*) en masas de *Pinus radiata* en Euskadi

Martín-Molina, David ¹ Pugni-Stanek, Tomás ¹ Merino-de-Miguel, Silvia ¹

¹ ETSI Montes, Forestal y del Medio Natural (Universidad Politécnica de Madrid) C/ José Antonio Novais, 10 Madrid 28040, España

ORCID: Martín-Molina 0009-0002-5746-348X Pugni-Stanek 0009-0000-2527-1657 Merino-de-Miguel 0000-0002-4764-5311

Correspondencia: dav.m.molina@gmail.com tomas.pugni@upm.es silvia.merino@upm.es

RESUMEN

Ante los daños persistentes sobre *Pinus radiata* ocasionados por la enfermedad de la banda marrón en Euskadi, la teledetección se presenta como una herramienta fundamental. Este trabajo explora la capacidad de los modelos de aprendizaje automático basados en datos Sentinel-2 para su monitorización. Se desarrolló un modelo Random Forest de clasificación para detectar los daños ocasionados según cuatro niveles de defoliación observada. La muestra de píxeles verdad-terreno se clasificó sobre ortofotografías de 25cm, calculándose como predictores un conjunto de índices de vegetación, así como variables topográficas. El modelo resultante quedó definido por tres índices, suficientes para un rendimiento del modelo estable con una precisión del 70%. Destacó su capacidad de identificar correctamente las clases extremas, sano y grave, en contraposición con la confusión en las clases intermedias, leve y moderado. Estos resultados son coherentes con la defoliación progresiva de los árboles afectados, lo que se traduce en un solape espectral entre los niveles de defoliación. Sin embargo, se ven penalizados por la decisión de definir una primera clase (sano) restrictiva, donde sólo se incluyeron píxeles sin afección. Estos resultados apuntan al potencial de los materiales y métodos empleados para aplicaciones futuras al seguimiento del estado de salud del pino.

Palabras clave: *Pinus radiata*, Sentinel-2, Random Forest, salud forestal


Fecha de recepción: 18 febrero 2026 · Fecha de aceptación: 18 febrero 2026

Aplicación de la teledetección y los modelos de aprendizaje automático para la detección y seguimiento de daños causados por la banda marrón (*Lecanosticta acicola*) en masas de *Pinus radiata* en Euskadi

Martín-Molina, David ⁽¹⁾, Pugni-Stanek, Tomás ⁽¹⁾, Merino-de-Miguel, Silvia ⁽¹⁾

⁽¹⁾ ETSI Montes, Forestal y del Medio Natural (Universidad Politécnica de Madrid), Madrid, España.

 0009-0002-5746-348X, dav.m.molina@gmail.com;  0009-0000-2527-1657, tomas.pugni@upm.es.

 0000-0002-4764-5311, silvia.merino@upm.es

Resumen: Ante los daños persistentes sobre *Pinus radiata* ocasionados por la enfermedad de la banda marrón en Euskadi, la teledetección se presenta como una herramienta fundamental. Este trabajo explora la capacidad de los modelos de aprendizaje automático basados en datos Sentinel-2 para su monitorización. Se desarrolló un modelo *Random Forest* de clasificación para detectar los daños ocasionados según cuatro niveles de defoliación observada. La muestra de píxeles verdad-terreno se clasificó sobre ortofotografías de 25cm, calculándose como predictores un conjunto de índices de vegetación, así como variables topográficas. El modelo resultante quedó definido por tres índices, suficientes para un rendimiento del modelo estable con una precisión del 70%. Destacó su capacidad de identificar correctamente las clases extremas, sano y grave, en contraposición con la confusión en las clases intermedias, leve y moderado. Estos resultados son coherentes con la defoliación progresiva de los árboles afectados, lo que se traduce en un solape espectral entre los niveles de defoliación. Sin embargo, se ven penalizados por la decisión de definir una primera clase (sano) restrictiva, donde sólo se incluyeron píxeles sin afección. Estos resultados apuntan al potencial de los materiales y métodos empleados para aplicaciones futuras al seguimiento del estado de salud del pino.

Palabras clave: *Pinus radiata*, Sentinel-2, Random Forest, salud forestal

Application of remote sensing and machine learning models for the detection and monitoring of damage caused by brown spot needle blight (*Lecanosticta acicola*) in *Pinus radiata* stands in the Basque Country

Abstract: Given the persistent damage on *Pinus radiata* caused by the disease of the brown spot needle blight in the Basque Country, remote sensing emerges as an essential monitoring tool. This study examines the potential of machine learning models based on Sentinel-2 data for monitoring it. A Random Forest classification model was developed to detect damage according to four levels of observed defoliation. The sample of ground-truth pixels was classified based on 25cm orthophotographs, calculating as predictors a set of vegetation indices and topographic variables. The resulting model was defined by three indices, which are enough for stable model performance of 70% accuracy. It stood out for its ability to successfully identify the extreme categories, healthy and severe, as opposed to the confusion in the intermediate categories, mild and moderate. These results are consistent with the progressive defoliation of the affected trees, which leads to spectral overlapping between the defoliation levels. However, they are negatively affected by the decision to establish a restrictive category (healthy), which only included unaffected pixels. The findings suggest the potential of the materials and methods used for future forest pine health applications.

Keywords: *Pinus radiata*, Sentinel-2, Random Forest, forest health.

1. INTRODUCCIÓN

La enfermedad de la banda marrón causada por el hongo *Lecanosticta acicola* es una de las micosis que forman parte de las denominadas bandas del pino. La infección se produce tras la germinación de las esporas asexuales, las cuales acceden a las acículas a través de los estomas. Las condiciones necesarias para que se produzca son una humedad relativa elevada (>75%),

acompañada de altas temperaturas (Mesanza *et al.*, 2021). El momento de máxima infección ocurre entre mayo y agosto. El hongo actúa en primer lugar generando manchas amarillentas de 1 a 3,5 mm en las acículas, las cuales terminan por ensancharse hasta que rodean completamente la acícula ocasionando la muerte de la parte distal. Las infecciones más intensas junto con la sucesión de episodios que se producen de manera continuada durante varios años pueden derivar en la

pérdida de ramillos, de parte de la copa o incluso causar la muerte del árbol (Muñoz López *et al.*, 2011).

La región más afectada por esta enfermedad hasta el momento ha sido Euskadi. Se estima que unas 40 000 hectáreas se vieron afectadas por el agresivo brote ocurrido en el año 2018 (Bueso, 2018), y su presencia se ha hecho permanente en las masas de pino radiata de la región, afectando con diferente intensidad, pero de manera recurrente desde entonces. La expansión de la banda marrón sobre *Pinus radiata* en España alcanza toda la cornisa cantábrica. También ha sido detectada recientemente por primera vez en Cataluña durante labores de seguimiento de la banda roja en *Pinus nigra*.

Todas las especies del género *Pinus* de la península ibérica son susceptibles de verse afectadas por la banda marrón. Análisis realizados en países de nuestro entorno han confirmado su presencia sobre *Pinus halepensis*, *P. nigra*, *P. sylvestris* y *P. uncinata*, así como en varias de las especies no nativas presentes en España (EPPO, n.d.). Por tanto, pese a que no hay reportes de presencia confirmada sobre *Pinus pinea* y *Pinus pinaster*, el potencial de propagación de la enfermedad es elevado y los efectos sobre los pinos locales aún desconocidos.

La teledetección, a través de la utilización de datos abiertos del programa Copernicus, representa una oportunidad para la monitorización del estado sanitario de los bosques, permitiendo el desarrollo de sistemas de seguimiento y predicción de la evolución de los bosques. El presente trabajo tiene como objetivo aplicar la teledetección al desarrollo de un modelo de *Random Forest* para la detección de la defoliación causada por la banda marrón y la clasificación de la superficie de pino radiata según el nivel de daño, evaluando y cuantificando la incidencia en Euskadi entre los años 2018 y 2024.

2. MATERIAL Y MÉTODOS

2.1. Área de estudio

El estudio se desarrolló tomando como localización Euskadi (España), empleando para el entrenamiento la provincia de Gipuzkoa y aplicando el modelo a la comunidad autónoma en su conjunto (Fig. 1).



Figura 1. Área de estudio. Fuente: elaboración propia a partir del MFCAE.

La superficie considerada se corresponde con la superficie forestal ocupada por *Pinus radiata*. Esta

especie, según datos oficiales para el año 2018, se extendía sobre 117 582 hectáreas, representando aproximadamente el 30% de la superficie forestal arbolada de Euskadi (HAZI Fundazioa, 2018).

2.2. Materiales

La cartografía utilizada para determinar el área objetivo fueron las sucesivas actualizaciones anuales del Mapa Forestal de España (MFE25), publicadas como Mapa Forestal del País Vasco (MFCAE) por el Eusko Jaurlaritz / Gobierno Vasco (n.d.). A partir de esta fuente, se obtuvo la superficie ocupada por las teselas cuya especie principal fuera el *Pinus radiata* (SP1=28) para cada año dentro del periodo de estudio (2018-2024).

La información satelital se derivó del programa Copernicus de la Unión Europea, concretamente de las misiones Sentinel-2. Las imágenes empleadas fueron las de nivel de procesado L2A, el mayor disponible, seleccionadas a su vez por la ausencia de nubes. La franja temporal de las imágenes, septiembre-octubre, se estableció por su idoneidad para la evaluación de los efectos de la enfermedad por coincidir con el momento en el que se observa la mayor severidad de la banda marrón. Se utilizó una imagen de baja nubosidad Sentinel-2 para cada año estudiado entre 2018-2024.

Los atributos topográficos evaluados como posibles variables del modelo se extrajeron de los productos MDT (Modelo Digital del Terreno), pendientes y orientaciones de 25m de resolución espacial, de máxima actualidad, que proporciona el Gobierno Vasco a través del portal GeoEuskadi y a los que se accedió como datos WCS. Estos datos provienen de la Base Topográfica Armonizada del Gobierno Vasco realizada en 2009, cuyo origen de datos es el vuelo fotogramétrico del año 2008, y que se actualiza allí donde se detectan cambios significativos a partir del vuelo que se realiza anualmente.

Se recurrió a la misma fuente de información para la generación de la muestra de verdad-terreno, clasificada según distintos niveles de afección. Para ello, se emplearon las ortofotografías anuales color verdadero y falso color infrarrojo de 25 cm de píxel (Fig. 2). La captura de estas imágenes coincide también temporalmente con el momento deseado, ya que su adquisición es a lo largo del mes de septiembre.

2.3. Variables de entrada

La muestra verdad-terreno se obtuvo a partir de los píxeles centrales de 20 m contenidos en ventanas de 3 × 3 píxeles (60 × 60 m) completamente incluidas dentro de las teselas de *Pinus radiata* del MFCAE actualizado a 2018, y localizadas exclusivamente en la provincia de Gipuzkoa.

El tamaño de muestra inicial se estableció siguiendo las recomendaciones para modelos de *machine learning* de disponer de un número mínimo de muestras por cada clase y variable considerada en el modelo. Utilizando un muestreo piloto para determinar la existencia o no de fuerte desbalanceo entre clases. Ello quedó descartado y se estableció, siendo conservadores, un valor “k” de 75 en la Ecuación 1.

$$n = k \times n^{\circ} \text{ clases} \times n^{\circ} \text{ predictores} \quad (1)$$

Esta muestra de píxeles se clasificó en cuatro niveles de afección según la defoliación observada dentro del píxel de 20x20m en la ortofotografía: sano (0%), leve (0-20%), moderado (20-60%) y grave (>60%) (Fig. 2). Se filtraron las nubes según la banda SCL (*Scene Classification*) de Sentinel-2, y se emplearon sólo aquellos píxeles con presencia de vegetación arbórea y sin elementos artificiales, caminos o construcciones.

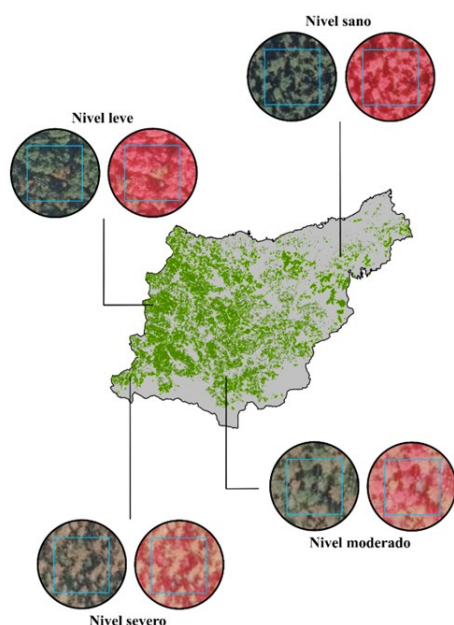


Figura 2. Ejemplos del criterio visual de clasificación: nivel sano, leve, moderado y severo sobre ortofotografía color natural y falso color infrarrojo.

Las variables consideradas fueron todas las bandas de Sentinel-2 MSI disponibles a 20 m de resolución espacial de la tesela 30TWN para la fecha 26-09-2018, las características topográficas y geográficas de cada píxel (altitud, pendiente, orientación y latitud) y los índices de vegetación: NDVI, NDRE, CI green, CCCI, EVI, NDMI, MSI, GVMI, NBR2, SIPI, PRI, TCARI, MCARI CI re, PSRI, NBR y MSAVI2, según las expresiones recogidas en la Index DataBase (Henrich *et al.*, s.f.). El entrenamiento se llevó a cabo únicamente en 2018.

2.4. Modelo *Random Forest*

El modelo de clasificación elegido fue *Random Forest* (RF), propuesto por Breiman (2001). Se desarrolló a través del código programado en Python utilizando la librería *Scikit learn v1.8.0*, así como otras librerías complementarias para la preparación de los datos. Pese a que RF no es sensible a la multicolinealidad de las variables, éstas fueron sometidas a un proceso de preselección basado en el cálculo del coeficiente de correlación de Spearman, estableciéndose un umbral de 0.95. Sin embargo, la decisión final de incorporar al modelo una variable predictora fue coherente con evitar la redundancia espectral sin reducir la información

disponible, evitando comprometer la interpretabilidad del modelo.

Se entrenó un primer modelo con las restantes variables. Este proceso de entrenamiento incluyó la optimización de los hiperparámetros críticos del modelo utilizando el método *RandomizedSearchCV*. Este procedimiento permitió determinar la combinación que maximizaba el rendimiento del modelo en términos de *F1-macro*, es decir, aquella que mejor detectara correctamente todas las clases asignando la misma relevancia a todas ellas.

El algoritmo ofrece una medida de la influencia de cada variable en las predicciones del modelo, denominada importancia, calculada como la reducción media del índice de Gini. Con esta medida, es posible establecer a partir de qué número de variables el rendimiento del modelo se estabiliza, y la incorporación de nuevas variables no reduce en una mejora significativa o incluso aporta ruido. Así se determinó el número definitivo de variables y se repitió el proceso de optimización de hiperparámetros ya que algunos son sensibles al número de predictores.

3. RESULTADOS Y DISCUSIÓN

La muestra verdad-terreno definitiva contó con 3500 píxeles válidos para el desarrollo del modelo, tras desechar los que no cumplían con las condiciones mínimas de calidad. Dicha muestra se distribuyó por niveles tal y como muestra la Tabla 1.

Tabla 1. Resultado del muestreo

Nivel de afección	Nº píxeles
Sano	781 (22,3%)
Leve	723 (20,7%)
Moderado	1002 (28,6%)
Grave	994 (28,4%)

El modelo RF quedó definido por los hiperparámetros optimizados y un total de tres variables a partir del cual su rendimiento se estabilizaba, tal y como se muestra en la Figura 3. Dichas variables fueron los índices de vegetación SIPI, NDGI y GVMI, en orden de importancia.

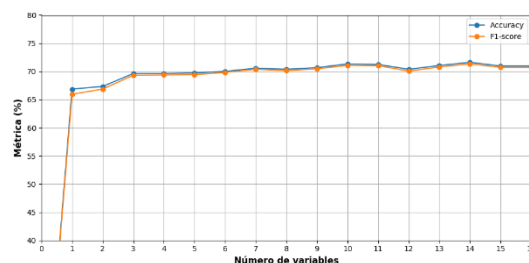


Figura 3. Influencia del número de variable en el rendimiento

El modelo RF, tal y como refleja la Figura 4, tiene un buen desempeño en las clases extremas, sano y grave, las cuáles además son perfectamente separables entre sí. Sin embargo, se diluye en las intermedias especialmente en la clase leve. Ello es coherente con la naturaleza gradual del proceso de infección, ya que las clases

establecidas representan un fenómeno continuo, y los resultados son el reflejo del solape espectral entre las clases. Al mismo tiempo refleja las limitaciones del modelo para la detección temprana de la infección. Por ese motivo algunos estudios equivalentes unifican en una clase los niveles sano y leve (Sun et al., 2025).

	Sano	Leve	Moderado	Grave		Media ponderada
Sano	175	45	13	1		
Leve	59	108	47	3		
Moderado	7	40	202	52		
Grave	0	4	44	250		
Precision	0,726	0,548	0,660	0,817	0,688	0,696
Recall	0,748	0,498	0,671	0,839	0,689	0,700
F1-score	0,737	0,522	0,666	0,828	0,688	0,698
Accuracy	0,700					

Figura 4. Matriz de confusión

Se realizó una comparación a nivel de píxel de los resultados obtenidos para el año 2021, con los facilitados por Hazi Fundazioa. Su metodología, basada en valores de NBR, contempla una división también por porcentaje de daño, pero con diferentes umbrales entre clases. Existiendo una equivalencia directa con la clasificación propuesta en este trabajo. Siendo: Sano + Leve = Tipo 5 (0-20%); Moderado = Tipo 4-3 (20-60%); y Grave = Tipo 2-1 (>60%). Este análisis reflejó una coincidencia espacial entre ambos métodos del 56%, produciéndose la mayor parte de las discrepancias en las clases que el modelo tiende a confundir, las intermedias.

El modelo finalmente sirvió para clasificar la superficie afectada por la banda marrón en masas de pino radiata en Euskadi durante el periodo 2018-2025. La Tabla 2 muestra la distribución por provincias.

Tabla 2. Evolución de la superficie afectada (ha)*

Año	Gipuzkoa	Bizkaia	Araba/Álava
2018	18819	21222	2711
2019	17708	17392	3548
2020	25736	33129	6869
2021	13832	18618	3208
2022	22640	34612	6858
2023	21220	31396	6202
2024	7100	7901	1607

*Superficie afectada incluye moderado y grave

Para la visualización del nivel de afección de la enfermedad fruto de la aplicación del modelo para todo Euskadi, se ha desarrollado el visor accesible a través del siguiente enlace: <https://mmolina-david.github.io/Mapa-BandaMarron-Euskadi/>

4. CONCLUSIONES

Mediante la aplicación de técnicas de aprendizaje automático utilizando datos Sentinel-2 MSI se consiguió desarrollar un modelo de clasificación para la banda marrón, el cual podría utilizarse para la identificación de superficies afectadas. No obstante, la definición

propuesta en este estudio de los distintos niveles de afección debe tenerse en cuenta para una correcta interpretación de su rendimiento.

Una futura continuación de este trabajo podría abordar un calibrado de la manera en que se asignan los píxeles a uno u otro nivel de afección, durante la elaboración de la muestra en base a datos de campo. También se podría dirigir el estudio hacia otros modelos, otras fuentes de información satelital, ampliar el área de estudio a otras regiones afectadas o hacer uso de la información temporal lo que podría ayudar en la detección en fases tempranas y en la detección de estados intermedios.

5. AGRADECIMIENTOS

Agradecemos la disponibilidad de datos abiertos del Inventario Forestal Nacional y del programa Copernicus de la Unión Europea. Así como al Eusko Jaurilaritza / Gobierno Vasco por hacer accesible la actualización anual del MFE25 para Euskadi, y por la disponibilidad de ortofotografías e información topográfica. Y en especial a Alejandro Cantero Amiano por compartir los resultados obtenidos para el año 2021, haciendo posible la comparación entre metodologías.

6. REFERENCIAS

- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- Bueso, G. (2018). Un tercio de los pinos del País Vasco, afectado por enfermedades fúngicas. *Phytoma*, 303.
- EPPO (n.d.). EPPO Global Database. Retrieved May 15, 2025, <https://gd.eppo.int/>
- Eusko Jaurilaritza / Gobierno Vasco. (n.d.). Mapa Forestal del País Vasco. <https://www.euskadi.eus/mapa-forestal-del-pais-vasco/web01-a3ingair/es/>
- HAZI Fundazioa (2018). El bosque vasco en cifras 2018. <http://www.nasdap.net/inventarioforestal>
- Henrich, V., Krauss, G., Götze, C., & Sandow, C. (s. f.). Index-Database: A database for remote sensing indices. Recuperado el 20 de octubre de 2025, de <https://www.indexdatabase.de/>
- Mesanza, N., García-García, D., Raposo, E. R., Raposo, R., Iturbide, M., Pascual, M. T., Barrena, I., Urkola, A., Berano, N., de Zerain, A. S., & Iturritxa, E. (2021). Weather variables associated with spore dispersal of *lecanosticta acicola* causing pine needle blight in Northern Spain. *Plants*, 10(12). <https://doi.org/10.3390/plants10122788>
- Muñoz López, C., Pérez Fortea, V., Cobos Suárez, P., Hernández Alonso, R., & Sánchez Peña, G. (2011). *Sanidad forestal* (3.a ed. rev.). Ediciones Mundi-Prensa.
- Sun, G., Huang, X., Dashzebeg, G., Ariunaa, M., Bao, Y., Bao, G., Tong, S., Dorjsuren, A., & Davaadorj, E. (2025). Discrimination of larch needle pest severity based on Sentinel-2 super-resolution and spectral derivatives—A case study of *Erannis jacobsoni* Djak. *Forests*, 16(1), 88. <https://doi.org/10.3390/f16010088>