

Canero, F.; Rodríguez-Galiano, V.; Mendes, M.; Ollega Caro, J. Teledetección para la cartografía de cromo y níquel en Sierra de la Nieves: una aproximación desde el aprendizaje automático interpretable

Teledetección para la cartografía de cromo y níquel en Sierra de la Nieves: una aproximación desde el aprendizaje automático interpretable

Canero, Francisco M. ¹ Rodríguez-Galiano, Víctor ¹ Mendes, Maria Paula ² Ollega Caro, Jose M. ¹

¹ Departamento de Geografía Física y AGR, Universidad de Sevilla, Sevilla, 41004, España

² CERENA - Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, 1049-001, Lisboa, Portugal

ORCID: Canero [0000-0002-7813-1487](https://orcid.org/0000-0002-7813-1487) Rodríguez-Galiano [0000-0002-5422-8305](https://orcid.org/0000-0002-5422-8305) Mendes [0000-0001-9451-2266](https://orcid.org/0000-0001-9451-2266)

Correspondencia: fcanero@us.es vrgaliano@us.es mpaulamendes@tecnico.ulisboa.pt jollega@us.es

RESUMEN

La cartografía de metales pesados (MP) en suelos ha ganado atención por su relevancia para la salud del suelo. Sin embargo, la información sobre los niveles de base en regiones no contaminadas sigue siendo limitada. Este estudio cartografía el contenido de cromo (Cr) y níquel (Ni) del suelo de la Sierra de las Nieves, España. Los conjuntos de variables predictivas incluyeron compuestos trimestrales, índices espectrales representativos de la cobertura del suelo, características fenológicas extraídas de la serie temporal de Sentinel 2 para 2018-2020, y características geológicas y del terreno. Se aplicó un enfoque basado en pilas de regresores (stacks), empleando el método de Selección Secuencial de Características para la reducción de dimensionalidad. La predicción de MP se evaluó utilizando los siguientes métodos de Aprendizaje Automático Interpretable (IML): Importancia de Características agnóstica del modelo y gráficos de Efecto Local Acumulado (ALE). El níquel mostró la mayor precisión ($R^2=0.6$, $RMSE=0.78$), seguido por el cromo ($R^2=0.41$, $RMSE=0.68$). La distancia a las peridotitas fue determinante, con varias características de teledetección, como SWIR, siendo altamente relevantes. Este marco metodológico ofrece mapas de referencia a escala de paisaje e información valiosa sobre los factores impulsores de la distribución espacial de MP del suelo.



Palabras clave: *cromo, níquel, Sentinel 2, feature importance, accumulated local effects, Sierra de las Nieves*

Fecha de recepción: 18 febrero 2026 · Fecha de aceptación: 19 marzo 2026


Teledetección para la cartografía de cromo y níquel en Sierra de las Nieves: una aproximación desde el aprendizaje automático interpretable

Canero, Francisco M. ⁽¹⁾, Rodríguez-Galiano, Víctor ⁽¹⁾, Mendes, Maria Paula ⁽²⁾, Ollega Caro, Jose M. ⁽¹⁾

⁽¹⁾ Departamento de Geografía Física y AGR, Universidad de Sevilla, España.

 0000-0002-7813-1487, fcanero@us.es ;  0000-0002-5422-8305, vrgaliano@us.es ; jollega@us.es.

⁽²⁾ CERENA - Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, 1049-001, Portugal.

 0000-0001-9451-2266, mpaulamendes@tecnico.ulisboa.pt.

Resumen: La cartografía de metales pesados (MP) en suelos ha ganado atención por su relevancia para la salud del suelo. Sin embargo, la información sobre los niveles de base en regiones no contaminadas sigue siendo limitada. Este estudio cartografía el contenido de cromo (Cr) y níquel (Ni) del suelo de la Sierra de las Nieves, España. Los conjuntos de variables predictivas incluyeron compuestos trimestrales, índices espectrales representativos de la cobertura del suelo, características fenológicas extraídas de la serie temporal de Sentinel 2 para 2018-2020, y características geológicas y del terreno. Se aplicó un enfoque basado en pilas de regresores (*stacks*), empleando el método de Selección Secuencial de Características para la reducción de dimensionalidad. La predicción de MP se evaluó utilizando los siguientes métodos de Aprendizaje Automático Interpretable (IML): Importancia de Características agnóstica del modelo y gráficos de Efecto Local Acumulado (ALE). El níquel mostró la mayor precisión ($R^2=0.6$, $RMSE=0.78$), seguido por el cromo ($R^2=0.41$, $RMSE=0.68$). La distancia a las peridotitas fue determinante, con varias características de teledetección, como SWIR, siendo altamente relevantes. Este marco metodológico ofrece mapas de referencia a escala de paisaje e información valiosa sobre los factores impulsores de la distribución espacial de MP del suelo.

Palabras clave: : cromo, níquel, Sentinel 2, feature importance, accumulated local effects, Sierra de las Nieves

Remote sensing for chromium and nickel mapping in Sierra de las Nieves: an Interpretable Machine Learning approach

Abstract: Mapping heavy metals (HM) in soil has gained attention for its relevance to soil health, yet background levels in unpolluted regions remain limited. This study maps soil chromium (Cr) and nickel (Ni) content in Sierra de las Nieves, Spain. Predictive feature sets included three-month composites, spectral indices representing land cover, and phenological features extracted from Sentinel 2 time series for 2018-2020, and geological and terrain features. An ensemble stacking approach was applied, with Sequential Feature Selection method for dimensionality reduction. The HM prediction was evaluated using the following Interpretable Machine Learning method: model-agnostic Feature Importance and Accumulated Local Effect (ALE) plots. Nickel showed the highest accuracy ($R^2 = 0.6$, $RMSE = 0.78$), followed by chromium ($R^2 = 0.41$, $RMSE = 0.68$). Distance to peridotites was the most important feature for both Cr and Ni, but several remote sensing features including SWIR composites played a highly relevant role. This framework offers baseline maps at the landscape scale and valuable insights into the distribution drivers.

Keywords: chromium, nickel, Sentinel 2, feature importance, accumulated local effects, Sierra de las Nieves

1. INTRODUCCIÓN

La presencia de metales pesados (MP) en suelos plantea riesgos ambientales significativos debido a su persistencia y origen tanto natural como antropogénico (Wang *et al.*, 2018). A pesar de los esfuerzos por parte

de la normativa europea a través de la novedosa Ley de Monitorización del Suelo y los mapas supranacionales existentes, las evaluaciones a gran escala omiten la heterogeneidad local. Dada la alta variabilidad espacial, es indispensable una cartografía de alta resolución para

diseñar estrategias de remediación efectivas (Moradpour *et al.*, 2023).

La teledetección ha emergido como una fuente prometedora para la predicción en estos entornos. La predicción espacial de MP basada en teledetección en suelos forestales se basa en métodos indirectos a partir del marco SCORPAN (McBratney *et al.*, 2003). En este enfoque, las imágenes de la misión Sentinel-2 de la ESA suponen un nuevo impulso para la cartografía de MP proporcionando información de cobertura, vegetación y suelo en diferentes canales espectrales o índices derivados (Azizi *et al.*, 2026). Este enfoque es valioso en áreas forestales donde el dosel obstruye la reflectancia del suelo, pero donde la contaminación pudiera alterar rasgos bioquímicos y respuestas espectrales de la vegetación (Kooistra *et al.*, 2004). Además, la resolución espacio-temporal de Sentinel-2 permite describir trayectorias fenológicas (Caparros-Santiago *et al.*, 2021) y extraer impulsores biofísicos como la dinámica vegetal o la actividad fotosintética acumulada.

La contribución innovadora de este trabajo reside en la combinación de técnicas de aprendizaje por apilamiento (*stacking learning*) y selección de características junto con aprendizaje automático interpretable (IML), junto a datos derivados de Sentinel-2, para el modelado espacial de cromo (Cr) y níquel (Ni) en la Sierra de las Nieves, sur de España.

2. MATERIAL Y MÉTODOS

2.1. Área de estudio y datos

Sierra de las Nieves (SN) es un área montañosa mediterránea ubicada en el sur de España que comprende aproximadamente 230 km². Sierra de las Nieves (SN) presenta un entorno geológico muy diverso, que puede dividirse a grandes rasgos en dos dominios geológicos principales dispuestos a lo largo de una orientación NE-SO: en el área sur, un sector de peridotita ultramáfica, y en la parte norte y central, sector de dolomía karstificada.

Se recolectaron cien muestras de suelo superficial en octubre de 2019 con una profundidad media de alrededor de 20 cm, basándose en un muestreo sistemático en una cuadrícula de 1 km, intensificado en áreas de interés. Las muestras se tamizaron y digirieron por microondas con ácidos bidedilados, para realizar una espectrometría de masas mediante Espectrometría de Masas con Plasma Acoplado Inductivamente (ICP-MS) para obtener el contenido en Cr y Ni, fuertemente correlacionadas ($R = 0.87$). Los valores observados se transformaron logarítmicamente, eliminando seis muestras con valores NA para las características de teledetección (tabla 1).

Tabla 1. Resumen estadístico de los metales pesados utilizados en este estudio (n=94).

Ud: log(mg/kg)	Media	Mediana	Min	Max	Rango IQ
Cr	4.23	3.91	1.39	7.84	1.00
Ni	4.54	4.08	2.40	8.09	1.17

2.2. Características predictoras

El conjunto de datos estuvo compuesto por noventa y cuatro características predictoras, que pueden resumirse en cinco subconjuntos: compuestos de Sentinel-2 y sus derivadas, características fenológicas, características del terreno y características geológicas. Se obtuvieron cuatro compuestos trimestrales para el año 2019: composiciones de enero-marzo (primavera), abril-junio (verano), julio-septiembre (otoño) y octubre-diciembre (invierno). Las imágenes de la tesela 30SUF se descargaron a nivel de reflectancia en el Techo de la Atmósfera (TOA) utilizando el paquete de R *sen2r*. Se utilizaron siete bandas originales con una distancia de muestreo del suelo de 20 m para cada composición, es decir, las bandas roja, verde y azul del visible (R, V y A), de borde rojo (RE1, RE2 y RE3), NIR, SWIR1 y SWIR2. Los índices estacionales se derivaron de las imágenes compuestas para representar los parámetros de la cobertura del suelo (Tabla 2), comprendiendo 68 características (9 bandas y 8 índices por composición).

Tabla 2. Índices extraídos de los compuestos de este estudio.

Índice	Abb.	Formula
Índice de vegetación de diferencia normalizada	NDVI	$(NIR-R)/(NIR+R)$
Índice de vegetación mejorado 2	EVI2	$2.5*(NIR-R)/(NIR+R+1)$
Índice de diferencia normalizada del agua	NDWI	$(V-NIR)/(V+NIR)$
Índice de clorofila verde	GCI	$(NIR/V)-1$
Índice Rojo/Verde	RG_FE	Red/Green
Índice de reflectancia de la antocianina	ARI	$(1/V)-(1/RE1)$
Silicatos ferrosos	FS	SWIR2/SWIR1
Óxidos ferrosos	FO	SWIR1/NIR

Las características fenológicas se obtuvieron de la serie temporal de Sentinel 2 para el período 2018-2020 (292 imágenes, 73 por año). Las series temporales de EVI2 se preprocesaron para reducir cualquier artefacto en las series temporales que pudiera alterar el patrón temporal derivado de los datos satelitales (Caparros-Santiago *et al.*, 2021). Las series temporales se suavizaron utilizando la función logística doble implementada en el software TIMESAT (Jonsson and Eklundh, 2002). Se extrajeron las trece fenométricas para la primera estación detectada del año 2019, utilizando un método basado en un umbral del 10 %.

Se extrajeron varias características del terreno derivadas del Modelo Digital de Elevaciones (MDE) del área de estudio. Estos MDE se obtuvieron por interpolación a partir de los vuelos LIDAR de la primera cobertura del Plan Nacional de Ortofotografía Aérea (PNOA), con una densidad de 0,5 puntos/m². La distancia de muestreo del suelo original fue de 5 metros, remuestreada a 20 m y a la malla de S-2. Las características del terreno se

calcularon en SAGA GIS 7.8.2 . Se derivaron dos características geológicas a partir de información geológica cualitativa del Instituto Geológico y Minero de España (IGME): la distancia a rocas peridotitas y la distancia a fracturas.

2.3. Modelado predictivo

El marco de modelado predictivo constó de tres etapas clave: (i) el entrenamiento de un modelo apilado basado en tres algoritmos base (Random Forest (RF) , Máquinas de Vector Soporte (SVR) y una Red Neuronal (NN)) usando RF como algoritmo de apilado, (ii) un método de Selección de Características utilizando la Selección Secuencial hacia Adelante (SFS) aplicado tras la selección de los hiperparámetros de los modelos apilados, y (iii) la aplicación de dos técnicas de IML para la interpretación del modelo a posteriori: importancia de características agnóstica al modelo y gráficas de efectos locales acumulados. Se usaron dos métricas para evaluar la exactitud, R^2 y raíz del error cuadrático medio, con las siguientes fórmulas:

$$R^2 = 1 - \frac{RSS}{TSS}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - o_i)^2}{n}}$$

donde: RSS = Suma de cuadrados residuales, TSS = Suma total de cuadrados, n = número de observaciones, p_i = valor i-ésimo predicho, o_i = valor observado para la observación i-ésima.

3. RESULTADOS

Los resultados del modelado de Cr y Ni se muestran en la Tabla 3, mostrando que el enfoque de apilamiento por stacks con el modelo de selección SFS superó a los algoritmos base utilizados en este estudio. Los mejores resultados comparados por MP fueron obtenidos para Ni, con $R^2 = 0,6$ y $RMSE = 0,78 \log(\text{mg/kg})$, mientras que Cr alcanzó un R^2 de 0,41. Se observó una clara jerarquía de rendimiento entre los enfoques de modelado. El modelo apilado con selección de características demostró una capacidad predictiva superior, seguido por el modelo apilado con todas las características, el cual a su vez superó a los mejores aprendices base individuales.

Tabla 3. Resultados del modelado.

Algoritmo	Métrica	Cr	Ni
SVR	R^2	-0,02	-0,34
	RMSE	1,13	1,52
RF	R^2	0,13	0,07
	RMSE	0,95	1,22
NN	R^2	-1,05	-0,03
	RMSE	1,03	1,24
STACK	R^2	0,13	-0,05
	RMSE	0,84	1,07
STACK+SFS	R^2	0,41	0,6
	RMSE	0,68	0,78

La Figura 2 resume los resultados de IML para el modelo apilado de Cr. La distancia a las peridotitas se identificó como el predictor más significativo (importancia media de 1,97); los gráficos ALE revelan que las muestras

próximas a estas litologías exhiben predicciones de Cr superiores a la media en más de 2 $\log(\text{mg/kg})$. Las bandas SWIR1_aut y SWIR2_aut siguieron en relevancia, mostrando correlaciones específicas entre la reflectancia y las variaciones en la predicción de concentración. Paralelamente, la Figura 3 presenta la clasificación para el modelo de Ni. La distancia a las peridotitas se confirmó nuevamente como el factor determinante (importancia de 2,67), mostrando en los gráficos ALE una tendencia análoga al Cr, con valores elevados en zonas próximas a peridotitas. La Elevación y SWIR2_sum destacaron como predictores secundarios; elevaciones superiores a 1000 m y baja reflectancia en SWIR2_sum se asociaron con desviaciones positivas en la concentración de Ni. Al igual que en el modelado de Cr, la mayoría de las variables seleccionadas provienen de datos de teledetección.

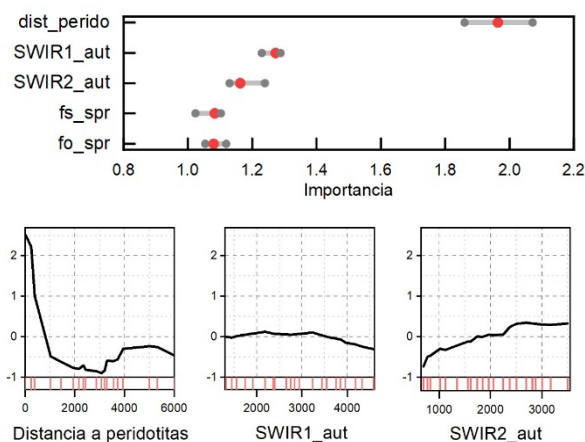


Figura 2. Resultados del IML para Cr. Arriba: clasificación de características para la predicción de Cr utilizando el modelo apilado con selección de características. Abajo: gráficos de efecto local acumulado para las tres características más importantes en la predicción de Cr.

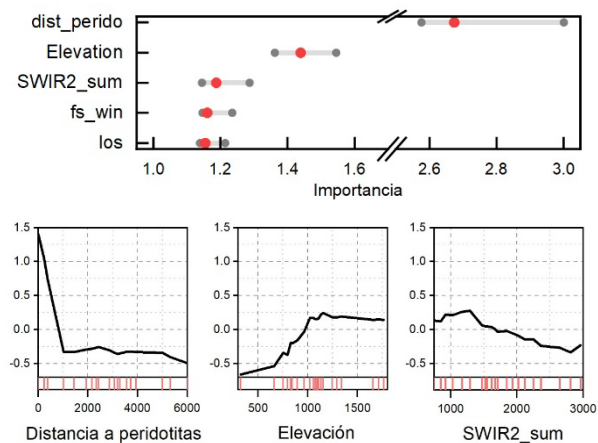


Figura 3. Resultados de IML para Ni. Arriba: clasificación de características para la predicción de Ni utilizando el modelo apilado con selección de características. Abajo: gráficos de efecto local acumulado para las tres características más importantes en la predicción de Ni.

La Figura 4 ilustra la distribución espacial de los MP del suelo. El Cr y el Ni exhibieron patrones espaciales similares, con un eje de alta concentración (más de 5,2

y 6,43 log(mg/kg) para Cr y Ni respectivamente) extendiéndose desde el este hacia el sur, correspondiendo a la geología de peridotitas. Sin embargo, en el resto del área de estudio, estas dos características mostraron patrones espaciales distintos.

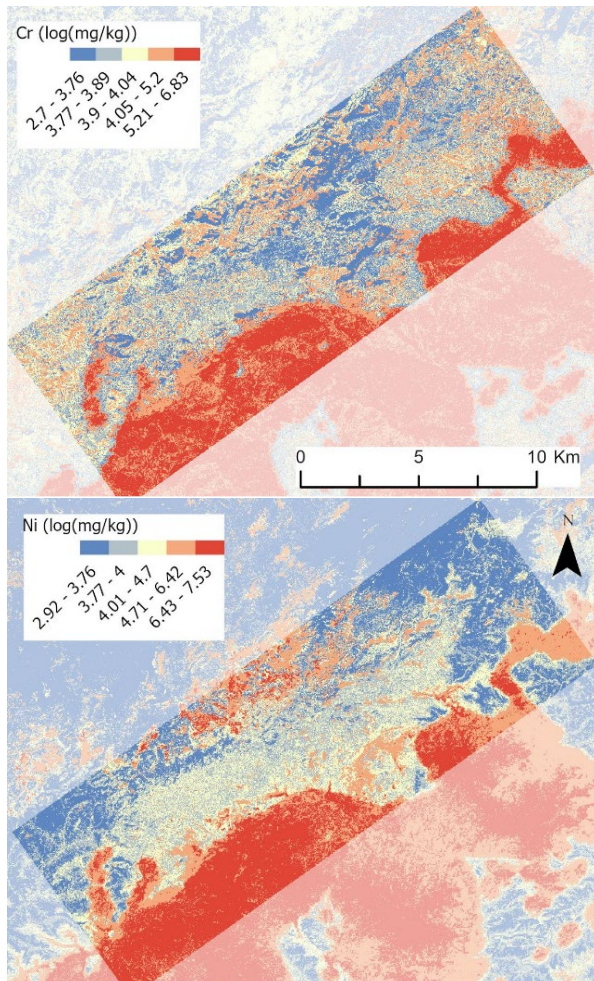


Figura 4. Cartografía de Cr (arriba) y Ni (abajo) a partir de la predicción del mejor modelo.

4. CONCLUSIONES

Este estudio modeló el contenido en cromo (Cr) y níquel (Ni) del suelo, integrando el aprendizaje por stacks con IML a partir de múltiples conjuntos de características predictoras de teledetección, terreno y características geológicas. La distancia a las peridotitas fue clave para las predicciones de Cr y Ni. Las observaciones dentro del área de peridotita se asociaron con concentraciones elevadas debido a depósitos localizados de mineral de Cr-Ni. Sin embargo, algunas características de teledetección, particularmente las composiciones SWIR, demostraron ser significativamente relevantes. En general, los modelos de apilamiento con selección de características superaron a los apilados con todas las características y a los modelos base. Los resultados destacaron el rendimiento del Ni, con el R^2 arrojando una precisión robusta ($R^2 = 0,6$; $RMSE = 0,78$), con resultados moderados logrados por los modelos de Cr ($R^2 = 0,41$; $RMSE = 0,68$). Estas metodologías de cartografía de suelos todavía pueden enfrentar limitaciones debido a la

complejidad geológica y de la cobertura del suelo mediterránea. Las futuras misiones hiperespectrales (p. ej., CHIME, SBG) que combinan altas resoluciones temporales y espectrales podrían abordar estos desafíos, mejorando la predicción indirecta de MP del suelo en entornos heterogéneos.

5. AGRADECIMIENTOS

El primer autor es beneficiario de una ayuda FPU financiada por el Ministerio de Universidades español (Referencia FPU18/04274). Los autores desean expresar su agradecimiento por el apoyo financiero proporcionado por el proyecto 2914/2022, financiado por el Organismo Autónomo de Parques Nacionales, el Ministerio para la Transición Ecológica y el Reto Demográfico. MPM agradece a la Fundación Portuguesa para la Ciencia y la Tecnología (FCT) la financiación estratégica de CERENA (UIDB/04028/2025) y a través del proyecto individual 2023.08653.CEECIND.

6. REFERENCIAS

- Azizi, K., Zhao, S., Ayoubi, S., Kamangarpour, A., Nabiollahi, K., Ge, Y., Zhao, K., Rodrigo-Comino, J., & Demattê, J. A. M. (2026). Integrating proximal sensing data (soil spectra and magnetic susceptibility) with the common covariates to predict soil carbon pools in semiarid regions. *JSTILL*, 256. <https://doi.org/10.1016/j.still.2025.106840>
- Caparros-Santiago, J. A., Rodriguez-Galiano, V., & Dash, J. (2021). Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review. *ISPRS/JPRS*, 171, 330-347. <https://doi.org/10.1016/j.isprsjprs.2020.11.019>
- Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1824-1832. <https://doi.org/10.1109/TGRS.2002.802519>
- Kooistra, L., Salas, E. A. L., Clevers, J. G. P. W., Wehrens, R., Leuven, R. S. E. W., Nienhuis, P. H., & Buydens, L. M. C. (2004). Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains. *Environmental Pollution*, 127(2), 281-290. [https://doi.org/10.1016/s0269-7491\(03\)00266-5](https://doi.org/10.1016/s0269-7491(03)00266-5)
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1-2), 3-52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Moradpour, S., Entezari, M., Ayoubi, S., Karimi, A., & Naimi, S. (2023). Digital exploration of selected heavy metals using Random Forest and a set of environmental covariates at the watershed scale. *Journal of Hazardous Materials*, 455. <https://doi.org/10.1016/j.jhazmat.2023.131609>
- Wang, F., Gao, J., & Zha, Y. (2018). Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges. *ISPRS/JPRS*, 136, 73-84. <https://doi.org/10.1016/j.isprsjprs.2017.12.003>