

Isgró, M.; Basallote Sánchez, M.; Moreno González, R.; Barbero, L. Predicción de la calidad del agua en entornos afectados por la minería: un estudio de teledetección hiperespectral en la faja pirítica Ibérica

Predicción de la calidad del agua en entornos afectados por la minería: un estudio de teledetección hiperespectral en la faja pirítica Ibérica

Isgró, Melisa ^{1,2} Basallote Sánchez, M Dolores ¹ Moreno González, Raúl ² Barbero, Luis ²

¹ Instituto de Ciencias Marinas de Andalucía (ICMAN), CSIC, Departamento de Ecología y Gestión Costera, E-11510, Puerto Real, Cádiz, España

² Departamento de Ciencias de la Tierra, Universidad de Cádiz, Av. República Saharaui s/n, Puerto Real, 11510, Cádiz, España

ORCID: Isgró 0000-0003-2194-8020 Basallote Sánchez 0000-0003-2011-3806 Moreno González 0000-0002-5679-2110 Barbero 0000-0002-3513-2025

Correspondencia: meli.isgro@gmail.com mdolores.basallote@csic.es raul.morenogonzalez@uca.es luis.barbero@uca.es

RESUMEN

Este estudio integra imágenes hiperespectrales obtenidas mediante vehículos aéreos no tripulados (UAV) con técnicas de aprendizaje automático para estimar concentraciones de pH, Fe y SO₄²⁻ en nueve masas de agua de las minas abandonadas de Tharsis y Lagunazo, en la Faja Pirítica Ibérica (SO España). Estos emplazamientos presentan un gradiente hidroquímico de contaminación extrema oscilando entre condiciones de extrema acidez y neutras. Se utilizaron mediciones in situ para entrenar diversos modelos de regresión basados en un ortomosaico hiperespectral de 221 bandas (398.8–887.1 nm). Se evaluaron cinco algoritmos, destacando el Perceptrón Multicapa (MLP) como el mejor para la predicción del pH (R² = 0.952) y Random Forest para el Fe (R² = 0.861) y el S (R² = 0.880). Los resultados subrayan el potencial de combinar sensores espectrales embarcados en drones y aprendizaje automático para el monitoreo operativo de alta resolución en entornos mineros impactados.

Palabras clave: Datos hiperespectrales, drones, pequeñas masas de agua, monitoreo, drenaje ácido de mina.

Fecha de recepción: 16 febrero 2026 · Fecha de aceptación: 17 febrero 2026

Predicción de la calidad del agua en entornos afectados por la minería: un estudio de teledetección hiperespectral en la faja pirítica Ibérica

Isgró, Melisa ^(1,2), Basallote Sánchez, M Dolores ⁽¹⁾, Moreno González, Raúl ⁽²⁾, Barbero, Luis ⁽²⁾

⁽¹⁾ Instituto de Ciencias Marinas de Andalucía (ICMAN), CSIC, Departamento de Ecología y Gestión Costera, E-11510, España.

 0000-0003-2194-8020, meli.isgro@gmail.com ;  0000-0003-2011-3806, mdolores.basallote@csic.es.

⁽²⁾ Departamento de Ciencias de la Tierra, Universidad de Cádiz, España.

 0000-0002-5679-2110, raul.morenogonzalez@uca.es ;  0000-0002-3513-2025, luis.barbero@uca.es.

Resumen: Este estudio integra imágenes hiperespectrales obtenidas mediante vehículos aéreos no tripulados (UAV) con técnicas de aprendizaje automático para estimar concentraciones de pH, Fe y SO₄²⁻ en nueve masas de agua de las minas abandonadas de Tharsis y Lagunazo, en la Faja Pirítica Ibérica (SO España). Estos emplazamientos presentan un gradiente hidroquímico de contaminación extrema oscilando entre condiciones de extrema acidez y neutras. Se utilizaron mediciones in situ para entrenar diversos modelos de regresión basados en un ortomosaico hiperespectral de 221 bandas (398.8–887.1 nm). Se evaluaron cinco algoritmos, destacando el Perceptrón Multicapa (MLP) como el mejor para la predicción del pH (R² = 0.952) y *Random Forest* para el Fe (R² = 0.861) y el S (R² = 0.880). Los resultados subrayan el potencial de combinar sensores espectrales embarcados en drones y aprendizaje automático para el monitoreo operativo de alta resolución en entornos mineros impactados.

Palabras clave: Datos hiperespectrales, drones, pequeñas masas de agua, monitoreo, drenaje ácido de mina.

Predicting Water Quality in Mine-Affected Environments: A Hyperspectral Remote Sensing Study in the Iberian Pyrite Belt

Abstract: *This study integrates hyperspectral images obtained by unmanned aerial vehicles (UAVs) with machine learning techniques to estimate pH, Fe, and SO₄²⁻ concentrations in nine water bodies of the abandoned Tharsis and Lagunazo mines in the Iberian Pyrite Belt (SW Spain). These sites present a contamination hydrochemical gradient ranging from extreme acidity to circum-neutral conditions. In situ measurements were used to train various regression models based on a 221-band (398.8–887.1 nm) hyperspectral orthomosaic. Five algorithms were evaluated, highlighting the Multilayer Perceptron (MLP) as the best for pH prediction (R² = 0.952) and Random Forest for Fe (R² = 0.861) and S (R² = 0.880). The results emphasize the potential of combining drone-mounted spectral sensors and machine learning for high-resolution operational monitoring in impacted mining environments.*

Keywords: *Hyperspectral data, drones, small water bodies, monitoring, acid mine drainage*

1. INTRODUCCIÓN

La Faja Pirítica Ibérica (IPB), en el suroeste de la Península Ibérica, alberga una de las mayores concentraciones de sulfuros masivos del mundo. La intensa actividad minera histórica desarrollada en la zona ha dejado tras de sí numerosas cortas abiertas, escombreras y depósitos de residuos que, al quedar expuestos a la meteorización, generan el fenómeno conocido como Drenaje Ácido de Mina (AMD) (Cánovas *et al.*, 2016; Huang *et al.*, 2023; Nieto *et al.*, 2007; Olías *et al.*, 2021; Sánchez España *et al.*, 2005). Estos lixiviados se caracterizan por una acidez extrema (a menudo con

pH < 3) y una carga masiva de elementos tóxicos disueltos.

La monitorización de estas masas de agua es fundamental para la protección de las cuencas fluviales receptoras (como los ríos Tinto y Odiel). Sin embargo, los métodos de muestreo convencionales se enfrentan a desafíos logísticos: la topografía escarpada de las minas abandonadas dificulta el acceso y la alta variabilidad espacial de las lagunas impide una caracterización completa mediante muestras puntuales. La teledetección satelital, aunque útil a escala regional,

carece de la resolución espacial necesaria para las dimensiones de estas cortas. En este contexto, los vehículos aéreos no tripulados (UAV) emergen como una solución intermedia ideal, proporcionando una "ultra-alta" resolución que permite capturar la heterogeneidad química superficial de estas aguas

(Koerting *et al.*, 2024; Ma *et al.*, 2023).

2. MATERIAL Y MÉTODOS

2.1. Área de estudio y toma de muestras

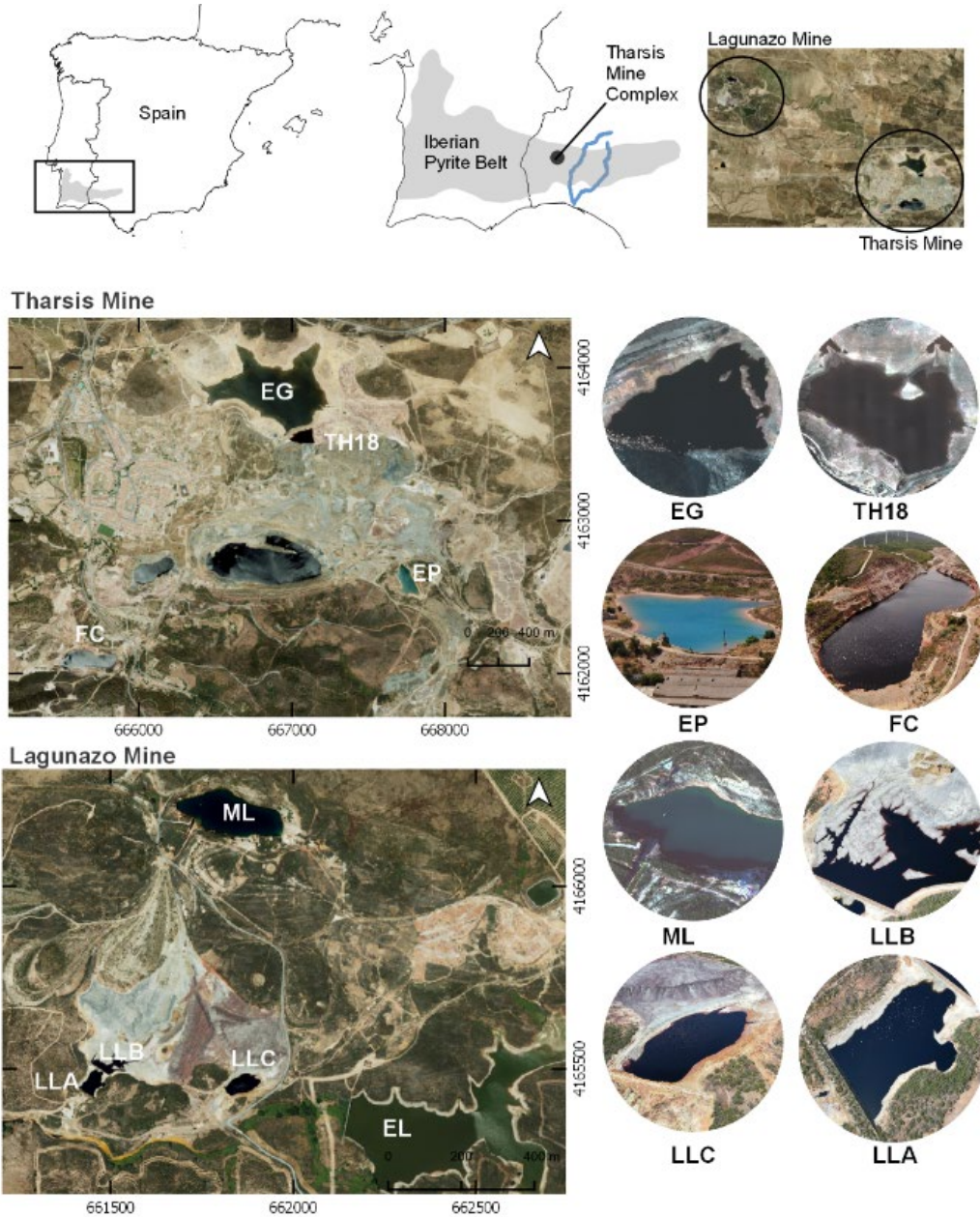


Figura 1. Mapa de localización de los puntos de muestreo. Se muestrearon cuatro masas de agua en la Mina de Tharsis: Filón Centro (FC), Embalse del Pino (EP), Embalse Grande (EG) y Th18. Adicionalmente, se muestrearon otras cinco masas de agua en el emplazamiento de la Mina Lagunazo: Mina Lagunazo (ML), Embalse Lagunazo (EL) y las Lagunas A, B y C (LLA, LLB y LLC, respectivamente).

El diseño experimental se basó en capturar el máximo gradiente hidroquímico posible para lo que se seleccionaron nueve masas de agua representativas en las minas de Tharsis y Lagunazo (Huelva, España; Fig 1). Paralelamente a los vuelos de dron, se recogieron muestras de agua a 10-20 cm de profundidad. El pH y la conductividad eléctrica se midieron in situ, mientras que la concentración de Fe y S se determinó en laboratorio

mediante Espectroscopía de Emisión Óptica de Plasma Acoplado Inductivamente (ICP-OES).

2.2. Adquisición y preprocesamiento hiperspectral

Se utilizó un hexacóptero DJI Matrice 600 equipado con el sensor Headwall Nano-Hyperspec (rango 400-1000 nm). Los vuelos se realizaron a una altura de 90 metros

con una velocidad de 4 m/s. El flujo de procesamiento incluyó:

- Corrección Radiométrica: Conversión de niveles digitales a radiancia.
- Corrección Atmosférica: Aplicación del Método de Línea Empírica (ELM) utilizando paneles de calibración con reflectancia conocida (2%, 12%, 24% y 50%) colocados en el área de estudio.
- Ortorrectificación: Integración de datos de la IMU y GPS para corregir distorsiones geométricas.

El resultado fue un ortomosaico de 221 bandas espectrales con un tamaño de píxel de 9 cm.

2.3. Pipeline de Aprendizaje Automático

Dada la alta colinealidad de las 221 bandas seleccionadas, se aplicaron dos estrategias de reducción de datos: 1) PCA (Análisis de Componentes Principales): Transformación de las bandas en componentes ortogonales que explican el 99% de la varianza. 2) LASSO (*Least Absolute Shrinkage and Selection Operator*): Selección de las bandas más influyentes eliminando aquellas con coeficientes nulos.

Se evaluaron cinco regresores: *Random Forest* (RF), *k-Nearest Neighbors* (KNN), *Partial Least Squares Regression* (PLSR), *XGBoost* y un Perceptrón Multicapa (MLP) de tres capas ocultas. El entrenamiento se realizó mediante una validación cruzada de 5 pliegues para asegurar la generalización de los modelos.

Se analizaron un total de 13 muestras derivadas de imágenes hiperespectrales (una por cuerpo de agua). Cada muestra consistió en firmas espectrales promediadas, obtenidas de seis regiones de interés (ROI) representativas por cada cuerpo de agua, lo que resultó en 78 espectros por parámetro. Los datos espectrales abarcaron el rango VNIR (aprox. 398–887 nm) con 221 bandas por firma. El entrenamiento de los modelos se realizó utilizando el conjunto de datos completo de 221 bandas VNIR, o bien subconjuntos multibanda seleccionados mediante LASSO o transformados a través de PCA. Se utilizó una validación cruzada de 5 iteraciones (K-Fold) para evaluar el rendimiento y la precisión del modelo. Este método es particularmente adecuado para conjuntos de datos pequeños, ya que optimiza el uso de muestras limitadas al tiempo que mitiga el riesgo de sobreajuste. El conjunto de datos se divide en cinco grupos (folds) de igual tamaño. En cada iteración, se utilizaron cuatro grupos para entrenar el modelo y el grupo restante se utilizó para la validación.

3. RESULTADOS Y DISCUSIÓN

Los cuerpos de agua seleccionados presentan un amplio gradiente hidroquímico, que abarca desde condiciones neutras hasta extremadamente ácidas ($\text{pH} < 1$ a 7.8) y desde soluciones diluidas a ricas en metales (e.g., Fe, Cu, Zn). Las masas de agua EP, FC, ML, TH18, LLA, LLB y LLC se clasifican como ácidas ($\text{pH} < 3.8$), mientras que EG y EL representan aguas circum-neutras ($\text{pH} 6.6\text{--}7.7$).

Las concentraciones de hierro oscilan entre 0.053 y 69 g/L. Del mismo modo, las concentraciones de sulfato varían de 49.33 a 84 g/L, notándose una variabilidad sustancial en el contenido de sulfato entre los cuerpos de agua.

3.1 Comportamiento espectral de las aguas de mina

Se observó una clara diferenciación espectral según el pH y la carga metálica. Los cuerpos de agua con pH neutro presentan una firma típica de agua limpia con baja reflectancia. Sin embargo, las aguas afectadas por AMD muestran una fuerte absorción en las longitudes de onda del azul (< 500 nm), causada por la presencia de hierro disuelto y la formación de complejos de sulfato de hierro. Además, se detectó un pico de reflectancia pronunciado entre 600 y 700 nm (región del rojo), coincidente con el color rojizo de las aguas más ácidas debido a la suspensión de precipitados minerales como la schwertmannita.

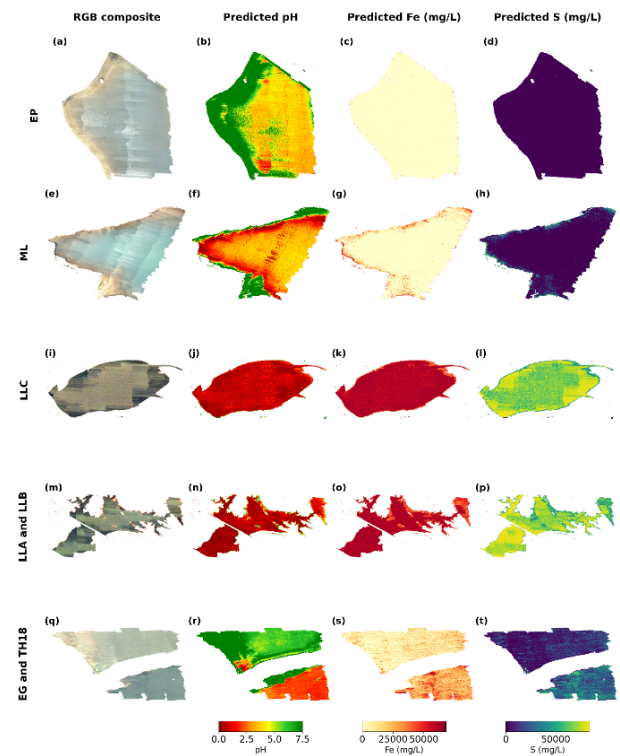


Figura 2. Imagen RGB hiperespectral compuesta y parámetros de calidad del agua predichos para EP, ML, LLC, LLA y LLB (2023), EG y TH18. a) Imagen compuesta RGB generada a partir de bandas VNIR seleccionadas del cubo hiperespectral; b) Predicción de la distribución de pH; c) Predicción de la concentración de Fe en mg/L; d) Predicción de la concentración de S en mg/L.

3.2 Evaluación del rendimiento de los modelos

Los resultados muestran una capacidad de predicción excelente para los tres parámetros clave (Tabla 1):

- pH: El modelo MLP fue el más preciso ($R^2=0.952$, $\text{RMSE}=0.51$). Esto sugiere que las redes neuronales captan mejor la relación no lineal entre la reflectancia y la escala logarítmica del pH.

- Hierro y Azufre: Los modelos de *Random Forest* (RF) superaron al resto cuando se utilizaron las variables seleccionadas por LASSO ($R^2=0.861$ para Fe y $R^2=0.880$ para S). La robustez de RF ante valores atípicos fue fundamental dada la presencia de concentraciones extremas en algunas lagunas.

Tabla 1. Resumen de los principales resultados obtenidos por parámetro.

| Parámetro | Mejor Modelo | R2 | RMSE |
|-----------|--------------|-------|-------|
| pH | MLP | 0.952 | 0.51 |
| Fe (mg/L) | RF | 0.861 | 185.2 |
| S (mg/L) | RF | 0.880 | 542.4 |

3.3 Análisis de importancia de variables (SHAP)

La aplicación de valores SHAP (*SHapley Additive exPlanations*) permitió interpretar los modelos de "caja negra". Para el pH y el Fe, las bandas más críticas se situaron en el rango de 450-480 nm y 670-710 nm (Fig. 3). Esto valida físicamente los modelos, ya que estas regiones espectrales están directamente ligadas a los procesos de absorción y dispersión del hierro en sus diferentes estados de oxidación.

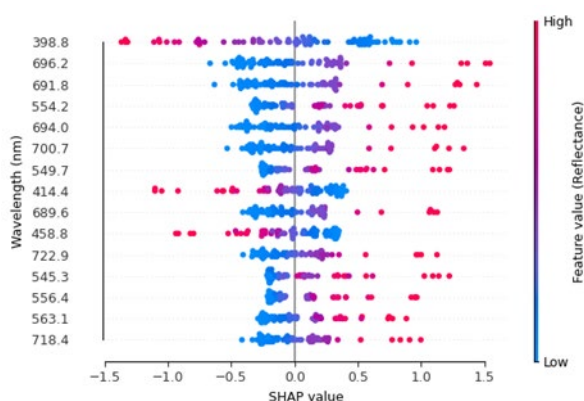


Figura 3. Gráfico de resumen SHAP para la predicción del pH mediante MLP con una representación espectral completa

4. CONCLUSIONES

La integración de sensores hiperespectrales en drones y técnicas de aprendizaje automático representa un avance disruptivo en el monitoreo ambiental minero. La metodología propuesta permite obtener mapas continuos de calidad del agua con una precisión cercana a los análisis de laboratorio, eliminando los riesgos del muestreo manual en zonas peligrosas. La capacidad de predecir el pH y la carga metálica de forma remota facilita una respuesta rápida ante posibles vertidos y permite evaluar la eficacia de los procesos de remediación y tratamiento de aguas de mina de manera operativa y sistemática.

5. AGRADECIMIENTOS

MD Basallote agradece a la Agencia Estatal de Investigación (AEI) la financiación del contrato RYC2022-035326-I, cofinanciado por

MICIU/AEI/10.13039/501100011033 y FSE+. R Moreno-gonzález agradece por su ayuda DGP_POST_2024_00813 (Junta de Andalucía/CUII y FSE+). El equipamiento científico ha sido financiado por los proyectos de infraestructuras EQC2018-004446-P (Gobierno de España) y PAIDI2020-INFRA/IN (Junta de Andalucía y Ministerio de Universidades). Los autores agradecen las facilidades otorgadas por la empresa minera *Tharsis Mining S.L.* para el acceso a las instalaciones de Tharsis y Lagunazo.

6. REFERENCIAS

- Cánovas, C. R., Olías, M., Macias, F., Torres, E., San Miguel, E. G., Galván, L., Ayora, C., & Nieto, J. M. (2016). Water acidification trends in a reservoir of the Iberian Pyrite Belt (SW Spain). *Science of the Total Environment*, *541*, 400–411. <https://doi.org/10.1016/j.scitotenv.2015.09.070>
- Huang, C., Chen, X. Y., & Lee, M. (2023). An improved hyperspectral sensing approach for the rapid determination of copper ion concentrations in water environment using short-wavelength infrared spectroscopy. *Environmental Pollution*, *333*(February), 121984. <https://doi.org/10.1016/j.envpol.2023.121984>
- Koerting, F. M., Asadzadeh, S., Hildebrand, J. C., Savinova, E., Kouzeli, E., Nikolakopoulos, K. G., Lindblom, D., Koellner, N., Buckley, S. J., Lehman, M. R., Schlöpfer, D., & Micklethwaite, S. (2024). *Imaging Spectroscopy for Mining: Insights for Hyperspectral Drone Applications*. <https://doi.org/10.20944/preprints202407.1407.v1>
- Ma, T., Zhang, D., Li, X., Huang, Y., Zhang, L., Zhu, Z., Sun, X., Lan, Z., & Guo, W. (2023). Hyperspectral remote sensing technology for water quality monitoring: knowledge graph analysis and Frontier trend. *Frontiers in Environmental Science (Vol. 11)*. Frontiers Media SA. <https://doi.org/10.3389/fenvs.2023.1133325>
- Nieto, J. M., Sarmiento, A. M., Olías, M., Canovas, C. R., Riba, I., Kalman, J., & Delvalls, T. A. (2007). Acid mine drainage pollution in the Tinto and Odiel rivers (Iberian Pyrite Belt, SW Spain) and bioavailability of the transported metals to the Huelva Estuary. *Environment International*, *33*(4), 445–455. <https://doi.org/10.1016/j.envint.2006.11.010>
- Olías, M., Cánovas, C. R., & Basallote, M. D. (2021). Surface and Groundwater Quality Evolution in the Agrio and Guadiamar Rivers After the Aznalcóllar Mine Spill (SW Spain): Lessons Learned. *Mine Water and the Environment*, *40*(1), 235–249. <https://doi.org/10.1007/s10230-020-00713-7>
- Sánchez España, J., Pamo, E. L., Santofimia, E., Aduvire, O., Reyes, J., & Baretino, D. (2005). Acid mine drainage in the Iberian Pyrite Belt (Odiel river watershed, Huelva, SW Spain): Geochemistry, mineralogy and environmental implications. *Applied Geochemistry*, *20*(7), 1320–1356. <https://doi.org/10.1016/j.apgeochem.2005.01.011>