

Barba Ceballos, P.; Heredia, S.; Navarro, G.; Rosado, B.; Roca Mora, M.; Caballero, I. Zonificación del Mar Menor mediante aprendizaje no supervisado sobre la serie histórica de Sentinel-2.

Zonificación del Mar Menor mediante aprendizaje no supervisado sobre la serie histórica de Sentinel-2.

Barba Ceballos, Paola ¹ Heredia, Sergio ¹ Navarro, Gabriel ¹ Rosado, Belén ² Roca Mora, Mar ¹ Caballero, Isabel ¹

¹ Instituto de Ciencias Marinas de Andalucía (ICMAN), España

² Laboratorio de Astronomía, Geodesia y Cartografía, Universidad de Cádiz

ORCID: Barba Ceballos 0000-0001-7726-6147 Heredia 0009-0003-9495-9625 Navarro 0000-0002-8919-0060 Rosado 0000-0003-1677-904X Roca Mora 0000-0001-7311-2618 Caballero 0000-0001-7485-0989

Correspondencia: paola.barba@csic.es sergio.h.c@csic.es gabriel.navarro@csic.es belen.rosado@uca.es
mar.roca@csic.es isabel.caballero@csic.es

RESUMEN

La inteligencia artificial (IA) está transformando el monitoreo ambiental al permitir el análisis automatizado de conjuntos de datos complejos y de alta dimensionalidad. En este estudio, se presenta un marco metodológico basado en IA, combinado con la descomposición estacional y de tendencia mediante Loess (STL), para caracterizar los patrones de calidad del agua y sus dinámicas temporales en el Mar Menor, identificar las áreas más afectadas por la degradación ecológica y analizar su dinámica espacio-temporal mediante imágenes satelitales multiespectrales. La serie histórica de Sentinel-2, desde julio de 2015 hasta junio de 2024, se procesó mediante la corrección atmosférica ACOLITE. Se han separado la tendencia a largo plazo, la variabilidad estacional y las anomalías residuales de la clorofila-a (Chl-a) y la turbidez, vinculando los eventos extremos ocurridos en la laguna con anomalías significativas observadas en los datos históricos de teledetección. Este enfoque híbrido permitió diferenciar regiones con comportamiento similar en cuatro zonas bien definidas. La integración de técnicas de clustering no supervisado basadas en IA y de descomposición estadística de series temporales ofrece una metodología potente, escalable y transferible para el monitoreo automatizado de regiones costeras dinámicas. Estos resultados proporcionan información útil y ofrecen una base analítica robusta para futuras estrategias de prevención y restauración.






Palabras clave: *k-means*, descomposición STL, zonificación, degradación.

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


Zonificación del Mar Menor mediante aprendizaje no supervisado sobre la serie histórica de Sentinel-2.

Barba Ceballos, Paola ⁽¹⁾, Heredia, Sergio ⁽¹⁾, Navarro, Gabriel ⁽¹⁾, Rosado, Belén ⁽²⁾, Roca Mora, Mar ⁽¹⁾, Caballero, Isabel ⁽¹⁾

⁽¹⁾ Instituto de Ciencias Marinas de Andalucía (ICMAN), España.

 0000-0001-7726-6147, paola.barba@csic.es ;  0009-0003-9495-9625, sergio.h.c@csic.es
 0000-0002-8919-0060, gabriel.navarro@csic.es ;  0000-0001-7311-2618, mar.roca@csic.es
 0000-0001-7485-0989, isabel.caballero@csic.es

⁽²⁾ Laboratorio de Astronomía, Geodesia y Cartografía, Universidad de Cádiz.

 0000-0003-1677-904X, belen.rosado@uca.es

Resumen: La inteligencia artificial (IA) está transformando el monitoreo ambiental al permitir el análisis automatizado de conjuntos de datos complejos y de alta dimensionalidad. En este estudio, se presenta un marco metodológico basado en IA, combinado con la descomposición estacional y de tendencia mediante Loess (STL), para caracterizar los patrones de calidad del agua y sus dinámicas temporales en el Mar Menor, identificar las áreas más afectadas por la degradación ecológica y analizar su dinámica espacio-temporal mediante imágenes satelitales multiespectrales. La serie histórica de Sentinel-2, desde julio de 2015 hasta junio de 2024, se procesó mediante la corrección atmosférica ACOLITE. Se han separado la tendencia a largo plazo, la variabilidad estacional y las anomalías residuales de la clorofila-a (Chl-a) y la turbidez, vinculando los eventos extremos ocurridos en la laguna con anomalías significativas observadas en los datos históricos de teledetección. Este enfoque híbrido permitió diferenciar regiones con comportamiento similar en cuatro zonas bien definidas. La integración de técnicas de clustering no supervisado basadas en IA y de descomposición estadística de series temporales ofrece una metodología potente, escalable y transferible para el monitoreo automatizado de regiones costeras dinámicas. Estos resultados proporcionan información útil y ofrecen una base analítica robusta para futuras estrategias de prevención y restauración.

Palabras clave: k-means, descomposición STL, zonificación, degradación.

Zoning of the Mar Menor using unsupervised learning on the Sentinel-2 historical series.

Abstract: Artificial intelligence (AI) is transforming environmental monitoring by enabling automated analysis based on complex, high-dimensional datasets. This study presents a methodological framework that combines AI with Seasonal and Trend decomposition using Loess (STL) to characterize water quality patterns and their temporal dynamics in the Mar Menor coastal lagoon, identify the areas most affected by ecological degradation, and analyze their spatio-temporal dynamics using multispectral satellite imagery. The Sentinel-2 time series, spanning from July 2015 to June 2024, was processed using ACOLITE atmospheric correction. The study clearly separated long-term trends, seasonal variability, and residual anomalies in chlorophyll-a (Chl-a) and turbidity, linking extreme events in the lagoon to significant anomalies observed in historical remote sensing data. This hybrid approach enabled the identification of four well-defined zones with similar behavior. The integration of unsupervised AI-based clustering techniques and statistical time-series decomposition provides a powerful, scalable, and transferable methodology for the automated monitoring of coastal and dynamic regions. These results provide useful information and offer a robust analytical basis for future prevention and restoration strategies.

Keywords: k-means, STL decomposition, zoning, degradation

1. INTRODUCCIÓN

La integración de técnicas de inteligencia artificial (IA), aprendizaje automático (Machine Learning, ML) y análisis de series temporales ha transformado las ciencias ambientales, permitiendo una gestión eficiente

de grandes volúmenes de datos satelitales y mejorando la capacidad de monitorización, predicción y evaluación de los ecosistemas acuáticos (White *et al.*, 2025). Estas metodologías facilitan la identificación de patrones complejos y relaciones no lineales, proporcionando una

comprensión más profunda de la dinámica espacio-temporal y apoyando el desarrollo de estrategias de gestión y conservación ambiental (Frincu, 2025). En este contexto, las lagunas costeras constituyen un entorno idóneo para su aplicación debido a su elevada heterogeneidad, compleja dinámica y alta vulnerabilidad frente a las presiones antrópicas (Pérez-Ruzafa *et al.*, 2019).

El Mar Menor, la mayor laguna hipersalina de Europa, ha sufrido una grave degradación ecológica desde 2015 causada principalmente por aportes masivos de nutrientes y sedimentos, lo que ha provocado procesos de eutrofización, incrementando los niveles de clorofila-a (Chl-a) y turbidez, deteriorando así la calidad del agua (Pérez-Ruzafa *et al.*, 2019). Estos procesos se intensifican durante episodios extremos de precipitación, favoreciendo la proliferación de algas, reduciendo la penetración de la luz, afectando negativamente a la vegetación sumergida y provocando episodios de anoxia con eventos de mortalidad masiva de fauna marina (Caballero *et al.*, 2022; Medina-López *et al.*, 2023). Por su importancia, el Mar Menor ha sido monitoreado mediante campañas de datos in situ y datos satelitales de Sentinel-3, pero las limitaciones espaciales, temporales y de resolución dificultan un análisis de la zona por su alta heterogeneidad (Giménez *et al.*, 2024; Salls *et al.*, 2024).

Para superar estas limitaciones, se propone un marco metodológico asistido por IA basado en imágenes Sentinel-2A/B de alta resolución espacial, aplicando técnicas de agrupamiento no supervisado (*k-means*) y descomposición STL de series temporales (Cleveland *et al.*, 1990; Medina-López *et al.*, 2023). Esto ha permitido definir zonas homogéneas y analizar la evolución temporal de Chl-a y turbidez, identificando tendencias, estacionalidades y anomalías asociadas a eventos extremos, proporcionando así una herramienta robusta y escalable para evaluar procesos de eutrofización y dinámicas en lagunas costeras.

2. MATERIAL Y MÉTODOS

2.1. Área de estudio

El Mar Menor es una laguna costera situada en Murcia, al SE de España. Su profundidad media es de 4 metros, siendo su máximo 7 metros. Tiene una extensión de 135 km² y está separado del Mar Mediterráneo mediante una barra de arena; cuenta con tres canales que comunican el Mar Menor y Mediterráneo entre sí, permitiendo la renovación del agua contenida en la laguna costera (Pérez-Ruzafa *et al.*, 2019). El Mar Menor se encuentra rodeado mayormente por cultivos agrícolas y cuenta con numerosas ramblas que desembocan en él, vertiendo sedimentos, productos químicos y otros contaminantes que afectan negativamente a la calidad del agua.

2.2. Descripción de los datos

Las imágenes de los satélites gemelos Sentinel-2A/B, desde julio de 2015 hasta junio de 2024, se usaron para evaluar la calidad del agua de la laguna. Estas imágenes contienen datos de 13 bandas espectrales de reflectancia, con resolución desde 60 metros hasta 10 metros, dependiendo de la banda, y un periodo de

revisita de 10 días de cada satélite; se tienen datos cada 5 días. Se procesaron las imágenes con el software ACOLITE, diseñado específicamente para zonas costeras, realizando correcciones atmosféricas y enmascarando nubes sobre los datos de reflectancia obtenidos (Vanhellemont y Ruddick, 2016). Se aplicó el software a un total de 596 imágenes y se calcularon sus índices de calidad del agua mediante combinación de bandas de reflectancia, Chl-a OC3 (O'Reilly y Werdell, 2019) y turbidez basada en la banda roja (665 nm) (Dogliotti *et al.*, 2015). Estos índices obtenidos mediante datos satelitales fueron validados con datos in situ (Caballero *et al.*, 2022).

2.3. Metodología

Para garantizar la representación temporal homogénea de los datos y evitar sesgos por el número desigual de imágenes mensuales, se promediaron los datos después del filtrado. Primero, se calculan las medias mensuales de cada mes y año del periodo de estudio. Posteriormente, se promediaron los meses durante los diferentes años, obteniendo 12 imágenes, una para cada mes del año, proporcionando información representativa sobre el comportamiento de cada mes. Por último, las 12 imágenes mensuales se promediaron entre sí para obtener una única imagen media, que representará el comportamiento de la laguna durante todo el periodo de estudio (2015-2024).

2.3.1. Filtrado k-sigma

Este filtrado se usa para la detección de valores atípicos, es simple y eficiente computacionalmente. Los valores que se desvían de la media en más de k veces la desviación estándar se consideraron valores atípicos (Rousseeuw y Hubert, 2011). En nuestro caso se usa $k=2$.

$$\mu - k\sigma \leq X \leq \mu + k\sigma \quad (1)$$

Donde X es el conjunto de datos no atípicos, μ es el valor medio y σ es la desviación típica del conjunto de datos inicial. Además, se añadió un umbral de valor 200 a ambos índices, para así descartar errores sistemáticos.

2.3.2. Análisis cluster: k-means

El análisis cluster, en concreto *k-means*, es una técnica de aprendizaje no supervisado usada para agrupar datos con valores similares, minimizando la suma de las distancias cuadráticas entre cada punto y el centroide de su grupo.

$$\sum_{i=1}^K \sum_{x_j \in C_i} \|x_j - c_i\|^2 \quad (2)$$

Donde K es el número de clústeres, $x_j \in C_i$ representa a los elementos que pertenecen al clúster C_i y c_i es el centroide del clúster i .

Se eligió este método por su alta eficacia y escalabilidad en grandes volúmenes de datos. El valor k se define mediante el método de Elbow (Everitt *et al.*, 2011).

2.3.3. Descomposición STL

STL es el acrónimo de "Seasonal and Trend decomposition using Loess"; esta descomposición separa la serie temporal de forma aditiva en sus componentes de estacionalidad, tendencia y residual.

$$Y_v = S_v + T_v + R_v \quad (3)$$

Siendo Y_v la serie temporal, S_v la componente estacional, T_v la componente tendencia y R_v la componente residual.

Los valores faltantes no influyen negativamente a la descomposición, estos se estiman durante el cálculo de las componentes tendencia y estacionalidad mediante la regresión loess. La componente residual se calcula como la diferencia de la serie y las componentes tendencia y estacionalidad (Cleveland *et al.*, 1990).

3. RESULTADOS Y DISCUSIÓN

3.1. Aplicación del análisis cluster

Mediante el método de Elbow se consigue obtener el valor óptimo de grupos, $k=4$, que deben formarse en el Mar Menor. El análisis de clúster (C^*), k -means, identifica los grupos homogéneos en función de las variables de Chl-a y turbidez (Fig. 1).

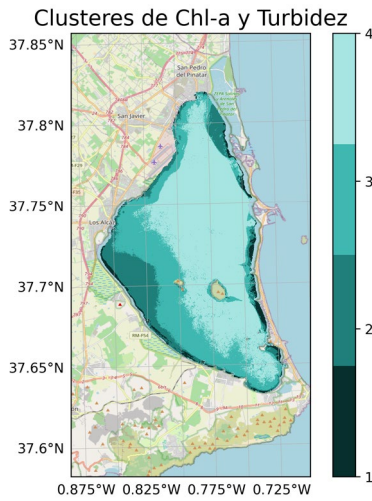


Figura 1. Resultado de k -means ($k=4$) para las variables de Chl-a y turbidez, identificando 4 grupos en el Mar Menor.

La degradación va disminuyendo del C1 al C4. Además, se analizó la compactación de cada cluster (C^*) y la separación entre los clústeres mediante la intravarianza y la distancia entre el centroide de cada clúster y el centroide global (Tabla 1). También se calculó el intervalo percentílico (5-95), C1 se encontraba completamente fuera (Tabla 2), decidiendo así eliminar C1 por su alta heterogeneidad, el cual se tomó como un clúster de valores atípicos.

Tabla 1. Intravarianza y distancia de los centroides de cada clúster y el centroide global, para las variables de Chl-a y turbidez. La intravarianza muestra una alta dispersión en C1, es decir, sus elementos son heterogéneos entre sí. Alta distancia entre C1 y el centroide global, los datos de C1 son los más alejados.

Cluster	Intravarianza	Distancia
		$ c_i - c_{global} ^2$
1	4.49	18.27
2	0.94	3.92

Cluster	Intravarianza	Distancia
		$ c_i - c_{global} ^2$
3	0.24	0.14
4	0.18	0.72

Tabla 2. Número de elementos por cada clúster. Se calculó el intervalo percentílico (IP) (5-95), indicando los puntos de cada clúster que se encuentran dentro y fuera de IP.

Cluster	Total de puntos	Dentro IP	Fuera IP	Porcentaje de puntos fuera
1	42474	0	42474	100.00 %
2	187608	125981	61627	32.85 %
3	390355	387366	2989	0.77 %
4	716613	592453	124160	17.33 %

3.2. Descomposición STL de las series temporales

Se construyeron series temporales de Chl-a y turbidez por separado, definidas por datos mensuales de cada variable en cada una de las zonas obtenidas mediante k -means, para luego aplicarse la descomposición STL de la variable Chl-a (Fig. 2). La componente tendencia muestra trayectorias similares entre zonas, con valores decrecientes de C2 a C4. La componente estacional revela dos ciclos diferenciados (2015-2020 y 2020-2024), con mayor estacionalidad en C2 durante el primer periodo y convergencia hacia C3 y C4 en el segundo. La componente residual destaca valores anómalos, principalmente en C2, asociados a eventos extremos que afectaron al Mar Menor, como eventos de “sopa verde”, anoxia, DANAs, etc.

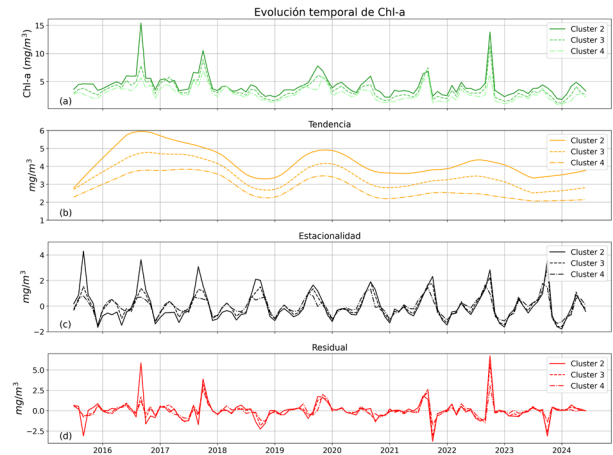


Figura 2. Descomposición STL de la serie temporal de Chl-a en (a) serie original, (b) tendencia, (c) componente estacional y (d) residual para C2, C3 y C4.

La aplicación de STL sobre la serie temporal de la variable turbidez (Fig. 3) muestra cómo la componente tendencia presenta tres ciclos principales (2016-2019, 2019-2021 y 2021-2022), con trayectorias similares en C3 y C4, mientras que C2 presenta un comportamiento diferenciado, con menores incrementos en el segundo ciclo y una tendencia decreciente desde el verano de

2021. La componente estacional presenta patrones anuales consistentes, con máximos entre primavera y mediados de verano. La componente residual identifica anomalías extremas asociadas a episodios de tormentas intensas (DANA), especialmente a finales de 2019 y en octubre de 2021, responsables de condiciones eutróficas e hipóxicas en la laguna.

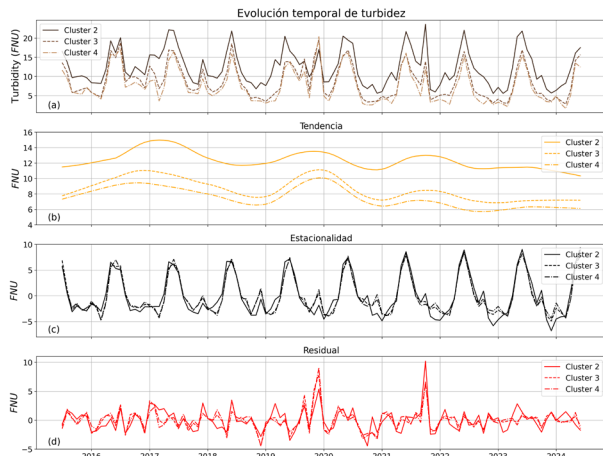


Figura 3. Descomposición STL de la serie temporal de turbidez en (a) serie original, (b) tendencia, (c) componente estacional y (d) residual para C2, C3 y C4.

La descomposición STL proporciona matemáticamente las componentes tendencia, estacionalidad y residual de las series temporales, ofreciendo información relevante sobre el comportamiento de Chl-a y turbidez en la laguna costera y relacionando valores significativos de las variables con eventos extremos o estacionales.

4. AGRADECIMIENTOS

Esta investigación ha sido financiada por el proyecto CNS2023-143630, concedido por el MICIU/AEI/10.13039/501100011033 y por la Unión Europea Next Generation EU/PRTR. Paola Barba Ceballos ha contado con el apoyo de la beca JAEPR23062, financiada por el programa JAEPRE del CSIC. Los autores agradecen al proyecto GOYAS, financiado por el proyecto OSCARS, que ha recibido financiación del programa de investigación e innovación «Horizonte Europa» de la Comisión Europea en virtud del acuerdo de subvención n.º 101129751.

5. REFERENCIAS

Caballero, I., Roca, M., Santos-Echeandía, J., Bernárdez, P., & Navarro, G. (2022). Use of the Sentinel-2 and Landsat-8 Satellites for Water Quality Monitoring: An Early Warning Tool in the Mar Menor Coastal Lagoon. *Remote Sensing*, 14(12), 2744. <https://doi.org/10.3390/rs14122744>

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*, 597 6(1), 3–73. <https://www.math.unm.edu/~ilj/Stat581/STL.pdf>

Dogliotti, A. I., Ruddick, K. G., Nechad, B., Doxaran, D., & Knaeps, E. (2015). A single algorithm to retrieve

turbidity from remotely-sensed data in all coastal and estuarine waters. *Remote Sensing of Environment*, 156, 157–168. <https://doi.org/10.1016/j.rse.2014.09.020>

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). Cluster analysis (5th ed.). Wiley. https://cicerocq.wordpress.com/wp-content/uploads/2019/05/cluster-analysis_5ed_everitt.pdf

Frincu, R. M. (2025). Artificial intelligence in water quality monitoring: A review of water quality assessment applications. *Water Quality Research Journal*, 60(1), 164–176. <https://doi.org/10.2166/wqrj.2024.049>

Giménez, J. G., Granero, A., Senent-Aparicio, J., Gómez-Jakobsen, F., Mercado, J. M., Blanco-Gómez, P., Ruiz, J. M., & Cecilia, J. M. (2024). Assessment of oceanographic services for the monitoring of highly anthropised coastal lagoons: The Mar Menor case study. *Ecological Informatics*, 81, 102554. <https://doi.org/10.1016/j.ecoinf.2024.102554>

Medina-López, E., Navarro, G., Santos-Echeandía, J., Bernárdez, P., Caballero, I. (2023). Machine Learning for Detection of Macroalgal Blooms in the Mar Menor Coastal Lagoon Using Sentinel-2. *Remote Sens.* 2023, 15, 1208. <https://doi.org/10.3390/rs15051208>

O'Reilly, J. E., & Werdell, P. J. (2019). Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6. *Remote Sensing of Environment*, 229, 32–47. <https://doi.org/10.1016/j.rse.2019.04.021>

Pérez-Ruzafa, A., Campillo, S., Fernández-Palacios, J. M., García-Lacunza, A., García Oliva, M., Ibañez, H., Navarro-Martínez, P. C., Pérez-Marcos, M., Pérez-Ruzafa, I. M., Quispe-Becerra, J. I., Sala-Mirete, A., Sánchez, O., & Marcos, C. (2019). Long-term dynamic in nutrients, chlorophyll a, and water quality parameters in a coastal lagoon during a process of eutrophication for decades, a sudden break and a relatively rapid recovery. *Frontiers in Marine Science*, 6, 26. <https://doi.org/10.3389/fmars.2019.00026>

Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 1(1), 73–79.

Salls, W. B., Schaeffer, B. A., Pahlevan, N., Coffey, M. M., Seegers, B. N., Werdell, P. J., Ferriby, H., Stumpf, R. P., Binding, C. E., & Keith, D. J. (2024). Expanding the application of Sentinel-2 chlorophyll monitoring across United States lakes. *Remote Sensing*, 16(11), 1977. <https://doi.org/10.3390/rs16111977>

Vanhellemont, Q., & Ruddick, K. (2016). ACOLITE for Sentinel-2: Aquatic applications of MSI imagery. In Proceedings of the 2016 ESA Living Planet Symposium (ESA Special Publication SP-740). Prague, Czech Republic, 9–13 May 2016.

White, S., Medina-Lopez, E., Silva, T., Spyarakos, E., Amoudry, L., & Martin, A. (2025). Classification-informed estimation: the role of water-type clustering to improve neural network generalization for salinity and temperature estimation in coastal waters. *Environmental Data Science*, 4, e32. <https://doi.org/10.1017/eds.2025.10005>