

Álvarez Ripado, A.; García Bruzón, A.; Arrogante Funes, F.; Arrogante Funes, P. Predicción espacial de la capacidad ecosistémica urbana a partir de datos de teledetección y aprendizaje automático

# Predicción espacial de la capacidad ecosistémica urbana a partir de datos de teledetección y aprendizaje automático

Álvarez Ripado, Ariadna <sup>1,2</sup> García Bruzón, Adrián <sup>1</sup> Arrogante Funes, Fátima <sup>1</sup>  
Arrogante Funes, Patricia <sup>1</sup>

<sup>1</sup> Universidad Rey Juan Carlos, C/Tulipán s/n, Móstoles, 28933, Madrid (España)

<sup>2</sup> Ecoacsa Reserva de Biodiversidad S.L, C/ Porto Cristo, 1 Esc. izq. 9º B, Alcorcón, 28924, Madrid (España)

ORCID: Álvarez Ripado [0000-0003-2113-2005](https://orcid.org/0000-0003-2113-2005) García Bruzón [0000-0002-5364-6388](https://orcid.org/0000-0002-5364-6388) Arrogante Funes [0000-0003-1724-6173](https://orcid.org/0000-0003-1724-6173) Arrogante Funes [0000-0003-1944-777X](https://orcid.org/0000-0003-1944-777X)

Correspondencia: [ariadna.ripado@urjc.es](mailto:ariadna.ripado@urjc.es) [adrian.bruzon@urjc.es](mailto:adrian.bruzon@urjc.es) [fatima.arrogante@urjc.es](mailto:fatima.arrogante@urjc.es) [patricia.arrogante@urjc.es](mailto:patricia.arrogante@urjc.es)

## RESUMEN

La capacidad de los ecosistemas urbanos para proporcionar servicios esenciales es clave para la sostenibilidad y resiliencia de las ciudades. Este estudio propone un enfoque basado en teledetección y modelización computacional para estimar espacialmente el índice de capacidad ecosistémica urbana en Madrid, integrando la condición ecosistémica y cuatro servicios urbanos: suministro de energía solar, refrigeración urbana, retención de aguas pluviales y acceso a la naturaleza. Se utilizaron principalmente variables derivadas de datos satelitales como predictores y el modelo CatBoost Regressor para estimar el índice con conjuntos de entrenamiento y validación (80/20). Los resultados mostraron alta concordancia entre valores predichos y reales (MAE = 0,0545, RMSE = 0,0852, R<sup>2</sup> = 0,8716), indicando que gran parte de la variabilidad espacial puede explicarse mediante datos de teledetección. La metodología es replicable y escalable, útil para apoyar la planificación urbana y la gestión de servicios ecosistémicos, y futuros desarrollos con datos hiperespectrales, LiDAR y aprendizaje profundo permitirán generar escenarios más realistas y robustos para la planificación sostenible. En conjunto, la integración de teledetección y modelización computacional ofrece un marco sólido para evaluar y anticipar la capacidad ecosistémica urbana, apoyando la toma de decisiones en ciudades resilientes y sostenibles.

**Palabras clave:** Teledetección, ecosistema urbano, capacidad ecosistémica, modelización predictiva, aprendizaje automático


Fecha de recepción: 11 febrero 2026 · Fecha de aceptación: 11 febrero 2026

# Predicción espacial de la capacidad ecosistémica urbana a partir de datos de teledetección y aprendizaje automático


Álvarez Ripado, Ariadna <sup>(1,2)</sup>, García Bruzón, Adrián <sup>(1)</sup>, Arrogante Funes, Fátima <sup>(1)</sup>,  
Arrogante Funes, Patricia <sup>(1)</sup>

<sup>(1)</sup> Universidad Rey Juan Carlos, Madrid (España).

<sup>(2)</sup> Ecoacsa Reserva de Biodiversidad S.L, Madrid (España).

 0000-0003-2113-2005, ariadna.ripado@urjc.es ; ariadna.alvarez@ecoacsa.com..

 0000-0002-5364-6388, adrian.bruzon@urjc.es ;  0000-0003-1724-6173, fatima.arrogante@urjc.es.

 0000-0003-1944-777X, patricia.arrogante@urjc.es.

**Resumen:** La capacidad de los ecosistemas urbanos para proporcionar servicios esenciales es clave para la sostenibilidad y resiliencia de las ciudades. Este estudio propone un enfoque basado en teledetección y modelización computacional para estimar espacialmente el índice de capacidad ecosistémica urbana en Madrid, integrando la condición ecosistémica y cuatro servicios urbanos: suministro de energía solar, refrigeración urbana, retención de aguas pluviales y acceso a la naturaleza. Se utilizaron principalmente variables derivadas de datos satelitales como predictores y el modelo *CatBoost Regressor* para estimar el índice con conjuntos de entrenamiento y validación (80/20). Los resultados mostraron alta concordancia entre valores predichos y reales (MAE = 0,0545, RMSE = 0,0852,  $R^2 = 0,8716$ ), indicando que gran parte de la variabilidad espacial puede explicarse mediante datos de teledetección. La metodología es replicable y escalable, útil para apoyar la planificación urbana y la gestión de servicios ecosistémicos, y futuros desarrollos con datos hiperespectrales, LiDAR y aprendizaje profundo permitirán generar escenarios más realistas y robustos para la planificación sostenible. En conjunto, la integración de teledetección y modelización computacional ofrece un marco sólido para evaluar y anticipar la capacidad ecosistémica urbana, apoyando la toma de decisiones en ciudades resilientes y sostenibles.

**Palabras clave:** Teledetección, ecosistema urbano, capacidad ecosistémica, modelización predictiva, aprendizaje automático

## ***Spatial prediction of urban ecosystem capacity using remote sensing data and machine learning***

**Abstract:** *The capacity of urban ecosystems to provide essential services is key to the sustainability and resilience of cities. This study proposes an approach based on remote sensing and computational modeling to spatially estimate the urban ecosystem capacity index in Madrid, integrating ecosystem condition and four urban services: solar energy supply, urban cooling, stormwater retention, and access to nature. Variables derived mainly from satellite data were used as predictors, and the CatBoost Regressor model was applied to estimate the index using training and validation sets (80/20). The results showed high agreement between predicted and observed values (MAE = 0.0545, RMSE = 0.0852,  $R^2 = 0.8716$ ), indicating that a large part of the spatial variability can be explained by remote sensing data. The methodology is replicable and scalable, making it useful for supporting urban planning and ecosystem service management. Future developments using hyperspectral data, LiDAR, and deep learning will enable the generation of more realistic and robust scenarios for sustainable planning. Overall, the integration of remote sensing and computational modeling provides a solid framework for assessing and anticipating urban ecosystem capacity, supporting decision-making in resilient and sustainable cities.*

**Keywords:** *Remote sensing, urban ecosystem, ecosystem capacity, predictive modelling, machine learning*

## 1. INTRODUCCIÓN

Los ecosistemas urbanos son esenciales para el bienestar humano, la resiliencia climática y la sostenibilidad de las ciudades, especialmente en áreas

densamente pobladas (Fang *et al.*, 2023). La evaluación integrada de la condición y de los servicios ecosistémicos que estos proporcionan resulta clave para apoyar la planificación urbana y la toma de decisiones basada en evidencia científica (Fang *et al.*, 2023).

En este contexto, los datos de teledetección y de Observación de la Tierra (*Earth Observation*, EO) se han consolidado como recursos fundamentales para el análisis de ecosistemas urbanos, al permitir una evaluación espacialmente explícita, consistente y replicable de variables biofísicas a distintas escalas espaciales y temporales (Di Palma *et al.*, 2024). En los últimos años, la integración de datos EO con enfoques de modelización computacional y métodos *data-driven* ha ampliado significativamente las capacidades para analizar y predecir patrones y procesos ecológicos en sistemas urbanos complejos (Moraci *et al.*, 2025).

El concepto de capacidad ecosistémica, formalizado en el Sistema de Contabilidad Ambiental-Económica – Contabilidad de Ecosistemas (SCAE-CE), permite evaluar el potencial de los ecosistemas para suministrar servicios ecosistémicos de manera sostenible, considerando su condición ecológica y las interacciones, sinergias y compensaciones entre distintos servicios (United Nations *et al.*, 2024). Sobre esta base conceptual, se desarrolló un índice de capacidad ecosistémica que integra las cuentas de condición y de servicios ecosistémicos en un marco coherente y compatible con los estándares internacionales de contabilidad ambiental.

A pesar de los avances en la cartografía y evaluación de servicios ecosistémicos urbanos (Di Palma *et al.*, 2024; Finizio *et al.*, 2024), persiste una limitada aplicación de enfoques orientados a la predicción de la capacidad ecosistémica mediante modelización basada en datos de teledetección (Akiner & Ghasri, 2025). En este contexto, el presente estudio tiene como objetivo modelizar y predecir el índice de capacidad ecosistémica urbana utilizando variables mayoritariamente derivadas de sensores satelitales y técnicas de modelización implementadas en Python. El enfoque se aplica a la ciudad de Madrid, explorando el potencial de los métodos *data-driven* basados en EO para apoyar la evaluación, monitorización y gestión de ecosistemas tan complejos como son los urbanos (Akiner & Ghasri, 2025; Moraci *et al.*, 2025).

## 2. MATERIAL Y MÉTODOS

### 2.1. Índice de capacidad ecosistémica

Este estudio se basa en un trabajo previo que desarrolló un índice de capacidad ecosistémica alineado con el marco del SCAE-CE, integrando cuentas de condición y de servicios ecosistémicos urbanos para representar espacialmente el potencial de los ecosistemas urbanos de suministrar servicios de manera sostenible (United Nations *et al.*, 2024).

Para la obtención del índice, se realizó un análisis de componentes principales (PCA) para calcular las sinergias entre cuatro servicios urbanos (suministro de energía solar, refrigeración urbana, retención de aguas

pluviales y acceso a la naturaleza urbana) y, posteriormente, estas sinergias se integraron con la condición mediante un enfoque de lógica difusa en ArcGIS Pro, permitiendo modelar relaciones no lineales relevantes en ecosistemas urbanos (Hanna *et al.*, 2023). Otros métodos, como la lógica clásica, se han empleado para generar índices de capacidad (Martini *et al.*, 2024).

Como predictores de la capacidad se incluyeron datos de Sentinel-2 tales como la permeabilidad del suelo, el Índice de Humedad de Diferencia Normalizada (NDMI), el verdor anual máximo (NDVI máximo), la densidad de cobertura arbórea y, por ende, la sombra. De Landsat 8 se obtuvieron la temperatura, el albedo y la evapotranspiración. La radiación y duración solar, así como la orientación y la pendiente del terreno, se obtuvieron a partir de la combinación de modelos digitales de superficies y datos satelitales. Los contaminantes atmosféricos (NO<sub>2</sub>, O<sub>3</sub>, CO, SO<sub>2</sub>, HCHO) y el vapor de agua se derivaron de Sentinel-5P. Además, se incluyeron variables derivadas del SIOSE AR, que forma parte del Plan Nacional de Observación del Territorio, como el índice de confort acústico, la red de carreteras y la fragmentación del territorio (*Guidos toolbox*) y otras, como la proximidad a zonas riparias, fueron calculadas a partir de “Copernicus VHR\_IMAGE\_2018” y Urban Atlas.

También se incluyeron el carbono orgánico y la densidad aparente del suelo del programa LUCAS (JRC), así como los grupos hidrológicos del suelo, la precipitación media anual, el área de los edificios y varios datos demográficos. Finalmente, se incluyó la riqueza de aves de interés comunitario, calculada a partir de la Directiva de Aves, como modelización de la presencia de estas y de otras variables ambientales y antrópicas. La resolución espacial final de los predictores es de 25 m.

### 2.2. Modelización predictiva

Tras generar la base de datos con todos los predictores, se realizó un análisis exploratorio de datos para limpiar y preparar los datos, identificando patrones, relaciones y posibles errores antes del modelado.

El proceso de modelización se realizó en Python, donde se buscó predecir el índice de capacidad ecosistémica utilizando como predictores las diferentes variables explicativas de las cuentas de condición y servicios de la ciudad de Madrid.

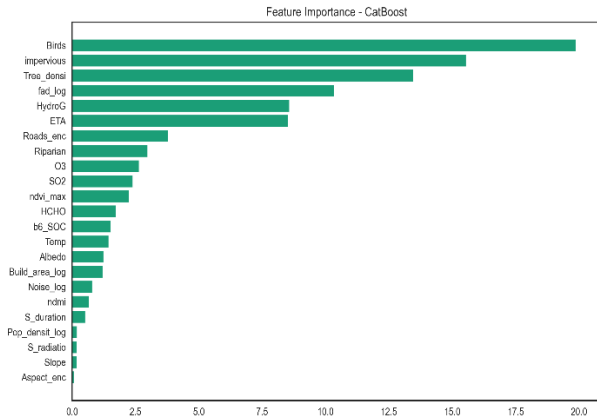
El conjunto de datos final se dividió en un 80% para entrenamiento y test y un 20% para validación final. Para la selección del modelo se utilizó la librería *PyCaret*, que permite evaluar de forma automatizada distintos algoritmos de aprendizaje automático bajo un marco común. El mejor modelo se seleccionó en función de su rendimiento predictivo sobre el conjunto de test.

El modelo seleccionado se utilizó posteriormente para la predicción espacial del índice de capacidad ecosistémica, permitiendo evaluar el potencial de los enfoques *data-driven* basados en teledetección para la estimación espacialmente explícita de la capacidad ecosistémica en entornos urbanos.

### 3. RESULTADOS

Se observa una gran concordancia visual entre la distribución espacial del índice de capacidad ecosistémica real y predicho (Figura 2 y Figura 3), reflejando la capacidad del modelo *CatBoost Regressor* para reproducir patrones espaciales complejos de la capacidad ecosistémica urbana a partir de variables derivadas mayoritariamente de teledetección (Di Palma *et al.*, 2024; Finizio *et al.*, 2024). Destaca especialmente la diferenciación clara entre las zonas con mayor capacidad ecosistémica, ubicadas en áreas periféricas y con mayor cobertura verde, y las zonas urbanas más densas con menor capacidad, situadas principalmente en el centro y sur de la ciudad.

La riqueza de aves, la permeabilidad del suelo, la densidad arbórea y la fragmentación del terreno destacan como las variables más influyentes en la predicción del índice de capacidad ecosistémica, integrando datos EO con variables ecológicas (Figura 1). Esto evidencia el valor de combinar sensores remotos con datos biológicos para evaluar la funcionalidad y el estado de ecosistemas urbanos.



**Figura 1.** Grado de influencia de los predictores empleados en el modelo *CatBoost Regressor*.

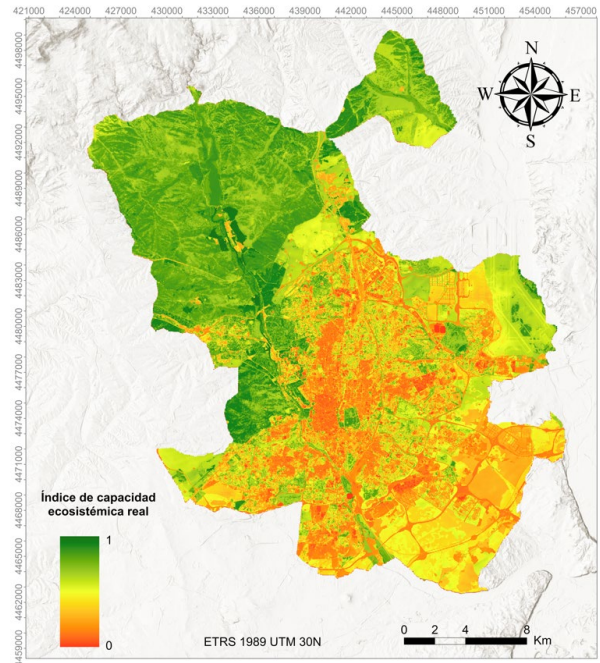
Las métricas obtenidas (Tabla 1) muestran un alto grado de exactitud y capacidad explicativa del modelo ( $R^2 = 0,8716$ ), indicando que la mayoría de la variabilidad espacial se explica con variables satelitales y confirmando la utilidad de los datos EO para modelización urbana (Di Palma *et al.*, 2024). Otras métricas, como RMSLE y MAPE, confirman además que el modelo mantiene un bajo error relativo en el conjunto de validación, lo que resulta especialmente relevante en aplicaciones prácticas de planificación urbana, donde la robustez frente a escalas variables es crítica (Di Palma *et al.*, 2024).

**Tabla 1.** Métricas del *dataset* de validación final del modelo seleccionado *CatBoost Regressor*.

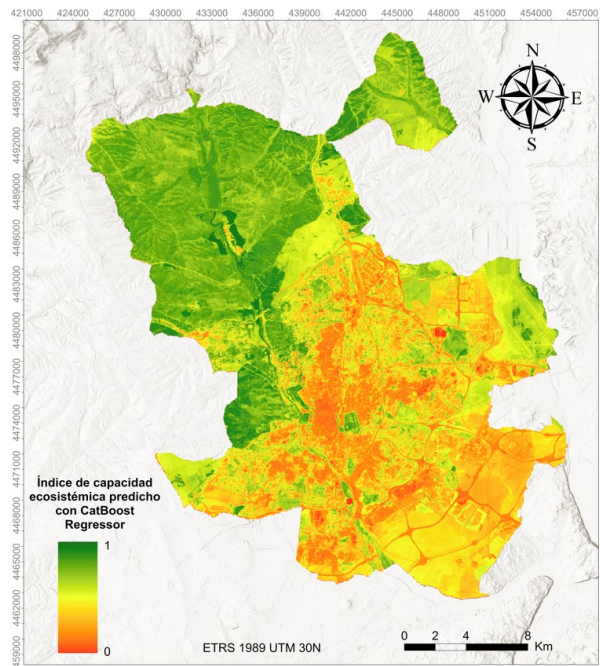
MAE	MSE	RMSE	$R^2$	RMSLE	MAPE
0,0545	0,0073	0,0852	0,8716	0,0615	0,2272

Estos resultados destacan la utilidad de la teledetección no solo para monitorear la condición de los ecosistemas urbanos, sino también para anticipar cambios en áreas urbanas complejas mediante técnicas de aprendizaje

automático (Akiner & Ghasri, 2025; Finizio *et al.*, 2024). La integración de técnicas avanzadas de modelización computacional, como *CatBoost*, con datos EO ofrece una vía prometedora para mejorar la gestión ambiental urbana y la toma de decisiones espaciales basadas en evidencia científica (Fang *et al.*, 2023).



**Figura 2.** Distribución del índice de capacidad ecosistémica real de la ciudad de Madrid.



**Figura 3.** Distribución del índice de capacidad ecosistémica predicho de la ciudad de Madrid.

Finalmente, el estudio reafirma el papel estratégico de la teledetección como herramienta clave para analizar ecosistemas urbanos, al proporcionar datos de alta resolución espacial y temporal que permiten capturar su heterogeneidad espacial y dinámica temporal. Esto es

especialmente relevante en ciudades como Madrid, con alta presión antropogénica y climática, donde marcos como el SCAE-CE apoyan la planificación sostenible y la resiliencia urbana (United Nations *et al.*, 2024).

La metodología desarrollada permite modelizar de forma eficiente y replicable el índice de capacidad ecosistémica a partir de datos satelitales multifuente, proporcionando información espacial útil para la planificación de infraestructuras verdes y servicios ecosistémicos urbanos (Hanna *et al.*, 2024; Di Palma *et al.*, 2024; Finizio *et al.*, 2024). Su combinación con aprendizaje automático facilita la captura de patrones complejos y la predicción de la variabilidad espacial, destacando el valor de integrar datos EO con información biológica, como la riqueza de aves. No obstante, la precisión del modelo depende de la calidad y resolución de los sensores y puede verse limitada por procesos no detectables mediante teledetección (Fang *et al.*, 2023; Martini *et al.*, 2024).

Como línea futura, se plantea aplicar el modelo a análisis contrafactuales para evaluar el impacto de las intervenciones urbanas sobre la capacidad ecosistémica mediante escenarios alternativos. La incorporación de datos hiperespectrales y LiDAR, junto con *deep learning*, permitirá reducir la incertidumbre asociada a variables *proxy* y generar escenarios más realistas, reforzando la planificación de infraestructuras verdes y decisiones alineadas con el SCAE-CE.

#### 4. CONCLUSIONES

El índice de capacidad ecosistémica puede predecirse con exactitud mediante datos EO y modelización computacional en Python, utilizando el algoritmo *CatBoost*. Los resultados muestran una elevada concordancia espacial entre valores predichos y reales, reflejando la heterogeneidad espacial de Madrid y las diferencias en el índice de capacidad ecosistémica.

La metodología es replicable, escalable y permite actualizaciones temporales, siendo una herramienta útil para la planificación de infraestructuras verdes y la gestión de servicios ecosistémicos urbanos. Además, el enfoque basado en datos EO permite un monitoreo continuo de la sostenibilidad, integrando la capacidad de los ecosistemas para mantener sus funciones y servicios frente a cambios ambientales o climáticos.

Entre sus limitaciones se encuentra la dependencia de algunas variables *proxy*, como el porcentaje de cobertura arbórea para representar la provisión de sombra y la densidad de población para reflejar la demanda de servicios ecosistémicos, así como de la resolución espacial y temporal de los datos empleados. Las futuras mejoras incluyen la incorporación de datos de sensores hiperespectrales y LiDAR, junto con técnicas de *deep learning*, para aumentar la precisión y captura de relaciones más complejas.

En resumen, la integración de datos de teledetección y modelización computacional proporciona un marco eficiente y escalable para evaluar y anticipar la capacidad ecosistémica urbana, apoyando la

planificación y gestión ambiental en ciudades densamente urbanizadas.

#### 5. AGRADECIMIENTOS

Esta investigación ha sido cofinanciada por la Comunidad de Madrid en el marco de ayudas para la realización de doctorados industriales 2023 (Ref.: IND2023/AMB-28820).

#### 6. REFERENCIAS

- Akiner, M. E., & Ghasri, M. (2025). Integrating remote sensing and machine learning to analyze urban growth and its environmental effects: A 30-year assessment. *Pure and Applied Geophysics*, 182, 1–20. <https://doi.org/10.1007/s00024-025-03727-w>.
- Di Palma, M., Rigillo, M., & Leone, M. F. (2024). Remote sensing technologies for mapping ecosystem services: An analytical approach for urban green infrastructure. *Sustainability*, 16(14), 6220. <https://doi.org/10.3390/su16146220>.
- Fang, X., Li, J., & Ma, Q. (2023). Integrating green infrastructure, ecosystem services and nature-based solutions for urban sustainability: A comprehensive literature review. *Sustainable Cities and Society*, 98, 104843. <https://doi.org/10.1016/j.scs.2023.104843>.
- Finizio, M., Pontieri, F., Bottaro, C., Di Febraro, M., Innangi, M., Sona, G., & Carranza, M. L. (2024). Remote sensing for urban biodiversity: A review and meta-analysis. *Remote Sensing*, 16(23), 4483. <https://doi.org/10.3390/rs16234483>.
- Hanna, E., Bruno, D. & Comín, F.A. (2024). The ecosystem services supplied by urban green infrastructure depend on their naturalness, functionality and imperviousness. *Urban Ecosyst*, 27, 187–202. <https://doi.org/10.1007/s11252-023-01442-9>.
- Martini, F., Conroy, K., King, E., Farrell, C. A., Kelly-Quinn, M., Obst, C., ... & Stout, J. C. (2024). A capacity index to connect ecosystem condition to ecosystem services accounts. *Ecological Indicators*, 167, 112731. <https://doi.org/10.1016/j.ecolind.2024.112731>.
- Moraci, F., Bevilacqua, C., Pizzimenti, P., Barrile, V., & Genovese, E. (2025). Data-Driven Planning Approaches for Measuring Urban Ecosystem Services: The Potential of Big Data and Urban Informatics for Planning the Transition of Cities. In INTERNATIONAL SYMPOSIUM: New Metropolitan Perspectives (pp. 18-34). Springer, Cham. [https://doi.org/10.1007/978-3-031-82927-7\\_2](https://doi.org/10.1007/978-3-031-82927-7_2).
- United Nations, European Commission, Food and Agriculture Organization of the United Nations, International Monetary Fund, Organisation for Economic Co-operation and Development, & World Bank. (2024). System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA) (ST/ESA/STAT/SER.F/124). United Nations.