





# Scalable approach for high-resolution land cover: a case study in the Mediterranean Basin\*

Antonio Manuel Burgueño<sup>1</sup>, José F. Aldana-Martín<sup>1</sup>, María Vázquez-Pendón<sup>1</sup>, Cristóbal Barba-González<sup>1</sup>, Yaiza Jiménez Gómez<sup>2</sup>, Virginia García Millán<sup>2</sup>, and Ismael Navas-Delgado<sup>1</sup>

<sup>1</sup> ITIS Software, University of Málaga, Málaga, Spain

[ambrbr@uma.es](mailto:ambrbr@uma.es)

<https://itis.uma.es/>

<sup>2</sup> ETC-UMA, University of Málaga, Málaga, Spain

<https://www.etc.uma.es/>

**Keywords:** Big Data, Land cover, Workflow, Explainable AI, Remote Sensing, Multispectral, Machine Learning

**Published in:** Journal of Big Data, Vol. 10, No. 1, pp. 01–22, 2023

**Impact Factor:** JCR 8.1 - Q1 - Position: 8/145 - Area: COMPUTER SCIENCE / THEORY & METHODS

**DOI:** <https://doi.org/10.1186/s40537-023-00770-z>

**Abstract.** The production of land cover maps is an everyday use of image classification applications on remote sensing. However, managing Earth observation satellite data for a large region of interest is challenging in the task of creating land cover maps. Since satellite imagery is getting more precise and extensive, Big Data techniques are becoming essential to handle the rising quantity of data. Furthermore, given the complexity of managing and analysing the data, defining a methodology that reduces the complexity of the process into different smaller steps is vital to data processing. This paper presents a Big Data methodology for creating land cover maps employing artificial intelligence algorithms. Machine Learning algorithms are contemplated for remote sensing and geodata classification, supported by explainable artificial intelligence. Furthermore, the process considers aspects related to downloading data from different satellites, Copernicus and ASTER, executing the pre-processing and processing of the data in a distributed environment, and depicting the visualisation of the result. The methodology is validated in a test case for creating a land cover map of the Mediterranean Basin.

\* This work has been partially funded by grants PID2020-112540RB-C41 funded by MCIN/AEI/10.13039/501100011033, and the e-infrastructure LifeWatch ERIC Project "EnBiC2-Lab", co-funded by the ERDF (Spain's Pluri-regional Operative Programme 2014-2020) through the Spanish Ministry for Research and Innovation. José F. Aldana-Martín is supported by Grant PRE2021-098594 (Spanish Ministry of Science, Innovation and Universities). Funding for open access charge: Universidad de Málaga / CBUA.