Unidad Docente de Geografía
Facultad de Filosofía y Letras
C/ Colegios, 2. 28801.
Alcalá de Henares, Madrid

VII Simposio de Doctorandos de la UAH en Investigación con Tecnologías de la Información Geográfica (SITIG-UAH)

Noviembre de 2025

Título (máximo 30 palabras):
Multi-scale integration of proximal and remote sensing data for the efficient implementation of artificial intelligence algorithms in environmental applications

Ponente (Nombre y apellidos):
Juan Ramón Bustos Caparrós

Resumen (máximo 400 palabras):

Remote sensing (RS) provides essential geospatial information to address global challenges such as climate change and resource scarcity. However, the requirements of high resolutions in different aspects (spatial, temporal, spectral) cannot be covered by single missions. Thus, complex problems must be addressed by integrating data from multiple sources: satellite, aerial (including drones), and *in situ*. Nevertheless, the combination of datasets with different resolutions and extents (e.g., remote and proximal sensing, destructive sampling, or eddy covariance measurements) encompass scale mismatches that compromise model calibration (Gamon et al. 2010). To mitigate these issues, both sampling and modelling strategies must explicitly address scale differences and account for key spatial and temporal dynamics of ecosystem processes (Gamon et al., 2006) through improved integration frameworks. This Ph.D. project aims to develop new methods to integrate multisource *in situ* and remote sensing data across scales improving modelling tools to increase model accuracy and reduce uncertainty. The research objectives will be addressed in three main phases: (1) simulation of virtual scenarios, (2) evaluation with real observations, and (3) development of open-source tools for broader application. Phase 1 relies on the Biodiversity Observing System Simulation Experiment (BOSSE; Pacheco-Labrador et al., 2025). BOSSE generates synthetic *in situ* and RS datasets by simulating multiple scenarios combining environmental conditions, sensor configurations, and data characteristics across spatial, spectral, and temporal resolutions. This allows to develop, benchmark and select optimal data integration methods. Currently, simple regression based on vegetation indices (VIs, NDVI and NIRv) and Partial Least Square Regression (PLSR) based on narrow bands from VIS, NIR and SWIR regions are being used for estimating Leaf Area Index (LAI) and Chlorophyll a and b content (Cab). Preliminary results indicate that PLSR performs better than VIs. Spatial patterns, climate conditions and spatial resolutions influence the model accuracy. In the next phase the methods selected using synthetic data will be validated using real observations from a *dehesa* ecosystem including spectral information at very high (centimetric), medium (10–100 m), and low (100 m–1 km) spatial resolution. The results will be compared against models neglecting data mismatches to quantify improvements in model training. Finally, in phase 3 AI algorithms will be implemented as open-source tools and integrated into geospatial catalogues (e.g., STAC) facilitating the application of the new integration algorithms.