



Machine Learning techniques for the detection of geomorphological features in badland landscapes

Angelo Sozio¹, Antonella Marsico¹, Rosa Colacicco¹, Marco La Salandra¹, Sandro Muscillo¹, Alberto Refice², and Domenico Capolongo^{1,2}

¹University of Bari Aldo Moro, Department of Earth and Geo-environmental Sciences, Bari, 70125, Italy

(a.sozio3@phd.uniba.it)

²IREA, National Research Council, Bari, 70125, Italy

Badland landscapes, characterised by intensely dissected slopes on unconsolidated sediments or soft rocks, are crucial hotspots for understanding soil erosion and sediment transport dynamics. Consequently, alongside their rapid climatic responses and links to anthropic land use, the geomorphological processes driving badlands morphogenesis are widely studied; recently, approaches combining multiplatform remote sensing and Machine Learning (ML) have been proposed due to their superior performance compared to other statistical models.

This study develops an integrated and multi-source approach using detailed geomorphic and hydrological parameters through a Random Forest (RF) algorithm to obtain a high-resolution land cover classification and erosion susceptibility maps of badland landscape in Basilicata Region (Italy). The workflow analyses geomorphological data at the micro-topography scale (3 cm/px) for geomorphometric landscape classification. Topographic and hydrological predictors were extracted from high-resolution Digital Elevation Models (DEMs) and orthomosaics were derived from optical images acquired in a drone survey conducted in May 2025. Spanning 0.025 km², the study area exhibits characteristic badland morphologies, located on poorly cemented silty clays from the Lower Pleistocene. Two different experiments were conducted. In the first one, ten topographic and hydrological predictors (e.g. Topographic Position Index, aspect, profile and tangent curvatures, Stream Power Index) were computed using open-source GDAL and GRASS GIS tools to assemble a multi-layer spatial dataset. In the second experiment, the R, G and B bands from the optical orthomosaics were also considered and included as three additional predictors. In both the experiments, 9,900 training points and 3,000 test points were extracted from the dataset to conduct a spatial cross-validation. Following the accuracy assessment, the algorithm was retrained on the full dataset to generate: **i**) a land cover map of three features: 'Badland', 'Vegetation' and 'Pediment'; and **ii**) an erosion susceptibility map based on the probability of a pixel belonging to the 'Badland' class. The first experiment using only morphometric predictors showed a global accuracy of 82.49%, while the second experiment integrating the three RGB bands increased accuracy to 97.43%.