GT 14 - 39 - O

SPATIOTEMPORAL PREDICTION OF SOIL ORGANIC CARBON DENSITY FOR EUROPE (2000--2022) BASED ON LANDSAT-BASED SPECTRAL INDICES TIME-SERIES

GT 14. SOIL DATA ACQUISITION, CURATION, SHARING & MODELLING / GT 10. SOIL CARBON DYNAMICS AND STABILIZATION

XUEMENG TIAN*[1,2], SYTZE DE BRUIN[2], ROLF SIMOES[1], MUSTAFA SERKAN ISIK[1], ROBERT MINARIK[1], YU-FENG HO[1], MURAT SAHIN[3], MARTIN HEROLD[4,2], DAVIDE CONSOLI[1], TOM HENGL[1]

[1.] *OpenGeoHub, Doorwerth, The Netherlands*

[2.] *Wageningen University and Research, Wageningen, The Netherlands*

[3.] *Delft University of Technology, Delft, The Netherlands*

[4.] *Helmholtz GFZ German Research Centre for Geosciences, Potsdam, Germany*

* *xuemeng.tian@opengeohub.org*

The paper describes a comprehensive framework for soil organic carbon density (SOCD) (kg/m3) modeling and mapping, based on spatiotemporal Random Forest (RF) and Quantile Regression Forests (QRF). A total of 45,616 SOCD measurements and various feature layers, particularly 30m Landsat-based spectral indices, were used to produce 30m SOCD maps for the EU at four-year intervals (2000--2022) and four soil depth intervals (0--20cm, 20--50cm, 50--100cm, and 100--200cm). Per-pixel 95% probability prediction intervals (PIs) and extrapolation probabilities are also provided. Model evaluation indicates consistent accuracy, with R2 between 0.53--0.67 and CCC 0.68--0.80 across cross-validations and independent tests. Prediction accuracy varies by land cover, depth interval and year of prediction with accuracy the worst for shrubland and deeper soils 100--200cm. PI validation confirmed effective uncertainty estimation, though with reduced accuracy for higher SOCD values. Shapley analysis identified soil depth as the most influential feature, followed by vegetation, long-term bioclimate, and topographic features. While pixel-level uncertainty is substantial, spatial aggregation reduces uncertainty by approximately 66%. Detecting SOCD changes remains challenging but offers a baseline for future improvements. Maps, based primarily on topsoil data from cropland, grassland, and woodland, are best suited for applications related to these land covers and depths. Users should interpret the maps with local knowledge and consider the uncertainty and extrapolation probability layers. All data and code are available under an open license at https://doi.org/10.5281/zenodo.13754343 and https://github.com/AI4SoilHealth/SoilHealthDataCube/.