



## **Using LUCAS topsoil database to estimate soil organic carbon content in local spectral libraries**

Fabio Castaldi (1), Bas van Wesemael (1), Sabine Chabrilat (2), and Caroline Chartin (1)

(1) Georges Lemaître Centre for Earth and Climate, Earth and Life Institute, Université Catholique de Louvain, (2) Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum GFZ

The quantification of the soil organic carbon (SOC) content over large areas is mandatory to obtain accurate soil characterization and classification, which can improve site specific management at local or regional scale exploiting the strong relationship between SOC and crop growth. The estimation of the SOC is not only important for agricultural purposes: in recent years, the increasing attention towards global warming highlighted the crucial role of the soil in the global carbon cycle. In this context, soil spectroscopy is a well consolidated and widespread method to estimate soil variables exploiting the interaction between chromophores and electromagnetic radiation. The importance of spectroscopy in soil science is reflected by the increasing number of large soil spectral libraries collected in the world. These large libraries contain soil samples derived from a consistent number of pedological regions and thus from different parent material and soil types; this heterogeneity entails, in turn, a large variability in terms of mineralogical and organic composition. In the light of the huge variability of the spectral responses to SOC content and composition, a rigorous classification process is necessary to subset large spectral libraries and to avoid the calibration of global models failing to predict local variation in SOC content. In this regard, this study proposes a method to subset the European LUCAS topsoil database into soil classes using a clustering analysis based on a large number of soil properties. The LUCAS database was chosen to apply a standardized multivariate calibration approach valid for large areas without the need for extensive field and laboratory work for calibration of local models. Seven soil classes were detected by the clustering analyses and the samples belonging to each class were used to calibrate specific partial least square regression (PLSR) models to estimate SOC content of three local libraries collected in Belgium (Loam belt and Wallonia) and Luxembourg. The three local libraries only consist of spectral data (199 samples) acquired using the same protocol as the one used for the LUCAS database. SOC was estimated with a good accuracy both within each local library (RMSE: 1.2 ÷ 5.4 g kg<sup>-1</sup>; RPD: 1.41 ÷ 2.06) and for the samples of the three libraries together (RMSE: 3.9 g kg<sup>-1</sup>; RPD: 2.47). The proposed approach could allow to estimate SOC everywhere in Europe only collecting spectra, without the need for chemical laboratory analyses, exploiting the potentiality of the LUCAS database and specific PLSR models.