EGU22-9669
https://doi.org/10.5194/egusphere-egu22-9669
EGU General Assembly 2022

A comparative study of the Hungarian Soil Monitoring System and LUCAS Topsoil dataset and their countrywide spatial predictivity

László Pásztor, Zsófia Bakacsi, Gábor Szatmári, Piroska Kassai, Brigitta Szabó, Annamária Laborczi, Mihály Kocsis, and András Benő
Institute for Soil Sciences, Centre for Agricultural Research, Department of Soil Mapping and Environmental Informatics, Budapest, Hungary (pasztor@rissac.hu)

EJP SOIL, which is a European Joint Programme Cofund on Agricultural Soil Management, has posed questions concerning the application of various soil observation datasets to account, monitor and map agricultural soil carbon, fertility and degradation. One of the most exciting issue is, how the continental LUCAS Topsoil dataset and national soil observation/monitoring databases could be harmonized and/or should complement or improve each other to produce Europe wide spatial soil information to support European contribution towards international reporting on soils, and the accuracy of European agri-environmental policies. Methods are being searched by a broad international team. In our paper we present a countrywide case study for the comparison of (i) the representativity of the Hungarian Soil Information and Monitoring System (SIMS) versus LUCAS Topsoil dataset and (ii) some map products, which were modelled by the usage of the two reference data sources.

The difference between national monitoring systems and LUCAS Topsoil dataset is mainly due to (i) the different measurement methods applied to determine soil properties, and (ii) the different sampling strategy, both in terms of sampling location and sampling depth. While transformation (using unit conversions, mass-preserving splines to derive soil properties for similar soil depth, pedotransfer functions for methods conversion, etc.) of the SIMS soil data into the units, methods and soil depth used in the LUCAS dataset could be carried out more or less straightforwardly, the spatial representativity, which strongly affects the performance of any digital soil map based on the given observation dataset, is a challenging feature, which could and should be checked.

From a statistical point of view, a sample is said to be representative if it reflects the characteristics of the population the best. First we analysed whether the two sets of soil data represent the same population applying two approaches: by the comparison of (i) empirical cumulative distribution function and (ii) the mean values aggregated for different land use categories of the selected soil properties coming from the two system. Although based on the Kolmogorov-Smirnov test we could conclude that the Hungarian subset of the LUCAS Topsoil dataset and the data of the Hungarian Soil Information and Monitoring System come from the same population in the case of particle size distribution, pH, organic carbon and carbonate content, the land use based comparison did not give satisfying results.

In the next round the countrywide spatial predictivity of the two datasets have been tested. Primary soil property map pairs have been compiled using the same ancillary datasets and digital mapping methods but the two different observation datasets. The two map products for the same property have been compared by both global measures and cell-to-cell statistics. In addition to pairwise comparison of basic statistical features (histograms, scatter plots), we have examined the spatial distribution of the differences. In our presentation our findings and experiences will be discussed.

Acknowledgment: Our research was supported by the Hungarian National Research, Development and Innovation Office (NKFIH; K-131820), European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 862695.