Mapping of topsoil organic carbon in agro-ecosystems of a flat terrain area (Lombardy) by means of legacy soil data, climatic data and NDVI time series predictors with machine learning methods

Calogero Schillaci (1,2), Sergio Saia (3), Andreas Braun (2), and Marco Acutis (1)

(1) University of Milan, Agricultural and Environmental Sciences Production, Landscape, Agroenergy Department, MILANO, Italy (calogero.schillaci@unimi.it), (2) Department of Geosciences, University of Tübingen, Germany, (3) Council for Agricultural Research and Economics (CREA), Cereal Research Centre (CREA-CER), Foggia, Italy

Topsoil organic carbon plays an important role in the agricultural yield, yield potential, and to deliver many ecosystem services, such as the potential to reduce greenhouse gas (GHG) emission from soil. In particular, SOC content sturdily affects soil properties, thus the precision of its estimation can support broad-scale agricultural and environmental management policy. Soils in temperate agro-ecosystem are generally highly productive and cropland occupies about 60% of their surface (Ramankutty et al 2008). In such contexts, lands is frequently subjected to SOC degrading operations, mostly ploughing, with drawbacks on soil fertility and erosion. In temperate agro-ecosystems, a strong role in SOC maintenance can be played by manure and residues inputs after husbandry and related activities and return of plant biomass to the soil (Acutis et al 2014). In this perspective, soil management can have a major role in SOC spatial distribution to maintain soil fertility and ecosystem services in a target area. Due to the considerable importance of SOC on both agronomical and ecological aspects of the agro-ecosystems, regional soil surveys over the years frequently take into account the measurement of SOC concentration and often stock. In the present study, we integrated a highly detailed legacy SOC dataset with climatic data and RS data to produce a reliable SOC maps from a farm to a district scale. In particular, the Normalized Difference Vegetation Index (NDVI)was used after the computation of its average value in a given pixel derived from several approximately cloud-free images. The input dataset was made of about 3000 Ap horizons implemented of SOC concentration, texture, bulk density and metadata. Climatic data (Worldclim), soil type (from the pedological map 1:250000 WRB), and a time series NDVI were applied. The NDVI data were derived from a set of Landsat 5 scenes (path 193, row 28,29) whereas the path 194, (row 28 and 29) contributes for less than one fourth of the study area. The use of machine learning approach for the generation of a SOC map of the flat terrain agricultural topsoil of Lombardy using the regional soil database relies on two assumptions: (1) the slow change in the content of the stabilised soil organic matter (SOM) fraction, which is almost everywhere the most represented SOM fraction; and (2) the intrinsic low erosion rates due to the low mean slope. In particular, NDVI, which is related land cover and to the amount of biomass returned the soil, can have drawbacks when applied in cultivated fields. These drawbacks mainly concern the variability on crop biomass within and across the year. Notwithstanding, this issue makes NDVI very suitable for differentiating contrasting land use (e.g. field crops vs. orchards) when computed from images captured in particular crop cycle moments (e.g. in summer). However, the same issue reduces NDVI suitability to estimate the amount of biomass within each land use or when aiming at highly detailed resolution.

Different grade of cloud cover were admitted to construct the average NDVI. Boosted regression trees were used to reveal the effect of each spatial covariate in predicting the SOC content. Preliminary results highlighted that the integration of the soil pedological classification and the mean NDVI improved the pixel classification in SOC classes according to crop type and management. As expected, climatic gradient played an important role in SOC modelling but did not affect the spatial resolution of the final map. In conclusion, SOC estimate strongly depends on sample density and homogeneity of distribution and the environmental heterogeneity. The lack of the strong topographical traits in flat terrain areas represents a challenge for soil mapping. In such conditions, the computation of a reliable biomass-related RS trait such as the mean NDVI can increase the prediction ability of the models and reduce the mapping biases.

References
Acutis, M., Alfieri, L., Giussani, A., Provolo, G., Di Guardo, A., Colombini, S., Bertoncini, G.,Castelnuovo, M., Sali, G., Moschini, M., Sanna, M., Perego, A., Carozzi, M., Chiodini, M.E., Fumagalli, M., 2014. ValorE: An integrated and GIS-based decision support system for livestock manure management in the Lombardy region

(northern Italy). Land use policy 41, 149–162. doi:10.1016/j.landusepol.2014.05.007

Ramankutty, N., A. T. Evan, C. Monfreda, and J. A. Foley (2008), Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, Global Biogeochem. Cycles , 22, GB1003, doi:10.1029/2007GB002952.