Geophysical Research Abstracts
Vol. 13, EGU2011-1020, 2011
EGU General Assembly 2011

Boosted regression trees models for estimating soil organic carbon stocks at the national scale

Manuel Martin (1), Martin Wattenbach (2), Pete Smith (3), Jeroen Meersmans (1), Claudy Jolivet (1), Line Boulonne (1), and Dominique Arrouays (1)

(1) INRA Orléans, InfoSol Unit, US 1106, CS 40001, Ardon, 45075, Orléans cedex 2, France, (2) Freie Universität Berlin, Institute of Meteorology, Carl-Heinrich-Becker-Weg 6–10, 12165 Berlin, Germany, (3) Institute of Biological & Environmental Sciences, University of Aberdeen, Cruickshank Building, St. Machar Drive, Aberdeen, AB24 3UU Scotland, UK

Soil organic carbon (SOC) is a key component of soil systems because it strongly influences other soil properties such as soil fertility or soil physical properties. At a larger scale, it participates in the global green house gas budget because soils may be a sink, or a source for gases such as CO2. The spatial distributions of SOC stocks are often characterized by great variability, since SOC concentrations in soils are a function of many environmental factors, which, alone or in interaction with other factors, influence the soil organic matter mineralization or carbon inputs into soils. However, being able to produce valid estimates of SOC spatial distribution is of major importance. These estimates can be used for initializing spatially distributed simulations of SOC stocks changes, which in turn are used for carbon accounting purposes for instance. It can be argued that minimizing the error onto the initial SOC distribution may increase the validity of modeled SOC dynamics.

Here we present a refined estimate of SOC stocks for the French territory. It results from the processing of data obtained through the first campaign of the French soil monitoring network (RMQS) using boosted regression tree (BRT) models. These models offered several advantages such as the handling of missing data, correlated predictors and the robustness to the presence of outliers within the dataset. Moreover, it enables the modeling of interactions between predictors with a varying degree of complexity. Several models were fitted, making use of pedologic (e.g. rock fragments, pH, clay content), climatic (rainfall, temperature, potential evapo-transpiration), biological (net primary productivity) or land use predictors. The calibrated models were evaluated through cross-validation. One model was eventually used for estimating SOC stocks for the whole of metropolitan France. Two other models were calibrated on forest and agricultural soils separately, in order to assess more precisely the influence of pedo-climatic variables on soil organic carbon for such land uses.

The boosted regression tree model showed a good predictive ability, and enabled quantification of relationships between SOC stocks and pedo-climatic variables over the French territory. These relationship strongly depended on the land use, and more specifically differed between forest soils and cultivated soil. The spatial estimate of SOC stocks yielded 3.260 ± 0.872 PgC for the first 30 cm. The later is remarkably lower as compared to another estimate, based on the previously published European soil organic carbon and bulk density maps (i.e. 5.303 PgC). We demonstrated that the estimate presented here might better represent the actual SOC stocks distributions of France, and consequently that the previously published approach at the European level greatly overestimates SOC stocks.