GLOBAL SYMPOSIUM ON SOIL ORGANIC CARBON, Rome, Italy, 21-23 March 2017

Update on Canada's contribution to the Global Soil Organic Carbon map

Chuck Bulmer[2], Jean-Daniel Sylvain[4], Julien Beguin[5], Brandon Heung[7], Robert MacMillan[8], Glenn Lelyk[1] Asim Biswas[6], Angela Bedard-Haughn[3], Margaret Schmidt[7], David Paré[5], Brian Titus[5], Bert VandenBygaart[1] and Scott Smith[1]*

[1]Science and Technology Branch, Agriculture and Agri-Food Canada, Ottawa, bert.vandenbygaart@agr.gc.ca
[2]BC Ministry of Forests, Lands and Natural Resource Operations
[3]Dept of Soil Science, University of Saskatchewan
[4] Ministère des Forêts, de la Faune et des Parcs Quebec
[5]Canadian Forest Service, Natural Resources Canada
[6]School of Environmental Science, University of Guelph
[7]Dept of Geography, Simon Fraser University
[8] LandMapper Environmental Solutions

Abstract

The Scientific and Technical Committee of the Canadian Digital Soil Mapping Working Group has taken on the task of producing the soil organic carbon map for Canada following the specifications outlined by the Global Soil Partnership. The working group is a collaborative of soil scientists from multiple government agencies and universities from across Canada with an interest in improving digital information in the realm of soil science, soil survey and interpretation. The concept is to use a bottom-up approach whereby digital soil maps with FAO specified soil organic carbon attributes are created by regional contributors using a range of methods. We will compile these regional contributions using an ensemble forecast to produce national soil map(s). Our work will use, as a base layer, soil organic carbon values derived through polygon averaging and disaggregation of legacy soil survey maps. This base layer will be upgraded through the integration of information from new maps produced through inference models calibrated from point datasets at regional scales. We plan to use model averaging to combine the attribute data derived from legacy soil maps and from our point datasets to produce our final product.

Keywords: soil carbon, digital soil mapping, Canada, disaggregation, model averaging

Introduction, scope and main objectives

The Scientific and Technical Committee of the Canadian Digital Soil Mapping Working Group (CDSM-WG) has taken on the task of producing the soil organic carbon map for Canada following the specifications outlined by the Global Soil Partnership. The working group is a collaborative formed within the Canadian Society of Soil Science and is composed of members from multiple government agencies and universities from across Canada. The CDSM-WG is a community of practice, where individual scientists with shared interests work together to advance common goals in the area of digital soil mapping. In this particular case, the focus is on the production of soil organic carbon maps using a range of innovative approaches including digital terrain analysis, machine learning, and knowledge extraction from legacy soil survey maps. We will utilize an existing soil organic carbon map for Canada (Tarnocai and Lacelle 1996) derived from the Soil Landscapes of Canada map series at 1:1,000,000 scale to serve as the basis four base layer for a global predictor for our soil carbon.

Methods

Our methods are currently being developed since the project is in its early stages. The concept for our new mapping is to use a bottom-up approach wherein digital soil maps and soil organic carbon attributes are created by regional contributors using a range of methods (Fig 1). In some cases, collaborators have been, or will be, working on a provincial basis, as is the case in British Columbia (Bulmer et al. 2014), Alberta, Saskatchewan, and Quebec or on an ecological region of the country such as the Boreal Forest (Mansuy et al. 2014) .

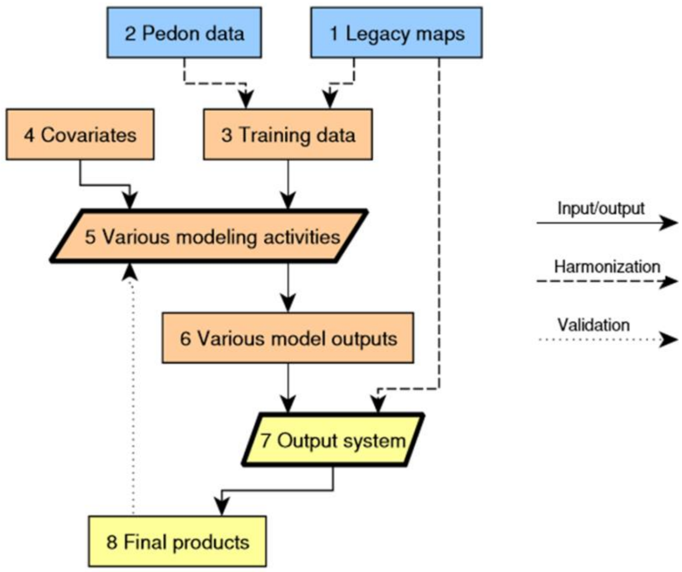

Fig. 1: A conceptual model for collaboration in map production was distributed in a flow chart format and reflected much of the discussion about how the Working Group expects to collaborate.

All of the regional contributions will require considerable effort to compile and harmonize available point profile (pedon) data in order to compute estimates of soil carbon stock individual soil types. It will also be necessary to compute a full suite of gridded covariates (DEM derivatives, climate, vegetation, parent material, remotely sensed imagery, etc) along with mask files of wetlands, urban areas, water bodies, snow, ice, and rock. The tools for making statistical predictions include neural network, support vector regression, Random Forest, Bayesian approaches and regression kriging. We will test the DSMART tool (Odgers et al. 2014) to disaggregate soil survey maps where that approach is required.

The following are some details of the proposed regional/provincial approaches.

British Columbia

Two approaches will be followed in the province of British Columbia (BC). The first will be based on point data (>10,000 profiles) extracted from several sources, primarily from the BC Soil Information System, that will provide reasonable coverage of the province. Compiling, editing, and formatting of these pedon datasets represents a significant undertaking but one that will facilitate on-going digital soil mapping efforts in the province. These point data together with covariates from several provincial sources will form the basis for statistically-driven predictions of soil carbon. A second approach will be to disaggregate existing legacy soil maps, principally the Soil Landscapes of Canada map for BC at 1:1,000,000 scale. Attributes for map unit components will be extracted from the National Soil Database. Some components will require new data to be assigned in order to calculate the required soil organic carbon stocks for the 1 km spatial resolution.

Alberta

Similar to BC, two approaches will be followed for the province of Alberta. The first task will be to extend and modify the existing 1: 100,000 scale regional soil mapping for the province and then use data held in provincial and national soil databases to calculate carbon stock values for each named soil component within map units. These values will be combined and summarized using polygon averaging. Where feasible, map unit polygons may be disaggregated by allocating individual soil components to different landform positions within each polygon. A final map would be produced by a vector to raster conversion of soil carbon stock estimates at 100 m grid size and then generalized to report carbon stock at the 1 km spatial resolution.

A second approach based on statistical modeling will be employed using all available point data for the province. Point data and covariates will be made freely available, and volunteered efforts will be solicited to create and apply DSM models to point data and covariates to predict soil organic carbon stock at 100 m for all of Alberta. These results will then be generalized and assembled at a 1 km spatial resolution for Canada.

Quebec

The approach is to make predictions entirely from point data. For the agricultural region, point data from the Canadian Soil Information System will form the basis to train the prediction model(s) used. For the forested region of Quebec, point data held by various forestry agencies from some 9,600 inventory plots will be used, along with other point data. In forested areas, national and provincial soil databases will be used to define relationships between predicted carbon stock values and soil attributes of texture, color, depth and pH. A hydrologically corrected DEM based on the digital surface model of SRTM at a 40 m resolution will form the spatial base for the modeling approaches described previously. The final product will be generalized to 1 km spatial resolution.

Boreal Forest region

The Boreal Forest region of Canada extends across the country at latitudes between 48°N in eastern Canada up to 65°N in western Canada (Fig 2). For this region, predictions will be generated from point data contained within the National Soil Carbon database and the National Forest Inventory database together providing around 3,000 high quality pedon records that contain SOC values. This work will be conducted at 250 m resolution, using a range of national scale covariates and DEM where the final product will be generalized to the 1 km spatial resolution.

Fig2. Highlighted area on the map of North America shows the extent of the Boreal Forest region in Canada.

Aggregating the inputs

We envisage an ensemble approach for aggregating the contributions received from the regional projects to compute final results for the national map. While the specific approach for aggregation is still being developed, the core concept of our methodology will be to first perform an objective assessment of model errors. Following this, either a simple weighting of the contributions or a more sophisticated approach will be used to combine all predictions into a single value with corresponding error estimates for each grid cell in the final map This 'output system' will be developed in such a way that it not only leads to the best prediction for any particular location, but it will also provide an important and objective source of feedback for the regional contributors. This provides the means to continuously improve estimates of carbon as new information comes available in the future.

Validation

The validation process for predicted map values will be done separately at regional and national levels. The accuracy for regional contributions will be reported using four metrics; determination coefficient (R2), mean error (ME), root mean squared error (RMSE) and bias. Two approaches will be used to report accuracy; k-fold cross-validation and/or bootstrapping procedures. The working group will provide a test dataset that will be used to provide an independent assessment of the errors for each regional contribution. This will also be useful to compare the applicability of individual covariates, inference models and maps.

Discussion

In regions where soil survey map disaggregation or point-based predictions are not possible, we will have to rely on gridded datasets that exist for the arctic and sub-arctic regions of Canada ((Hulervig et al. 2014) and for much of the agricultural region of the country (Hempel et al. 2012). These have been derived through weighted averaging of component attribute values from legacy soil survey map polygons. While useful to fill gaps in our coverage, these reported values are not spatially explicit (many adjacent cells have the same value) nor do they currently include any measures of uncertainty. Compiling a comprehensive gridded soil organic carbon map for all of Canada with a land area greater than 9 million km^2 spanning climatic regimes from temperature to polar will present many challenges. In particular, the dynamic 'forest floor' layers of decomposed and semi-decomposed organic matter (i.e. the F and H layers) that overlie the mineral soil material in our forests are an important carbon pool in our forests that can respond quickly to impacts like harvesting or fire and therefore it is essential that global mapping approaches take account of it effectively. The extensive peatlands and permafrost-affected soils in Canada's north contain very large carbon stocks and unusual organic carbon depth profiles. Finally, in agricultural regions, changing land management practices and soil redistribution have altered soil organic carbon concentrations in the upper soil horizons of cultivated soils in the time since the legacy soil surveys were conducted many decades ago. Correctly quantifying the carbon stocks in all of these different soil environments will be critical to generating enhanced, scientifically-credible soil carbon maps.

References

Bulmer, C., M.G. Schmidt, B. Heung, C. Scarpone, J. Zhang, D. Filatow, M. Finvers, S. Berch, and S. Smith. 2016. Improved soil mapping in British Columbia with legacy soil data and random forest. Ch 24 **in** G. Zhang et al. (eds.) Digital Soil Mapping Across Paradigms, Scales and Boundaries. Springer, Singapore.

Hempel, J. W., Z. Libohova, N. P. Odgers, J.A. Thompson, C.A.S. Smith and G.L. Lelyk. 2012. Versioning of GlobalSoilMap.net raster property maps for the North America Node. Pages 429-434 **in** B. Minasny, B.P. Malone and A.B. McBratney (eds). Digital Soil Assessments and Beyond. CRC Press, Leiden, The Netherlands. 466pp.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., and Swanson, D. K. 2013. The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, Earth Syst. Sci. Data, 5, 3-13, doi:10.5194/essd-5-3-2013.

Mansuy, N., Thiffault, E., Paré, D., Bernier, P., Guindon L., Villemaire, P., Poirier, V. and Beaudoin, A. 2014. Digital mapping of soil properties in Canadian managed forests at 250 m of resolution using the k-nearest neighbor method. Geoderma 235–236: 59–73.

Malone, B.P., Minasny, B., Odgers, N. P.,McBratney, A. B. 2014. Using model averaging to combine soil property rasters from legacy soil maps and from point data. Geoderma 232–234: 34–44.

Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., Clifford, D., 2014. Disaggregating and harmonising soil map units through resampled classification trees. Geoderma 214, 91–100.

Tarnocai, C. and Lacelle, B. 1996. Soil organic carbon map of Canada, Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Research Branch, Ottawa, Ontario, Canada.