A first survey of glomalin-related soil protein using digital soil mapping in France.

Sangchao Chen, Wartini. Ng, Dominique. Arrouays, Siobhan Staunton, Gaoussou Cissé, Hervé Quiquampoix, Buduman Minasny, Nicholas P.A. Saby

A first survey of glomalin-related soil protein using digital soil mapping in France

Songchao Chen[1,2], Wartini Ng[4], Dominique Arrouays[1], Siobhan Staunton[3], Gaoussou Cissé[3], Hervé Quiquampoix[3], Budiman Minasny[4], Nicolas P.A. Saby[1]

[1]Unité InfoSol, INRA, France; [2]UMR SAS, INRA, Agrocampus Ouest, France; [3] Eco&Sols, INRA, IRD, Cirad, SupAgro, University of Montpellier, Montpellier, France; [4]Sydney Institute of Agriculture, the University of Sydney, Sydney, Australia

2 - 6th June 2019
Guelph, Canada

1. Introduction and objectives

- Organic matter plays essential roles in soil. Total soil carbon content alone is insufficient to predict the dynamics of organic matter.
- An operationally defined fraction of soil organic matter known as **glomalin** or glomalin-related soil protein (GRSP) which is obtained by autoclaving soil in citrate solution, is a very stable compound and responsible for enhanced soil physical stability.
- The use of infrared spectroscopy for soil analysis has been thriving over the past decade (Bellon-Maurel and McBratney, 2011). However, authors showed that these soil spectral inference studies are largely dedicated to predicting soil properties for point locations. Soil spectral inference studies could be further expanded into a spatial context for soil mapping purposes.
- We explored the feasibility of using a combination of laboratory measured, and mid infrared (MIR) spectra estimated GRSP data for predicting GRSP content across the French national territory using an extensive dataset.

2. Materials and methods

2.1. Soil data

- The soil data were obtained from the French Soil Monitoring Network (RMQS) between 2001 and 2009. The RMQS is based on a 16 km × 16 km square grid (which results in around 2200 sites in mainland France)
- On basis of an unaligned sampling design with a 20 m × 20 m square, 25 individual core samples were collected from topsoil (0-30 cm) and subsoil (30-50 cm) by a hand auger. In this study, we only considered topsoil.
- Total GRSP was obtained by a single cycle of autoclaving of soil in 50 mM sodium citrate at pH 8.
- The GRSP content across the whole territory was predicted using the MIR spectra (Fig.1).

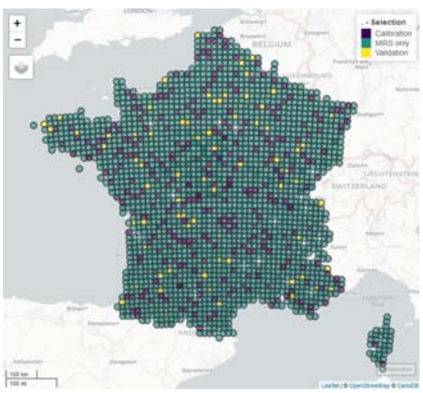

Fig.1 Study area and sampling sites for glomalin survey across the French territory

2.2. General workflow

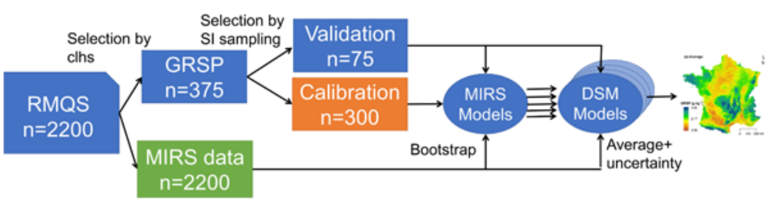

Fig.2 General workflow of the process in predicting GRSP content across French

2.3. MIRS model

GRSP prediction is predicted using Cubist model with 50 bootstrap based on the calibration samples. Each bootstrap works independently. The predictions of the GRSP content across the mainland France were made based on each estimates of the bootstrap model.

2.4. Digital soil mapping

Digital soil mapping (DSM) is based on the framework of *scorpan* - SSPFe (soil spatial prediction function with spatially autocorrelated errors, McBratney *et al.*, 2003). This method fits quantitative relationships between soil properties or classes and seven *scorpan* factors, which can be written as:

S = f(s, c, o, r, p, a, n) + e

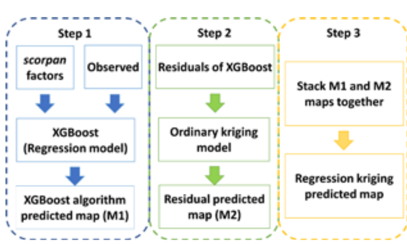

Fig.3 Workflow of the GRSP mapping

where *S* is soil classes or soil properties. The *s* refers to soil information either from prior maps, or from remote or proximal sensing data. The *c* is climatic properties of the environment at a point.
The *o* is organisms including vegetation or fauna. The *r* refers to relief. The *p* is parent material or lithology. The *a* is age, which is regarded as time factor. The *n* refers to space or spatial position. The *e* is spatially correlated residual.

2.4. Digital soil mapping (continued)

The *scorpan* factors include aspect, compound topographic index, curvature, elevation, exposition, roughness, topographic wetness index, erosion rates, soil type (WRB), parent material, net primary productivity, mean annual temperature and mean annual precipitation, land use/land cover and climatic zone.

We applied the Regression Kriging (RK) approach. It consists of three steps: 1) fit a regression model using XGBoost algorithm; 2) fit a Ordinary Kriging model on the regression residuals; 3) sum up the prediction of step 1 and 2 to get the final output.

Thus, we obtained 50 RK models for each bootstrap, which resulted in 50 GRSP maps. We used the average GRSP map as the final output. The 90% prediction intervals were determined by the formula below:

$$90\%\ \mathrm{PIs} = \mathrm{GRSP}_{\mathrm{average}} \pm 1.645 \times \sqrt{\sigma^2 + \mathrm{MSE}}$$

where $\mathrm{GRSP}_{\mathrm{average}}$ is the average GRSP prediction in 50 bootstrap maps, σ is the standard deviation of GRSP prediction in 50 bootstrap maps, and MSE is the averaged mean square error of GRSP prediction in 50 MIR models using validation dataset.

3. Results

3.1. Models' performance

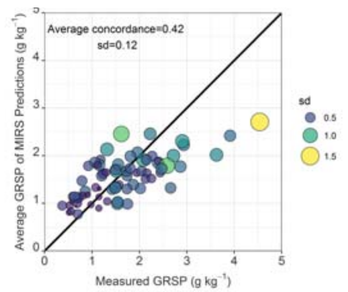

Fig.4 MIR predicted versus measured GRSP concentrations for validation dataset .

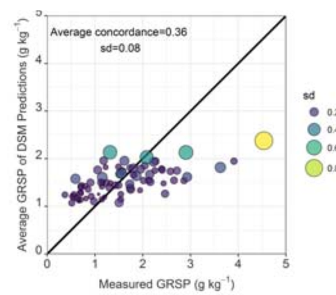

Fig.5 DSM predicted versus measured GRSP concentrations for validation dataset .

3.2. Mapping

The national GRSP spatial distribution was estimated. Patterns of GRSP exhibited a strong link with elevation and land use. Besides, these patterns were strongly linked the geographical distribution of soil organic carbon content (Chen *et al.*, 2019).

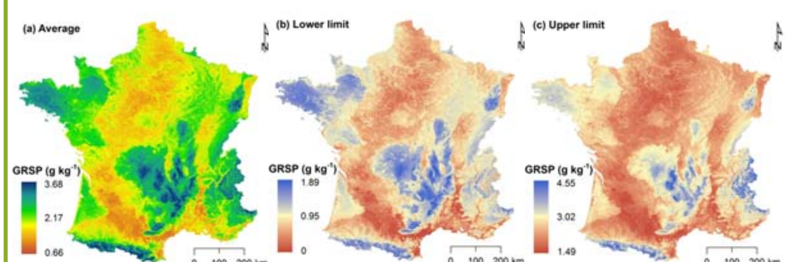

Fig.6 Spatial distribution of GRSP (a) and 90% prediction intervals (b and c)

4. Discussion and Conclusions

There are two limitations about these preliminary results.

1. We only used 374 out of 2200 RMQS sites.
2. To achieve an optimal prediction in a spatial modelling exercise, measurement errors should be filtered out (Cressie, 1991).

This large scale screening study indicates that operationally defined GRSP constitutes distinct fractions of soil organic matter with different dynamics to average organic matter.

Integrating Soil Sensing and Digital Soil Mapping provides a great potential for predicting and mapping GRSP at a national scale.

References

Bellon-Maurel, V., & McBratney, A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils–Critical review and research perspectives. Soil Biology and Biochemistry, 43(7), 1398-1410.

Chen, S., Arrouays, D., Angers, D. A., Chenu, C., Barré, P., Martin, M. P., ... & Walter, C. (2019). National estimation of soil organic carbon storage potential for arable soils: A data-driven approach coupled with carbon-landscape zones. Science of The Total Environment, 666, 355-367.

Cressie, N.A.C. 1991. Statistics for Spatial Data. New York: John Wiley & Sons.

McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. Geoderma, 117(1-2), 3-52.

Communications are welcome, please contact Songchao Chen (songchao.chen@inra.fr)