Geoderma 337 (2019) 99–110

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Merging country, continental and global predictions of soil texture: Lessons from ensemble modelling in France

Manon Caubet[a], Mercedes Román Dobarco[a], Dominique Arrouays[a], Budiman Minasny[b], Nicolas P.A. Saby[a,*]

[a] INRA, Centre de Recherche d'Orléans, US1106 Unité InfoSol, F-45075 Orléans, France
[b] Sydney Institute of Agriculture, School of Life and Environmental Sciences, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Handling Editor: A.B. McBratney

Keywords:
Digital soil mapping
Model ensemble
Granger-Ramanathan
Variance-weighted
Stratification
Soil texture

ABSTRACT

The increasing demand for soil information has led to the rapid development of Digital Soil Mapping (DSM) products. As a consequence, multiple soil maps are sometimes available for a particular area. Rather than selecting the best map, model ensemble offers a way to capitalize on existing soil information, and to improve the map accuracy. In this study we ensemble four topsoil texture maps of France with different resolution made by different organizations at the national, European, and global scale. We investigated two methods of model ensemble: the Granger-Ramanathan (GR) and Variance-Weighted (VW) methods. Ensemble methods based on area stratification were also tested to take into account local soil information. We also assessed the impact of the number of calibration points on the evaluation indicators. Both ensemble methods improved the accuracy of the map compared to the best of the primary maps, while the GR method outperformed the VW method. We found that the different stratification strategies did not improve the accuracy significantly when compared to the global methods. Finally, we showed that a relatively low number of calibration points is required in the merging process if the sampling is well designed. This study demonstrates that digital soil mapping products at various scales from various data sources can be combined with the ensemble method taking advantage of all existing efforts and taking care of harmonization issues.

1. Introduction

The increasing demand for soil information has led to the rapid development of Digital Soil Mapping (DSM; McBratney et al., 2003) techniques and products around the globe. In many parts of the world, DSM has now moved from an academic activity to operational (Arrouays et al., 2017; Minasny and McBratney, 2016). Global initiatives, such as *GlobalSoilMap* (Arrouays et al., 2014; Sanchez et al., 2009) and SoilGrids (Hengl et al., 2017) are now producing gridded prediction of soil attributes both at local and global scales. They fully recognize that both bottom-up and top-down products are complementary (Arrouays et al., 2017). As a consequence, several spatial data sources (maps, grids, or point data) referring to the same soil properties are often available for a particular area. In this context, one may wonder which source of data we should use. Indeed, each model and data used to create maps has its strengths and weaknesses. Therefore, on the one hand, it can be difficult for the end user to choose the best or most suitable map among those available; but on the other hand, the information provided by different sources may be complementary and merging them may be a way to capitalize existing information to create the most accurate map possible. These issues can be addressed by applying model ensemble (or model averaging). The principle is to gather predictions from different maps to create a final map which is at least as accurate as the most accurate map assembled (Diks and Vrugt, 2010). This method also allows combining data from different extent at different grid spacing.

Several methods of model ensemble have already been compared in soil science to predict soil properties: e.g., pH, particle size fraction, available water capacity, depth to the water table (Clifford and Guo, 2015; Heuvelink and Bierkens, 1992; Malone et al., 2014; Padarian et al., 2014; Román Dobarco et al., 2017). The ensemble methods improved the accuracy of the final maps compared to the primary maps (i.e., assembled maps). In France, model ensemble has been used at regional extent to predict topsoil texture combining maps at regional, national, and continental extent (Román Dobarco et al., 2017). The final map of clay content had greater accuracy than the primary maps

* Corresponding author.
E-mail addresses: manon.caubet@inra.fr (M. Caubet), mercedes.roman-dobarco@inra.fr (M. Román Dobarco), dominique.arrouays@inra.fr (D. Arrouays), budiman.minasny@sydney.edu.au (B. Minasny), nicolas.saby@inra.fr (N.P.A. Saby).

https://doi.org/10.1016/j.geoderma.2018.09.007
Received 9 March 2018; Received in revised form 7 August 2018; Accepted 4 September 2018

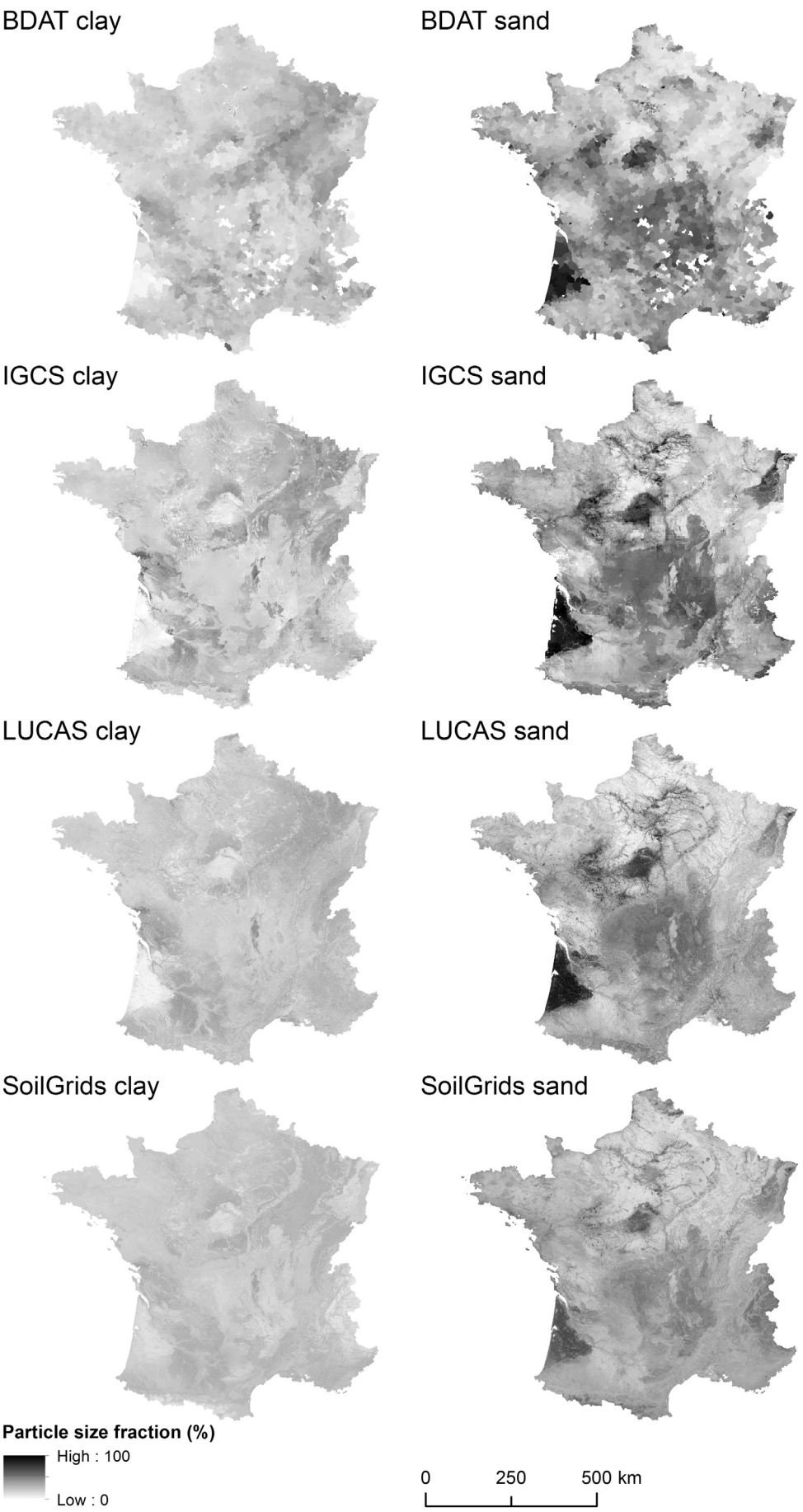

Fig. 1. Primary maps used for assemblage.