# Sentinel-2/1 bare soil temporal mosaics of 6-year periods for Soil Organic Carbon Content mapping in La Beauce, Central France

Diego URBINA-SALAZAR [1,2], Emmanuelle VAUDOUR [1], Anne C. RICHER-DE-FORGES [2], Songchao CHEN [3,4], Guillaume MARTELET [5], Nicolas BAGHDADI [6], Dominique ARROUAYS [2]

[1] Université Paris-Saclay, INRAE, AgroParisTech, UMR EcoSys, 91120, Palaiseau, France

[2] INRAE, Info&Sols, Orléans, France

[3] ZJU-Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 311215, China

[4] Institute of Agriculture Remote Sensing and Information Technology, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

[5] BRGM, UMR 7327, 45060 Orléans, France

[6] TETIS, Univ. Montpellier, INRAE, CIRAD, AgroparisTech, France

Abstract:
Satellite-based soil organic carbon content (SOC) mapping over wide regions is generally hampered by the low soil sampling density and the diversity of soil sampling periods. Some unfavorable topsoil conditions, such as high moisture, rugosity, the presence of crop residues, the limited amplitude of SOC values and the limited area of bare soil when a single image is used, are also among the influencing factors. To generate a reliable SOC map, this study addresses the use of Sentinel-2 (S2) temporal mosaics of bare soil (S2Bsoil) over 6 years jointly with soil moisture products (SMPs) derived from Sentinel 1 and 2 images, SOC measurement data and other environmental covariates derived from digital elevation models, lithology maps and airborne gamma-ray data. In this study (Urbina-Salazar et al., 2023 doi.org/10.3390/ rs15092410), we explore (i) the dates and periods that are preferable to construct temporal mosaics of bare soils while accounting for soil moisture and soil management; (ii) which set of covariates is more relevant to explain the SOC variability. From four sets of covariates (Figure 1, table 1), the best contributing set was selected, and the median SOC content along with uncertainty at 90% prediction intervals were mapped at a 25-m resolution from quantile regression forest models. The accuracy of predictions was assessed by 10-fold cross-validation, repeated five times. The models using all the covariates had the best model performance. Airborne gamma-ray thorium, slope and S2 bands (e.g., bands 6, 7, 8, 8a) and indices (e.g., calcareous sedimentary rocks, “calcl”) from the “late winter–spring” time series were the most important covariates in this model. Our results also indicated the important role of neighboring topographic distances and oblique geographic coordinates between remote sensing data and parent material. These data contributed not only to optimizing SOC mapping performance but also provided information related to long-range gradients of SOC spatial variability, which makes sense from a pedological point of view.

This research was supported by the POLYPHEME project through the TOSCA program of the Centre national d'études spatiales (CNES), STEROPES project of the EJP-SOIL Horizon H2020 and the MELICERTES project of the French National Research Agency (France2030, PEPR “agroécologie et numérique”).

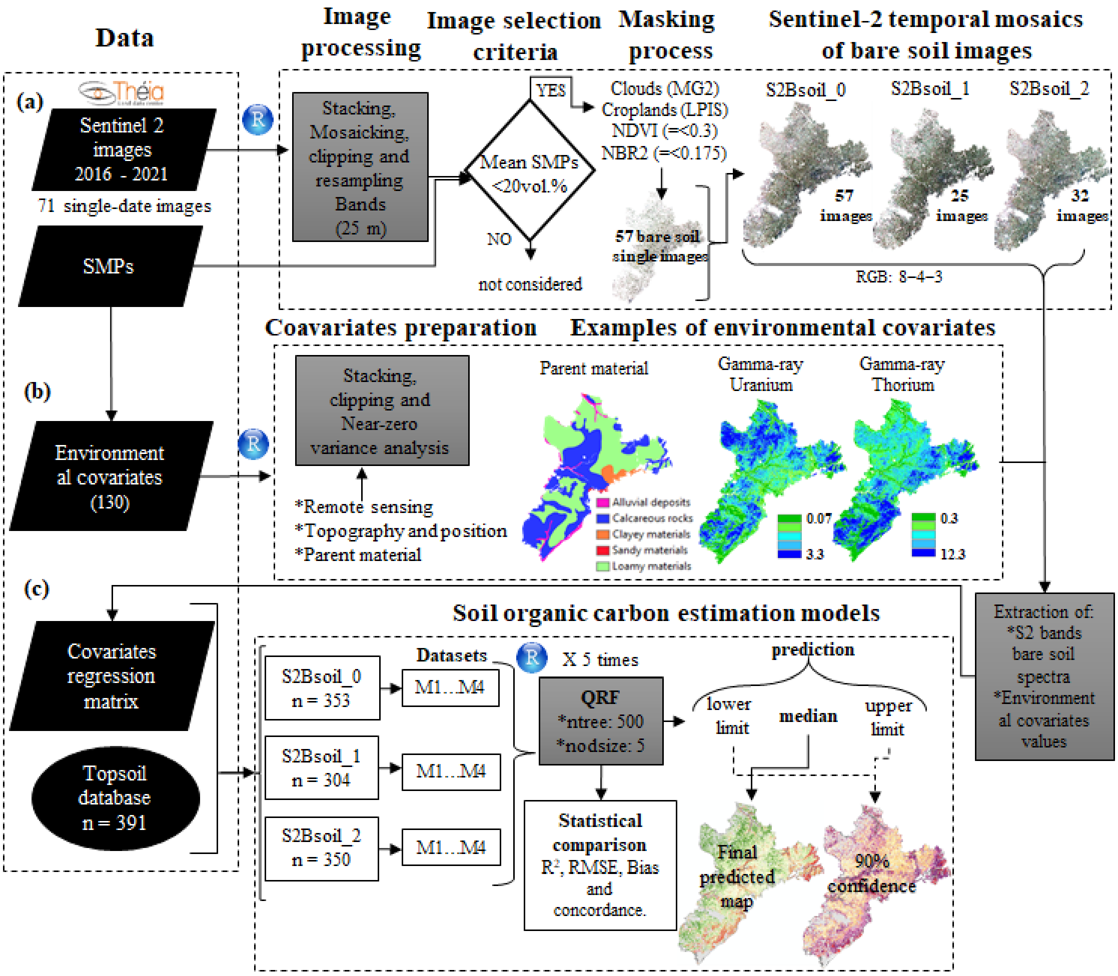


**Figure 1.** Flowchart methodology. Sentinel-2 temporal mosaics of bare soil (S2Bsoil) over 6 years: S2Bsoil_0, 57 images used; S2Bsoil_1, 25 images between February and May; S2Bsoil_2, 32 images between July and November. Sets of covariates: M1, the 10 S2Bsoil bands; M2, the 10 S2Bsoil bands plus spectral indices were considered (24 covariates); M3, the same covariates used in (M2) plus soil moisture were used (25 covariates); M4, all covariates used in (M1 to M3) plus covariates of topography and position and parent material (85 covariates).

**Table 1.** Model performance of a Quantile Random Forest to predict soil organic carbon by mosaics of bare soil over a 6-year period

| S2Bsoil | Modeling dataset | $R^2$ | RMSE ($g.kg^{-1}$) | Bias | Concordance |
|---|---|---|---|---|---|
| S2Bsoil_0 | M1 | 0.18 | 3.00 | −0.33 | 0.32 |
| | M2 | 0.19 | 2.98 | −0.31 | 0.33 |
| | M3 | 0.15 | 2.98 | −0.30 | 0.29 |
| | M4 | 0.26 | 2.75 | −0.20 | 0.40 |
| S2Bsoil_1 | M1 | 0.19 | 2.97 | −0.32 | 0.35 |
| | M2 | 0.22 | 2.90 | −0.30 | 0.35 |
| | M3 | 0.22 | 2.79 | −0.28 | 0.34 |
| | M4 | 0.33 | 2.59 | −0.22 | 0.42 |
| S2Bsoil_2 | M1 | 0.11 | 3.17 | −0.35 | 0.25 |
| | M2 | 0.11 | 3.14 | −0.30 | 0.24 |
| | M3 | 0.12 | 3.00 | −0.29 | 0.25 |
| | M4 | 0.27 | 2.71 | −0.21 | 0.39 |

$R^2$, coefficient of determination; RMSE, root mean square error.