



## Investigating the complementarity of thermal and physical soil organic carbon fractions

Amicie A. Delahaie<sup>1</sup>, Lauric Cécillon<sup>1</sup>, Marija Stojanova<sup>1</sup>, Samuel Abiven<sup>1</sup>, Pierre Arbelet<sup>2</sup>, Dominique Arrouays<sup>3</sup>, François Baudin<sup>4</sup>, Antonio Bispo<sup>3</sup>, Line Boulonne<sup>3</sup>, Claire Chenu<sup>5</sup>, Jussi Heinonsalo<sup>6</sup>,  
5 Claudy Jolivet<sup>3</sup>, Kristiina Karhu<sup>6</sup>, Manuel Martin<sup>3</sup>, Lorenza Pacini<sup>1,2</sup>, Christopher Poeplau<sup>7</sup>, Céline Ratié<sup>3</sup>, Pierre Roudier<sup>8</sup>, Nicolas P. A. Saby<sup>3</sup>, Florence Savignac<sup>4</sup>, Pierre Barré<sup>1</sup>

<sup>1</sup> École Normale Supérieure de Paris, Laboratoire de Géologie, France

<sup>2</sup> Greenback (commercial name: Genesis), Paris, France

<sup>3</sup> INRAE, Info&Sols, 45075, Orléans, France

10 <sup>4</sup> UMR ISTeP 7193, Sorbonne Université, CNRS, Paris, France

<sup>5</sup> UMR ECOSYS, INRAE, AgroParisTech, Université Paris Saclay, 91123 Palaiseau, France

<sup>6</sup> Department of Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland

<sup>7</sup> Thünen Institute of Climate-Smart Agriculture, Braunschweig, Germany

<sup>8</sup> Manaaki Whenua - Landcare Research, Palmerston North, New Zealand

15 *Correspondence to:* Amicie A. Delahaie (amicie.delahaie@ens.fr)

**Abstract.** Partitioning soil organic carbon (SOC) in fractions with different biogeochemical stability is useful to better understand and predict SOC dynamics, and provide information related to soil health. Multiple SOC partition schemes exist but few of them can be implemented on large sample sets and therefore be considered as relevant options for soil monitoring. The well-established particulate- (POC) vs. mineral-associated organic carbon (MAOC) physical fractionation scheme is one  
20 of them. Introduced more recently, Rock-Eval® thermal analysis coupled with the PARTY<sub>SOC</sub> machine-learning model can also fractionate SOC into active ( $C_a$ ) and stable SOC ( $C_s$ ). A debate is emerging as to which of these methods should be recommended for soil monitoring. To investigate the complementarity or redundancy of these two fractionation schemes, we compared the quantity and environmental drivers of SOC fractions obtained on an unprecedented dataset from mainland France. About 2,000 topsoil samples were recovered all over the country, presenting contrasting land covers and  
25 pedoclimatic characteristics, and analysed. We found that the environmental drivers of the fractions were clearly different, the more stable MAOC and  $C_s$  fractions being mainly driven by soil characteristics, whereas land cover and climate had a greater influence on more labile POC and  $C_a$  fractions. The stable and labile SOC fractions provided by the two methods strongly differed in quantity ( $MAOC/C_s = 1.88 \pm 0.46$  and  $POC/C_a = 0.36 \pm 0.17$ ;  $n = 843$ ) and drivers, suggesting that they correspond to fractions with different biogeochemical stability. We argue that, at this stage, both methods can be seen as  
30 complementary and potentially relevant for soil monitoring. As future developments, we recommend comparing how they relate to indicators of soil health such as nutrient availability or soil structural stability, and how their measurements can improve the accuracy of SOC dynamics models.



## 1 Introduction

Evaluating the biogeochemical stability of soil organic carbon (SOC) is crucial for predicting future SOC stock changes and assessing soil health. SOC biogeochemical stability depends on many interacting factors such as soil organic matter (SOM) molecular composition, and interactions with the mineral matrix (von Lützow et al., 2006; Schmidt et al., 2011). For a given soil, SOC represents a continuum of mean residence times (MRT) ranging from days to millennia (Balesdent, 1996). However, this continuum cannot be measured directly and then used efficiently for evaluating SOC biogeochemical stability. For this reason, many studies have proposed SOC fractionation schemes to distinguish fractions with contrasting residence times, enabling a practical assessment of SOC biogeochemical stability (Poeplau et al., 2018). Nevertheless, many of these fractionation methods are expensive and time-consuming, making their use on large datasets almost impossible.

Recently, Lavallee et al. (2020) proposed a drastic simplification of the SOC biogeochemical stability continuum by dividing SOC into two fractions of contrasted stability: particulate (POC) and mineral-associated (MAOC) fractions, following on early work by Cambardella & Elliott (1992). This physical fractionation scheme is relatively quick, can be implemented on hundreds of samples (Lugato et al., 2021), and recent studies have underlined the potential interest of such a dualistic view of the SOC persistence continuum (Cécillon et al., 2021b; Angst et al., 2023; Lugato et al., 2021).

Less popular than physical fractionation, thermal fractionation has also been proposed as an efficient method to evaluate SOC quality (Plante et al., 2009). In particular, Rock-Eval® thermal analysis has been the subject of a growing interest in recent years for assessing SOC biogeochemical stability (Saenger et al., 2013; Barré et al., 2016; Sebag et al., 2016; Soucémariadin et al., 2018). This method is relatively fast and can be used to analyse a series of thousands of samples (Delahaie et al., 2023). Moreover, Cécillon et al. (2018; 2021a) developed a machine learning model, PARTY<sub>SOC</sub>, which uses Rock-Eval® thermal analyses results as input variables to estimate the proportion of SOC that is stable at a centennial scale and, by difference, the proportion of SOC that is active at this timescale. Kanari et al. (2022) showed that the fractions determined by PARTY<sub>SOC</sub> match the “stable” and “active” fractions of the AMG model (Clivot et al., 2019), improving its simulations of SOC stock evolutions in croplands. As a result, Rock-Eval® thermal analysis associated with PARTY<sub>SOC</sub> allows partitioning SOC in a more labile fraction ( $C_a$ ) with an MRT ranging from 20 to 40 years, and a stable fraction ( $C_s$ ) which can be considered inert at a centennial timescale.

The POC/MAOC physical fractionation and the  $C_a/C_s$  thermal fractionation are therefore two methods that can potentially be used to split SOC in fractions with contrasted biogeochemical stability and be implemented on large sample sets. With the growing interest in monitoring programs of soil health and the need for better initialization methods able to improve the accuracy of SOC dynamics models, it is necessary to assess the extent to which these two fractionation approaches are complementary or redundant.

Our hypotheses were that as they do not target the same SOC pools (Balesdent, 1996; Poeplau et al., 2018; Kanari et al., 2022), POC and  $C_a$  as well as MAOC and  $C_s$  fractions may represent different quantities and have different environmental drivers (soil characteristics, land cover, and climate variables) and can therefore be considered as complementary. To test our



hypotheses, we used an unprecedented dataset comprising ca. 2,000 Rock-Eval® thermal analyses and ca. 1,000 POC/MAOC physical fractionation data from the analysis of topsoil (0–30 cm) samples that are part of the French soil monitoring network (RMQS).

## 2 Material and methods

### 70 2.1 RMQS soil samples

The soil samples used in this article are part of the French “Réseau de mesures de la qualité des sols” (RMQS) network and were previously described in Gogé et al. (2012) and Delahaie et al. (2023). A complete description of this monitoring network is available in Jolivet et al. (2006; 2022). Briefly, French mainland soils are monitored every 15 years following a 16 km × 16 km regular square grid, resulting in 2,170 sites. When possible, the sampling site is set at the centre of the cell; alternatively, another site is selected if needed within a 1 km radius from the centre of the cell. At each sampling site, 25 topsoil samples (0 to 30 cm or tilled layer depth, whichever depth is smaller) are taken from a 20 m × 20 m area using a spiral auger, and then mixed, resulting in a composite sample of 5 to 10 kg.

The composite samples are then placed in trays and air-dried at 30 °C for 8 to 10 days, and quartered according to NF ISO 11464, resulting in a subsample of ca. 650 g. They are then hand-crushed and sieved at 2 mm, and an aliquot is ground under 250 µm by a Cyclotec 1093 (FOSS) (Gogé et al., 2012). The remains of the samples are stored in plastic buckets. A total of 2,037 samples from the 1<sup>st</sup> sampling campaign (2000–2009) out of 2170 were recovered and analysed by Rock-Eval® thermal analysis in Delahaie et al. (2023).

### 2.2 Soil and environmental data associated to each RMQS site and topsoil sample

#### 85 2.2.1 Soil data

Physical and chemical analyses were carried out on the composite soil samples at the Laboratoire d’Analyse des Sols (INRAE, Arras, France). The inorganic carbon content ( $C_{inorg}$ ) was derived from the total carbonate content, in grams per kilogram of sample (volumetric method, NF EN ISO 10693), and calculated as  $C_{inorg} = \text{total carbonate} \times 0.12$ ; the total carbon content, in grams per kilogram of sample, was determined by elemental analysis using dry combustion on non-decarbonated soil; the organic carbon content was derived from the elemental analysis (TOC<sub>ea</sub>), in grams per kilogram of sample, and calculated as total carbon content minus inorganic carbon content:  $TOC_{ea} - C_{inorg}$  (NF ISO 10694 ; “NF” standing for French standard); the total nitrogen, in grams per kilogram of sample, was determined by dry combustion (NF ISO 13878); the particle size distribution was measured without decarbonation, in grams per kilogram of sample (Robinson pipette and underwater sieving, method validated in relation to standard NF X31-107); pH was measured in a suspension of soil diluted with water (dilution 1 : 5, NF ISO 10390); the exchangeable calcium content, in centimoles per kilogram of sample, was measured by cobaltihexammine chloride extraction (NF X31-130); the exchangeable magnesium content, in centimoles per kilogram of sample, was measured by cobaltihexammine chloride extraction (NF X31-130); the exchangeable

potassium content, in centimoles per kilogram of sample, was measured by cobalthexammine chloride extraction (NF X31-130); the free iron oxides, in grams per 100 g, measured with the Tamm method in the dark (amorphous oxides) and Mehra–  
100 Jackson method (crystalline oxides) (INRA standard/NF ISO 22036).

### 2.2.2 Climate data

We allocated to each sampling site the climatic data corresponding to the SAFRAN 8 km x 8 km grid-cell based on where the cell was located (<https://publitheque.meteo.fr/okapi/accueil/okapiWebPubli/index.jsp>, last access: 31 March 2022). The  
105 daily data were averaged over the 1969–1999 period (i.e. the 30-year common period before the first sampling campaign starting in 2000) in order to compute the mean annual temperature (MAT) and mean annual precipitation (MAP) for each site.

### 2.2.3 Land cover data

Land cover data were recorded during sampling. Four main categories of land cover were considered for this study:  
110 “croplands”, “forests”, “grasslands”, and “vineyards & orchards”. A few samples were collected in “wastelands”, “urban parks”, and “sites with little human disturbance”. Considering the very small number of samples, “wastelands” (ca. 10) and “gardens” (n = 3) were not included in this study. The number of samples from environments with little human disturbance (ca. 30) could potentially be considered sufficient for statistical treatment; however, these samples represent a very heterogeneous set (10 miscellaneous subclasses, such as peatlands, alpine grasslands, water edge vegetation, heath, and dry  
115 siliceous meadows). Thus, those sites were also discarded.

## 2.3 Thermal SOC fractionation

### 2.3.1 Rock-Eval® thermal analyses

In total, 2,037 samples were analysed by Rock-Eval® thermal analysis (Disnar et al., 2003; Baudin et al., 2015). For each  
120 sample, ca. 60 mg of finely ground matter (< 250 µm) was placed in a special high-temperature-resistant stainless-steel pod, allowing the transport gas to pass through, and then placed inside a Rock-Eval® 6 (RE6) Turbo device (Vinci Technologies). There, it underwent a first phase of pyrolysis under an inert atmosphere (N<sub>2</sub>) from ambient temperature to 650 °C (three-minute isotherm at 200°C and then a temperature ramp of 30°C min<sup>-1</sup>), and a second phase of oxidation under the laboratory atmosphere purged from water and CO<sub>2</sub>, from 300 °C to 850 °C (one-minute isotherm at 300°C and then a temperature ramp  
125 of 20°C min<sup>-1</sup>). During the pyrolysis phase, hydrocarbon effluents were monitored by a flame ionisation detector, and CO and CO<sub>2</sub> were monitored by infrared detectors. During the oxidation phase, CO and CO<sub>2</sub> were monitored by infrared detectors. The resulting thermograms were processed using the Geoworks software (Geoworks V1.6R2, Vinci Technologies, 2021).

The organic carbon yield was defined as the ratio of the total organic carbon amount measured by Rock-Eval® thermal  
130 analysis (TOCre6, calculated from thermogram area integration) over the total organic carbon amount measured by elemental analysis (TOCea). We chose to apply a quality criterion on this yield: further study was conducted only on



135 samples with an organic carbon yield ranging from 0.7 to 1.3. This range was set to identify the acceptable yields ensuring the quality of the Rock-Eval® analysis as well as the identity of the sample. Of the 2,037 samples analysed by Rock-Eval® thermal analysis, 1,891 presented an organic carbon yield ranging from 0.7 to 1.3 (Delahaie et al., 2023). We also removed 12 samples with TOC<sub>ea</sub> > 120 g kg<sup>-1</sup> to avoid organic soils (Eggleston et al., 2006), resulting in a dataset of 1,879 samples.

### 2.3.2 The PARTY<sub>SOC</sub> fractionation

140 The PARTY<sub>SOC</sub> model (Cécillon et al., 2018; 2021a) is a machine-learning model using the results of the Rock-Eval® thermal analysis of mineral topsoils as entry variables. This model, trained on data from long-term agronomic experiments, uses 18 Rock-Eval® parameters as input variables and determines the proportion of the centennially-persistent organic carbon pool in a mineral topsoil sample. The stable SOC proportion is now calculated routinely by the Geoworks software (Geoworks V1.6R2, Vinci Technologies, 2021) using the model PARTY<sub>SOC</sub> v2.0EU published in Cécillon et al. (2021a). For each site, we multiply the proportion of stable C by the TOC<sub>ea</sub> to calculate the C<sub>s</sub> pool (g C kg<sup>-1</sup> sample); the active pool C<sub>a</sub> (g C kg<sup>-1</sup> sample) is obtained by difference: C<sub>a</sub> = TOC<sub>ea</sub> - C<sub>s</sub>.

145

### 2.4 Physical SOC fractionation

The physical SOC fractionation, i.e. particle size fractionation was conducted by the SADEF laboratory (Aspach-Le-Bas, France) on a subset of the RMQS (997 sites) following a protocol based on the norm NF X 31-516, itself based on Balesdent et al. (1991; 1998). The dispersion was carried out with a solution of sodium hexametaphosphate at a concentration of 5 g L<sup>-1</sup>. 50 g of soil sieved at 2 mm was stirred in 180 ml of hexametaphosphate solution with 10 glass beads of 5 mm diameter and underwent rotary agitation for 16 hours at 20 °C at 45 rpm in a 250 mL flask. The fractions were then sieved by hand at 0.2 mm with rotative movements, and sprays of demineralised water were used to complete the sieving process. The matter remaining on the sieve was transferred in a capsule, dried, and crushed. The suspension containing the fine particles (< 0.2 mm) and rinsing water were collected for further sieving to 0.05 mm. The same principle was then applied for the sieving at 0.05 mm, but it was conducted on three or four successive fractions of the suspension to avoid clogging the sieve. The liquid fraction containing the particles below 0.05 mm was recovered in a 1 L crystallizer and dried. The drying of the fractions was carried out in a ventilated oven at 105 °C.

155 This fractionation process thus resulted in three fractions: the mineral-associated organic matter fraction (MAOM) corresponds to the 0–50 µm fraction, the fine particulate organic matter fraction (fPOM) corresponds to the 50–200 µm fraction, and the coarse particulate organic matter fraction (cPOM) corresponds to the 200–2000 µm fraction. The carbon contained in the MAOM fraction constitutes the MAOC while the carbon contained in both the fPOM and cPOM constitutes the POC.

160 After drying, all the dry matter in each fraction was recovered. Each fraction was introduced into a corundum bowl and ground with corundum balls (Retch PM400 planetary ball mill) at 400 rpm for 5 minutes to ensure the final matter is ground at < 250 µm and homogenised.

165



Carbon and nitrogen measurements were carried out on a Flash 2000 Elemental Analyzer for soils without carbonates (determined by acid test) following the norms NF ISO 10694 and NF ISO 13878, respectively. For carbonated soils, only nitrogen was measured on the Flash 2000 Elemental Analyzer (Dumas method NF ISO 13878). Organic carbon was analysed by chemical oxidation (NF ISO 14235). The total organic carbon retrieved after physical fractionation is noted  
170 TOC<sub>fr</sub> and the organic carbon yield for this fractionation was defined as the ratio of TOC<sub>fr</sub> over TOC<sub>ca</sub>.

The C yield was on average 93.4% for the 997 samples. For the same reasons as above, we also introduced a quality criterion on this yield, identical to the one for the Rock-Eval® results (0.7 to 1.3). Following this rule, 33 samples were removed from the dataset. Then, four samples with TOC<sub>ca</sub> > 120 g kg<sup>-1</sup> were also removed to avoid organic soils (Eggleston et al., 2006),  
175 resulting in a final dataset comprising 960 fractionation results. The removed samples showed no particular pedoclimatic characteristics (Fig. A1, Appendix).

As the physical and thermal fractionations were not conducted on all the samples, there are samples for which data of only one method was available. The intersection of the physical fractionation dataset and thermal fractionation dataset consists of 843 samples and is thereafter designated as the “intersection dataset”.

180

## 2.5 Statistical analysis

### 2.5.1 The determination of drivers of the POC, MAOC, C<sub>a</sub>, C<sub>s</sub>, and TOC<sub>ca</sub> quantities

In this study, we tested the influence of different environmental variables on the C<sub>s</sub>, C<sub>a</sub>, MAOC, and POC content by using Random Forest regression models based on the method used by Georgiou et al. (2022). One advantage of Random Forest  
185 models is that they can cope with non-normally distributed and correlated variables (Breiman, 2001). The considered environmental drivers were related to soil characteristics (particle size distribution, pH, inorganic carbon content, exchangeable cations contents (calcium, magnesium, potassium), amorphous and crystalline iron oxyhydroxides contents), climate (mean annual precipitation, mean annual temperature), and land cover. The relative importance of each of these features as estimated by the Random Forest model allowed us to evaluate the main drivers of the quantity of each fraction.

The modelling pipeline was divided into two steps. The first preprocessing step used one-hot encoding (creating one boolean  
190 column per class for any categorical variable), while all the numeric variables were standardised by removing their mean and dividing them by their variances. In the second step, a bootstrapped Random Forest regressor was used to calibrate the actual model. Grid search and cross-validation were used to choose the model's hyper-parameters (Table A1, Appendix), using cross-validated R<sup>2</sup> (determination coefficient) as a performance metric. These hyper-parameters define the structure and  
195 construction of the forest's trees, and as such they are crucial and must be chosen wisely to ensure that the model does not overfit or underfit the training data.

After the model was calibrated, the importance of the environmental predictors was calculated using two different methods: the mean decrease in impurity (MDI) and the Permutation Importance (PI) score (Louppe, 2014). Both methods gave a measurement of the importance of the environmental variables selected in the model. MDI aims at selecting the predictors



200 that, on average, produce trees with the purest nodes and leaves. In this case, purity refers to the similarity of the samples contained in a single leaf. MDI is calculated using the training set and it is the default decision metric used in constructing scikit-learn's random forests. However, the MDI has a known tendency to lower the importance of the low-cardinality variables and possibly has a bias towards highly-correlated variables (Louppe et al. 2013). The PI score measures the impact of each individual predictor by randomly permuting it and then calculating the increase in prediction error as opposed to  
205 using the non-permuted predictor. By construction, PI can be calculated on both the training and the test set. Its main advantage over MDI is that it shows no bias towards high-cardinality predictors. Highly-correlated variables can be detected using PI as their permutation will have little to no impact on the model's prediction accuracy.

In order to analyse the variable importance results, the environmental variables were grouped into three broad categories: "land cover", "pedology" (particle size distribution, pH, inorganic carbon content, exchangeable calcium content,  
210 exchangeable magnesium content, exchangeable potassium content, amorphous, and crystalline iron oxyhydroxides contents), and "climate" (mean annual temperature and mean annual precipitation). These groupings were used to guide visual analysis of the variable importance plots generated for the  $C_s$ ,  $C_a$ , MAOC, and POC models.

The Random Forest modelling (and associated metrics) was done in Python as implemented in the the scikit-learn v1.3.0 library (Pedregosa et al, 2011), while the least-squares linear regression used the scipy v1.10.1 library (Virtanen et al., 2020).

215

### 2.5.2 Assessments of the effect of land-cover on fractions

To assess the effect of land cover on the different fractions, we performed pairwise comparisons of medians using non-parametric Kruskal–Wallis tests ( $p < 0.05$ ) followed by Wilcoxon tests, with  $p < 0.05$  for each pair. The correction of  $p$  values within the framework of the multiple comparisons was done using the Holm–Bonferroni method. Correlations  
220 between parameters were calculated using the Spearman method.

The data processing and statistical analysis were carried out using R software (V4.1.2; R Core Team 2021): the corrplot (Wei and Simko, 2021), car (Fox and Weisberg, 2019), ggplot2 (Wickham, 2016), ggpubr (Kassambara, 2023a), factoextra (Kassambara and Mundt, 2020), plot3D (Soetaert, 2021), and rstatix (Kassambara, 2023b) packages were added.

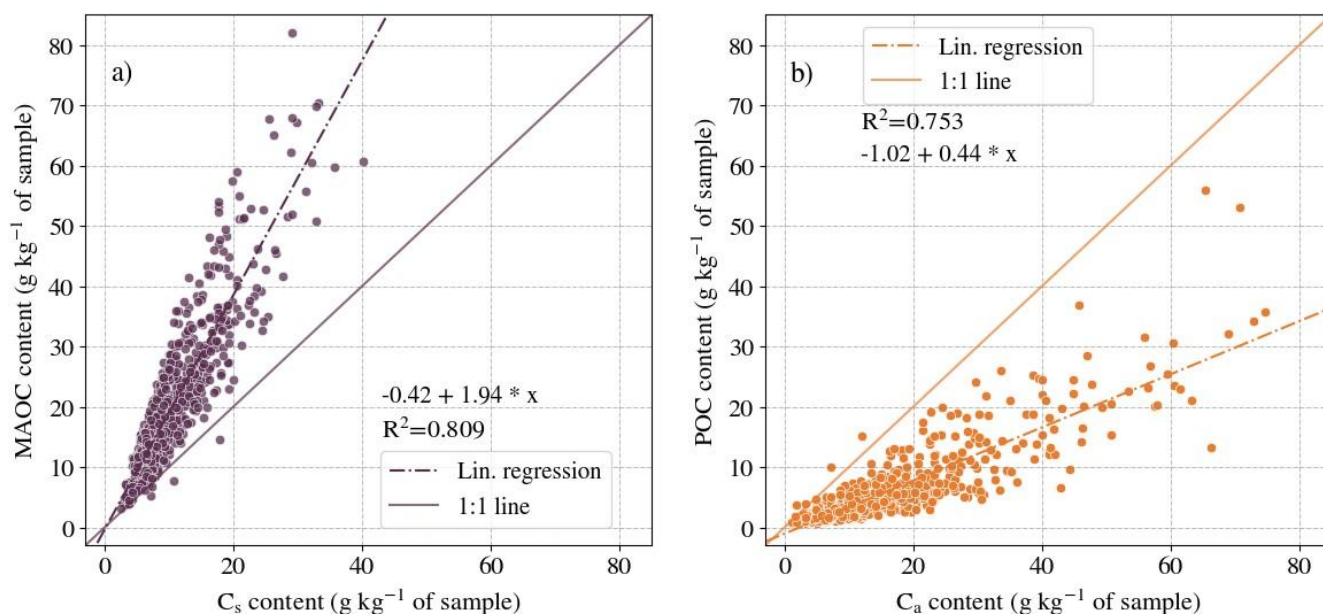
## 225 3 Results

### 3.1 POC vs. $C_a$ , MAOC vs. $C_s$ : fractions in different quantities

Figure 1 shows the quantities of  $C_s$  plotted against MAOC and the quantities of  $C_a$  plotted against POC, comparing two by two the more stable and more labile fractions for each fractionation scheme for the intersection dataset (samples having been subjected to both thermal and physical fractionation schemes). Regarding the more stable fractions (Fig. 1a), the MAOC content was much higher than the  $C_s$  content (on average  $19.13 \text{ g kg}^{-1}$  of sample versus  $10.06 \text{ g kg}^{-1}$  of sample for the 843  
230 samples of the intersection dataset). The average ratio of MAOC/ $C_s$  for the samples of the intersection dataset was  $1.88 \pm 0.46$ . For the more labile fractions (Fig. 1b), the POC content was much lower than the  $C_a$  content (on average  $5.32 \text{ g kg}^{-1}$  of sample versus  $14.40 \text{ g kg}^{-1}$  of sample for the 843 samples of the intersection dataset). The average ratio of POC/ $C_a$  for the



235 samples of the intersection dataset was  $0.36 \pm 0.17$ . The correlations between  $C_s$  and MAOC on one side (0.90) and between  $C_a$  and POC on the other side (0.87) are both very significant. In comparison, the correlation coefficients between  $C_s$ , MAOC,  $C_a$  and POC on the one hand, and TOC on the other hand are 0.91, 0.96, 0.98, and 0.86 respectively.



240 **Figure 1: Comparison of the quantities of the more stable and more labile fractions for the physical and thermal SOC fractionation schemes, with their correlation coefficient  $R^2$  and linear regression. Panel (a) shows the quantities of MAOC plotted against  $C_s$ . Panel (b) shows the quantities of POC plotted against  $C_a$ . The dataset is the intersection dataset, i.e. samples for which thermal and physical data are available ( $n = 843$ ).**

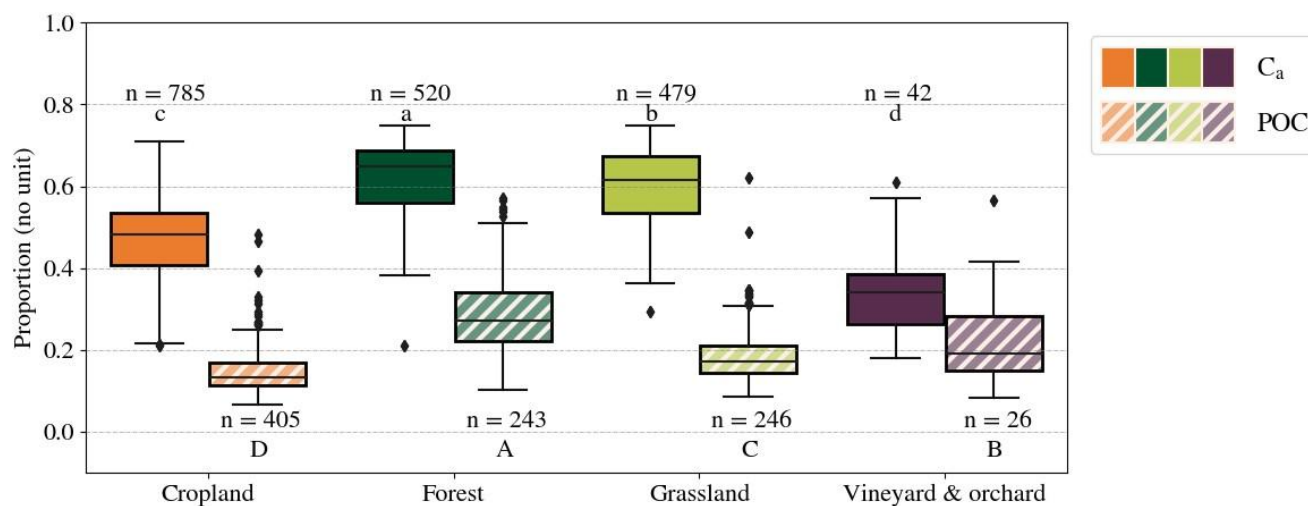
### 245 3.2 Differences in SOC fractions' proportions under different land covers

Figure 2 shows  $C_a$  and POC as a proportion of TOC for the four considered land covers. As  $C_a + C_s$  and  $POC + MAOC$  were equal to TOC, the analysis on  $C_s$  or MAOC proportions would have given the same information. Figure 2 shows that the proportion of  $C_a$  of the TOC followed the order vineyards & orchards < croplands < grasslands < forests (and similarly for POC). The mean values of the proportion of  $C_a$  were 0.48, 0.60, and 0.62 in croplands, grasslands and forests, respectively. It also shows that POC as a proportion of TOC was significantly smaller in croplands compared to forests, grasslands, and vineyards & orchards, but the median value of POC in croplands was close to the median value in grasslands (0.13 in croplands, 0.17 in grasslands, 0.27 in forests, and 0.19 in vineyards & orchards). As for the median values, the mean values showed a similar difference between croplands ( $0.15 \pm 0.05$ ) and grasslands ( $0.19 \pm 0.07$ ), with forests displaying a higher value ( $0.28 \pm 0.09$ ). While the  $C_a$  fraction generally represents a high proportion of TOC (up to 0.75 in forests; proportion of





255  $C_a > 0.5$  in 1277 out of 1879 samples), POC most often represented only a minority of TOC (proportion of POC  $> 0.5$  in 17 out of 960 samples).



260 **Figure 2: Proportion of the  $C_a$  and POC fractions depending on the land cover. The black line in each box is the median, the lower and upper edges of the black rectangle are the respective first (Q1) and third (Q3) quartiles, and the lower and upper whiskers are the maximum between the minimum value or the first quartile minus 1.5 times the interquartile range ( $\max [\min; Q_1 - 1.5 \times (Q_3 - Q_1)]$ ) and the minimum between the maximum or the third quartile plus 1.5 times the interquartile range ( $\min [\max; Q_3 + 1.5 \times (Q_3 - Q_1)]$ ), respectively. Different letters indicate significant differences in the distribution of the values for the land covers according to a Kruskal–Wallis test ( $p < 0.05$ ) and a pairwise Wilcoxon rank sum test ( $p < 0.05$ ); lowercase letters are used for  $C_a$  and uppercase for POC.**

265

Additionally, different ratios ( $MAOC/C_s$ ,  $POC/C_a$  and  $(MAOC/C_s)/(POC/C_a)$ ) are given in Figure A2, Appendix. Croplands and grasslands exhibited a similar and small  $POC/C_a$  ratio, while forests and vineyards & orchards show a higher  $POC/C_a$  ratio. On the contrary, the  $MAOC/C_s$  ratio is very small in the vineyards & orchards, and the highest in grasslands, with a significant difference between croplands and grasslands.

270

### 3.3 Drivers of the different SOC fractions quantities

The Random Forest models fitted for the four different SOC fractions aimed at elucidating the relative importance of the soil and environmental variables. Their explanatory capacity was evaluated based on the  $R^2$  values obtained for each fraction. The Random Forest models'  $R^2$  scores on the test set were 61 % for  $C_s$ , 53 % for MAOC, 57 % for  $C_a$ , 36 % for POC, and 58 % for TOC<sub>ca</sub> (Fig. A2, Appendix). The  $R^2$  values obtained on cross-validation for the train set were higher but close (Table

275



A1, Appendix), except for the POC which has a strongly higher train set score. Overall, the models' performance was satisfactory, and allowed their variable importance to be analysed.

Figure 3 shows the importance of the different categories of environmental variables on the quantities of POC,  $C_a$ , MAOC, and  $C_s$ , evaluated using two different methods. The results with both the MDI and the PI showed a higher importance of the pedological features and a smaller importance of climate and land cover in the more stable fractions ( $C_s$  and MAOC). On the contrary, the more labile fractions ( $C_a$  and POC) showed a higher importance of land cover and, to a lesser extent, climate, compared to  $C_s$  and MAOC.  $C_a$  tended to show a slightly stronger influence of climate and weaker influence of land cover, while POC rather showed the opposite. Among the more stable fractions, climate and land cover tended to have a slightly higher importance for MAOC than for  $C_s$ . The results for the TOC<sub>ea</sub> showed a mixture of drivers, in between stable and labile fractions.

The results for the Spearman correlation coefficients between all environmental variables and fractions quantities are given in Appendix in Table A2. Soil variables favouring organic matter/mineral interactions (clay, metallic oxides, CEC, exchangeable calcium) were positively correlated with fractions. Overall, these correlations were stronger for the  $C_s$  and MAOC fractions. On average, iron oxyhydroxides and exchangeable cations are the most important factors influencing the size of the fractions (Figure 3). Carbonates and pH little influenced the size of the fractions and texture had a minor role but for  $C_s$ . Regarding climate variables, MAT had a higher influence than MAP except for  $C_s$  fractions.

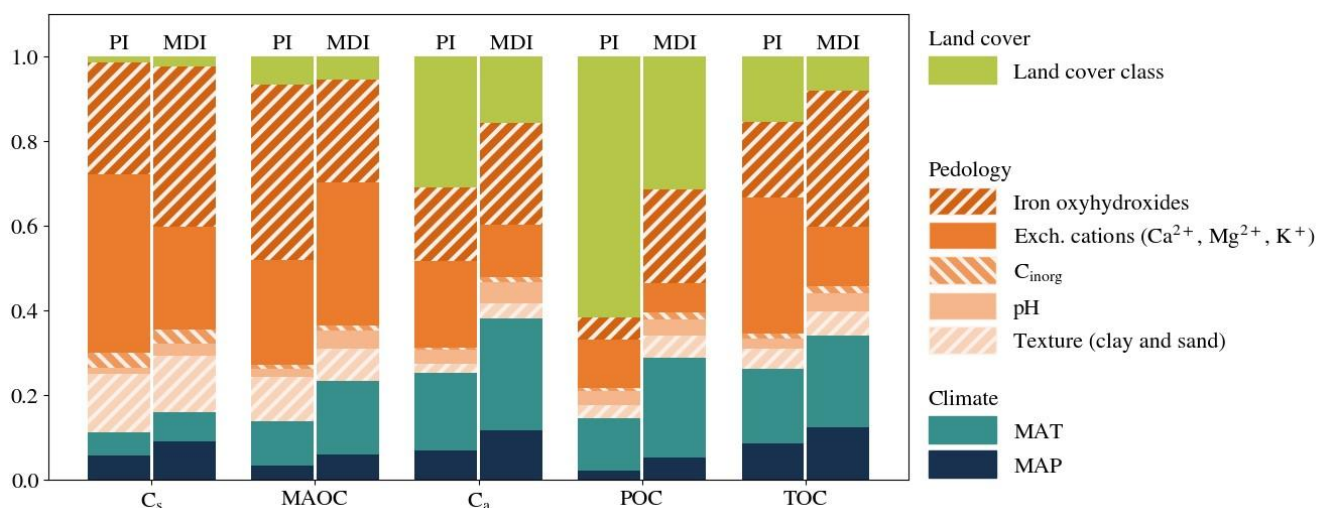


Figure 3: Importance of the different categories of soil and environmental variables (climate, pedology, and land cover) for the four fractions  $C_s$ , MAOC,  $C_a$ , and POC, and TOC<sub>ea</sub> as a comparison (in g C kg<sup>-1</sup> sample), assessed using MDI and PI.



## 4 Discussion

### 4.1 A strong influence of land cover on the more labile SOC fractions

300

This study, based on an unprecedented number of measurements ( $n = 960$  for POC/MAOC and  $n = 1,879$  for  $C_a/C_s$ ), shows the high influence of land cover on the relative quantity of SOC labile and stable fractions (Fig. 2).

The higher proportion of POC in forest than in grassland and cropland soils (0.13, 0.17, and 0.27 in croplands, grasslands, and forests, respectively) was also observed in Lugato et al. (2021) who used results from 352 physical fractionations obtained using a fractionation protocol similar to ours on samples from the LUCAS soil monitoring network. It was also observed by Hansen et al. (2024) on a global dataset combining physical fractions obtained with different protocols. Other studies showed similar results for POC, but comparisons are less straightforward as the physical fractionation protocols used are significantly different from those used in our study. For instance, some studies used the protocol of Zimmerman et al. (2007), where the sorting size for POC is  $> 63 \mu\text{m}$  (compared with  $> 50 \mu\text{m}$  in our protocol) and part of POC is also recovered after the disruption of sand-sized aggregates. When combining free POC and occluded POC (which corresponds to what is called POC in our study), Poeplau & Don (2013) found proportions of POC of 0.15 in croplands, 0.21 in grasslands, and 0.27 in forests in a variety of sites across Europe which is very close to the values observed here for French soils. The value of 0.13 for the proportion of POC in croplands is also close to the 0.15 value used in Angers et al. (2011) to estimate the proportion of POC from TOC. Chen et al. (2019) gathered data from multiple previous studies to derive POC proportion values of 0.15 for croplands, 0.31 for grasslands and 0.34 for forests.

Regarding the  $C_a$  fraction, we found a higher proportion of  $C_a$  in forest than in grassland and cropland soils (0.48, 0.62, and 0.65 in croplands, grasslands, and forests, respectively). This was also observed in previous works, but the body of literature available for discussion is much more limited than for physical fractionation. We can note that the median value of  $C_a$  for croplands is close to the mean value (0.48 vs 0.42) obtained by Kanari et al. (2022) on nine long-term field experiments in mainland France. Moreover, the proportion of  $C_a$  is set to 0.60 in the AMG model for grasslands (Clivot et al., 2019) in its default parameterisation which is also in line with the  $C_a$  proportions observed for grassland sites of the RMQS (0.62 on average). The difference is bigger regarding croplands as the default parameterisation for  $C_a$  is set to 0.35 while the value found for the RMQS is 0.48. This may be explained by the fact that the AMG model was developed on long-term experimental sites that have been cropped for several decades, whereas French agricultural soils have probably on average undergone more changes in vegetation cover. The higher  $C_a$  value in RMQS cropland soils could thus be due to a difference in land-cover history in French cropland soils compared to cropland sites of long-term experiments.

The similar proportions of POC in vineyards & orchards and grasslands were less expected as C inputs and contents are reduced in vineyards & orchards (grass cover was very sparse in vineyards & orchards at the time of the sampling). This may be explained by differences in the composition of the particulate organic matter for the different land covers, as the particulate organic matter groups together particles that can be biochemically quite heterogeneous (Schrumpf et al., 2013; Soucémariadin et al., 2019). In our study, the POC fraction in vineyards & orchards topsoils might be much more

330



biogeochemically stable than the same fraction in cropland and grassland topsoils because of the presence of more lignified woody debris and pyrogenic C derived from the combustion of vine shoots. An extreme example of soil with high proportion of POC and  $C_s$  can be found at the Versailles long-term bare fallow site where topsoils have a proportion of  $C_s$  close to 1 and a POC proportion around 0.30, constituted essentially of charcoal (Chassé et al., 2021). Our results therefore suggest that POC/MAOC fractionation gives an erroneous view of biogeochemical stability for vineyards & orchards, which have the same POC proportion as grassland but are very C depleted (9.5 vs 24 g C kg<sup>-1</sup>). The biochemical nature of the fractions is unaccounted for in our study as this data is not available, but it is probably an important driver for POC quantities. The fact that such an important driver is not taken into account may explain the poorer performance of the model for POC compared to the other fractions (36% vs 55-60% for the other fractions).

#### 4.2 SOC fractions with different quantities, drivers and biogeochemical stabilities

The significant differences between the  $C_s$  and MAOC quantities on one side, and  $C_a$  and POC quantities on the other side (Fig. 1), show that they do not correspond to the same fractions. This result was expected due to the definition of the four fractions, but it is evidenced for the first time on a large dataset. Indeed, previous studies using isotopic measurements observed that the mean residence times for MAOC ranged from decades to centuries (Anderson & Paul, 1984; Balesdent et al. 1987; Balesdent 1996; von Lützow et al., 2007; Kleber et al., 2015). By definition,  $C_s$  corresponds to centennially stable SOC (Cécillon et al., 2018; 2021a). This implies that MAOC encompasses also a certain amount of relatively labile SOC and therefore explains its larger quantity compared to  $C_s$ . Potentially, this could partly be related to the fractionation method itself, in which only size fractionation is employed to separate fractions, so that the MAOC fraction might also contain a certain amount of fine POC or soluble compounds (Lavallee et al., 2020; Cotrufo et al., 2023). Conversely, the  $C_a$  fraction which corresponds to SOC fraction with a mean residence time of 20-40 years (Kanari et al., 2022) is larger than the POC fractions dominated by SOC with a generally shorter MRT (Balesdent, 1996; von Lützow et al., 2007).

This ranking in terms of MRT,  $C_s > MAOC > C_a > POC$ , is also in line with the different environmental drivers explaining the quantities of these four fractions, whatever the selected method (MDI or permutation) (Fig. 3). Indeed, we observed that the importance of land cover is much higher for POC and  $C_a$  than it is for  $C_s$  and MAOC, whereas pedological variables are much more important for  $C_s$  and MAOC fractions. The importance of land cover for SOC with lower MRT has already been documented in several studies. For instance, Poeplau and Don (2013) observed that POC fractions are very sensitive to land cover in topsoils and Balesdent et al. (2018) showed that land cover is a major driver of the incorporation of “young” C in topsoils indicating on average a smaller portion of SOC with low MRT in croplands compared to grassland and forest topsoils. The fact that SOC with higher MRT is mostly driven by soil variables and that SOC with lower MRT is mainly explained by land cover and climate was also evidenced by Mathieu et al. (2015) using <sup>14</sup>C in 122 profiles of mineral soil across the world. They observed that the age of topsoil organic carbon, which is on average less biogeochemically stable than deep SOC, was primarily affected by climate and land cover whereas, the age of deep soil carbon was affected more by soil type and soil characteristics such as clay content and mineralogy. The little influence of land cover on the stable



fractions was somehow expected as most of French soils have undergone a series of land cover changes during the last millenia during which the stable fraction formed. Regarding the bulk SOC, our results are in line with those of Edlinger et al. (2023), who used a similar methodology on a smaller dataset to investigate the drivers of SOC. While their features were not strictly identical to ours (no iron oxyhydroxides for instance, but more climatic features), the main categories that stand out as drivers of the SOC are pedology (mostly exchangeable calcium) and climate, which is what we observed.

Among pedological variables, iron oxyhydroxides (mostly crystalline oxides) and exchangeable cations (mainly calcium cations) were the factors with the greatest influence on the size of the most stable  $C_s$  and MAOC fractions. The strong influence of exchangeable cations and oxides on MAOC has also been recently documented in a study involving 16 agricultural sites in the United States (King et al., 2023). The influence of iron oxides and hydroxides on SOC biogeochemical stability is a well-known fact (Kögel-Knabner et al., 2008), although it was pointed out that crystalline oxides were less efficient at providing bonding sites for SOM. The importance of exchangeable cations, notably calcium, on SOC biogeochemical stability was previously documented (Rowley et al., 2018; 2021). Indeed, calcium cations can strengthen the interactions between 2:1 clay minerals and SOM, both negatively charged, or enable the formation of co-precipitates with SOM.

Overall, our study confirms results from recent studies conducted at regional and global scales showing that physical fractions have different drivers (King et al., 2023 ; Hansen et al., 2024) and strongly supports the idea that it is relevant and informative to consider SOC fractions (either physical or thermal) instead of TOC alone. However, it is difficult to compare our results directly with those of these recent studies. Indeed, the data processing strategies are different and the explanatory variables considered are not the same. Notably, we observed that land cover and soil variables such as exchangeable calcium and iron oxides were explanatory variables of primary importance in explaining the quantities of POC and MAOC respectively. These variables were not considered as explanatory variables in the path analyses developed by Hansen et al. (2024) which likely explains the low explained variations provided by these path analyses. Conversely, we did not consider NPP as explanatory variable which was found to be a significant driver of MAOC quantities by Hansen et al. (2024). We therefore consider that new data that will probably arrive in the next few years will enable us to refine the drivers of physical and thermal fractions at different spatial scales.

#### **4.3 Which fractionation method to use to assess SOC biogeochemical stability in soil monitoring networks?**

Several recent high-level political initiatives have highlighted the importance of soil health for food security and climate change mitigation and adaptation. These include, for instance, the UNFCCC Koronivia joint work action, the 4p1000 initiative ([www.4p1000.org](http://www.4p1000.org), Rumpel et al., 2020) and the new “Soil Monitoring Law” proposed by the EU. These initiatives emphasize the importance of SOC by highlighting the C sequestration potential of soils, or by stressing the strong influence of SOC on soil health. This general political context is favourable to the development and support of soil monitoring networks. In these networks, SOC content is always measured. While this data is important, information on SOC biogeochemical stability would be particularly useful. Indeed, most soil functions related to SOM, such as nitrogen



400 mineralization, actually depend on its decay (Janzen, 2006) and assessing biogeochemical stability is also of primary importance to simulate SOC stock evolution (Luo et al., 2016). In this context, the development of indicators of SOC biogeochemical stability that can be implemented on large sample sets is of particular relevance.

Both physical and thermal fractionation methods are good candidates for this. Indeed, they both split SOC in two fractions of different biogeochemical stability using protocols that can be applied to large sample sets. Each has its advantages and drawbacks regarding its large-scale implementation for soil monitoring. The thermal method is faster (one hour per sample), highly reproducible (Pacini et al., 2023) and, at least in France, less costly than the physical method (< 50€ per sample vs > 100€ per sample in commercial laboratories). POC/MAOC fractionation, on the other hand, requires no expensive equipment, and is already used worldwide. Moreover some studies have proposed to predict POC/MAOC fractions using a machine learning model to make the method faster (Cotrufo et al., 2019; Lugato et al., 2021). However, a recent study 405 showed that the results of such prediction methods can be questionable and even misleading (Begill et al., 2023).

In this context, the question arises as to what method should be used to determine biogeochemical stability in soil monitoring networks. Our study, based on an unprecedented sample set, reveals that the POC vs  $C_a$  and MAOC vs  $C_s$  fractions are significantly different in size and do not have exactly the same environmental drivers, meaning that they are not biogeochemically equivalent. This suggests that the two fractionation methods provide, at least partly, different information, and could be, at that stage, be seen as complementary. Furthermore, they can also be used to answer different questions: For example, physical fractionation methods can be used in combination with isotopic measurements (e.g.  $^{13}C$ ), to study transformation and stabilisation processes of organic matter in soils (Cotrufo et al., 2015). Moreover, it is probably premature to assess the relevance of the two protocols at this stage, as interesting data will be provided by the monitoring networks over the next few years. For instance, with new SOC stock measurement campaigns, it will be possible to have 415 measurements of SOC stock evolution allowing the proper evaluation of model SOC stock projections at network scales (Le Noë et al., 2023). In addition, several indicators of soil functions are to be measured at large scale for soil health assessment. All these new data will enable us to assess the extent to which the information on SOC biogeochemical stability provided by fractionation results can be used to improve the accuracy of SOC stock evolution simulations, and to gain a better understanding of soil functioning. Such upcoming studies are likely to bring new key elements to the emerging question of 425 the redundancy or complementarity of physical and thermal fractionation schemes.

## 5 Conclusion

This study allowed us to compare the POC/MAOC physical fractionation and thermal fractionation on an unprecedented amount of samples with an interesting diversity with respect to pedological characteristics, climatic characteristics and land 430 covers. We showed that both the stable ( $C_s$  and MAOC) and labile ( $C_a$  and POC) fractions strongly differ in quantities. While the environmental drivers were close for the two stable fractions (respectively the two labile fractions) with a predominance of the soil characteristics (respectively the climate and land cover), they still presented differences suggesting that  $C_s$  and MAOC (respectively  $C_a$  and POC) correspond to different fractions with different biogeochemical stability. This



435 means that both fractionation techniques display different thus complementary information. Future work will enable us to discuss the relevance of one technique rather than the other on a case-by-case basis, depending on the soil properties studied.

### Code availability

The codes for the Random forest and the plots are available on Zenodo:

Stojanova, M., & Delahaie, A. (2024). SOC fractions drivers. Zenodo. <https://doi.org/10.5281/zenodo.10551240>.

### Data availability

440 Data on basic soil properties are freely available from the GIS Sol dataverse website:  
<https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/BNCXYB>.

### Author contributions

Amicie Delahaie, Lauric Cécillon, Pierre Barré, François Baudin and Claire Chenu designed the study. Florence Savignac and François Baudin produced the Rock-Eval® thermal analyses. Dominique Arrouays, Antonio Bispo, Line Boulonne,  
445 Claudy Jolivet, Manuel Martin, Céline Ratié and Nicolas Saby provided the detailed pedoclimatic data. Amicie Delahaie processed and interpreted the data with the contribution of all co-authors. Marija Stojanova wrote the Random Forest algorithm and provided technical support for the use of the code. Amicie Delahaie and Pierre Barré wrote the manuscript with contribution of all the co-authors.

### Competing interests

450 Some authors are members of the editorial board of SOIL.

### Acknowledgements

The École Normale Supérieure of Paris is greatly acknowledged for the funding of the PhD thesis grant of Amicie Delahaie. The ADEME (Rock-Eval®-RMQS project, convention n°2003C0017) is acknowledged for their support. Pierre Roudier is funded by the New Zealand Government to support the objectives of the Global Research Alliance on Agricultural  
455 Greenhouse Gases. The authors thank the FREACS project funded by the external call of the EJP Soil (ANR-22-SOIL-0001). We thank Sophie Cornu for helping us to recover the RMQS sample collection.



## 460 References

- Anderson, D. W., and Paul, E. A. (1984): Organo-mineral complexes and their study by radiocarbon dating, *Soil Science Society of America Journal*, 48(2), 298-301, <https://doi.org/10.2136/sssaj1984.03615995004800020014x>, 1984.
- Angers, D. A., Arrouays, D., Saby, N. P. A., and Walter, C.: Estimating and mapping the carbon saturation deficit of French agricultural topsoils, *Soil Use and Management*, 27(4), 448-452, <https://doi.org/10.1111/j.1475-2743.2011.00366.x>, 2011.
- 465 Angst, G., Mueller, K. E., Castellano, M. J., Vogel, C., Wiesmeier, M., and Mueller, C. W.: Unlocking complex soil systems as carbon sinks: multi-pool management as the key, *Nature Communications*, 14(1), 2967, <https://doi-org.insu.bib.cnrs.fr/10.1038/s41467-023-38700-5>, 2023.
- Balesdent, J., Mariotti, A., and Guillet, B.: Natural  $^{13}\text{C}$  abundance as a tracer for studies of soil organic matter dynamics, *Soil Biology and Biochemistry*, 19(1), 25-30, [https://doi.org/10.1016/0038-0717\(87\)90120-9](https://doi.org/10.1016/0038-0717(87)90120-9), 1987.
- 470 Balesdent, J., Pétraud, J. P., and Feller, C.: Effets des ultrasons sur la distribution granulométrique des matières organiques des sols, *Science du sol*, 29(2), 95-106, 1991.
- Balesdent, J.: The significance of organic separates to carbon dynamics and its modelling in some cultivated soils, *European Journal of soil science*, 47(4), 485-493, <https://doi.org/10.1111/j.1365-2389.1996.tb01848.x>, 1996.
- Balesdent, J., Besnard, E., Arrouays, D., and Chenu, C.: The dynamics of carbon in particle-size fractions of soil in a forest-  
475 cultivation sequence, *Plant and soil*, 201, 49-57, <https://doi.org/10.1023/A:1004337314970>, 1998.
- Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Derrien, D., Fekiacova, Z., and Hatté, C.: Atmosphere–soil carbon transfer as a function of soil depth, *Nature*, 559(7715), 599-602, <https://doi-org.insu.bib.cnrs.fr/10.1038/s41586-018-0328-3>, 2018.
- Barré, P., Plante, A. F., Cécillon, L., Lutfalla, S., Baudin, F., Bernard, S., Christensen, B. T., Eglin, T., Fernandez, J. M.,  
480 Houot, S., Kätterer, T., Le Guillou, C., Macdonald, A., van Oort, F., and Chenu, C.: The energetic and chemical signatures of persistent soil organic matter, *Biogeochemistry*, 130, 1-12, <https://doi.org/10.1007/s10533-016-0246-0>, 2016.
- Baudin, F., Disnar, J. R., Aboussou, A., and Savignac, F.: Guidelines for Rock–Eval analysis of recent marine sediments, *Organic Geochemistry*, 86, 71-80, <https://doi.org/10.1016/j.orggeochem.2015.06.009>, 2015.
- Begill, N., Don, A., and Poeplau, C.: No detectable upper limit of mineral-associated organic carbon in temperate  
485 agricultural soils, *Global Change Biology*, 29(16), 4662-4669, <https://doi.org/10.1111/gcb.16804>, 2023.
- Breiman, L.: Random forests, *Machine learning*, 45, 5-32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- Cambardella, C. A., and Elliott, E. T.: Particulate soil organic-matter changes across a grassland cultivation sequence, *Soil science society of America journal*, 56(3), 777-783, <https://doi.org/10.2136/sssaj1992.03615995005600030017x>, 1992.
- Cécillon, L., Baudin, F., Chenu, C., Houot, S., Jolivet, R., Kätterer, T., Lutfalla, S., Macdonald, A., van Oort, F., Plante, A.,  
490 F., Savignac, F., Soucémarianadin, L. N., and Barré, P.: A model based on Rock-Eval thermal analysis to quantify the size





- of the centennially persistent organic carbon pool in temperate soils, *Biogeosciences*, 15(9), 2835-2849,  
<https://doi.org/10.5194/bg-15-2835-2018>, 2018.
- Cécillon, L., Baudin, F., Chenu, C., Christensen, B. T., Franko, U., Houot, S., Kanari, E., Kätterer, T., Merbach, I., van Oort, F., Poeplau, C., Quezada, J. C., Savignac, F., Soucémariadin, L. N., and Barré, P.: Partitioning soil organic carbon into its  
495 centennially stable and active fractions with machine-learning models based on Rock-Eval® thermal analysis  
(PARTYSOCv2. 0 and PARTYSOCv2. 0EU), *Geoscientific Model Development*, 14(6), 3879-3898,  
<https://doi.org/10.5194/gmd-14-3879-2021>, 2021a.
- Cécillon, L. : A dual response, *Nature Geoscience*, 14(5), 262-263, Cécillon, L. : A dual response, *Nature Geoscience*, 14(5),  
262-263, <https://doi.org/insu.bib.cnrs.fr/10.1038/s41561-021-00749-6>, 2021b.
- 500 Chassé, M., Lutfalla, S., Cecillon, L., Baudin, F., Abiven, S., Chenu, C., and Barré, P.: Long-term bare-fallow soil fractions  
reveal thermo-chemical properties controlling soil organic carbon dynamics, *Biogeosciences*, 18(5), 1703-1718,  
<https://doi.org/10.5194/bg-18-1703-2021>, 2021.
- Chen, S., Arrouays, D., Angers, D. A., Martin, M. P., and Walter, C.: Soil carbon stocks under different land uses and the  
applicability of the soil carbon saturation concept, *Soil and Tillage Research*, 188, 53-58,  
505 <https://doi.org/10.1016/j.still.2018.11.001>, 2019.
- Clivot, H., Mouny, J. C., Duparque, A., Dinh, J. L., Denoroy, P., Houot, S., Vertès, F., Trochard, R., Bouthier, A., Sagot, S.,  
and Mary, B.: Modeling soil organic carbon evolution in long-term arable experiments with AMG model. *Environmental  
modelling and software*, 118, 99-113, <https://doi.org/10.1016/j.envsoft.2019.04.004>, 2019.
- Cotrufo, M. F., Soong, J. L., Horton, A. J., Campbell, E. E., Haddix, M. L., Wall, D. H., and Parton, W. J.: Formation of soil  
510 organic matter via biochemical and physical pathways of litter mass loss, *Nature Geoscience*, 8(10), 776-779,  
<https://doi.org/10.1038/ngeo2520>, 2015.
- Cotrufo, M. F., Ranalli, M. G., Haddix, M. L., Six, J., and Lugato, E.: Soil carbon storage informed by particulate and  
mineral-associated organic matter, *Nature Geoscience*, 12(12), 989-994, <https://doi-org.insu.bib.cnrs.fr/10.1038/s41561-019-0484-6>, 2019.
- 515 Cotrufo, M. F., Lavalley, J. M., Six, J., and Lugato, E.: The robust concept of mineral-associated organic matter saturation: A  
letter to Begill et al., 2023, *Global Change Biology*, 29(21), 5986-5987, <https://doi.org/10.1111/gcb.16921>, 2023.
- Delahaie, A. A., Barré, P., Baudin, F., Arrouays, D., Bispo, A., Boulonne, L., Chenu, C., Jolivet, C., Martin, M. P., Ratié,  
C., Saby, N. P. A., Savignac, F., and Cécillon, L.: Elemental stoichiometry and Rock-Eval® thermal stability of organic  
matter in French topsoils, *Soil*, 9(1), 209-229, <https://doi.org/10.5194/soil-9-209-2023>, 2023.
- 520 Disnar, J. R., Guillet, B., Kérais, D., Di-Giovanni, C., and Sebag, D.: Soil organic matter (SOM) characterization by Rock-  
Eval pyrolysis: scope and limitations, *Organic geochemistry*, 34(3), 327-343, [https://doi.org/10.1016/S0146-6380\(02\)00239-5](https://doi.org/10.1016/S0146-6380(02)00239-5), 2003.
- Edlinger, A., Garland, G., Banerjee, S., Degruene, F., García-Palacios, P., Herzog, C., Pescador, D. S., Romdhane, S., Ryo,  
M., Saghaï, A., Hallin, S., Maestre, F. T., Philippot, L., Rillig, M. C., and van Der Heijden, M. G.: The impact of agricultural



- 525 management on soil aggregation and carbon storage is regulated by climatic thresholds across a 3000 km European gradient, *Global Change Biology*, 29(11), 3177-3192, <https://doi.org/10.1111/gcb.16677>, 2023.
- Fox, J., and Weisberg, S.: An R Companion to Applied Regression, Third edition. Sage, Thousand Oaks CA.  
<https://socialsciences.mcmaster.ca/jfox/Books/Companion/>, 2019.
- Gogé, F., Joffre, R., Jolivet, C., Ross, I., and Ranjard, L.: Optimization criteria in sample selection step of local regression  
530 for quantitative analysis of large soil NIRS database. *Chemometrics and Intelligent Laboratory Systems*, 110(1), 168-176,  
<https://doi.org/10.1016/j.chemolab.2011.11.003>, 2012.
- Georgiou, K., Jackson, R. B., Vindušková, O., Abramoff, R. Z., Ahlström, A., Feng, W., Harden, J.W., Pellegrini, A. F. A.,  
Polley, H. W., Soong, J. L., Riley, W. J., and Torn, M. S.: Global stocks and capacity of mineral-associated soil organic  
carbon, *Nature communications*, 13(1), 3797, <https://doi-org.insu.bib.cnrs.fr/10.1038/s41467-022-31540-9>, 2022.
- 535 Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., and Tanabe, K.: 2006 IPCC guidelines for national greenhouse gas  
inventories, 2006.
- Hansen, P.S., Even, R., King, A.E., Lavallee, J., Schipanski, M., and Cotrufo, M.F.: Distinct, direct and climate-mediated  
environmental controls on global particulate and mineral-associated organic carbon storage, *Global Change Biology*, 30(1),  
<https://doi.org/10.1111/gcb.17080>, 2024.
- 540 Janzen, H. H.: The soil carbon dilemma: shall we hoard it or use it?, *Soil Biology and Biochemistry*, 38(3), 419-424,  
<https://doi.org/10.1016/j.soilbio.2005.10.008>, 2006.
- Jolivet, C., Boulonne, L., and Ratié, C.: Manuel du Réseau de Mesures de la Qualité des Sols, édition 2006, Unité InfoSol,  
INRA Orléans, France, 190 pp., ISBN: 2-73-80-1235-3, 2006.
- Jolivet, C., Almeida Falcon, J.-L., Berché, P., Boulonne, L., Fontaine, M., Gouny, L., Lehmann, S., Maître, B., Ratié, C.,  
545 Schellennberger, E., Soler-Dominguez, N.: French Soil Quality Monitoring Network Manual RMQS2: second metropolitan  
campaign 2016–2027. 2-7380-1451-8, <https://doi.org/10.17180/KC64-NY88>, 2022.
- Kanari, E., Cécillon, L., Baudin, F., Clivot, H., Ferchaud, F., Houot, S., Levvasseur, F., Mary, B., Soucémariadin, L. N.,  
Chenu, C., and Barré, P.: A robust initialization method for accurate soil organic carbon simulations. *Biogeosciences*, 19(2),  
375-387, <https://doi.org/10.5194/bg-19-375-2022>, 2022.
- 550 Kassambara, A. and Mundt, F.: Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package  
Version 1.0.7., <https://CRAN.R-project.org/package=factoextra>, 2020.
- Kassambara, A.: ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.6.0,  
<https://rpkgs.datanovia.com/ggpubr/>, 2023a.
- Kassambara, A.: rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2,  
555 <https://rpkgs.datanovia.com/rstatix/>, 2023b.
- King, A. E., Amsili, J. P., Córdova, S. C., Culman, S., Fonte, S. J., Kotcon, J., Liebig, M., Masters, M. D., McVay, K., Olk,  
D. C., Schipanski, M., Schneider, S. K., Stewart, C. E., and Cotrufo, M. F.: A soil matrix capacity index to predict mineral-



- associated but not particulate organic carbon across a range of climate and soil pH, *Biogeochemistry*, 165(1), 1–14, <https://doi.org/10.1007/s10533-023-01066-3>, 2023.
- 560 Kleber, M., Eusterhues, K., Keiluweit, M., Mikutta, C., Mikutta, R., and Nico, P. S.: Mineral–organic associations: formation, properties, and relevance in soil environments, *Advances in agronomy*, 130, 1-140, <https://doi.org/10.1016/bs.agron.2014.10.005>, 2015.
- Kögel-Knabner, I., Guggenberger, G., Kleber, M., Kandeler, E., Kalbitz, K., Scheu, S., Eusterhues, K., and Leinweber, P.: Organo-mineral associations in temperate soils: Integrating biology, mineralogy, and organic matter chemistry, *Journal of Plant Nutrition and Soil Science*, 171(1), 61-82, <https://doi.org/10.1002/jpln.200700048>, 2008.
- 565 Lavalley, J. M., Soong, J. L., and Cotrufo, M. F.: Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century, *Global change biology*, 26(1), 261-273, <https://doi.org/10.1002/jpln.200700048>, 2020.
- Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., Ciais, P., Chenu, C., Clivot, H., Derrien, D., 570 Ferchaud, F., Garnier, P., Goll, D., Lashermes, G., Martin, M., Rasse, D., Rees, F., Sainte-Marie, J., Salmon, E., Schiedung, M., Schimel, J., Wieder, W., Abiven, S., Barré, P., Cécillon, L., and Guenet, B.: Soil organic carbon models need independent time-series validation for reliable prediction, *Communications Earth and Environment*, 4(1), 158, <https://doi-org.insu.bib.cnrs.fr/10.1038/s43247-023-00830-5>, 2023.
- Louppe, G., Wehenkel, L., Sutura, A., and Geurts, P.: Understanding variable importances in forests of randomized trees, 575 *Advances in neural information processing systems*, 26, 2013.
- Louppe, G.: Understanding random forests: From theory to practice, arXiv preprint arXiv:1407.7502, <https://doi.org/10.48550/arXiv.1407.7502>, 2014.
- Lugato, E., Lavalley, J. M., Haddix, M. L., Panagos, P., and Cotrufo, M. F.: Different climate sensitivity of particulate and mineral-associated soil organic matter, *Nature Geoscience*, 14(5), 295-300, [https://doi-org.insu.bib.cnrs.fr/10.1038/s41561-021-00744-x](https://doi-org.insu.bib.cnrs.fr/10.1038/s41561-580 021-00744-x), 2021.
- Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E. A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J. W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R. B., Jones, C. D., Lara, M. J., Liang, J., McGuire, A. D., Parton, W., Peng, C., Randerson, J. T., Salazar, A., Sierra, C. A., Smith, M. J., Tian, H., Todd-Brown, K. E. O., Torn, M., Van Groenigen, K. J., Wang, Y. P., West, T. O., Wei, Y., Wieder, W. R., 585 Xia, J., Xu, X., Xu, X. and Zhou, T.: Toward more realistic projections of soil carbon dynamics by Earth system models, *Global Biogeochemical Cycles*, 30(1), 40-56, <https://doi.org/10.1002/2015GB005239>, 2016.
- Mathieu, J. A., Hatté, C., Balesdent, J., and Parent, É.: Deep soil carbon dynamics are driven more by soil type than by climate: a worldwide meta-analysis of radiocarbon profiles, *Global change biology*, 21(11), 4278-4292, <https://doi.org/10.1111/gcb.13012>, 2015.



- 590 Pacini, L., Adatte, T., Barré, P., Boussafir, M., Bouton, N., Cécillon, L., Lamoureux-Var, V., Sebag, D., Verrecchia, E.,  
Wattripont, A., and Baudin, F.: Reproducibility of Rock-Eval® thermal analysis for soil organic matter characterization,  
*Organic Geochemistry*, 186, 104687, <https://doi.org/10.1016/j.orggeochem.2023.104687>, 2023.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R.,  
Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: Machine  
595 learning in Python, *The Journal of machine Learning research*, 12, 2825-2830, 2011.
- Plante, A. F., Fernández, J. M., and Leifeld, J.: Application of thermal analysis techniques in soil science, *Geoderma*, 153(1-  
2), 1-10, <https://doi.org/10.1016/j.geoderma.2009.08.016>, 2009.
- Poeplau, C., and Don, A.: Sensitivity of soil organic carbon stocks and fractions to different land-use changes across Europe,  
*Geoderma*, 192, 189-201, <https://doi.org/10.1016/j.geoderma.2012.08.003>, 2013.
- 600 Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M. F., Derrien, D., Gioacchini, P., Grand, S.,  
Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L. M., Soong, J., Trigalet,  
S., Vermeire, M.-L., Rovira, P., van Wesemael, B., Wiesmeier, M., Yeasmin, S., Yevdokimov, I., and Nieder, R.: Isolating  
organic carbon fractions with varying turnover rates in temperate agricultural soils—A comprehensive method comparison,  
*Soil Biology and Biochemistry*, 125, 10-26, <https://doi.org/10.1016/j.soilbio.2018.06.025>, 2018.
- 605 Rowley, M. C., Grand, S., and Verrecchia, É. P.: Calcium-mediated stabilisation of soil organic carbon, *Biogeochemistry*,  
137(1-2), 27-49, <https://doi.org/10.1007/s10533-017-0410-1>, 2018.
- Rowley, M. C., Grand, S., Spangenberg, J. E., and Verrecchia, E. P.: Evidence linking calcium to increased organo-mineral  
association in soils, *Biogeochemistry*, 153(3), 223-241, <https://doi.org/10.1007/s10533-021-00779-7>, 2021.
- Rumpel, C., Amiraslani, F., Chenu, C., Garcia Cardenas, M., Kaonga, M., Koutika, L. S., Ladha, J., Madari, B., Shirato, Y.,  
610 Smith, P., Soudi, B., Soussana, J.-F., Whitehead, D., and Wollenberg, E.: The 4p1000 initiative: Opportunities, limitations  
and challenges for implementing soil organic carbon sequestration as a sustainable development strategy, *Ambio*, 49, 350-  
360, <https://doi.org/10.1007/s13280-019-01165-2>, 2020.
- Saenger, A., Cécillon, L., Sebag, D., and Brun, J. J.: Soil organic carbon quantity, chemistry and thermal stability in a  
mountainous landscape: A Rock-Eval pyrolysis survey, *Organic Geochemistry*, 54, 101-114,  
615 <https://doi.org/10.1016/j.orggeochem.2012.10.008>, 2013.
- Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I.,  
Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic  
matter as an ecosystem property, *Nature*, 478(7367), 49-56, <https://doi.org/insu.bib.cnrs.fr/10.1038/nature10386>, 2011.
- Schrumpf, M., Kaiser, K., Guggenberger, G., Persson, T., Kögel-Knabner, I., and Schulze, E. D.: Storage and stability of  
620 organic carbon in soils as related to depth, occlusion within aggregates, and attachment to minerals, *Biogeosciences*, 10(3),  
1675-1691, <https://doi.org/10.5194/bg-10-1675-2013>, 2013.

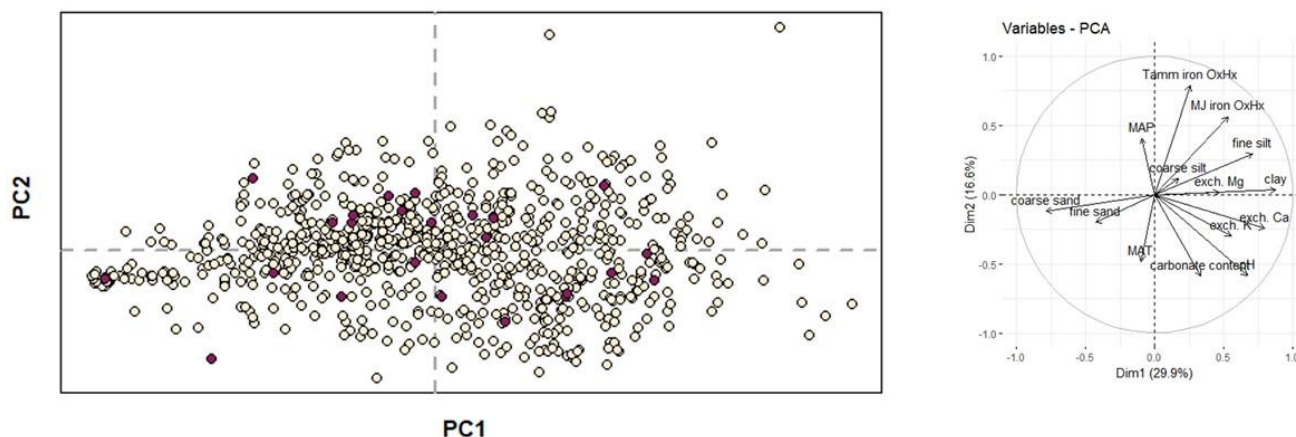


- Sebag, D., Verrecchia, E. P., Cécillon, L., Adatte, T., Albrecht, R., Aubert, M., Bureau, F., Cailleau, G., Copard, Y., Decaens, T., Disnar, J.-R., Hetényi, M., Nyilas, T., and Trombino, L.: Dynamics of soil organic matter based on new Rock-Eval indices, *Geoderma*, 284, 185-203, <https://doi.org/10.1016/j.geoderma.2016.08.025>, 2016.
- 625 Soetaert, K.: plot3D: Plotting Multi-Dimensional Data. (Version 1.4), <https://cran.rproject.org/web/packages/plot3D/plot3D.pdf>, 2021.
- Soucémariadin, L., Cécillon, L., Chenu, C., Baudin, F., Nicolas, M., Girardin, C., and Barré, P.: Is Rock-Eval 6 thermal analysis a good indicator of soil organic carbon lability?—A method-comparison study in forest soils, *Soil Biology and Biochemistry*, 117, 108-116, <https://doi.org/10.1016/j.soilbio.2017.10.025>, 2018.
- 630 Soucémariadin, L., Cécillon, L., Chenu, C., Baudin, F., Nicolas, M., Girardin, C., Delahaie, A. A., and Barré, P.: Heterogeneity of the chemical composition and thermal stability of particulate organic matter in French forest soils, *Geoderma*, 342, 65-74, <https://doi.org/10.1016/j.geoderma.2019.02.008>, 2019.
- Vinci Technologies (Nanterre, France): Geoworks V1.6R2, <https://www.vinci-technologies.com/rocks-and-fluids/geology/organic-geochemistry/geoworks-geochemical-software/113335/> (last access: 12 June 2023), 2021.
- 635 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., Vanderplas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedragosa, F., and Van Mulbregt, P.: SciPy 1.0: fundamental algorithms for scientific computing in Python, *Nature methods*, 17(3), 261-272, [https://doi-](https://doi.org/insu.bib.cnrs.fr/10.1038/s41592-019-0686-2)
- 640 [org.insu.bib.cnrs.fr/10.1038/s41592-019-0686-2](https://doi.org/insu.bib.cnrs.fr/10.1038/s41592-019-0686-2), 2020.
- von Lützw, M. V., Kögel-Knabner, I., Ekschmitt, K., Matzner, E., Guggenberger, G., Marschner, B., and Flessa, H.: Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions—a review, *European journal of soil science*, 57(4), 426-445, <https://doi.org/10.1111/j.1365-2389.2006.00809.x>, 2006.
- von Lützw, M., Kögel-Knabner, I., Ekschmitt, K., Flessa, H., Guggenberger, G., Matzner, E., and Marschner, B.: SOM
- 645 fractionation methods: relevance to functional pools and to stabilization mechanisms, *Soil biology and biochemistry*, 39(9), 2183-2207, <https://doi.org/10.1016/j.soilbio.2007.03.007>, 2007.
- Wei, T., Simko, V.: R package 'corrplot': Visualization of a Correlation Matrix. (Version 0.92), <https://github.com/taiyun/corrplot>, 2021.
- Wickham, H.: *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York. ISBN 978-3-319-24277-4,
- 650 <https://ggplot2.tidyverse.org>, 2016.
- Zimmermann, M., Leifeld, J., Schmidt, M. W. I., Smith, P., and Fuhrer, J.: Measured soil organic matter fractions can be related to pools in the RothC model, *European Journal of Soil Science*, 58(3), 658-667, <https://doi.org/10.1111/j.1365-2389.2006.00855.x>, 2007.

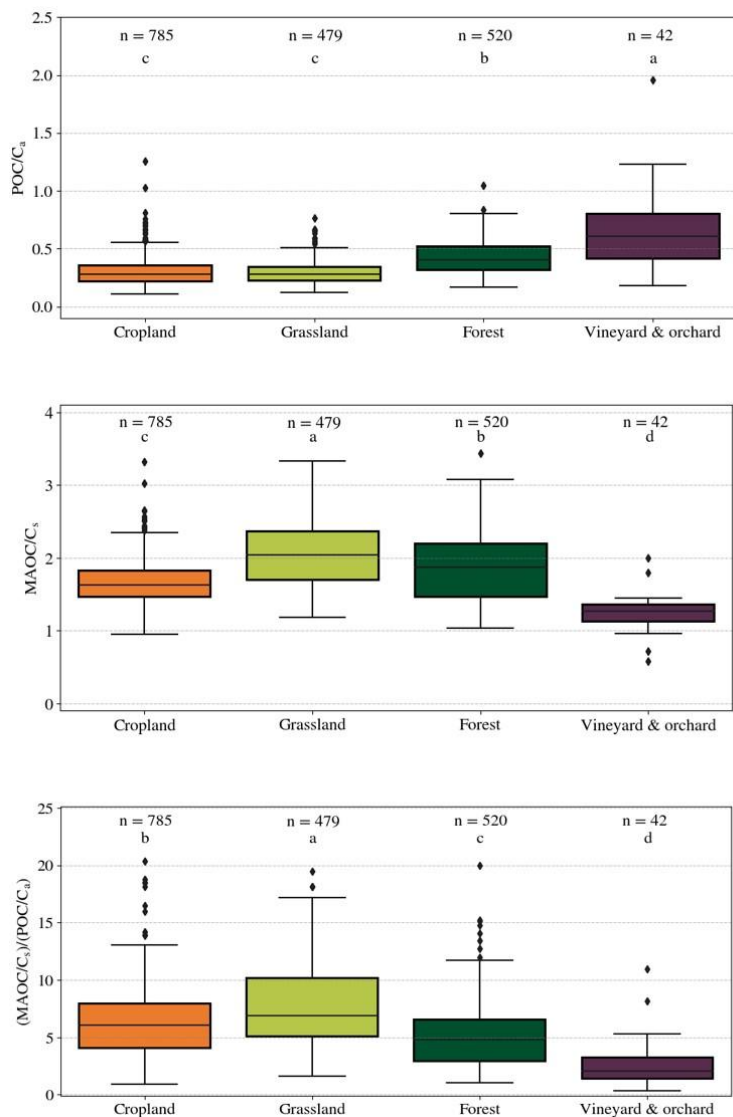


## Appendices

### PCA on 14 pedoclimatic parameters (n=993)



660 **Figure A1: Score of the 993 samples on axes 1 and 2 of the principal component analysis on 14 pedoclimatic parameters: clay, fine silt, coarse silt, fine sand and coarse sand contents; pH in water; carbonate content; mean annual temperature and mean annual precipitation; Tamm and Mehra–Jackson iron oxyhydroxide contents; exchangeable calcium, magnesium and potassium ions. Samples with an organic carbon yield between 0.7 and 1.3 are plotted in light yellow, whereas samples with an organic carbon yield <0.7 or >1.3 are plotted in dark red.**



665

670

**Figure A2: Ratios of the fractions for the four major land covers.** The black line in each box is the median, the lower and upper edges of the black rectangle are the respective first (Q1) and third (Q3) quartiles, and the lower and upper whiskers are the maximum between the minimum value or the first quartile minus 1.5 times the interquartile range ( $\max[\min; Q1-1,5 \times (Q3-Q1)]$ ) and the minimum between the maximum or the third quartile plus 1.5 times the interquartile range ( $\min[\max; Q3+1,5 \times (Q3-Q1)]$ ), respectively. Different letters indicate significant differences in the distribution of the values for the land covers according to a Kruskal–Wallis test ( $p < 0.05$ ) and a pairwise Wilcoxon rank sum test ( $p < 0.05$ ).



	<b>C<sub>a</sub> quantity</b>	<b>C<sub>s</sub> quantity</b>	<b>TOC</b>	<b>MAOC quantity</b>	<b>POC quantity</b>
<b>RF train accuracy</b>	0.639	0.707	0.667	0.690	0.668
<b>RF test accuracy</b>	0.574	0.616	0.583	0.525	0.352
<b>Number of estimators</b>	200	600	400	600	200
<b>Maximum number of features per tree</b>	0.33	0.33	0.33	0.33	0.33
<b>Maximum tree depth</b>	4	4	4	4	4
<b>Minimum number of samples per node</b>	5	3	5	5	3
<b>Minimum number of samples per leaf</b>	3	3	3	5	3

675 **Table A1: Hyper-parameters used by the Random Forest model for TOC and the four fractions of SOC.**





		C <sub>s</sub> content	C <sub>s</sub> content	C <sub>s</sub> proportion		MaOC content	POC content	POC proportion		C <sub>s</sub> /MaOC	MaOM - C <sub>s</sub>	POM/C <sub>s</sub>	C <sub>s</sub> - POM
clay	n=1,879	<b>0.52</b> ***	<b>0.23</b> ***	<b>0.15</b> ***	n=958	<b>0.39</b> ***	<b>0.18</b> ***	<b>-0.15</b> ***	n=843	<b>0.17</b> ***	<b>0.17</b> ***	<b>-0.06</b> b	<b>0.22</b> ***
total silt	n=1,879	<b>-0.04</b> b	<b>-0.12</b> ***	<b>0.17</b> ***	n=958	<b>-0.04</b> b	<b>-0.18</b> ***	<b>-0.37</b> ***	n=843	<b>-0.06</b> a	<b>-0.05</b> b	<b>-0.22</b> ***	<b>-0.12</b> ***
total sand	n=1,879	<b>-0.28</b> ***	<b>-0.04</b> a	<b>-0.22</b> ***	n=958	<b>-0.20</b> ***	<b>0.03</b> b	<b>0.36</b> ***	n=843	<b>-0.06</b> b	<b>-0.06</b> a	<b>0.20</b> ***	<b>-0.04</b> b
TOC <sub>ea</sub>	n=1,879	<b>0.91</b> ***	<b>0.98</b> ***	<b>-0.53</b> ***	n=960	<b>0.96</b> ***	<b>0.86</b> ***	<b>0.32</b> ***	n=843	<b>-0.27</b> ***	<b>0.82</b> ***	<b>0.01</b> ***	<b>0.90</b> ***
C/N	n=1,879	<b>0.11</b> ***	<b>0.29</b> ***	<b>-0.40</b> ***	n=959	<b>0.12</b> ***	<b>0.39</b> ***	<b>0.60</b> ***	n=843	<b>-0.03</b> b	<b>0.11</b> **	<b>0.30</b> ***	<b>0.17</b> ***
MAT 1969–1999	n=1,879	<b>-0.30</b> ***	<b>-0.45</b> ***	<b>-0.39</b> ***	n=960	<b>-0.38</b> ***	<b>-0.30</b> ***	<b>-0.08</b> ***	n=843	<b>0.33</b> ***	<b>-0.41</b> ***	<b>0.11</b> **	<b>-0.40</b> ***
MAP 1969–1999	n=1,879	<b>0.34</b> ***	<b>0.37</b> ***	<b>-0.27</b> ***	n=960	<b>0.37</b> ***	<b>0.29</b> ***	<b>0.14</b> ***	n=843	<b>-0.16</b> ***	<b>0.35</b> ***	<b>-0.02</b> b	<b>0.38</b> ***
pH	n=1,879	<b>0.15</b> ***	<b>-0.20</b> ***	<b>0.55</b> ***	n=960	<b>-0.12</b> ***	<b>-0.17</b> ***	<b>-0.21</b> ***	n=843	<b>0.43</b> ***	<b>-0.26</b> ***	<b>0.04</b> b	<b>-0.25</b> ***
InoC	n=1,879	<b>0.11</b> ***	<b>-0.03</b> b	<b>0.16</b> ***	n=958	<b>0.00</b> b	<b>-0.04</b> b	<b>-0.05</b> b	n=843	<b>0.14</b> ***	<b>-0.04</b> b	<b>-0.03</b> b	<b>-0.03</b> b
Tamm iron	n=1,610	<b>0.38</b> ***	<b>0.43</b> ***	<b>-0.25</b> ***	n=823	<b>0.51</b> ***	<b>0.28</b> ***	<b>-0.09</b> *	n=714	<b>-0.24</b> ***	<b>0.49</b> ***	<b>-0.16</b> ***	<b>0.49</b> ***
Mehra–Jackson iron	n=1,609	<b>0.49</b> ***	<b>0.27</b> ***	<b>0.06</b> *	n=822	<b>0.39</b> ***	<b>0.23</b> ***	<b>-0.05</b> b	n=713	<b>0.09</b> *	<b>0.22</b> ***	<b>0.03</b> b	<b>0.25</b> ***
CEC	n=1,879	<b>0.61</b> ***	<b>0.31</b> ***	<b>0.13</b> ***	n=960	<b>0.39</b> ***	<b>0.27</b> ***	<b>-0.05</b> b	n=843	<b>0.21</b> ***	<b>0.16</b> ***	<b>0.02</b> b	<b>0.21</b> ***
exch. Ca	n=1,879	<b>0.56</b> ***	<b>0.25</b> ***	<b>0.18</b> ***	n=960	<b>0.33</b> ***	<b>0.20</b> ***	<b>-0.08</b> *	n=843	<b>0.24</b> ***	<b>0.10</b> **	<b>0.02</b> b	<b>0.15</b> ***
exch. Mg	n=1,879	<b>0.25</b> ***	<b>0.15</b> ***	<b>0.04</b> a	n=960	<b>0.15</b> ***	<b>0.15</b> ***	<b>0.07</b> *	n=843	<b>0.15</b> ***	<b>0.02</b> b	<b>0.07</b> *	<b>0.07</b> *
exch. K	n=1,879	<b>0.10</b> ***	<b>-0.04</b> a	<b>0.18</b> ***	n=960	<b>0.03</b> b	<b>-0.08</b> *	<b>-0.21</b> ***	n=843	<b>0.15</b> ***	<b>-0.06</b> b	<b>-0.07</b> *	<b>-0.05</b> b
exch. Al	n=1,879	<b>0.11</b> ***	<b>0.39</b> ***	<b>-0.45</b> ***	n=960	<b>0.37</b> ***	<b>0.33</b> ***	<b>0.21</b> ***	n=843	<b>-0.33</b> ***	<b>0.43</b> ***	<b>-0.01</b> b	<b>0.45</b> ***
exch. Fe	n=1,879	<b>0.26</b> ***	<b>0.35</b> ***	<b>-0.22</b> ***	n=960	<b>0.28</b> ***	<b>0.31</b> ***	<b>0.17</b> ***	n=843	<b>-0.10</b> **	<b>0.25</b> ***	<b>0.08</b> *	<b>0.29</b> ***
exch. Na	n=1,879	<b>-0.01</b> b	<b>-0.02</b> a	<b>0.08</b> ***	n=958	<b>-0.02</b> b	<b>0.00</b> b	<b>0.09</b> **	n=843	<b>0.04</b> b	<b>0.01</b> b	<b>0.04</b> b	<b>0.02</b> b
exch. Mn	n=1,851	<b>-0.04</b> a	<b>0.09</b> ***	<b>-0.18</b> ***	n=947	<b>-0.04</b> b	<b>0.05</b> b	<b>0.12</b> ***	n=832	<b>-0.15</b> ***	<b>-0.00</b> b	<b>0.04</b> b	<b>-0.01</b> b

680 Table A2: Spearman correlation coefficients of the Ca content, Cs content, Cs proportion, POC content, MAOC content, POC proportion, Cs/MaOC, MAOC - Cs, POC/Ca, and Ca - POC with the following pedoclimatic variables  
 685 for the RMQS topsoil (0–30 cm) samples: clay, total silt, total sand, TOC<sub>ea</sub>, C/N ratio, mean annual temperature (MAT) averaged over 1969–1999, mean annual precipitation (MAP) averaged over 1969–1999, pH in water, carbonate content, Tamm iron oxyhydroxides (crystalline), Mehra–Jackson iron oxyhydroxides (amorphous), cation exchange capacity (CEC), exchangeable calcium, exchangeable magnesium, exchangeable potassium, exchangeable aluminium, exchangeable iron, exchangeable sodium, and exchangeable manganese. The analysis was limited to samples meeting the required criterion for Rock-Eval® thermal analysis and/or physical fractionation. The number of samples presenting data for each calculation is indicated. Absolute values ≥ 0.3 are in bold. The asterisks and superscript letters indicate the p value: \*\*\* between 0 and 0.001, \*\* between 0.001 and 0.01, \* between 0.01 and 0.05, a between 0.05 and 0.1, and b > 0.1. The values in grey have non significant p values.