







## RESEARCH ARTICLE OPEN ACCESS

# Influence of Soil Texture on the Estimation of Soil Organic Carbon From Sentinel-2 Temporal Mosaics at 34 European Sites

J. Wetterlind<sup>1</sup>  | M. Simmler<sup>2</sup>  | F. Castaldi<sup>3</sup> | L. Borůvka<sup>4</sup> | J. L. Gabriel<sup>5</sup> | L. C. Gomes<sup>6</sup>  | V. Khosravi<sup>4</sup>  | C. Kıvrak<sup>7</sup> | M. H. Koparan<sup>7</sup> | A. Lázaro-López<sup>5</sup> | A. Łopatka<sup>8</sup> | F. Liebisch<sup>9</sup>  | J. A. Rodriguez<sup>5</sup> | A. Ö. Savaş<sup>7</sup> | B. Stenberg<sup>1</sup> | T. Tunçay<sup>7</sup> | I. Vinci<sup>10</sup> | J. Volungevičius<sup>11</sup> | R. Žydelis<sup>11</sup> | E. Vaudour<sup>12</sup> 

<sup>1</sup>Swedish University of Agricultural Science (SLU), Department of Soil and Environment, Skara, Sweden | <sup>2</sup>Agroscope, Sustainability Assessment and Agricultural Management, Ettenhausen, Switzerland | <sup>3</sup>National Research Council of Italy (CNR) Institute of BioEconomy, Firenze, Italy | <sup>4</sup>Czech University of Life Sciences Prague (CZU), Faculty of Agrobiology, Food and Natural Resources, Department of Soil Science and Soil Protection, Prague, Czech Republic | <sup>5</sup>Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC), Madrid, Spain | <sup>6</sup>Aarhus University (AU), Department of Agroecology, Tjele, Denmark | <sup>7</sup>Republic of Turkey Ministry of Agriculture and Forestry, General Directorate of Agricultural Research and Policies (TAGEM), Ankara, Türkiye | <sup>8</sup>Institute of Soil Science and Plant Cultivation – National Research Institute (IUNG), Puławy, Poland | <sup>9</sup>Agroscope, Agroecology and Environment, Zürich, Switzerland | <sup>10</sup>Environmental Protection Agency of the Veneto Region (ARPAV), Treviso, Italy | <sup>11</sup>Lithuanian Research Centre for Agriculture and Forestry (LAMMC), Institute of Agriculture, Kėdainiai, Lithuania | <sup>12</sup>Université Paris-Saclay, INRAE, AgroParisTech, UMR EcoSys, Palaiseau, France

**Correspondence:** J. Wetterlind ([johanna.wetterlind@slu.se](mailto:johanna.wetterlind@slu.se))

**Received:** 28 June 2024 | **Revised:** 11 December 2024 | **Accepted:** 14 January 2025

**Funding:** This work was supported by European Union's Horizon 2020 Research and Innovation Programme, 862695.

**Keywords:** clay | field scale | remote sensing | satellite | SOC | soil moisture | time series

## ABSTRACT

Multispectral imaging satellites such as Sentinel-2 are considered a possible tool to assist in the mapping of soil organic carbon (SOC) using images of bare soil. However, the reported results are variable. The measured reflectance of the soil surface is not only related to SOC but also to several other environmental and edaphic factors. Soil texture is one such factor that strongly affects soil reflectance. Depending on the spatial correlation with SOC, the influence of soil texture may improve or hinder the estimation of SOC from spectral data. This study aimed to investigate these influences using local models at 34 sites in different pedo-climatic zones across 10 European countries. The study sites were individual agricultural fields or a few fields in close proximity. For each site, local models to predict SOC and the clay particle size fraction were developed using the Sentinel-2 temporal mosaics of bare soil images. Overall, predicting SOC and clay was difficult, and prediction performances with a ratio of performance to deviation (RPD) > 1.5 were observed at 8 and 12 of the 34 sites for SOC and clay, respectively. A general relationship between SOC prediction performance and the correlation of SOC and clay in soil was evident but explained only a small part of the large variability we observed in SOC prediction performance across the sites. Adding information on soil texture as additional predictors improved SOC prediction on average, but the additional benefit varied strongly between the sites. The average relative importance of the different Sentinel-2 bands in the SOC and clay models indicated that spectral information in the red and far-red regions of the visible spectrum was more important for SOC prediction than for clay prediction. The opposite was true for the region around 2200 nm, which was more important in the clay models.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *European Journal of Soil Science* published by John Wiley & Sons Ltd on behalf of British Society of Soil Science.

## Summary

- Study of how soil texture influences the performance of SOC prediction models using Sentinel-2 data.
- Local models were developed at 34 European sites with a large range of pedo-climatic conditions.
- There were large differences in prediction performance between sites for both SOC and clay content.
- Soil texture showed significant but weak and variable influence on SOC prediction performance.

## 1 | Introduction

Soil organic carbon (SOC) affects soil spectra in the visible region (Vis, 350–750 nm) by absorption of energy through excitation of electrons, and in the near infrared (NIR, 750–1100 nm) and short wave infrared (SWIR, 1100–2500 nm) through overtones and combinations of fundamental vibrations in the mid-infrared region (Stenberg et al. 2010). This has been successfully used for predicting SOC content under laboratory conditions with high spectral resolution instruments (e.g., Viscarra Rossel et al. 2006; Nocita et al. 2013; Barthès and Chotte 2021). Encouraging performances have also been obtained using field hyperspectral measurements (Gras et al. 2014; Wetterlind et al. 2015; Piccini et al. 2024), airborne and satellite hyperspectral images (Angelopoulou et al. 2019; Mzid et al. 2022), and multispectral satellite data on a local scale (e.g., Gholizadeh et al. 2018) or on a regional scale (Vaudour et al. 2019; Dvorakova et al. 2020; Urbina-Salazar et al. 2021).

Satellite remote sensing is frequently suggested as a possible tool for estimating and monitoring soil properties and soil health indicators on a large scale (Smith et al. 2020; Soubry et al. 2021; Pandey and Pandey 2023; Wang et al. 2023). The launch of the Sentinel-2 multispectral imaging satellites Sentinel-2A and Sentinel-2B in 2015 and 2017, respectively, brought an increased interest in using Earth observations to retrieve soil information, especially on SOC (Castaldi 2021; Vaudour et al. 2022; Yuzugullu et al. 2024). The spatial resolution of 10–20 m (10 of the 13 bands) makes the Sentinel-2 satellites relevant for applications guiding soil and crop management at the field and farm scale, and the short revisiting time of 5 days, or even less in areas with overlapping orbits further away from the Equator, increases the availability of images with bare soil conditions for gathering soil information.

However, the results of SOC prediction from satellite data are variable (Vaudour et al. 2022). Soil property mapping from satellite images is based on the interaction between sunlight and soil surfaces as detected by satellite-borne sensors. Many factors affect the reflectance received by the satellite sensors, including soil surface conditions (e.g., surface roughness, crop residues, and emerging plants), soil moisture, soil type, and soil texture. In this regard, satellite imagery with a short revisit time offers the possibility of collecting many bare soil images within a given time period. However, for satellite imagery with medium (10–30 m) or coarser spatial resolution, pure bare soil pixels are difficult to obtain. Therefore, over the last few years, some studies have focused on the influence of vegetation, either

green (Bartholomeus et al. 2011; Ouergemmi et al. 2016) or dry (Castaldi et al. 2019), to describe the decrease in performance due to vegetation coverage and provide spectral index thresholds for removing these effects. In addition to the vegetation index Normalised Difference Vegetation Index (NDVI), which is used to discard green vegetated pixels (e.g., Dematté et al. 2018; Loiseau et al. 2019), a now commonly used approach for removing the effect of dry vegetation consists of thresholding with the Normalised Burn Ratio 2 (NBR2) spectral index (Castaldi et al. 2019; Dvorakova et al. 2020). The effect of soil moisture on SOC prediction performance has mainly been studied using laboratory spectroscopy (Minasny et al. 2011; Nocita et al. 2013; Knadel et al. 2022; Metzger et al. 2024) or synthetic satellite spectra simulated from laboratory spectroradiometric measurements (Castaldi et al. 2015). More recently, radar-derived soil moisture information has also been considered (Urbina-Salazar et al. 2021, 2023; Zayani et al. 2023). Another approach reduces the varying effects of soil surface and retrieves a representative value of soil surface condition by considering a pixelwise median reflectance value over several years, instead of single-date information. This temporal mosaicking also allows the extension of the mappable bare soil area (Castaldi 2021; Heiden et al. 2022; Urbina-Salazar et al. 2023; Žižala et al. 2022).

Although soil texture is sometimes mentioned as a possible reason for deviations in SOC prediction results by Vis–NIR–SWIR spectroscopy (Stenberg 2010; Wight, Ashworth, and Allen 2016; Ogen et al. 2018), only a few studies have focused on disentangling the influence of soil texture on SOC prediction performance using satellite data (Khosravi et al. 2024). Therefore, the overall aim of this study was to better understand the spectral information used for modelling SOC from satellite data, focusing on the influence of soil texture. The study analyses the specific influence of texture on satellite-based SOC prediction instead of targeting the overall improvement of SOC prediction from texture, as achieved by Khosravi et al. (2024), with texture-based stratification of SOC models.

Particle size distribution, especially the clay fraction, strongly affects soil reflectance through clay minerals and the iron (hydr)oxide minerals contained in the clay fraction. Clay minerals absorb light at specific wavelengths, but the wavelength regions that hold information on clay minerals and organic molecules are located close together in the spectra and sometimes overlap (Stenberg et al. 2010; Wight, Ashworth, and Allen 2016; Ogen et al. 2018), with the risk of masking important information. This could be especially problematic when using low spectral resolution multispectral data, such as from the Sentinel mission. Soil texture also affects the visibility of SOC, with lower visibility of SOC in clayey soils, where SOC can be hidden in soil aggregates (Balesdent, Chenu, and Balabane 2000). Accordingly, Stenberg et al. (2010) found SOC overpredicted for sandy soils in a Swedish dataset. This was interpreted as an effect of the smaller specific surface area in sandy soil and, therefore, a higher concentration of organic matter on the exposed mineral particle surfaces. However, one of the strongest influences of soil texture related to reflectance in satellite data is probably its influence on soil moisture (Bousbih et al. 2019). Both organic matter and clay content increase soil water-holding capacity and soil moisture (Manns,

Parkin, and Martin 2016; Soltani et al. 2019). Water largely influences soil reflectance through strong absorbance at specific wavelengths (Babet et al. 2018; Soltani et al. 2019). However, these regions are not included in the multispectral bands registered by the satellites due to interference with water in the atmosphere, so the main effect is in the darkening of the soil with increased soil moisture.

The influence of soil texture on the possibility of predicting SOC using multispectral satellite data could be expected to be both favourable and unfavourable, depending on the nature of the relationship between SOC and soil texture. If both soil properties follow the same (or the opposite) spatial pattern, the combined effect on soil reflectance might help in modelling SOC. However, models relying on a relationship between SOC and soil texture might not generalise well over larger areas, as the SOC-to-soil texture relationship might change, and it will be difficult to detect changes in SOC content over time that are not correlated to soil texture. If, on the other hand, the variation in SOC does not follow the variability in soil texture, variability related to soil texture might hide a more subtle variation in SOC, causing difficulties in predicting SOC.

The goal of this study was to investigate the influence of soil texture on SOC predictions using Sentinel-2 temporal mosaics. The study therefore analysed how local within-site variability and correlations in SOC and soil texture influence the possibility of predicting SOC from satellite data using local models at sites in different pedo-climatic zones across Europe. Analyses of 34 individual sites in 10 European countries were carried out within the framework of the STEROPES project of the European Joint H2020 Programme, EJP SOIL. The study focused on the field and farm scale, developing site-specific models to further understand the basis for SOC models using satellite data and the role of soil texture in explaining variations in SOC modelling results. The field or farm scale has the advantage of reducing differences in management within the study sites, which might otherwise interfere with the interpretations. Whereas most previous studies aimed at predicting SOC typically considered a more reduced set of agropedological contexts and, most of the time, one study area within one country, this study relied on an unprecedented set of local-scale study areas from across Europe.

The study aimed to answer four specific questions related to the influence of soil texture on the estimation of soil organic carbon using Sentinel-2 temporal mosaics: (1) Does a stronger correlation between SOC and soil texture favour better SOC prediction performance? (2) Does the variation and overall amount of clay content affect SOC prediction performance? (3) Does adding information on soil texture as additional predictors to the models improve SOC predictions? and (4) Are different spectral bands important for predicting SOC compared with predicting clay?

## 2 | Materials and Methods

### 2.1 | Study Sites and Soil Datasets

To investigate the influence of soil texture on SOC models using satellite data, soil samples and satellite data were collected from 34 agricultural sites in 10 European countries

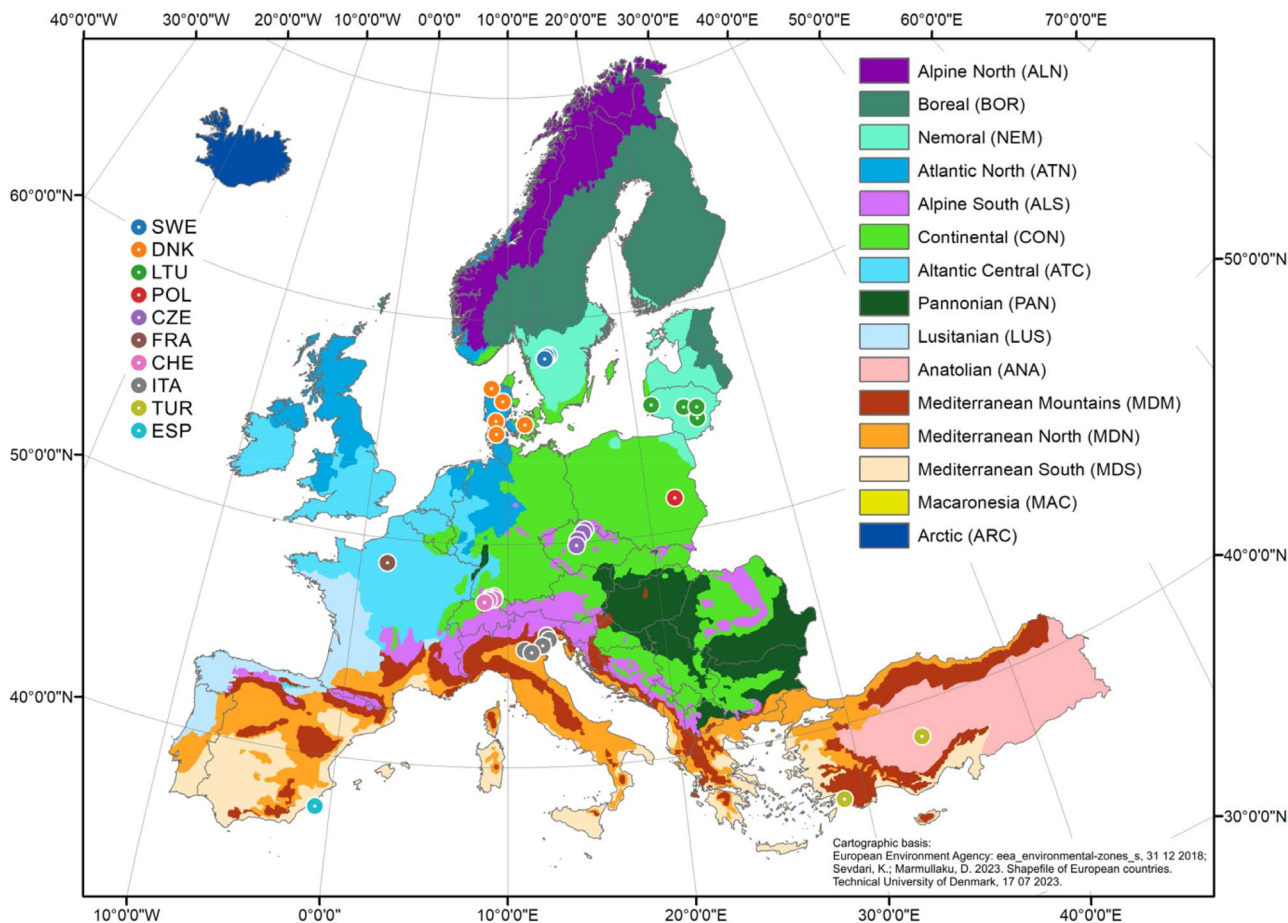
covering different pedoclimatic conditions and cropping systems (Figure 1, Table S1) by decreasing latitude from Sweden, Denmark, Lithuania, Poland, Czech Republic, France, Switzerland, Italy, Turkey, and Spain. The study sites were not selected to represent specific conditions but rather to serve as a variable sample of possible conditions to increase the external validity of the findings. The sites covered a range of soil types according to the FAO classification (IUSS Working Group WRB 2015), from Cambisols and Luvisols of the temperate Atlantic central and continental regions and Chernozems, Phaeozems, Histosols, Gleysols, and Arenosols in the Atlantic north and Nemoral regions to Vertisols, Fluvisols, Calcisols, and Gypsisols in the Anatolian and Mediterranean regions.

The sites were individual or neighbouring fields ranging from a few hectares to more than 450 ha at two sites, with a median size of 17 ha (Table S2). Soil samples were taken from 0 to 20 cm soil depths at all sites, except for the Italian sites, which were sampled at 0–30 cm, and the French and Spanish sites, which were sampled at 0–10 cm. The sampled area at each sampling point varied from 0.25 to 80 m<sup>2</sup>. About half of the sites used a target sampling design based on soil maps and/or satellite data, and the other half used a regular grid. The sampling design and methodology were usually the same for sites within the same country. The number of samples and sample density, however, also varied between individual sites from the same country, ranging from 16 to 280 samples per site and 0.2 to 60 samples ha<sup>-1</sup>. The median number of samples per site was 54, and the median sample density was 3 samples ha<sup>-1</sup>. The soil sampling dates also varied between the sites. Nine sites from Sweden and Denmark were sampled before 2015, with the oldest sampling from 1992 at one of the Danish sites. Most of the sites were sampled between 2018 and 2022. SOC contents and soil texture (clay, silt, and sand particle size fractions) of the soil samples were determined using the methods listed in Table S3. The silt-sand particle size limit was 50 μm in all countries, except Sweden, Lithuania, and Spain, where the limit was 63 μm. For these countries, soil texture data were therefore converted to values that corresponded to a silt-sand limit of 50 μm using a log-linear interpolation (see, e.g., Nemes et al. 1999).

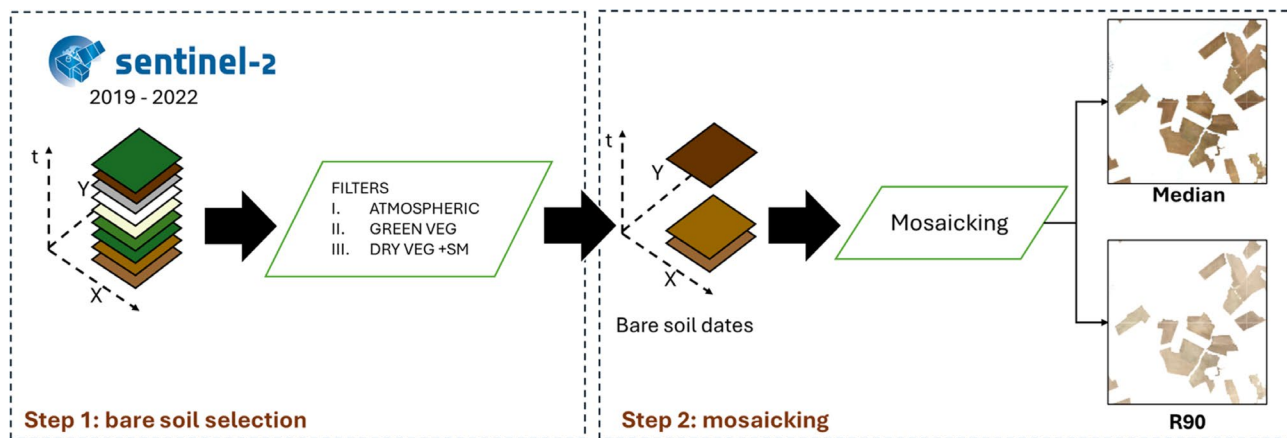
### 2.2 | Satellite Data

A pixelwise temporal mosaicking approach was applied to select bare soil images using Sentinel-2 image collections and to obtain composite bare soil layers for each region of interest (Castaldi et al. 2023) from the time series 2019–2022. The mosaicking approach was implemented by using the rgee R package (Aybar et al. 2020) to wrap the Earth Engine Python API for working in the Google Earth Engine environment (Gorelick et al. 2017) from within R (R Core Team 2024). Google Earth Engine is a cloud-based platform that was used here to gain access to Copernicus Sentinel-2 Multispectral Instrument (MSI) level 2A dataset provided by the European Space Agency (ESA).

The approach consisted of two steps (Figure 2). In the first step, three filters were applied for each pixel and acquisition date to identify and select bare soil pixels.



**FIGURE 1** | Map showing the location of the 34 sites included in the study and the environmental stratification of Europe (Metzger et al. 2005).



**FIGURE 2** | General flowchart of the two temporal mosaicking approaches (adapted from Castaldi et al. 2023). SM indicates soil moisture.

I. *Atmospheric filter.* This includes a cloud filter to remove dates affected by clouds and cloud shadows. The cs quality band from the combination of Cloud Score + and the S2\_HARMONIZED dataset was used with a cs threshold of 0.7 to remove thin clouds, haze, and cirrus shadows. The cs value is the similarity between the observed pixel and the theoretical clear reference pixel based on their spectral distance. A snow mask was also applied using the MSK\_SNOWPRB quality band provided by

ESA, removing records with > 10% probability of snow on the ground.

II. *Green vegetation filter.* The NDVI was calculated from bands 4 and 8 as follows:  $(B8 - B4)/(B8 + B4)$  was computed and pixels with NDVI > 0.35 were filtered out because they were likely affected by green vegetation.

III. *Dry vegetation + soil moisture filter.* The NBR2 index was calculated from bands 11 and 12 as  $(B11 - B12)/(B11 + B12)$

and pixels having  $NBR2 > 0.125$  (Castaldi et al. 2019) were filtered out because they were likely affected by dry vegetation and/or by high soil moisture content.

In the second step, if at least three bare soil dates were left after applying the filters for each pixel, the available data were combined using two mosaicking approaches to return bare soil composite images (temporal mosaics):

- A. The first approach involves selecting the *median* (50th percentile) reflectance values for each band and pixel of the time series, with the aim of obtaining spectral data representative of the average bare soil conditions not affected by extreme reflectance values.
- B. The second approach, called *R90*, involves selecting the 90th percentile of reflectance values of the time series for each band and pixel, with the aim of selecting dry conditions, assuming that higher reflectance corresponds to lower soil moisture content. Using the 90th percentile instead of the highest reflectance aims at excluding anomalously high reflectance values that could arise, for example, when cloud masking fails (Castaldi et al. 2023).

The outcome of these two mosaicking approaches was two composite bare soil layers of 10 m resolution for each study site. The layers include reflectance information of the following bands of the MSI instrument (central wavelength): B2 (492 nm), B3 (559 nm), B4 (664 nm), B5 (704 nm), B6 (740 nm), B7 (782 nm), B8 (842 nm), B8A (865 nm), B11 (1613 nm), and B12 (2202 nm; Sentinel-2 User Handbook 2015). Bands with a 20-m resolution (B5, B6, B7, B8A, B11, and B12) were resampled to 10 m. These layers were used to extract spectral data from soil sampling locations and then to develop soil property prediction models.

## 2.3 | Model Development and Validation

Local models for predicting SOC or clay content were developed for each site using partial least squares regression (PLSR) and random forest algorithms, as implemented in the Python machine learning framework scikit-learn v. 1.3.0 (Pedregosa et al. 2011). Predictors included the Sentinel-2 bare soil spectral information processed as described in the previous section (10 bands; median or *R90*) and/or, to predict SOC, information on the measured soil texture. Soil texture information from the soil samples was included either as the clay fraction or as information on all three particle size fractions (clay, silt, and sand). In the latter case, the three fractions were subjected to additive log ratio transformation to the two log ratios  $\log(\text{clay/sand})$  and  $\log(\text{silt/sand})$ . This transformation accounts for the compositional nature of the data (e.g., Filzmoser, Hron, and Templ 2018). Soil texture information was added as a predictor in the SOC prediction models to study the influence of soil texture on SOC prediction performance.

The optimal number of components to keep in the PLSR models was tuned based on the minimum root mean squared error (RMSE) in leave-one-out cross-validation (LOO-CV). Random forest models were fitted with default values for all

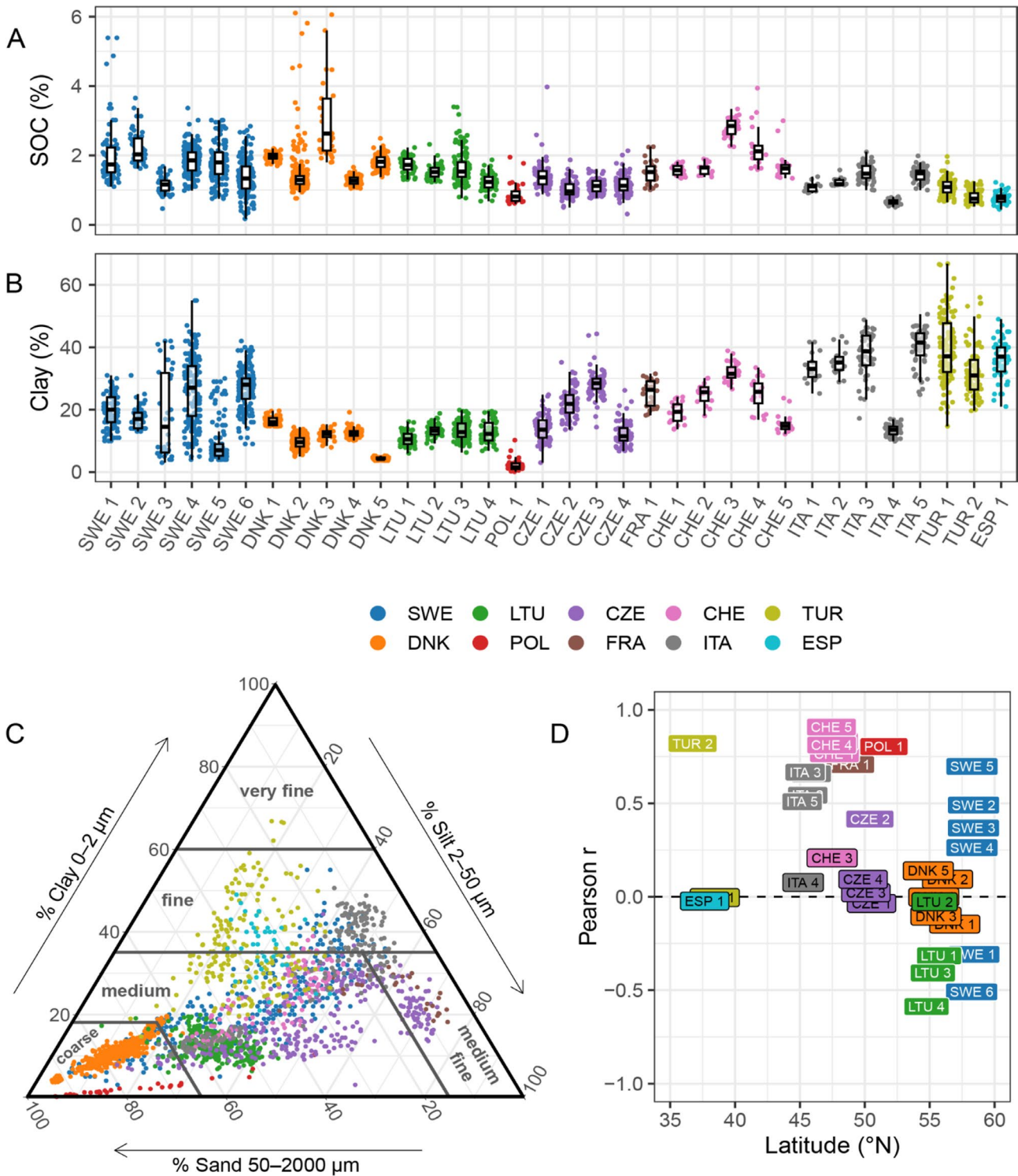
hyperparameters apart from the number of trees, which was set to 200 (no hyperparameter tuning was applied). Model performance was evaluated in LOO-CV using the RMSE, the ratio of performance to deviation ( $RPD = SD/RMSE$ ), the ratio of performance to interquartile distance ( $RPIQ = IQR/RMSE$ ), and Lin's concordance correlation coefficient (ccc). See Appendix 3 of Piikki et al. (2021) for the definition and interpretation of these metrics. As a simple baseline, we fitted 'dummy models' that always predict the mean of the target variable in the training data, that is, the mean of the actual measured SOC or clay content in the training data. These baseline models thus predict a constant value (calculated during model fitting) and use no predictors. The performance of these dummy models was evaluated using the same LOO-CV procedure and metrics as for the real models. To investigate the importance of the predictors, we calculated SHAP (SHapley Additive exPlanations) values using the Python package SHAP (Lundberg and Lee 2017). The importance of each predictor was calculated as the mean absolute SHAP value. Across predictors, these were normalised to sum to 100% to provide a measure of relative importance.

## 3 | Results

### 3.1 | SOC and Texture Data

Across all sites, SOC content ranged from 0.2% to 12%. The highest SOC levels (above 7.5%) and the largest range ( $> 6\%$ ) were recorded at two Danish sites (DNK 2 and 3) and one Swedish site (SWE 1; Figure 3A). Apart from these sites, large ranges in SOC content ( $> 2\%$ ) were also found at four of the other Swedish sites, and at one Czech (CZE 1), one Lithuanian (LTU 3), and one Swiss (CHE 4) site. The lowest average SOC contents were found in Spain, at a site in Italy (ITA 4), at a Turkish site (TUR 2), and at the Polish site (POL 1), all not exceeding 1%, which has been established as a critical level below which serious decline in soil quality might occur (desertification; Loveland and Webb 2003). Most other sites had average SOC content between 1% and 2%. Mainly, except for two Swiss sites (CHE 3 and 4), average SOC content tended to be lower in the southern countries. An opposite tendency was observed in clay content, except for the Swedish sites (Figure 3B).

Across all sites, clay content ranged from 0.1% to 67%. In general, the Swedish and Turkish sites had the largest values and the largest ranges of clay content. The average clay content ranged from 2% (Polish site, POL 1) to 40% (one of the Italian sites, ITA 5). Average contents  $> 30\%$  were found in some Italian, Spanish, Turkish, and Swiss sites. The range in clay content tended to asymptotically increase with the area of the sites (log-linear Pearson,  $r = 0.72$ ,  $p < 0.001$ ), while there was no statistical support for a similar relationship for the range in SOC ( $r = 0.05$ ,  $p = 0.78$ ). Texture classes across the 34 sites were quite varied and covered the entire texture triangle (Figure 3C), from coarse (such as the Danish sites) to fine (some Italian, Turkish, and Swedish sites). Silt contents across all sites ranged from 3% to 76%. The highest average contents (62% and 64%) were found at a Czech (CZE 2) and a French site (FRA 1). Across all sites, sand content ranged from 4% to 94%. The highest average sand

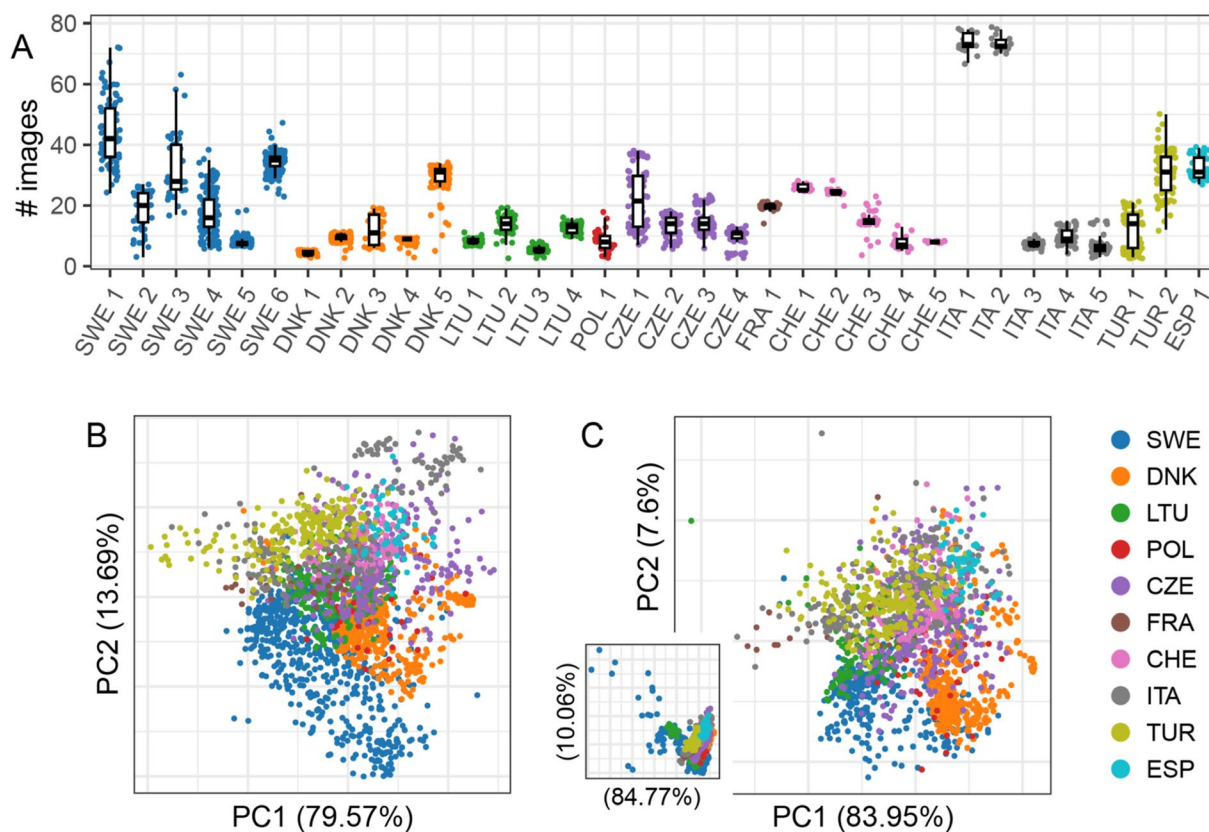


**FIGURE 3** | SOC and texture analyses. (A) Boxplots showing SOC content in the soil samples at each site. (B) Boxplots showing clay content in the soil samples at each site. Boxplots with lower and upper hinges representing 25th and 75th percentiles, line in between representing median, and whiskers extending to  $1.5 \times \text{IQR}$  from both hinges. Jittered points represent individual sampling points. For SWE 1, DNK 2, and DNK 3, a few points fall outside the range shown for SOC and are not visible. (C) Texture triangle showing individual sampling points coloured according to country. Texture classification is as used in the database of Hydraulic Properties of European Soils (HYPRES, Wösten et al. 1999). (D) Correlation between SOC and clay content per site. White text indicates a significant correlation (Pearson) at the 0.05 level.

contents, at above 70%, were found for Danish and Polish sites. SOC content showed positive correlations with clay content at 16 of the 34 sites, and negative correlations at 5 of 34 sites (Figure 3D). Interestingly, all sites with negative correlations were from the northernmost countries of Lithuania and Sweden.

### 3.2 | Satellite Data

The average number of bare soil images available per soil sampling point at the individual sites varied from 4 to 74 (Figure 4A). Two Italian sites stood out, with on average more than 70 bare



**FIGURE 4** | Satellite data. (A) Boxplot showing the number of available bare soil images per soil sampling point after applying the cloud and vegetation filters in the time series from 2019 to 2022. Boxplots with lower and upper hinges representing 25th and 75th percentiles, line in between representing median, and whiskers extending to  $1.5 \times \text{IQR}$  from both hinges. Jittered points represent individual sampling points. (B) Principal component analysis (PCA) of the median satellite data. (C) PCA of the R90 satellite data with (small inset) and without sites SWE 3, SWE 4, SWE 5, and LTU 3 (for these sites, the cloud filtering failed). Score plots representing the information in the first two principal components (PCs). The percentage of total variance in the data (10 satellite bands) explained by the PCs is given in parentheses. Each point in the score plot represents a sampling point.

soil images per point, and three Swedish, the Spanish, and one Turkish site had more than 30 images on average per point. However, the average number of bare soil images per point was below 20 for most of the sites. The minimum number of bare soil images for a sampling point of three images was observed for 22 points, mainly in Turkey, Denmark, Lithuania, and Poland. The similar clustering in the texture triangle in Figure 3C as in the PCA score plots of the satellite data in Figure 4B,C indicates an influence of soil texture on the satellite data. However, sites close in geography also tended to be close in the PCA score plot, suggesting influences of factors acting on a larger scale related to more general differences between the sites.

The PCA score plots in Figure 4B,C for median and R90 satellite data show very similar patterns (median and R90 temporal mosaics from three sites as an example are shown in Figure S1). However, at sites in two countries, Sweden and, to a lesser extent, Lithuania, using R90 resulted in several anomalously high reflectance values (resulting in the outliers in the inset in Figure 4C). Further inspection of the spectral data at these sampling points suggested that satellite images were affected by thin clouds that were not detected by the cloud mask. We thus excluded four sites (SWE 3, 4, and 5, and LTU 3) from further analyses using the R90 method.

### 3.3 | SOC and Clay Predictions Using Satellite Data

The performance of prediction models for both SOC and clay content varied largely between the sites and was poor in many cases (Table 1 and Figure 5 for PLSR, Table S4 and Figure S2 for random forest). Overall, PLSR models performed comparably with random forest models. As comparing the algorithms is not the aim of this study, we henceforth focus on the results of the PLSR models.

For each site and target variable (SOC or clay), we selected either the median or R90 satellite data as predictors, depending on which resulted in a model with a smaller RMSE. On average, across the 34 sites, there was no statistical support for a difference in prediction performance (RPD) for predicting SOC using satellite data compared to predicting clay using satellite data (Wilcoxon,  $p=0.39$ ; paired Wilcoxon,  $p=0.56$ ). The differences between the clay and SOC prediction performances varied greatly between the sites, with clay easier to predict at some sites and SOC on others. SOC content could be predicted with an RPD  $> 1.5$  at 8 sites, and with an RPD  $> 2.0$  at 2 sites. For SOC predictions, 8 sites (24%) had RPD values of  $< 1.1$ , which was only slightly better than the RPD of close to, or just below 1.0 of the dummy models, which are simple baseline models

**TABLE 1** | Performance of site-specific PLSR models for predicting SOC and clay as determined in leave-one-out cross-validation.

Site	SOC predictions												Clay predictions											
	Satellite				Satellite + texture				Texture				Satellite				Satellite							
	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc				
SWE 1	med	0.64	1.66	1.12	0.78	med	clay	0.65	1.65	1.12	0.78	allTex	1.04	1.02	0.69	0.14	med	3.43	1.54	2.33	0.75			
SWE 2	r90	0.47	0.97	1.35	-0.05	med	clay	0.43	1.05	1.47	0.29	clay	0.42	1.10	1.53	0.34	med	3.02	1.19	1.65	0.55			
SWE 3	med	0.27	0.97	1.08	0.09	med	clay	0.27	0.98	1.09	0.13	clay	0.26	1.02	1.13	0.15	med	4.67	2.95	5.46	0.94			
SWE 4	med	0.30	1.29	1.72	0.58	med	allTex	0.27	1.43	1.91	0.69	clay	0.38	1.02	1.36	0.10	med	5.83	1.84	2.74	0.83			
SWE 5	med	0.49	1.13	1.28	0.42	med	clay	0.33	1.69	1.90	0.79	clay	0.40	1.37	1.54	0.64	med	3.25	2.14	1.15	0.88			
SWE 6	med	0.34	1.58	1.89	0.76	med	clay	0.34	1.58	1.90	0.76	clay	0.46	1.15	1.38	0.40	r90	4.61	1.42	1.41	0.70			
DNK 1	med	0.08	1.45	1.83	0.71	med	clay	0.07	1.51	1.90	0.73	clay	0.12	0.96	1.21	-0.05	med	0.94	1.44	2.49	0.72			
DNK 2	med	1.15	1.50	0.20	0.72	med	allTex	1.03	1.68	0.23	0.79	allTex	1.71	1.00	0.14	0.05	med	1.30	1.33	2.07	0.62			
DNK 3	r90	1.01	1.61	1.48	0.77	p90	clay	1.03	1.59	1.46	0.76	allTex	1.44	1.13	1.04	0.43	med	1.64	1.04	1.15	0.21			
DNK 4	r90	0.11	1.18	1.63	0.50	p90	allTex	0.10	1.29	1.78	0.60	allTex	0.12	1.06	1.46	0.26	med	1.43	0.97	1.02	0.04			
DNK 5	med	0.18	1.15	1.52	0.42	p90	allTex	0.17	1.24	1.64	0.56	allTex	0.21	1.01	1.34	0.11	med	0.28	1.44	1.86	0.70			
LTU 1	med	0.19	1.24	1.65	0.58	med	allTex	0.16	1.48	1.98	0.72	allTex	0.17	1.34	1.78	0.62	med	2.02	1.08	1.64	0.45			
LTU 2	med	0.21	0.99	1.21	0.04	med	allTex	0.20	0.99	1.22	0.20	allTex	0.20	1.01	1.24	0.16	med	2.17	0.96	1.04	0.19			
LTU 3	med	0.42	1.17	1.04	0.49	med	allTex	0.36	1.35	1.19	0.64	allTex	0.37	1.31	1.16	0.60	med	2.71	1.06	1.59	0.26			
LTU 4	med	0.24	1.00	1.26	0.06	med	clay	0.20	1.18	1.48	0.48	clay	0.20	1.19	1.50	0.47	med	3.23	1.06	1.75	0.35			
POL 1	med	0.21	1.44	1.33	0.74	med	clay	0.19	1.61	1.48	0.79	clay	0.20	1.52	1.41	0.74	med	1.73	1.19	1.07	0.58			
CZE 1	med	0.39	1.06	0.97	0.24	med	clay	0.39	1.06	0.97	0.24	allTex	0.42	0.98	0.90	-0.03	r90	4.05	0.99	1.39	0.04			
CZE 2	med	0.17	1.34	1.70	0.63	med	allTex	0.17	1.39	1.77	0.67	allTex	0.21	1.11	1.41	0.37	med	2.05	2.30	2.77	0.90			
CZE 3	med	0.16	1.23	1.85	0.56	med	allTex	0.16	1.22	1.85	0.56	allTex	0.20	1.00	1.52	0.09	med	3.51	1.19	0.91	0.49			
CZE 4	r90	0.27	1.10	1.21	0.37	p90	clay	0.27	1.10	1.21	0.36	allTex	0.30	0.99	1.09	0.00	med	3.68	1.00	1.09	0.26			
FRA 1	r90	0.22	1.51	1.84	0.74	p90	clay	0.22	1.51	1.84	0.75	clay	0.25	1.31	1.59	0.61	med	1.74	2.34	4.50	0.90			
CHE 1	r90	0.10	1.44	2.42	0.73	p90	clay	0.10	1.55	2.59	0.78	clay	0.11	1.42	2.38	0.69	r90	2.36	1.36	2.24	0.68			
CHE 2	med	0.07	2.15	2.92	0.88	p90	clay	0.06	2.68	3.64	0.93	clay	0.10	1.61	2.19	0.78	med	2.31	1.43	1.90	0.75			
CHE 3	med	0.24	1.10	1.58	0.39	med	allTex	0.24	1.11	1.59	0.40	allTex	0.24	1.13	1.63	0.43	med	2.86	1.01	1.19	0.24			

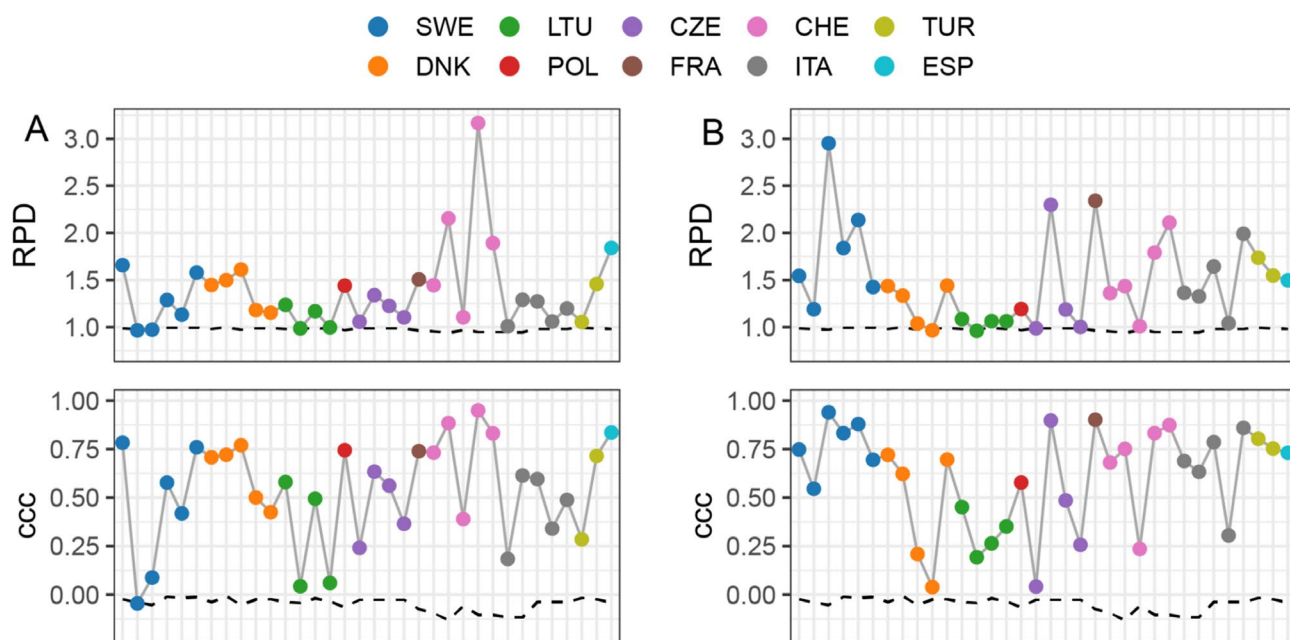
(Continues)



TABLE 1 | (Continued)

Site	SOC predictions												Clay predictions												
	Satellite						Satellite + texture						Texture						Satellite						
	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	RMSE	RPD	RPIQ	ccc	
CHE 4	r90	0.19	3.17	2.07	0.95	p90	clay	0.21	2.80	1.83	0.94	clay	0.39	1.54	1.00	0.75	med	2.82	1.79	2.27	0.83				
CHE 5	r90	0.21	1.89	1.13	0.83	med	clay	0.11	3.81	2.27	0.97	clay	0.20	2.04	1.21	0.86	r90	1.38	2.11	1.51	0.87				
ITA 1	r90	0.12	1.01	1.51	0.18	p90	clay	0.11	1.15	1.72	0.58	clay	0.11	1.17	1.75	0.50	r90	3.22	1.36	1.52	0.69				
ITA 2	r90	0.09	1.29	1.63	0.61	p90	allTex	0.07	1.57	1.97	0.78	allTex	0.10	1.10	1.39	0.43	med	3.01	1.33	1.40	0.63				
ITA 3	med	0.19	1.27	1.79	0.60	med	clay	0.15	1.55	2.19	0.76	allTex	0.18	1.30	1.84	0.60	r90	4.26	1.64	2.22	0.79				
ITA 4	med	0.06	1.06	1.55	0.34	med	clay	0.06	1.06	1.56	0.34	allTex	0.07	0.97	1.42	-0.04	r90	1.67	1.04	1.48	0.31				
ITA 5	med	0.16	1.20	1.55	0.49	med	allTex	0.16	1.20	1.55	0.49	allTex	0.17	1.13	1.46	0.39	med	3.04	1.99	2.34	0.86				
TUR 1	r90	0.21	1.05	1.40	0.28	p90	clay	0.21	1.09	1.45	0.31	allTex	0.23	0.99	1.32	0.03	med	6.17	1.74	2.54	0.80				
TUR 2	med	0.13	1.46	2.09	0.71	med	allTex	0.09	2.07	2.96	0.87	clay	0.11	1.71	2.45	0.80	r90	5.28	1.55	1.79	0.75				
ESP 1	med	0.09	1.84	2.00	0.84	med	allTex	0.09	1.09	2.06	0.85	allTex	0.14	1.14	1.24	0.43	med	4.30	1.50	1.80	0.73				
Median		0.21	1.26	1.55	0.58			0.20	1.46	1.75	0.71		0.21	1.13	1.40	0.42		2.94	1.39	1.70	0.70				
Mean		0.28	1.37	1.55	0.53			0.26	1.53	1.73	0.63		0.33	1.20	1.40	0.38		2.94	1.48	1.92	0.60				

Note: Predictors in the models for SOC include satellite data only, satellite data and texture information, or texture information only. The best models (smallest RMSE) using either the median or R90 satellite data, as well as using either clay only or information on all three granulometric fractions (clay, silt, sand; 'allTex') are listed.



**FIGURE 5** | (A) Performance of site-specific PLSR models for predicting SOC using satellite data. (B) Performance of site-specific PLSR models for predicting clay using satellite data. RPD and ccc were determined in leave-one-out cross-validation. The results are shown for the best models (smallest RMSE) using either the median or R90 satellite data (corresponding to Table 1). The dotted lines indicate the performance of the dummy models, which are simple baseline models, always predicting the mean of the measured SOC or clay content in the training data (ignoring the satellite information). The solid lines are only eye guides. The order of sites (x-axis) corresponds to Figures 3A,B and 4A.

that always predict the mean of the measured value in the training data (see methods). Clay content could be predicted with RPD values  $> 1.5$  at 12 of the 34 sites, and at 5 of those sites, the RPD was  $> 2.0$ . At 10 sites (29%), the models had RPD values  $< 1.1$ . At 10 sites, the models had RPIQ values  $> 1.7$  for both SOC and clay.

### 3.4 | Influence of Soil Texture

There was a significant positive correlation (Pearson  $r=0.42$ ,  $p<0.05$ ) between the performance (RPD) of the models predicting SOC using satellite data and the absolute value of the Pearson correlation coefficient  $|r|$  between SOC and clay in the soil. However, the correlation was weak, and for sites with a strong correlation between SOC and clay in soil, both good and poor SOC models were found (Figure 6). Examples of locations where the correlation between SOC and clay content seemed to be important were the two Turkish sites. Both sites had large variations in clay content and fairly good prediction models for clay. The variation in SOC was similar at the two sites, but only one of the sites (TUR 2) had significant and strong correlations between SOC and clay content (TUR 1  $r=0.00$ ,  $p=0.96$ ; TUR 2  $r=0.82$ ,  $p<0.001$ ), and it was at that site that SOC could be predicted considerably better (TUR 1 RPD 1.05; TUR 2 RPD 1.46).

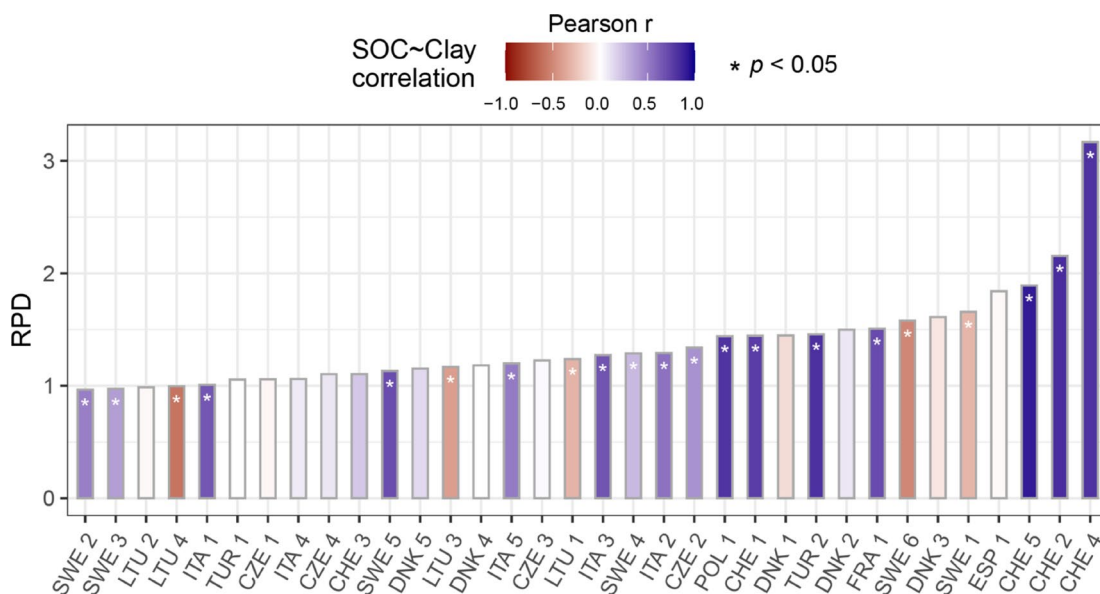
There was no statistical support for an overall correlation between SOC prediction performance (RPD) from satellite data and the average amount (Pearson,  $p=0.51$  and  $0.17$ ) or range ( $p=0.55$  and  $0.36$ ) of clay or SOC content, respectively. Similarly, there was no statistical support for an overall correlation between the clay prediction performance (RPD) from satellite data and the

average (Pearson,  $p=0.46$ ) and range ( $p=0.69$ ) of SOC. Neither did the clay prediction performance correlate with the average amount of clay content (Pearson,  $p=0.10$ ). However, there was a significant, but weak, correlation between the clay prediction performance and the range in clay content (Pearson  $r=0.48$ ,  $p<0.01$ ).

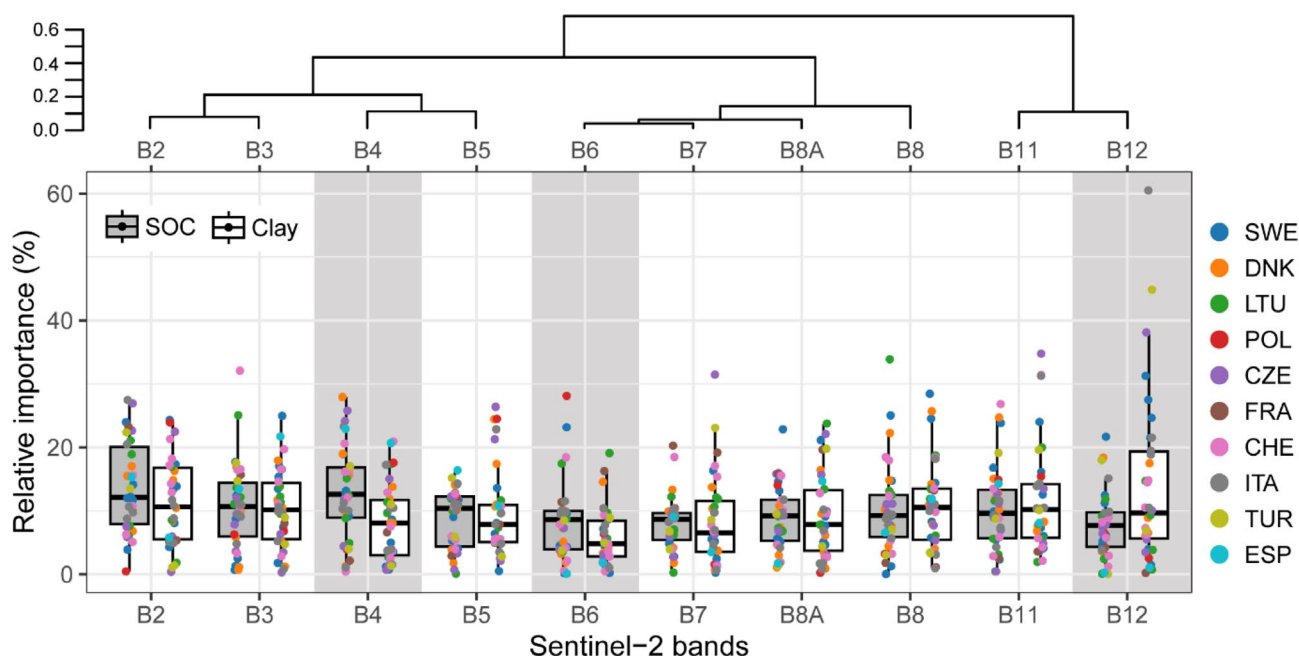
Soil texture information was added as a predictor in the SOC prediction models to study the influence of soil texture on SOC prediction performance. It was added as either clay or as information on all three particle size fractions, depending on which combination of predictors resulted in the lowest RMSE. On average, combining satellite data and soil texture data resulted in SOC prediction models with slightly, but statistically significant, higher RPD (mean difference of 0.16; paired Wilcoxon,  $p<0.001$ ), RPIQ (mean difference of 0.18;  $p<0.001$ ), and ccc (mean difference of 0.1;  $p<0.001$ ) values compared to using satellite data only (Table 1). There was a weak but statistically significant positive relationship between the change in RPD when adding soil texture to the SOC prediction models and the absolute value of the Pearson correlation coefficient  $|r|$  between SOC and clay in soil (Pearson  $r=0.45$ ,  $p<0.01$ ). However, the variability between the sites was large, and at several sites with a high correlation between SOC and clay in soil, the SOC prediction performance did not improve when soil texture was added as additional predictor.

### 3.5 | Relative Importance of Satellite Bands

The relative importance of the satellite bands in the PLSR models for predicting SOC and clay content is presented in Figure 7 (and more detailed in Figure S3). The relative importance of the



**FIGURE 6** | RPD values of SOC prediction PLSR models using satellite data (median or R90, whichever gave the smaller RMSE for each site; corresponding to Table 1). The sites are sorted according to the RPD values and coloured according to the correlation between SOC and clay content at the sites (significant correlations at the 0.05 level are indicated with an asterisk (\*)).



**FIGURE 7** | Relative importance of predictors in PLSR models for predicting SOC and clay using satellite data (median or R90, whichever resulted in the smallest RMSE for each site, corresponding to Table 1). Box plots with lower and upper hinges representing 25th and 75th percentiles, line in between representing median, and whiskers extending to  $1.5 \times \text{IQR}$  from both hinges. Jittered points represent individual sites. Shaded areas indicate the bands with importance that are significantly different in SOC models compared to clay models (paired Wilcoxon, 0.05 level). The dendrogram shows the redundancy structure of the set of predictors. It is based on a (complete-linkage) hierarchical clustering of the median satellite dataset, with the distance between the bands defined as  $1 - |r|$  (Pearson).

different bands varied largely between the models and sites. On average, across all sites, the visible range (especially the red to far-red wavelength range, B4 and B6) was more important in the SOC models compared with the clay models (paired Wilcoxon,  $p < 0.05$ ), whereas the wavelength range around 2200 nm (B12) was more important in the clay models (paired Wilcoxon,  $p < 0.05$ ).

## 4 | Discussion

### 4.1 | Availability and Quality of Satellite Data

Differences between satellite images from different dates can be considerable and influence model performances, both for SOC (Vaudour et al. 2019; Urbina-Salazar et al. 2021) and

for clay (Gomez et al. 2022). Using time series for individual fields to find the best possible or representative conditions can therefore be advantageous. In the present study, using satellite data from 4 years (2019 to 2022), however, resulted in less than on average 10 images per sampling point at 13 of the 34 sites, with 22 soil sampling points having as few as 3 images (Figure 4A). A low number of satellite images can result in less stable median and R90 data and thus less robust models when using these temporal mosaicking approaches. However, there was no statistical support for a correlation between the performance of the SOC and clay models and the number of available bare soil satellite images (Pearson  $p=0.68$  and  $p=0.72$  for SOC and clay, respectively).

A possible drawback of the mosaic is that it might disrupt otherwise continuous patterns within the fields, as possibly observed from a single-date shot, such as short time variations in soil moisture patterns. In general, the temporal mosaics using both median and R90 showed continuous and very similar spatial patterns (Figure S1 shows fields FRA 1, POL 1, and ESP 1). However, the R90 mosaic at the Polish site, POL 1, had a small area showing inconsistencies in the general pattern and sharper differences between satellite information from different dates. Although the small area in the Polish field did not include any of the sampling points and therefore did not affect the models, similar artefacts on other sites might be part of the explanation for the varying results when using temporal mosaics. At the French site, FRA 1, R90 enhanced the contrast between soils with different soil conditions compared with the median mosaic, highlighting the differences between the parts of the field with Calcaric Cambisols and the Luvisols, and a recently moved field boundary.

The selected method for collecting satellite data aimed to find the best possible bare soil conditions while using an automated procedure. The automated procedure made it easy to apply the same methods and criteria to multiple sites, which is important for use on a larger scale or by non-experts. However, the evident outliers at some of the sites using the R90 method despite the applied filters (inset in Figure 4C) highlight the difficulties with addressing all disturbing factors and the importance of validating the quality of the collected data.

## 4.2 | Between-Site Differences in Prediction Performance for SOC and Clay Content Using Satellite Data

The performance for predicting SOC from satellite data at the 34 sites included in the study varied largely, ranging from an RPD  $> 1.5$  for 8 sites to as low as  $< 1.1$  for another 8 sites (Table 1 and Figure 5). Similar variations in SOC prediction performance are also presented in a recent review when considering SOC prediction models from satellite data without additional auxiliary variables (Vaudour et al. 2022). A relationship between soil texture and the satellite data was evident when comparing the PCA score plot of the satellite data in Figure 4 with the texture triangle in Figure 3. However, it was obvious that it did not explain all the variations in the satellite data. Only at 10 sites the models for predicting clay from satellite data had an RPD above 1.5 (Table 1 and Figure 5).

The variable results in the present study correspond fairly well with previously published results on clay predictions and texture classification that also include both poor and good prediction models (e.g., Nanni and Dematté 2006; Gomez et al. 2018, 2019; Loiseau et al. 2019).

Clay content is a mineral particle size fraction, whereas its spectrally active components are related to clay and iron (hydr)oxide minerals, with different absorption patterns by different minerals (Hunt and Salisbury 1970). However, the narrow absorption features of different clay minerals cannot be detected with a multispectral sensor. Potentially detectable differences in the absorption spectra of the clay fraction could be related to differences in the proportions of clay minerals to iron (hydr)oxides, which absorb in different parts of the spectra (Stenberg et al. 2010). The models presented in this study were developed for field and farm scale, where variation in mineralogy could be expected to be small. However, different mineralogy between the fields might still be part of the explanation for the varying results between sites.

Several soil spectroscopy studies have shown the benefit of geographically small-scale calibration models due to the larger similarities in soil type and mineralogy (e.g., Kuang and Mouazen 2011; Araújo et al. 2014). In the present study with field scale models (the largest site being 1100ha and the median size 17ha) there was no statistical support for the prediction performance of SOC or clay models being related to the size of the sites. Rather than the geographic size per se, it is the variation in the modelled soil property which often explains large parts of the differences seen in model performance (Stenberg et al. 2010). A larger variation in the soil properties to be predicted is often desirable when developing prediction models (Ramirez-Lopez et al. 2014). In this study, there was a positive correlation between clay prediction performance and the range in clay content, but no statistical support for a correlation between range of SOC prediction performance and the range in SOC content. However, the three sites with the largest ranges in SOC content, SWE 1, DNK 2, and DNK 3, all had SOC models with RPD  $> 1.5$  (Figures 3A and 5 and Table 1).

## 4.3 | Influence of Soil Texture on SOC Predictions

A relationship of soil texture with the SOC predictions from satellite data was evident in the weak, but statistically significant, trend of increased SOC prediction performance with higher SOC to clay correlation in the soils of the 34 sites. However, it was obvious that this positive relationship could explain only a small part of the large variation in SOC prediction performance across the sites. The Danish sites DNK 1–3, for example, had comparably good SOC predictions but no significant correlation between SOC and clay content in soil (Figures 3D and 6). This might partly be explained by the high content and large variation in SOC, combined with among the lowest content and variation in clay. This makes variation in SOC the dominant soil property affecting water-holding capacity, and the main varying soil property affecting the soil spectra at those sites. In cases where variation in SOC is more pronounced than variation in texture, a strong correlation between SOC and clay content might instead help in clay predictions. It might be difficult to identify whether one or the other

of the two soil properties is dominant, but the correlation between SOC and clay content might likely support models for both soil properties, as may be the case at, for example, the Swiss sites CHE 2, 4, and 5 (Figures 3D and 6).

The influence of soil texture on the SOC predictions was also evident by the average improvement in the model performance when adding information on soil texture as predictors to the models, in addition to the satellite data. However, adding information on soil texture to prediction models for SOC is not practical, as soil texture analyses are both time consuming and expensive. Adding information on soil texture through available soil texture maps or other sensor measurements, such as gamma ray measurements using proximal or remote sensing, could be an alternative. In particular,  $^{232}\text{Th}$  concentration has been shown to be highly correlated with clay content (Piiikki and Söderström 2019) and adding airborne gamma ray measurements to a regional satellite-based model in central France improved SOC predictions (Urbina-Salazar et al. 2023).

Over all sites, the satellite bands in the red to far-red region of the visible spectrum (particularly B4, B6) were more important for SOC than for clay prediction, while the opposite was the case for the B12 band at around 2200nm (Figure 7). This is consistent with what has been reported as important spectral regions for SOC and clay predictions. Absorbance in the spectral region covered by Sentinel-2 B12, around 2200nm, is a dominant feature of several common clay minerals (including Gibbsite, Illite, Smectite, and Kaolinite; Stenberg et al. 2010). Although organic matter also absorbs light in this region, its absorbance is often reported to be weak. Evidence of the influence of organic matter in the visible region is well established, and soil colour has been used to estimate SOC in a number of situations (e.g., Gholizadeh et al. 2020).

#### 4.4 | Large and Variable Set of Study Sites

The large number of study sites with diverse pedo-climatic conditions from across Europe captured in this study made it possible to assess general patterns in performance of local (field- and farm-scale) models related to variations in SOC and soil texture. However, this diversity could also present challenges, as a wide range of conditions may have masked correlations between model performance and the SOC and soil texture relationship. The differences in SOC and clay content discussed earlier in this paper indicate differences in soil type. However, available data on dominant soil types (Table S1), diagnostic horizons of topsoils, and other related soil properties did not allow for a more thorough investigation.

In this study, there was no evidence of especially poor results at sites with a long time between soil sampling and the collection of satellite data. There was also no statistical support for a relationship between model performance and the number of samples (Pearson  $p=0.42$  and  $p=0.97$  for SOC and clay) and no or weak statistical support for a relationship between model performance and the sample density ( $p=0.37$  and  $p=0.06$ ,  $r=-0.33$ , for SOC and clay). Another difference between the sites in relation to the soil sampling was the differences in the areas that the individual soil samples represented. At several of the sites, the sampling was not designed for satellite modelling applications,

with sampling points unadjusted to the satellite pixels in terms of location inside the pixel or in terms of the size of the area sampled for each point. Sampling points from a very small area might, for example, lead to low representability of the 100 or 400m<sup>2</sup> satellite pixels. However, some of the best models were developed at sites where the individual soil sampling point represented only 0.25m<sup>2</sup> (CHE 2, 4, and 5), and no general patterns in model performance could be observed related to the area represented by a soil sample.

In soil remote sensing, it is generally assumed that soil property variations within the topsoil are minimised by repeated tillage operations. However, mixing depth and the degree of vertical homogenisation depend on the tillage system (Priori et al. 2024). In this study, 0–20 cm sampling depth was by far the most common (27 of the 43 sites), with only two sites with samples from 0 to 10cm (French and Spanish fields) and the 5 Italian fields with 0–30cm soil sampling (Table S2). The very unbalanced dataset and a lack of information on tillage practice do not allow a meaningful discussion of possible effects of sampling depth on model performance.

## 5 | Conclusion

This study investigated the influence of soil texture on SOC predictions from Sentinel-2 data with an unprecedented set of 34 local-scale study areas across 10 European countries. The prediction performance for both SOC and clay content varied largely between the sites when only satellite data was used as predictors, with RPD values ranging from around 1 to 3.2 and RPD values > 1.5 observed at 8 and 12 of the 34 sites for SOC and clay, respectively.

In response to the research questions, we draw the following conclusions:

- There was statistical support for better SOC prediction performance at sites with a stronger correlation between SOC and clay content in the soil. However, these relationships could explain only a small part of the overall variation in SOC model performances across the sites.
- Differences in the range and overall amount of clay content in soil could not explain the variation in SOC prediction performance between the sites.
- Adding information on soil texture as additional predictors to the models improved the SOC prediction on average, but, as with the modelling results, the improvement varied largely between the sites.
- The relative importance of the different spectral bands in the SOC and clay models indicated that the models for the two soil properties used different information in the satellite data to some degree, with the red and far-red regions of the visible spectrum more important for SOC prediction and the region around 2200nm more important for clay prediction.

The influence of soil texture on SOC prediction performance, although weak and variable across the studied sites, emphasises

the need for caution when interpreting SOC predictions from models relying only on satellite data and might explain some of the difficulties in expanding models to cover larger areas.

## Acknowledgements

The authors thank the local support teams taking the field data, and Tereza Zádorová for help with the soil classification. This study was produced within the activities of the STEROPES project (<https://ejpsoil.eu/soil-research/STEROPES>), of the EJP SOIL funded by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 862695.

## Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## References

- Angelopoulou, T., N. Tziolas, A. Balafoutis, G. Zalidis, and D. Bochtis. 2019. "Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review." *Remote Sensing* 11: 676. <https://doi.org/10.3390/rs11060676>.
- Araújo, S. R., J. Wetterlind, J. A. M. Demattê, and B. Stenberg. 2014. "Improving the Prediction Performance of a Large Tropical Vis-NIR Spectroscopic Soil Library From Brazil by Clustering Into Smaller Subsets or Use of Data Mining Calibration Techniques." *European Journal of Soil Science* 65: 718–729. <https://doi.org/10.1111/ejss.12165>.
- Aybar, C., Q. Wu, L. Bautista, R. Yali, and A. Barja. 2020. "rgee: An R Package for Interacting With Google Earth Engine." *Journal of Open Source Software* 5, no. 51: 2272. <https://github.com/r-spatial/rgee/>.
- Bablet, A., P. V. H. Vu, S. Jacquemoud, et al. 2018. "MARMIT: A Multilayer Radiative Transfer Model of Soil Reflectance to Estimate Surface Soil Moisture Content in the Solar Domain (400–2500 nm)." *Remote Sensing of Environment* 217: 1–17. <https://doi.org/10.1016/j.rse.2018.07.031>.
- Balesdent, J., C. Chenu, and M. Balabane. 2000. "Relationship of Soil Organic Matter Dynamics to Physical Protection and Tillage." *Soil and Tillage Research* 53: 215–230. [https://doi.org/10.1016/S0167-1987\(99\)00107-5](https://doi.org/10.1016/S0167-1987(99)00107-5).
- Barthès, B. G., and J. Chotte. 2021. "Infrared Spectroscopy Approaches Support Soil Organic Carbon Estimations to Evaluate Land Degradation." *Land Degradation and Development* 32: 310–322. <https://doi.org/10.1002/ldr.3718>.
- Bartholomeus, H., L. Kooistra, A. Stevens, et al. 2011. "Soil Organic Carbon Mapping of Partially Vegetated Agricultural Fields With Imaging Spectroscopy." *International Journal of Applied Earth Observation and Geoinformation* 13: 81–88. <https://doi.org/10.1016/j.jag.2010.06.009>.
- Bousbih, S., M. Zribi, C. Pelletier, et al. 2019. "Soil Texture Estimation Using Radar and Optical Data From Sentinel-1 and Sentinel-2." *Remote Sensing* 11: 1520. <https://doi.org/10.3390/rs11131520>.
- Castaldi, F. 2021. "Sentinel-2 and Landsat-8 Multi-Temporal Series to Estimate Topsoil Properties on Cropland." *Remote Sensing* 13: 3345. <https://doi.org/10.3390/rs13173345>.
- Castaldi, F., S. Chabrillat, A. Don, and B. van Wesemael. 2019. "Soil Organic Carbon Mapping Using LUCAS Topsoil Database and Sentinel-2 Data: An Approach to Reduce Soil Moisture and Crop Residue Effects." *Remote Sensing* 11, no. 18: 2121. <https://doi.org/10.3390/rs11182121>.
- Castaldi, F., M. H. Koparan, J. Wetterlind, et al. 2023. "Assessing the Capability of Sentinel-2 Time-Series to Estimate Soil Organic Carbon and Clay Content at Local Scale in Croplands." *ISPRS Journal of Photogrammetry and Remote Sensing* 199: 40–60. <https://doi.org/10.1016/j.isprsjprs.2023.03.016>.
- Castaldi, F., A. Palombo, S. Pascucci, S. Pignatti, F. Santini, and R. Casa. 2015. "Reducing the Influence of Soil Moisture on the Estimation of Clay From Hyperspectral Data: A Case Study Using Simulated PRISMA Data." *Remote Sensing* 7: 15561–15582. <https://doi.org/10.3390/rs71115561>.
- Demattê, J. A. M., C. T. Fongaro, R. Rizzo, and J. L. Safanelli. 2018. "Geospatial Soil Sensing System (GEOS3): A Powerful Data Mining Procedure to Retrieve Soil Spectral Reflectance From Satellite Images." *Remote Sensing of Environment* 212: 161–175. <https://doi.org/10.1016/j.rse.2018.04.047>.
- Dvorakova, K., P. Shi, Q. Limbourg, and B. Van Wesemael. 2020. "Soil Organic Carbon Mapping From Remote Sensing: The Effect of Crop Residues." *Remote Sensing* 12: 1913. <https://doi.org/10.3390/rs12121913>.
- Filzmoser, P., K. Hron, and M. Templ. 2018. *Applied Compositional Data Analysis*. Cham, Switzerland: Springer. <https://doi.org/10.1007/978-3-319-96422-5>.
- Gholizadeh, A., M. Saberioon, R. A. Viscarra Rossel, L. Boruvka, and A. Klement. 2020. "Spectroscopic Measurements and Imaging of Soil Colour for Field Scale Estimation of Soil Organic Carbon." *Geoderma* 357: 113972. <https://doi.org/10.1016/j.geoderma.2019.113972>.
- Gholizadeh, A., D. Žižala, M. Saberioon, and L. Borůvka. 2018. "Soil Organic Carbon and Texture Retrieving and Mapping Using Proximal, Airborne and Sentinel-2 Spectral Imaging." *Remote Sensing of Environment* 218: 89–103. <https://doi.org/10.1016/j.rse.2018.09.015>.
- Gomez, C., K. Adeline, S. Bacha, et al. 2018. "Sensitivity of Clay Content Prediction to Spectral Configuration of VNIR/SWIR Imaging Data, From Multispectral to Hyperspectral Scenarios." *Remote Sensing of Environment* 204: 18–30. <https://doi.org/10.1016/j.rse.2017.10.047>.
- Gomez, C., S. Dharumarajan, J. B. Féret, P. Lagacherie, L. Ruiz, and M. Sekhar. 2019. "Use of Sentinel-2 Time-Series Images for Classification and Uncertainty Analysis of Inherent Biophysical Property: Case of Soil Texture Mapping." *Remote Sensing* 11: 565. <https://doi.org/10.3390/rs11050565>.
- Gomez, C., E. Vaudour, J. B. Féret, F. De Boissieu, and S. Dharumarajan. 2022. "Topsoil Clay Content Mapping in Croplands From Sentinel-2 Data: Influence of Atmospheric Correction Methods Across a Season Time Series." *Geoderma* 423: 115959. <https://doi.org/10.1016/j.geoderma.2022.115959>.
- Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202: 18–27.
- Gras, J.-P., B. G. Barthès, B. Mahaut, and S. Trupin. 2014. "Best Practices for Obtaining and Processing Field Visible and Near Infrared (VNIR) Spectra of Topsoils." *Geoderma* 214–215: 126–134. <https://doi.org/10.1016/j.geoderma.2013.09.021>.
- Heiden, U., P. d'Angelo, P. Schwind, et al. 2022. "Soil Reflectance Composites—Improved Thresholding and Performance Evaluation." *Remote Sensing* 14: 4526. <https://doi.org/10.3390/rs14184526>.
- Hunt, G. R., and J. W. Salisbury. 1970. "Visible and Near-Infrared Spectra of Minerals and Rocks: I Silicate Minerals." *Modern Geology* 1: 283–300.
- IUSS Working Group WRB. 2015. "International Soil Classification System for Naming Soils and Creating Legends for Soil Maps." In *World Reference Base for Soil Resources 2014, Update 2015*. FAO, Rome: World Soil Resources Reports No. 106.
- Khosravi, V., A. Gholizadeh, D. Žižala, et al. 2024. "On the Impact of Soil Texture on Local Scale Organic Carbon Quantification: From Airborne to Spaceborne Sensing Domains." *Soil and Tillage Research* 241: 106125. <https://doi.org/10.1016/j.still.2024.106125>.

- Knadel, M., F. Castaldi, R. Barbetti, E. Ben-Dor, A. Gholizadeh, and R. Lorenzetti. 2022. "Mathematical Techniques to Remove Moisture Effects From Visible-Near-Infrared-Shortwave-Infrared Soil Spectra—Review." *Applied Spectroscopy Reviews* 58: 629–662. <https://doi.org/10.1080/05704928.2022.2128365>.
- Kuang, B., and A. M. Mouazen. 2011. "Calibration of Visible and Near Infrared Spectroscopy for Soil Analysis at the Field Scale on Three European Farms." *European Journal of Soil Science* 62: 629–636.
- Loiseau, T., S. Chen, V. L. Mulder, et al. 2019. "Satellite Data Integration for Soil Clay Content Modelling at a National Scale." *International Journal of Applied Earth Observation and Geoinformation* 82: 101905. <https://doi.org/10.1016/j.jag.2019.101905>.
- Loveland, P., and J. Webb. 2003. "Is There a Critical Level of Organic Matter in the Agricultural Soils of Temperate Regions: A Review." *Soil and Tillage Research* 70: 1–18. [https://doi.org/10.1016/S0167-1987\(02\)00139-3](https://doi.org/10.1016/S0167-1987(02)00139-3).
- Lundberg, S. M., and S. I. Lee. 2017. "A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems." In *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA: NIPS Papera.
- Manns, H. R., G. W. Parkin, and R. C. Martin. 2016. "Evidence of a Union Between Organic Carbon and Water Content in Soil." *Canadian Journal of Soil Science* 96: 305–316. <https://doi.org/10.1139/cjss-2015-0084>.
- Metzger, K., F. Liebisch, J. M. Herrera, T. Guillaume, and L. Bragazza. 2024. "Prediction Accuracy of Soil Chemical Parameters by Field- and Laboratory-Obtained Vis-NIR Spectra After External Parameter Orthogonalization." *Sensors* 24: 3556. <https://doi.org/10.3390/s24113556>.
- Metzger, M. J., R. G. H. Bunce, R. H. G. Jongman, C. A. Muecher, and J. W. Watkins. 2005. "A Climatic Stratification of the Environment of Europe." *Global Ecology and Biogeography* 14: 549–563. <https://doi.org/10.1111/j.1466-822x.2005.00190.x>.
- Minasny, B., A. B. McBratney, V. Bellon-Maurel, et al. 2011. "Removing the Effect of Soil Moisture From NIR Diffuse Reflectance Spectra for the Prediction of Soil Organic Carbon." *Geoderma* 167–168: 118–124. <https://doi.org/10.1016/j.geoderma.2011.09.008>.
- Mzid, N., F. Castaldi, M. Tolomio, S. Pascucci, R. Casa, and S. Pignatti. 2022. "Evaluation of Agricultural Bare Soil Properties Retrieval From Landsat 8, Sentinel-2 and PRISMA Satellite Data." *Remote Sensing* 14: 714. <https://doi.org/10.3390/rs14030714>.
- Nanni, M. R., and J. A. M. Demattê. 2006. "Spectral Reflectance Methodology in Comparison to Traditional Soil Analysis." *Soil Science Society of America Journal* 70: 393–407. <https://doi.org/10.2136/sssaj2003.0285>.
- Nemes, A., J. H. M. Wösten, A. Lilly, and J. H. Oude Voshaar. 1999. "Evaluation of Different Procedures to Interpolate Particle-Size Distributions to Achieve Compatibility Within Soil Databases." *Geoderma* 90: 187–202. [https://doi.org/10.1016/S0016-7061\(99\)00014-2](https://doi.org/10.1016/S0016-7061(99)00014-2).
- Nocita, M., A. Stevens, C. Noon, and B. Van Wesemael. 2013. "Prediction of Soil Organic Carbon for Different Levels of Soil Moisture Using Vis-NIR Spectroscopy." *Geoderma* 199: 37–42. <https://doi.org/10.1016/j.geoderma.2012.07.020>.
- Ogen, Y., C. Neumann, S. Chabrilat, N. Goldshleger, and E. Ben Dor. 2018. "Evaluating the Detection Limit of Organic Matter Using Point and Imaging Spectroscopy." *Geoderma* 321: 100–109. <https://doi.org/10.1016/j.geoderma.2018.02.011>.
- Ouerghemmi, W., C. Gomez, S. Naceur, and P. Lagacherie. 2016. "Semi-Blind Source Separation for the Estimation of the Clay Content Over Semi-Vegetated Areas Using VNIR/SWIR Hyperspectral Airborne Data." *Remote Sensing of Environment* 181: 251–263. <https://doi.org/10.1016/j.rse.2016.04.013>.
- Pandey, P. C., and M. Pandey. 2023. "Highlighting the Role of Agriculture and Geospatial Technology in Food Security and Sustainable Development Goals." *Sustainable Development* 31: 3175–3195. <https://doi.org/10.1002/sd.2600>.
- Pedregosa, F., G. Varoquaux, A. Gramfort, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12: 2825–2830.
- Piccini, C., K. Metzger, G. Debaene, et al. 2024. "In-Field Soil Spectroscopy in Vis-NIR Range for Fast and Reliable Soil Analysis: A Review." *European Journal of Soil Science* 75: e13481. <https://doi.org/10.1111/ejss.13481>.
- Piikki, K., and M. Söderström. 2019. "Digital Soil Mapping of Arable Land in Sweden—Validation of Performance at Multiple Scales." *Geoderma* 352: 342–350. <https://doi.org/10.1016/j.geoderma.2017.10.049>.
- Piikki, K., J. Wetterlind, M. Söderström, and B. Stenberg. 2021. "Perspectives on Validation in Digital Soil Mapping of Continuous Attributes—A Review." *Soil Use and Management* 37: 7–21.
- Priori, S., M. Zanini, V. Falcioni, and R. Casa. 2024. "Topsoil Vertical Gradient in Different Tillage Systems: An Analytical Review." *Soil and Tillage Research* 236: 105947. <https://doi.org/10.1016/j.still.2023.105947>.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Ramirez-Lopez, L., K. Schmidt, T. Behrens, B. van Wesemael, J. A. M. Demattê, and T. Scholten. 2014. "Sampling Optimal Calibration Sets in Soil Infrared Spectroscopy." *Geoderma* 226–227: 140–150. <https://doi.org/10.1016/j.geoderma.2014.02.002>.
- SENTINEL-2 User Handbook. 2015. "ESA Standard Document, Date 24/07/2015 Issue 1 Rev 2." [https://sentinel.esa.int/documents/247904/685211/Sentinel-2\\_User\\_Handbook](https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook).
- Smith, P., J. Soussana, D. Angers, et al. 2020. "How to Measure, Report and Verify Soil Carbon Change to Realize the Potential of Soil Carbon Sequestration for Atmospheric Greenhouse Gas Removal." *Global Change Biology* 26: 219–241. <https://doi.org/10.1111/gcb.14815>.
- Soltani, I., Y. Fouad, D. Michot, P. Bréger, R. Dubois, and C. Cudennec. 2019. "A Near Infrared Index to Assess Effects of Soil Texture and Organic Carbon Content on Soil Water Content." *European Journal of Soil Science* 70: 151–161. <https://doi.org/10.1111/ejss.12725>.
- Soubry, I., T. Doan, T. Chu, and X. Guo. 2021. "A Systematic Review on the Integration of Remote Sensing and GIS to Forest and Grassland Ecosystem Health Attributes, Indicators, and Measures." *Remote Sensing* 13: 3262. <https://doi.org/10.3390/rs13163262>.
- Stenberg, B. 2010. "Effects of Soil Sample Pretreatments and Standardised Rewetting as Interacted With Sand Classes on Vis-NIR Predictions of Clay and Soil Organic Carbon." *Geoderma* 158: 15–22. <https://doi.org/10.1016/j.geoderma.2010.04.008>.
- Stenberg, B., R. A. Viscarra Rossel, A. M. Mouazen, and J. Wetterlind. 2010. "Chapter Five—Visible and Near Infrared Spectroscopy in Soil Science." In *Advances in Agronomy*, edited by D. L. Sparks, 163–215. Academic Press. [https://doi.org/10.1016/S0065-2113\(10\)07005-7](https://doi.org/10.1016/S0065-2113(10)07005-7).
- Urbina-Salazar, D., E. Vaudour, N. Baghdadi, et al. 2021. "Using Sentinel-2 Images for Soil Organic Carbon Content Mapping in Croplands of Southwestern France. The Usefulness of Sentinel-1/2 Derived Moisture Maps and Mismatches Between Sentinel Images and Sampling Dates." *Remote Sensing* 13: 5115. <https://doi.org/10.3390/rs13245115>.
- Urbina-Salazar, D., E. Vaudour, A. C. Richer-de-Forges, et al. 2023. "Sentinel-2 and Sentinel-1 Bare Soil Temporal Mosaics of 6-Year Periods for Soil Organic Carbon Content Mapping in Central France." *Remote Sensing* 15: 2410. <https://doi.org/10.3390/rs15092410>.
- Vaudour, E., A. Gholizadeh, F. Castaldi, et al. 2022. "Satellite Imagery to Map Topsoil Organic Carbon Content Over Cultivated Areas: An Overview." *Remote Sensing* 14: 2917. <https://doi.org/10.3390/rs14122917>.

- Vaudour, E., C. Gomez, T. Loiseau, et al. 2019. "The Impact of Acquisition Date on the Prediction Performance of Topsoil Organic Carbon From Sentinel-2 for Croplands." *Remote Sensing* 11: 2143. <https://doi.org/10.3390/rs11182143>.
- Viscarra Rossel, R. A., D. J. J. Walvoort, A. B. McBratney, L. J. Janik, and J. O. Skjemstad. 2006. "Visible, Near Infrared, Mid Infrared or Combined Diffuse Reflectance Spectroscopy for Simultaneous Assessment of Various Soil Properties." *Geoderma* 131: 59–75. <https://doi.org/10.1016/j.geoderma.2005.03.007>.
- Wang, J., J. Zhen, W. Hu, et al. 2023. "Remote Sensing of Soil Degradation: Progress and Perspective." *International Soil and Water Conservation Research* 11: 429–454. <https://doi.org/10.1016/j.iswcr.2023.03.002>.
- Wetterlind, J., K. Piikki, M. Söderström, and B. Stenberg. 2015. "Exploring the Predictability of Soil Texture and Organic Matter Content With a Commercial Integrated Soil Profiling Tool." *European Journal of Soil Science* 66: 631–638.
- Wight, J. P., A. J. Ashworth, and F. L. Allen. 2016. "Organic Substrate, Clay Type, Texture, and Water Influence on NIR Carbon Measurements." *Geoderma* 261: 36–43. <https://doi.org/10.1016/j.geoderma.2015.06.021>.
- Wösten, J. H. M., A. Lilly, A. Nemes, and C. Le Bas. 1999. "Development and Use of a Database of Hydraulic Properties of European Soils." *Geoderma* 90: 169–185. [https://doi.org/10.1016/S0016-7061\(98\)00132-3](https://doi.org/10.1016/S0016-7061(98)00132-3).
- Yuzugullu, O., N. Fajraoui, A. Don, and F. Liebisch. 2024. "Satellite-Based Soil Organic Carbon Mapping on European Soils Using Available Datasets and Support Sampling." *Science of Remote Sensing* 9: 100118. <https://doi.org/10.1016/j.srs.2024.100118>.
- Zayani, H., Y. Fouad, D. Michot, et al. 2023. "Using Machine-Learning Algorithms to Predict Soil Organic Carbon Content From Combined Remote Sensing Imagery and Laboratory Vis-NIR Spectral Datasets." *Remote Sensing* 15: 4264. <https://doi.org/10.3390/rs15174264>.
- Žížala, D., R. Minařík, J. Skála, et al. 2022. "High-Resolution Agriculture Soil Property Maps From Digital Soil Mapping Methods, Czech Republic." *Catena* 212: 106024. <https://doi.org/10.1016/j.catena.2022.106024>.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.