Author(s):	Raisa Mäkipää, Lorenzo Menichetti, Eduardo Martínez-García, Tiina Törmänen, Aleksi Lehtonen
Title:	Is the organic carbon-to-clay ratio a reliable indicator of soil health?
Year:	2024
Version:	Published version
Copyright:	The Author(s) 2024
Rights:	CC BY 4.0
Rights url:	https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Raisa Mäkipää, Lorenzo Menichetti, Eduardo Martínez-García, Tiina Törmänen, Aleksi Lehtonen, Is the organic carbon-to-clay ratio a reliable indicator of soil health?, Geoderma, Volume 444, 2024, 116862, ISSN 0016-7061, https://doi.org/10.1016/j.geoderma.2024.116862.

Geoderma 444 (2024) 116862

Contents lists available at ScienceDirect

Geoderma

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

Is the organic carbon-to-clay ratio a reliable indicator of soil health?

Raisa Mäkipää, Lorenzo Menichetti *, Eduardo Martínez-García, Tiina Törmänen, Aleksi Lehtonen

Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-00790 Helsinki, Finland

ARTICLE INFO

Handling Editor: L. Morgan Cristine

Keywords:
Soil organic carbon (SOC)
SOC:Clay ratio
European mineral soils
LUCAS soil survey
Forest soil, agricultural soil
Soil monitoring

ABSTRACT

Climate action plans under the Paris Climate Agreement and other national commitments aimed at improving soil-based ecosystem services require the operational monitoring of soil carbon (C). The European Union is aiming to enhance soil health, and as part of the proposed Soil Monitoring Law, the European Commission recommends the monitoring of the soil C loss indicator among other soil health indicators. In this study, we evaluate the feasibility of the proposed soil C loss indicator by assessing its performance using the EU-wide 2009 LUCAS soil survey data. The proposed indicator is the soil organic carbon (SOC) to clay ratio, with a threshold value of 1:13. The results are also compared with the C stock changes reported by countries to the climate convention (UNFCCC). Our results reveal that the variation in SOC and clay content at European scale exceeds that of the data used to develop the proposed indicator. We also found that the variation in the SOC content was influenced not only by clay content but also by climate and land-use reflecting C input levels. Therefore, the defined threshold is inadequate for detecting degraded soils if the SOC and clay content are beyond the conditions used to establish the criteria. Furthermore, major discrepancies were observed between the soil carbon stock changes reported by the national greenhouse gas (GHG) inventories and the proportions of degraded soils identified by using the soil C loss indicator. We conclude that employing a single indicator such as SOC:Clay ratio with one threshold value for all soils across various land covers, management practices, and climatic conditions, as defined by the European Commission for the Soil Monitoring Law, is inappropriate for monitoring soil C loss.

1. Introduction

Anthropogenic activities have considerably depleted initial soil organic carbon (SOC) stocks, thus directly and negatively affecting soil health, its functioning, and capacity to deliver ecosystem services such as food security, water security, climate regulation or biodiversity protection (Kopittke et al., 2023). Soil health refers to ‘the ability of the soil to sustain the productivity, diversity, and environmental services of terrestrial ecosystems’ (FAO ITPS, 2020). Yet, the concept remains elusive and cannot be quantitatively measured (Janzen et al., 2021). As a result, considerable efforts have concentrated on identifying a set of measurable biological, chemical, and/or physical indicators to facilitate the assessment of soil health. These indicators should represent soil processes, properties, and ecosystem services, but they must also be sensitive to management and be feasible for practical use (Doran and Zeiss, 2000). In this context, the SOC content is the most widely used indicator (Bünemann et al., 2018). In developing a plan for restoring soil health, the loss of SOC has been considered as one of the measures of soil degradation by the European Commission in the proposed Soil Monitoring Law, which aims to turn all soils within the European Union into healthy state by 2050 (Panagos et al., 2022). In the proposed law, the SOC content and SOC:Clay ratio were suggested to be an indicator of the soil C loss (European Commission, 2023).

Monitoring of the soil C serves both international climate policy and national targets set for soil productivity and C stocks. Currently, monitoring of the SOC stock changes is a part of the national greenhouse gas inventories (IPCC, 2006; European Environment Agency, 2023) and countries may have included soil C sequestration in their climate action plans as set in the Paris Climate Agreement (UNFCCC, 2023). Loss of SOC on mineral soils can be measured by repeated sampling (e.g., Grüneberg et al., 2014; Ortiz et al., 2013). However, most countries use modeling to report changes in SOC stocks in their national greenhouse gas (GHG) inventories following IPCC guidelines (IPCC, 2019), since repeated measurements for soil carbon monitoring are both labor-intensive and expensive (Mäkipää et al., 2008). On the other hand, the soil C stability and direction of change of the SOC stock might be

---

* Corresponding author.
E-mail address: lorenzo.menichetti@luke.fi (L. Menichetti).

https://doi.org/10.1016/j.geoderma.2024.116862
Received 16 December 2023; Received in revised form 8 March 2024; Accepted 15 March 2024

assessed using indicators based on measured SOC content and other soil variables. According to Drexler et al. (2022), the interpretation of measured SOC contents as a C loss indicator relies on i) the comparison of SOC values with a set of baseline values derived from representative datasets (De Rosa et al., 2023; De Vos et al., 2015; Nunes et al., 2021), and/or ii) the use of desirable SOC thresholds or reference values. These thresholds can vary according to relationships between SOC and soil properties (Carter, 2002; Loveland and Webb, 2003), the use of SOC values of undisturbed control sites values (Maharjan et al., 2020; Poeplau and Don, 2013), and/or the use of proxies such the proposed SOC:Clay ratio (Johannes et al., 2017; Prout et al., 2021).

The amount of clay is proposed to be a major factor influencing the ability of a soil to retain SOC (Hassink, 1997), related to the SOC-stabilizing properties of clay (Krause et al., 2018; Siebers et al., 2018). Dexter et al. (2008) suggested that a SOC:Clay ratio of 1:10 is the limit for clay to protect SOC. This approach was further reassessed by Johannes et al. (2017) and Prout et al. (2021) defining thresholds for SOC:Clay ratios of 1:8, 1:10, and 1:13 to indicate the boundaries between 'very good', 'good', 'moderate', and 'degraded' levels of soil structural conditions, respectively. In view of this, the European Soil Monitoring Law proposes to use a SOC:Clay ratio of 1:13 as the threshold for identifying degraded European mineral soils (European Commission, 2023).

In this study, we tested the performance of SOC:Clay ratio as an indicator of soil C loss on the soil data obtained from the Soil Module of 2009 Land Use and Coverage Area Frame Survey (LUCAS) survey in combination with other datasets for soil class, climate zone, and soil C stock changes. The LUCAS soil survey dataset covers a wide range of land covers, soil types, and SOC-clay content data across the EU. The principal objective was to evaluate applicability of the SOC:Clay ratio and its 1:13 threshold as a normalized 'soil health' metric for identifying carbon loss of European mineral soils and to compare this indicator with the C stock changes reported by countries to the climate convention (UNFCCC). Secondary objectives were to analyze distribution of the degraded soils across land cover and climatic zones, and to examine the potential of additional variables in predicting variation of the SOC content.

2. Material and methods

We relied on the 2009 LUCAS soil survey (https://esdac.jrc.ec.europa.eu/content/lucas-2009-topsoil-data) comprising a total of 21,859 data points in Europe. Since soil texture data (incl clay content) were collected during the LUCAS soil sampling in 2009, a more recent LUCAS soil survey was not used in this study. The data points are georeferenced and represent a composite soil sample of four subsamples to a depth of approximately 20 cm (measured after removing the organic layer and vegetation residues), collected in a radius of 100 m from the relative LUCAS point. Further details can be found in the documentation of the dataset (Tóth et al., 2013).

The land cover data from each data point was obtained from the primary 2009 LUCAS dataset (https://ec.europa.eu/eurostat/web/lucas/data/primary-data/2009). This dataset was cross-referred with several datasets using spatial join, with the aim of determining both the soil class and climatic zone of each point. For determining climate zones, we relied on the current Köppen-Geiger climate classification from Beck et al. (2023). The soil classification, which was based on the World Reference Base (WRB) for Soil Resources (IUSS Working Group WRB, 2015), was obtained from the European Soil Database v2.0 to define the soil class (https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data). In further selection, we excluded all points with SOC content above 200 g kg^{-1} dw, identified as Histosols by the WRB. This is because the high organic content in Histosols is likely a result of either extremely high inputs or extremely reduced decomposition, and as such, Histosols may not be sensitive to processes related to clay protection mechanisms.

We considered coarse land cover classes, namely Artificial (A in 2009 LUCAS classification), Cropland (B), Forest (C), Shrubland (D), Grassland (G), Bare land (F), Water land (G), and Wetlands (H). We studied the distribution of 'healthy/degraded soils' in this dataset, defined by the 1:13 (or 0.077) SOC:Clay ratio, according to land cover, countries, and climatic zones. In addition, we assessed the explanatory potential of the different variables included in the dataset to explain the SOC variance. For this, we trained a random forest model on 70 % of the data points, with the remaining 30 % reserved for model validation. Subsequently, we used this model to determine the importance of each variable in predicting SOC. The implementation used for this analysis was the *paRF* command (Zhang et al., 2021) in the *BAGofT* package in R software (version 4.3.1), and the metaparameter *mtry* (i.e., number of variables to sample at each split) was optimized via *caret*. Three random forests (RF) models were trained (Eqs. (1)-(3)):

$$SOC \approx f\left(r_{coarse}, r_{clay}, r_{sand}, r_{silt}, pH_{CaCl_2}, CaCO_3, Country, LC, lat, long,\right) \tag{1}$$

$$SOC \approx f\left(r_{coarse}, r_{clay}, r_{sand}, r_{silt}, pH_{CaCl_2}, CaCO_3, Country, LC, lat, long, WRB\right) \tag{2}$$

$$SOC \approx f\left(r_{coarse}, r_{clay}, r_{sand}, r_{silt}, pH_{CaCl_2}, CaCO_3, Country, LC, lat, long, CEC\right) \tag{3}$$

where 'r' is each specific soil fraction, 'pH_{CaCl_2}' is pH (measured in 0.01 M $CaCl_2$), '$CaCO_3$' is calcium carbonate content, 'Country' is the individual country, 'LC' is the land cover class, while 'lat' and 'long' are the latitude and longitude, respectively. We note that variables 'N' nitrogen, 'P' phosphorous, 'K' and potassium, which are also included in the 2009 LUCAS soil survey, were excluded from Eq. (1) because they are always strongly correlated with SOC, being also elements present in organic matter. The parameter WRB in Eq. (2) represents the soil class according to WRB, while the parameter 'CEC' in Eq. (3) represents cation exchange capacity of soil.

Variable importance is calculated as mean decrease in impurity, which is averaged across the trees in the forest. This is based on the reduction in impurity (Gini impurity) that each feature brings to the trees within the model. This metric quantifies the contribution of each feature to the overall homogeneity of the nodes and leaves by averaging the impurity decrease attributed to that feature across all trees in the forest. The more a feature decreases impurity, the more important it is considered for the model's predictive accuracy. The metric has only a value relative to each trained model.

Finally, we also compared the proportion of healthy/degraded soils (based on criteria SOC:Clay with threshold of 1:13) in each country according to land cover types with the emission data obtained from the annual EU greenhouse gas inventory as submitted to the UNFCCC (European Environment Agency, 2023). These data are developed based on various methodologies. Greenhouse gas inventories report their soil C stock changes based on methods that are in line with IPCC good practice guidelines (Penman et al., 2003). These methods vary according to their Tier levels from default emission factors to state-of-art soil models and to repeated soil inventories (European Environment Agency, 2023).

3. Results

3.1. Land cover and indicated soil C loss

Over the two dimensions of SOC and clay content, which were proposed to be a basis of the soil C loss indicator, the distribution of the data points in what is classified as degraded space was clearly influenced by land cover. In this analysis, we focused on the three main land cover classes, i.e., cropland, grassland, and forest, which represent 44.6 %, 21.8 % and 28.6 % of the total points, respectively (see Supplementary Table S1). Our results show a much higher proportion of degraded soils among croplands in comparison to grasslands and forests (Fig. 1). The share of degraded soil increased in the order forests (4.2 %) < grasslands (15.7 %) < croplands (51.0 %; Supplementary Table S2). Within other land cover classes (Supplementary Table S2), bare lands showed the highest proportion of degraded soils (69.7 %), while these soils were less represented in shrublands and wetlands (18.8 % and 4.0 %, respectively).

The median of the SOC content in cropland, grassland, and forest on mineral soils (when observations above 200 g kg^{-1} dw were excluded) were 14.2, 15.8, and 18.2 g kg^{-1} dw, respectively, whereas their median of clay content 210, 150, and 70 g kg^{-1} dw, respectively. Thus, grassland and forest contained more SOC in comparison to cropland, while the clay content of the soil was highest in cropland (Fig. 1, Supplementary Table S3).

3.2. SOC:Clay as soil C loss indicator and reported C losses by countries

According to the SOC:Clay ratio, it is estimated that on average 29.1 % of the mineral soils analyzed in the 2009 LUCAS survey are non-healthy (i.e., degraded; Fig. 2). However, there is a large variation in the proportion of degraded soils between studied countries, ranging from 1.2 % in Ireland to 68.5 % in Spain (Fig. 2b).

The SOC:Clay ratio is proposed as an indicator of the soil C loss. Therefore, the results that were based on the indicator values were compared to the soil C stock changes reported in the national GHG inventories (Fig. 2a). Based on the last 10 years of United Nations Framework Convention on Climate Change (UNFCCC) data for mineral soils (and excluding countries within a ± 0.5 t C ha^{-1} which were considered neutral), all grasslands and most of the forests (except Ireland) acted as C sinks. Croplands in several countries (Belgium, Finland, France, Germany, Netherlands, Poland, Portugal, and Slovenia) acted instead as C sources. In those countries, proportion of healthy croplands according to SOC:Clay ratio varied from 55 % in Slovenia to as high as 71 % in Germany. Interestingly, our findings reveal that Spain has neutral C soils in all reported land covers according to their GHG reporting to the UNFCCC, but only around 30 % of soils were classified as healthy with C loss indicator. In Estonia, the proportion healthy soils vary between 20 and 50 % depending on land-use, but simultaneously all these land-uses have been reported as carbon sinks by the UNFCCC reporting. The comparison of results between reporting of soil C stock changes to the UNFCCC and the proposed soil health indicator highlights inconsistencies between measures that are aiming to monitor soil C losses.

Our analysis also draws attention to the latitudinal gradient and its correspondence with climatic conditions on the dominance of healthy soils across the 2009 LUCAS survey. The healthiest soils are concentrated in northern European countries, while more degraded soils are found in southern countries (Fig. 2b). Among the climatic zones represented in the dataset (Fig. 3), the colder climates, i.e., Continental subartic (Dfc), and Continental humid with warm summers (Dfb), presented 98.5 % and 85.7 % of healthy soils, respectively. These regions are situated in northern Europe. The northern part of the boreal region has a Dfc climate, while the southern boreal region and the majority of eastern Europe are characterized by Dfb climate (Supplementary Fig. S1).

Temperate oceanic (Cfb) regions, mainly in UK, France, certain parts in Germany, as well as mountain ranges in Italy, had 76 % of soils classified as healthy (Fig. 3). Moreover, the Mediterranean climate with warm summers (Csb) had 68 % of the points classified as healthy. These regions are characterized by a Mediterranean climate that has oceanic influences, such as those found in northern Portugal and Spain. In contrast, Mediterranean areas with hot summers (Csa) only 45 % of the soils was classified as healthy. These regions are characterized by Mediterranean macchia vegetation with some xeric adaptation, as in the south-west part of Italy and Spain. Similarly, the Hot humid continental climate (Dfa) only included 48 % of healthy soils. It should be noted that soils from this climate are primarily present in Hungary, representing 98 % of all data points recorded for this climate class.

The Oceanic subtropical (Cfa) climate accounts for only 42 % of healthy soils (Fig. 3). This climate is common in highly productive agricultural regions around the Mediterranean basin, such as Pianura Padana in Italy. While the climate is hot, it is also typically quite humid. Finally, Arid (BWk and BWh) and Semi-arid steppe climates (BSk and BSh) are predominantly situated in central Spain, with smaller proportions in southern Italy and in Greece. These regions only included 30 %, 25 %, 20 %, and 0 % of healthy soils, respectively. They represent the driest Mediterranean areas, such as most of Puglia in Italy and the western part of central Spain, where agriculture requires significant adaptation to cope with the arid conditions.

3.3. Importance of variables in explaining the SOC variance

The 10 most important predictors in the first model (Eq. (1), Fig. 4) determining variation of the SOC content were: 1) clay, 2) pH, 3) if land

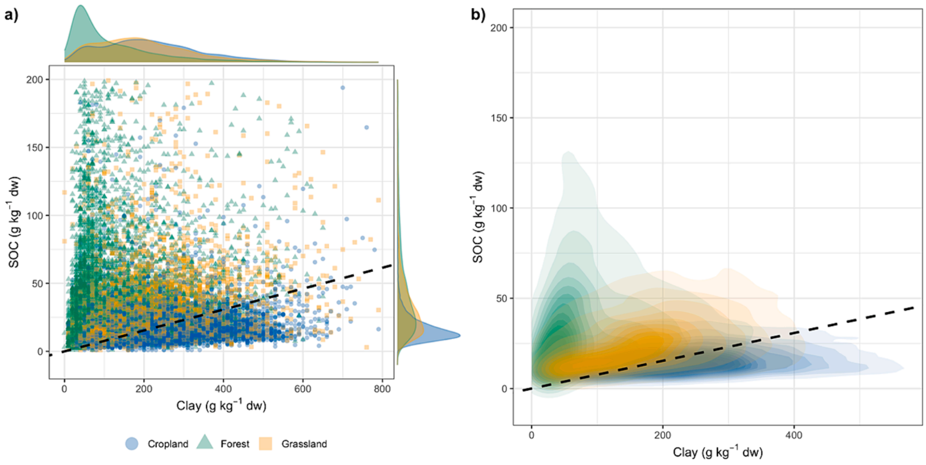

Fig. 1. Distribution of the soil organic carbon (SOC) and clay content for different land covers (cropland, forest, and grassland). Density plots of SOC and clay contents for each land cover are shown in panel a). Panel b) shows the 2-dimension density distribution of the data points of the 2009 LUCAS soil survey on croplands, croplands, and grasslands in the same space. Dashed line in panels a) and b) represents the SOC:Clay threshold of 1:13. Less represented land-use classes are not plotted for readability and SOC contents above 200 g kg^{-1} dry weight (dw) were excluded, since they were considered to represent organic soils where the SOC:Clay content is not proposed to be a relevant indicator.

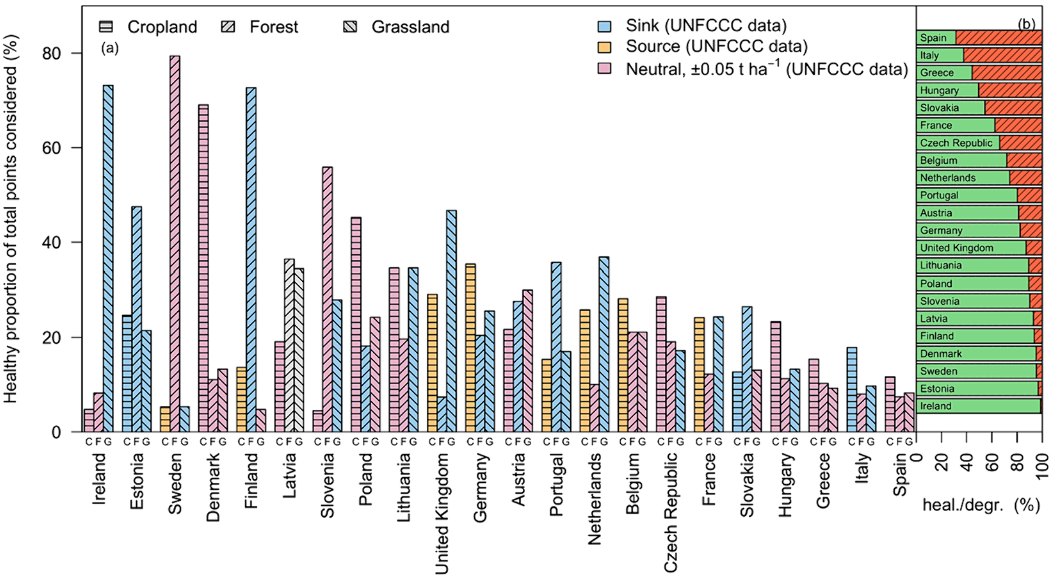

Fig. 2. Proportions of healthy (non-degraded) mineral soils by the three major land cover classes and country included in the 2009 LUCAS soil survey (a) and the proportion of healthy/degraded soils by country (b). The colors in panel (a) show the carbon stock change for each land cover based on the last 10 years of greenhouse gas inventory reporting under United Nations Framework Convention on Climate Change (UNFCCC) for mineral soils. Blue bars denote carbon sink, while orange bars indicate carbon source. Pink bars denote “neutral” countries, within a ± 0.5 t C ha^{-1} change. Grey bars denote missing data. The small labels under each bar denote the land covers cropland (C), forest (F), and grassland (G). The colors in panel (b) correspond to degraded (red, shaded) and healthy (green) soils.

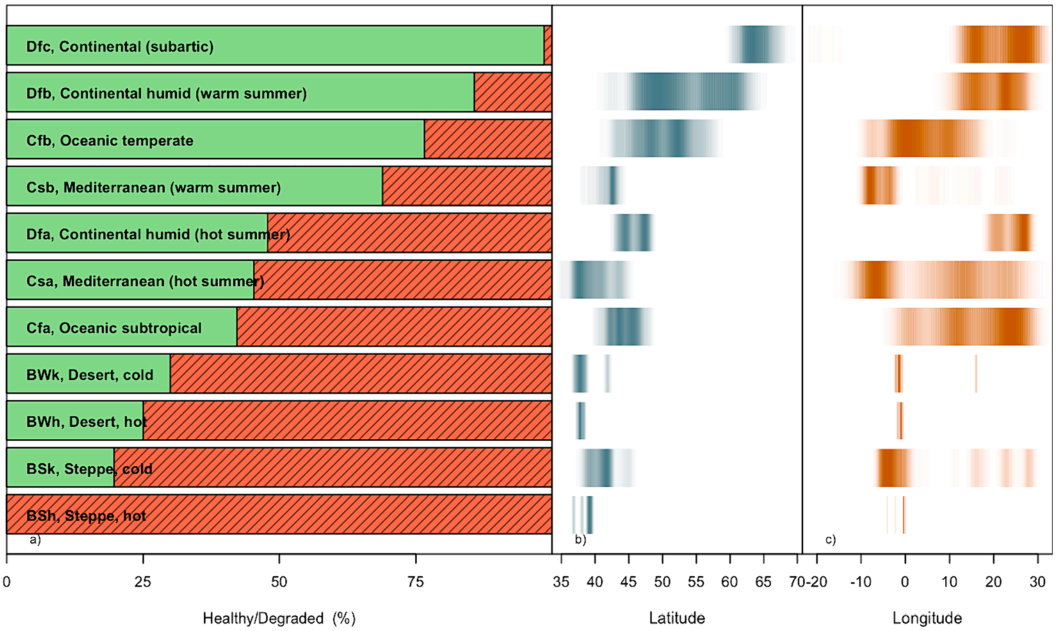

Fig. 3. Proportion of soil data classified as healthy (green) or degraded (red shaded) based on the SOC:Clay 1:13 threshold for different Köppen climatic zones (a), with the density distribution of their latitude (b) and longitude (c). Density distributions are calculated on all the points considered with kernel density estimation, darker color corresponds to higher point density.

cover was a cropland or any other class, 4) latitude, 5) longitude, 6) sand fraction, 7) calcium carbonate ($CaCO_3$) content, 8) coarse fraction, 9) silt fraction, and 10) if land cover was grassland or any other class.

In terms of linear correlations (i.e., Pearson’s correlation and R^2 of a linear regression), from the lowest correlation to the highest, the variables correlated with organic C as follows: Clay (Pearson’s = 0, $R^2 = 0$), silt (Pearson’s = -0.01, $R^2 = 0$), sand (Pearson’s = 0.01, $R^2 = 0$), K (Pearson’s = -0.01, $R^2 = 0$), P (Pearson’s = 0.02, $R^2 = 0$), coarse (Pearson’s = 0.11, $R^2 = 0.01$), $CaCO_3$ (Pearson’s = -0.12, $R^2 = 0.02$), Longitude (Pearson’s = 0.12, $R^2 = 0.01$), CEC (Pearson’s = 0.27, $R^2 = 0.08$), Latitude (Pearson’s = 0.32, $R^2 = 0.11$), pH($CaCl_2$) (Pearson’s = -0.36, $R^2 = 0.13$), and N (Pearson’s = 0.84, $R^2 = 0.71$).

The first random Forest (RF) model (Eq. (1)) could explain 50 % of the variance on the independent validation dataset (Supplementary Fig. S2), which even fall to 49 % when including the WRB soil classes as predictors (Eq. (2)). However, the introduction of CEC (Eq. (3)) significantly contributed to explaining a large portion of the SOC variability, as the only available information on C inputs to the soil were the land cover classes. On the independent validation dataset, this model explained 73 % of the variance.

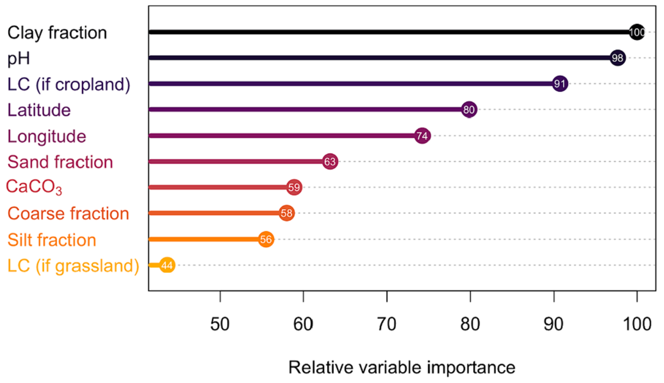

Fig. 4. Relative importance of the first 10 explanatory variables defined by the first Random Forest model (Eq. (1)). Note that the relative variable importance is proportional to the fraction of explained variance for each variable, but not the same thing. Colour is proportional to the relative importance and is added only for readability purposes.

4. Discussion

4.1. The SOC:Clay ratio as indicator of soil carbon loss

Observed discrepancies between soil carbon stock changes reported by the national GHG inventories and proportions of degraded soils identified by using the soil C loss indicator showed that a simple indicator such as SOC:Clay ratio with one threshold value for all soils across land cover classes, management practices and climate conditions cannot serve soil health monitoring as proposed by the European Commission for the Soil Monitoring Law. Although, monitoring of a simple indicator of the soil C loss cannot replace soil carbon inventory that builds on repeated soil measurements or comprehensive soil C modelling (European Environment Agency, 2023), the SOC:Clay ratio may be a valid indicator for soil structural quality that protects SOC on local scale as proposed in the studies where approach was developed (Dexter et al., 2008; Johannes et al., 2017; Prout et al., 2021). In soils with SOC:Clay $<$ 1:13, majority of the SOC is assumed to be protected by mineral matter and when the ratio is $>$ 1:13 there is assumed to be unprotected SOC which may be relatively easily mineralized in case of management change or changes in environmental conditions. For a soil health indicator that is relevant on a larger scale, variation in the SOC and clay content should be considered and threshold values specific to different land cover classes, edaphic and climatic conditions should be developed (Rabot et al., 2024). For example, in a dataset encompassing 3800 sites of croplands, grasslands, and forest soils throughout the UK where indicator based on SOC:Clay ratio was tested, the median clay content varied from 242 g kg^{-1} dw in forest soils to 260 g kg^{-1} dw in permanent grassland soils (Prout et al., 2021). In the European wide 2009 LUCAS soil survey data, however, the median clay contents of forest and grassland soils across Europe were 80 and 180 g kg^{-1} dw, respectively. Thus, the SOC:Clay threshold values derived by Prout et al (2021) cannot be applicable for soils across different land-use classes in Europe.

The SOC:Clay ratio in the 2009 LUCAS soil survey data indicate that a large proportion of the cropland soils is non-healthy (degraded), while the soils of forests and grasslands exhibited higher SOC content and lower clay content being therefore considered healthy if 1:13 threshold is applied (Fig. 1). There are, however, several limitations to consider when using the SOC:Clay ratio and its associated degradation threshold as a soil health indicator across different land cover types. For instance, clay content is not directly affected by the land cover (vegetation) and management history, which are drivers of C inputs to the soil and together with microbial activity determine SOC dynamics (Wiesmeier et al., 2013; Prout et al., 2021). Croplands tend to have, since they are usually located in the bottom part of the soil catena, higher clay content but simultaneously lower SOC content when compared to forests (Fig. 1) due to lower C inputs that forest and grassland soils. Due to their higher inputs, forest and grassland soils have usually more C than cropland soils (Morais et al., 2019). This suggests that land management is much more intensive in agriculture, which has annually repeated agricultural practices, compared to forestry, where there can be several decades between operations (e.g., soil preparation, thinning, and final felling). In any case, it appears clear the importance of C inputs to determine levels of SOC stocks. On the other hand, relatively low clay content of forest soils may relate to land cover history, i.e., areas with higher clay content were more fertile and converted to croplands, while soils with lower clay content remained as forests and grasslands. All this means that SOC:Clay ratio alone without also considering C inputs to the soil, is not a consistently sensitive indicator for detecting soil carbon loss on different land-use classes. Consequently, it is necessary that such soil health indicators are tailored to specific conditions and land covers. The problem can of course be mitigated by subsetting the SOC variability into main classes, for example cropland/forest/grassland, but still it is likely that there will be a significant part of the SOC variability explained by C inputs, which are variable also within each major class. The finer classification, based on some proxy of C inputs, we use, the more we would mitigate the problems of the SOC:Clay indicator at the expense of reducing its general applicability.

Our study found an influence of soil types (Supplementary Table S4), but often in interaction with land cover and climate. It is well known that different soil types inherently contain varying levels of SOC and clay (Prout et al., 2021). According to our analyses, Vertisols were the most degraded class, despite being rich in high-active clay minerals (Kögel-Knabner and Amelung, 2021), likely due to cumulative loss of C during their intensive and long-term use in agriculture. The second and third most degraded soil classes where Calcisols and Solonchaks, which are soils with high concentration of calcium and sodium salts (respectively) and are found in arid and semi-arid regions, which suggests an impact of climate and limited productivity on their degradation status. On the other hand, Podzols, which are usually soils having sandy to loamy sand texture and not particularly rich in clay, where the second least degraded class. This is likely due to Podzols being usually characterized by slow microbial activity in cool climate and forest’s high C input, partly persistent towards decomposition, to the soil. Cryosols were the less degraded class, and these soils are usually on marginal lands such as tundra but here cold climate associated with cryogenic processes (Ping et al., 2008; Kögel-Knabner and Amelung, 2021) is likely to play a role slowing down decomposition. Thus, influence of soil classes seemed to be consequence of many different processes.

According to applied SOC:Clay ratio as an indicator for soil C loss, the proportion on non-degraded healthy mineral soils was highest in the cool climate of northern Europe and Ireland, while in the Mediterranean region a large proportion of the soils were degraded (Fig. 2b and Fig. 3). Relying solely on the SOC:Clay ratio oversimplify the multiple processes involved in soil C dynamics and penalizes regions or countries with land covers and soil properties that lead to lower SOC:Clay ratios. Recently, Poeplau and Don (2023) carried out a study on German soils, in which they criticized the SOC:Clay ratio as an inappropriate metric for measuring SOC loss sensitivity. The authors argued that this indicator is highly reliant on clay and is highly susceptible to strong bias, but it is also being partially insensitive to SOC changes. Our research expands on the analyses of the use of this indicator in the wider EU region, highlighting the limitations of the proposed indicator. We suggest that sensitivity of a certain soil to C losses is more strongly influenced by land cover, intensity of land management, and climate rather than clay content (Fig. 1 and Fig. 3). Soil C stocks are affected by land-use and land management practices, since intensive management and frequent disturbances reduce both mean biomass stocks over rotations and annual organic matter inputs to soil (Lorenz and Lal, 2010). Climate and soil fertility define productivity, which affects biomass and organic matter inputs to soil. In addition, while the ability of soils to store organic C into soil particles is partly determined by their clay content, the type of clay minerals present also affects this process (Georgiou et al., 2022). There are regional patterns in this respect, e.g., northern EU regions are richer than southern regions in vermiculite and less in smectite (Ito and Wagai, 2017). As soil formation proceeds, clay composition undergoes changes. Soils that are enriched in subsoil by certain type of clay minerals are considered in international soil classification systems (IUSS Working Group WRB 2022). Thus, existing soil maps are connected to clay mineralogy. However, it is hard to define a precise geographical pattern of clay activity. Including clay mineralogy into the predictive models of SOC might contribute to explain more variance. However, it is difficult to speculate on the extent of this contribution without high spatial resolution soil data or clay mineralogy data related to each specific LUCAS sample.

4.2. *What variables are the most important in determining soil organic carbon content?*

Although the SOC content does not necessarily determine the fragility of SOC and trends of SOC stock changes, the same variables that impact SOC stocks can assist in identifying potential future trends and the risk of SOC losses. Therefore, a good predictor of SOC content may also be a good predictor of soil C health.

According to our analysis, the relationship between SOC and independent variables appears to be rather complex and likely with interactions between the main drivers: soil C protection/stabilization, climate, and C inputs to the soil (and therefore land cover). We showed that an important predictor of SOC was whether a sampling point is on cropland or not, which is the most impacting land cover for soil C.

Based on the statistical analyses (random forest) conducted in this study, the clay content was identified as one of the main predictors of the variation of soil organic matter when CEC was not considered. However, the model only accounted for 50 % of the total variance in the validation dataset, so there is still a lot of unexplained variation. Moreover, in terms of linear correlation, the explanatory power of clay fraction was among the lowest of all the variable tested, suggesting that although clay can indeed contain information about soil C, the relationship is not linear and is influenced by other variables (e.g., land cover). Therefore, although clay is known as a good proxy for some of the processes related to soil C loss (in particular, all the processes influencing the protection of soil C through adsorption; Xu et al., 2016), the clay fraction alone is a poor indicator of soil C content.

When also cation exchange capacity (CEC) was included in the model (Eq. (3)), it emerged as more important than clay, and the predictive power of the model increased substantially to a 73 % of the variance of the validation dataset explained by it. This might be related to clay exchange sites, but may also be attributed to the ability of organic matter to retain soluble ions because of its abundance of weak bonding sites (Hallsworth and Wilkinson, 1958). This makes hard to determine if the accumulation of SOC might be a result of the effect of organic matter on CEC or the cause. The good predictive power of pH in all models could have a similar explanation.

Silt, coarse, and sand fractions were all identified as good predictors, among the texture descriptors, together with clay, suggesting that part of the variance due to soil physical processes is related also to coarser fractions. These may also provide information regarding hydrology or other processes associated with soil C, which is not present in clay.

Additionally, another important predictor was latitude. Latitude has a strong influence on temperature, which in turn strongly affects SOC decomposition (Lloyd and Taylor, 1994). Similarly, longitude is also important, particularly in combination with latitude as it contains information about the distance from the sea and/or large mountain ranges and therefore the precipitation patterns. The interaction between latitude and longitude collectively provides information on climate. According to our analysis, soil class (WRB) was not a good predictor of SOC. We can only speculate about the reason for that, but it is possible that the information contained in the soil classification, which is highly multivariate and not quantitative, may not be suitable for describing the processes associated with SOC protection in a general enough way. Another possibility is soil edaphic variability, which is always a concern when using soil maps due to their relatively coarse resolution compared to soil variability, which may result in inaccurate data.

4.3. *Limitations of the study: Representativity of the LUCAS sampling*

The LUCAS soil survey has proven crucial in mapping the variability of European soils (Orgiazzi et al., 2017). However, as an extensive sampling effort requires coordination of multiple actors, the survey presents potential sources of bias. For example, there is a documented spatial bias in the exclusion of soils located at an altitude higher than 1500 m.a.s.l. (Schmidt et al., 2018). Furthermore, sampling protocol can be challenged, since the organic layer of the boreal forest soils, which is highly sensitive to disturbance induced loss of soil C, is not considered. Additionally, a sampling protocol, where sampling of the organic layer is not considered and separation of organic layer and uppermost mineral soil layer is not specified, may result in varying quantities of organic layer being included in the soil sample that represents the upper mineral soil layer. This can lead to biased results on the soil carbon content of that layer (see Tóth et al., 2013). Ziche et al. (2022) compared the 2015 LUCAS soil with German national soil inventory data. They found much higher variability and a bias towards higher values of C and N in the LUCAS data, possibly because of imprecision in how the organic layer was separated from the mineral layer. Panagos et al. (2020) had reported the same issue, again concerning forest soils. Also, Tóth et al., (2013) reported that in some cases part of the organic layer might have been sampled together with mineral soils in the 2009 sampling. The identification of the organic layer in forest soils is difficult also for trained soil scientists, and it is reasonable to imagine that the LUCAS sampling, relying on more general professionals, could have faced this problem. This might have increased the variability and reduced the significance of any predictive relationship of SOC, including the SOC: Clay ratio.

5. Conclusions

The SOC:Clay ratio with a threshold value of larger than 1:13 for healthy soils as proposed by the European Commission in the Soil Monitoring Law might be a very speditive indicator of soil C loss, but its general applicability as such is only within narrow ranges of SOC values, clay contents, land-use, and climate. Although there is evidence of the linkage between SOC:Clay ratio and soil structural quality, which indicates soil functioning and protection of SOC by mineral matter (Johannes et al., 2017; Prout et al., 2021), this process may not be the most important. By examining the points of the 2009 LUCAS soil dataset, we found that the variation in the SOC content was influenced not only by clay content but also by climate and land-use reflecting C input levels. Thus, the defined threshold is inadequate for detecting degraded soils if the SOC and clay content are beyond the conditions used to establish the criteria.

More efficient monitoring of the soil C and the risk of soil degradation could be achieved by utilizing other potential indicators beside SOC:Clay ratio, such as more strictly process-based approach that examines the ratio between particulate and mineral associated organic matter (POM and MAOM, respectively) as a warning signal of soil C fragility (Whalen et al., 2022). However, such a method requires relatively expensive analyses and monitoring network. A more cost-efficient strategy could rely on SOC models, which already account land-use and climate effects, to determine the reference maximum achievable C stocks for the entire EU. Thereafter, the measured levels may be compared to the reference. However, the lessons learned from the model-based setting of the forest reference level for 2021–2025 under the Land Use, Land Use Change, and Forestry (LULUCF) regulation and then determining LULUCF 2030 target based on mean values of the reported C sinks suggest that it may be challenging to develop and agree method that could provide solid scientific basis for the threshold of the healthy soil.

Currently, the national GHG inventories report SOC loss (or positive change in the SOC stock) through repeated SOC stock measurements or modelling (e.g., European Environment Agency, 2023). Given the inconsistency between the soil C stock change estimates of the GHG inventories and the SOC:Clay ratio as an indicator of soil C loss (Fig. 2), we conclude that the ratio is unsuitable for the monitoring of the SOC changes. Since many countries do not estimate soil C stock changes for all land-use classes or they continue using default emissions factors in their reporting under the UNFCCC, it is evident that both identification of non-healthy soils with a set of indicators and reporting of soil C stock changes under UNFCCC with higher Tier methods need repeated soil measurements and/or advanced soil modelling. Therefore, our study highlights the importance of implementing repeated soil measurements by national soil monitoring frameworks and further developing soil health indicators, scope, and relevant thresholds for SOC status evaluation. This will provide solid data for the local implementation of the European Soil Monitoring Law across different land-uses and biomes.

CRediT authorship contribution statement

Raisa Mäkipää: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Lorenzo Menichetti:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Eduardo Martínez-García:** Writing – review & editing, Writing – original draft. **Tiina Törmänen:** Writing – review & editing, Writing – original draft. **Aleksi Lehtonen:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data used in this study is available as OpenData as described in the manuscript

Acknowledgements

We acknowledge the LUCAS Soil survey for providing the European soil data and the national GHG inventories that created the UNFCCC data on soil carbon stock changes by the countries and land-use types.

This study received funding from the European Union's Horizon 2020 and Horizon Europe research and innovation programs under Grant Agreements number 101000289 for the HoliSoils project and number 101091010 for the Soil Mission project Benchmarks. The study has been done with affiliation to the UNITE Flagship funded by the Finnish Research Council (decision no 337655).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2024.116862.

References

Beck, H.E., McVicar, T.R., Vergopolan, N., Berg, A., Lutsko, N.J., Dufour, A., Zeng, Z., Jiang, X., van Dijk, A.I.J.M., Miralles, D.G., 2023. High-resolution (1 km) köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. Sci. Data 10, 724. https://doi.org/10.1038/s41597-023-02549-6.

Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T. W., Mäder, P., Pulleman, M., Sukkel, W., van Groenigen, J. W., Brussaard, L. (2018). Soil quality – A critical review. Soil Biol. Biochem., 120, 105–125. 10.1016/j.soilbio.2018.01.030.

Carter, M.R., 2002. Soil quality for sustainable land Management. Agron. J. 94, 38–47. https://doi.org/10.2134/agronj2002.3800.

De Rosa, D., Ballabio, C., Lugato, E., Fasiolo, M., Jones, A., Panagos, P., 2023. Soil organic carbon stocks in european croplands and grasslands: how much have we lost in the past decade? Glob. Chang. Biol. 00, e16992.

De Vos, B., Cools, N., Ilvesniemi, H., Vesterdal, L., Vanguelova, E., Carnicelli, E., 2015. Benchmark values for forest soil carbon stocks in Europe: results from a large scale forest soil survey. Geoderma 251–252, 33–46. https://doi.org/10.1016/j.geoderma.2015.03.008.

Dexter, A.R., Richard, G., Arrouays, D., Czyż, E.A., Jolivet, C., Duval, O., 2008. Complexed organic matter controls soil physical properties. Geoderma 144 (3–4), 620–627. https://doi.org/10.1016/j.geoderma.2008.01.022.

Doran, J.W., Zeiss, M.R., 2000. Soil health and sustainability: managing the biotic component of soil quality. Appl. Soil Ecol. 15 (1), 3–11. https://doi.org/10.1016/S0929-1393(00)00067-6.

Drexler, S., Broll, G., Flessa, H., Don, A., 2022. Benchmarking soil organic carbon to support agricultural carbon management: a german case study. J. Plant Nutr. Soil Sci. 185, 427–440. https://doi.org/10.1002/jpln.202200007.

Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Soil Monitoring and Resilience (Soil Monitoring Law). Brussels, 5.7.2023, COM (2023) 416 final, 2023/0232 (COD).

European Environment Agency (2023). Annual European Union greenhouse gas inventory 1990–2021 and inventory report 2023. Submission to the UNFCCC Secretariat. EEA/PUBL/2023/044. 2023.

FAO ITPS (2020). Towards a Definition of Soil Health. Soil Letters. 2020. Food and Agriculture Organization of the United Nations (FAO) and Intergovernmental Technical Panel on Soils (ITPS).

Georgiou, K., Jackson, R.B., Vindušková, O., Abramoff, R.Z., Ahlström, A., Feng, W., Harden, J.H., Pellegrini, A.F.A., Polley, H.W., Soong, J.L., Riley, W.J., Torn, M.S., 2022. Global stocks and capacity of mineral-associated soil organic carbon. Nat. Commun. 13, 3797. https://doi.org/10.1038/s41467-022-31540-9.

Grüneberg, E., Ziche, D., Wellbrock, N., 2014. Organic carbon stocks and sequestration rates of forest soils in Germany. Glob. Chang. Biol. 20, 2644–2662. https://doi.org/10.1111/gcb.12558.

Hallsworth, E.G., Wilkinson, G.K., 1958. The contribution of clay and organic matter to the cation-exchange capacity of the soil. J. Agric. Sci. 51, 1–3. https://doi.org/10.1017/S0021859600032743.

Hassink, J., 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant Soil 191, 77–87. https://doi.org/10.1023/A:1004213929699.

IPCC (2006) 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Available at: http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.htm.

IPCC (2019) Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Available at: https://www.ipcc-nggip.iges.or.jp/public/2019rf/index.

IUSS Working Group WRB (2015) World Reference Base for Soil Resources 2014, Update 2015. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps. World Soil Resources Reports N^{o}. 106, Rome. FAO.

Ito, A., Wagai, R., 2017. Global distribution of clay-size minerals on land surface for biogeochemical and climatological studies. Sci. Data **4**, 170103. https://doi.org/10.1038/sdata.2017.103.

Janzen, H.H., Janzen, D.W., Gregorich, E.G., 2021. The 'soil health' metaphor: Illuminating or illusory? Soil Biol. Biochem. 159, 108167 https://doi.org/10.1016/j.soilbio.2021.108167.

Johannes, A., Matter, A., Schulin, R., Weisskopf, P., Baveye, P.C., Boivin, P., 2017. Optimal organic carbon values for soil structure quality of arable soils. Does clay content matter? Geoderma 302, 14–21. https://doi.org/10.1016/j.geoderma.2017.04.021.

Kögel-Knabner, I., Amelung, W., 2021. Soil organic matter in major pedogenic soil groups. Geoderma 384, 114785. https://doi.org/10.1016/j.geoderma.2020.114785.

Kopittke, P.M., Minasny, B., Pendall, E., Rumpel, C., McKenna, B.A., 2023. Healthy soil for healthy humans and a healthy planet. Crit. Rev. Environ. Sci. Technol. 52, 4308–4324. https://doi.org/10.1080/10643389.2021.2024484.

Krause, L., Rodionov, A., Schweizer, S.A., Siebers, N., Lehndorff, E., Klumpp, E., Amelung, W., 2018. Microaggregate stability and storage of organic carbon is affected by clay content in arable Luvisols. Soil Till. Res. 182, 123–129. https://doi.org/10.1016/j.still.2018.05.003.

Lloyd, J., Taylor, J.A., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323. https://doi.org/10.2307/2389824.

Lorenz, K., Lal R. (2010) Carbon sequestration in forest Ecosystem. Springer Press. .

Loveland, P., Webb, J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil Tillage Res. 70 (1), 1–18. https://doi.org/10.1016/S0167-1987(02)00139-3.

Maharjan, B., Das, S., Acharya, B.S., 2020. Soil health gap: a concept to establish a benchmark for soil health management. Global Ecol. Conserv. 23, e01116.

Mäkipää, R., Häkkinen, M., Muukkonen, P., Peltoniemi, M., 2008. The costs of monitoring changes in forest soil carbon stocks. Boreal Environ. Res. 13, 120–130.

Morais, T.G., Teixeira, R.F.M., Domingos, T., 2019. Detailed global modelling of soil organic carbon in cropland, grassland and forest soils. PLoS One 14 (9), e0222604.

Nunes, M.R., Veum, K.S., Parker, P.A., Holan, S.H., Karlen, D.L., Amsili, J.P., van Es, H. M., Wills, S.A., Seybold, C.A., Moorman, T.B., 2021. The soil health assessment protocol and evaluation applied to soil organic C. Soil Sci. Soc. Am. J. 85, 1196–1213. https://doi.org/10.1002/saj2.20244.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2017. LUCAS soil, the largest expandable soil dataset for Europe: a review. Eur. J. Soil Sci. 69 (1), 140–153. https://doi.org/10.1111/ejss.12499.

Ortiz, C.A., Liski, J., Gärdenäs, A.I., Lehtonen, A., Lundblad, M., Stendahl, J., Ågren, G.I., Karltun, E., 2013. Soil organic carbon stock changes in swedish forest soils – a comparison of uncertainties and their sources through a national inventory and two simulation models. Ecol. Model. 251, 221–231. https://doi.org/10.1016/j.ecolmodel.2012.12.017.

Panagos, P., Ballabio, C., Scarpa, S., Borrelli, P., Lugato, E., Montanarella, L. (2020) Soil related indicators to support agro-environmental policies. EUR 30090 EN. Publications Office of the European Union, Luxembourg. 10.2760/011194.

Panagos, P., Montanarella, L., Barbero, M., Schneegans, A., Aguglia, L., Jones, A., 2022. Soil priorities in the European Union. Geoderma Reg. 29, e00510.

Penman, J., Gytarsky, M., Hiraishi, T., Krug, T., Kruger, D., Pipatti, R., Buendia, L., Miwa, K., Ngara, T., Tanabe, K. & Wagner, F. (2003). IPCC Good practice guidance for land use, land-use change and forestry, Institute for Global Environmental Strategies (IGES), Japan.

Ping, C.L., Michaelson, G.L., Kimble, J.M., Romanovsky, V.E., Shur, Y.L., Swanson, D.K., Walker, D.A., 2008. Cryogenesis and soil formation along a bioclimate gradient in Arctic North America. J. Geophys. Res. 113, G3. https://doi.org/10.1029/2008JG000744.

Poeplau, C., Don, A., 2013. 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.

Poeplau, C., Don, A., 2023. A simple soil organic carbon level metric beyond the organic carbon-to-clay ratio. Soil Use Manag. 39, 1057–1067. https://doi.org/10.1111/sum.12921.

Prout, J.M., Shepherd, K.D., McGrath, S.P., Kirk, G.J.D., Haefele, S.M., 2021. What is a good level of soil organic matter? an index based on organic carbon to clay ratio. Eur. J. Soil Sci. 72, 2493–2503. https://doi.org/10.1111/ejss.13012.

Rabot, E., Saby, N.P.A., Martin, M.P., Barré, P., Chenu, C., Cousin, I., Arrouays, D., Angers, D., Bispo, A., 2024. Relevance of the organic carbon to clay ratio as a national soil health indicator. Geoderma 443, 116829. https://doi.org/10.1016/j.geoderma.2024.116829.

Schmidt, S., Ballabio, C., Alewell, C., Panagos, P., Meusburger, K., 2018. Filling the European blank spot – Swiss soil erodibility assessment with topsoil samples. J. Plant Nutr. Soil Sci. 181 (5), 737–748. https://doi.org/10.1002/jpln.201800128.

Siebers, N., Bauke, S.L., Tamburini, F., Amelung, W., 2018. Short-term impacts of forest clear-cut on P accessibility in soil microaggregates: an oxygen isotope study. Geoderma 315, 59–64. https://doi.org/10.1016/j.geoderma.2017.11.024.

Tóth, G., Jones, A., Montanarella, L. (2013). LUCAS Topsoil Survey. Methodology, data and results. JRC Technical Reports. Luxembourg. Publications Office of the European Union, EUR26102 – Scientific and Technical Research series. Doi: 10.2788/97922.

UNFCCC (2023) Nationally determined contributions under the Paris Agreement. Synthesis report by the secretariat, UN Climate Change Conference 14 Nov 2023 FCCC/PA/CMA/2023/12 https://unfccc.int/documents/632334.

Whalen, E.D., Grandy, A.S., Sokol, N.W., Keiluweit, M., Ernakovich, J., Smith, R.G., Frey, S.D., 2022. Clarifying the evidence for microbial- and plant-derived soil organic matter, and the path towards a more quantitative understanding. Glob. Chang. Biol. 28, 7167–7185. https://doi.org/10.1111/gcb.16413.

Wiesmeier, M., Hübner, R., Barthold, F., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Schilling, B., von Lützow, M., Kögel-Knabner, I., 2013. Amount, distribution and driving factors of soil organic carbon and nitrogen in cropland and grassland soils of southeast Germany (Bavaria). Agr. Ecosyst. Environ. 176, 39–52. https://doi.org/10.1016/j.agee.2013.05.012.

Xu, X., Shi, Z., Li, D., Rey, A., Ruan, H., Craine, J.M., Liang, J., Zhou, J., Luo, Y., 2016. Soil properties control decomposition of soil organic carbon: results from data-assimilation analysis. Geoderma 262, 235–242. https://doi.org/10.1016/j.geoderma.2015.08.038.

Zhang, J., Ding, J., Yang, Y., 2021. Is a classification procedure good enough? a goodness-of-fit assessment tool for classification Learning. J. Am. Stat. Assoc. 118 (542), 1115–1125. https://doi.org/10.1080/01621459.2021.1979010.

Ziche, D., Grüneberg, E., Riek, W., Wellbrock, N. (2022) Comparison of the LUCAS 2015 inventory with the second National Forest Soil Inventory: Comparability and representativeness of two soil inventories conducted in Germany. Johann Heinrich von Thünen Institut, Thünen-Report 94, 62 S.