Science of the Total Environment 951 (2024) 175642

Contents lists available at ScienceDirect

Science of the Total Environment

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

Benchmarking soil organic carbon (SOC) concentration provides more robust soil health assessment than the SOC/clay ratio at European scale

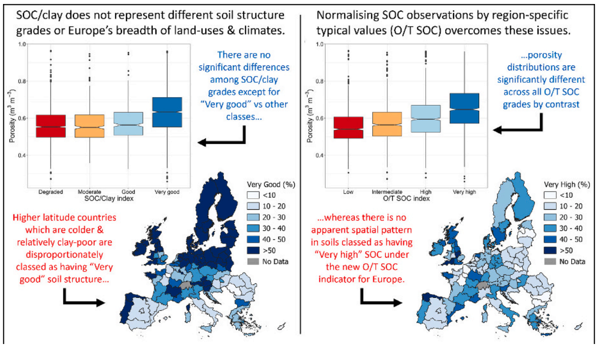
Christopher J. Feeney [a,*], Laura Bentley [a], Daniele De Rosa [b], Panos Panagos [c], Bridget A. Emmett [a], Amy Thomas [a], David A. Robinson [a]

[a] UK Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
[b] School of Agriculture, Forestry, Food and Environmental Sciences, University of Basilicata, 85100 Potenza, Italy
[c] European Commission – Joint Research Centre, Via Enrico Fermi, 2749, 21027 Ispra, VA, Italy

HIGHLIGHTS

- Developed the O/T SOC indicator accounting for climate, land cover and soil texture
- We model 12 pedo-climate zones across Europe with associated typical SOC values.
- O/T SOC is less sensitive to soil texture, land-use and climate than SOC/clay.
- O/T SOC better reflects differing porosity and SOC stock grades than SOC/clay does.
- O/T SOC is sensitive to changing SOC under land management over years to decades.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Fang Wang

Keywords:
Soil carbon
Clay
Pedo-climate zones
Soil structure
Soil monitoring
Sustainable Development Goals

ABSTRACT

Increasing soil organic carbon (SOC) confers benefits to soil health, biodiversity, underpins carbon sequestration and ameliorates land degradation. One recommendation is to increase SOC such that the SOC to clay ratio (SOC/clay) exceeds 1/13, yet normalising SOC levels based on clay alone gives misleading indications of soil structure and the potential to store additional carbon. Building on work by Poeplau & Don (2023) to benchmark observed against predicted SOC, we advance an alternative indicator: the ratio between observed and "typical" SOC (O/T SOC) for pan-European application. Here, "typical" SOC is the average concentration in different pedo-climate zones, PCZs (which, unlike existing SOC indicators, incorporate land cover and climate, alongside soil texture) across Europe, determined from mineral (<20 % organic matter) topsoils (0–20 cm) sampled during 2009–2018 in LUCAS, Europe's largest soil monitoring scheme (n = 19,855). Regression tree modelling derived 12 PCZs, with typical SOC values ranging 5.99–39.65 g kg^{-1}. New index classes for comparison with SOC/clay grades were established from the quartiles of each PCZ's O/T SOC distribution; these were termed: "Low" (below the 25th percentile), "Intermediate" (between the 25th and 50th percentiles), "High" (between the 50th and 75th percentiles), and "Very high" (above the 75th percentile). Compared with SOC/clay, O/T SOC was less sensitive to clay content, land cover, and climate, less geographically skewed, and better reflected differences in soil porosity and SOC stock, supporting 2 EU Soil Health Mission objectives (consolidating SOC stocks; improving soil

---

* Corresponding author.
E-mail address: chrfee@ceh.ac.uk (C.J. Feeney).

https://doi.org/10.1016/j.scitotenv.2024.175642
Received 10 March 2024; Received in revised form 19 July 2024; Accepted 17 August 2024

structure for crops and biota). These patterns held for 2 independent datasets, and O/T SOC grades were sensitive enough to reflect land management differences across several long-term field experiments. O/T SOC used in conjunction with several other physical, chemical and biological soil health indicators can help support the EU Soil Monitoring Law and achieve several United Nations Sustainable Development Goals.

1. Introduction

Land degradation poses one of the greatest environmental challenges to humanity, with threats to "soil health" (defined by the EU Soil Health Mission as the physical, chemical and biological condition of the soil, determining its continued capacity to function as a vital living system and support ecosystem services (Veerman et al., 2020)) representing the bulk of major land degradation pathways (Prăvălie, 2021). The EU Soil Observatory (EUSO) has identified >60 % of soils across Europe as "unhealthy", with at least one land degradation process occurring in affected areas (Panagos et al., 2024). In response to this, the European Union (EU) proposed a Soil Monitoring Law in July 2023 to provide a legal framework to achieve healthy soils by 2050 and ensure their long-term sustainable use (Directorate-General for Environment, 2023). More broadly, improving soil health is regarded by the EU Mission Board for Soil Health and Food, among others, as essential for achieving several United Nations Sustainable Development Goals (SDGs) – notably SDGs 2 (zero hunger), 6 (clean water and sanitation), 13 (climate action) and 15 (life on land) (Bonfante et al., 2020; Bouma, 2014; Keesstra et al., 2018, 2016; Veerman et al., 2020; Visser et al., 2019).

Given the myriad benefits soil organic carbon (SOC) confers upon soil, particularly aggregation, which in turn promotes water retention capacity and fluxes of water, air and gas, reduces risks of compaction, sealing and erosion, and through creation of more stable aggregates, promotes further long-term SOC storage (Blanco-Canqui et al., 2013), it is widely regarded as one of the most important metrics for indicating soil health (Kopittke et al., 2022). Additionally, soils store the most organic carbon in the terrestrial biosphere (Lal, 2004; Schmidt et al., 2011) and it is estimated that a third of soils globally are losing SOC rapidly (FAO and ITPS, 2015). As a result, there have been growing calls to promote not just "climate-smart soils" through initiatives such as "4per1000" to promote SOC storage for climate change mitigation, but also "soil-smart agriculture" to promote sustainable land management for the betterment of soils generally (Moinet et al., 2023). Because of these concerns, soil properties including topsoil SOC concentration are commonly monitored in national field surveys (Reynolds et al., 2013) to assess status and trends. A persistent challenge however is deriving an optimal or target SOC content for policymakers. Topsoil SOC content is sensitive to several environmental and management factors and may thus have individual ranges that are the product of site-specific configurations of land cover and inherent conditions such as soil type and climate that are beyond the realm of manageable intervention. Furthermore, the particular ecosystem services of interest will be an important determining factor in defining thresholds or optimal levels of SOC.

It has been established that intimate association with fine silt and clay particles is a key storage mechanism for SOC (Hassink, 1997; Matus, 2021; Six et al., 2002). Furthermore, it has been posited that the physical behaviour of mineral soils is controlled by the amount of complexed organic carbon with clay (COCC) (Dexter et al., 2008). These observations have spurred the development of a SOC content metric that is normalised by the relative clay content of the soil: the SOC/clay ratio. Analysis of Swiss arable soils yielded 3 SOC/clay ratio thresholds: 1/13, 1/10 and 1/8, marking boundaries between the index classes, "Degraded", "Moderate", "Good" and "Very good", respectively (Johannes et al., 2017). These index classes were found to reflect a gradient of soil structural quality scores, with recent analysis from England and Wales appearing to validate their wider applicability within Europe (Prout et al., 2021, 2022). This has led to the inclusion of the SOC/clay ratio as a recommended indicator to monitor SOC concentrations in the proposed EU Soil Monitoring Law (Directorate-General for Environment, 2023). However, the suitability of the SOC/clay ratio as a soil health indicator has been challenged. Analysis of German agricultural soils revealed that soils with relatively low clay (< 20 %) were disproportionately likely to have a "Very good" SOC/clay ratio, despite showing clear potential to store significantly more carbon, while clay-rich soils (> 60 %) were overwhelmingly likely to be classed as "Degraded", even where other metrics would imply adequate soil structure or fertility (Poeplau and Don, 2023). Additionally, inconsistencies have been shown between the soil carbon stock changes reported by national greenhouse gas inventories and the proportions of degraded soils according to SOC/clay across Europe (Mäkipää et al., 2024), while Rabot et al. (2024) recommended an update to the SOC/clay indicator that accounted for different soil types and climates.

As an alternative to the SOC/clay ratio, Poeplau and Don (2023) developed a normalised SOC content metric based on the ratio of observed SOC to the "expected" concentration. This "observed/expected SOC ratio" was found to be less sensitive to clay content and more adequately reflect soil porosity differences. However, Poeplau and Don (2023) recommended that modifications are necessary for application at European scale because they predicted SOC as a function of clay only, and for soils from agricultural settings (i.e. no woodlands or semi-natural grasslands) within a single country, which does not capture the spectrum of climates across the continent (from the Boreal north to the Mediterranean south). A new SOC-based soil health indicator ought to apply across the breadth of possible mineral soils, from sandy soils under arable land-use in relatively warm and dry regions (where SOC is likely low), to clay-rich soils in semi-natural habitats in relatively cold and wet regions (where SOC is likely high). Thus, the consideration of such a high diversity of environmental conditions will likely require multiple reference SOC content values (Poeplau and Don, 2023; Rabot et al., 2024).

The aim of this study was to improve the alternative normalised SOC indicator proposed by Poeplau and Don (2023) for the assessment of soil health for all mineral soils across Europe. The objectives included: (i) deriving the "typical" SOC as a function of key environmental and management controls on observed SOC concentrations across Europe; (ii) defining index categories that usefully summarise differing grades of topsoil porosity and SOC stocks; and (iii) demonstrate how the new indicator is less sensitive to geography, clay content, land cover and climate factors than SOC/clay and thus, more suitable for soil health monitoring at large scale.

2. Methods

2.1. Study area and selected datasets

The analysis used EU-wide soil survey data from the Land Use/Cover Area Frame Survey (LUCAS) (Orgiazzi et al., 2018; Tóth et al., 2013b). Soil property data, including soil organic carbon (SOC) concentration, texture and land cover from the first three surveys (2009/12, 2015 and 2018) are publicly available from the European Soil Data Centre (Panagos et al., 2012, 2022) and were selected for data analysis. Soil bulk density data for a subset of the 2018 LUCAS survey sites ($n = 6059$) are also available (Pacini et al., 2023) and used in this study. All details of data collection in the field, laboratory analysis and quality assurance procedures can be found in the associated technical documentation (Fernandes-Ugalde et al., 2022; Jones et al., 2020; Tóth et al., 2013a).

Table 1

Proportional make-up of land cover classes and biogeographical regions (BGRs) in the LUCAS dataset. Land cover refers specifically to areas that have been the same class for all 3 LUCAS surveys (2009/12, 2015 & 2018) whereas percentage of EU+UK area represents means (covering CORINE land cover mapping years 2012 & 2018; no map exists for 2009 or 2015). Land cover percentages for EU+UK area are calculated after removing other land cover classes (e.g. bogs, urban areas, open water) to allow fair comparison against LUCAS.

	Percentage of LUCAS dataset	Percentage of EU+UK area
Cropland land cover	52.7	43.6
Grassland land cover	24.4	16.5
Woodland land cover	22.9	40.0
Alpine BGR	4.4	8.7
Atlantic BGR	18.0	17.9
Boreal BGR	15.0	19.3
Continental BGR	33.6	29.5
Mediterranean BGR	24.2	20.7
Pannonian BGR	3.3	2.9
Other (Black Sea and Steppic) BGRs	1.4	1.1

For the purposes of our analysis, we used an aggregated dataset that combined data from all 3 extant LUCAS surveys, which had been prepared for a recent analysis of SOC content changes across Europe (De Rosa et al., 2024). As part of the development of this aggregated dataset, rigorous checks were undertaken to validate recorded land cover classes against a database of 874,646 geo-tagged close-up photos with land cover and plant species labels (D'Andrimont et al., 2022). Where SOC concentration data were available for multiple years, an average value was calculated. The dataset was also filtered to exclude sites with either missing SOC or texture information, or a change in land cover class between survey years. The dataset was then filtered further to include just cropland, grassland and forest sites because other land cover classes were judged to be too rare for use in our modelling. After further refining our data to exclude organic soils (≥ 120 g kg^{-1} SOC), we were left with a total of 19,855 data points. Cropland and the Continental biogeographical region represent the dominant land cover and climate zone, respectively in the LUCAS dataset (Table 1). More broadly, the proportions of land cover classes and biogeographical regions in the LUCAS dataset are generally similar to those categories that have been mapped for the EU + UK land area (Table 1), demonstrating that the LUCAS data closely reflects the environment across Europe.

2.2. Creating a new normalised soil carbon content metric for soil health assessment

Before deriving a new soil carbon-based indicator of soil health, we sought to establish whether the SOC/clay ratio exhibited a strong clay content bias as had been previously identified for German agricultural topsoils (Poeplau and Don, 2023) and across Europe more widely (Mäkipää et al., 2024; Rabot et al., 2024). The SOC/clay ratio was calculated by dividing SOC by 10 to convert the units from g kg^{-1} to % and then dividing by percentage clay content. Next, SOC/clay ratio was plotted as a function of clay content to determine whether a clear bias was observable (i.e. a decay curve from high SOC/clay ratios for sites with minimal clay content to low SOC/clay ratios for sites much richer in clay content, e.g. >35 % clay).

As an alternative to the SOC/clay ratio, Poeplau and Don (2023) proposed the ratio between observed and "expected" SOC, with expected SOC estimated using a linear regression model to predict the SOC concentration as a function of clay content. Developing a similar metric for the LUCAS dataset required an alternative approach however, as the broad range of land cover types and climate zones encompassed by Europe means there is unlikely to be a single linear function of SOC and clay content that would be universally applicable.

Here, we applied univariate regression tree modelling to stratify the mineral soils from the LUCAS dataset into several discrete subsets – referred to herein as "pedo-climate zones" (PCZs) – each characterised by a combination of key environmental predictors of SOC concentration. Regression tree modelling was adopted here because it allows predictions to be generated from multiple potential predictors, including categorical as well as continuous variables, and produces a transparent decision tree which conveniently characterises discrete environments (Elith et al., 2008). Potential predictor variables included elevation, geomorphon landform class (Jasiewicz and Stepinski, 2013), land cover, calcium carbonate content, coarse fragment content, biogeographical region (as a proxy for climate) and soil type. Biogeographical regions were assigned to the point locations from all LUCAS survey years through an intersect operation performed in QGIS 3.32 [Lima] with a polygon shapefile of the 2016 map of the biogeographical regions of Europe (European Environment Agency, 2017). Elevation values from the 30 m digital elevation model of Europe from EcoDataCube and geomorphon land classes from a 90 m global map of landform classes (Amatulli et al., 2020) were extracted to LUCAS point locations using the "Sample Raster Values" tool in QGIS. Soil types were defined by categorising the LUCAS observed mineral soils using the European Soil Map (HYPRES) (Wösten, 2000) soil texture classification according to clay, silt and sand content. Soil type classification was implemented with the "soiltexture" package in R (Moeys et al., 2018; R Core Team, 2023).

Once all predictor variables were assigned, a training dataset was created by extracting a random sample of 80 % of the LUCAS dataset; the remaining 20 % was kept as testing data. Univariate regression tree modelling via recursive partitioning was implemented in R using the "rpart" package (R Core Team, 2023; Therneau et al., 2022) with the training dataset as input and the natural logarithm of SOC concentration as the response variable. SOC concentrations were log-transformed to ensure values were approximately normally distributed because ANOVA is used to compare the means of multiple groups in the "rpart()" function (Therneau et al., 2022). Tree depth was controlled by optimising the complexity parameter value via 10-fold cross-validation with the "caret" package in R (Kuhn et al., 2023; R Core Team, 2023), which ensured the usual step of "pruning" decision tree sub-nodes to reduce model overfitting was unnecessary. Regression tree accuracy was assessed by applying the model fitted to the training dataset to the testing dataset and calculating the root mean square error (RMSE). A key disadvantage of regression tree modelling is the tendency for a small change in the dataset to induce a large change in the final tree (Elith et al., 2008). Normally, this is overcome by fitting models to several unique training samples of the original dataset and "bagging" (aggregating) the results. However, for interpretive clarity, we wanted to produce a single decision tree that could be used to identify average or "typical" SOC concentrations based on identifiable combinations of environmental variables. To overcome this issue, we modelled regression trees for 5000 unique permutations of training datasets and selected the one with the lowest RMSE when evaluated against testing data. Additionally, the number of times each complete branch to a "leaf" (terminal node) of the selected regression tree occurred out of the 5000 models was calculated and expressed as a percentage. Means and standard deviations were recorded for each leaf and back transformed from logarithmic-scaled SOC concentrations. Coefficients of variation were calculated for each PCZ-level SOC distribution by dividing standard deviation by the mean and multiplying by 100 to express as percentages. This was done to illustrate the level of variability around the "typical" (mean) SOC value in each PCZ.

Using the newly created decision tree, each of the LUCAS mineral soil records was assigned a PCZ based on their specific combination of relevant environmental characteristics. We then divided each observed SOC concentration by the mean value predicted for the relevant PCZ to derive our new normalised SOC metric: the observed/typical SOC ratio (herein shortened to O/T SOC). There are 4 index classes established for the SOC/clay ratio ("Degraded", "Moderate", "Good" and "Very good", separated by thresholds of 1/13, 1/10 and 1/8) (Johannes et al., 2017).

To compare our new O/T SOC indicator with the established SOC/clay indicator, we calculated the 25th, 50th and 75th percentiles of O/T SOC for each PCZ, to define index classes as follows: "Low" = below the 25th percentile; "Intermediate" = between the 25th and 50th percentiles; "High" = between the 50th and 75th percentiles; "Very high" = above the 75th percentile. Here, we opted to use descriptors like "low" and "high" because to judge a soil as "degraded" or "good" depends on contexts that we do not capture in our methodology, such as which ecosystem service(s) – as well as the balance of trade-offs and synergies between these – that stakeholders wish to prioritise.

Classifications, based on the new O/T SOC metric were compared with the established SOC/clay classes in two main ways. First, using the LUCAS data, O/T SOC was plotted as a function of clay content to determine whether a clear relationship with clay content exists. Second, distributions of total soil porosity and SOC stocks from the 2018 LUCAS survey were plotted for each of the SOC/clay and O/T SOC classes using the "tidyverse" package in R (R Core Team, 2023; Wickham, 2023). Effective soil porosity, which considers aggregation, connectivity between pores and pore morphology than total porosity, would likely be a more suitable indicator of soil structure. However, it was not possible to derive this metric from LUCAS data. Nevertheless, total soil porosity is still an important factor influencing fluxes of water and gas through soils (Assouline and Or, 2013), and microbial communities (Or et al., 2007), and has been shown to be highly sensitive to changes in SOC with environmental change (Robinson et al., 2022; Thomas et al., 2024). By assessing total porosity and SOC stocks in relation to O/T SOC and SOC/clay, we aim to show that O/T SOC index classes can more usefully separate differing grades of physical structural quality and carbon storage than would be the case for SOC/clay index classes.

Porosity ($m^3\ m^{-3}$) was calculated following the approach of Breuning-Madsen et al. (2015):

$$Porosity = 1 - \frac{Bulk\ Density}{Particle\ Density} \tag{1}$$

Bulk density ($g\ cm^{-3}$) refers to the fine fraction of soil specifically (<2 mm). For this, we used the corrected to fine earth (<2 mm) bulk densities presented in Pacini et al. (2023), and first removed extreme bulk density values and any records associated with organic soils (leaving $n = 4642$). Particle density ($g\ cm^{-3}$) was calculated as:

$$Particle\ density = \frac{OM_{pd} \times Si_{pd} \times CaCO_{3pd}}{\left(Si_{pd} \times CaCO_{3pd} \times OM_{prop}\right) + \left(OM_{pd} \times CaCO_{3pd} \times Si_{prop}\right) + \left(OM_{pd} \times Si_{pd} \times CaCO_{3prop}\right)} \tag{2}$$

OM_{pd}, Si_{pd} and $CaCO_{3.pd}$ refer to the particle densities of organic matter (1.30 $g\ cm^{-3}$), silicates (2.65 $g\ cm^{-3}$) and calcium carbonate fractions (2.80 $g\ cm^{-3}$), respectively. OM_{prop}, Si_{prop} and $CaCO_{3.prop}$ refer to the relative proportions of organic matter, silicates and calcium carbonate contents, respectively.

SOC stocks were calculated following the approach of Poeplau et al. (2017):

$$SOC\ Stock = SOC\ (\%) \times Bulk\ Density \times Depth \times \left(1 - Coarse_{prop}\right) \tag{3}$$

Depth (cm) is set to 20 to align with the LUCAS survey design; $Coarse_{prop}$ refers to the fraction by volume of coarse fragments in the soil according to Pacini et al. (2023). Differences in porosity and SOC stocks between SOC level classes for each metric were tested using the Kruskall-Wallis and the post hoc pairwise Wilcoxon tests, with significance assessed at $\alpha = 0.05$ for both tests.

As a further comparison of the 2 SOC-based soil health indicators, percentages of observations with different SOC/clay and O/T SOC index classes were summarised by HYPRES soil texture class, land cover and biogeographical region. We also developed maps of the percentages of LUCAS points in each NUTS 1 region with extreme SOC/clay and extreme O/T SOC conditions. First, the number of LUCAS observations in each NUTS 1 region was calculated using the "Count points in polygons" tool in QGIS. This same step was repeated 4 more times, once for each of the following conditions: (i) observations classed as "Degraded" SOC/clay, (ii) observations classed as "Very Good" SOC/clay, (iii) observations classed as "Low" O/T SOC, and (iv) observations classed as "Very High" O/T SOC. Counts from this second step were divided by the total number of LUCAS observations in each NUTS 1 region, calculated in step 1, and multiplied by 100 to express as percentages. Heavily urbanised NUTS 1 regions in the UK, Belgium and Germany were excluded from the analysis as these had very few or no LUCAS observations.

2.3. *Demonstrating the applicability of O/T SOC beyond the LUCAS dataset*

We used 2 independent datasets to evaluate the suitability of our O/T SOC index classes for individual EU member states and regions to benchmark their soil health status against comparable landscapes across Europe. These 2 datasets include the Glastir Monitoring and Evaluation Programme (GMEP) for Wales from UKCEH, and the Bodenzustandserhebung (BZE) survey of agricultural soils for Germany from the Thünen Institute. Both datasets were generated within the 2009–2018 timeframe of LUCAS and cover similar topsoil depths (0–15 cm for GMEP; 0–30 cm for BZE) to LUCAS, so they should provide adequate comparison. The 2 surveys also reflect contrasting land-use mixes, with BZE predominantly covering cropland and GMEP covering pastures and semi-natural grasslands and forests. Published location data for GMEP and BZE datapoints were overlain with the biogeographical regions map of Europe to extract climate zone information. These extracted data were combined with recorded land cover and soil texture information in the GMEP and BZE surveys to assign PCZ classes to all datapoints. Next, O/T SOC ratios were calculated, porosities estimated according to Eqs. (1) and (2), and the porosity distributions modelled for each index class. Published location information in BZE and GMEP represent approximate positions to protect landowner confidentiality. This may mean that a small minority of BZE sites close to the Atlantic-Continental boundary are assigned to an incorrect biogeographical region (and ultimately, PCZ class). GMEP should be unaffected however, as all of Wales is covered by the Atlantic biogeographical region. All data analyses and display items were generated with R (4.3.0) (R Core Team, 2023) except where this is stated otherwise.

To demonstrate the sensitivity of the O/T SOC indicator to changes in land management, several datasets with data on SOC changes over time under long-term field experiments (LTFEs) were sourced. These LTFEs span multiple decades and represent semi-arid, temperate, and cold-weather climates across Europe. Recorded SOC values were converted to O/T SOC values based on the typical SOC value that corresponds to each LTFE's PCZ, determined from associated information on climate, land cover and soil texture.

Table 2

Pedo-climate zones (PCZs, $n = 12$) derived from univariate regression tree modelling of SOC concentration recorded in the LUCAS data. Columns 2–4 outline the biogeographical region(s), land cover class(es) and soil texture class(es) that constitute each PCZ. Column 5 indicates the predicted mean, standard deviation and coefficient of variation SOC concentration for each PCZ. Columns 6–8 indicate the O/T SOC index category thresholds derived from the LUCAS data.

Pedo-climate zones derived from regression tree modelling	EEA Biogeographical regions (2016 classification)	Land cover classes (LUCAS scheme)	Soil texture classes (HYPRES scheme)	Typical (mean) SOC, g kg^{-1}; [standard deviation, g kg^{-1}]; {coefficient of variation, %}	Low/ intermediate threshold	Intermediate/ high threshold	High/very high threshold
Mediterranean arable sandy soils	Mediterranean	Cropland	Coarse	5.99 [1.53] {25.54}	0.68	0.96	1.53
Mediterranean arable loamy and clayey soils	Mediterranean	Cropland	Fine; Medium; Medium fine; Very fine	12.18 [1.71] {14.04}	0.71	0.99	1.38
Central arable sandy soils	Black Sea; Continental; Pannonian; Steppic	Cropland	Coarse	10.59 [1.75] {16.53}	0.76	0.93	1.28
Atlantic arable sandy soils	Atlantic	Cropland	Coarse	19.11 [1.68] {8.79}	0.7	0.97	1.41
Atlantic and Central arable loamy and clayey soils	Atlantic; Black Sea; Continental; Pannonian; Steppic	Cropland	Fine; Medium; Medium fine; Very fine	16.61 [1.55] {9.33}	0.75	0.96	1.26
Alpine and Boreal arable soils	Alpine; Boreal	Cropland	Coarse; Fine; Medium; Medium fine; Very fine	23.34 [1.87] {8.01}	0.71	0.95	1.31
Southern semi-natural sandy soils	Black Sea; Mediterranean; Pannonian; Steppic	Forest; Grassland	Coarse	12.55 [1.94] {15.46}	0.58	0.96	1.57
Continental semi-natural sandy soils	Continental	Forest; Grassland	Coarse	19.11 [2.07] {10.83}	0.62	0.96	1.57
Southern semi-natural loamy and clayey soils	Black Sea; Mediterranean; Pannonian; Steppic	Forest; Grassland	Fine; Medium; Medium fine; Very fine	23.81 [1.81] {7.6}	0.64	0.99	1.57
Continental semi-natural loamy and clayey soils	Continental	Forest; Grassland	Fine; Medium; Medium fine; Very fine	29.96 [1.76] {5.87}	0.7	1.02	1.45
Cold climate semi-natural sandy soils	Alpine; Atlantic; Boreal	Forest; Grassland	Coarse	29.96 [2.28] {7.61}	0.61	1.01	1.72
Cold climate semi-natural loamy and clayey soils	Alpine; Atlantic; Boreal	Forest; Grassland	Fine; Medium; Medium fine; Very fine	39.65 [2.08] {5.25}	0.69	1.02	1.5

3. Results and discussion

3.1. *Observed/typical SOC ratio as an alternative soil health indicator to SOC/clay ratio*

Plotting SOC/clay against clay content for mineral soils aggregated across the LUCAS surveys from 2009/12, 2015 and 2018 revealed a clear relationship between the 2 variables (Fig. S1a), comparable to that found for German agricultural soils (Poeplau and Don, 2023). This led to coarse-textured soils (clay <18 %) being disproportionately classified as likely to be “Very good” (SOC/clay $\geq$1/8), while clay-rich soils (clay >35 %) were most likely to be “Degraded” (SOC/clay <1/13) (Fig. S1a). This necessitated the development of a new normalised SOC content ratio which would be less dependent on clay content (see Methods).

A total of 12 PCZs were defined from unique combinations of land cover, biogeographical region (as a proxy for climate), and soil texture – the 3 variables determined to be significant predictors according to our regression tree modelling (Table 2). The PCZs are characterised by binary splits in land cover (Cropland vs. Grassland and Woodland) and soil texture (Coarse vs all other texture categories according to the HYPRES (Wösten, 2000) classification scheme), with biogeographical regions generally clustered into variations of wet and cold vs warm and dry groups. Two thirds of PCZs occur in >60 % of all 5000 modelled decision trees, with the “Mediterranean arable sandy soils” and “Mediterranean arable loamy and clayey soils” groups occurring 100 % of the time (Fig. S2). This demonstrates a high level of overall confidence in our delineation of PCZ classes. “Typical SOC” (predicted PCZ-level averages) ranges from 5.99 to 39.65 g kg^{-1}, with higher SOC tending to occur for combinations of semi-natural habitats, wetter and colder climates, and more clay-rich soil textures (Table 2), reflecting similar topsoil trends in Germany and the UK (Drexler et al., 2022; Feeney et al., 2023; Vos et al., 2019). Coefficients of variation of SOC concentrations in each PCZ range from 5.25 % (Cold climate semi-natural loamy and clayey soils) to 25.54 % (Mediterranean arable sandy soils), with most PCZs associated with a coefficient of variation <10 %, suggesting that typical SOC contents are generally well constrained (Table 2). Dividing observed SOC concentration by the PCZ-typical reference value yields a new “Observed/Typical SOC” (O/T SOC) indicator. The O/T SOC thresholds, separating the 4 index classes per PCZ are summarised in Table 2 and Fig. S3.

Plotting O/T SOC as a function of clay reveals no apparent sensitivity to clay content in this new metric (Fig. S1b). This suggests that O/T SOC should be applicable across mineral soils irrespective of their texture. SOC is a key control on soil structure, which together can benefit soil biogeochemical cycling and microbial activity (Mattila, 2023). Therefore, a SOC-based indicator ought to be useful for differentiating between soils of differing structural condition. A comparison of the porosity (as an indicator of soil structure) distributions reveals no differences among the bottom 3 SOC/clay index classes (Fig. 1a). Only when comparing the bottom 3 grades with the top class, “Very good”, is there a statistically significant difference ($p < 0.05$; Table S1). By contrast, porosity distributions appear to be much more clearly differentiated by O/T SOC index classes (Fig. 1b), with statistically significant differences ($p < 0.05$) between all classes (Table S1).

Similarly, O/T SOC index classes appear to capture sharper contrasts in SOC stocks (medians range from 22.38 to 62.54 t ha^{-1}) than their SOC/clay counterparts (medians range from 26.83 to 52.59 t ha^{-1}; see Fig. 1 c & d and Table S2). Additionally, there is no statistically significant difference ($p < 0.05$ level) between the “Moderate” and “Good” SOC stock distributions for SOC/clay (Table S1). Together, these results demonstrate that compared to the SOC/clay ratio, the O/T SOC ratio

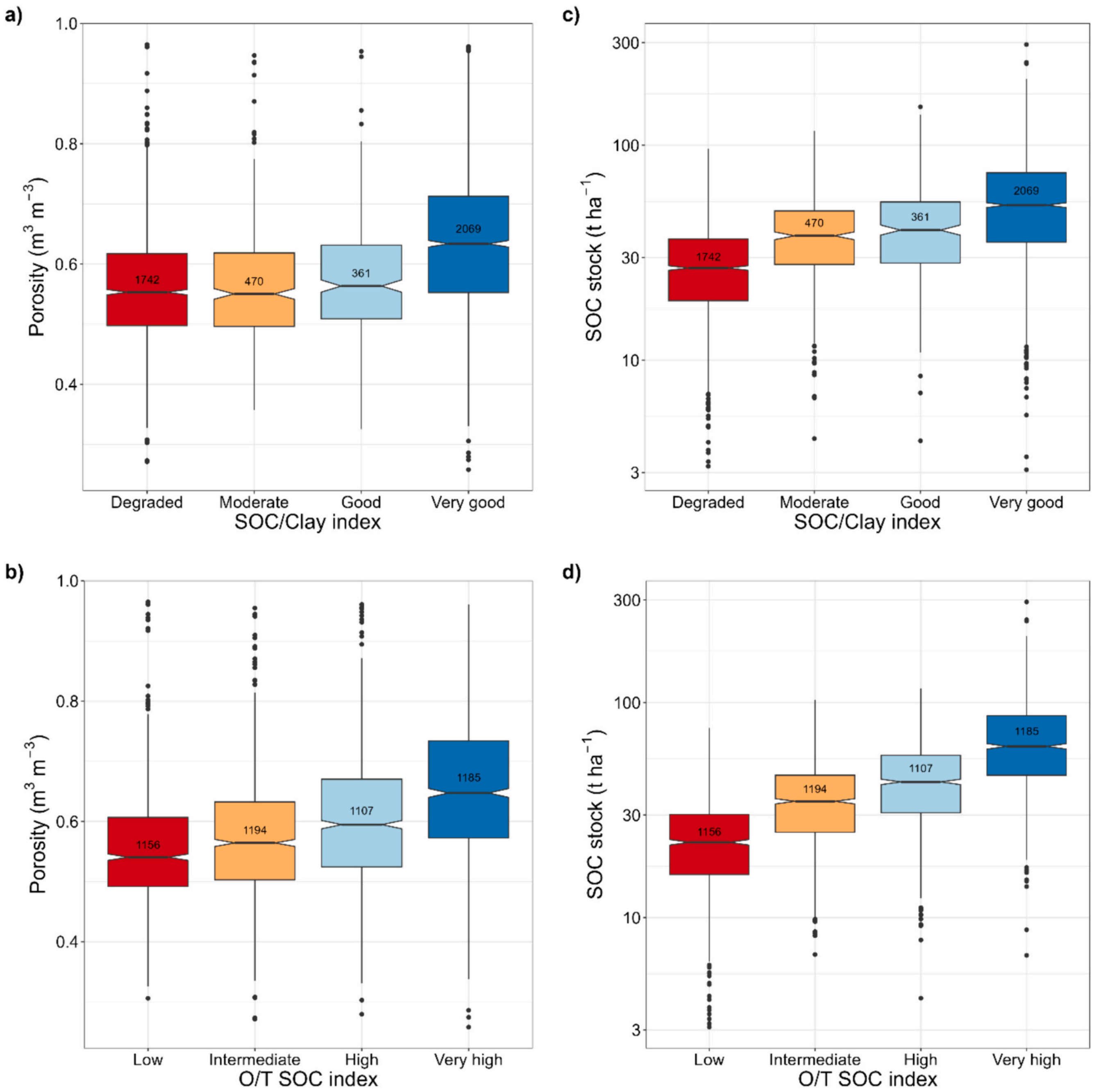


Fig. 1. Porosity distributions by (a) SOC/clay index and (b) O/T SOC index; SOC stocks by (c) SOC/clay index and (d) O/T SOC index. All plots were generated from LUCAS 2018 data with available bulk density measurements. The 25th, 50th and 75th percentiles of all 16 boxplots are listed in Table S2). Numbers on the boxplots refer to the number of LUCAS observations with bulk density measurements in each category.

should be: (i) less sensitive to clay content, and (ii) more appropriately reflect gradients in soil structure and carbon stocks with increasing SOC concentration. Thus, the O/T SOC indicator should be more effective for monitoring progress to 2 of the EU's Soil Health Mission objectives for 2030: conservation and enhancement of SOC stocks; improving soil structure to support soil biodiversity and crops (European Commission, 2023).

LUCAS mineral soils classed as "coarse" under the HYPRES texture classification are disproportionately categorised (>80 %) as having "very good" soil structure with SOC/clay >1/8, whereas similar proportions of "fine" and "very fine" soils are described as having "degraded" soil structure with SOC/clay <1/13 (Fig. 2a). By contrast, there are roughly equal proportions of all 4 O/T SOC index classes for "coarse", "medium" and "medium fine" soils, with a slight skew towards higher proportions of "high" and "very high" O/T SOC ratios for "fine" and "very fine" soils (Fig. 2b). This suggests that soil texture (which cannot be readily modified by landowners) is not the prevailing factor that determines soil health status according to O/T SOC. Furthermore, a critical weakness of the SOC/clay indicator is that croplands are disproportionately likely to be classed as "degraded", while forests are disproportionately likely to be classed as having "very good" soil structure – trends which are clearly visible from LUCAS data (Fig. 3a). The absence of any apparent bias in O/T SOC ratios by land cover class (Fig. 3b), coupled with similar patterns among biogeographical regions (Fig. S4), suggests that the O/T SOC ratio ought to capture soil health differences more effectively within common environments. This means

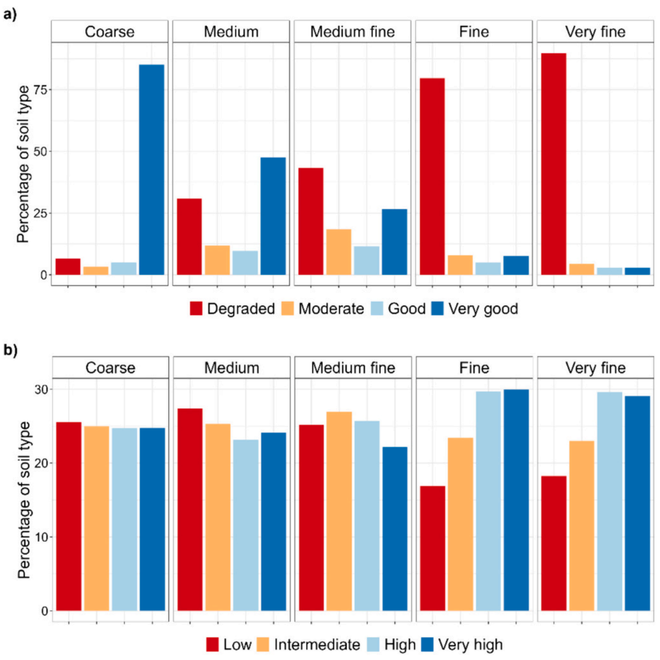


Fig. 2. Percentages of soil types in the LUCAS mineral soils data per SOC/clay index (a) and O/T SOC index (b).

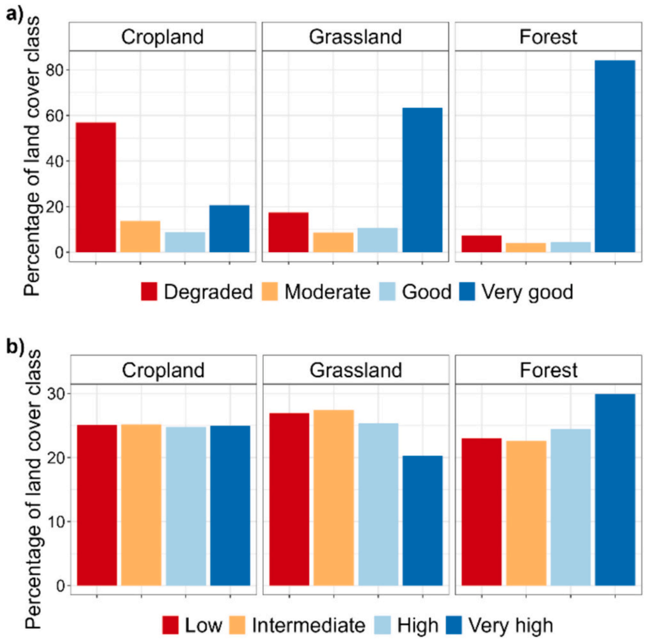


Fig. 3. Percentages of land cover classes in the LUCAS mineral soils data per SOC/clay index (a) and O/T SOC index (b).

that the soil health of arable land in the Mediterranean region (where SOC is likely to be lower than Europe-wide average levels) can be benchmarked more fairly against comparable landscapes; meanwhile, the soil health of forest in cold climates (where SOC is likely to be relatively high) can also be benchmarked against comparable landscapes as well.

3.2. *Spatial patterns and further evaluation of the O/T SOC ratio across Europe*

Mapping the proportions of LUCAS samples per NUTS 1 region highlights the degree to which the 2 extreme SOC/clay index classes (“Degraded” and “Very Good”) are skewed across Europe spatially. >50 % of LUCAS points per NUTS 1 region are classed as having “Degraded” SOC/clay in most regions of southern Europe (Fig. 4a), with 3 of the most extreme examples found in Romania (Table S3). By contrast, northern Europe is overwhelmingly represented by “Very Good” SOC/clay observations (Fig. 4b), with nearly 100 % of observations in Sweden’s “Norra Sverige” region classed as “Very Good” and >90 % of observations in 3 regions of the British Isles classed as “Very Good” (Table S3). These geographical patterns partly reflect the contrast between relatively warm and dry climate conditions in the south and wetter and colder conditions farther north, which are not accounted for in the established SOC/clay index classes. These patterns also seem to reflect the distributions of clay, silt and sand mapped for Europe, with the authors of that study attributing these patterns to the influence of Late Glacial loess deposition (Ballabio et al., 2016).

Repeating the same mapping process for O/T SOC classes reveals no clear geographical trends in either “Low” (Fig. 4c) or “Very Good” (Fig. 4d) O/T SOC condition. Compared to the SOC/clay maps, there are far fewer regions where the number of observations with extreme O/T SOC exceeds 50 %. Additionally, the 5 regions with the highest proportions of “Low” O/T SOC levels have between 48.1 and 63 % of observations rated as “Low”, well below the equivalent range for “Degraded” SOC/clay: 72.7–81.5 % rated as “Degraded” (Table S3). Similarly, the 5 regions most dominated by “Very High” O/T SOC observations have between 47.8 and 60 % of samples rated as “Very High” compared to the 91.8–97.3 % range for the top 5 regions with “Very Good” SOC/clay (Table S3). These figures illustrate both the clustering of soils with extreme SOC/clay ratios compared to extreme O/T SOC ratios, as well as the narrowness between the “Degraded” and “Very Good” SOC/clay categories compared to the gap between “Low” and “Very High” O/T SOC. Thus, incremental improvements in SOC storage and soil structure should be much clearer to track over time and with much less geographical sensitivity using the O/T SOC indicator than the SOC/clay indicator.

As a further demonstration, data from national soil monitoring programmes in Germany (BZE) and Wales (GMEP) were used to calculate O/T SOC and model porosity distributions by index class in these 2 countries. Results presented in Fig. 5a illustrate that porosity increases with O/T SOC and differences in porosity between index classes are all statistically significant ($p < 0.05$ level) for both countries. Median porosity increases from 0.45 $m^3 m^{-3}$ in “Low”-rated soils to 0.52 $m^3 m^{-3}$ in “Very high”-rated soils in Germany, while in Wales, the medians range from 0.54 to 0.72 $m^3 m^{-3}$ across the same index classes (Fig. 5a). These results illustrate how porosity can continue to improve from relatively SOC-poor, largely arable-dominated soils through to relatively SOC-rich, predominantly semi-natural habitat covered soils (Fig. 5b). These patterns also illustrate further the importance of benchmarking soil health by appropriately defined landscape contexts, as what makes a “Very high”-rated soil in German agricultural landscapes would most likely be classed among the “Low”-rated mineral soils occurring in Wales. Importantly, we have demonstrated how O/T SOC status within smaller European regions could be assessed against EU-wide data. This potentially offers EU policymakers an indicator which can be used within future “Soil Health Districts” (discrete geographical regions within Europe for soil health monitoring and assessment) for soil health reporting.

For a soil health indicator to be useful, it should be sensitive enough to detect change over time, including responses to land management interventions (Bünemann et al., 2018). Several long-term field experiments (LTFEs), beginning as early as the 19th Century in some cases, have been undertaken throughout Europe. For those LTFEs where SOC concentration data are available, we can test the sensitivity of the O/T SOC status to change over time under land management interventions. Rothamsted Research contains a rich archive of long-term data from LTFEs in the UK. The LTFEs of Broadbalk and Highfield (Fig. 6) present

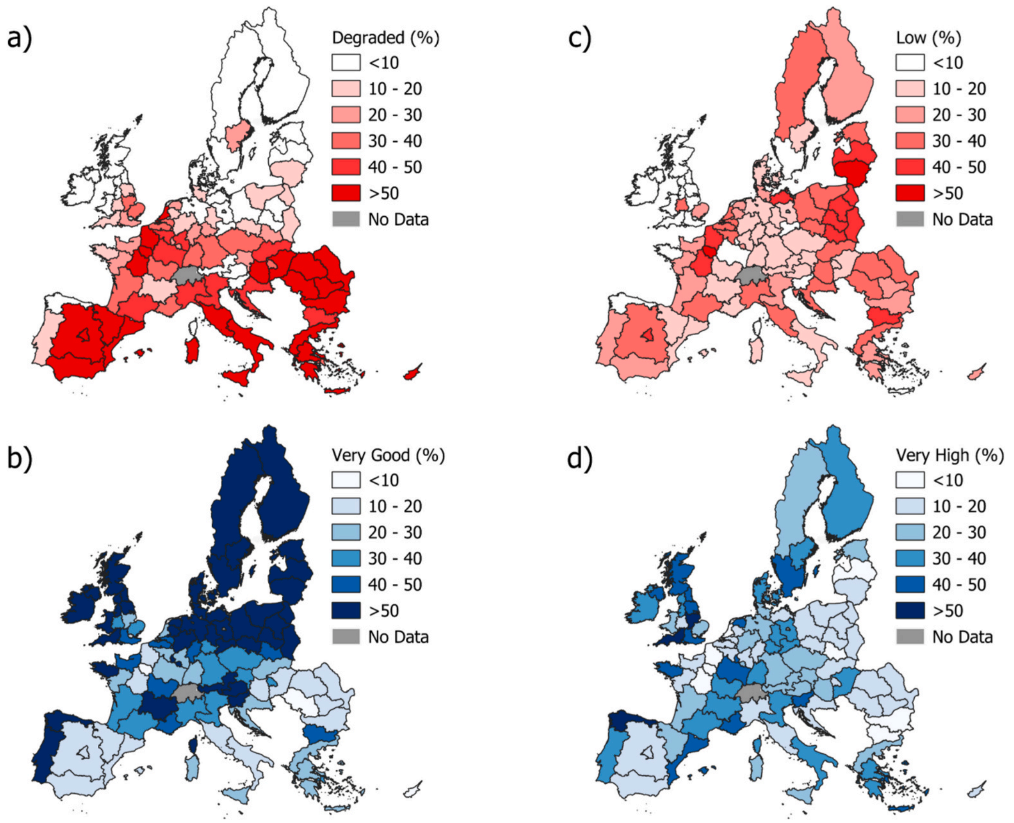


Fig. 4. Percentages of "Degraded" SOC/clay (a), "Very Good" SOC/clay (b), "Low" O/T SOC (c) and "Very High" O/T SOC (d) mapped across the EU + UK by NUTS 1 region. Percentages refer to proportions of LUCAS samples points per NUTS 1 region.

contrasting types of land management and trends in SOC contents over time. As well as being agricultural plots from the same country, both LTFEs consist of similar soil textures (~23–33 % clay), so share the same PCZ class, "Atlantic and Central arable loamy and clayey soils" that allows the 2 LTFEs to be more easily compared.

In the Broadbalk example (Fig. 6a), 6 fertiliser treatments (including a control with no fertiliser application, "nil") were trialled, with treatments starting in 1843 (or later, if stated) with SOC concentration measured at regular intervals (Rothamsted Research, 2021). In this example, the O/T SOC status can shift from "low" to "very high" O/T SOC status over a few decades of farmyard manure (FYM) application. Under FYM treatments, the O/T SOC status of the soil can improve in as little as 9 years in the case of FYM since 1885 and decline in as little as 5 years when FYM applications cease, as in the case of FYM + N inputs (Fig. 6a). The Highfield bare fallow LTFE by contrast recorded the decline of SOC content in multiple plots where long-term grassland (since 1838) was ploughed up in 1959 and left as bare fallow and sampled at regular intervals for multiple decades thereafter (Poulton et al., 2022). Here, O/T SOC declined from "very high" to "high" status across all 4 monitored plots in 12 years (1959–1971), with plot 2 declining to "intermediate" status in another 8 years after that (Fig. 6b). By 2000, all 4 plots had declined to "low" status and remained there for the additional 14 years of the experiment (Fig. 6b).

Changes in O/T SOC status are also visible in other environments across Europe. The Ultuna LTFE in Sweden (Kätterer et al., 2011), representing the "Alpine and Boreal arable soils" PCZ, shows that over 60 years, experimental plots under 8 different organic amendments exhibited a shift from "low" O/T SOC to one of the higher O/T SOC grades (Fig. S5). Poeplau and Don (2023) identified a similar result for their observed/expected SOC ratio in the Ultuna LTFE and remarked that by contrast, only 3 treatments produced a shift in SOC/clay status from "degraded" (and in these cases, only up to "Moderate" status). Monitoring of forest succession following abandonment of cropland in northeast Spain (Bell et al., 2021) shows that after early-stage succession, O/T SOC reaches at least "high" status on average (Fig. S6).

3.3. Potential future applications and limitations

The derivation of an "expected" or "typical" SOC value depends entirely on the population of data used for prediction (Poeplau and Don, 2023). In this sense, the higher and lower recorded SOC concentration values do not necessarily indicate a good or poor-quality soil, but instead, what might reflect typical or atypical conditions for a particular environment, following previous benchmarking studies (Drexler et al., 2022; Feeney et al., 2023). This population-driven issue applies to the O/T SOC indicator too, but the main advantages of the benchmarking approach used here are that newly monitored soils can in future can be assessed more fairly against landscapes with similar characteristics, and that each of the derived O/T SOC thresholds separated differing grades of porosity and SOC stocks across Europe much more clearly than established SOC/clay thresholds did. Thus, O/T SOC index classes should be more meaningful than SOC/clay index classes in describing soil structure and below ground carbon stocks across Europe. Furthermore, a benchmarking approach offers much flexibility, especially for metrics where larger values are generally better than lower values (e.g. SOC concentration) or vice versa (e.g. in the case of bulk density). If a landowner for instance, learns that their soil shows atypical conditions against representative benchmarks (e.g. abnormally low SOC or abnormally high bulk density for their environment), they can better understand which aspects of soil health to improve most urgently. Policymakers meanwhile can set objectives and appropriate timetables (e.g. X percentage of soils with a "low" O/T SOC status should improve by at least 1 O/T SOC grade by year, Y). Additionally, provided that grades still reflect differences in soil condition (e.g. porosity), >4 O/T

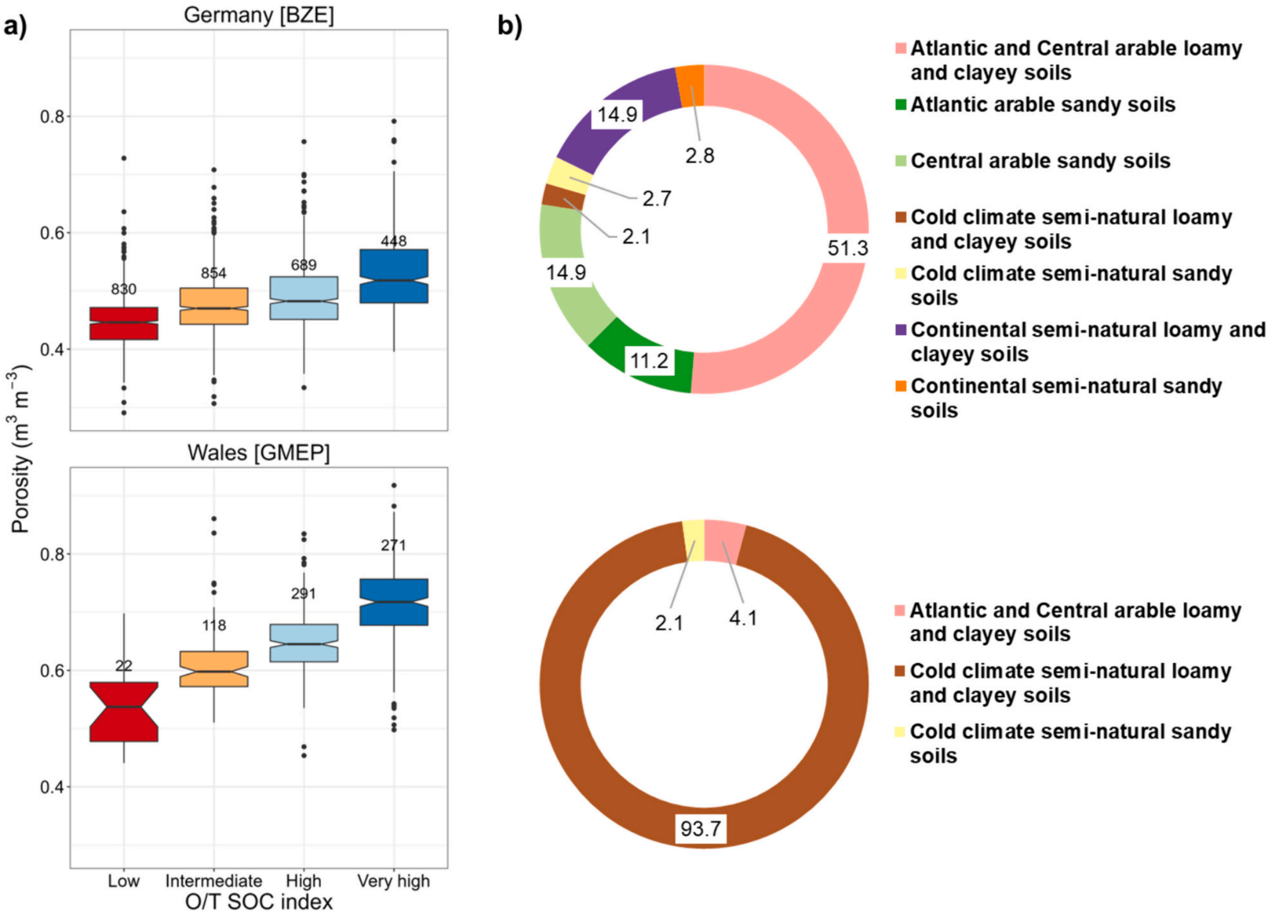


Fig. 5. O/T SOC index classes and associated porosity distributions in Germany and Wales according to national monitoring datasets, BZE and GMEP, respectively (a); PCZs represented in each dataset broken down into percentages (b). Numbers on the boxplots refer to the number of observations in each O/T SOC index class.

SOC grades could be derived to increase the sensitivity of O/T SOC to detect change in response to land management influences. Determining the maximum number of potential O/T SOC grades will require a more thorough sensitivity analysis of soil monitoring datasets and is suggested as an important direction for future research.

Developing a harmonised approach for the EU using the 3 currently available LUCAS surveys (2009/12, 2015 & 2018) as a baseline dataset is arguably important given that ultimately, LUCAS is the most comprehensive and up to date dataset publicly available at European scale. Additionally, there is ample scope to apply our approach to generate PCZs and associated O/T SOC thresholds that are appropriate to other geographical contexts beyond Europe. For instance, datasets, including national soil survey data and a map of North American Ecoregions to represent different bioclimates (Koop et al., 2023), combined with land cover mapping and other contextual spatial data (e.g. elevation) could be applied to determine O/T SOC ratios and establish benchmarks across the USA. Key challenges for any application beyond Europe will include ensuring that PCZs are i) representative of all mineral soil environments across the region; ii) O/T SOC grades are sufficiently constrained to be sensitive enough to detect change over time in response to land management interventions; iii) can be readily identified by stakeholders to benchmark third-party data against regionally defined O/T SOC grades (e.g. following the decision tree approach used here for Europe). More widespread monitoring of O/T SOC ratios internationally could contribute to realising key UN SDGs worldwide (e.g. climate action and life on land), so adapting methods to determine O/T SOC beyond Europe is an important direction for further research.

The EUSO recently introduced the “Soil Health Dashboard”, an interactive web portal of map layers, which provides a convergence of evidence assessment of threats to soil health across Europe (Panagos et al., 2024). Several soil health indicators have been mapped at pan-European scale and are fed into the Soil Health Dashboard to highlight where soils may be classed as either “healthy” (no soil health threats present) or “degraded” (at least 1 soil health threat is present). This system provides a powerful tool for policymakers to see which soils are most under threat and is the basis of the estimate that >60 % soils across Europe are rated “degraded” (Panagos et al., 2024). The O/T SOC could be incorporated in this system and used in conjunction with maps of other soil health indicators to help refine understanding of which soils are degraded across Europe. Details of how O/T SOC could be mapped are included in Text S1 (Supplementary Materials) and we have produced maps of PCZs and O/T SOC grades using high-resolution digital soil maps for Europe from EcoDataCube for illustration (Fig. S7). While in principle, it is straightforward to map O/T SOC grades from digital soil maps, in practice, key limitations should be considered, including: i) the distribution of PCZs (and with them, O/T SOC thresholds) will change over time as land cover changes; ii) predicted PCZ classes and O/T SOC levels are sensitive to the accuracy of the digital soil maps used. Digital soil maps often do not predict extreme values very effectively (Feeney et al., 2022), so areas of PCZs defined by coarse textured soils for instance, may be significantly under-estimated because predicted clay contents at the lower end of the distribution may be over-estimated; conversely, many soils that are mapped as having high mineral SOC concentrations may in reality be organic soils (and therefore should not be considered in O/T SOC assessment) because SOC values at the upper end of the distribution have been under-estimated. Maps of any soil health indicator should be sufficiently robust to provide reasonably accurate assessments of soil health at regional scales, otherwise users

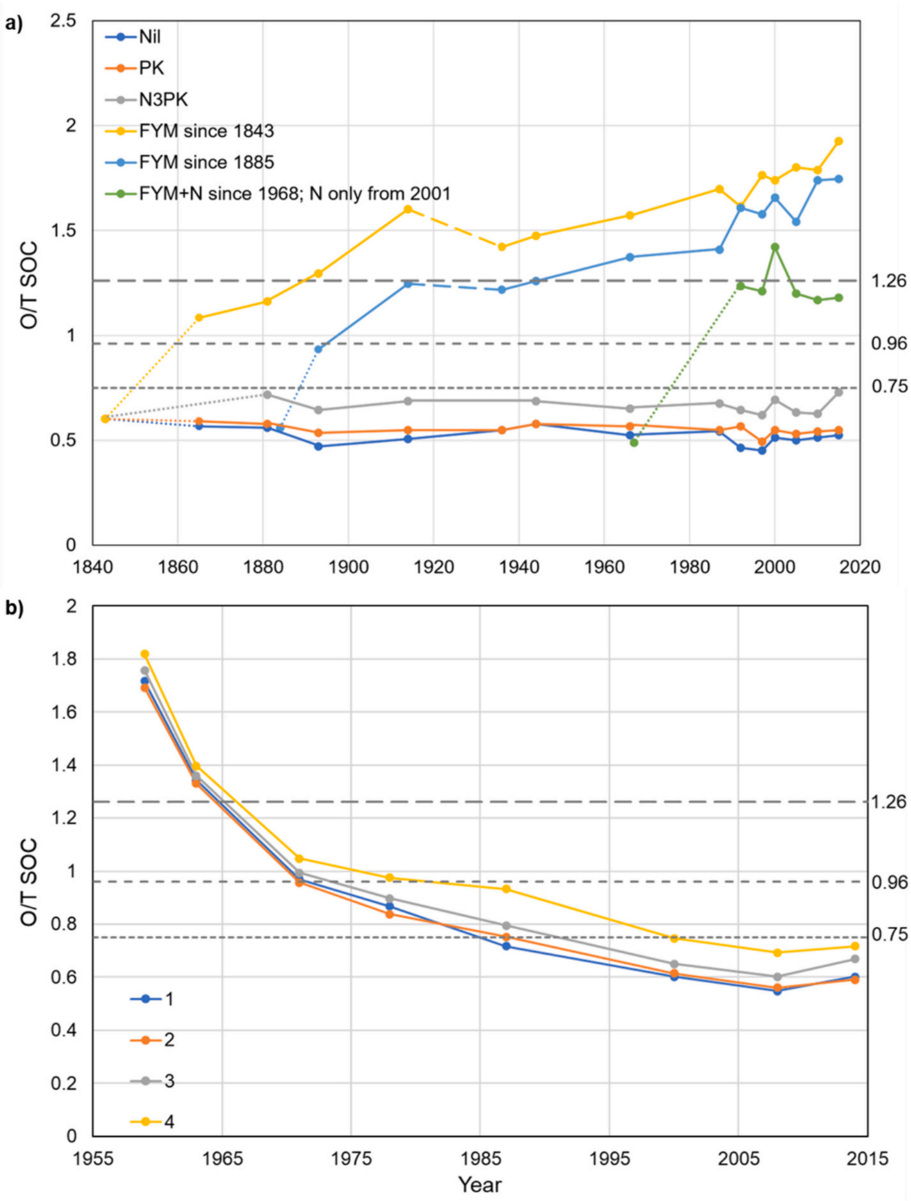


Fig. 6. Changes in O/T SOC over multiple decades under contrasting land management interventions for 2 LTFEs in the UK: Broadbalk (a) and Highfield (b). The Broadbalk LTFE consisted of a series of field plots under winter wheat cover, where different forms of fertiliser application were trialled. Dotted lines indicate where starting SOC contents were estimated; dashed lines show where farmyard manure application paused during fallow periods (Rothamsted Research, 2021). The Highfield LTFE represents a bare fallow experiment across 4 plots along a north-south gradient (1 is most northerly; 4 is most southerly); here long-term grass (since 1838) was ploughed out in 1959 and kept in bare fallow by cultivation to suppress weeds (Poulton et al., 2022). The PCZ for both LTFE examples is “Atlantic and Central arable loamy and clayey soils”; limits between the 4 O/T SOC categories for this PCZ are indicated by the horizontal grey lines with thresholds labelled.

may be presented with misleading assessments that reflect weaknesses in the maps themselves rather than current soil health status. Judgements of what makes a soil health map “robust”, alongside other considerations including the flexibility to choose different combinations of soil health indicators and rank order their relative importance depending on user priorities, are all important areas for further research to support tools for policymakers and other stakeholders.

Some limitations should be noted for O/T SOC. First, this new metric only applies to mineral topsoils. In large part, this was done for comparison with the SOC/clay ratio, which was also designed for mineral topsoils. There is also much wider variability of SOC concentrations for organic (≥120 g kg^{-1}) compared to mineral soils (<120 g kg^{-1}), and it is unclear whether O/T SOC index classes could usefully separate differing grades in organic soil structure. Robinson et al. (2022) identified a reduction in the rate of porosity increase with SOM content around the mineral-organic soil transition (20 % soil organic matter). It is also likely that a different configuration of environmental factors will best predict SOC concentration deeper in the soil (Vos et al., 2019); thus, a different set of “typical” SOC concentrations may be necessary for subsoils.

Second, data from all 3 extant LUCAS surveys were aggregated together to develop the O/T SOC indicator. This was done primarily to maximise the data available to generate PCZs and “typical” (average) SOC concentrations in as statistically a robust way as possible. While this does prevent any change over time analysis from being performed with LUCAS, it has been demonstrated from several LTFE plots, spanning semi-arid to cold-weather environments, that changes in SOC content can be reflected by changes in O/T SOC status over management timescales (multiple years to decades).

Third, additional data are required to confirm whether increases in O/T SOC would translate into increases in soil porosity and SOC stocks.

Moreover, additional metrics related to soil structure (e.g. aggregation and effective porosity) should be compared across O/T SOC grades to further validate their robustness. This will require different datasets to LUCAS, potentially from a wide range of pilot sites across Europe.

Lastly, SOC is not a universal indicator of soil health. Indeed, soils can have a “Very High” O/T SOC condition but also be afflicted with other issues such as excess nitrogen and phosphorous or high concentrations of metallic elements such as lead and copper. O/T SOC should therefore be complimented with other indicators, reflecting various physical, chemical and biological aspects of soils to assess soil health, similar to the CASH framework (Moebius-Clune et al., 2017) in the USA.

4. Conclusion

The EU Soil Strategy 2030 aims to increase SOC to improve soil health and promote carbon sequestration to help mitigate climate change. Tracking progress towards these aims under the proposed EU Soil Monitoring Law requires a SOC-based indicator that is applicable across Europe, irrespective of environmental and management factors, can separate differing grades of soil structure and carbon storage effectively, and is sensitive enough to reflect the effects of land management interventions. Building on an existing blueprint to benchmark observed SOC against a predicted value, which hitherto was based on using clay alone as a predictor and for agricultural landscapes (Poeplau and Don, 2023), we present the “observed/typical SOC” (O/T SOC) ratio. O/T SOC represents an important advancement, as it is derived from modelling that considers land cover and climate, in addition to soil texture, as important predictors, which allows appropriate SOC levels to be determined in different pedo-climatic zones (PCZs). Applying regression tree modelling to Europe’s largest active soil monitoring dataset (LUCAS), 12 PCZs, representing the full breadth of mineral soil environments across Europe, were delineated, each with a predicted average SOC concentration which defined the “typical” SOC. Dividing observed SOC values in each PCZ by the corresponding typical SOC value and splitting each population into quartiles defined the O/T SOC grades, which were termed “low”, “intermediate”, “high” and “very high”.

We successfully demonstrated that O/T SOC better reflects differing grades of soil porosity, as a proxy for physical structure, and SOC stocks than the SOC/clay ratio, establishing O/T SOC as a more useful indicator of the technical potential for future SOC storage for both soil structure enhancement and land-based carbon sequestration. The flexibility of our indicator design means that O/T SOC is applicable to datasets throughout Europe, which we demonstrated by assessing porosity distributions from national soil monitoring data in 2 European countries. Analysis of multi-decadal time-series of SOC concentration changes from several long-term field experiment plots in temperate, semi-arid and cold-weather climate environments shows that O/T SOC can be useful for detecting responses to various land management interventions. This opens the possibility of incorporating O/T SOC as a soil health indicator for the EU and its member states to track progress under the proposed EU Soil Monitoring Law, or for guidance to landowners to engage in voluntary monitoring. With adaptation to other datasets, the O/T SOC ratio could be incorporated into decision support tools such as the EUSO Soil Health Dashboard for policymakers. It could also serve as an important indicator to guide efforts to enhance soil health beyond Europe, ultimately contributing to the delivery of several UN SDGs worldwide.

CRediT authorship contribution statement

Christopher J. Feeney: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laura Bentley:** Writing – review & editing, Methodology. **Daniele De Rosa:** Writing – review & editing, Data curation. **Panos Panagos:** Writing – review & editing, Data curation. **Bridget A. Emmett:** Writing – review & editing. **Amy Thomas:** Writing – review & editing. **David A. Robinson:** Writing – review & editing, Supervision, Project administration, Methodology, 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

LUCAS soils data are available from the JRC European Soil Data Centre (ESDAC) for 2009–12 from https://esdac.jrc.ec.europa.eu/content/lucas-2009-topsoil-data, 2015 from https://esdac.jrc.ec.europa.eu/content/lucas2015-topsoil-data and 2018 including bulk density data from https://esdac.jrc.ec.europa.eu/content/lucas-2018-topsoil-data. The EcoDataCube digital elevation map and Geomorphon90 landform classes can be accessed from OpenGeoHub: https://opengeohub.org/ and OpenTopography https://portal.opentopography.org/dataspace/dataset?opentopoID=OTDS.012020.4326.1, respectively. The land cover, biogeographical regions and EU NUTS administrative boundaries are available from https://land.copernicus.eu/en/products/corine-land-cover/clc2018, https://www.eea.europa.eu/data-and-maps/figures/biogeographical-regions-in-europe-2 and https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts, respectively. GMEP soil survey data for Wales is available from the Environmental Information Data Centre: https://catalogue.ceh.ac.uk/documents/0fa51dc6-1537-4ad6-9d06-e476c137ed09 and BZE soil survey data for Germany is available from the Thünen Institute: https://www.openagrar.de/receive/openagrar_mods_00054877. Data for the Broadbalk and Highfield long-term field experiments are available from the Electronic Rothamsted Archive: https://www.era.rothamsted.ac.uk/index.php.

Acknowledgements

This analysis was undertaken as part of the Horizon Europe project AI4SoilHealth (Grant No. 101086179). We would like to thank the developers of the datasets used in our analysis, particularly the LUCAS, BZE and GMEP soil monitoring teams. Additional comments on the original version of the manuscript from Dr. Fabrizio Albanito (UKCEH) and from 4 anonymous reviewers are also gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.175642.

References

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., Domisch, S., 2020. Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. Sci. Data 7 (1). https://doi.org/10.1038/s41597-020-0479-6.

Assouline, S., Or, D., 2013. Conceptual and parametric representation of soil hydraulic properties: a review. Vadose Zone J. 12 (4), vzj2013-07 https://doi.org/10.2136/vzj2013.07.0121.

Ballabio, C., Panagos, P., Monatanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma 261, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006.

Bell, S.M., Terrer, C., Barriocanal, C., Jackson, R.B., Rosell-Melé, A., 2021. Soil organic carbon accumulation rates on Mediterranean abandoned agricultural lands. Sci. Total Environ. 759 https://doi.org/10.1016/j.scitotenv.2020.143535.

Blanco-Canqui, H., Shapiro, C.A., Wortmann, C.S., Drijber, R.A., Mamo, M., Shaver, T.M., Ferguson, R.B., 2013. Soil organic carbon: the value to soil properties. J. Soil Water Conserv. 68 (5) https://doi.org/10.2489/jswc.68.5.129A.

Bonfante, A., Basile, A., Bouma, J., 2020. Targeting the soil quality and soil health concepts when aiming for the United Nations Sustainable Development Goals and the EU Green Deal. Soil 6 (2), 453–466. https://doi.org/10.5194/soil-6-453-2020.

Bouma, J., 2014. Soil science contributions towards Sustainable Development Goals and their implementation: linking soil functions with ecosystem services. J. Plant Nutr. Soil Sci. 177 (2), 111–120. https://doi.org/10.1002/jpln.201300646.

Breuning-Madsen, H., Kristensen, J. A., & Balstrom, T. (2015). Final report on the establishment of a spade 14 soil profile analytical database connected to the EU Soil Map at Scale 1*:1.000.000*.

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. https://doi.org/10.1016/j.soilbio.2018.01.030.

D'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Haub, P., Buck, O., Haub, C., Eiselt, B., Van Der Velde, M., 2022. LUCAS cover photos 2006-2018 over the EU: 874 646 spatially distributed geo-tagged close-up photos with land cover and plant species label. Earth System Science Data 14 (10), 4463–4472. https://doi.org/10.5194/essd-14-4463-2022.

De Rosa, D., Ballabio, C., Lugato, E., Fasiolo, M., Jones, A., Panagos, P., 2024. Soil organic carbon stocks in European croplands and grasslands: how much have we lost in the past decade? Glob. Chang. Biol. 30 (1) https://doi.org/10.1111/gcb.16992.

Dexter, A.R., Richard, G., Arrouays, D., Czyz, 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.

Directorate-General for Environment, 2023. Proposal for a Directive on Soil Monitoring and Resilience. https://environment.ec.europa.eu/publications/proposal-directive-soil-monitoring-and-resilience_en.

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 (3), 427–440. https://doi.org/10.1002/jpln.202200007.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77 (4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

European Commission, 2023. A soil deal for Europe. Https://Research-and-Innovation.Ec.Europa.Eu/Funding/Funding-Opportunities/Funding-Programmes-and-Open-Calls/Horizon-Europe/Eu-Missions-Horizon-Europe/Soil-Deal-Europe_en.

European Environment Agency, 2017. Biogeographical regions in Europe. Environmental Information Systems. https://www.eea.europa.eu/data-and-maps/figures/biogeographical-regions-in-europe-2.

FAO, ITPS, 2015. Status of the World's Soil Resources (SWSR) - Main Report.

Feeney, C.J., Cosby, B.J., Robinson, D.A., Thomas, A., Emmett, B.A., Henrys, P., 2022. Multiple soil map comparison highlights challenges for predicting topsoil organic carbon concentration at national scale. Sci. Rep. 12, 1379. https://doi.org/10.1038/s41598-022-05476-5.

Feeney, C.J., Robinson, D.A., Keith, A.M., Vigier, A., Bentley, L., Smith, R.P., Garbutt, A., Maskell, L.C., Norton, L., Wood, C.M., Cosby, B.J., Emmett, B.A., 2023. Development of soil health benchmarks for managed and semi-natural landscapes. Sci. Total Environ. 886 (April), 163973 https://doi.org/10.1016/j.scitotenv.2023.163973.

Fernandes-Ugalde, O., Scarpa, S., Orgiazzi, A., 2022. LUCAS 2018 Soil Module. https://doi.org/10.2760/215013.

Hassink, J., 1997. A model of the physical protection of organic matter in soils the capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant and Soil 191, 77–87. https://www.researchgate.net/publication/40154117.

Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons-a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156. https://doi.org/10.1016/j.geomorph.2012.11.005.

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 (September 2016), 14–21. https://doi.org/10.1016/j.geoderma.2017.04.021.

Jones, A., Fernandez-Ugalde, O., Scarpa, S., 2020. LUCAS 2015 Topsoil Survey. https://doi.org/10.2760/616084.

Kätterer, T., Bolinder, M.A., Andrén, O., Kirchmann, H., Menichetti, L., 2011. Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment. Agric. Ecosyst. Environ. 141 (1–2), 184–192. https://doi.org/10.1016/j.agee.2011.02.029.

Keesstra, S., Mol, G., de Leeuw, J., Okx, J., Molenaar, C., de Cleen, M., Visser, S., 2018. Soil-related sustainable development goals: four concepts to make land degradation neutrality and restoration work. Land 7 (4). https://doi.org/10.3390/land7040133.

Keesstra, Saskia, Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., Montanarella, L., Quinton, J.N., Pachepsky, Y., Van Der Putten, W.H., Bardgett, R.D., Moolenaar, S., Mol, G., Jansen, B., Fresco, L.O., 2016. The significance of soils and soil science towards realization of the United Nations sustainable development goals. Soil 2 (2), 111–128. https://doi.org/10.5194/soil-2-111-2016.

Koop, A.N., Hirmas, D.R., Billings, S.A., Li, L., Cueva, A., Zhang, X., Wen, H., Nemes, A., Souza, L.F.T., Ajami, H., Flores, A.N., Rudick, A.K., Guthrie, A., Klamm, L.M., Unruh, M., Sullivan, P.L., 2023. Is macroporosity controlled by complexed clay and soil organic carbon? Geoderma 437. https://doi.org/10.1016/j.geoderma.2023.116565.

Kopittke, P.M., Berhe, A.A., Carrillo, Y., Cavagnaro, T.R., Chen, D., Chen, Q.-L., Román Dobarco, M., Dijkstra, F.A., Field, D.J., Grundy, M.J., He, J.-Z., Hoyle, F.C., Kögel-Knabner, I., Lam, S.K., Marschner, P., Martinez, C., McBratney, A.B., McDonald-Madden, E., Menzies, N.W., Minasny, B., 2022. Ensuring planetary survival: the centrality of organic carbon in balancing the multifunctional nature of soils. Crit. Rev. Environ. Sci. Technol. 52 (23), 4308–4324. https://doi.org/10.1080/10643389.2021.2024484.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., 2023. caret: Classification and Regression Training (6.0-94), CRAN. https://cran.r-project.org/web/packages/caret/index.html.

Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science 304 (5677), 1623–1627. https://doi.org/10.1126/science.1097396.

Mäkipää, R., Menichetti, L., Martinez-Garcia, E., Törmänen, T., Lehtonen, A., 2024. Is the organic carbon-to-clay ratio a reliable indicator of soil health? Geoderma 444, 116862.

Mattila, T.J., 2023. Redox potential as a soil health indicator – how does it compare to microbial activity and soil structure? Plant and Soil. https://doi.org/10.1007/s11104-023-06305-y.

Matus, F.J., 2021. Fine silt and clay content is the main factor defining maximal C and N accumulations in soils: a meta-analysis. Sci. Rep. 11 (1), 1–17. https://doi.org/10.1038/s41598-021-84821-6.

Moebius-Clune, B.N., Moebius-Clune, D.J., Gugino, B.K., Idowu, O.J., Schindelbeck, R.R., Ristow, A.J., van Es, H.M., Thies, J.E., Shayler, H.A., McBride, M.B., Kurtz, K.S.M., Wolfe, D.W., Abawi, G.S., 2017. Comprehensive Assessment of Soil Health - The Cornell Framework (3.2). Cornell University.

Moeys, J., Shangguan, W., Petzold, R., Minasny, B., Rosca, B., Jelinski, N., Zelazny, W., Souza, R.M.S., Safanelli, J.L., Caten, A. Ten, 2018. soiltexture: functions for soil texture plot, classification and transformation (1.5.1), CRAN. https://cran.r-project.org/web/packages/soiltexture/index.html.

Moinet, G.Y.K., Hijbeek, R., van Vuuren, D.P., Giller, K.E., 2023. Carbon for soils, not soils for carbon. Glob. Chang. Biol. 29 (9), 2384–2398. https://doi.org/10.1111/gcb.16570.

Or, D., Smets, B.F., Wraith, J.M., Dechesne, A., Friedman, S.P., 2007. Physical constraints affecting bacterial habitats and activity in unsaturated porous media – a review. Adv. Water Resour. 30 (6–7), 1505–1527. https://doi.org/10.1016/j.advwatres.2006.05.025.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2018. 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.

Pacini, L., Yunta, F., Jones, A., Montanarella, L., Barrè, P., Saia, S., Chen, S., Schillaci, C., 2023. Fine earth soil bulk density at 0.2m depth from Land Use and Coverage Area frame Survey (LUCAS) soil 2018. Eur. J. Soil Sci. May, 1–11. https://doi.org/10.1111/ejss.13391.

Panagos, P., Van Liedekerke, M., Jones, A., Montanarella, L., 2012. European Soil Data Centre: response to European policy support and public data requirements. Land Use Policy 29 (2), 329–338. https://doi.org/10.1016/j.landusepol.2011.07.003.

Panagos, P., Van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., Montanarella, L., 2022. European Soil Data Centre 2.0: soil data and knowledge in support of the EU policies. Eur. J. Soil Sci. 73 (6), 1–18. https://doi.org/10.1111/ejss.13315.

Panagos, P., Borrelli, P., Jones, A., Robinson, D.A., 2024. A 1 billion euro mission: a Soil Deal for Europe. Eur. J. Soil Sci. 75 (1) https://doi.org/10.1111/ejss.13466.

Poeplau, C., Don, A., 2023. A simple soil organic carbon level metric beyond the organic carbon-to-clay ratio. Soil Use Manag. 1057–1067. https://doi.org/10.1111/sum.12921 (December 2022).

Poeplau, C., Vos, C., Don, A., 2017. Soil organic carbon stocks are systematically overestimated by misuse of the parameters bulk density and rock fragment content. Soil 3 (1), 61–66. https://doi.org/10.5194/soil-3-61-2017.

Poulton, P.R., MacDonald, A.J., Gregory, A.S., Glendining, M.J., Ostler, R.J., 2022. Highfield Bare Fallow Soil Chemical Properties 1959–2014. Electronic Rothamsted Archive, Rothamsted Research.

Prăvălie, R., 2021. Exploring the multiple land degradation pathways across the planet. Earth Sci. Rev. 220 (May) https://doi.org/10.1016/j.earscirev.2021.103689.

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 (6), 2493–2503. https://doi.org/10.1111/ejss.13012.

Prout, J.M., Shepherd, K.D., McGrath, S.P., Kirk, G.J.D., Hassall, K.L., Haefele, S.M., 2022. Changes in organic carbon to clay ratios in different soils and land uses in England and Wales over time. Sci. Rep. 12 (1), 1–13. https://doi.org/10.1038/s41598-022-09101-3.

R Core Team, 2023. R: A language and environment for statistical computing (4.3.0), CRAN. https://www.r-project.org/.

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. https://doi.org/10.1016/j.geoderma.2024.116829.

Reynolds, B., Chamberlain, P.M., Poskitt, J., Woods, C., Scott, W.A., Rowe, E.C., Robinson, D.A., Frogbrook, Z.L., Keith, A.M., Henrys, P.A., Black, H.I.J., Emmett, B.A., 2013. Countryside survey: national "Soil Change" 1978–2007 for topsoils in Great Britain-acidity, carbon, and total nitrogen status. Vadose Zone J. 12 (2), vzj2012.0114 https://doi.org/10.2136/vzj2012.0114.

Robinson, D.A., Thomas, A., Reinsch, S., Lebron, I., Feeney, C.J., Maskell, L.C., Wood, C.M., Seaton, F.M., Emmett, B.A., Cosby, B.J., 2022. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. Sci. Rep. 12 (1) https://doi.org/10.1038/s41598-022-11099-7.

Rothamsted Research, 2021. Dataset: Broadbalk Soil Organic Carbon Content 1843–2015. Electronic Rothamsted Archive, Rothamsted Research.

Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature 478 (7367), 49–56. https://doi.org/10.1038/nature10386.

Six, J., Conant, R.T., Paul, E.A., Paustian, K., 2002. Stabilization mechanisms of SOM implications for C saturation of soils. Plant and Soil 241 (2), 155–176.

Therneau, T., Atkinson, B., Ripley, B., 2022. rpart: Recursive Partitioning and Regression Trees (4.1.19), CRAN. https://cran.r-project.org/package=rpart.

Thomas, A., Seaton, F., Dhiedt, E., Cosby, B.J., Feeney, C., Lebron, I., Maskell, L., Wood, C., Reinsch, S., Emmett, B.A., Robinson, D.A., 2024. Topsoil porosity prediction across habitats at large scales using environmental variables. Sci. Total Environ. 922 https://doi.org/10.1016/j.scitotenv.2024.171158.

Tóth, G., Jones, A., Montanarella, L., 2013a. LUCAS Topsoil Survey: Methodology, Data, and Results. https://doi.org/10.2788/97922.

Tóth, G., Jones, A., Montanarella, L., 2013b. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. Environ. Monit. Assess. 185 (9), 7409–7425. https://doi.org/10.1007/s10661-013-3109-3.

Veerman, C., Correia, T.P., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E.A., Grand, A., Filchew, L.H., Kriaučiūnienė, Z., Pogrzeba, M., Soussana, J. F., Olmo, C.V., Wittkowski, R., 2020. Caring for Soil Is Caring for Life – Ensure 75% of Soils Are Healthy by 2030 for Food, People, Nature and Climate. https://doi.org/10.2777/611303.

Visser, S., Keesstra, S., Maas, G., de Cleen, M., Molenaar, C., 2019. Soil as a basis to create enabling conditions for transitions towards sustainable land management as a key to achieve the SDGs by 2030. Sustainability 11 (23). https://doi.org/10.3390/su11236792.

Vos, C., Don, A., Hobley, E.U., Prietz, R., Heidkamp, A., Freibauer, A., 2019. Factors controlling the variation in organic carbon stocks in agricultural soils of Germany. Eur. J. Soil Sci. 70 (3), 550–564. https://doi.org/10.1111/ejss.12787.

Wickham, H., 2023. tidyverse: Easily Install and Load the 'Tidyverse' (2.0.0), CRAN. https://cran.r-project.org/web/packages/tidyverse/index.html.

Wösten, J.H.M., 2000. The HYPRES database of hydraulic properties of European soils. Advances in GeoEcology 32, 135–143.