



OPEN

DATA DESCRIPTOR

A global database on land use and management change effects on soil KMnO₄-oxidisable organic carbon (POXC)

Cécile Chéron-Bessou^{1,2,3}✉, Damien Beillouin^{4,5}, Alexis Thoumazeau^{1,2},
Lydie Chapuis-Lardy⁶, Tiphaine Chevallier⁶, Julien Demenois^{7,8} & Paul N. Nelson³

Soil carbon transformation is vital for ecosystem functions like food production and climate regulation. While soil organic carbon is a key soil health indicator, its sensitivity to management changes is debated. Alternative indicators, such as permanganate-oxidisable carbon (POXC), are being explored. This database compiles 10,068 comparisons of soil POXC content from 284 peer-reviewed studies published up to 2023, covering 45 countries and 63 land use types, including arable land, grassland, agroforestry, and forests. Most studies focused on arable land ($n = 7,809$), examining input changes ($n > 500$) and tillage intensity ($n > 200$). The most studied land-use changes were grassland conversion to arable land ($n = 324$) and vice versa ($n = 261$). The dataset includes rich metadata on geographical context, soil types, key properties (pH, clay content), POXC protocols, and data quality scores. This resource supports scientific and policy discussions on POXC's potential as a practical indicator for improving land use and soil health management.

Background & Summary

Understanding and preserving soil health is critical to ensuring its continued ability to support life. Soils are at the core of most provisioning and regulating ecosystem services^{1–3}. Despite its crucial role in both the ecosystem and the economy, soil degradation not only continues but is accelerating^{4,5}. Human activities, such as deforestation, unsustainable agricultural practices, and urbanization, degrade soil health. Worldwide, at least 33% of croplands are moderately or highly degraded^{6,7}. Land degradation affects 3.2 billion people, and results in an economic loss of approximately 10% of annual global gross product. These numbers stress that “an urgent step change in effort is needed to prevent irreversible land degradation and to accelerate the implementation of restoration measures”⁸. Both scientific research and policy discussions emphasise that achieving this goal requires standardised methodologies and regulations for monitoring, reporting and verifying (MRV) soil measurements that are sensitive to management practices^{9–12}.

Soils are multifunctional, and it is impossible to maximize the provision of all ecosystem services simultaneously^{13,14}. In agricultural soil management, trade-offs are unavoidable —whether between different services or across spatial and temporal scales— making their analysis, assessment and valuation challenging^{13–15}. The provision of ecosystem services is driven, among other factors, by carbon transformation^{16–18}. Soil carbon transformation is a continuous process driven by living organisms, influencing a wide range of ecosystem services, from food production to climate regulation. As a result, soil organic matter and embedded carbon have been central to many studies on soil health. In recent years, soil organic carbon (SOC) storage has also attracted increasing attention in relation to climate change mitigation strategies¹⁹. Various studies and databases on SOC have been developed^{20–22}, and SOC has been used as one of the proxy indicators for soil health^{23–25} or to assess links between soil quality and agronomic outputs such as crop yields²⁶. However, authors have also concluded that

¹CIRAD, UMR ABSys, ELSA Group, F-34398, Montpellier, France. ²ABSys, Univ Montpellier, CIRAD, Montpellier, France. ³JCU, College of Science and Engineering, QLD-, 4878, Smithfield, Australia. ⁴CIRAD, UPR HortSys, F-34398, Montpellier, France. ⁵HortSys, Univ Montpellier, CIRAD, Montpellier, France. ⁶IRD, UMR Eco&Sols, Montpellier, France. ⁷CIRAD, UPR AIDA, F-34398, Montpellier, France. ⁸AIDA, Univ Montpellier, CIRAD, Montpellier, France. ✉e-mail: cecile.bessou@cirad.fr

SOC may lack sufficient sensitivity to management practices to effectively guide soil health protection and land restoration strategies^{27,28}. SOC comprises a diverse range of organic compounds with varying functions, mineral associations, and turnover rates, including highly stable compounds with long residence times. Therefore, SOC has limited utility for detecting rapid changes in soil quality or identifying functional shifts^{29–31}, highlighting the need for more sensitive indicators.

Labile or active carbon pools represent a significant fraction of SOC, typically accounting for 10–20% of SOC, and play a key role in soil microbial activity and nutrient cycling. By definition, such pools have a rapid turnover, as they are more readily available to the soil food web as an energy source^{32–34}. Changes in pool size may hence indicate variations in the activity of soil organisms driven by changes in soil management^{30,35,36}. Permanganate oxidisable carbon (POXC) is one of the labile carbon pools quantified through oxidation by potassium permanganate (KMnO₄). First measurements of POXC date back to 1987, but the most widely used protocols were established in 1995³⁷, followed by further adjustments^{38–41}, with the POXC acronym first introduced in 2003⁴². Several studies have shown that POXC is more sensitive to management changes than SOC, suggesting that POXC can be a valuable metric for quickly tracking management-induced changes in soil organic matter^{33,43–47}. Moreover, POXC measurement relies on a cost-effective, straightforward protocol, which has contributed to its popularity as a soil health indicator^{48,49}. Nevertheless, other studies have pointed to discrepancies among protocols^{50–52} and questioned the direct links between chemical oxidisability, biological availability, and their relationship to soil processes^{49,53}. There is still a critical knowledge gap regarding the nature of the organic materials identified by the KMnO₄ method, and how to interpret the sensitivity of POXC to changing management practices compared to other SOC parameters^{38,40,54}. This database⁵⁵ aims at providing comprehensive POXC datasets and metadata to facilitate further investigations on how POXC relates to soil management and health.

This database⁵⁵ provides a worldwide comprehensive and quantitative assessment of available datasets on the effects of major soil use and management interventions on POXC. It contains 10,068 quantified effect sizes, representing changes in POXC due to land use or land management changes, compared to a control. Where applicable, we considered both the original controls of the studies and/or compared parallel treatments against each other within the same studies as they also inform management changes. Data were published in 284 studies across 45 countries. We provide the locations of the primary studies (WGS84 datum), the type of interventions, applied protocols, and expert-cross-checked information on soil types, offering an opportunity to identify knowledge gaps. Through an iterative process, we defined archetypes for the various interventions in order to cover the main land uses, forestry managements and agricultural practices encountered across the studies. These archetypes were necessary to harmonise the assessment of land uses and management practices across the studies. In total, we covered 63 archetypes including 53 for agricultural management (embedding grassland and agroforestry), 4 for forest plantation management, 4 for natural vegetation stands, then fallow/marginal land and bare soil. Finally, to guide future work, we also identified transparency and quality issues based on a harmonised quality appraisal grid aggregating three criteria into a final quality score. This database is intended to advance the scientific and policy debate on the interest of using POXC as a tool to help and guide improved land use and management practices aimed at improving or maintaining soil health. In particular, it can support the refinement of characterisation factors within life cycle assessment as a component of soil health assessment. An overview of the database content is shown in Fig. 1.

Methods

Search process. The data collection followed Reporting Standards for Systematic Evidence Syntheses (ROSES) guidelines v3.2018⁵⁶, and adhered to the gold standard of systematic review methods^{57,58}. Our methodology is based on a repeatable paper selection procedure, provides a complete list of references, and ensures dataset availability.

The literature search was initially conducted on February 03, 2023 and updated on February 02, 2024 (Fig. 2), and aimed to identify relevant peer-reviewed articles without language restriction. The following searching equation was used: “Soil AND carbon AND (permanganate OR POXC OR KMnO₄)”, applied to the “topic words” (titles, abstracts and keywords). The search was carried out in 4 databases: Web of Science (WoS) core collection; Scopus; OVID; and Google Scholar. In Google Scholar, only the first 200 search results in English, ranked by relevance, were reviewed, as this engine is highly precise primarily within the top-ranked results⁵⁹. In addition to the database searches, 15 primary studies or datasets identified in meta-analyses and systematic syntheses within the initial corpus were manually added during the review process.

No restriction was applied regarding the year of publication, except for Google Scholar, where the search was limited to studies published between 2003 and 2023. This constraint aimed at reducing the number of duplicates potentially added through Google Scholar, while still incorporating the most prolific publication period (Fig. 1a). All climatic zones and countries were considered. We prioritised sensitivity over specificity, meaning the search strategy aimed to collect the largest selection of potentially relevant studies at the cost of retrieving a larger number of non-relevant studies. This strategy increased the duration of the screening step and the number of final rejected studies. In total, 1,091 studies were identified, including 696 unique records, which were compiled as candidates for inclusion (Fig. 2).

Article screening. All identified articles were screened based on their abstract and material and methods sections. When necessary, the full text was reviewed. The inclusion and exclusion criteria were defined using the PICOC framework: Population (i), Intervention (ii), Comparator (iii), Outcomes (iv), Context (v). For a study to be included, the following criteria had to be met:

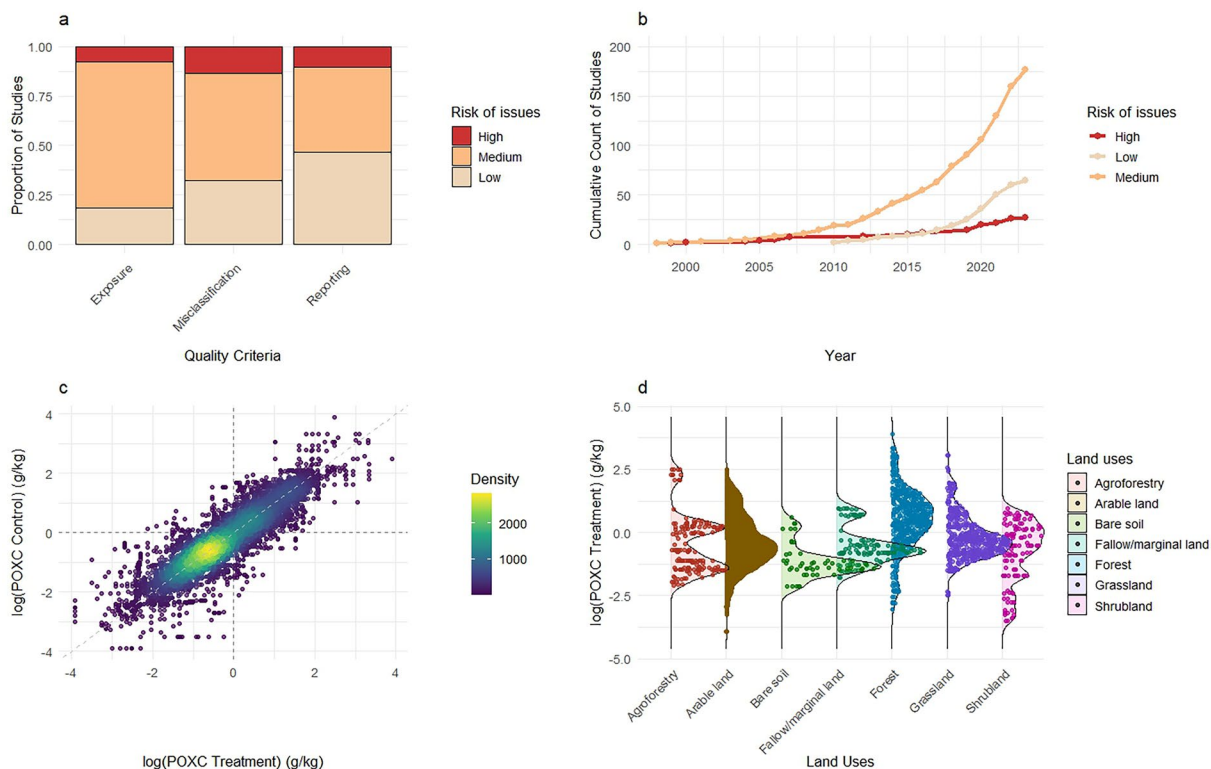


Fig. 1 Overview of the POXC database V2 showing the study quality appraisal (a) and final score with tendencies along publication years (b), as well as a quantitative display of POXC measurements according to treatment and control (c), and over the various studied land uses (d).

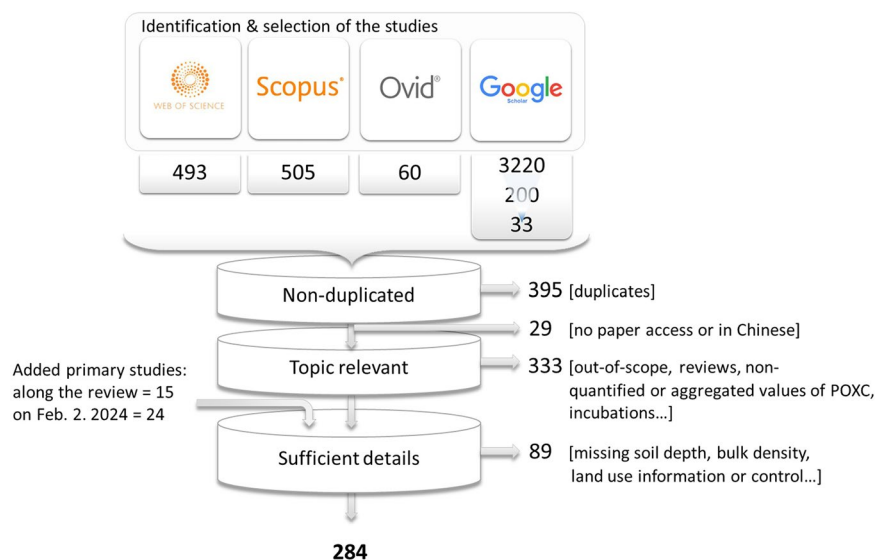


Fig. 2 Identification and selection of the studies. For Google Scholar: 3,220 were the search total results; 200 were displayed on the first 10 pages that are the relevant ones; and 33 were the pre-selected studies remaining after rejection of the clearly out-of-scope search results. Added primary studies consisted in i) 15 primary studies with consistent POXC measurements identified through digging into meta-analysis and syntheses; and ii) 24 studies identified through the same literature search updated on February 2nd, 2024.

- i. **Effect on what?** The primary study examined the amount of *in situ* soil carbon that reacted with potassium permanganate (KMnO₄) following a published protocol (protocols were recorded);
- ii. **Effect of what?** The primary study analysed the effect of land use changes, encompassing shifts in land use type (natural or more or less exploited) or shifts in agricultural and forestry land management practices.

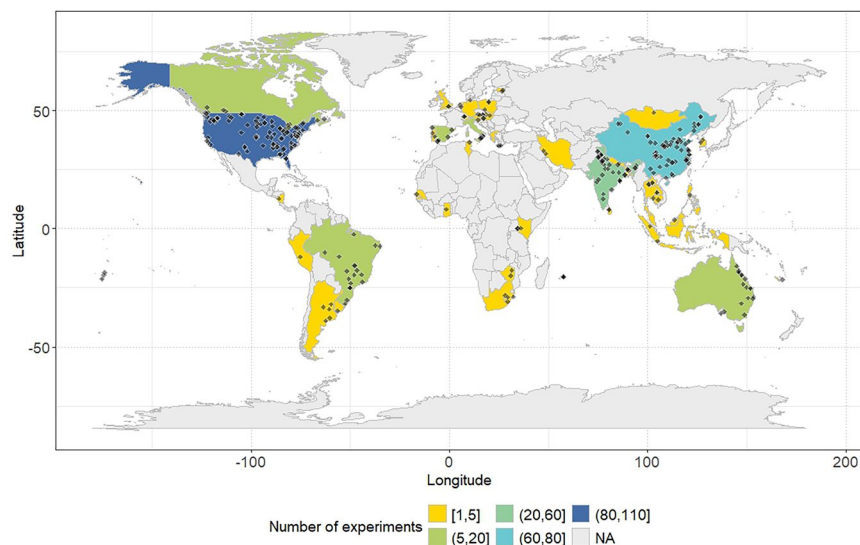


Fig. 3 World map indicating the location of experiments where POXC data were recorded at country level.

- Industrial or urban land uses were not included in the study. Studies using KMnO_4 for bioremediation, rather than for assessing soil carbon stock related to a specific land use or management, were also excluded. Additionally, studies lacking any information on land use type or management were not considered;
- iii. **Effect compared to what?** The change in land use or land management was compared to a control in a similar pedoclimatic context;
 - iv. **What is extracted?** The primary study provided the mean POXC concentration (in g.kg^{-1}) over a given soil depth, along with a measure of dispersion (e.g., standard deviation (SD) or standard error of the mean (SE or SEM)), and the number of replicates for both the control and the treatments. Studies that expressed POXC in units other than g.kg^{-1} were included only if sufficient information (e.g., bulk density) was available for conversion to g.kg^{-1} .
 - v. **In which contexts is data extracted?** POXC measurements should be done in undisturbed soil to reflect actual *in situ* soil functioning. Therefore, lab incubations or measurements on chemically pre-treated soils were excluded.

The exclusion criteria were cross-checked by two authors on a subset of the recorded studies and refined where necessary. Each study was then screened by a single author, and discussions were held among co-authors to resolve any uncertainties regarding specific studies. Ultimately, the article authors were contacted directly via emails and Research Gate messages to shed some light on uncertainties or missing information, yielding a response rate of 40% (10 out of 25 inquiries). All final extractions were reviewed by the lead author (*see* Data validation).

Finally, POXC data were extracted from 284 studies and included in the database (Fig. 2). The rejected studies were compiled in an exclusion sheet, with reasons for their exclusion.

Characterisation of the studies and extracted information. For each study, we documented author names, year of publication, title, journal name, abstract, keywords, DOI and URL when available. Additionally for each included study, we further extracted the following meta-data: KMnO_4 protocol reference, location measurement name and/or GPS coordinates, mean annual precipitation and temperature (when reported), soil type and some initial soil variables such as SOC, pH, and clay content (if available), the type of field trial, number of replicates and number of POXC measurement dates.

Some studies presented results for different contexts, or different experimental set-up. All information was extracted as independent experiment, and resulted in multiple entries in the *C_RETAINED.STUDIES_POXC_DB.v1.1.csv* file. Similarly, some studies presented experimental data for different treatments against a control plot. When comparisons between these plots were conducted on the same soil type, consistent soil management histories, and identical protocols, all possible combinations of comparisons were included in the database, even if they were not explicitly presented in the original paper. This led to multiple entries in the *D_EFFECT_SIZES_POXC_DB.v1.1.csv* file.

Coordinates were harmonised and, when necessary, converted to the WGS84 projection, using the free tool “Online converter to all coordinate systems” (<https://coordinates-converter.com/en>). Our database encompasses data from 45 countries, with highest numbers of experiments in the United States, followed by China, India, and Australia, Brazil and Canada. In contrast, Africa and Europe were the least investigated regions (Fig. 3). Based on the coordinates and information on historical mean annual precipitation and temperature from World Clim⁶⁰ (<https://www.worldclim.org/data/worldclim21.html>), we plotted the recorded data points on a Whittaker biome plot. This analysis reveals that both tropical and temperate biomes (except for the seasonal forest) are well represented in the dataset, while boreal and tundra biomes are notably underrepresented (Fig. 4).

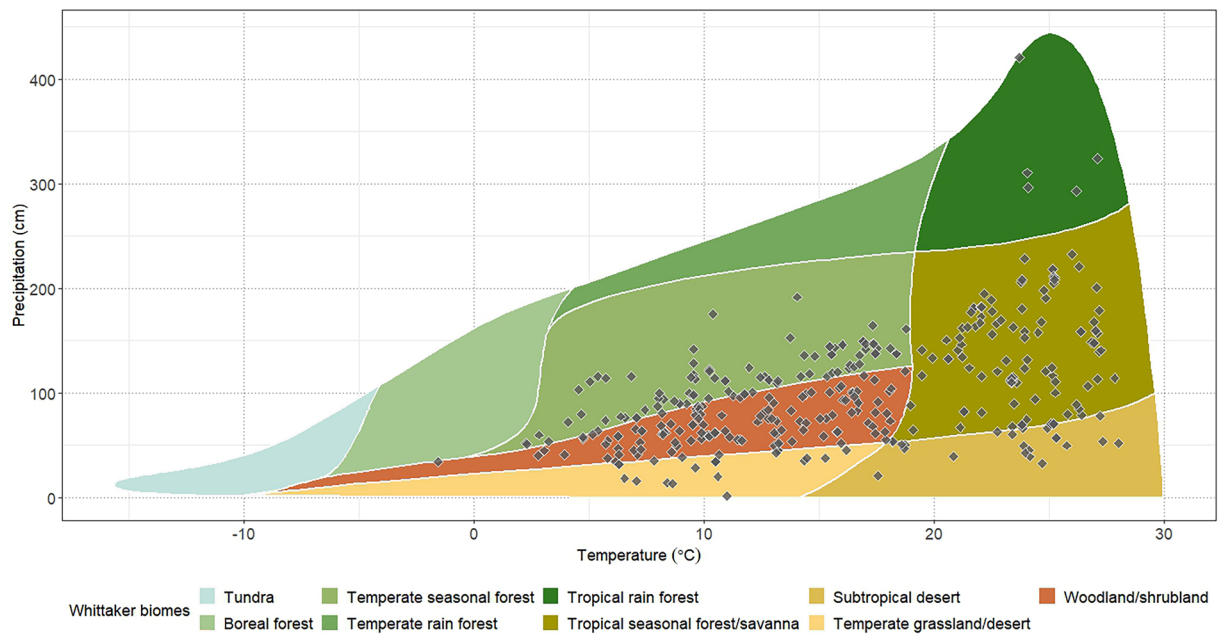


Fig. 4 Whittaker plot showing the biomes represented by POXC recorded data points based on local mean annual precipitation and temperature over 1970–2020.

Mean POXC values per treatment and corresponding replicate numbers, at specified soil depth, were collected manually from tables or using the free online tool “Extractable” (<https://www.extractable.com/>), and from figures using another free online tool “WebPlot Digitizer software” (<https://automeris.io/>). Information was reported under various terms including POXC, permanganate oxidisable carbon (POC), labile carbon (LC), active carbon (AC), or reactive organic carbon (ROC), among others. For consistency, all these measurements are hereafter referred to as POXC.

We collected data on the precision of the POXC means, mainly using standard deviation (SD) or standard error of the mean (SEM). In cases where the nature of the dispersion value was unspecified, we assumed it to be SEM by default to avoid underestimating variability, while also keeping track of the added uncertainty. In a few cases, we also estimated SD from 95% confidence intervals⁶¹ ($n = 12$ studies). In other cases where dispersion information was not provided but results of statistical tests were available (e.g., letters indicating significance from Tukey’s test or Least Significant Difference values), we used the free EX-TRACT tool_1.2 to estimate pooled standard deviations⁶² ($n = 75$ studies) (Supplementary Table 1). For the estimated SD obtained from this tool, we recorded the maximum pooled SD value to avoid overconfidence in the results. Furthermore, in cases of uneven replicate numbers among treatments and controls, assumptions were made to maximize pooled SD and mitigate error underestimation. This global strategy may help balance bias and account for underrepresented studies that show no effect⁶³. Eventually, for the last studies lacking any dispersion information, we calculated the median coefficients of variation within subsets of studies with similar soil types, soil depths and land uses using only data where the original SD was available. We then applied these median coefficients of variation to the means of treatments to compute missing SD values.

We developed a comprehensive land use classification designed to analyse effects in ways that facilitate practical application. Land uses and management practices can be described in varying ways according to the studies, which may hamper any cross-cutting analysis. To address this issue, we classified land uses and land management practices into classes and subclasses. Building on the IPCC^{20,64} classification, we first established higher-order classes: cropland, forest, grassland, wetland, settlement and other land. Settlement and other land categories were excluded as they fell outside the study’s scope, while shrubland, fallow/marginal land and bare soil were added. Wetland was not treated as a separate class to avoid overlapping with soil type classification. To further refine our classification, cropland was subdivided into annual crops, perennial crops and agroforestry systems. By default, “arable land” refers exclusively to annual crops, while specific terms such as “arable land, perennial” and “agroforestry” are used for the latter categories.

Within each land use class, we then defined sub-classes to capture the diversity of managements practices. The sub-classes were iteratively refined during the review process in order to cover all the case studies, while maintaining sufficient number of occurrences within each sub-class for meaningful analysis. Arable land classes were not further disaggregated by crop type, due to the large variety of crops. The only exception was rice, which was covered in many studies and required specific practices related to irrigation mode. Forest types and forest plantation managements were also further differentiated. In total, we defined and applied 63 combinations of land use classes and sub-classes, also called archetypes (Table 1).

Archetypes were defined for each treatment and each control based on the primary management factor investigated. The data dictionary provides all the details about these archetypes. In some cases, a single treatment (or control) could fit into multiple archetypes due to the combination of different management practices.

High-order land use classes	Management types	Management sub-classes	Archetypes numbers
arable land arable land, perennial crop agroforestry forest forest, plantation grassland shrubland	Not specified or not relevant for the land use comparison	unspecified	7
bare soil	None	None	1
fallow/marginal land	None	None	1
agroforestry	Input management	<ul style="list-style-type: none"> ■ no input ■ low inputs ■ medium inputs ■ high inputs 	4
agroforestry	Intercropping (spatial)	<ul style="list-style-type: none"> ■ with legumes ■ without legumes 	2
arable land	Soil physical disturbance	<ul style="list-style-type: none"> ■ no tillage ■ intermediate intensity tillage ■ high intensity tillage 	3
arable land	Water management Specific to rice	<ul style="list-style-type: none"> ■ Irrigated ■ non-irrigated ■ flooded ■ non/alternatively flooded 	4
arable land	Rotations (temporal)	<ul style="list-style-type: none"> ■ with legumes/cover crop ■ without legumes/cover crop 	2
arable land	Intercropping (spatial)	<ul style="list-style-type: none"> ■ intercropping ■ no intercropping 	2
arable land	Integrated management systems	<ul style="list-style-type: none"> ■ conventional ■ organic ■ integrated/agroecological 	3
arable land arable land, perennial crop grassland	Input management	<ul style="list-style-type: none"> ■ no input ■ low mineral inputs ■ medium mineral inputs ■ high mineral inputs ■ low organic inputs ■ medium organic inputs ■ high organic inputs ■ mineral and organic inputs 	24
forest	Natural stand Semi-natural stand	<ul style="list-style-type: none"> ■ primary undisturbed ■ secondary or regrowth 	2
forest, plantation	Intensity of intervention given production objectives	<ul style="list-style-type: none"> ■ low productivity ■ medium productivity ■ high productivity 	3
grassland	Intercropping (spatial)	<ul style="list-style-type: none"> ■ intercropping ■ with legumes ■ without legumes 	3
grassland	Grazing Intensity	<ul style="list-style-type: none"> ■ grazed ■ ungrazed 	2
Total number of archetypes			63

Table 1. Archetypes detailed by high-order land use classes and management sub-classes.

For these cases, other investigated factors were recorded as bias risks. In some cases, one treatment needed to be duplicated and attributed different archetypes in order to treat management factors separately. In such cases, records kept track of duplicates and ensured consistent land use or land management comparisons.

Overall, this database contains 10,068 comparisons of soil POXC content across various land uses and managements practices. Most studies focused on management changes within arable land ($n = 7,809$) (Fig. 5). Among these studies, change in input types and application rates were the most frequently studied practices ($n > 500$), followed by tillage intensity and cover crop (both $n > 200$) (Fig. 6). The most frequently studied land-use changes were from grassland to arable land ($n = 324$) and vice-versa ($n = 261$) (Fig. 5). There was a similar number of studies on deforestation to arable land ($n = 177$) and afforestation from arable land to forest ($n = 174$) (Fig. 5).

Finally, we paid a meticulous attention to soil types and properties, which are, intrinsically, closely related to soil health and soil carbon pools. We compiled all soil information provided by the authors and, based on the recorded coordinates, completed it with information from the Harmonized World Soils Database version 2.0⁶⁵ (HWSD2), World Soil Information referred to as ISRIC maps (soilgrids), and expert knowledge. ISRIC maps were downloaded from the webpage (<https://files.isric.org/soilgrids/latest/data/>). We focused on maps of clay content and pH (H₂O) at 0-5 cm and 5-15 cm soil depths to check and, where necessary, complete missing author-reported data. For each site, we then compiled the best estimates for the reference soil group (RSG) following the harmonised World Reference Base⁶⁶, the clay content (%) and the pH (depending on the method used by the authors, see Supplementary Table 2). We also proposed a grouping of these three variables based on their suitability or main constraints for biological activities to help with the data treatment (Supplementary Table 2).

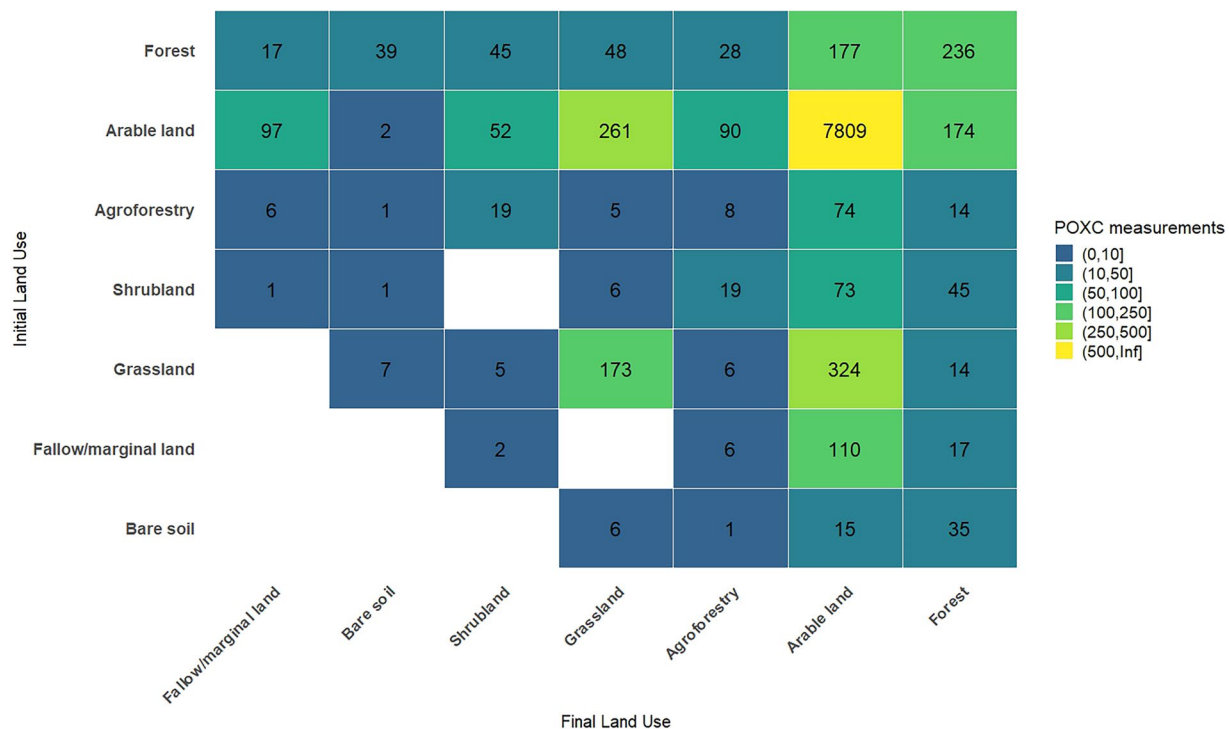


Fig. 5 Matrix of covered land use and land use changes in the POXC database V2.

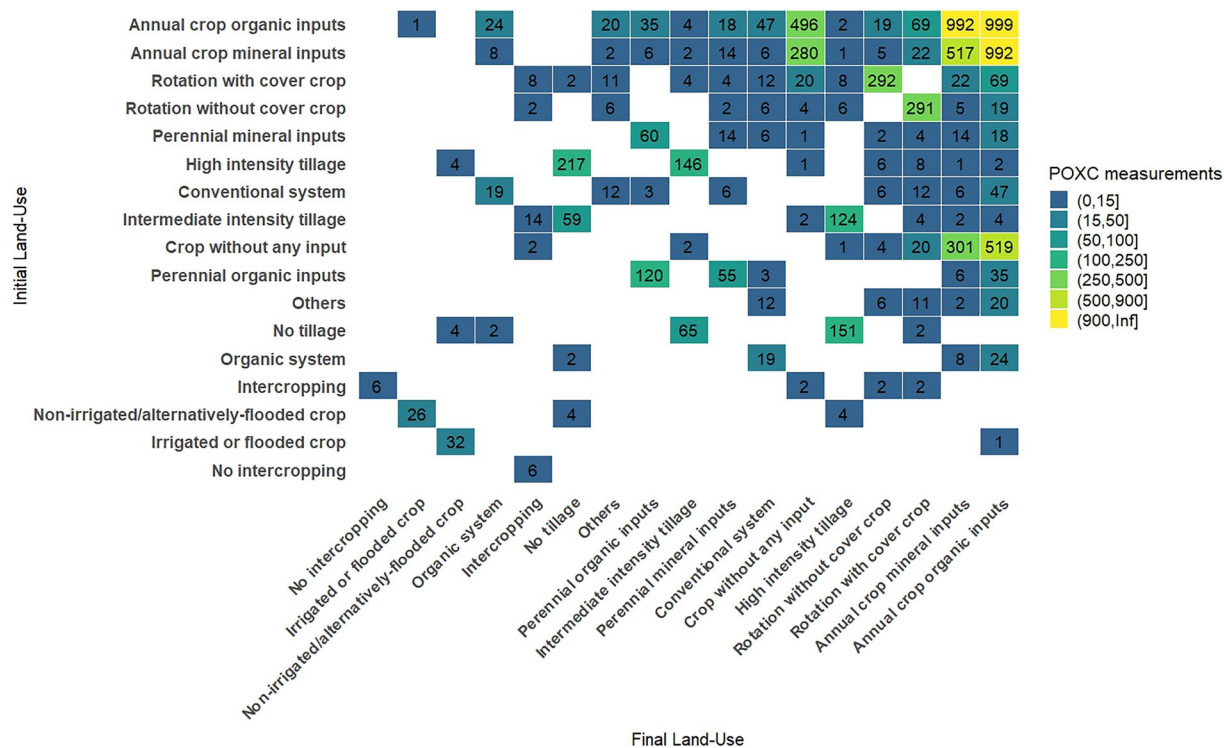


Fig. 6 Matrix of covered land management and management changes for arable land in the POXC database V2.

Despite all our efforts, there remain some uncertainties related to clay content and pH, which one should bear in mind when analysing these co-variables together with POXC measurements.

Data Records

This database⁵⁵ adheres to the FAIR principles, ensuring that all information is accessible, robust and transparent for reuse and integration into future analytical frameworks (Supplementary Table 3).

The files are grouped into two categories: i) The database file starting with the number 1, is the one file to be directly used to work with the data, as it contains all the useful variables for any analysis; ii) Files starting with a capital letter are the ones providing complementary information to delve into the individual studies and the database establishment.

All files are provided in .csv format except for the E_QUALITY_APPRAISAL_TOOL_POXC_DB.v1.0 Excel spreadsheet due to formatted decision trees as well as the supplementary material provided as a pdf. The “README” file outlines the study authors, context and lists all associated files as follows:

- 1_EFFECT.SIZES_POXC_DB.v1.1.csv: This file contains all quantitative comparisons between the treatments and the controls. Each row corresponds to a POXC measurement from a specific pedo-climatic location and soil depth; one article may yield multiple rows if it includes measurements from several locations or soil depths;
- A_STUDIES_POXC_DB.v1.1.csv: This file provides the complete list of the 699 studies identified in the February 2023 search, excluding duplicates (except for 3 discovered later), plus 24 additional studies identified in February 2024. The metadata include details for each study;
- B_REJECTED.STUDIES_POXC_DB.v1.1.csv: This file lists the 454 studies discarded due to one or more exclusion criteria. If a study ‘inclusion was uncertain, it was initially retained and further assessed;
- C_RETAINED.STUDIES_POXC_DB.v1.1.csv: This file lists all the retained experimental sites and provides detailed metadata on the POXC protocol, geographical coordinates, soil type, field trial characteristics, and the number of POXC measurements. From the 284 studies, a total of 372 experimental sites were identified, as some studies included multiple locations;
- D_QUALITY.APPRAISAL.TOOL_POXC_DB.v1.0.xlsx: This file details the study quality appraisal approach developed for the database;
- E_QUALITY.SCORES_POXC_DB.v1.0.csv: This file lists the detailed quality scores by study, based on meticulous quality review process described in the “Technical validation” section;
- F_ADDED.STUDIES_POXC_DB.v1.0.csv: This file includes all studies extracted from secondary sources (*e.g.*, some initial references were reviews listing other studies). In total, 53 studies were identified: 17 were already present in the list from the original search, 5 were inaccessible, and 15 were retained after applying the selection criteria. DOIs or URLs are provided where available;
- G_DATA.DICTIONARY_POXC_DB.v1.1.csv: This file defines all terms and archetypes used in the database.
- H_SUPPLEMENTARY.MAT_POXC_DB.v1.1.pdf: This file contains three tables with supplementary information to the datapaper on (1) the estimation of dispersion based on the free EX-TRACT tool⁶², (2) the grouping of some soil variables, and (3) the applied FAIR principles.

Technical Validation

The data extraction process was carefully reviewed to minimise errors and ensure the highest possible level of accuracy. Data quality assurance and control involved three complementary steps: 1) using plots to visually identify potential errors in both in POXC measurements and geographical coordinates; 2) reviewing extracted data from the retained studies to verify consistency through iterative discussions between the lead author and each co-author; and 3) conducting an overall data quality screening using a robust critical appraisal framework.

For the first step, following the initial data extraction, we used RStudio (2024.04.2) with R-4.3.0 to carefully inspect for extreme values, potential unit discrepancies and other inconsistencies. We also plotted the geographical coordinates to identify and correct potential errors.

For the second step, the authors discussed a total of 198 potential issues (*e.g.*, uncertainties regarding POXC measurements or replicates, or ambiguities in selecting the archetypes). These detailed discussions led to the exclusion of 20 studies —representing 522 POXC measurements— that had been initially retained. These discarded studies are included in the total number of rejections provided in Fig. 2. In a few cases ($n = 11$ studies), not all POXC measurements (*e.g.*, treatments or soil depths) were extracted, primarily for three reasons: i) some data could not be extracted from figures due to insufficient resolution or details; ii) some data were not relevant to land use or management analysis; or iii) some data were redundant with already extracted data.

Finally, for the third step, we implemented a robust critical appraisal strategy to identify potential sources of bias and enhance the reliability of the dataset. Following the Collaboration for Environmental Evidence (CEE) approach but adjusted to agronomy, we defined three key bias risk criteria, each supported by a decision tree. To ensure objectivity in the critical appraisal process, the lead author conducted a consistent review of all retained studies and collaborated with the primary reviewers whenever ambiguities arose. The key bias risk criteria were as following:

- Misclassification risks:** These risks refer to cases where land use classification was ambiguous—due to unclear treatments or land use definitions, land uses not matching the defined archetypes, imprecise control definitions, or potential confounding effects;
- Length/exposure risks:** These risks address situations where the experimental duration was unclear, where the control might not have matched the age of the treatments plot, or where exposure was likely insufficient to ensure the origin of observed effects;
- Outcome reporting risks:** These risks focus on uncertainty regarding POXC unit, soil depth, or data extraction from figures; or cases where dispersion metrics were not provided and had to be estimated.

Final quality scores were categorised as “Low”, “Medium”, and “High” risk of issues. The overall score was determined based on the most frequent risk level (*i.e.*, at least two out of three criteria). When each level

occurred once, the score was set to “Medium risk of issues”. The only exception was for cases where treatments and controls occurred on different soil types, indicating a significant risk based on expert knowledge. In these cases ($n = 5$ studies), the scores were set to “High risk of issues”. Both initial and final reviewers’ comments were recorded to document the main concerns and justify the decision. These quality scores are intended to help to weight studies in meta-analyses, and to adjust the mean overall calculated effect for varying degree of bias.

Usage Notes

To our knowledge, this is the most comprehensive database⁵⁵ of soil POXC measurements related to changes in land use and land management practices so far. It is designed as a valuable resource for both scientists and decision-makers to analyse the effectiveness and uncertainty associated with various land use strategies and land management practices aiming at preserving or improving soil health, with a particular focus on the soil carbon transformation function. In addition to POXC data, the database includes a wide range of complementary variables, such as pH, clay content, and crop yields where available, allowing for deeper analysis of POXC sensitivity as a soil health indicator under varying conditions.

Crucially, this database is designed to be dynamic and can be easily updated with new data as the number of studies on this topic continues to grow. We encourage researchers and practitioners to contribute relevant studies or data not yet included. Submissions can be sent directly to the corresponding author to support the ongoing expansion and utility of this resource.

Data availability

All data files are available in the CIRAD repository Version 2, and can be accessed at: <https://doi.org/10.18167/DVN1/78A2II>.

Code availability

R code was used to:

- Extract HWSO2 data by coordinates, compile and check soil types and initial properties.
- Figures and maps
- Data initial check

The code to upload data and draw maps and figures is made available on GitHub: https://github.com/bessou-web/POXC_Database

Received: 11 February 2025; Accepted: 11 September 2025;

Published online: 27 October 2025

References

1. Adhikari, K. & Hartemink, A. E. Linking soils to ecosystem services — A global review. *Geoderma* **262**, 101–111 (2016).
2. Dominati, E., Patterson, M. & Mackay, A. A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecol. Econ.* **69**, 1858–1868 (2010).
3. Paul, C., Kuhn, K., Steinhoff-Knopp, B., Weißhuhn, P. & Helming, K. Towards a standardization of soil-related ecosystem service assessments. *Eur. J. Soil Sci.* **72**, 1543–1558 (2021).
4. Marmo, L. The role of the Soil Thematic Strategy in promoting soil fertility. in *Fertilitas agrorum* vol. 2 (Roma, 2007).
5. Jiang, K. *et al.* Global land degradation hotspots based on multiple methods and indicators. *Ecol. Indic.* **158**, 111462 (2024).
6. FAO & ITPS. *Status of the World’s Soil Resources: Main Report*. (Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy, 2015).
7. Montanarella, L. *et al.* World’s soils are under threat. *SOIL* **2**, 79–82 (2016).
8. IPBES. *Summary for Policymakers of the Assessment Report on Land Degradation and Restoration of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. 48 (2018).
9. Smith, P. *et al.* How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Change Biol.* **26**, 219–241 (2020).
10. Nortcliff, S. Standardisation of soil quality attributes. *Agric. Ecosyst. Environ.* **88**, 161–168 (2002).
11. Moebius-Clune, B. N., *et al.* Developing standard protocols for soil quality monitoring and assessment. in *Innovations as Key to the Green Revolution in Africa* 833–842 (Springer, The Netherlands, 2011).
12. COM. *Communication from the Commission to the European Parliament and the Council: Sustainable Carbon Cycles. 800 final* (2021).
13. Bartkowski, B. *et al.* Potential of the economic valuation of soil-based ecosystem services to inform sustainable soil management and policy. *PeerJ* **8**, e8749 (2020).
14. Kopittke, P. M. *et al.* Ensuring planetary survival: the centrality of organic carbon in balancing the multifunctional nature of soils. *Crit. Rev. Environ. Sci. Technol.* **52**, 4308–4324 (2022).
15. Zhao, R., *et al.* Understanding trade-offs and synergies among soil functions to support decision-making for sustainable cultivated land use. *Front. Environ. Sci.* **10**, (2022).
16. Batjes, N. h. Total carbon and nitrogen in the soils of the world. *Eur. J. Soil Sci.* **47**, 151–163 (1996).
17. Lal, R. Soil carbon sequestration to mitigate climate change. *Geoderma* **123**, 1–22 (2004).
18. Kibblewhite, M. G., Ritz, K. & Swift, M. J. Soil health in agricultural systems. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 685–701 (2008).
19. Minasny, B. *et al.* Soil carbon 4 per mille. *Geoderma* **292**, 59–86 (2017).
20. Beillouin, D. *et al.* A global database of land management, land-use change and climate change effects on soil organic carbon. *Sci. Data* **9**, 228 (2022).
21. De Rosa, D. *et al.* Soil organic carbon stocks in European croplands and grasslands: How much have we lost in the past decade? *Glob. Change Biol.* **30**, e16992 (2024).
22. FAO & IIASA. *Harmonized World Soil Database Version 2.0*. <https://doi.org/10.4060/cc3823en> (FAO; International Institute for Applied Systems Analysis (IIASA), Rome and Laxenburg, 2023).
23. Brandao, M. & Canals, L. Global characterisation factors to assess land use impacts on biotic production. *Int. J. LIFE CYCLE Assess.* **18**, 1243–1252 (2013).
24. Cowie, A. L. *et al.* Land in balance: The scientific conceptual framework for Land Degradation Neutrality. *Environ. Sci. Policy* **79**, 25–35 (2018).

25. UNEP Life Cycle Initiative. *Global Guidance for Life Cycle Impact Assessment Indicators Volume 2*. (UNEP Life Cycle Initiative, 2019).
26. Oldfield, E. E., Bradford, M. A. & Wood, S. A. Global meta-analysis of the relationship between soil organic matter and crop yields. *SOIL* **5**, 15–32 (2019).
27. Bongiorno, G. Novel soil quality indicators for the evaluation of agricultural management practices: a biological perspective. *Frontiers of Agricultural Science and Engineering* **7**, 257–274 (2020).
28. Hurisso, T. T., Culman, S. W., Zone, P. & Sharma, S. Absolute values and precision of emerging soil health indicators as affected by soil sieve size. *Communications in Soil Science And Plant Analysis* **49**, 1934–1942 (2018).
29. Chenu, C. *et al.* Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil Tillage Res.* **188**, 41–52 (2019).
30. Haynes. Labile organic matter as central components of the quality of agricultural soils: an overview. *Adv. Agron.* **85**, (2005).
31. Weil, R. *et al.* Estimating active carbon for soil quality assessment: a simplified method for laboratory and field use. *Am J Alternative Agric* **18**, 3–17 (2003).
32. Chan, K. Y., Bowman, A. & Oates, A. Oxidizable organic carbon fractions and soil quality changes in an oxic paleustalf under different pasture leys. *Soil Sci.* **166**, 61–67 (2001).
33. Benbi, D. K., Brar, K., Toor, A. S. & Sharma, S. Sensitivity of Labile Soil Organic Carbon Pools to Long-Term Fertilizer, Straw and Manure Management in Rice-Wheat System. *Pedosphere* **25**, 534–545 (2015).
34. Brady, N. C. & Weil, R. R. *The Nature and Properties of Soils*. (Prentice Hall, Upper Saddle River, NJ, 2003).
35. Soil Organic Matter Fractions and Their Relevance to Soil Function. in *Soil Organic Matter in Sustainable Agriculture* (eds. Weil, R. R. & Magdoff, F.) (CRC Press, 2004).
36. Strosser, E. Methods for determination of labile soil organic matter: An overview. *J. Agrobiol.* **27**, 49–60 (2010).
37. Blair, G., Lefroy, R. & Lise, L. Soil carbon fractions based on their degree of oxidation, and the development of a carbon management index for agricultural systems. *Australian Journal of Agricultural Research* **46**, 1459–1466 (1995).
38. Culman, S. W. *et al.* Permanganate Oxidizable Carbon Reflects a Processed Soil Fraction that is Sensitive to Management. *Soil Sci. Soc. Am. J.* **76**, 494–504 (2012).
39. Tatzber, M. *et al.* KMnO₄ determination of active carbon for laboratory routines: three long-term field experiments in Austria. *Soil Research* **53**, 190–204 (2015).
40. Tirol-Padre, A. & Ladha, J. Assessing the reliability of permanganate-oxidizable carbon as an index of soil labile carbon. *Soil Science Society Of America Journal* **68**, 969–978 (2004).
41. Vieira, F. *et al.* Carbon management index based on physical fractionation of soil organic matter in an Acrisol under long-term no-till cropping systems. *SOIL TILLAGE Res.* **96**, 195–204 (2007).
42. Weil, R. R., Islam, K. R., Stine, M. A., Gruver, J. B. & Samson-Liebig, S. E. Estimating active carbon for soil quality assessment: A simplified method for laboratory and field use. *Am. J. Altern. Agric.* **18**, 3–17 (2003).
43. Awale, R., Emeson, M. A. & Machado, S. Soil Organic Carbon Pools as Early Indicators for Soil Organic Matter Stock Changes under Different Tillage Practices in Inland Pacific Northwest. *Frontiers In Ecology And Evolution* vol. **5** (2017).
44. Bongiorno, G. *et al.* Sensitivity of labile carbon fractions to tillage and organic matter management and their potential as comprehensive soil quality indicators across pedoclimatic conditions in Europe. *Ecological Indicators* **99**, 38–50 (2019).
45. Hurisso, T. T. *et al.* Comparison of Permanganate-Oxidizable Carbon and Mineralizable Carbon for Assessment of Organic Matter Stabilization and Mineralization. *Soil Sci. Soc. Am. J.* **80**, 1352–1364 (2016).
46. Zhao, H., Lv, Y., Wang, X., Zhang, H. & Yang, X. Tillage impacts on the fractions and compositions of soil organic carbon. *Geoderma* **189–190**, 397–403 (2012).
47. Liptzin, D. *et al.* An evaluation of carbon indicators of soil health in long-term agricultural experiments. *Soil biology & Biochemistry* vol. **172** (2022).
48. Thoumazeau, A. *et al.* Biofunctool[®]: a new framework to assess the impact of land management on soil quality. Part A: concept and validation of the set of indicators. *Ecol. Indic.* **97**, 100–110 (2019).
49. Woodings, F. S. & Margenot, A. J. Revisiting the permanganate oxidizable carbon (POXC) assay assumptions: POXC is lignin sensitive. *Agric. Environ. Lett.* **8**, e20108 (2023).
50. Gasch, C., Mathews, S., Deschene, A., Butcher, K. & DeSutter, T. Permanganate oxidizable carbon for soil health: Does drying temperature matter? *Agric. Environ. Lett.* **5**, e20019 (2020).
51. Gruver, J. Evaluating the Sensitivity and Linearity of a Permanganate-Oxidizable Carbon Method. *Communications in Soil Science and Plant Analysis* **46**, 490–510 (2015).
52. Wade, J. *et al.* To standardize by mass of soil or organic carbon? A comparison of permanganate oxidizable carbon (POXC) assay methods. *GEODERMA* vol. **404** (2021).
53. Christy, I., Moore, A., Myrold, D. & Kleber, M. A mechanistic inquiry into the applicability of permanganate oxidizable carbon as a soil health indicator. *Soil Sci. Soc. Am. J.* **87**, 1083–1095 (2023).
54. Hendricks, S. *et al.* Agricultural management affects active carbon and nitrogen mineralisation potential in soils. *Journal of Plant Nutrition and Soil Science* **185**, 513–528 (2022).
55. Chéron-Bessou, C., *et al.* A global database on land use and management change effects on soil KMnO₄-oxidisable organic carbon (POXC). <https://doi.org/10.18167/DVN1/78A2II>. *CIRAD Dataverse*, V2 (2025).
56. Haddaway, N., Macura, B., Whaley, P. & Pullin, A. ROSES for Systematic Review Protocols. Version 1.0. *figshare* <https://doi.org/10.6084/m9.figshare.5897269.v3> (2018).
57. Pullin, A. S. *et al.* Standards of conduct and reporting in evidence syntheses that could inform environmental policy and management decisions. *Environ. Evid.* **11**, 16 (2022).
58. Petrokofsky, G. Guidelines and Standards for Evidence Synthesis in Environmental Management: Version 5.0. <https://eprints.soton.ac.uk/420195/> (2018).
59. Jansen, B. J. & Spink, A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Inf. Process. Manag.* **42**, 248–263 (2006).
60. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
61. Wan, X., Wang, W., Liu, J. & Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.* **14**, 135 (2014).
62. Acutis, M. *et al.* An excel tool for the estimation of standard deviations from published articles. *Environ. Model. Softw.* **147**, 105236 (2022).
63. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis*. <https://doi.org/10.1002/9780470743386> (John Wiley & Sons, Ltd, 2009).
64. De Laurentiis, V. *et al.* Soil organic carbon as an indicator of land use impacts in life cycle assessment. *Int. J. Life Cycle Assess.* **29**, 1190–1208 (2024).
65. FAO & IIASA. *Harmonized World Soil Database Version 2.0*. (FAO; International Institute for Applied Systems Analysis (IIASA), 2023).
66. IUSS Working Group WRB. *World Reference Base for Soil Resources. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps*. (International Union of Soil Sciences (IUSS), Vienna, Austria, 2022).

Acknowledgements

This work is part of the project Soil quality Assessment in Agriculture For life cycle assessment-based Eco-design (SAAFE), funded by the European Commission under the Global Individual Marie Skłodowska-Curie Fellowship (Grant Number 843845) obtained by Cécile Chéron-Bessou. Complementary funds were provided by ANR-20-CE03-0004 PALMEXPAND. The authors would like to thank the editors and reviewers for their thorough work in reviewing the paper and allowing for a significant improvement of the information clarity.

Author contributions

Conceptualization ideas; formulation or evolution of overarching research goals and aims: C.C.-B. Formal analysis, application of statistical, mathematical, computational, or other formal techniques to analyse or synthesise study data C.C.-B., D.B., P.N.N. Investigation process, specifically performing the data/evidence collection: all the authors. Data curation management activities to annotate, scrub data and maintain research data for initial use and later reuse: C.C.-B., D.B. Writing original draft preparation, creation or presentation of the published work, specifically writing the initial draft: C.C.-B., D.B., P.N.N. Writing, review & editing preparation, creation or presentation of the published work by those from the original research group, specifically critical review, commentary or revision, including pre-or post- publication stages: all the authors. Visualisation preparation, creation or presentation of the published work, specifically visualisation/data presentation: C.C.-B., D.B. Supervision oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team: C.C.-B. Project administration management and coordination responsibility for the research activity planning and execution: C.C.-B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05976-9>.

Correspondence and requests for materials should be addressed to C.C.-B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025