SR4
A Systematic Map

EviEM Scientific Report 2015

Neal R. Haddaway
Katarina Hedlund
Louise E. Jackson
Thomas Kätterer
Emanuele Lugato
Ingrid K. Thomsen
Helene Bracht Jørgensen
Bo Söderström

What are the effects of agricultural management on soil organic carbon in boreo-temperate systems?

EviEM SR4 Scientific Report

Publication date: 11 December 2015

Cover photo: Jan Håkan Dahlström / Bildhuset / IT

This review follows the CEE *Guidelines for systematic reviews in environmental management*,
issued by the Collaboration for Environmental Evidence.

It has also been published in the open-access journal *Environmental Evidence*. Please cite as: Haddaway *et al.*:
What are the effects of agricultural management on soil organic carbon in boreo-temperate systems? A systematic map.
Environmental Evidence 2015 **4:**23

What are the effects of agricultural management on soil organic carbon in boreo-temperate systems?

A systematic map

Neal R Haddaway[1*] **(NRH)**
*Corresponding author
Email: neal.haddaway@eviem.se; neal_haddaway@hotmail.com

Katarina Hedlund[2]
Email: katarina.hedlund@biol.lu.se

Louise E Jackson[3]
Email: lejackson@ucdavis.edu

Thomas Kätterer[4]
Email: Thomas.Katterer@slu.se

Emanuele Lugato[5]
Email: emanuele.lugato@jrc.ec.europa.eu

Ingrid K Thomsen[6]
Email: ingrid.thomsen@agro.au.dk

Helene Bracht Jørgensen[2] **(HBJ)**
Email: helene.bracht_jorgensen@biol.lu.se

Bo Söderström[1]
Email: bo.soderstrom@kva.se

1 Mistra Council for Evidence-Based Environmental Management, Royal Swedish Academy of Sciences, P.O. Box 50005, SE-104 05 Stockholm, Sweden.

2 Department of Biology, Lund University, SE-223 62 Lund, Sweden.

3 Department of Land, Air and Water Resources, University of California Davis, One Shields Avenue, Davis, CA 95616, USA.

4 Department Ecology, SLU, P.O. Box 7044, SE-750 07 Uppsala, Sweden.

5 Joint Research Centre, Land Resource Management, Institute for Environment and Sustainability (IES), European Commission, Ispra, VA, Italy.

6 Department of Agroecology, Aarhus University, P.O. Box 50, DK-8830 Tjele, Denmark

Abstract

Background

Soils contain the largest stock of organic carbon (C) in terrestrial ecosystems and changes in soil C stocks may significantly affect atmospheric CO_2. A significant part of soil C is present in cultivated soils that occupy about 35% of the global land surface. Agricultural intensification has led to practices that may decrease soil organic carbon (SOC), and agricultural management has the potential to be a powerful tool for climate change mitigation and increased soil fertility through SOC sequestration. Here, we systematically map evidence relating to the impacts of agricultural management on SOC in arable systems of the warm temperate and snow climate zones (subset of temperate and continental climates: Köppen-Geiger Classification).

Methods

Seventeen academic citation databases, 3 search engines and 25 organisational websites were searched for literature (academic and grey) using search strings translated into a range of languages relevant to the included geographical scope of the topic. Stakeholders were also contacted with requests for evidence. Bibliographic checking of 127 relevant reviews was undertaken to check for missing articles. Screening for relevance against predefined inclusion criteria was undertaken at title, abstract and full text levels according to a published protocol. All relevant studies were coded in a meta-database describing the citation, study settings, methods and quantitative data available (without extraction of the study findings). A basic critical appraisal of included studies was also performed. A geographical information system (GIS) presenting the map database on a physical, online map was also produced.

Results

A total of 735 studies from 553 articles was included in the systematic map database. Studies investigated one or more of five broad categories of interventions: amendments (286 studies), crop rotations (238), fertilisers (307), tillage (306), and multiple interventions (55). Studies were identified from across the includible climate zones, with the notable underrepresentation from Russia. The majority of studies employed only point sampling of SOC, low levels of true spatial replication and moderate study periods (i.e. 10-20 years). Missing key methodological information was found in 28% of studies.

Conclusions

Long-term study sites identified in this map provide a useful addition to existing databases of long-term experiments (LTEs). The identification of knowledge gaps, such as studies from Russia, also identify a need for improved cataloguing or reporting of existing and on-going research. This systematic map database represents a useful resource for decision-makers wishing to identify knowledge gaps warranting further primary research, knowledge gluts warranting further secondary research, and deficiencies and best practice in research methodology. In addition to the systematic map database, we have also produced two further resources: i) a database of LTE sites investigating agricultural management and SOC, and ii) a database of reviews and meta-analyses. To our knowledge, this is the first systematic review or map that utilises a GIS for presentation of an evidence base, which we believe substantially increases the utility of the map outputs.

Keywords

Soil carbon, Carbon storage, Carbon sequestration, Conservation agriculture, Agricultural practices, Long-term, Amendments, Crop rotation, Fertilisation, Tillage

Background

Soils contain the largest stock of organic carbon (C) in terrestrial ecosystems (2500 Pg of C to 2 m depth); about double that stored in the atmosphere [1-3]. As a result, changes in soil C stocks may significantly affect the concentration of carbon dioxide (CO_2) in the atmosphere. A significant part of this carbon (12%) is present in agricultural soils [3] that occupy about 35% of the global land surface [4]. Agricultural soils are often depleted in SOC, which means that they have a potential to sequester soil organic carbon (SOC) [5]. Agricultural management has the potential to be a powerful tool for climate change mitigation through C sequestration in soil [6, 7]. Pressures from: international development of food and feedstock markets; increasing global population; and, changes towards more bio-based and fossil-free economies have led to the intensification of agriculture, which has been achieved by simplified crop rotations, increased soil tillage, and a lack of organic amendments all of which may decrease SOC [8-10]. The combined effects of actions mitigating climate change by increasing area of crops for bioenergy may concurrently deplete soil carbon stocks and intensify climate change.

Despite the different reasons for C loss, there are measures in addition to land-use change that can potentially slow down or reverse the trend for C depletion in cropland. Several measures have been proposed in the literature: i) crop rotations including, for example, leys [11] and cover crops [12], ii) organic amendments [13] or crop residues [14], iii) organic fertilisers such as farmyard manure [15] or inorganic fertilisers [16], and iv) tillage type and intensity [17]. Additional benefits of increasing C stocks in agricultural soils include increased soil fertility [11, 18] and improved physical and biological properties [19] through a reduction in bulk density, increased water-holding capacity, improved soil structure and enhanced microbial activity [20]. However, a change in the soil C stock may not imply a reciprocal change in the atmospheric C stock by the same amount, since the management employed to achieve increased SOC stocks may consume energy from non-renewable resources and cause changes in the atmospheric C stock (i.e. net carbon release) [3,8]. Furthermore, interventions that aim to increase soil C may result in increases in GHG emissions, which should be taken into account in a systems-scale approach (although this is not the focus of this review). The work herein focuses on the effects of agricultural cropland management on SOC and how the soil C can be increased or depleted by different types of management techniques used across temperate agricultural regions. The net effect of land-use change or soil management practices on GHG budgets should be considered in a broader context [21] and deserves a more complete life cycle assessment approach [22] than what is intended here.

The evidence of management on organic carbon in agricultural soils is extensive and links to soil fertility, erosion prevention, nutrient retention and above and below-ground biodiversity [23-25]. SOC responds slowly to changes in agricultural management [26], which means that these changes require many years to be detectable due to the large amounts SOC present in the soil profile compared to the much smaller proportion of organic C being sequestered or lost from the soil annually [27]. As a result long-term experiments are required to quantify the effect of management on SOC. Despite this restriction, a substantial number of studies have been performed and a number of traditional literature reviews have been published [7, 28-31]. For example, Gonzalez-Sanchez et al. [7] concluded from a meta-analysis of data from 29 publications from Spain that some forms of conservation agriculture (i.e., no tillage and implementing cover crops) can have positive effects on SOC. Govaerts et al. [30] reviewed three aspects of conservation agriculture: reduction in tillage intensity, retention of crop residues and use of crop rotations. The data (mainly from the Americas) indicated that the largest contribution of conservation agriculture to reducing emissions from farming activities from the reduction of tillage operations. Whilst a number of reviews have considered the relative impacts of different farming systems on SOC [e.g. 7, 28, 29], it is very difficult to accurately compare multiple systems that differ greatly in the management interventions involved. The purpose of the work herein is to synthesise evidence pertaining to individual management interventions under a wide variety of conditions in order to provide specific advice for land managers.

Although several meta-analyses and literature reviews have been published on the impacts of agricultural management on SOC, a systematic map listing and describing published studies of the efficacy of different management techniques to increase SOC stocks in agricultural areas has not yet been published. Such a systematically produced database of evidence would provide a range of uses. Systematic mapping is an emerging field in environmental management evidence synthesis [32], having been adapted to agricultural topics from the social sciences [33]. Systematic maps have previously been stated as a tool for the identification of knowledge gaps (areas lacking published research that may be suitable for primary research) and knowledge gluts (areas with sufficient published evidence to allow secondary synthesis via systematic review) [34-36]. We have identified several major additional benefits of our systematic map on the subject of agricultural management impacts on SOC as follows. We believe the systematic map database can be used for:

- The calibration and validation of models used for simulating the effect of agricultural management on soil carbon stocks
- Validating existing maps of soil carbon and monitoring (e.g. European soil carbon map, http://eusoils.jrc.ec.europa.eu/esdb_archive/octop/octop_data.html)
- Including soil carbon when designing actions on climate mitigation
- Designing agricultural policies based on existing evidence from a specific subset of the evidence base
- Making recommendations for research design, both in terms of maximising the usability of statistics and improving or standardising experimental design
- Analysis of keywords used across the evidence base to increase the findability of future research of relevance to the topic
- Contacting researchers that have worked on specific long-term study sites with requests for data or further information. This is also an important resource if the user is interested in outcomes other than SOC
- Identifying a comprehensive list of research published on a subset of the evidence included, for example a specific country, soil type or climate zone
- Finding relevant citations and links to their full texts
- Obtaining summary and raw data where the systematic map is linked to future systematic reviews

Objective of the map

This systematic map is intended to provide a catalogue of academic and grey literature on the impacts of cropland management interventions on SOC across temperate regions. Included studies are described within a searchable database with full details of study setting and experimental design. We have also undertaken critical appraisal of study susceptibility to bias (internal validity) and relevance to the topic in hand (external validity). Furthermore, we have also produced several additional outputs that aim to maximise utility and interactivity (using a web-based geographical information system (GIS)), and ensure legacy, updatability and ongoing relevance.

This topic was originally proposed as a systematic review [37], but once searches were underway it was recognised that a systematic map would better suit the scope of the question and needs of the stakeholders, since the subject of interest was broad and knowledge of the state of evidence across the subject was limited. It is hoped that the map will be easily updated as new research becomes available, and it is recommended that the database be revised at regular intervals to ensure it is kept up-to-date.

Identification of topic and identified stakeholders

The topic was suggested by Karin Hjerpe (Swedish Board of Agriculture; May 4, 2012 and September 20, 2012) and Olof Johansson (Swedish Board of Agriculture; September 24, 2012). The following stakeholder groups were identified as having a potential interest in the findings of the review:

- The Swedish Board of Agriculture is responsible for the national environmental quality objective “A varied agricultural landscape”. One expected outcome within this goal is that arable land will have a well-balanced nutrient status, good soil structure and organic matter content. Another expected outcome is that the land will be cultivated in such a way as to sustain the long-term productivity of the soil. These outcomes are closely related to SOC. The Swedish Board of Agriculture also addresses issues relating to climate change.

- The Swedish Environmental Protection Agency (EPA) is responsible for the environmental quality objective “Reduced Climate Impact”. In this context, the Swedish Parliament has adopted a vision of zero net emissions of greenhouse gases to the atmosphere in Sweden by 2050.

- The Federation of Swedish Farmers (LRF) is interested in both the environmental issues and the productivity aspect. In their Climate Policy it is stated that increased soil organic matter (SOM) content in cropland potentially can reduce concentrations of GHGs in the atmosphere and that such opportunities should be seized. The Federation of Swedish Farmers is also taking part in Focus on Nutrients (“Greppa Näringen” in Swedish), which is a joint venture between LRF, the Swedish Board of Agriculture, the County Administrative Boards and a number of companies in the farming sector. Focus on Nutrients offers advice to farmers, for example on climatic issues and SOC management.

During a stakeholder meeting at the EviEM secretariat (June 4, 2013), representatives from the Swedish Board of Agriculture, Swedish Environmental Protection Agency, Federation of Swedish Farmers, and Swedish University of Agricultural Sciences discussed the formulation of the review question and exclusion/inclusion criteria. It was suggested that the focus should be on long-term studies of how agricultural management affect SOC stocks within the temperate climate zone (humid and summer dry) as well as the snow climate zone (northern Sweden). The stakeholders suggested that cereal grains such as wheat and barley were of particular interest, but also other crops that could become more important in Sweden in a changing climate (such as maize). All agricultural management types and soil types within these agricultural regions were of interest. Greenhouse gases other than CO_2, such as methane (CH_4) and nitrous oxide (N_2O), and studies solely focusing on soil phosphorus and nitrogen were considered to be outside the systematic review's scope. There is a lack of data on CH_4 since it is infrequently measured in upland soils. Similarly, there is little data on long-term changes in N_2O in which contrasting treatments have reached a new equilibrium. It is therefore difficult to integrate short-term N_2O processes with long-term trends in SOC changes. Stakeholders also emphasised that although the review question by definition must be fairly narrow, the narrative synthesis should have a broad contextual scope. For example, SOC may increase under bioenergy crops, but if the total cropped area remains constant, less food may be produced (i.e. a blanket switch to bioenergy crops would be a poor policy recommendation). Certain interventions may also require increased use of non-renewable energy leading to a reduced net effect on carbon emissions.

In addition to the meeting described above, stakeholders were invited to provide comments and suggestions on a draft protocol and a draft of this map report prior to submission for publication.

Primary question: *What are the effects of agricultural management on SOC stocks?*

Components of the primary question:

Population: Arable soils in agricultural regions from boreo-temperate systems, more specifically defined as a subset of the temperate/mesothermal and the continental/microthermal climate zones (according to the Köppen-Geiger climate classification; see *Relevant subjects* text below).

Within these climate zones, agricultural management systems in which wheat, barley, rye, oats, maize or oilseed rape can grow in the crop rotation were selected.

Intervention: A range of soil management practices relating to tillage, addition of crop residues, mineral fertiliser, manure or other organic "wastes", and different crop rotation schemes.

Comparator: Alternative intervention or no intervention.

Outcome: SOC stocks/concentration, quantifiable as a change relative to the spatial or temporal comparator.

Methods

Development of the review question

This systematic map was undertaken according to a protocol published in Environmental Evidence [37]. Initially a systematic review was proposed in the published protocol, but after searches were performed and the volume of evidence was revealed to be substantial, it was proposed that a systematic map be produced in the first instance, since a detailed catalogue of this large evidence base was perceived to of value to stakeholders. Thus, our methods described below follow the original systematic review protocol closely until the point of critical appraisal and synthesis. This systematic map includes a basic critical appraisal coding but does not extract study findings (i.e. quantitative data) and does not attempt any form of quantitative synthesis. All other activities proceeded according to the published protocol as detailed below.

Searches for literature

Search terms

Search terms were developed based on population, exposure and outcome question elements as follows (* indicates a wildcard):

Population terms: soil*, arable, agricult*, farm*, crop*, cultivat*

Exposure/intervention terms: till*, direct drill*, fertili*, bio*solid*, organic, manur*, sewage, compost*, amendment*, biochar*, digestate*, crop residue*, crop straw*, mulch*, crop rotat*, break crop*, grass, clover ley*, legume*, bioenergy crop*, cover crop*, grass clover, crop* system*, winter crop*, spring crop*, summer fallow*, catch-crop*, intercrop*, conservation

Outcome terms: soil organic carbon, soil carbon, soil C, soil organic C, SOC, carbon pool, carbon stock, carbon storage, soil organic matter, SOM, carbon sequestrat*, C sequestrat*"

Search terms were tested and a final search string produced. Details of the development of the search string can be found in Additional File 1. These terms were adapted as necessary for the different resources detailed below. The Web of Science equivalent Boolean search string was therefore:

> soil* AND (arable OR agricult* OR farm* OR crop* OR cultivat*) AND (till* OR "direct drill*" OR fertili* OR bio*solid* OR organic OR manur* OR sewage OR compost* OR amendment* OR biochar* OR digestate* OR crop residue* OR crop straw* OR mulch* OR crop rotat* OR break crop* OR (grass OR clover) ley* OR legume* OR bioenergy crop* OR cover crop* OR "grass clover" OR "crop* system*" OR winter crop* OR spring crop* OR summer fallow* OR "catch-crop*" OR intercrop* OR conservation) AND ("soil organic carbon" OR "soil carbon" OR "soil C" OR "soil organic C" OR SOC OR "carbon pool" OR "carbon stock" OR "carbon storage" OR "soil organic matter" OR SOM OR "carbon sequestrat*" OR "C sequestrat*")

Academic databases

Searches of academic databases were performed between the 16th and 19th September 2013. Additional File 2 details the search strings used in each of the 17 academic databases along with any optional restrictions or limitations employed and the numbers of results obtained.

Search engines

A search was undertaken in Google and DogPile in English on 17th March 2014 using the following search string; (carbon AND sequestration AND soil AND agriculture). The first 100 results from each search engine were screened for relevance. A search using the search engine Scirus was not performed as stated in the protocol, since the facility was retired in January 2014.

Searches in Google Scholar were performed in English, German, Italian, French, Danish and Swedish as described in Additional File 3.

Specialist websites

A total of 25 specialist websites were searched for grey literature both manually (navigating to publication pages and searching by eye for relevant studies) and using automated search facilities within each website. Details of the search terms and methods employed for each website are provided in Additional File 4:
Aarhus University Department of Agroecology, 24 March 2014
African Network for Soil Biology and Fertility, 24 March 2014
Columbia Basin Agricultural Research Center, 24 March 2014
European Environment Agency, 24 March 2014
European Soil Portal, 24 March 2014
Eusomnet, 19 May 2014
GCTE SOMNET, 19 May 2014
GRACEnet, USDA Agricultural Research Service, 19 May 2014
Indian Agricultural Statistics Research Institute, 24 March 2014
National Soil Carbon Network (NSCN) of the US Forest Service, 24 March 2014
Rapid Assessment of US Soil Carbon (RaCA), 19 May 2014
Rothamsted Research, 25 March 2014
Soil Carbon Center at Kansas State University, 25 March 2014
Soilservice, 25 March 2014
Swedish Board of Agriculture, 25 March 2014
Swedish Environmental Protection Agency, 25 March 2014
Swedish University of Agricultural Sciences, 25 March 2014
UC Davis, Agricultural Sustainability Institute, 25 March 2014
University of Copenhagen, 25 March 2014
University of Illinois, Department of Crop Sciences, 25 March 2014
USDA Agricultural Research Service, 25 March 2014
Victorian Long Term Agro-ecological Experiments, 25 March 2014
Videncentret for Landbrug, 25 March 2014
Working Group for Long-term Experiments (LTE), 24 March 2014
World Bank, 24 March 2014

Bibliographies

Following assessment of full texts a total of 127 reviews and meta-analyses (123 from full text screening and a further 4 from secondary sources) were identified as relevant to the subject. A database describing these reviews and meta-analyses is included as an additional resource from the systematic map (see Additional File 5). The bibliographies of all of these reviews were assessed for additional relevant publications at title level, with potentially relevant citations screened at abstract and then full text levels. Any relevant articles from the reviews were checked against the final list of included studies and studies not previously identified were added to the systematic map database. Studies excluded at full text were added to the list of excluded studies described with the screening process below.

Comprehensiveness assessment

Members of the review team selected 83 key references based on their own knowledge of the subject (e.g. often cited in their own publications) that were checked against search results to assess comprehensiveness of the search strategy. Only four of these articles had not been included in the previous searches of the academic databases listed above. These articles were included in screening at full-text.

Comprehensiveness of searches was also assessed using seven reviews suggested by the experts in the review team that contained a high proportion of relevant references within their bibliographies. These bibliographies were checked against the combined search results and any missing articles noted. A total of 74 articles relevant at title level (excluding duplicates) were found and, among these, 10 had not previously been included in the searches. These were added to the database.

Screening

Screening process

Screening was undertaken in a three-stage process: at title-, abstract- and full-text- level. Those articles included at one stage proceeded to the next. After abstract-level screening, articles were obtained in full text (see Article Retrieval, below). Where no abstract was available articles proceeded automatically to the next stage of assessment.

Screening of titles of search results was carried out by one reviewer, with a subset of 200 articles (1 % of search results) screened independently by a second reviewer (formed from a random subset of 100 articles from the entire results and a random selection of 100 articles from the journal *Agriculture, Ecosystems and Environment*) on 1st October 2013. Consistency between reviewers was checked with Kappa testing [38], which revealed moderate agreement (kappa=0.51). A second Kappa test of 333 articles from the journal *Applied Soil Ecology* was performed on 4th October 2013, also showing moderate agreement (0.50) between the first reviewer and a third reviewer. These two journals were chosen because they were believed to contain a relatively high proportion of relevant studies reported to a similar degree of detail in abstracts, facilitating Kappa testing. All disagreements between reviewers were discussed, understandings improved and inclusion criteria clarified. Abstract-level screening was performed by two reviewers. An initial validation screening showed a moderate agreement between reviewers (0.49) on a subset of 115 abstracts (2 % of screened abstracts). Disagreements were again discussed and clarified. Following discussion, further Kappa tests on subsets of 109 and 89 abstracts showed high agreement (0.67 and 0.82 respectively). Kappa testing of screening was also performed at full-text stage, showing high level of agreement (0.72) between two reviewers on a subset of 120 articles (7 % of screened full texts).

Uncertainties were discussed within the review group and doubtful cases were included, proceeding to the next stage of assessment. Numbers of included and excluded articles at all stages of the screening process were recorded and reasons for exclusion were documented for all articles assessed at full-text.

Article retrieval

Articles believed to be potentially relevant after abstract screening were sought in full text using subscriptions using a number of institutional access credentials, including: Stockholm University, Lund University, and Bangor University. A total of 208 articles (11 % of relevant abstracts) could not be obtained in full text either digitally or in print (see Additional File 6).

Inclusion criteria

Articles were screened and included/excluded according to the following criteria:

Relevant subject(s)

Arable (cropland) soils in boreo-temperate systems lying within the following the Köppen-Geiger climate classification zones [39] (Figure 1): agricultural regions from the warm temperate climate zone (fully humid and summer dry, i.e., Cfa, Cfb, Cfc, Csa, Csb, Csc); the snow climate zone (fully humid, i.e., Dfa, Dfb, Dfc). Permanent grasslands, paddy rice systems, agroforestry systems and orchards were excluded.

Relevant types of study design

Investigated interventions must have been in place for 10 years or more, since detectable changes in SOC typically do not occur over shorter time periods [40].

Relevant intervention(s)

Any type of agricultural cropland management falling into the following broad categories: amendments, crop rotation, fertiliser and tillage. Study data was included irrespective of the focus of the article (e.g. C sequestration to counteract climate change or management intended to increase soil fertility). Multiple interventions without information on specific interventions made in the same crop field preclude the opportunity to assess the effect of each intervention separately and were classified as ‘multiple’ where described in detail. For example, comparisons of organic and conventional farming may not always clearly state the management differences. Where sufficient details on multiple interventions were lacking, for example in comparisons of farming ‘systems’, these studies were excluded.

Relevant comparator(s)

Two types of comparator were included; spatial and temporal. Spatial comparators are control treatments or areas where the intensity or type of intervention differed (comparator-intervention; CI).

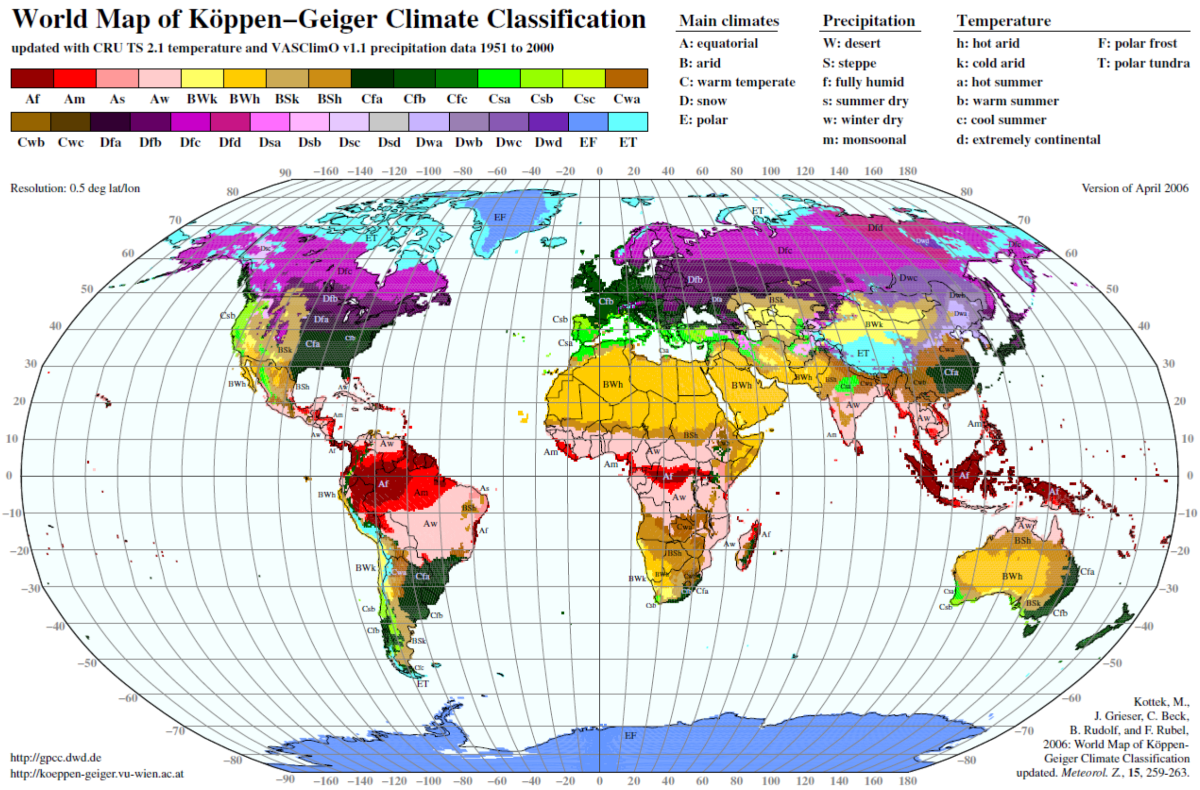


Figure 1. Köppen-Geiger climate classification zones included within this systematic map. Zones included are the warm temperate climate zone (fully humid and summer dry, i.e., Cfa, Cfb, Cfc, Csa, Csb, Csc): the snow climate zone (fully humid, i.e., Dfa, Dfb, Dfc).

Temporal comparators are present where data have been recorded for one intervention at multiple time points (either before-after; BA or time series). Individual studies could possess both spatial and temporal comparators as part of the same design (before-after-comparator-intervention; BACI).

Relevant outcome(s)

Measures of SOC as concentration (e.g. g/kg or %) or as stock (e.g. g/m^2). The measure may be reported as Soil Organic Carbon (SOC), Total Organic Carbon (TOC), Total Carbon (TC) or Soil Organic Matter (SOM).

Effect modifiers/sources of heterogeneity

The following effect modifiers were identified *a priori* and iteratively during screening and entered into the systematic map database:

Type of crop (i.e. annual or perennial)
Soil type, soil texture class (description or quantified percentages of silt, clay and sand)
Latitude and longitude
Köppen-Geiger climate zone
Intervention/study duration
Soil depth sampled

Coding and data extraction

Meta-data (descriptive categorical information regarding citations, study setting, design and methods) was extracted from included studies following full text assessment. Information was extracted according to the framework described in Table 1.

Table 1. List and description of meta-data variables extracted during coding.

Meta-data	Description
Title	Source article title
Authors	Source article authors
Journal	Source article journal
Year	Source article publication year
Volume	Source publication volume
Page Start-Page End	Source publication pages (start page to end page)
Author E-Mail	Source article corresponding author contact email address
Study Country	Country in which study was undertaken
Location Site/Name	Name or location of study site
Latitude	Latitude cited within article (converted to decimal degrees)
Longitude	Longitude cited within article (converted to decimal degrees)
Köppen Climate Zone	Stated climate zone (according to [39]) or established from latitude and longitude using Köppen Climate Zone GIS layer
Reference to Previous Articles	Stated reference to previously published work regarding experimental design: likely source of missing information
Study ID	Reviewer-assigned study identification code
Soil Texture, % Clay	Stated soil clay content (per cent) or stated soil type
Soil Texture, % Silt	Stated soil silt content (per cent)
Soil Texture, % Sand	Stated soil sand content (per cent)
Intervention, Start	Intervention start date (year)
Intervention, End	Intervention end date (year) or soil sampling date (year)
Duration of Intervention, Years	Time period intervention was in place before sampling (years)
Intervention	Intervention category (see Coding Framework for details of included sub-categories) (Crop Rotation / Fertiliser / Amendments / Catch Crop / Tillage / Multiple / Farming System)
Number of Treatments	Total number of treatment and control groups
Which Treatments	Stated treatment groups (either by intervention group when factorial designs or stated individual intervention combinations)
Crop	Crop type (annual, perennial, single crop, double crop)
Study Design	Design of the comparison within the study (Before-After / Comparator-Intervention / Before-After-Comparator-Intervention)
Comparator Type	Stated comparator
Experimental Design	Stated design of the experiment (Randomised Complete Block / Split-Plot / Time Series / Etc.)
Level of Replication	Level at which true replication was undertaken
Dimension of Plots	Plot (subplot) dimensions (metres or hectares)
Number of Spatial Replicates	Number of true spatial replicates per treatment group
Temporal Replicates	Number of temporal replicates (repeated measures)
Sampling Precision	Number and type of pseudoreplicates (within true replicate sampling)
Soil Sampling Depth	Soil sampling depth categories for which data are provided independently
C Measurement Method	Carbon quantification method
Outcome	Type of outcome measured (SOC / TOC / SOM / TC /OC)
Unit	Units of measurement
Data Location	Location of relevant outcome data within the article
Bulk Density	Presence of bulk density measurements (Measured / Calculated / Not Present)

Variance Presented	Presence and location of variance relating to true replication (standard deviation, standard error or confidence interval). Description of 'next best variance' if no true replicate variance reported.

Critical appraisal

All included studies were critically appraised for internal validity (susceptibility to bias) using a predefined framework. This framework is loosely based upon the Cochrane Critical Appraisal Tool (http://bmg.cochrane.org/sites/bmg.cochrane.org/files/uploads/TTT%20June%202010.pdf) and aims to; i) exclude studies with unacceptable risk of bias, ii) assess the risk of bias across a range of variables for each remaining study, and iii) assign each study with a critical appraisal category based on these variables. The framework calls for studies to be coarsely categorised for each variable (high [0], medium [1], low [2], or unclear [?] susceptibility to bias), which is supplemented with a brief textual justification of the appraisal.

Studies included through critical appraisal were required to meet the following characteristics:

- For comparator-intervention studies; true replication (i.e. 1 replicate studies excluded)
- For time series studies; more than two temporal replicates (3 or more samples) over time after intervention AND measurement period should exceed 30 years
- Comparator cannot be forest/woodlot or native grassland (i.e. permanent, ungrazed, unmanaged, or non-intervention)

Where insufficient information is presented to enable a full assessment a '?' is assigned. The five variables assessed are detailed in Table 2.

Table 2. Variables (domains) assessed during critical appraisal of primary study validity. Details of sources assigned to different values are also given.

Variable	Value	Score
Spatial (true) replication	2 replicates	0
	3-4 replicates	1
	>4 replicates	2
Temporal replication	<=3 replicates	0
	4-6 replicates	1
	>6 replicates	2
Treatment allocation (as described for the full experimental design)	Purposive (selective)	0
	Split-/strip-plot / Latin square / blocked / randomised / exhaustive	2
Duration of experiment	10-19 years	0
	20-29 years	1
	≥30 years	2
Soil sampling depth	Shallow (<15 cm) single or multiple sampling	0
	Plough layer (0-25 cm) single or multiple sampling, or deep (>25 cm) single sampling	1
	Multiple deep sampling (>25 cm)	2

Critical appraisal was carried out on meta-data by NRH and HBJ, with a subset of 10% of studies being appraised by both prior to assessment of the remaining studies by one reviewer. All disagreements were discussed and a common approach decided before continuing with appraisal.

All studies included in the systematic map were given a specific code for each of the individual intervention comparisons investigated. This coding can be used to filter similar studies that could be synthesised together. Interventions were identified iteratively within included studies, and once all studies had been assessed, a coding framework was developed to describe the intervention comparisons investigated. The framework describes four individual intervention groups (amendments, crop rotation, fertiliser and tillage) and one multiple comparison group (where described interventions were combined and their effects could not be separated). According to this framework, each study is assigned one or more codes to indicate the comparisons within it. Codes indicate a specific comparison between interventions or an intervention relative to a control treatment. Where information is lacking a higher code within the coding tree can be assigned; e.g. 1.1.4 where details on the type of fallow are not known, precluding assignment of 1.1.4.1 or 1.1.4.2.

1. **Crop rotations**
    1.1. Monoculture vs:
        1.1.1. Complex rotation (4 yrs)
        1.1.2. Rotations involving perennials
        1.1.3. 2 or 3 yrs rotations
        1.1.4. Rotations involving 1 yr "perennials"
            1.1.4.1. Sown fallow or grass
            1.1.4.2. Chemical- or bare- fallow
    1.2. Comparison of different crops
        1.2.1. Legumes +/- or legumes vs legumes
        1.2.2. Grass/ hay/ green fallow +/- or vs. each other
        1.2.3. Only crops
    1.3. Annual crop vs perennial crop (2 yrs or more)
    1.4. Multi cropping
        1.4.1. Under-sown/intercropping (legumes or other)
        1.4.2. Double cropping
        1.4.3. Cover- or catch crops
    1.5. Energy crops
        1.5.1. Energy crops vs. energy crops
        1.5.2. Energy crops vs. other crops
    1.6. Simple rotation (2 yrs) vs. complex rotations (3 yrs or more)

(NB: when grown for more than 1 yr in a row the following will be defined as perennials: hay, grass, green fallow, unmanaged fallow)

2. **Amendments** (with vs without)
    2.1. Crop residue
        2.1.1. Composted
        2.1.2. Uncomposted
            2.1.2.1. Green (mulch)
            2.1.2.2. Yellow
            2.1.2.3. Burned
    2.2. Manure
        2.2.1. Liquid (slurry)
        2.2.2. Solid

        - 2.2.2.1. Cattle
        - 2.2.2.2. Poultry
        - 2.2.2.3. Composted
  - 2.3. Lime
  - 2.4. Sewage sludge
  - 2.5. Green manure
  - 2.6. Peat/sediment
  - 2.7. Domestic compost/waste
  - 2.8. Processed wood (sawdust/woodchips)
  - 2.9. Bone meal/animal products
  - 2.10. Different application methods

3. **Nitrogen fertiliser** (with vs without)
   - 3.1. Inorganic vs organic/other amendment (organic also given an amendment code to describe the type of organic used)
   - 3.2. Different amounts of N
   - 3.3. Different application methods

4. **Tillage** (one versus another)
   - 4.1. No tillage (zero) vs:
     - 4.1.1. Reduced tillage (minimum/conservation/disc/chisel/harrow/mulch/ridge)
     - 4.1.2. Conventional (traditional) tillage (mouldboard)
     - 4.1.3. Rotational tillage (occasional)
     - 4.1.4. Subsoiling (>30 cm depth conventional tillage)
   - 4.2. Reduced tillage (minimum/conservation/disc/chisel/harrow/mulch/ridge) vs:
     - 4.2.1. Conventional (traditional) tillage (mouldboard)
     - 4.2.2. Rotational tillage (occasional)
     - 4.2.3. Subsoiling (>30 cm depth conventional tillage)
   - 4.3. Conventional (traditional) tillage (mouldboard) vs:
     - 4.3.1. Rotational tillage (occasional)
     - 4.3.2. Subsoiling (>30 cm depth conventional tillage)
   - 4.4. Rotational tillage (occasional) vs:
     - 4.4.1. Subsoiling (>30 cm depth conventional tillage)

5. **Multiple interventions** (just coded to pairs of confounded interventions)
   - 5.1. Crop rotation and:
     - 5.1.1. Amendment
     - 5.1.2. Fertiliser
     - 5.1.3. Tillage
   - 5.2. Amendment and:
     - 5.2.1. Crop rotation
     - 5.2.2. Fertiliser
     - 5.2.3. Tillage
   - 5.3. Fertiliser and:
     - 5.3.1. Tillage
   - 5.4. Multiple confounded

Studies were coded by the review team subject experts following an in-depth discussion and trial of the framework. Coding was checked by NRH and HBJ.

Mapping

Systematic map database

Identified studies that met the inclusion criteria stated above were described in detail in a database that documented citation information, study setting, design and methods. This database formed the systematic map provided here (see Additional File 7).

Overlapping studies

Many articles reported results of experiments at the same study site. Some of these articles reported results that overlapped in time or space with other studies, and a smaller number seemingly reported the same complete results in multiple formats – either in the same way or subject to further statistical analysis.

A study is a complex concept to define. Previous systematic reviews have defined a study as a particular experimental method used in a specific place at a specific time [e.g. 41]. Studies in our systematic map challenged this definition, since we used a 10-year minimum study period in our inclusion criteria, and the risk of overlapping, yet distinct studies is therefore increased. Our efforts to differentiate between studies are further hampered by the low level of methodological detail typical in primary studies, particularly with respect to the precise location of study sites. Linked to this are problems caused by different measurement methods for the same experimental study unit – measurement methods cannot therefore be used to differentiate independent studies. Figure 2 displays this issue visually.

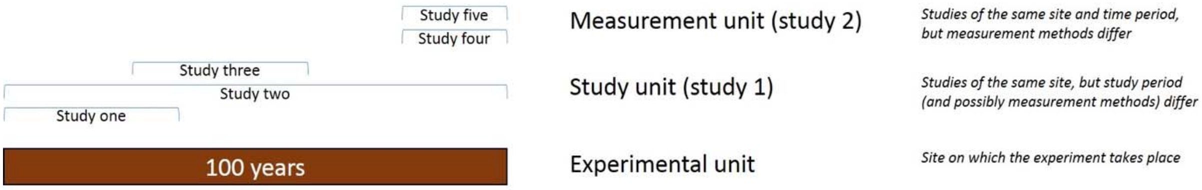


Figure 2. The complex relationship between experimental unit, and study units (study and measurement)

In our systematic map, where studies were clearly published more than once (i.e. the same data reported in multiple articles), the most comprehensive study was retained in the database whilst others were excluded (classified as 'superseded by another study'). In this instance, meta-data (descriptive information, not study findings) were aggregated from all reports of the same experiment to ensure no loss of information from excluding some articles. In all other cases each study has been retained in the systematic map for clarity. Users should therefore exercise caution to avoid double counting if selecting studies for full synthesis.

Results

The process

Number of articles at various stages

A total of 24,547 records were identified through searching academic citation databases (after removal of 17,328 duplicates). These records were screened at title level (5,735 included) and abstract level (1,814 included), and with the addition of 88 articles from grey literature searches 1,902 articles were identified as potentially relevant and needing full text assessment. Some 11 % of these articles (208) could not be retrieved, and a total of 1,694 articles were subsequently screened at full text level. After exclusion of a further 842 non-relevant articles, a total of 740 primary research articles and 127 reviews were then subjected to meta-data extraction, where information regarding study setting and experimental design were extracted and used to populate the systematic map database. During this stage a further 258 studies were excluded for a variety of reasons. The references of the 127 reviews (listed in Additional File 5) were screened in full at title, abstract and full text stage and any relevant studies were added to the database. Further reviews identified whilst reading full texts were identified iteratively and their references were also screened in the same way. All articles excluded following full text assessment (including full text screening, meta-data extraction and bibliographic checking) are listed in Additional File 8 along with reasons for exclusion. These processes are described visually in Figure 3.

Studies versus articles

The final number of included lines of data in the systematic map database was 735, corresponding to the number of individual studies identified. These studies were sourced from a total of 553 published articles. For the purposes of this systematic map, here we define a *study* as a discrete experiment (i.e. reported measurement) undertaken over a specific time period. Where consistent methodology was used across a series of sites these were classified as a single study.

Since many systems studied in agronomy are established long-term experiments (LTEs), there is a disparity between interventions (i.e. the physical system being modified) and studies (i.e. the period of observation and measurement). Typically a study is defined as a discrete place, period and mode of experimentation or observation [e.g. 42], however, LTEs are often measured and reported on multiple times. As a result, independent data points in any subsequent meta-analysis must take the form of individual interventions at specific study sites. In a systematic map, however, the subject of interest is the evidence base rather than the results contained therein. Identification of independent intervention units (i.e. experimental manipulations) for synthesis is complicated by the lack of reporting detail (discussed below), since study sites are rarely described well enough to identify the precise study location. We therefore include all published study reports. Furthermore, our novel approach to visualising our systematic map database through use of a GIS helps users to identify single sites with multiple studies.

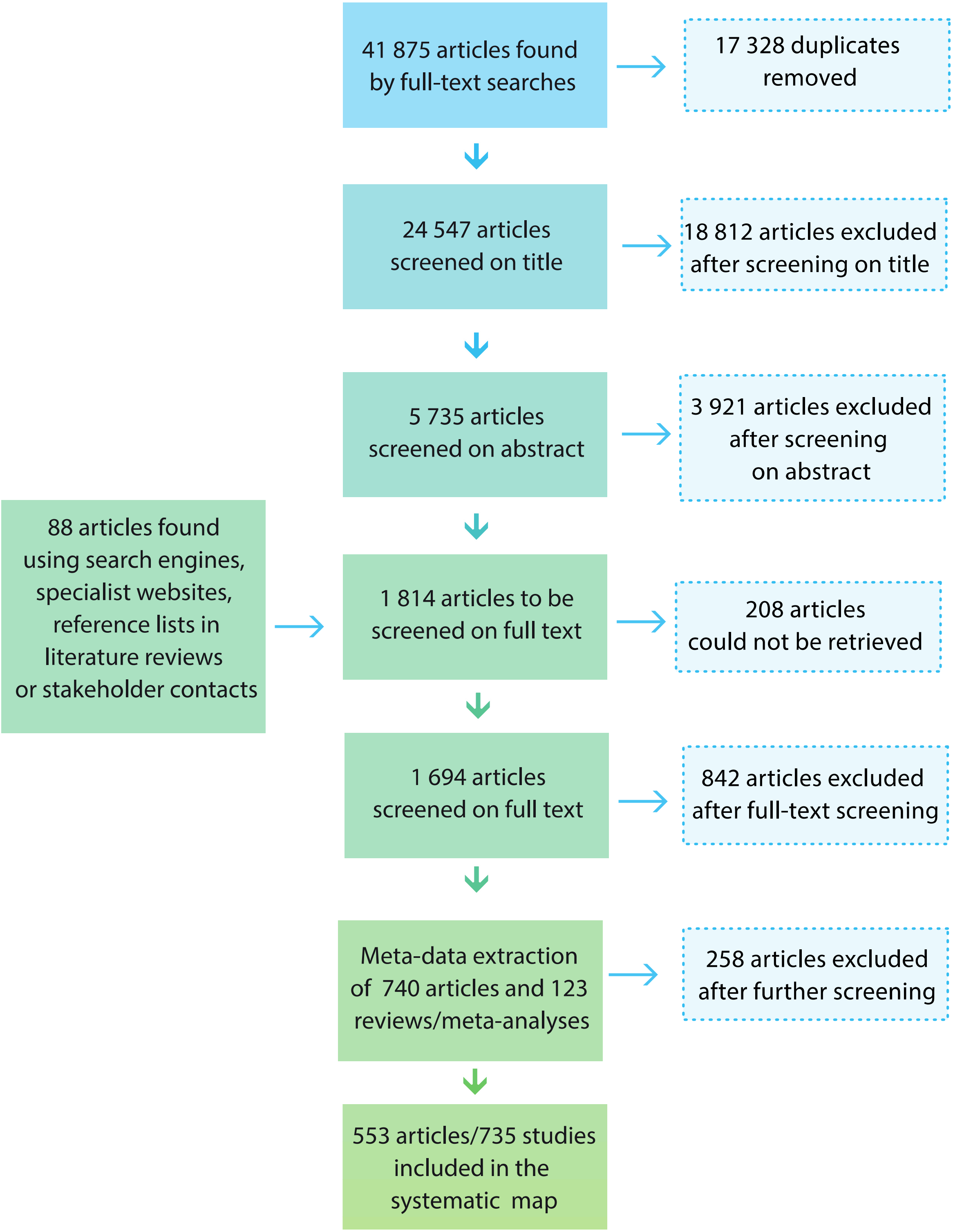
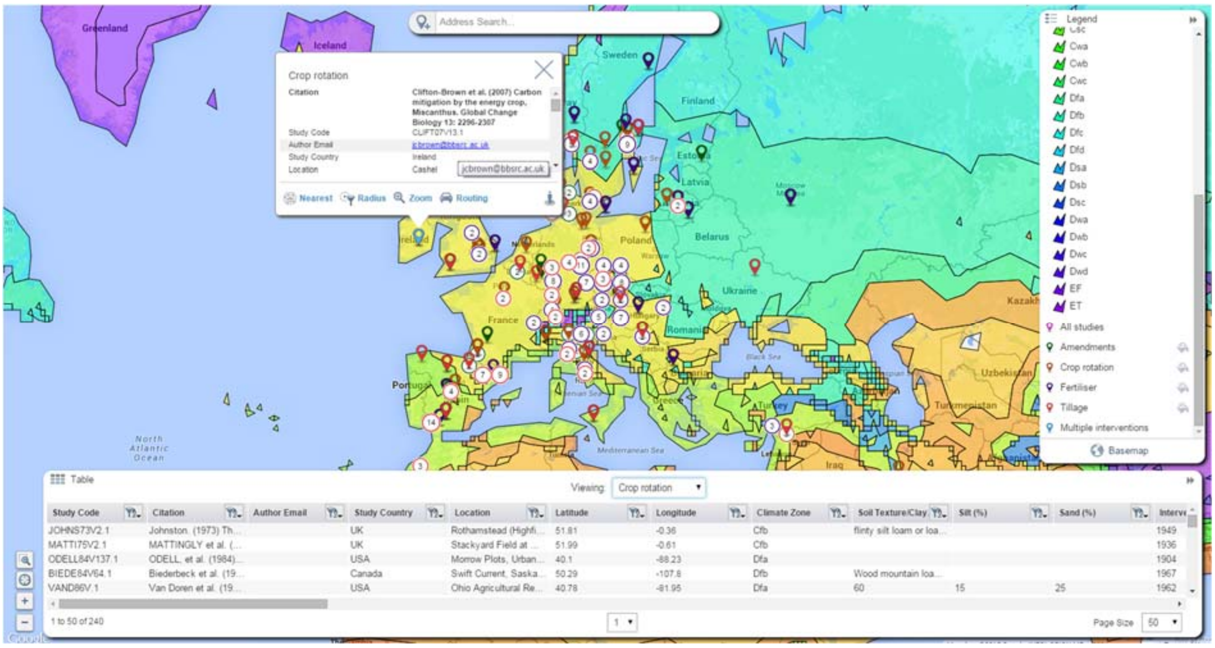


Figure 3. Schematic of the mapping process.
Displayed are the numbers of articles and studies retained and excluded at various stages of the mapping process.

Along with the presence of multiple measurements of single interventions, there were indications within the evidence base of multiple publication of the same data. Some of this appears to be re-analysis of existing datasets, which is not generally viewed as a questionable research practice, but a small number of included studies presented identical results in different platforms. These duplications were included in the systematic map database for completeness but should be screened by reviewers where further synthesis is employed on the outputs of this map in order to avoid double-counting of studies.

The map and its contents

An interactive version of this GIS is available at the following URL: http://www.eviem.se/en/projects/Soil-organic-carbon-stocks/. This map (see an example snapshot in Figure 4) displays various layers of data for all included studies ('All studies' layer), and for studies investigating each intervention group ('Amendments', 'Crop rotation', 'Fertiliser', 'Tillage' and 'Multiple interventions' layers). The map displays the systematic map database visually and allows the user to identify, collate, filter and examine meta-data for single or grouped studies based on a wide variety of variables from within the database. Also displayed are the Köppen-Geiger climate classification zones. Finally, users can locate citations and full texts by following a Google Scholar link, identify other research on soil C by the review lead authors, and email corresponding authors directly (where email addresses were available) directly from within the map. For examination and manipulation of bulk records from the systematic map database, however, the accompanying database file (Additional File 7) should be used. Full instructions of how to use the systematic map database and the map GIS can be found at the above web address.



Crop rotation

Citation: Clifton-Brown et al. (2007) Carbon mitigation by the energy crop, Miscanthus. Global Change Biology 13: 2296-2307
Study Code: CLIFT07V13.1
Author Email: jcbrown@bbsrc.ac.uk
Study Country: Ireland
Location: Cashel

Viewing: Crop rotation

Study Code	Citation	Author Email	Study Country	Location	Latitude	Longitude	Climate Zone	Soil Texture/Clay	Silt (%)	Sand (%)	Interv
JOHNS73V2.1	Johnston. (1973) Th...		UK	Rothamstead (Highfi...	51.81	-0.36	Cfb	flinty silt loam or loa...			1949
MATTI75V2.1	MATTINGLY et al. (...		UK	Stackyard Field at ...	51.99	-0.61	Cfb				1936
ODELL84V137.1	ODELL, et al. (1984)...		USA	Morrow Plots, Urban...	40.1	-88.23	Dfa				1904
BIEDE84V64.1	Biederbeck et al. (19...		Canada	Swift Current, Saska...	50.29	-107.8	Dfb	Wood mountain loa...			1967
VAND86V.1	Van Doren et al. (19...		USA	Ohio Agricultural Re...	40.78	-81.95	Dfa	60	15	25	1962

1 to 50 of 240

1

Page Size 50

Figure 4. Screenshot of the systematic map geographical information system (GIS). Figure demonstrates interrogation of the 'Crop rotation' layer for a study in Ireland (Clifton-Brown et al. 2007) and showing the filtered systematic map database presented below the map. Also shown in this screenshot is the Köppen-Geiger Climate Classification.

Study systems

Population

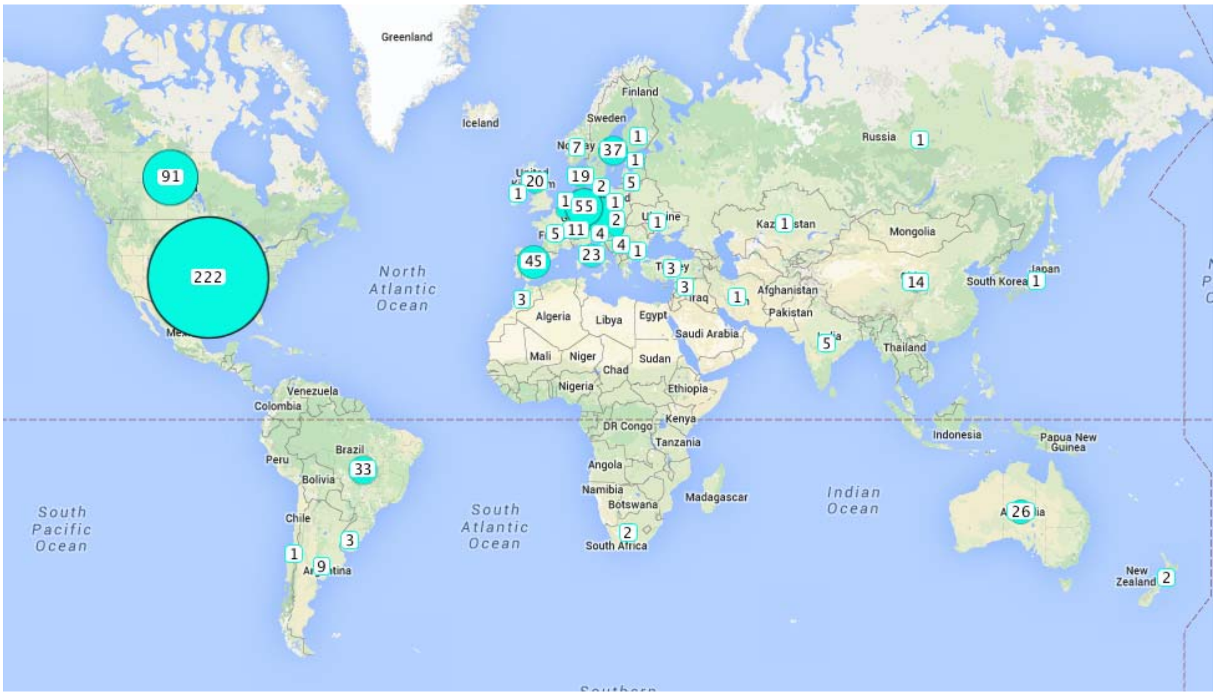


Figure 5. World map showing the number of studies per country.

Studies were reported from 43 countries (Figure 5). In the following text, numbers in brackets indicate the number of individual studies in the categories described. Europe (319) and North America (313) were the most frequently studied continents, with the USA (221), Canada (91), Germany (55), Czech Republic (47) and Spain (45) being the most frequently studied countries. The most commonly studiedclimate zones were Cfb (213) with a marine mild climate with no dry season and warm summer; Cfa (191) with a humid subtropical climate with no dry season, and hot summer; and Dfb (159) with a humid continental climate with severe winter, and mild summer (Table 3).

Soil types were commonly described by classification, but soil textures classes were also reported. These ranged from 3 to 87 % clay, 5 to 85 % silt, and 1 to 87 % sand.

Table 3. The number of studies included in the systematic map per Köppen-Geiger climate zone. Multiple zones are for large scale studies ('and') or where the exact location is not given and may span several zones ('or'). The names of the climate zones are from Encyclopedia of the Earth (http://www.eoearth.org/view/article/162263)

Köppen-Geiger climate zone	Name of the climate zone	No. of studies
Cfa	Cfa - Humid Subtropical	191
Cfb	Cfb - Marine - Mild Winter	213
Cfb or Dfb	Cfb - Marine - Mild Winter or Dfb- Humid Continental Mild Summer, Wet All Year	2
Cfb and Csa	Cfb - Marine - Mild Winter and Csa- Interior Mediterranean	1
Cfb and Dfb	Cfb - Marine - Mild Winter and Dfb - Humid Continental Mild Summer, Wet All Year	1
Cfb or Cfc	Cfb - Marine - Mild Winter or Marine - Cool Winter	1
Csa	Csa- Interior Mediterranean	57
Csb	Csb- Coastal Mediterranean	
Dfa	Dfa- Hot Summer, Wet All Year	82
Dfa and Cfa	Dfa- Hot Summer, Wet All Year and Cfa - Humid Subtropical	1
Dfa and Dfb	Dfa- Hot Summer, Wet All Year and Dfb - Humid Continental Mild Summer, Wet All Year	1
Dfb	Dfb - Humid Continental Mild Summer, Wet All Year	159
Dfc	Dfc - Subarctic With Cool Summer, Wet All Year	13
Dfc or Dsc	Dfc - Subarctic With Cool Summer, Wet All Year or Dsc - Summer Dry With Cool Summer	1
Undetermined		6

Interventions

Five main groups of interventions were identified iteratively during screening: amendments (286), crop rotations (238), fertilisers (307), tillage (306), and multiple interventions (55). The numbers of studies reporting investigations of the subcategories of each group of intervention are shown in Tables 4 to 8. There was a degree of variability in the level of detail concerning interventions provided within articles, which accounts for the number of studies in the less specific intervention categories.

Table 4. The number of studies included in the systematic map investigating amendment interventions. Interventions are catalogued in a hierarchical way, such that column one is broad intervention groups, whilst columns two and three are more specific sub-categories from within the category in the preceding column.

Intervention			**No. of studies**
Crop residue	Unspecified		31
	Composted		4
	Uncomposted		40
		Green (mulch)	16
		Yellow	36
		Burned	18
Manure	Unspecified		63
	Liquid (slurry)		18
	Solid		112
		Cattle	9
		Poultry	3
		Composted	14
Lime			9
Sewage sludge			17
Green manure			24
Peat/sediment			10
Domestic compost/waste			4
Processed wood (sawdust/woodchips)			8
Different application methods			1
Bone meal/animal products			1

Table 5. The number of studies included in the systematic map investigating amendment interventions. Interventions are catalogued in a hierarchical way, such that column one is broad intervention groups, whilst columns two and three are more specific sub-categories from within the category in the preceding column.

Intervention			**No. of studies**
Monoculture vs:	Unspecified		1
	Complex rotation (4yrs)		32
	Rotations involving perennials		27
	2 or 3 yrs rotations		105
	Rotations involving 1 yr "perennials"		26
		Sewn fallow or grass	26
		Chemical or bare fallow	4
	Comparison of different crops	Unspecified	1
		Legumes +/- or legumes vs legumes	144
		Grass/ hay/ green fallow +/- or vs. each other	20
		Only crops	32
	Annual crop vs perennial crop (2 yrs or more)	Unspecified	61
	Multi cropping	Under-sown/intercropping (legumes or other)	20
		Double cropping	28
		Cover- or catch crops	17
	Energy crops	Energy crops vs. energy crops	9
		Energy crops vs. other crops	7
	Simple rotation (2 yr) vs. complex rotations (3 yrs or more)		46
Total			240

Table 6. The number of studies included in the systematic map investigating fertiliser interventions.

Intervention	**No. of studies**
Inorganic vs organic/other amendment (organic also given an amendment code to describe the type of organic used)	187
Different amounts of N	299
Different application methods	5

Table 7. The number of studies included in the systematic map investigating tillage interventions.

Intervention		**No. of studies**
No tillage (zero) vs:	Reduced tillage (minimum / conservation / disc / chisel / harrow / mulch / ridge)	103
	Conventional (traditional) tillage (mouldboard)	255
	Rotational tillage (occasional)	9
	Subsoiling (>30 cm depth conventional tillage)	7
Reduced tillage (minimum / conservation / disc / chisel / harrow / mulch / ridge) vs:	Conventional (traditional) tillage (mouldboard)	129
	Rotational tillage (occasional)	2
	Subsoiling (>30 cm depth conventional tillage)	7
Conventional (traditional) tillage (mouldboard) vs:	Rotational tillage (occasional)	1
	Subsoiling (>30 cm depth conventional tillage)	8
Rotational tillage (occasional) vs:	Subsoiling (>30 cm depth conventional tillage)	0

Table 8. The number of studies included in the systematic map investigating multiple simultaneous interventions. Multiple confounded interventions involve more than two interventions simultaneously.

Multiple Interventions		**No. of studies**
Crop rotation and:	Amendments	4
	Fertiliser	2
	Tillage	1
Amendments and:	Crop rotation	5
	Fertiliser	6
	Tillage	3
Fertiliser and:	Tillage	0
Multiple confounded		35

Within crop rotation, studies comparing monoculture with rotations (221) and comparing similar rotations with different crops (197) were most common. Within these, monocultures versus 2- or 3-year rotations were particularly common (104), as were comparisons with or without legumes (143).

Within amendments, the most frequently studied interventions were manure (219) and handling of crop residues (152). Some 40 studies investigated residue burning, whilst other common amendments studied included uncomposted (33) and yellow (i.e. straw) (31) residues. Solid manure (112) was the most common form of manure application.

Within fertilisers, comparisons between different levels of inorganic fertiliser were most common (299), whilst comparisons between organic and inorganic fertilisers were slightly less common (187): only 5 studies investigated different methods of nitrogen application.

Within tillage, comparisons between no tillage and conventional tillage were most common (254). Conventional tillage versus reduced tillage (129) and no tillage versus reduced tillage (103) were also common. A small number of studies (between 2 and 9) considered rotational and subsoil tillage relative to conventional or reduced tillage.

For multiple interventions, crop rotation was combined with other single interventions in 7 studies, amendments were combined with other interventions in 9 studies (excluding those combined with crop rotation (4)), and multiple confounded studies numbered 35 (including 3 or more combined interventions). These combined studies did not attempt to identify the effect of individual interventions.

Figure 6 displays the distribution of interventions across the continents. Amendments dominate in Asia, Australasia and Europe. Tillage dominates in Australasia (equally with amendments), North America and South America. Crop rotation dominates in South America. Fertiliser dominates in Asia and Europe.

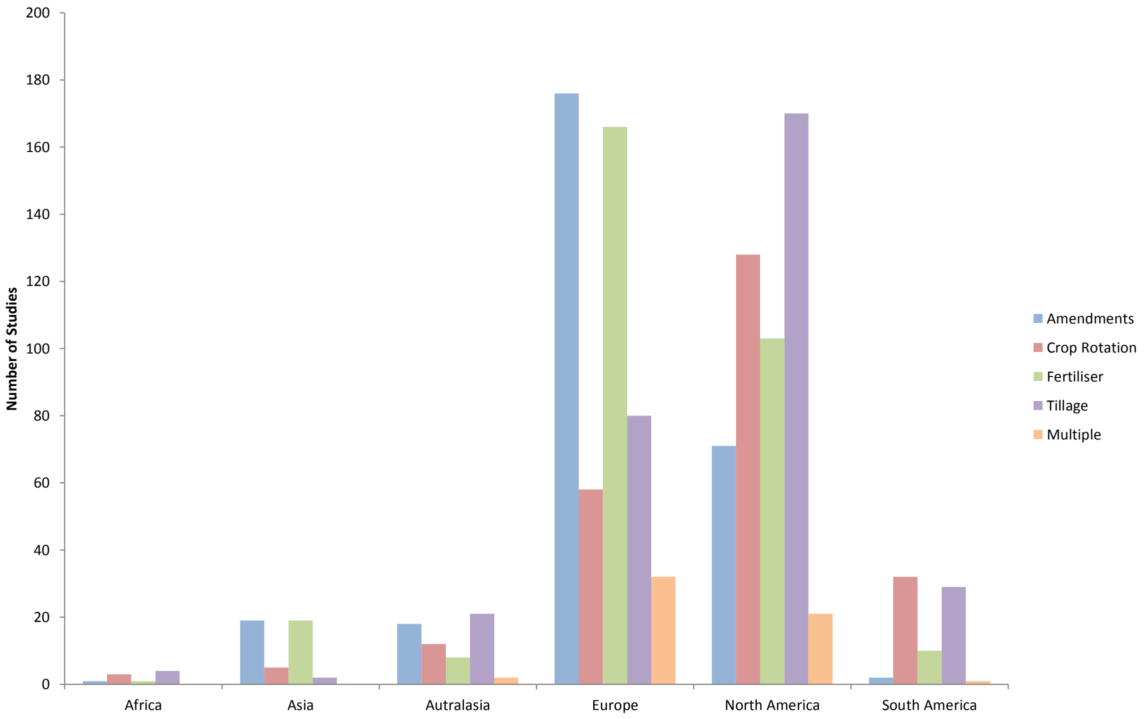


Figure 6. The number of studies undertaken across continents. Numbers are separated by intervention group investigated within each study. Studies may be present in more than one intervention category.

Comparators

The majority of studies (645) were performed as comparator-intervention (CI) experiments, comparing an intervention with a spatial control (i.e. a nearby similar set of samples that had not experienced the intervention of interest). Before-after-control-intervention (BACI) studies were less frequent (81), reflecting the lack of baseline measurements in this evidence base. Only 3 studies were before-after experiments, comparing baseline conditions with those after an intervention, in the absence of a spatial control. A small number of studies failed to describe their designs in sufficient detail to ascertain the type of comparator used (5). A total of 53 studies met our criteria of ‘time series’ (i.e. 30 years or more with 3 or more sampling time-points).

Experimental designs

Study duration

A large number of studies lasted between 10 and 19 years (340) (Figure 7). Some 190 studies lasted between 20 and 29 years, 144 were between 30 and 59 years, and 23 were in excess of 60 years duration. For 6 studies it was not possible to ascertain study duration from the written report. Study starting dates range from 1843 to 2000 (Figure 8), with the majority of studies commencing after the 1950s (93% of 703 studies provided start dates). The slight drop in studies after the peak in the 1980s results from our requirement for a minimum intervention period of 10 years. The longest studies occurred in Europe (153 years) and North America (114 years), with other continents showing lower maximum study lengths (38-59 years).

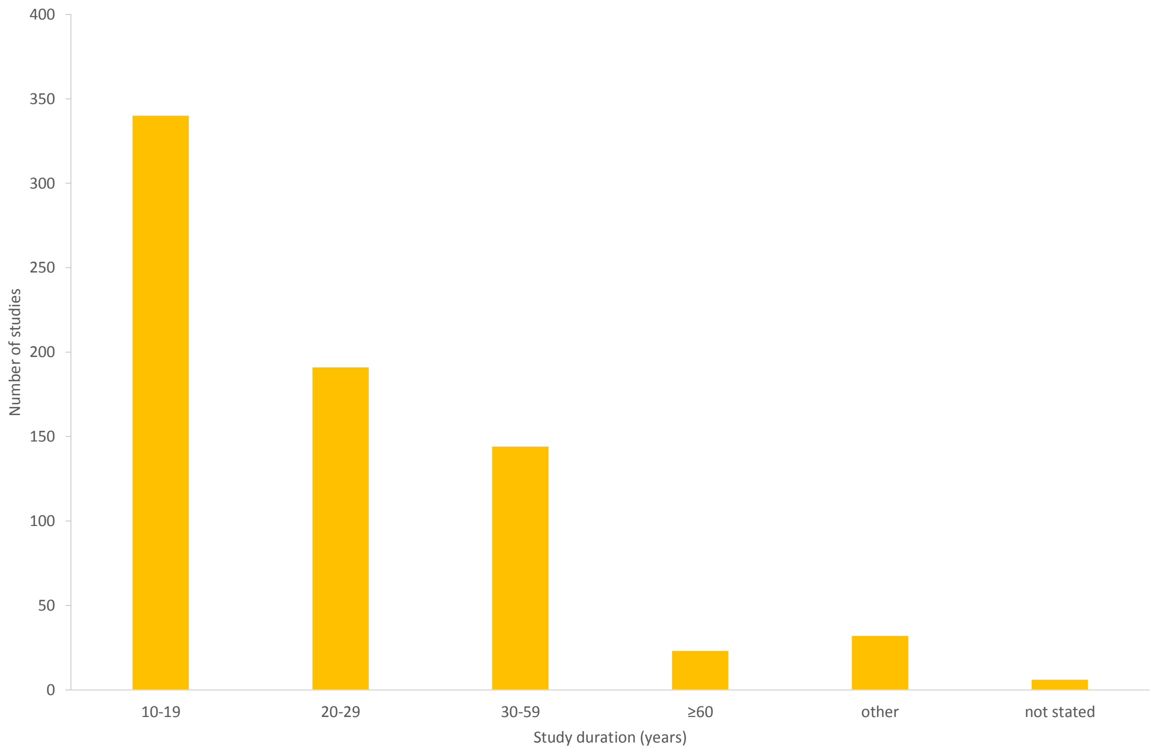


Figure 7. Duration of studies included in the systematic map.
‘Other’ includes complex, multiple study periods that span categories.

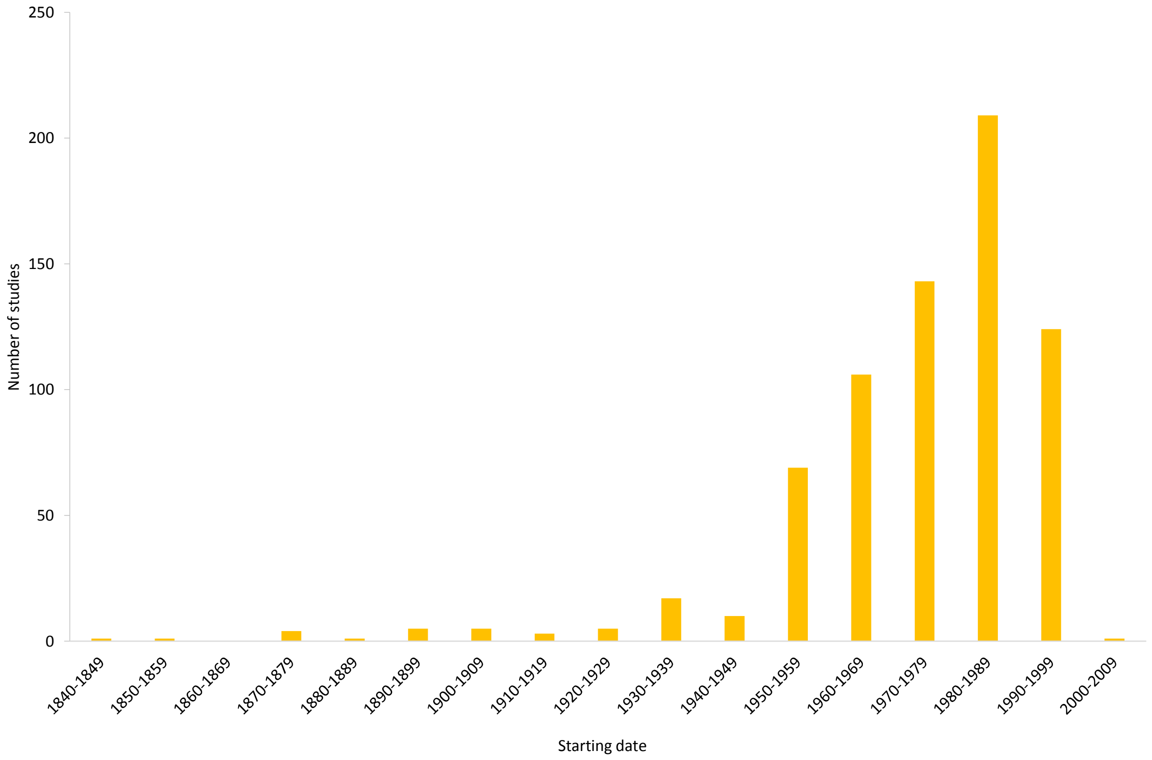


Figure 8. Intervention start year for studies included in the systematic map.
Thirty-three studies failed to report this information.

Treatment allocation (i.e. from randomised to purposeful)

Table 9 details the various experimental designs that were used across the evidence base and the number of different studies reporting them. A total of 377 studies used some form of randomised treatment allocation: 276 of these were randomised blocked designs, 76 were randomised split plot designs, and 25 used randomisation alone. Replication is described below, but randomisation should be considered alongside replication. At low sampling intensity randomisation can give a false indication of bias mitigation: systematic bias may be introduced by chance allocation of all samples with a particular characteristic to one treatment. Replication was low throughout the evidence base: for both randomised (i.e. explicitly stated as randomised) and blocked (i.e. employing any form of non-randomised blocking) designs there were 1 to 14 spatial replicates (median = 4). Purposeful and paired designs showed a higher upper range of replication in general (1 to several thousand replicates) (median = 3).

Table 9. The number of studies included in the systematic map employing various categories of experimental design. Studies could be represented by multiple categories where mixed designs were used.

Experimental Design	No. of studies
Randomised design (blocked)	276
Not stated	126
Split-plot design	95
Randomised design (split-plot)	77
Time series	53
Purposeful	45
Blocked design	33
Randomised design	25
Factorial design	14
Paired design	12
Strip-plot design	10
Latin square	8

Sampling designs

Spatial replication

The majority of studies sampled the median of 4 true spatial replicates or fewer (580), with 73 studies sampling between 5 and 10 replicates (Figure 9). Some 74 studies failed to report the level of replication used or reported conflicting details. True replicates are defined as sample units that are made at the same level as that at which the intervention or exposure is applied.

For the purposes of this review question, sampling within true spatial replicates (sub-sampling) does not contribute to measures of variability in the effects of interventions (an assessment of accuracy), but it does contribute to increased precision. The term *pseudo-replication* describes the use of within sample replication as a measure of between sample variability in analyses and is typically a criticism of statistical analysis. However, in a synthesis, pseudo-replication only occurs if reviewers treat sub-sampling within true replicates as true replication in a meta-analysis. Thus we use the terms pseudo-replication and sub-sampling somewhat interchangeably but this is not necessarily a criticism of the primary research.

A large number of studies did not report whether sub-sampling was undertaken (282). Of those that reported sub-sampling, 3 sub-samples was the most common degree of sampling effort (94), with 286 studies taking 2 to 6 soil samples per replicate. Subsampling ranged from 1 to 100 subsamples per replicate (Figure 10).

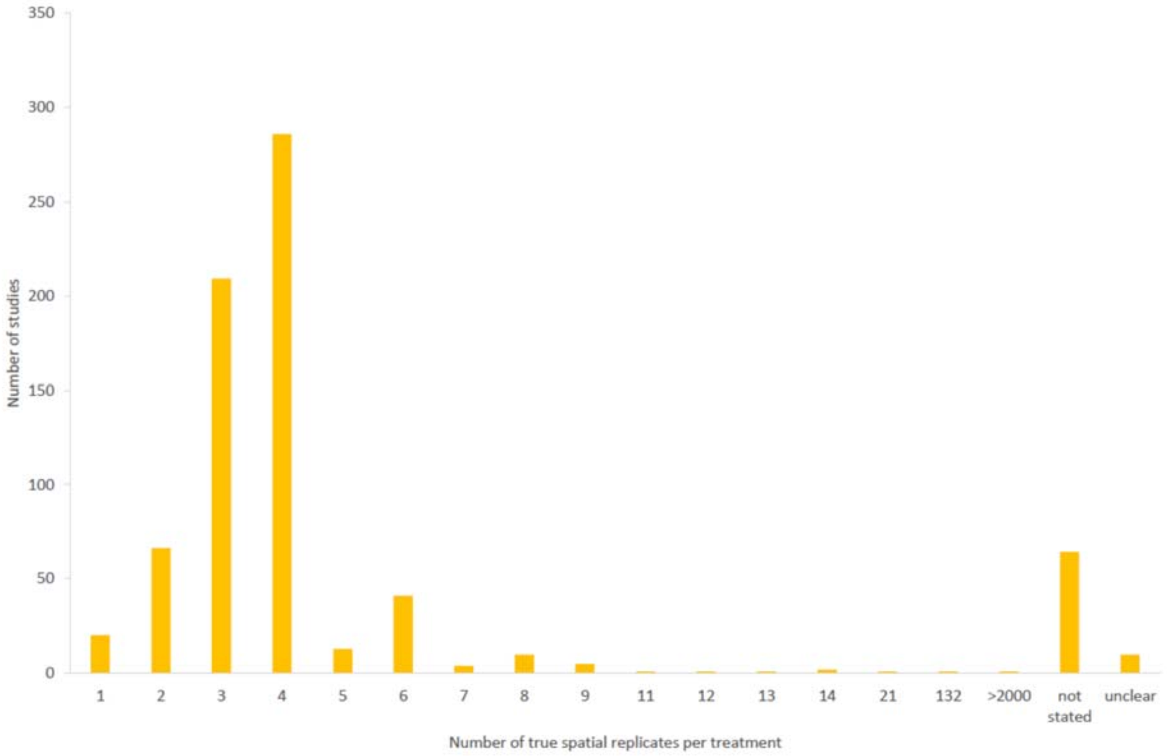


Figure 9. Spatial replication across the studies included in the systematic map.

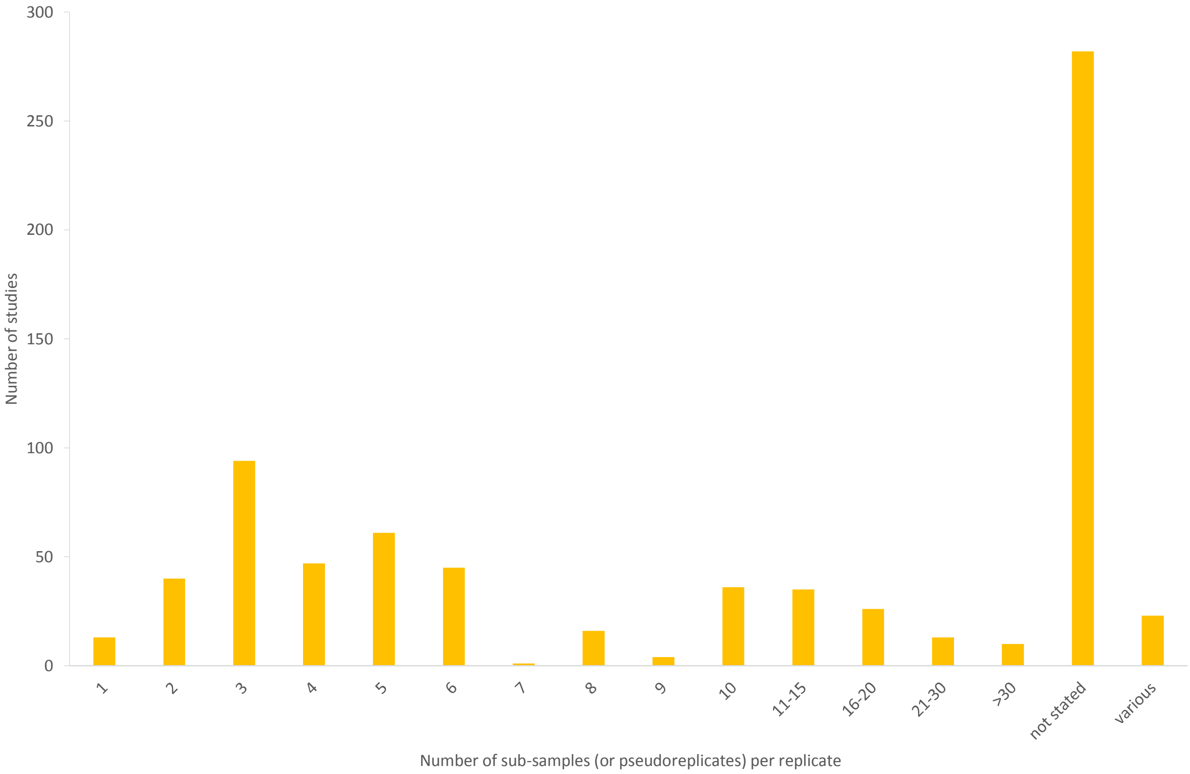


Figure 10. Subsampling (or pseudoreplication) within studies included in the systematic map.

Temporal replication

The majority of studies performed sampling at only one point in time (446), whilst a relatively small number failed to report the number of sampling occasions (9). The remaining studies in the evidence base reported collecting samples between 2 and 108 times.

Sampling depth

A total of 704 studies reported their sampling depth. The majority of studies reported just one (382) or two (90) carbon measurements across the depth of each soil profile (i.e. one or two samples per core), with other studies measuring up to 20 independent layers within each core (mean=2.3 ± 2.0 (SD)). Some 52 studies did not report sampling depth. Soil cores were taken down to maximum depths of between 2.5 and 330 cm, with a median maximum depth (i.e. the lowest point sampled) of 20 cm (mean=29.8 ± 25.5 cm (SD)) (Table 10). A total of 257 studies reported carbon measurements below 25 cm depth.

Table 10. Soil sampling depth statistics for studies included in the systematic map. Number of samples per soil profile refers to the number of individual subsamples taken from each core across different depths where data are reported separately for each sample. Maximum depth sampled refers to the deepest extent to which soil samples were taken where data are reported.

Number of samples per soil profile (cm)	
Statistic	**Value**
Minimum	1
Maximum	20
Mean	2.28
Standard deviation	1.97
Studies failing to mention number of soil depths sampled	33
Total number of studies	737
Studies taking 1 sample per soil profile	382
Studies taking 2 samples per soil profile	90

Maximum depth sampled (cm)	
Statistic	**Value**
Minimum	2.5
Maximum	330
Median	20
Mean	29.84
Standard deviation	25.50
Studies failing to mention maximum sampling depth	54
Total number of studies	737
Studies sampling to 25cm depth or more	257

Data

Most studies reported results in the form of tables (470) and figures (180) or both (81), with very few reporting data solely within the text (4). A large proportion of studies failed to report any measure of variability along with treatment means (380), whilst 17 studies did not report variability but provided sufficient data to allow measures to be calculated for each intervention group. Standard errors (118), standard deviations (55), coefficients of variation (5), and 95% confidence intervals (6) were reported across the evidence base. Some studies (56) reported variability measures for study means rather than treatment means (i.e. aggregating the variability for individual treatments). For a further 120 studies an overall significance threshold was provided (least squared difference, LSD) that can be used to generate an overall variability measure for the study. Some 41 studies reported a variability measure but failed to describe what the measure was.

Critical appraisal

Figure 11 displays the critical appraisal scores for the five domains investigated. In general, the evidence available scored moderately for spatial (true) replication, with 488 studies sampling 3 or 4 replicates, although 83 studies failed to describe replication. The majority of studies sampled at only one point in time, thus resulting in 588 scores of 0 for temporal replication. Treatment allocation (i.e. experimental design) was generally of high validity (i.e. split-plot, strip-plot, Latin square, blocked, randomised, or exhaustive designs), with 533 studies demonstrating some form of blocking or randomisation. Considering the low level of replication across the evidence base it was decided to combine scores for blocked and randomised designs. A large number of studies failed to report sufficient detail to be scored for this domain, however (126). Study duration was generally low-scoring, with 353 studies lasting 10 to 19 years and a roughly equal proportion in the middle (20-29 years) and higher (≥30 years) categories combined (204 and 172, respectively). The evidence scored moderately for sampling depth, with 320 studies sampling the plough layer (i.e. 0-25 cm) in single or multiple sections, or deeper (>25 cm) single section samples.

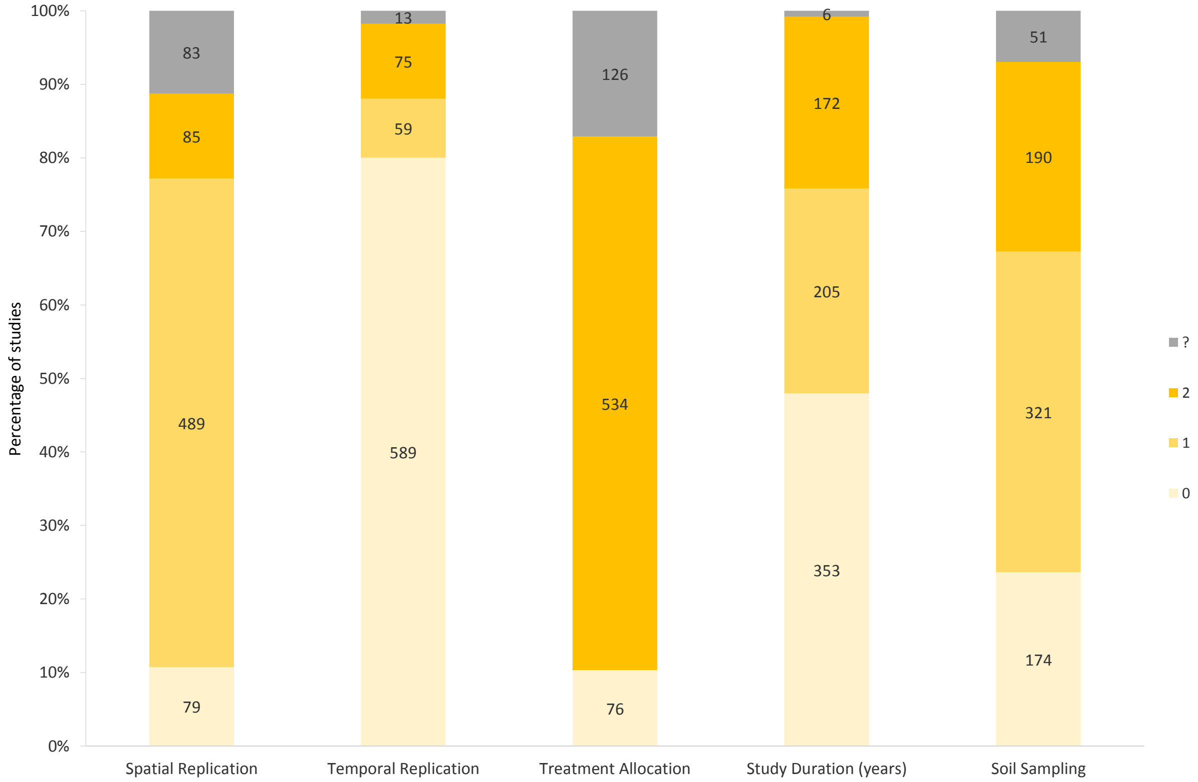


Figure 11. Critical appraisal summary scores for studies included in the systematic map.
Scores are displayed across 5 key domains (see Methods for details).

Level of reporting detail

Between 11 and 24% of studies in each intervention category failed to report information pertaining to one of the critical appraisal domains described above and 4 to 14 % of studies had missing information for multiple domains. A total of 28 % of studies across the entire evidence base failed to report some critical information. Three studies were excluded during meta-data extraction and four during bibliographic checking for reasons of methodological detail (see Additional File 8).

Further outputs from the systematic map

Along with the systematic map database and interactive online GIS, two further databases have been produced. As discussed above, a database of reviews and meta-analyses pertinent to soil organic carbon and agricultural management has been produced (see Additional File 5). This database is relevant for all climate zones and study durations, since they were identified prior to application of the full inclusion criteria.

In addition, a further database of long-term experimental (LTE) study sites has been produced, which summarises the geographical locations described across the evidence base included within the systematic map (see Additional File 9). This LTE database lists all locations for which research has been published that falls within the inclusion criteria for this systematic map. As with the systematic map database, latitude and longitude are taken from publications where provided, or based on the cited location. Specific site names were standardised across publications that referred to the same site with slightly different names. Non-specific site names, e.g. Lethbridge, that were associated with known sites (i.e. sharing identical or very similar [<0.3 degrees separation] geographical coordinates) were removed where known sites were already featured in the database.

Discussion

The systematic map presented here details studies within the Köppen warm temperate and snow climate zones [39] that investigate the impact of agricultural practices on SOC. The map database is an interactive, freely available resource for further synthesis of evidence. This report highlights potential synthesis topics along with research areas that would benefit from further investigation. We detail these knowledge gluts and gaps below (see *Implications for Policy, Practice and Research*, below).

Evidence on how agricultural management may change SOC is a timely issue both regarding food security issues as well as global agreements on how to mitigate climate change. Recent suggestions have promoted discussions at the climate change convention (COP21) in 2015: that if soils can sequester 4 g/kg SOC every year globally this could largely counteract the current climate change, at least on a temporary basis [43]. However, the evidence on how this might be achieved and how large the potential is for storing more carbon needs to be summarised and evaluated. Generally, agricultural soils contain 25 to 75 % less SOC than soils in undisturbed or natural ecosystems [44], though regional climatic conditions will also regulate the rate of carbon sequestration. Mapping evidence of SOC and agricultural management on a global scale will provide decision makers and researchers dealing with climate change issues with a database where evaluated evidence can be found. Such a database also provides information on typical management of soils in different regions that is important for climate change policy. The map presented herein is restricted to current management practices in the published literature and does not contain novel practices that combines both agricultural output as well as climate change mitigation, such as biochar application [45], inclusion of biomass crops, or novel crops such as perennial cereals [46].

Potential increase in food production and other ecosystem services, such as nutrient and water retention, can be linked to agricultural management that promotes soil carbon, although evidence is needed to quantify these relationships in order to provide solid policy recommendations [47]. The map can be used both on regional and local scales to find evidence and quantify changes of soil carbon under local management practices. The metadata in this database does not include data on yields specifically, but a substantial part of the published evidence contains yield measurements, since this is a key agroecosystem outcome. The maintenance of sequestered C may be costly, and evidence is needed to provide incentives for farmers to adopt practices that increase not only SOC but also many other ecosystem services [48, 49].

The required minimum duration of investigation for studies in this systematic map was 10 years and the majority of studies were initiated after the 1950s. Also included were 23 studies in excess of 60 years in duration, with the earliest starting date of 1843. Some of these studies do not comply with current standards in terms of replication and statistical design. Often, the original plots in experiments commenced in the 19th century were laid out in large strips without replication, and these strips may have been later subdivided to allow new treatments to be introduced [50, 51]. Other long-term experiments also lack a well-defined statistical design but do have replicates [52].

While long-term experiments are essential for evaluation of resilient soil properties, such as soil pH and soil organic matter, their long duration also poses challenges in keeping the experiments relevant [50]. For obvious reasons, these studies have often been altered to accommodate new crop varieties, plant protection, machinery and nutrient supply, for example. For each situation, a change in design must be balanced between the need for continuity on the one hand and relevance to the prevailing current agricultural management on the other.

An example of a conflict between the desire for continuity and relevance is the use of animal manure. At the time of initiating the early long-term experiments, solid farmyard manure (FYM) including bedding was traditionally used in common agriculture, and therefore also in published experiments. In the last few decades there has been a general shift to more intensive farming systems based on animal slurry in many countries. This transition in agriculture resulted in a decision to change to slurry application in the Askov long-term experiment after the manure plots had been based on FYM for almost 80 years [52]. As the slurry was applied to plots previously treated with FYM, the change to slurry application would not allow a direct experimental comparison between slurry and mineral fertiliser: the treatment history of the experimental groups would differ, causing significant confounding. Animal slurry is not a common intervention in the systematic map, which likely reflects that the inclusion of this type of manure has not been undertaken for the required time span to be included (10 years or more).

Depending on previous land use history, climate, or site conditions, identical experimental treatments can result in either increases or decreases in SOC [e.g. 11] and among the experimental designs used within studies in the map there are some that are more powerful than others in determining differences among management interventions. We found 81 before-after-control-intervention (BACI) studies, which are considered to be superior to comparator-intervention (CI). BACI studies contain quantitative information both on the starting condition of intervention and control groups, and also on the changes following intervention. Furthermore, BACI studies presenting time-series are also valuable for parameterising dynamic SOC-models [53]. CI studies, which are most frequent in this map, can only compare different types of management after a certain time of intervention and no dynamic information can be obtained from these studies. Limitations of CI studies are also related to field-scale variability of SOC, which can be caused by prior land-use, pedology, topography and patchiness of organic amendment additions or spots of charcoal derived from burning [54]. Such variability can lead to misinterpretation of results when differences in initial conditions are not accounted for and the whole effect is interpreted as treatment effect. Although locations for field experiments are most often carefully selected, spatial variability is still present [55]. To account for this problem, randomised designs are often used, which are sometimes considered superior to BACI designs. However, initial differences between treatments are often not controlled for, even in replicated long-term field experiments [14]. Large experimental plots, poor experimental design and low replication usually exacerbate this problem [56].

Approximately half of the studies in the systematic map did not report bulk density (soil dry weight divided by its volume) (365 of 735 in total). Whilst many studies measured bulk density, the data were not reported in many instances. Indeed, this parameter may be considered a good qualitative indicator of SOC changes, especially where repeated samples are taken across the soil profile. The underlying hypothesis is that the soil density (i.e. bulk density) variation is negligible among treatments; therefore SOC measurements of disturbed samples taken at the same depth do not introduce large bias. This approach is justifiable for experiments addressing soil quality, and considering the cost and time required to carry out the additional bulk density sampling and analysis. For example, studies reporting bulk density in the systematic map constituted 35, 40 and 56 % of the evidence for amendment, fertiliser and crop rotation intervention categories, respectively.

However, SOC concentrations alone may be inadequate if the focus is on a quantitative SOC balance, such as the carbon sequestration capacity for climate change mitigation, in which SOC stock should be accurately assessed. In particular, when the management under investigation could significantly alter soil density, as with tillage interventions [57], bulk density becomes a fundamental parameter for calculating SOC stock. This is likely the reason why 67% of the studies investigating tillage effects reported this parameter, the highest percentage among all the intervention groups.

Although bulk density measurements give more transparency to the experimental results they may not guarantee the greatest accuracy if depth is not properly considered. For example, soil with the same SOC concentration but with a different density in relation to treatments, may be erroneously considered as having different SOC stock if the same depth is taken into account.

In much of the evidence base, most of the comparisons among treatments have been made by the simple multiplication of SOC concentration and bulk density, considering fixed depths. This method often introduces significant errors when soil bulk density differs among treatments under study, such as between tillage and no-tillage experiment or change in land uses [58, 59]. In order to undertake more rigorous quantitative SOC estimation, both bulk density and calculations based on equivalent soil mass (ESM) must be reported [60, 61]. Although the ESM approach increases the complexity of SOC assessment, it is a necessary step for reporting and accounting of CO_2 emissions and removal under potential post-Kyoto agreements (UNFCCC). It is certainly recommendable that precise guidelines for SOC sampling and analysis could be developed by recognised international bodies such as the IPCC, and adopted as standards in field experiments. As demonstrated in the systematic map, substantial heterogeneity in the reported outcome exists among studies that may confound the interpretation of results.

Regional differences exist in the types of agricultural practices that were emphasized in research studies. In North America a central focus was on how tillage and crop rotations affected soil C. A similar trend was observed in South America, although many fewer studies occurred there. In contrast, European studies tended to emphasize effects of amendments and fertilisers, as did the few studies included for Asia.

The differences between America and Europe may reflect differences in agricultural systems and in the practices deemed most relevant to agricultural policy. In the USA, large federal investments were made for conservation tillage systems (defined in the USA as ≥30 % of the soil surface covered after planting) as a way to reduce erosion in corn-soybean-small grain rotations, which occupy millions of ha of often uneven topography in the central and south-eastern part of the country. Soil erosion on cropland decreased 43 % between 1982 and 2007, largely due to adoption of conservation tillage [62]. Sequestration of C is a co-benefit that has also received substantial federal funding, partly for the establishment and maintenance of long-term research plots [63]. Research results on different forms of tillage, and on permutations on crop rotations, are fairly transferable across the large expanses of cropland in the Great Plains and Corn Belt of the USA.

With the wide variety of soil types, land use and climatic conditions across Europe, specific sets of tillage and crop rotations are not as widely appropriate as in the USA. In addition, agricultural policy in the EU has rewarded a wider range of options for increasing soil C [5]. In Europe, application of C-rich inputs, or higher fertiliser rates to increase production and incorporated residue, are frequently used to increase soil C on cultivated land. One reason may be that the availability of manures and composts is more locally available than in much of the USA. There is closer proximity of cropland to farms where from manure is produced and to large urban areas where products like sewage sludge and compost are derived. Moreover, reduced tillage in wet humid climates in Europe reduces yields [64], is considered to bring more risk of fungal attack, reduced emergence, crop failure [5], and N loss via dentitrification [65].

Only 12 of the 736 studies in our systematic map were grey literature, published in the form of organisation reports. This is somewhat surprising, since long-term study sites frequently document their activities in the form of reports (e.g. the Rothamstead reports archive; http://www.era.rothamsted.ac.uk/eradoc/books/1). Grey literature was sought via multi-language searches of organisational websites, web-based search engines and knowledge of the expert review team. Research from Eastern Europe and Russia was conspicuous in its absence (a notable 'knowledge gap'), possibly resulting from the lack of digital cataloguing and the use of native-languages making them difficult to find.

Conclusions

Limitations of the evidence base

Our systematic map details the setting and design for a large number of studies spanning a broad range of interventions, experimental designs, and geographical locations. Across this diverse evidence base, the level of reporting was typically good, with 72 % of studies reporting all information for each of the five critical appraisal domains (spatial replication, temporal replication, treatment allocation, study duration, sampling depth). A very small number of studies were missing multiple pieces of information (8 %). A small number of studies cited other articles for their methods that were not accessible because they were grey literature (e.g. dissertations or reports) or were not written in a language covered by this review.

Of those studies with a good level of reporting of methodology, designs were typically of a moderate level of reliability. Spatial replication was typically low; not a surprising finding for field-scale studies, but somewhat disappointing for studies undertaken on smaller scales as identified herein. The majority of studies (61 %) performed ‘snapshot’ style measurements recording C only once. Only 53 studies met our criteria of providing ‘time series’ data; with over 30 years study duration (i.e. time between start of intervention and final measurement) and more than three time point measurements across this period. Experiments were typically well-designed, with the vast majority involving blocked and/or randomised designs. It is worth noting that a randomised design with a low level of replication may not be as reliable as a blocked design at the same level of replication. Study duration was fairly evenly distributed across the three categories; 10- to 20- year experiments were common, whilst less than 30 studies lasted over 60 years. Similarly, studies sampled across a variety of study depths. Some studies were included that sampled to shallow depths (minimum of 2.5 cm) because their aim was to investigate soil microbial activity, but the average soil sampling depth was 30 cm (median 20 cm). A lower number of studies with data from depths below the plough layer, roughly 30 cm, means that when comparing effects of management on total carbon stocks, there will be less evidence for changes in the whole profile, and a possible bias could occur if this is not considered in future reviews [66].

A large proportion of the evidence base did not provide measures of variability for mean values (51.8 %). Of the remaining studies, measures of variability could be calculated (n=17) (i.e. calculated for each treatment compared) or reliably estimated (n=176) (i.e. an overall mean across multiple treatment groups) from the information provided. This means that over 50 % of the evidence base would not be included in any form of reliable quantitative synthesis; a surprisingly high proportion.

In addition to missing information, a number of studies described aspects of their study design that were not reported in their results. This included selective reporting of treatments (for example, where a subset of treatments were described from a larger experiment). In some cases this missing information may be traceable through citations and references to other studies, but in other cases it is impossible to assess reasons for this selectivity. Selection of more significant results could be causing significant publication bias, for example. A further problem for synthesis occurs where authors do not report non-significant results; for example, stating purely that statistics were non-significant or that presented results are averaged across non-significant variables. Non-significant results are vital parts of synthesis, including meta-analysis. By omitting non-significant results the evidence base is overemphasising significance in the same way that non-publication of non-significant studies causes publication bias.

Limitations of the systematic map

As with the majority of systematic reviews and systematic maps undertaken at present, our map relied heavily upon the powerful search abilities of digital search engines and academic databases. However, we also employed physical article retrieval where databases identified potentially relevant articles that were not available in digital form, strengthening our methods. Some 208 articles could not be accessed in full text. This was caused by a combination of factors, including an inability to locate the reference (due to citation transcription mistakes or spurious citations) and a lack of journal subscriptions. This rate of retrieval is high for systematic reviews, which often have much lower rates [e.g. 67].

It is, of course, possible that evidence may have been missed by our systematic map. However, the high numbers of academic databases, organisational web sites and research languages used in our review reduce the likelihood of missed evidence. Additionally, the substantial efforts to include evidence through bibliographic checking of 127 relevant reviews mean that very little relevant evidence should exist that has not been catalogued.

Our review was restricted to research published in Danish, English, French, German, Italian and Swedish. Although this is a relatively comprehensive list of languages we did find a small number of articles (27) that were not included at full text because they were in other languages, namely: Chinese, Hungarian, Portuguese, and Russian. Given that Brazil forms a large proportion of the relevant climate zones in South America, future updates of the review may benefit from the inclusion of Portuguese language expertise. Similarly, Russia forms a large area of relevant climate zones that was represented by only one study in our map. Russian subject and language expertise may add value if additional sources of evidence from Russia are identified.

One of the inclusion criteria that was critical to the reliability of our systematic map was the requirement for interventions to have been in place for a minimum of 10 years. This was used to increase the probability of identifying statistical differences in SOC across land use and management over time. However, this restriction means that more recent forms of land management were not included: for example the use of biochar application [45], inclusion of biomass crops, or novel crops such as perennial cereals [46].

Our map includes a basic coding of variables related to validity of the evidence base, including extensiveness of soil sampling, spatial and temporal replication, study design, and study duration. In addition, we have excluded studies with unacceptable risk of bias. This basic critical appraisal must be extended before further synthesis in the form of meta-analysis can be undertaken.

Implications for policy, practice and research

Our systematic map, whilst not synthesising evidence to provide summary effect estimates of particular interventions, can be used in a range of different ways by decision-makers (i.e. policy and management stakeholders) and primary and secondary researchers.

Systematic maps can be used to investigate existing policies to examine whether current practices are supported by the evidence. For example, prescribed burning of crop residues is currently banned in Europe due to the cross compliance regulation of the Common Agricultural Policy (EC 1306/2013) and yet there is a fairly limited evidence base with respect to the impacts of burning on soil C, particularly from Europe. This limited evidence base could be rapidly synthesised in an attempt to calculate summary effect estimates for the practice in relation to SOC to guide future policy.

This systematic map may have important implications in light of the recent EU decision (EC 529/2013), which introduces mandatory accounting of cropland and grazing land management for EU Member States. This policy implies that SOC changes related to management should be accounted for, which itself suggests the need for quantitative data from field experiments to aid Member States in developing more accurate accounting methods.

Many studies included in our systematic map did not primarily set out to quantify the impacts of agricultural practices on SOC, but are included since they provide background carbon data, for example as part of an investigation of soil microbial activity. However, in the process of screening and appraising studies for the systematic map we are able to provide the following key points as suggestions for 'gold standards' in investigations of agricultural practices on SOC.

1. **Study design:** baselines should be measured before an intervention begins and controls should be sampled alongside management treatments (i.e. BACI design). Where baselines cannot be measured, long-term data series should be recorded (see Study duration, below). Control and treatment sites should be matched so that they are as similar as possible.
2. **Replication:** precision can be maximised by increasing the level of subsampling (within-sample variability) and increasing the number of samples taken over time (see Study duration, below). Accuracy can be maximised by increasing the number of true replicates (i.e. samples taken at the same level at which the treatment is applied). In practice, this means more fields and/or farms rather than quadrats within fields, if treatments are applied at the field level.
3. **Study duration:** since changes in soil C take many years to become detectable studies should aim to be long-term (i.e. ≥10 years). Time-series studies, which sample repeatedly over time, are more reliable than before-and-after studies, since rates of change can be assessed and measurement precision quantified and accounted for. Furthermore, sample processing and analysis should be consistent, which may prove problematic for multi-decadal studies.
4. **Soil sampling:** soil profiles that consist of multiple samples across depth and that sample well beyond the plough layer are more reliable than those that involve a small number of samples or that are shallow (i.e. <15 cm), particularly in studies including tillage and deep-rooted crops.
5. **External validity (generalisability):** the scale at which a study is undertaken affects the scale at which conclusions based on the findings can be made: study results cannot be reliably extrapolated beyond the system investigated. High levels of generalisability may be favourable where authors seek high impact or where broad policy and practice advice is sought. In these cases, studies should aim to match the scale of desired conclusions with the scale of their sampling. In practice, this may mean reallocating sampling effort across a broader geographical area or more diverse soil types. Where this reallocation occurs using a fixed budget this may mean a decrease in the signal to noise ratio of the findings: i.e. increased generalisability may introduce more heterogeneity in the sample population, and reduce the likelihood of finding a significant pattern. This payoff between external validity and precision should be considered carefully when designing experiments with limited resources. Authors should also be clear about the extent to which their study results can be generalised.

Knowledge gaps

Our systematic map can be used by funders and primary researchers to identify areas of research from within the map that are under-represented by evidence and then undertake the research necessary to help fill them. For example, there was a conspicuous lack of published research from parts of Russia that fell within our relevant climate zones. There are around 300 long term experiments in Russia mentioned by Shevtsova et al. [68] but rarely is evidence published in English. This means that a large volume of data produced from long-term soil experiments is collected in the Russian geographical network on field experiments (e.g. http://www.geo-set.ru/site/57) and also in the International Soil Carbon Network (ISCN) (e.g. http://iscn.fluxdata.org/Data/LTSEs/Pages/Map.aspx) (although these are not all soil C agricultural research sites). It is not clear whether this evidence is published as academic literature, or whether it is not presented in English or in a way that is accessible through the scientific databases used herein. We have linked our mapped evidence to the ISCN long-term experimental site database to increase the overlap between different meta-databases (Additional File 9).

A number of novel practices, such as applications of biochar, organic composting of urban waste or integration of biomass production, are not present in the evidence base and demonstrate that carbon sequestration research necessitates long-term studies, while management practices change over shorter time scales. The relative importance of SOC contribution through roots or harvest residues (or other types of biomass amendments) is another gap, as evidence generally concerns the effect of each of the management practices individually but without a quantitative comparison over longer periods [69].

In the review we have focused on SOC concentration as this is generally a common and methodologically well-defined parameter, while there is a gap of knowledge concerning C stocks versus flows of functions or, for example, CO_2 from the soil. Soil C concentration and total amount of soil C may be relevant from a C sequestration point of view but we are fully aware that for other aspects, such as soil functioning and plant growth, the organic matter of which SOC is a part is the carrier of the energy in the system [70]. Thus, we have mapped evidence of the C 'stocks' that are less critical to functioning of the ecosystems than maintaining 'flows' of functions. A stock and flow comparison can show that soil C sequestration is not always correlated to the flows of energy and functions in the system [71]. Evidence regarding stocks of sequestered C in relation to the flows of functions and energy, for example, increased GHG fluxes caused by raised SOC levels [72], have not been included in the map and therefore constitute another knowledge gap.

Rates of SOC change is another subject that could be developed further from data of long-term experiments, in order to undertake quantitative comparisons between different types of management on a large spatial scale. Here, there is a possibility to use global data from long-term experiments, such as those in the *Somnet* database (http://iscn.fluxdata.org/Data/LTSEs/Pages/Map.aspx), and calculate rates of change that can be further used in models of climate change or food production studies.

Knowledge gluts

Secondary researchers and funders can use our systematic map to identify 'knowledge gluts'; topics for which substantial evidence exists but that have not yet been synthesised via systematic review. As with knowledge gaps, not all knowledge gluts are worthy of the resource requirements needed for a full and reliable synthesis. Prioritisation may be necessary, for example based on the cost, controversy/interest, and applicability of the findings [73]. For example, 308 studies within our systematic map have investigated a narrow range of interventions relating to tillage. The adoption of no-till agriculture is not controversial, but the benefits of reduced tillage on C sequestration relative to conventional intensive tillage has been the subject of debate [17, 72]. A systematic review on the impacts of tillage on soil C would therefore be worthwhile since it is somewhat controversial, the adoption of reduced tillage may be costly in terms of short-term yield, and the findings of the review (particularly where heterogeneity was investigated) would be of direct use to stakeholders.

Other examples of viable systematic review questions that arise from our systematic map are:

1. What is the effect of untreated crop residue and green manure amendments on SOC?
2. What is the impact of multi-cropping (under-sowing/intercropping/catch cropping/cover cropping/double cropping) on SOC?
3. What is the effect of manure application on SOC (solid and liquid)?
4. What is the difference in impacts on SOC between crop rotations that include perennials versus annual cropping systems?
5. What is the impact of the inclusion of legumes in crop rotations on SOC?
6. How do 2- and 3- year crop rotations differ in SOC impacts relative to monocultures, and how does the timing of sowing affect soil organic carbon?
7. Which interventions are most influential on SOC as demonstrated by long-term time series data?

Encouraging researchers to submit their data so that the map is updated

Systematic reviews and systematic maps are powerful tools that can aid decision-making. However, the resource-intensive processes often take considerable time to ensure a high reliability in the end product. As a result, the review may be several years old when published. In fact, the review becomes essentially out-of-date as soon as searches are undertaken, although in practice this depends heavily on the rate of publication for the subject area. Whilst systematic reviews can and are updated to include research made available since their publication, these updates can be as resource intensive as the original reviews, particularly where some of the original work must be redone.

Separate from updating reviews, there has been considerable discussion of methods for creating and managing ‘living’ systematic reviews that can extend their period of relevance by incorporating new research as it is published [e.g. 74]. For example, a paper on *Drosophila* walking mechanisms published in the journal F1000 Research [75] employed the use of an updatable online figure that has been modified since publication to include more recent research [76]. There exist many significant challenges to establishing a living systematic review or systematic map. Firstly, technical difficulties must be dealt with to allow for figures and tables to be updated, along with any quantitative synthesis such as meta-analysis and associated forest plots. Most significant, however, is the maintenance of a systematic approach to continual updating. Reviewers go to great lengths to perform a comprehensive and systematic search using predefined search strings. Any iterative updating of a living review would either require continued manual checking or would be open to non-systematic inclusion of research.

However, we advocate an intermediate approach whereby members of the public can submit research to a living ‘unverified’ list of resources that could be used to highlight novel research that may be useful for systematic updates to the map. This unverified database, clearly identified as a list of potentially relevant studies, could then be used as a basis for checking that any systematic search update had included all relevant research. A second benefit relates to the ability to include a greater volume of grey literature that might otherwise be missed using a closed review system. As such, we have established a multi-language web-based submission system for members of the public to submit research to an unverified map database, which can be accessed via the MISTRA EviEM website: http://www.eviem.se/en/projects/Soil-organic-carbon-stocks/. We encourage readers to submit relevant research via this portal.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors participated in the drafting, revision and approval of the manuscript.

Acknowledgements

We are grateful for valuable comments and advice received from a considerable number of stakeholders, both at meetings in the initial phase of this project, during the public review of the protocol and during our work with the systematic map.

This review was financed by the Mistra Council for Evidence-Based Environmental Management (EviEM). EviEM is funded by the Swedish Foundation for Strategic Environmental Research (Mistra) and hosted by the Royal Swedish Academy of Sciences. The review process has been approved by the EviEM Executive Committee, but the authors are solely responsible for the contents and conclusions of the review.

References

1. Batjes N. **Total carbon and nitrogen in the soils of the world**. European Journal of Soil Science. 1996;47(2):151-63.

2. Batjes N, Dijkshoorn J. **Carbon and nitrogen stocks in the soils of the Amazon Region.** Geoderma. 1999;89(3):273-86.

3. Schlesinger WH. Biogeochemistry: **An Analysis of Global Change.** San Diego, California: Academic Press, Inc.; 1998. 443 p.

4. Betts RA, Falloon PD, Goldewijk KK, Ramankutty N. **Biogeophysical effects of land use on climate: Model simulations of radiative forcing and large-scale temperature change.** Agricultural and Forest Meteorology. 2007;142(2):216-33.

5. Freibauer A, Rounsevell MD, Smith P, Verhagen J. **Carbon sequestration in the agricultural soils of Europe.** Geoderma. 2004;122(1):1-23.

6. Lal R, Delgado J, Groffman P, Millar N, Dell C, Rotz A. **Management to mitigate and adapt to climate change**. Journal of Soil and Water Conservation. 2011;66(4):276-85.

7. González-Sánchez E, Ordóñez-Fernández R, Carbonell-Bojollo R, Veroz-González O, Gil-Ribes J. **Meta-analysis on atmospheric carbon capture in Spain through the use of conservation agriculture.** Soil and Tillage Research. 2012;122:52-60.

8. Paul EA, Paustian K, Elliott ET, C.V.Cole. **Soil Organic Matter in Temperate Agroecosystems: Long-Term Experiments in North America.** Boca Raton: CRC Press; 1997.

9. Luo Z, Wang E, Sun OJ. **Soil carbon change and its responses to agricultural practices in Australian agro-ecosystems: A review and synthesis.** Geoderma. 2010;155(3–4):211-23.

10. Pan G, Smith P, Pan W. **The role of soil organic matter in maintaining the productivity and yield stability of cereals in China.** Agriculture, Ecosystems and Environment. 2009;129(1–3):344-8.

11. Bolinder M, Kätterer T, Andrén O, Ericson L, Parent L-E, Kirchmann H. **Long-term soil organic carbon and nitrogen dynamics in forage-based crop rotations in Northern Sweden** (63–64 N). Agriculture, Ecosystems & Environment. 2010;138(3):335-42.

12. Poeplau C, Aronsson H, Myrbeck Å, Kätterer T. **Effect of perennial ryegrass cover crop on soil organic carbon stocks in southern Sweden.** Geoderma Regional. 2015;4:126-33.

13. Diacono M, Montemurro F. **Long-term effects of organic amendments on soil fertility. Sustainable Agriculture Volume 2**: Springer; 2011. p. 761-86.

14. Poeplau C, Kätterer T, Bolinder MA, Börjesson G, Berti A, Lugato E. **Low stabilization of aboveground crop residue carbon in sandy soils of Swedish long-term experiments.** Geoderma. 2015;237:246-55.

15. Blair N, Faulkner R, Till A, Korschens M, Schulz E. **Long-term management impacts on soil C, N and physical fertility: Part II: Bad Lauchstadt static and extreme FYM experiments.** Soil and Tillage Research. 2006;91(1):39-47.

16. Khan S, Mulvaney R, Ellsworth T, Boast C. **The myth of nitrogen fertilization for soil carbon sequestration.** Journal of Environmental Quality. 2007;36(6):1821-32.

17. Baker JM, Ochsner TE, Venterea RT, Griffis TJ. **Tillage and soil carbon sequestration—What do we really know?** Agriculture, Ecosystems & Environment. 2007;118(1):1-5.

18. Lal R, Follett R. Soils and climate change. In: R L, F. F, editors. **Soil carbon sequestration and the greenhouse effect.** Madison, Wisconsin: SSSA Special Publication; 2009.

19. Hati KM, Swarup A, Dwivedi A, Misra A, Bandyopadhyay K. **Changes in soil physical properties and organic carbon status at the topsoil horizon of a vertisol of central India after 28 years of continuous cropping, fertilization and manuring.** Agriculture, Ecosystems & Environment. 2007;119(1):127-34.

20.Yang X, Li P, Zhang S, Sun B, Xinping C. **Long-term-fertilization effects on soil organic carbon, physical properties, and wheat yield of a loess soil.** Journal of Plant Nutrition and Soil Science. 2011;174(5):775-84.

21.Kätterer T, Bolinder M, Berglund K, Kirchmann H. **Strategies for carbon sequestration in agricultural soils in northern Europe.** Acta Agriculturae Scandinavica, Section A–Animal Science. 2012;62(4):181-98.

22. Adler PR, Grosso SJD, Parton WJ. **Life-cycle assessment of net greenhouse-gas flux for bioenergy cropping systems.** Ecological Applications. 2007;17(3):675-91.

23. Johnston AE, Poulton PR, Coleman K. **Soil organic matter: its importance in sustainable agriculture and carbon dioxide fluxes.** Advances in agronomy. 2009;101:1-57.

24. Van Oost K, Quine T, Govers G, De Gryze S, Six J, Harden J, et al. **The impact of agricultural soil erosion on the global carbon cycle.** Science. 2007;318(5850):626-9.

25. Tsiafouli MA, Thébault E, Sgardelis SP, Ruiter PC, Putten WH, Birkhofer K, et al. **Intensive agriculture reduces soil biodiversity across Europe**. Global change biology. 2015;21(2):973-85.

26. Ludwig B, Geisseler D, Michel K, Joergensen R, Schulz E, Merbach I, et al. **Effects of fertilization and soil management on crop yields and carbon stabilization in soils. A review.** Agronomy for Sustainable Development. 2010.

27. Rasmussen PE, Goulding KW, Brown JR, Grace PR, Janzen HH, Körschens M. **Long-term agroecosystem experiments: Assessing agricultural sustainability and global change.** Science. 1998;282(5390):893-6.

28. Guo LB, Gifford R. **Soil carbon stocks and land use change: a meta analysis**. Global change biology. 2002;8(4):345-60.

29. Lal R. **Carbon sequestration**. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2008;363(1492):815-30.

30. Govaerts B, Verhulst N, Castellanos-Navarrete A, Sayre K, Dixon J, Dendooven L. **Conservation agriculture and soil carbon sequestration: between myth and farmer reality.** Critical Reviews in Plant Science. 2009;28(3):97-122.

31. SOILSERVICE. **Conflicting demands of land use, soil biodiversity and the sustainable delivery of ecosystem goods and services in Europe.** 2012.

32. Collaboration for Environmental Evidence. **Guidelines for Systematic Review and Evidence Synthesis in Environmental Management.** Version 4.2: Collaboration for Environmental Evidence: www.environmentalevidence.org/Documents/Guidelines/Guidelines4.2.pdf; 2013.

33. Randall NP, James KL. **The effectiveness of integrated farm management, organic farming and agri-environment schemes for conserving biodiversity in temperate Europe-A systematic map.** Environmental Evidence. 2012;1(4).

34. Bates S, Clapton J, Coren E. **Systematic maps to support the evidence base in social care.** Evidence & Policy: A Journal of Research, Debate and Practice. 2007;3(4):539-51.

35. Grant MJ, Booth A. **A typology of reviews: an analysis of 14 review types and associated methodologies.** Health Information & Libraries Journal. 2009;26(2):91-108.

36. Clapton J, Rutter D, Sharif N. **SCIE Systematic mapping guidance**. London: SCIE. 2009.

37. Söderström B, Hedlund K, Jackson LE, Kätterer T, Lugato E, Thomsen IK, et al. **What are the effects of agricultural management on soil organic carbon (SOC) stocks?** Environmental Evidence. 2014;3(2).

38. Cohen J. **A coefficient of agreement for nominal scales.** Educational and psychological measurement. 1960;20(1):37-46.

39. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. **World map of the Köppen-Geiger climate classification updated.** Meteorologische Zeitschrift. 2006;15(3):259-63.

40. Smith P. **How long before a change in soil organic carbon can be detected?** Global Change Biology. 2004;10(11):1878-83.

41. Haddaway NR, Burden A, Evans CD, Healey JR, Jones DL, Dalrymple SE, et al. **Evaluating effects of land management on greenhouse gas fluxes and carbon balances in boreo-temperate lowland peatland systems.** Environmental Evidence. 2014;3(1):5.

42. Pullin AS, Bangpan M, Dalrymple S, Dickson K, Haddaway NR, Healey JR, et al. **Human well-being impacts of terrestrial protected areas.** Environmental Evidence. 2013;2(1):19.

43. Our Common Future Under Climate Change. **SOIL CARBON SEQUESTRATION: TOWARDS AN INTERNATIONAL RESEARCH PROGRAM AND ACTION PLAN** 2014 [cited 2015 7th September]. Available from: http://www.commonfuture-paris2015.org/Programme/Side-Events/170-Soil-carbon-sequestration.htm.

44. Lal R. **Soil carbon sequestration impacts on global climate change and food security.** Science. 2004;304(5677):1623-7.

45. Jeffery S, Verheijen F, Van Der Velde M, Bastos A. **A quantitative review of the effects of biochar application to soils on crop productivity using meta-analysis.** Agriculture, ecosystems & environment. 2011;144(1):175-87.

46. Glover J, Reganold J, Bell L, Borevitz J, Brummer E, Buckler E, et al. **Increasing food and ecosystem security through perennial grain breeding.** Science. 2010;328(5986).

47. Lal R. **Managing soils and ecosystems for mitigating anthropogenic carbon emissions and advancing global food security.** BioScience. 2010;60(9):708-21.

48. Barrios E. **Soil biota, ecosystem services and land productivity.** Ecological economics. 2007;64(2):269-85.

49. Antle JM, Stoorvogel JJ. **Payments for ecosystem services, poverty and sustainability: The case of agricultural soil carbon sequestration.** Payment for Environmental Services in Agricultural Landscapes: Springer; 2009. p. 133-61.

50. Jenkinson D. **The Rothamsted long-term experiments: are they still of use?** Agronomy Journal. 1991;83(1):2-10.

51. Mitchell C, Westerman R, Brown J, Peck T. **Overview of long-term agronomic research.** Agronomy Journal. 1991;83(1):24-9.

52. Christensen BT, Petersen J, Trentemøller U. **The Askov Long-Term Experiments on Animal Manure and Mineral Fertilizers: The Lermarken site 1894-2004.** DIAS report Plant production no. 121, 2006.

53. Smith P, Smith J, Powlson D, McGill W, Arah J, Chertov O, et al. **A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments.** Geoderma. 1997;81(1):153-225.

54. Conant RT, Smith GR, Paustian K. **Spatial variability of soil carbon in forested and cultivated sites.** Journal of Environmental Quality. 2003;32(1):278-86.

55. Lück E, Ruehlmann J, Kirchmann H. **Properties of soils from the Swedish long-term fertility experiments: VI. Mapping soil electrical conductivity with different geophysical methods.** Acta Agriculturae Scandinavica, Section B-Soil & Plant Science. 2011;61(5):438-47.

56. Janzen H. **The role of long-term sites in agroecological research: A case study.** Canadian Journal of Soil Science. 1995;75(1):123-33.

57. Carter M. **Relative measures of soil bulk density to characterize compaction in tillage studies on fine sandy loams.** Canadian Journal of Soil Science. 1990;70(3):425-33.

58. Ellert B, Janzen H, Entz T. **Assessment of a method to measure temporal change in soil carbon storage.** Soil Science Society of America Journal. 2002;66(5):1687-95.

59. Wuest SB. **Correction of bulk density and sampling method biases using soil mass per unit area.** Soil Science Society of America Journal. 2009;73(1):312-6.

60. Lee J, Hopmans JW, Rolston DE, Baer SG, Six J. **Determining soil carbon stock changes: simple bulk density corrections fail.** Agriculture, Ecosystems & Environment. 2009;134(3):251-6.

61. Wendt J, Hauser S. **An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers.** European Journal of Soil Science. 2013;64(1):58-65.

62. US Department of Agriculture. **Summary Report: 2007 National Resources Inventory.** Natural Resources Conservation Service, Washington, DC, and Center for Survey Statistics and Methodology, Iowa State University, Ames, Iowa; 2009. p. 123.

63. Paul EA, Paustian KH, Elliott E, Cole CV. **Soil Organic Matter in Temperate Agroecosystems: Long Term Experiments in North America**. CRC Press; 1997.

64. Pittelkow CM, Liang X, Linquist BA, Van Groenigen KJ, Lee J, Lundy ME, et al. **Productivity limits and potentials of the principles of conservation agriculture.** Nature. 2015;517(7534):365-8.

65. Smith P, Goulding KW, Smith KA, Powlson DS, Smith JU, Falloon P, et al. **Enhancing the carbon sink in European agricultural soils: including trace gas fluxes in estimates of carbon mitigation potential.** Nutrient Cycling in Agroecosystems. 2001;60(1-3):237-52.

66. VandenBygaart A, Bremer E, McConkey B, Ellert B, Janzen H, Angers D, et al. **Impact of sampling depth on differences in soil carbon stocks in long-term agroecosystem experiments.** Soil Science Society of America Journal. 2011;75(1):226-34.

67. Haddaway NR, Styles D, Pullin AS. **Evidence on the environmental impacts of farm land abandonment in high altitude/mountain regions: a systematic map.** Environmental Evidence. 2014;3(1):17.

68. Shevtsova L, Romanenkov V, Sirotenko O, Kanzyvaa S. **Transformation study in arable soils based on long-term experiments in russia: Historical experience and international co-OPERATION.** Archives of Agronomy and Soil Science. 2003;49(5):485-502.

69. Kätterer T, Bolinder MA, Andrén O, Kirchmann H, Menichetti L. **Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment.** Agriculture, Ecosystems & Environment. 2011;141(1):184-92.

70. Janzen H. **Beyond carbon sequestration: soil as conduit of solar energy.** European Journal of Soil Science. 2015;66(1):19-32.

71. Olson KR, Al-Kaisi M, Lal R, Lowery B. **Examining the paired comparison method approach for determining soil organic carbon sequestration rates.** Journal of Soil and Water Conservation. 2014;69(6):193A-7A.

72. Powlson DS, Stirling CM, Jat M, Gerard BG, Palm CA, Sanchez PA, et al. **Limited potential of no-till agriculture for climate change mitigation.** Nature Climate Change. 2014;4(8):678-83.

73. Dicks LV, Walsh JC, Sutherland WJ. **Organising evidence for environmental management decisions: a '4S'hierarchy.** Trends in ecology & evolution. 2014;29(11):607-13.

74. Badgett RG, Vindhyal M, Stirnaman JT, Gibson CM, Halaby R. **A Living Systematic Review of Nebulized Hypertonic Saline for Acute Bronchiolitis in Infants.** JAMA pediatrics. 2015;169(8):788-9.

75. Colomb J, Brembs B. **Sub-strains of Drosophila Canton-S differ markedly in their locomotor behavior**. F1000Research. 2014;3.

76. Singh CD. **'Living figures' make their debut.** Nature. 2015;521(7550):112.

List of additional files

Additional File 1. Search String Development.

Additional File 2. Academic Database Search Results.

Additional File 3. Internet Search Engine Search Results.

Additional File 4. Specialist Websites Search Results.

Additional File 5. Database of Reviews and Meta-analyses.

Additional File 6. Articles Unobtainable as Full Texts.

Additional File 7. Systematic Map Database.

Additional File 8. Articles Excluded at Full Text.

Additional File 9. Long-Term Experimental (LTE) Sites Database.

EviEM Scientific Report 2015

How does agricultural management affect the amount of carbon stored in soils? Some farming practices, such as reduced tillage, can improve carbon storage, potentially helping to reduce global warming. In the systematic map that is presented here MISTRA EviEM has catalogued and described relevant research from areas with a temperate or boreal climate.

www.eviem.se

EviEM conducts systematic reviews of environmental issues identified as important by public agencies and other stakeholders. These provide an overall assessment of the state of scientific knowledge and help to improve the basis for environmental decision-making in Sweden.

Mistra EviEM
Royal Swedish Academy of Sciences
Box 50005, SE-104 05 Stockholm, Sweden
Visit/Deliveries
Lilla Frescativägen 4A, SE-114 18 Stockholm
Telephone + 46 8-673 95 00
E-mail info@eviem.se