Haddaway *et al. Environ Evid (2017) 6:30*
DOI 10.1186/s13750-017-0108-9

Environmental Evidence

SYSTEMATIC REVIEW

Open Access

How does tillage intensity affect soil organic carbon? A systematic review

Neal R. Haddaway[1*], Katarina Hedlund[2], Louise E. Jackson[3], Thomas Kätterer[4], Emanuele Lugato[5], Ingrid K. Thomsen[6], Helene B. Jørgensen[2] and Per-Erik Isberg[7]

Abstract

Background: The loss of carbon (C) from agricultural soils has been, in part, attributed to tillage, a common practice providing a number of benefits to farmers. The promotion of less intensive tillage practices and no tillage (NT) (the absence of mechanical soil disturbance) aims to mitigate negative impacts on soil quality and to preserve soil organic carbon (SOC). Several reviews and meta-analyses have shown both beneficial and null effects on SOC due to no tillage relative to conventional tillage, hence there is a need for a comprehensive systematic review to answer the question: what is the impact of reduced tillage intensity on SOC?

Methods: We systematically reviewed relevant research in boreo-temperate regions using, as a basis, evidence identified within a recently completed systematic map on the impacts of farming on SOC. We performed an update of the original searches to include studies published since the map search. We screened all evidence for relevance according to predetermined inclusion criteria. Studies were appraised and subject to data extraction. Meta-analyses were performed to investigate the impact of reducing tillage [from high (HT) to intermediate intensity (IT), HT to NT, and from IT to NT] for SOC concentration and SOC stock in the upper soil and at lower depths.

Results: A total of 351 studies were included in the systematic review: 18% from an update of research published in the 2 years since the systematic map. SOC concentration was significantly higher in NT relative to both IT [1.18 g/kg ± 0.34 (SE)] and HT [2.09 g/kg ± 0.34 (SE)] in the upper soil layer (0–15 cm). IT was also found to be significant higher [1.30 g/kg ± 0.22 (SE)] in SOC concentration than HT for the upper soil layer (0–15 cm). At lower depths, only IT SOC compared with HT at 15–30 cm showed a significant difference; being 0.89 g/kg [± 0.20 (SE)] lower in intermediate intensity tillage. For stock data NT had significantly higher SOC stocks down to 30 cm than either HT [4.61 Mg/ha ± 1.95 (SE)] or IT [3.85 Mg/ha ± 1.64 (SE)]. No other comparisons were significant.

Conclusions: The transition of tilled croplands to NT and conservation tillage has been credited with substantial potential to mitigate climate change via C storage. Based on our results, C stock increase under NT compared to HT was in the upper soil (0–30 cm) around 4.6 Mg/ha (0.78–8.43 Mg/ha, 95% CI) over ≥ 10 years, while no effect was detected in the full soil profile. The results support those from several previous studies and reviews that NT and IT increase SOC in the topsoil. Higher SOC stocks or concentrations in the upper soil not only promote a more productive soil with higher biological activity but also provide resilience to extreme weather conditions. The effect of tillage practices on total SOC stocks will be further evaluated in a forthcoming project accounting for soil bulk densities and crop yields. Our findings can hopefully be used to guide policies for sustainable management of agricultural soils.

Keywords: Agriculture, Conservation, Till, Plough, Farming, Land management, Climate change, Land use change, Carbon sequestration

*Correspondence: neal.haddaway@eviem.se; neal_haddaway@hotmail.com
[1] Mistra Council for Evidence-Based Environmental Management (EviEM), Stockholm Environment Institute, Box 24218, 10451 Stockholm, Sweden
Full list of author information is available at the end of the article

Background

Soils contain the largest terrestrial carbon (C) pool that is sensitive to changes in land use and agricultural management practices. Indeed, soils could provide a vital ecosystem service by acting as a C sink, potentially mitigating climate change [1–3]. Consequently, changes in soil C could affect atmospheric CO_2 concentration. Approximately 12% of soil C is held in cultivated soils [4], which cover around 35% of the terrestrial land area of the planet [5].

Arable soils are under considerable threat due to unsustainable cultivation practices. It has been estimated that US soils may have lost between 30 and 50% of the SOC that they contained prior to the establishment of agriculture there [6]. This has been attributed to loss of C from agricultural soils due to the advent of the plough [e.g. 7], indicating that agricultural soils may have a potential to mitigate climate change through C sequestration [8, 9]. Besides climate change, SOC has a number of potential associated benefits, including: increased soil fertility [10, 11]; improved biological and physical soil characteristics [12] via a reduction in bulk density, improved water-holding capacity and enhanced activity of soil microbes [13] (although this may increase CO_2 emission). Promoting SOC also often increases soil biodiversity and ecosystem functions that can enhance agricultural productivity by mediating nutrient cycling, soil structure formation, and crop resistance to pests and diseases [14].

Historically, tillage has been performed because of a number of benefits associated with the practice. These benefits include: loosening and aeration of topsoil, facilitating planting and seedbed preparation; mixing of crop residues into the soil; mechanical destruction of weeds; drying wetter soils prior to seeding; allowing frost-induced disturbance of the soil when undertaken prior to winter.

However, conventional tillage may increase compaction of soil below the depth of tillage (i.e., formation of a plough pan), the susceptibility to water and wind erosion and the energy costs for the mechanical operations [15]. In recent years, the promotion of less intensive tillage practices (also referred to as conservation tillage or reduced tillage) and no tillage (NT) (the absence of mechanical soil disturbance) agricultural management has sought to mitigate some of these negative impacts on soil quality and to preserve SOC. These practices aim at maintaining organic matter on the surface or in the upper soil layer thereby increasing SOC concentration especially in the topsoil [16, 17]. A reduction in the need for mechanical tillage practices reduces energy consumption and C emissions through the use of fossil fuels [18], whilst also reducing labour requirements [19], but this benefit may be outweighed to a certain extent by the increased requirements for pesticides, especially herbicides. Furthermore, reduction of tillage activities has been associated with a loss of yield by a number of authors [20]; in one case, 8.5% lower yield for NT relative to conventional tillage [21]. Moreover, higher N_2O emissions can occur with reduced or NT, due to moister and denser soil conditions, which may eventually offset positive effects on SOC balances [22, 23].

Alvarez [24] recognised the need for a broad synthetic approach to assess the impact of agricultural management. As such, a number of authors have reviewed the impact of tillage on soil C [e.g. 8, 17, 24–28]. These reviews and meta-analyses have shown both beneficial [8, 17] and null [29, 30] effects on SOC due to NT relative to conventional tillage. Furthermore, the efficacy of reduced tillage relative to NT is also unclear [24, 26]. Discrepancies may depend on whether total SOC stocks are measured or only presented as the SOC concentration, and also whether they are measured only in the upper soil layers or are reported accounting for the full soil profile [31]. Whilst some advantages of conservation tillage are clear (e.g. reduced erosion and reduced fuel consumption), other impacts (e.g. N_2O emission, crop yield, SOC sequestration) can be variable [31]. What seems to be decisive for the direction of SOC changes is the effect of tillage on net primary production (NPP). If NPP increases due to certain tillage practices, SOC stocks are more likely to increase and vice versa [32]. The purpose of this systematic review is to identify the state-of-the-art results regarding the so far inconclusive effects of tillage on SOC in a comprehensive, transparent and objective manner.

Identification of the topic

The subject of tillage was originally identified and included in the previously published systematic map [33] following in depth discussion with Swedish stakeholders, including the Swedish Board of Agriculture. Following completion of the systematic map, tillage was identified as a candidate topic for full systematic review based on a number of key criteria: the presence of sufficient reliable evidence, the relevance of the topic for stakeholders, the applicability of the topic for the Swedish environment, the benefit of a systematic approach to a topic that has received some attention via traditional reviews, and the added value of investigating effect modifiers and sources of heterogeneity across studies via a large meta-analysis. The topic was proposed and accepted during a meeting of the authors in May 2015.

Objective of the review

We hypothesise that reduced or NT will mitigate losses of soil carbon as compared to more intensive ploughing

[16, 17]. However, reduced tillage is assumed to have effects on SOC in the surface of the soil but not always through deeper soil layers [31]. Hence, we also test effects of reduced tillage from experiments with measurements in the upper 15 cm and deeper in the soil profile.

The effects of tillage on SOC have previously been reviewed [e.g. 8, 17, 24–28, 34] but as yet none of these reviews has been systematic in nature. The objective of this review is to systematically review and synthesise existing research pertinent to tillage practices in warm temperate and boreal regions (see *Relevant subject* below for details) using, as a basis, the evidence identified within a recently completed systematic map [35, 36]. This systematic map aimed to collate evidence relating to the impacts of all agricultural management on soil organic carbon in boreo-temperate regions.

Primary Question: What is the effect of tillage intensity on soil organic carbon (SOC)?

Secondary Question: How do other factors interact with tillage to affect SOC?

Subject: Arable soils in agricultural regions from the warm temperate climate zone (fully humid and summer dry, i.e., Köppen–Geiger climate classification; Cfa, Cfb, Cfc, Csa, Csb, Csc) and the snow climate zone (fully humid, i.e., Köppen–Geiger climate classification; Dfa, Dfb, Dfc).

Interventions: Any described reduced tillage practice (including NT, reduced tillage, rotational tillage, conventional tillage).

Comparators: More intensive tillage practice (including the above tillage practices along with subsoiling). Also before/after comparisons for single tillage treatments.

Outcomes: SOC (measured as either concentration or stock).

Methods

This systematic review was conducted in accordance with a CEE systematic review protocol [37].

Searches

Original systematic map search

Searches of 17 academic databases were undertaken as part of the published systematic map between the 16th and 19th September 2013 [see 33]. This search was broader than just tillage, including also interventions relating to amendments, fertilisers and crop rotations (some 750 studies in total). These academic database searches were supplemented by searches for grey literature via web search engines and organisational websites, and by searches of the bibliographies of 127 relevant reviews and meta-analyses identified during the course of the systematic map. Full details for all searches can be found in Additional files accompanying the systematic map described in Haddaway et al. [37].

Search update

A search update was undertaken in September 2015 to capture research published since the original search in September 2013. The update was restricted to four academic databases, Academic Search Premier, Pub Med, Scopus, Web of Science (Web of Science Core Collection, BIOSIS Citation Index, Chinese Science Citation Database, Data Citation Index, SciELO Citation Index), and one academic search engine, Google Scholar, which has been shown to be effective at identifying both academic and grey literature [38]. The choice to reduce the number of citation databases was driven by observations made during the undertaking of the systematic map, where a large number of duplicates was identified in many of the databases used. Only English language search terms was used for the update, but any articles identified in Danish, English, French, German, Italian, and Swedish were included.

Search strategy

The following search string was used in the academic databases mentioned above to search on 'topic words' (i.e. titles, abstracts and keywords). This search string has been adapted from the original string used in the published systematic map [36] to identify specifically tillage research and restricted to the period since the original search was undertaken (September 2013):

> *soil* AND (arable OR agricult* OR farm* OR crop* OR cultivat*) AND (till* OR "no till*" OR "reduced till*" OR "direct drill*" OR "conservation till*" OR "minimum till*") 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*")*

[the underlined text indicates modifications to the original systematic map search string]

In Google Scholar the following search string was used and the first 1000 records for full text searches and all 163 title searches were downloaded:

> *soil AND carbon AND (till OR tillage OR "reduced tillage" OR "conservation tillage" OR "no tillage" OR "direct drill" OR "minimum till*")*

Searches were restricted to 2013–2015 and downloaded using web crawling software [38, 39].

Additional bibliographic checking

One review was identified through screening of search results from the search update [40]. The bibliography of this review article was screened for potentially relevant articles that may have been missed by the searches. Six additional articles were sourced from this checking and all articles screened at full text and excluded are listed in Additional file 1.

Study inclusion criteria

A total of 311 studies were already identified as part of the recent systematic map [33]. These studies were originally assessed according to predefined inclusion criteria [see 36] as part of the systematic map. These original inclusion criteria were modified for the purposes of this systematic review by the inclusion of a requirement for studies to have investigated tillage interventions. The inclusion criteria used to screen all studies (including the original 311 studies and the updated search results) were as follows:

Relevant subject:	Arable soils in agricultural regions from the warm temperate climate zone (fully humid and summer dry, i.e., Köppen–Geiger climate classification; Cfa, Cfb, Cfc, Csa, Csb, Csc) and the snow climate zone (fully humid, i.e., Köppen–Geiger climate classification; Dfa, Dfb, Dfc). These zones were selected due to their relative homogeneity and relevance to the Swedish environment. Studies involving agroforestry, paddy or rice cropping systems were excluded.
Relevant interventions:	All tillage practices identified iteratively within the evidence base. Such practices include: NT (also described as direct drill); reduced, minimum or conservation tillage (i.e. chisel plough, disc plough, harrow, mulch plough, ridge till); rotational tillage (i.e. non-annual, regular tillage); conventional tillage (i.e. mouldboard plough); subsoiling. We appreciate that some tillage practices classified above as reduced tillage may be intensive, and all described tillage practices will be assessed on an individual basis before classifying them broadly as *NT*, *intermediate intensity tillage* (IT) (any non-inversion tillage performed above 40 cm depth), and *high intensity tillage* (HT) (any inversion tillage or non-inversion tillage performed to 40 cm or below).
Relevant comparators:	Any comparison between different intensities of tillage from NT to intensive tillage. Additionally, studies will be included that make comparisons of single interventions from before relative to after the intervention.
Relevant outcomes:	Soil C measures, including: soil organic carbon (SOC), total organic carbon (TOC), total carbon (TC) (where soils are shown to be free of carbonates), and soil organic matter (SOM). This may be expressed either as a concentration (e.g. g/kg or %) or as a stock (e.g. Mg/ha).
Relevant study types:	Field studies examining interventions that have lasted at least 10 years to ensure that changes in soil C are detectable [41].

Only research written in Danish, English, French, German, Italian, Norwegian, and Swedish were included in the review. Potentially relevant research identified in other languages was reported in Additional file. Every study identified via the update was screened through three stages: title, abstract and full text. At each level, records containing or likely to contain relevant information were retained and taken to the next stage. Where information was lacking (for example where abstracts are missing), the record was retained in order to be conservative. Following abstract screening full texts were retrieved and those that could not be obtained were documented as such (see Additional file 1: Bibliographic database

search record.xlsx, Additional file 2: Unobtainable articles.xlsx, Additional file 3). Screening was performed by one reviewer (NRH), immediately following screening of full texts for the systematic map [33]. A Kappa tests [42] for consistency checking were performed to assess the level of agreement amongst members of the review team (NRH, KH and HBJ), indicating high agreement at abstract (kappa = 0.75) and full text (kappa = 0.72) using a subset of 198 and 120 records at each level, respectively.

Potential effect modifiers and reasons for heterogeneity

All studies included in this review were subject to extraction of meta-data (see Data Extraction, below), which included the extraction of data regarding key sources of heterogeneity, namely: climate zone, latitude, longitude, and soil type (classification or texture). These potential modifiers were used in meta-analyses to account for significant differences between studies, as described below in synthesis. All studies used in this review were long-term agricultural sites, and so the impacts of interventions were investigated in relation to implementation of alternative agricultural practices on similar land-use types.

Critical appraisal of study validity

Critical appraisal undertaken in the completed systematic map

The completed systematic map undertook critical appraisal of the included studies for the purposes of excluding unreliable studies that were highly susceptible to bias (such as those lacking details on methods, or those with no replication) or non-generalisable and to assess the reliability of the evidence base. Reasons for exclusion were transparently recorded for all studies [see additional information in 33]. In addition to excluding studies that were highly susceptible to bias, five domains were assessed for study reliability for those studies passing the initial assessment: spatial replication (number of spatial replicates); temporal replication (number of time samples); treatment allocation (e.g. randomised, blocked, purposeful); study duration (length of the experimental period); soil sampling depth (the number and extent of soil depth samples taken). For each of these domains, studies were awarded a 0, 1, or 2 for the degree of reliability as described in Table 1. Where insufficient information was reported a '?' was awarded. See Haddaway et al. [33] for full details of the methods used and results from the systematic map.

For the purposes of critically appraising studies in this systematic review, two of the domains described above (spatial replication and treatment allocation) were summed and scores of 3 or 4 (maximum of 4) were given an appraisal category of 'high' validity, whilst those of 2 or below were assigned a 'low' validity category. Temporal replication was excluded from the final critical appraisal categorisation, since the majority of studies were single time point studies. Duration of the experiment and sampling depth were excluded because they will be accounted for during statistical modelling within meta-analyses. Where any of the original 5 domains assessed in the systematic map had been awarded a '?', indicating a lack of information, these studies were assigned a category of 'unclear'. Following critical appraisal, 3 studies were excluded on account of unacceptable susceptibility to bias (see Additional file 3).

Table 1 Critical appraisal criteria for five domains used in the systematic map by Haddaway et al. [33]

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 (years)	10–19	0
	20–29	1
	≥ 30	2
Soil sampling depth	Shallow (maximum depth ≤ 15 cm) single or multiple sampling	0
	Plough layer (maximum depth 15–25 cm) single or multiple sampling, or deep (maximum depth > 25 cm) single sampling	1
	Multiple deep sampling (maximum depth > 25 cm)	2

Italics text indicates the domains used for critical appraisal in the full systematic review described herein

Data extraction strategy

Meta-data were extracted for all studies. This information included the following: citation; study location (country, site, climate zone, latitude and longitude); soil type (classification or percent clay/silt/sand); study description (start year, duration, treatments investigated, cropping system, experimental design); sampling strategy (spatial and temporal replication, subsampling, soil sampling depth, C measurement method). In addition, quantitative data (i.e. study findings) were described (outcome type, units, data location, measure of variability, presence of bulk density) and extracted. Tillage categories for further synthesis were assessed as belonging to one of the following three categories: NT, IT and HT. As discussed above, IT corresponds to methods that do not invert the soil profile and that are performed above 40 cm depth (e.g. disk and chisel tillage). HT corresponds to methods that invert the soil profile (e.g. mouldboard plough and ridge tillage), along with very deep non-inversion tillage performed to 40 cm depth or below (i.e. very deep chisel tillage or subsoiling). This assessment was undertaken by extracting all interventions in the evidence base (machinery, tillage depth and timing) and building a coding tool iteratively. Where information was insufficient to readily allow coding, information gaps were filled using meta-data from other articles based at the same experimental site or using consensus during a meeting of the review team. Where consensus could not be reached, studies were excluded for a lack of information regarding the intervention (see Additional file 3). This coding tool is described in Table 2. Tillage machinery and depth were also extracted, and depth was categorised as shallow ($\leq$ 15 cm) or deep (> 15 cm).

Data synthesis and presentation

Effect size calculation

All quantitative data (i.e. study results) were extracted from each study as separate spreadsheets (see Additional file 4). Data were pooled across non-target treatments and exposures (such as slope position) using an a priori protocol (see Additional file 5). Data were analysed separately as concentrations and stocks (see Synthesis, below). Where studies reported bulk density and stocks, data were back-transformed into concentration data. Where concentrations were reported with bulk densities that were separated by depth and by treatment, data were converted into stocks using the equation in Additional file 5 (see http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4138211/) and included in both concentration and stocks meta-analyses (n = 55 studies, see systematic map database in Additional file 6).

Effect sizes All studies reported data in comparable units, and as a result, raw mean difference (RMD) was used as the effect size for all studies, preserving original units (g/kg and Mg/ha) and facilitating understanding of meta-analysis outputs. Data were grouped into three paired comparisons: no till versus HT, no till versus IT, and intermediate tillage versus HT. In each case, effect sizes were calculated as the less intensive intervention SOC value minus the more intensive intervention SOC value. Thus, a positive effect size indicates a greater SOC value in the more conservative tillage intervention (i.e. tillage reduction).

All effect sizes were initially calculated by one reviewer, with double-checking of calculations and all extracted data by the same reviewer and subsequently by a second reviewer.

Measures of variability Standard deviations were pooled across treatments, after coefficients of variation, standard errors and confidence intervals were converted to standard deviations where necessary. Studies that reported overall measures of variability (i.e. standard deviations, standard errors, coefficients of variation, confidence intervals) were converted to overall standard deviations and identified as estimated measures of treatment variability (since they do not precisely reflect variability within each treatment). These estimated variability measures were used in sensitivity analysis to examine the importance of accuracy in variability measures during meta-analysis. The following measures were also converted to overall standard deviations: least square difference, p values, and F-statistics. Additional files 4 and 5

Table 2 Coding tool for tillage intervention categories

Tillage category	Description
No tillage	No tillage activities undertaken. Also described as direct drill. Some machine activity is present where crops are planted or seeds broadcast, or where fertilisers and pesticides are applied. Light scarification of the soil surface may also occur
Intermediate intensity tillage	Soil is disturbed to a maximum depth not exceeding 40 cm and using only non-inversion tillage machinery; i.e. the soil profile is not inverted by means of turning apparatus such as a mouldboard. Examples include; chisel plough, disking, heavy duty cultivator, rotavator, rotary harrow, grubber, and field cultivator
High intensity tillage	Full inversion tillage occurs where the soil profile is inverted by means of turning apparatus such as a mouldboard. Non-inversion tillage to depths of 40 cm or more is also included. Ridge tillage, where soil ridges are built up using sets of opposing mouldboards and planted into, is also included

transparently document all processes involved with calculation of effect sizes and measures of variability.

Soil depth profiles Since studies reported soil depth across a variety of different depth layer thicknesses, soil profiles were split into two or three separate layers for independent analysis for stocks and concentration data respectively (see Synthesis, below).

For concentration data, these layers were defined as: 0–15, 15–30, and > 30 cm. In this way, study data were aggregated where provided in smaller increments by calculating a weighted mean for concentration. Where data overlapped one of the above soil layer boundaries (i.e. 15 or 30 cm), data were included in the layer above if the overlapping layer thickness was no more than 5 cm deeper than the specified layer. Similarly, data were included in the lower layer if the overlap was 3 cm or less. This distinction was made in order to remain conservative when separating data into three layers, since SOC concentration differences between tillage treatments are likely to be more pronounced at shallower depths (therefore, including data in layers above that which it belongs to decreases the chance of finding a significant difference). This process is shown in Fig. 1.

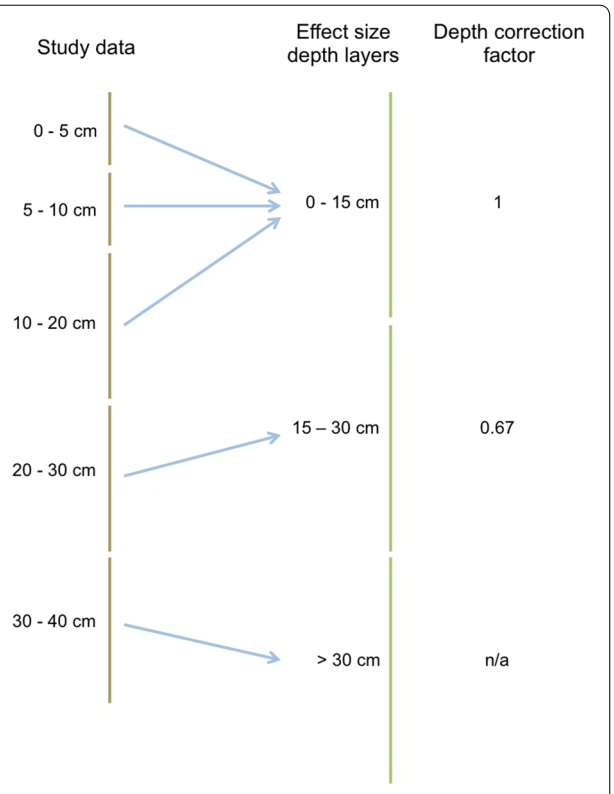

Fig. 1 Soil depth profiles and depth correction factors used in effect size calculation and meta-analysis

All studies reporting SOC concentrations were given a depth correction factor for data belonging to each of the three depth layers that was used in meta-analysis to weight data that came from incomplete soil layers. This number was calculated as the fraction of the profile covered by the data (e.g. a value of 0.67 for 0–10 cm depth). Where data overlapped one full layer a maximum value of 1 was calculated. No depth correction factors were calculated for the > 30 cm depth layer, however. This correction was avoided for > 30 cm depths since there was no lower boundary for this layer relevant to all studies, making a weighting disproportionate across studies, and since the correlation between SOC concentration and depth below this point was deemed to be inconsequential.

For stocks data, these layers were: upper layer (0–30 cm) and full profile (0–150 cm). These layers were chosen since it was felt that there was likely to be a significant difference in the impact on SOC stocks based on activities in the upper 30 cm that would manifest differently in the full profile (the maximum measures depth was 150 cm). For each of these two layers the full carbon content of the soil was calculated down to the maximum depth. Studies were either classed as reporting upper or lower maximum depths.

Other calculations Soil USDA texture classifications [43] were calculated for studies reporting clay, silt and sand percentages, and all comparable USDA soil texture data was used to describe soil texture in meta-analyses.

Narrative synthesis

An update of the systematic map containing only tillage studies was produced and included as an additional file, along with a dedicated geographical information system (GIS) (see Additional file 6). All studies in the evidence base were also included in tables describing the tillage comparisons and quantitative studies results in the form of effect sizes and pooled standard deviations. Those studies reporting measures of variability or providing data from which variability measures could be calculated were included in meta-analysis (see below). Studies reporting only means could not be meta-analysed and for these studies and all others, key descriptive characteristics of the evidence base were summarised using series of tables and figures.

Meta-analysis

We have performed (and hence report) models in the following order: (1) we have plotted meta-analyses without moderators and tested for heterogeneity; (2) where significant heterogeneity exists, we have included a complete list of moderators that we believe to be biologically significant (see below) and tested for heterogeneity again; (3) where significant heterogeneity still

remained we have tested for key significant interactions (see below) and included these where they proved to be significant.

Model fitting

Meta-analyses were conducted in R [44] using the *rma.mv* function the *metafor* package [45], which allows moderators to be declared as nested random factors. A total of 15 separate analyses were undertaken; 9 for concentration data (separated by 0–15, 15–30, and > 30 cm depth layers for each of the three tillage level comparisons [NT-vs-IT, NT-vs-HT, IT-vs-HT]) and 6 for stocks data (total sampling depth across the upper profile (0–30 cm), or the full profile (0–150 cm) for each of the three tillage comparisons). For all models, study ID (a unique code for each independent study) was nested within study site and declared as a random factor. All models used maximum likelihood (ML) to estimate random effects, which has been shown to be appropriate for comparisons between like models (unlike restricted maximum likelihood, ReML) [46]. In all cases, the following basic model was used for both concentration and stock analyses:

$$SOC_{ES} \sim SOC_{-ref} + duration + \text{depth}_{\text{till}} + soil + \underline{latitude + climate} + (\sim study|site)$$

where SOC_{ES}, raw mean difference SOC; SOC_{ref}, reference (i.e. the comparator) SOC value; tillage, paired tillage comparison (NT-vs-IT, NT-vs-HT, IT-vs-HT); duration, study duration; latitude, decimal latitudinal study location; climate, Köppen-Geiger climate zone; $depth_{till}$, comparator tillage depth category; soil, soil texture class; study, study code; site, study site.

The following key moderators were included and retained in all models where the data allowed: SOC_{ref}, duration, $depth_{till}$, and soil class. Latitude and climate zone were included individually and only retained in the models if they were significant. Moderators were chosen because they have been widely used by previous authors as factors influencing C sequestration, particularly climate, soil types and texture [47–49].

For comparisons between two different tillage types (i.e. IT–HT) an additional moderator ($depth_{till-B}$, intervention tillage depth) was included, and four additional interactions were tested between $depth_{till-B}$ and the following three moderators: SOC_{HI}, duration, and soil.

As described above, for each meta-analysis, the full model with moderators was tested for residual heterogeneity. Where significant residual heterogeneity existed in concentration meta-analyses, the models were then tested for the significance of interactions one by one. A list of important two-way interactions was assembled a priori and tested as follows:

1. Concentration meta-analysis
    - 1.1. NT–IT/NT–HT comparisons
        - 1.1.1. $depth_{till}$ * SOC_{ref}
        - 1.1.2. $depth_{till}$ * duration
        - 1.1.3. $depth_{till}$ * soil
    - 1.2. IT–HT comparisons
        - 1.2.1. $depth_{till}$ * SOC_{ref}
        - 1.2.2. $depth_{till}$ * duration
        - 1.2.3. $depth_{till}$ * soil
        - 1.2.4. $depth_{till-B}$ * $depth_{TILL}$
        - 1.2.5. $depth_{till-B}$ * SOC_{ref}
        - 1.2.6. $depth_{till-B}$ * duration
        - 1.2.7. $depth_{till-B}$ * soil

where these interactions were significant they were retained in the models. Interactions were not tested for in stocks data meta-analyses due to low sample size and underrepresented subgroups.

When we present our results we present first the results of a basic meta-analysis (i.e. models without moderators). We then tested for the presence of heterogeneity. Where there was no heterogeneity we did not attempt to include moderators and finished by testing for bias (publication, validity, and variability). Where significant heterogeneity existed, we attempted to include moderators as described above. We then test for residual heterogeneity. If residual heterogeneity remains, we then tested for significance of interaction terms. Because of unexplained heterogeneity and the risks of overparameterisation, we choose to present all models (i.e. both unmoderated and moderated) in an attempt to increase transparency. We avoid using models with heterogeneity or overparameterisation when making conclusions, since these models are not reliable.

Sensitivity analyses

Sensitivity analyses were carried out for each model to investigate the influence of critical appraisal categories and types of variability measures used. Firstly, for each of the 15 models above additional models were fitted using just those studies assessed as being 'high' validity. Secondly, separate models were fitted using only those studies that reported individual variability measures (i.e. separated by treatment group). For both sets of analyses, the results were compared to the overall model fit to examine significant differences in mean effect sizes.

Duplicate studies

Study site was denoted as a random factor in the model, accounting for multiple studies being undertaken on

some sites. There is no clear distinction in the evidence base between studies and experiments, since the physical experiments exist independently of studies that measure their outcomes. Often, single experiments are measured multiple times, since they are long-term experimental set-ups. Similarly, at any one site experiments can be established independent of one another, whilst research authors do not typically identify on which fields or plots the experiments were undertaken. In order to remain conservative in our analysis we could remove all duplicate studies, but this is an inherently challenging task due to the lack of detail in the study reports. Therefore, we have chosen to retain all studies in our analysis and treat each study as a random factor nested within study site locations.

Assumptions and other tests Heterogeneity was tested for amongst the evidence base by calculating τ^2 and performing Q/QE tests for heterogeneity/residual heterogeneity [50], integrated into the rma functions within metafor. Significant heterogeneity indicates the presence of a moderator that has not been accounted for in the model. Heterogeneity was tested for in simple univariate meta-analysis models (declaring study code and site as nested random factors) and again following the addition of moderators to examine the influence of including moderators on residual heterogeneity.

The presence of publication bias was investigated by performing an Egger's regression test, and by plotting funnel plots (effect sizes against standard errors) and looking for asymmetry, which is indicative of publication bias.

The influence of individual studies was examined by plotting Cook's Distance for each study [51], pointing out small groups of studies with considerable influence in the models.

Visualisations

All meta-analyses were plotted as forest plots (provided in aditional files) and the summary effect estimates and 95% confidence intervals combined into single plots for each of the concentration and stock sets of analyses. Where categorical moderators were significant, boxplots for these subgroup analyses were produced using coefficients from full moderated models (having tested and then removed the moderators climate zone and latitude where necessary). Where continuous moderators and interactions were significant, scatterplots for these meta-regressions were produced using coefficients from full moderated models (having tested and then removed the moderators climate zone and latitude where necessary). Regression lines are plotted from model coefficients that account for moderators (and climate zone or latitude where significant).

Results

Review descriptive statistics

Numbers of relevant articles/studies and their sources

A total of 288 articles and 351 studies were included in the systematic review (see Additional file 3). The search update returned 2338 relevant records, with 1376 remaining after removal of duplicates (see Fig. 2 for flow diagram; Additional file 1 for database search records). Following title screening 636 records were excluded, and following abstract screening a further 455 were excluded, leaving 312 articles to be retrieved for full text screening. Some 20 articles could not be retrieved for various reasons (see Additional file 2). Full text screening resulted in the inclusion of 56 articles and 64 studies, with 232 articles and 288 studies being included from the systematic map (see Additional file 3 for a list of studies excluded from the systematic map with reasons, respectively).

Articles and studies

The publication rate of articles within the review demonstrates an exponential increase over time, with a relatively recent history of only 25 years (Fig. 3). The 57 articles identified through the update demonstrate that a high proportion of the evidence base (20%) was published in the 2 years since the original search was performed (September 2013).

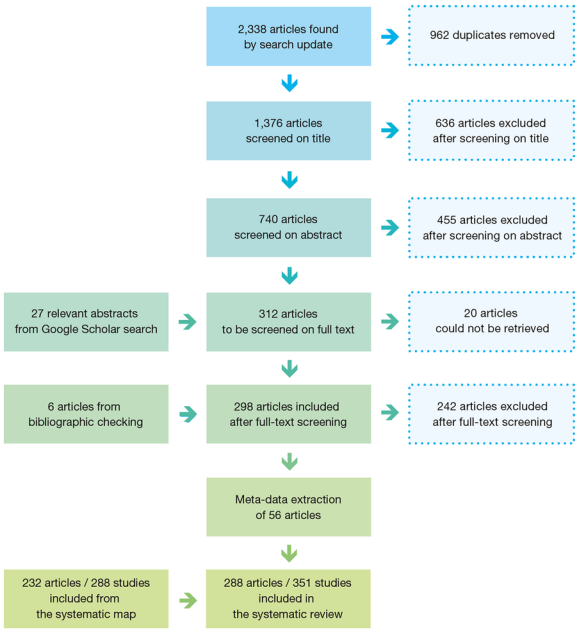

Fig. 2 Flow diagram showing sources of studies in the systematic review

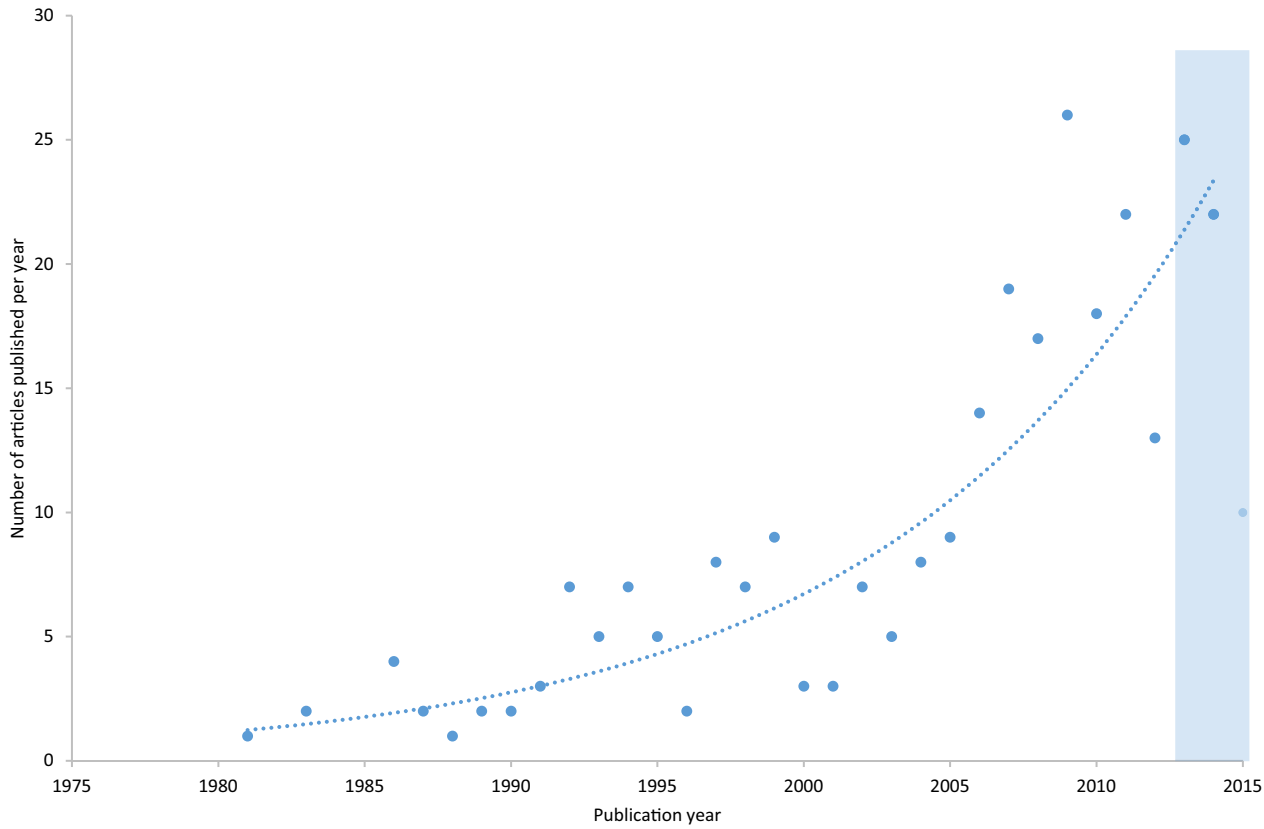

Fig. 3 Publication rate of articles in the systematic review. The dotted line represents an exponential curve fit. The shaded area represents studies identified mainly within the search update, rather than the original systematic map published in 2015

Study sites

Across the 351 studies in the review, the most commonly studied country was the USA (142 studies), followed by Canada (46), and Spain (42) (Table 3). Figure 4 displays the discrepancies between the area of arable land and the number of studies identified during this systematic review. This identifies several countries that are well studied relative to the area of arable land: Switzerland, Spain and Denmark. This data should be viewed with caution because it does not take into account the area of arable land within included climate zones. Table 4 displays the number of studies per climate zone, and shows that Cfa (humid subtropical, such as the southeaster USA) was the most commonly studied zone (123 studies), with Dfb, Cfb and Dfa (which have humid climates year-round or nearly so) equally represented (63, 60 and 50 studies, respectively). A total of 213 of the 351 studies (61%) allowed common USDA soil texture classes to be calculated, and for 82 of these 213 studies soil texture classes were estimated from sand, silt and clay percentages. A further 83 studies that did not provide enough information to calculate USDA soil texture classes reported some other form of description of the soil type, whilst 37 studies failed to report any description of the soil at the study site.

Study designs and experimental layout

A total of 179 studies (51% of 351 studies) were focused purely on investigations of the impacts of tillage, whilst the remaining 172 studies included combined paired, factorial, blocked or split plot assessments of other interventions, including: amendments, crop rotation, fertiliser, and irrigation. Studies ranged in duration from 10 years (the minimum required for inclusion in the review) to 100 years (Fig. 5). Only 1 study failed to provide information about its duration, whilst 20 studies out of 351 (6%) reported study duration but not the years the study took place. Randomisation was common in experimental designs (228 studies), with blocking (160 studies) and split-plot (117 studies) designs also common (Fig. 6). Some 29 studies failed to report their study design. Figures 7 and 8 show the number of true spatial replicates and temporal replicates used across the evidence base. The median level of spatial replication was 4 (151 studies), with 3 replicates also very common (113 studies), together forming 75% of the evidence base. Temporal

Table 3 Number of studies per country in the review

Continent	Study country	Number of studies
Africa	Morocco	2
Africa	South Africa	1
Asia	China	1
Asia	Syria	1
Australasia	Australia	21
Australasia	New Zealand	1
Europe	Spain	42
Europe	Germany	17
Europe	Denmark	7
Europe	Italy	6
Europe	Finland	4
Europe	UK	4
Europe	France	3
Europe	Sweden	3
Europe	Switzerland	3
Europe	Czech Republic	2
Europe	Lithuania	2
Europe	Austria	1
Europe	Belgium	1
Europe	Ireland	1
Europe	Norway	1
Europe	Poland	1
Europe	Serbia	1
Europe	Slovakia	1
Europe	Ukraine	1
North America	USA	142
North America	Canada	46
South America	Brazil	25
South America	Argentina	6
South America	Uruguay	3
South America	Chile	1

replication was not common, with the majority of studies (267: 76%) not reporting any repeated sampling. Some 18 studies failed to report the level of spatial replication, whilst only 3 studies failed to report temporal replication.

Soil sampling

A large proportion of the evidence base only sampled one soil layer (105 studies), whilst 149 studies (42%) sampled 3 or more layers (Fig. 9). Only 1 study failed to report the sampling depth measured. The soil sampling depth critical appraisal scoring was undertaken as follows: '0', shallow (maximum depth $\leq$ 15 cm) single or multiple sampling; '1', plough layer (maximum depth 15–25 cm) single or multiple sampling, or deep (maximum depth > 25 cm) single sampling; '2', multiple deep sampling (maximum depth > 25 cm). A total of 95 studies were given a score of 0, 118 studies a score of 1, and 138 studies a score of 2, demonstrating a relatively even distribution of soil sampling strategies. For studies reporting concentration data (see 'Outcome reporting' below), 265 reported data for the 0–15 cm layer, 112 for the 15–30 cm layer, and 66 for the > 30 cm layer. For studies reporting stocks data, the median soil depth sampled was between 15 and 30 cm (67 studies), whilst other depths were less common: 0–15 cm, 34 studies; 30–75 cm, 29 studies; > 75 cm, 16 studies.

Outcome reporting

Virtually all studies reported SOC (336 studies), whilst a small number reported SOM that was converted to SOC as described previously (15 studies). Over half of the studies reported concentration data alone (195 studies: 56%), 92 studies reported only stocks (26%), and 64 studies (18%) reported both together. Just over half of all studies reported bulk densities (183 studies: 52%), with similar rates of reporting for studies with concentration data (56%) as for studies with stocks data (58%) (Fig. 10). A large number of studies failed to provide measures of variability (i.e. standard deviation, standard error, 95% confidence intervals and coefficients of variation) around their means (111 concentration studies, 41%; 85 stocks studies, 58%), precluding them from inclusion in any form of meta-analysis. Relatively few studies provided variability measures separated by tillage treatment groups (30 and 24% for concentration and stocks studies, respectively), however, the body of evidence that was meta-analysable was greater than these numbers, since some studies provided overall variability measures (for treatments groups combined, some form of pooled measure), some studies provided raw data, and some studies provided p values and least square difference (LSD) values that permitted pooled or individual variability measures to be calculated (Table 5). The use of these other forms of variability measure allowed us to increase the meta-analysable body of evidence from 81 to 160 studies for concentration data meta-analyses, and from 35 to 61 studies for stocks data meta-analyses.

Tillage treatment comparisons

Comparisons between NT and HT were the most common (200 studies: 57%), with NT versus IT studied in 101 studies (29%), and IT versus HT studied in just 50 studies (14%). Tillage depth for HT studies was most commonly deep (148 studies), with relatively few shallow (19 studies), and a large number of undescribed tillage treatments (51 studies). Mouldboard ploughing (169 studies), very deep ($\geq$ 40 cm) chisel tillage (24 studies) also referred to as sub-soiling, and ridge tillage (17 studies) were the most frequently described methods for HT (Table 6). Tillage

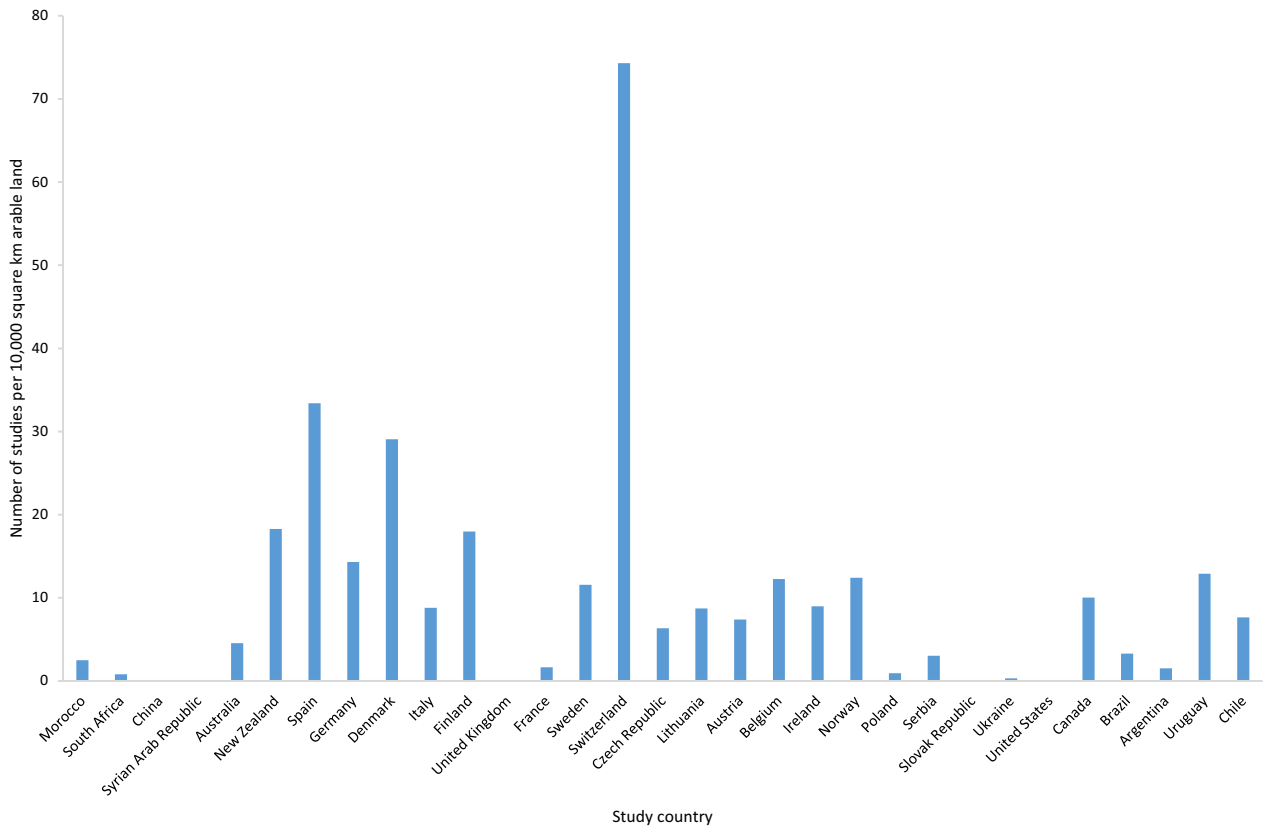

Fig. 4 Studies per 10,000 km^2 of arable land, separated by country. Arable land percentage and land area per country data for 2013 from The World Bank (http://data.worldbank.org/indicator/AG.LND.ARBL.ZS, accessed 15/06/2016)

Table 4 Number of studies per climate zone in the systematic review

Köppen climate zone	Number of studies
Cfa	123
Dfb	63
Cfb	60
Dfa	50
Csa	35
Dfc	8
Csb	6
Cfb and Csa	1
Cfb and Dfb	1
Cfb/Cfc	1
Dfa and Cfa	1
Dfa and Dfb	1
Dfa/Dfb	1

depth for IT studies (studies comparing NT with IT) was most commonly shallow (71 studies), with slightly fewer deep (50 studies), and a large number of non-described tillage treatments (57 studies). A wider range of tillage types was investigated in IT comparisons than for HT comparisons (see Table 7).

Systematic map

In the process of undertaking this review we have produced an updated systematic map (relative to the systematic map published in 2015 [33]) for studies that purely focus on tillage interventions (Additional file 7). The studies in this map have also been visualised in an updated geographic information system (GIS) that can be accessed through the following: http://www.eviem.se/en/projects/SOC-Tillage/). A help file has been produced to assist with use of the online GIS (Additional file 7).

Narrative synthesis

Descriptive meta-data and coding for all included studies and their effect size data for concentration and stocks reporting studies are available in Additional files 6, 8, and 9, respectively.

Validity of the evidence

Figure 11 displays the critical appraisal scores that were awarded to all studies in the review. Although only spatial

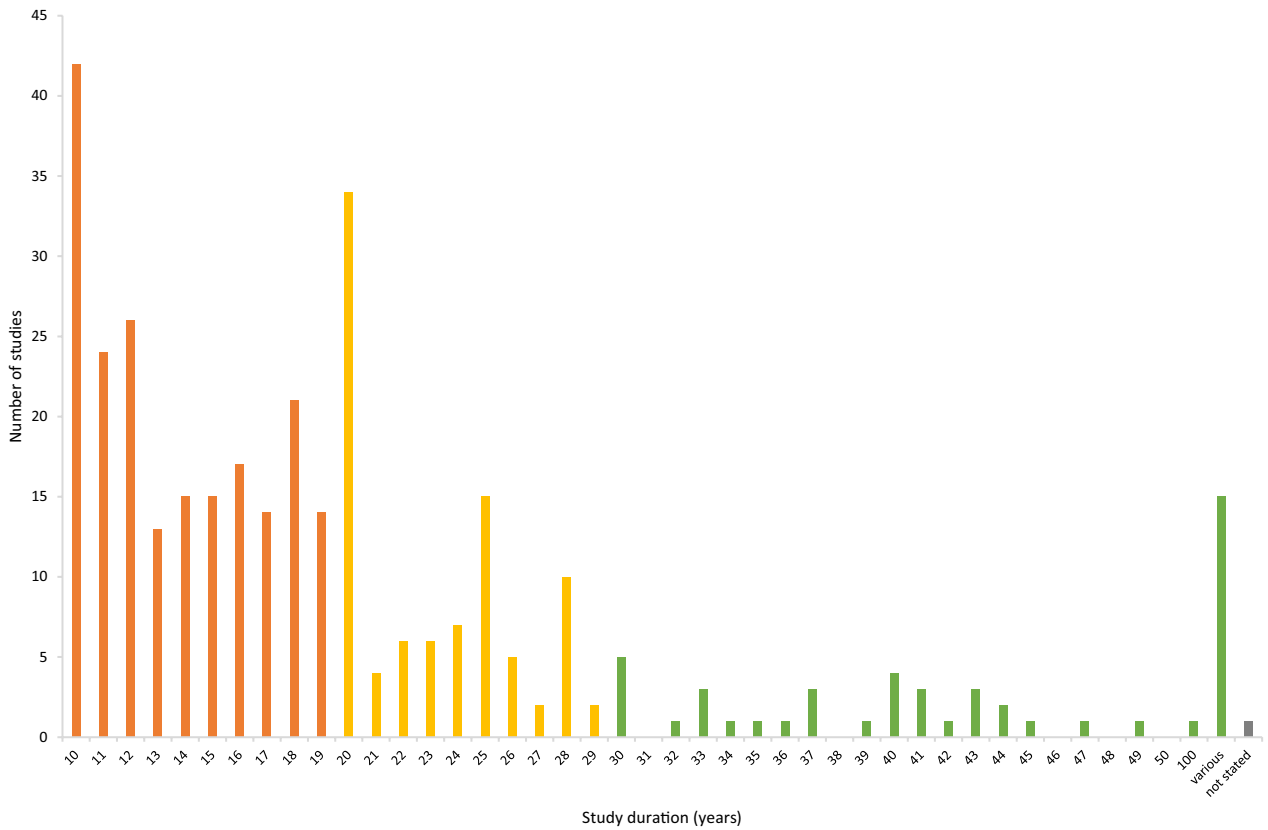

Fig. 5 Duration of studies included in the review. Green bars represent a critical appraisal score of 2, yellow bars 1, orange bars 0, and grey bars 'unclear'

replication and treatment allocation domains were used in the meta-analysis, we will discuss the general patterns across the evidence base here. As mentioned above, spatial replication was relatively low (82% studies with a score of '0' or '1'). Temporal replication was also low, with the majority of studies conducting sampling at one time point. In general treatment allocation was of high validity, with the majority of studies (85%) employing some form of blocking (typically also employing randomisation, see above). The majority of studies scored poorly for experimental duration (69% with a score of '0'), being conducted over 10–20 years. Soil sampling was generally of moderate validity, with most studies scoring '2' in this domain: these studies performed deep sampling with multiple layers sampled separately.

Meta-analysis

For all analyses reported here, detailed statistical outputs (including all non-significant tests) and models used are provided in Additional files 10, 11 for concentration and stock meta-analyses, respectively. Copies of the R-scripts used (Additional files 12, 13), along with the data files used (Additional files 8, 9) are also provided.

We present results first for simple models lacking moderators. Where significant heterogeneity exists we then present results for moderated models before checking for residual heterogeneity. Finally, if significant heterogeneity still remains, we then present results for significance of interactions. Due to the complex structure of moderators and the relatively low sample size, we must be cautious about the risk of overparameterisation, but must also be careful not to base conclusions on models with substantial unexplained heterogeneity. We therefore choose to present all model results for transparency.

Concentration data

Figure 12 and Table 8 display the summary effect estimates for all of the nine meta-analyses on concentration data. These estimates are for the basic models and do not account for moderators, discussed below. Their purpose is to identify clear patterns. A lack of significance does not indicate no significant patterns within the evidence and can only be interpreted as a lack of evidence for an effect if there is no indication of heterogeneity. Where heterogeneity exists, moderators may be significantly driving different patterns within the evidence. As such,

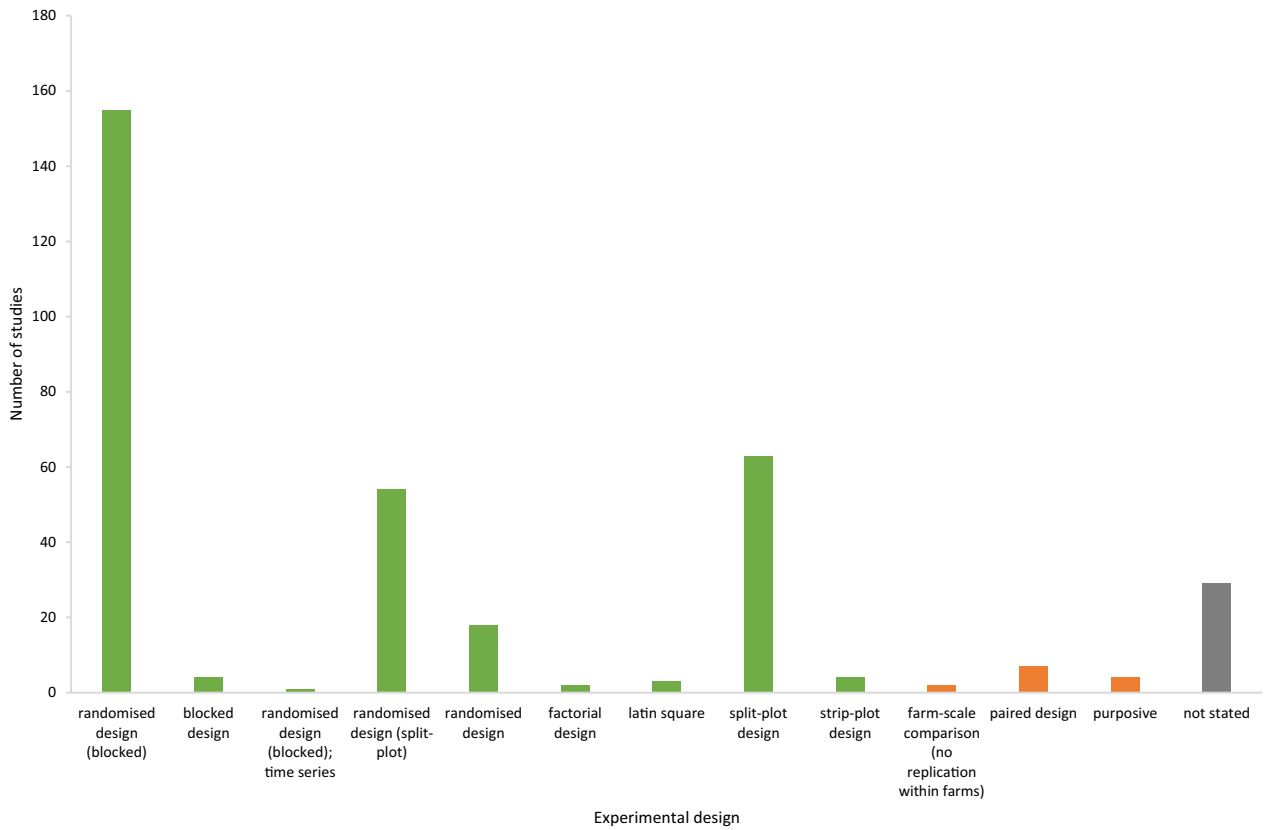

Fig. 6 Number of studies with different study designs. Green bars represent a critical appraisal score of 2, orange bars 0, and grey bars 'unclear'

we will not discuss this plot further, but rather examine each meta-analysis in detail in the following pages.

NT–HT 0–15 cm A significant positive difference in SOC in NT relative to HT can be seen for the simple model at 0–15 cm (Fig. 13). There was significant heterogeneity in this model ($Q_{101} = 554.631$ $p < 0.001$), which remained following the addition of moderators ($Q_{88} = 297.256$ $p < 0.001$). No interaction terms were significant, nor were the single moderators, latitude and climate zone (see Additional file 10). Study duration, soil class, and HT depth category were significant ($LRT_{15} = 12.605$ $p < 0.001$, $LRT_{7} = 19.005$ $p = 0.025$, and $LRT_{14} = 7.923$ $p = 0.019$, respectively), whilst reference SOC was not ($LRT_{15} = 0.329$ $p = 0.566$). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias, with two studies exerting high influence on the model (see Additional file 10).

Figure 14 demonstrates the significant positive relationship between study duration and SOC difference in NT relative to HT at 0–15 cm: the regression line intercepts the y-axis at around 10 years, indicating that studies longer than 10 years are needed to detect a difference in SOC. Figure 15 shows the effect of HT depth on the SOC difference in NT relative to HT, suggesting that a change from deep HT to NT would result in a greater SOC increase near the surface than a change from shallow HT. Figure 16 displays the effect of soil texture class on the SOC difference in NT relative to HT, with some soil classes appearing to demonstrate greater effects of NT than others: sandy clay loam (SaClLo) and silty clay (SiCl), in particular.

NT–HT 15–30 cm There was no significant difference in SOC in NT relative to HT at 15–30 cm observed in the simple model (Fig. 17). There was significant heterogeneity amongst studies ($Q_{48} = 224.173$ $p < 0.001$), which was not present in the moderated model ($Q_{35} = 30.502$ $p = 0.685$). The single moderator climate zone was not significant (see Additional file 10). The single moderators latitude, soil class and HT depth category were significant ($LRT_{15} = 14.642$ $p < 0.001$, $LRT_{8} = 21.399$ $p = 0.006$, and $LRT_{14} = 16.524$ $p < 0.001$, respectively), whilst study duration and reference SOC were not ($LRT_{15} = 0.024$ $p = 0.878$ and $LRT_{15} = 0.016$ $p = 0.900$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and

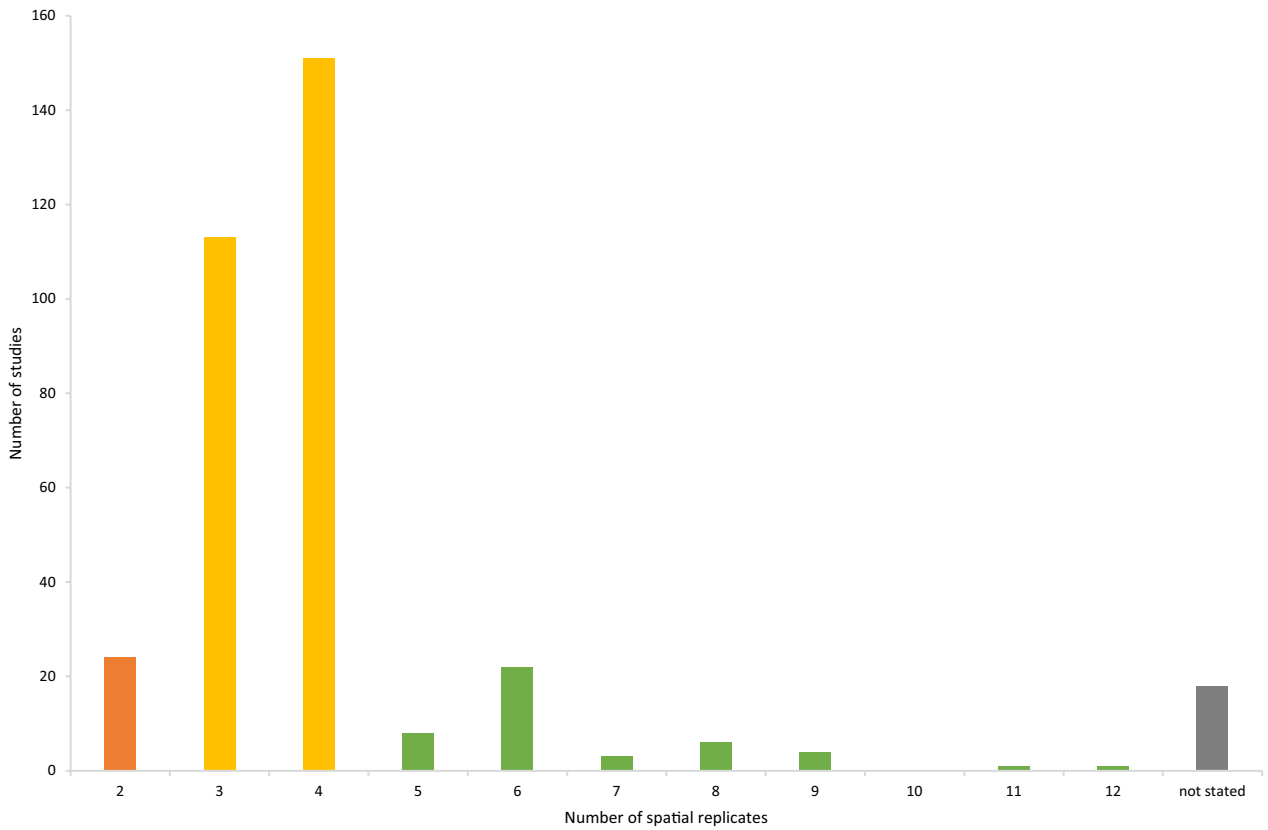

Fig. 7 Level of true spatial replication across studies. Green bars represent a critical appraisal score of 2, orange bars 0, and grey bars 'unclear'

there was no evidence of publication bias, whilst one study appeared to have a high influence in the model (see Additional file 10).

Figure 18 displays the significant negative relationship between latitude and SOC difference in NT relative to HT at 15–30 cm, showing that there is a change in direction of effect from positive at latitudes below c. 38° and negative at latitudes above 38°. The impact of soil texture class is shown in Fig. 19, and suggests that soil types may differ in their responses to a reduction in tillage: loams (Lo) and sandy clay loams (SaClLo) show a negative response (i.e. a reduction in SOC), whilst silty clay loams (SiClLo) show a positive response. Figure 20 shows the difference in SOC in NT relative to HT, NT results in a loss of SOC relative to both shallow and deep HT, with a change from deep HT showing a greater loss (and greater variability around the mean) than shallow HT.

NT–HT > 30 cm No significant difference in SOC in NT relative to HT was apparent from the simple model (Fig. 21). Significant heterogeneity was present in this model ($Q_{30} = 68.217$ $p < 0.001$), which was not present in the moderated model ($QE_{20} = 17.363$ $p = 0.629$). Neither latitude nor climate zone were significant (see Additional file 10). Reference SOC and HT depth category were significant ($LRT_{12} = 28.451$ $p < 0.001$, $LRT_{11} = 18.1137$ $p < 0.001$, respectively), whilst duration and soil class were not ($LRT_{12} = 1.739$ $p = 0.187$ and $LRT_{7} = 12.513$ $p = 0.052$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, although there was evidence of publication bias: more precise studies appear to show negative effect sizes, whilst less precise studies had positive findings. Three studies appeared to contribute strongly to the models (see Additional file 10).

Figure 22 displays the significant negative relationship between reference SOC and the difference in SOC in NT relative to HT in depths below 30 cm, showing that soils with a starting SOC of c. 5 g/kg and below respond with an increase in SOC in NT, whilst soils with SOC concentration greater than 5 g/kg demonstrate a reduction in SOC following conversion to NT. Figure 23 shows the difference in SOC in NT relative to HT for different HT depth categories, and indicates that the significant result for this moderator is likely spurious, since the shallow group is represented by only 1 study, and it is the 'not stated' group that does not overlap the line of no effect.

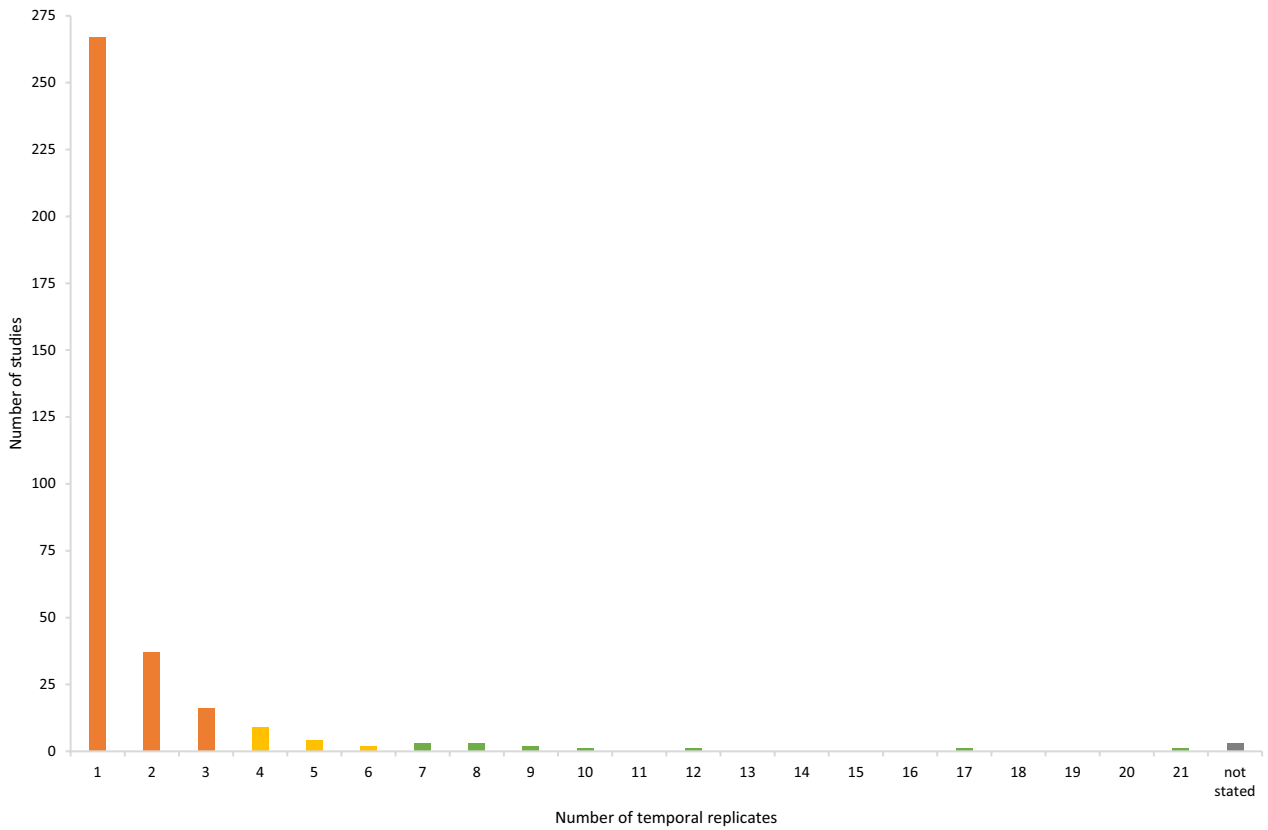

Fig. 8 Level of temporal replication across studies. Green bars represent a critical appraisal score of 2, orange bars 0, and grey bars 'unclear'

NT–IT 0–15 cm A significant positive overall pattern can be observed in the simple model of NT versus IT at 0–15 cm (Fig. 24). Significant heterogeneity was present in this model ($Q_{94} = 364.884$ $p < 0.001$), which remained in the moderated model ($Q_{94} = 364.884$ $p < 0.001$). There was a significant interaction between IT depth category and study duration ($LRT_{16} = 19.987$ $p < 0.001$). All other interactions terms were not significant, nor were the single moderators, latitude and climate zone. Soil class and reference SOC were also not significant ($LRT_7 = 2.957$ $p = 0.996$ and $LRT_{15} = 0.764$ $p = 0.382$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias. OIne study was more influential than others, but many studies contributed with moderate influence (see Additional file 10).

Figure 25 shows the interaction between IT depth category and study duration, demonstrating that a conversion to NT from deep IT increases SOC linearly over time to a greater extent than a conversion from shallow IT.

NT–IT 15–30 cm No significant overall summary effect was identified in the simple model of NT versus IT at 15–30 cm (Fig. 26). Significant heterogeneity was present ($Q_{44} = 512.163$ $p < 0.001$), which was still present in the moderated model ($QE_{30} = 256.097$ $p < 0.001$). The interactions between soil class and IT depth category and study duration and IT depth category were not significant, nor were the single moderators, latitude and climate zone. The interaction between IT depth category and reference SOC was significant ($LRT_{15} = 17.473$ $p < 0.001$). Soil class and study duration were not significant ($LRT_9 = 1.509$ $p = 0.993$ and $LRT_{16} = 0.025$ $p = 0.874$). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias. Two studies were particularly influential in these models (see Additional file 10).

Figure 27 shows a negative relationship between reference SOC at 15–30 cm and difference in SOC between NT and IT at shallow IT depths, whilst there is no relationship for deep IT depths: soils with a greater starting SOC concentration demonstrate a greater loss of SOC in shallow IT, whilst reference SOC has no impact on difference in SOC for deep IT.

NT–IT > 30 cm The simple model did not identify a clear significant pattern within the evidence base

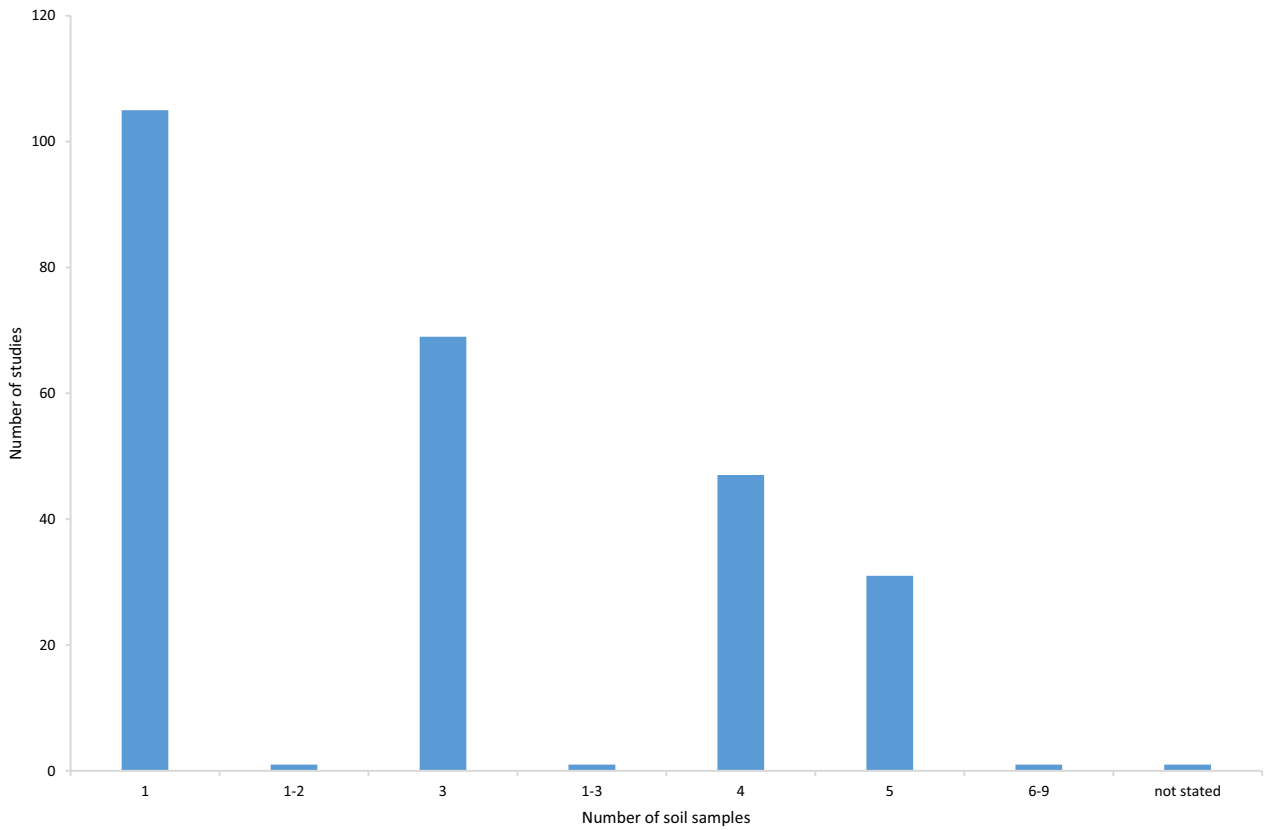

Fig. 9 Number of soil samples measured within studies in the review

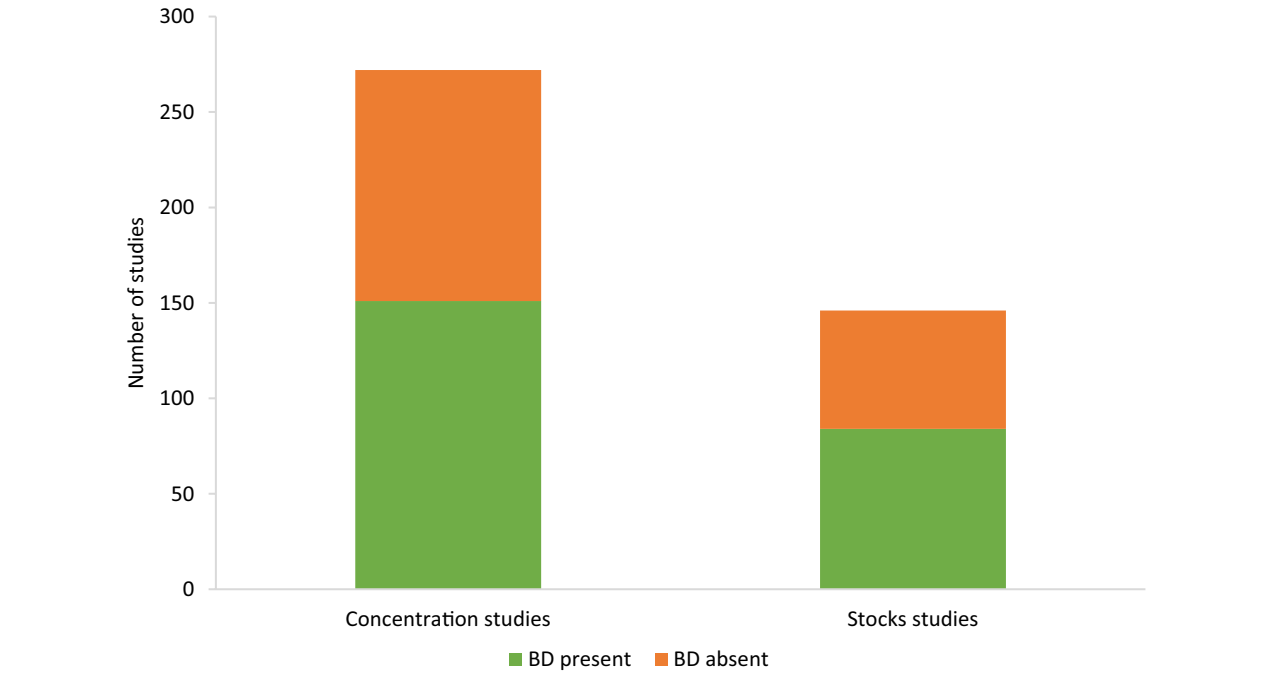

Fig. 10 Number of studies reporting concentration and stock data that also report bulk density

Table 5 Number of studies reporting different forms of variability measures around their means. See text for explanation of terms

Concentration study variability measure	Number of studies	Stock study variability measure	Number of studies
None	111	None	85
Undescribed measure	6	Undescribed measure	1
Overall 95% CI	1	Overall 95% CI	1
Overall CV	10	Overall CV	6
Overall LSD	44	Overall LSD	12
Overall LSD/overall p values	1	Overall p values	1
Overall p values	8	Overall SE	2
Overall SE	2	Raw data	3
Overall SED	1	Individual CI	1
Raw data	6	Individual LSD	2
Individual 95% CI	0	Individual SD	15
Individual LSD	4	Individual SE	17
Individual p values	3	Total	146
Individual SD	23		
Individual SE	50		
Individual SED	1		
Total	271		

Table 6 High intensity tillage descriptions

High intensity tillage descriptions	Number of studies
Mouldboard plough	156
Mouldboard and described secondary tillage	12
Not stated	9
Subsoiler	9
Chisel, disk and ridge tillage	9
Ridge tillage	8
Plough	7
Subsoiler and field cultivator	2
Straight shanks, deep rip	2
Three tillage passes	1
Reversible plough	1
Chisel plough and occasional mouldboard plough	1
Subsoiler and chisel tillage	1

(Fig. 28). There was no significant heterogeneity present in this model ($Q_{19} = 16.044$ $p = 0.654$). As expected, the interaction terms and the single moderators, latitude and climate zone, were therefore not significant. Similarly, study duration, soil class, reference SOC and IT depth

Table 7 Intermediate intensity tillage descriptions (from no tillage versus intermediate tillage comparisons)

Intermediate intensity tillage descriptions	Number of studies
Chisel tillage	56
Disk tillage	40
Chisel and disk	16
Not stated	7
Cultivator	6
Heavy duty cultivator	4
Disk and harrow	4
Heavy disk plough, harrow and rotavator	3
Cultivator or disk harrow	3
Rotavator	2
Chisel, disk and field cultivator	2
Offset disk and scarifier	2
Tine cultivator	2
Spring-tined cultivator	1
Multiple disk tillage	1
Sweep plough	1
Rototill	1
One cultivation	1
Blade cultivator, single pass	1
Heavy duty cultivator, 5–6 passes	1
Disk and mulch treader	1
Chisel, disk and rotary tiller	1
Double disk tillage	1
Heavy duty cultivator, disk and harrow	1
Disk and scarifier	1
Rotary harrow	1
Rotary field cultivator	1
Chisel, disk and spring tine	1
Cultivator (grubber)	1
Tandem disk	1
Undercutter and fallow master	1
Cultivator (disk coulters in combination with a spiker roller)	1
Rotary tiller (aggressive tillage)	1
Tined implements	1
Roto-till, sweep, disk and chisel	1
3 cultivations	1
Dual-layer tillage with a chisel and light secondary tillage (15 cm)	1
Heavy duty cultivator, sweep and harrow	1
Harrow	1
No till with one chisel every 3 years	1
Mouldboard plough or rotovator	1
Rigid tine cultivator	1
Chisel and cultivator	1
Grubber	1
Chisel and field cultivator	1

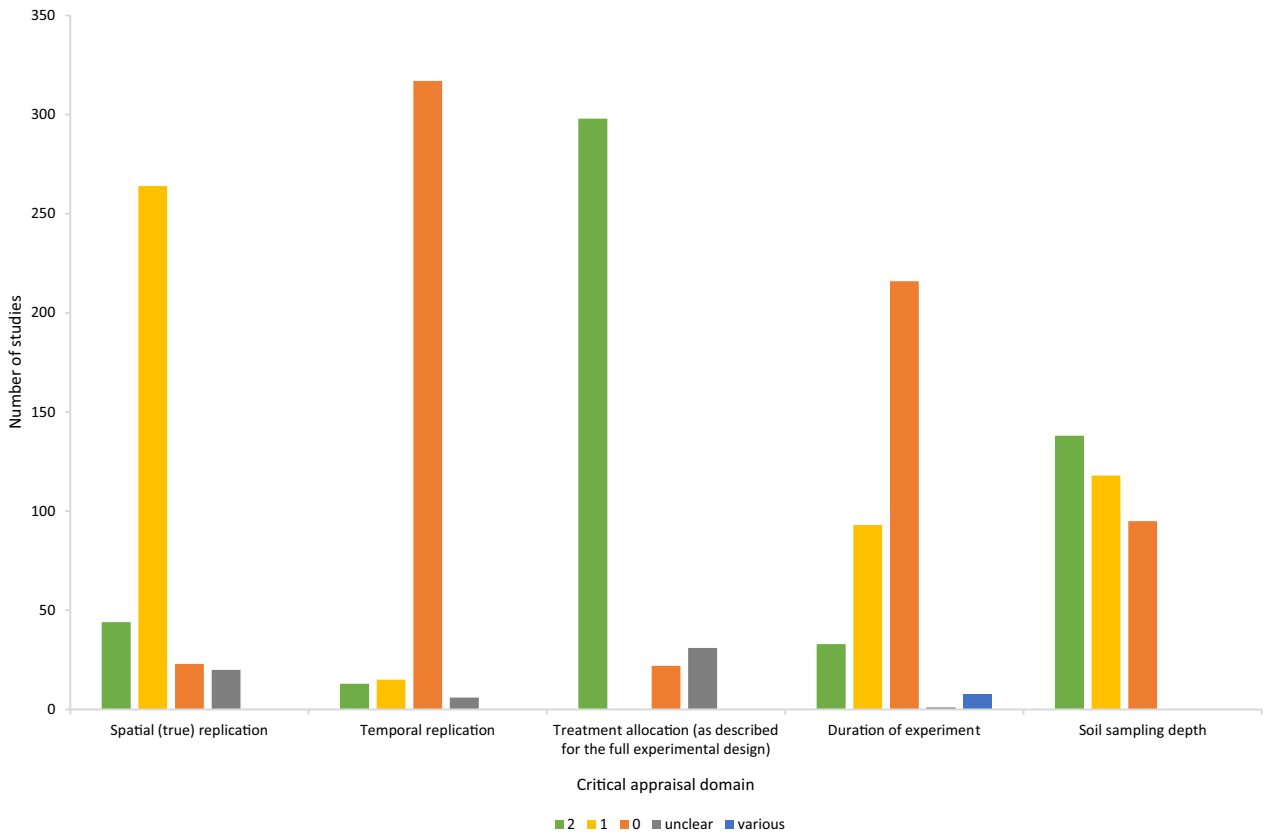

Fig. 11 Critical appraisal scores across the evidence base for five assessed domains. See text for explanation

category were not significant ($LRT_{12} = 1.170$ $p = 0.279$, $LRT_7 = 1.447$ $p = 0.963$, $LRT_{12} = 0.063$ $p = 0.801$, $LRT_{11} = 5.091$ $p = 0.078$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias. One study was more influential than others, although the sample size is low (see Additional file 10).

IT–HT 0–15 cm A significant positive pattern was detected across the evidence in the simple model (Fig. 29). Significant heterogeneity was also present ($Q_{76} = 168.336$ $p < 0.001$), which was not present in the moderated model ($QE_{48} = 60.681$ $p = 0.219$). There was a significant interaction between IT depth category and soil class ($LRT_{24} = 22.009$ $p = 0.003$). No other interaction term was significant, nor were the single moderators, latitude and climate zone. Study duration and reference SOC were also not significant ($LRT_{30} = 1.124$ $p = 0.289$ and $LRT_{30} = 0.203$ $p = 0.653$). HT depth category was marginally not significant ($LRT_{29} = 5.1506$ $p = 0.076$). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, whilst there was some statistical evidence of publication bias, indicated in the funnel plot by a slight positive tendency in studies with lower precision. A large number of studies contributed to the models, with no single study showing strong influence (see Additional file 10).

Figure 30 shows the impact of soil class on SOC difference at 0–15 cm between IT and HT for deep and shallow IT depth categories. The significance of this interaction term may have come about due to low sample sizes in certain subgroups, but it demonstrates that some soils are consistently greater in SOC difference than others (e.g. sandy clay loams [SaClLo]), whilst other soils differ between deep and shallow IT (e.g. silty clay loams [SiClLo] and silt loams [SiLo]).

IT–HT 15–30 cm A significant negative pattern was detected in the simple model of IT versus HT for 15–30 cm (Fig. 31). Significant heterogeneity existed in this model ($Q_{41} = 198.235$ $p < 0.001$), which remained after including moderators ($Q_{26} = 159.521$ $p < 0.001$). Interactions were not run due to low sample size and overparameterisation, and the single moderators, latitude and climate zone, were also not significant (see Additional file 14). The moderators soil class, study duration, reference SOC, HT depth category and IT depth

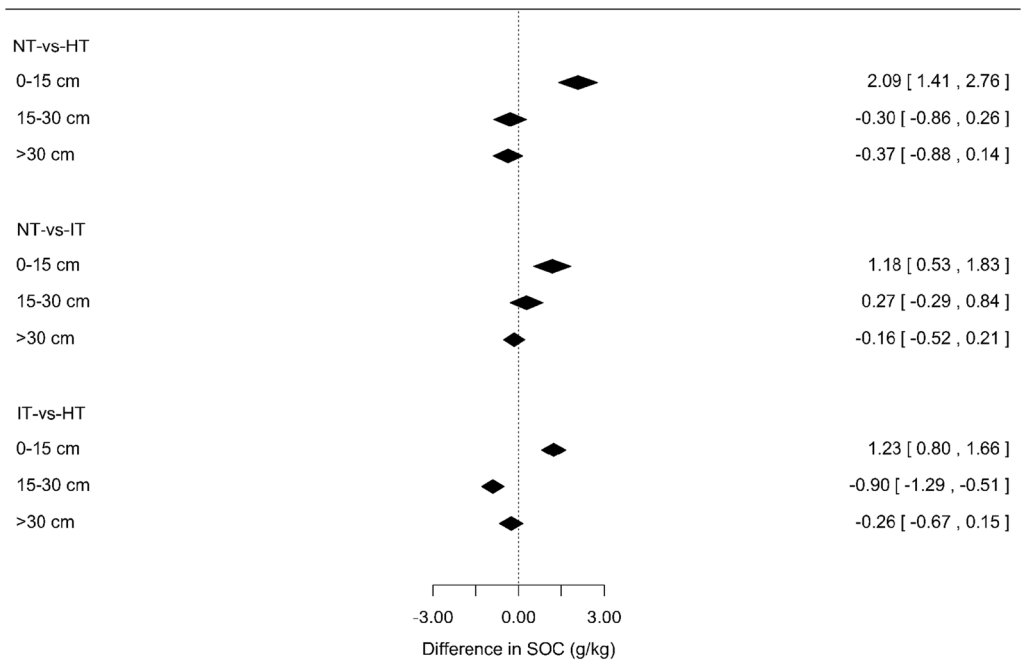

Fig. 12 Summary effect estimates (difference in SOC, g/kg) for concentration data meta-analyses. Three tillage comparisons are shown: *NT* no tillage, *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Diamonds are centred on the summary effect estimate for each meta-analyses, with the points of the diamonds representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

Table 8 Summary effect estimates for meta-analyses of concentration data

Tillage comparison	Soil depth (cm)	Summary effect estimate (g/kg)	Standard error	z statistic	p value	Lower 95% confidence interval	Upper 95% confidence interval	Number of studies
NT–HT	0–15	2.086	0.344	6.073	< 0.001	1.413	2.760	102
	15–30	− 0.298	0.286	− 1.039	0.299	− 0.859	0.264	49
	> 30	− 0.371	0.261	− 1.424	0.155	− 0.882	0.134	31
NT–IT	0–15	1.179	0.334	3.537	< 0.001	0.526	1.833	95
	15–30	0.273	0.289	0.944	0.345	− 0.293	0.839	45
	> 30	− 0.157	0.186	− 0.845	0.398	− 0.522	0.208	20
IT–HT	0–15	1.229	0.221	5.575	< 0.001	0.797	1.662	77
	15–30	− 0.890	0.198	− 4.539	< 0.001	− 1.288	− 0.511	42
	> 30	− 0.261	0.209	− 1.250	0.211	− 0.671	0.148	16

NT no tillage, *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation)

category were not significant ($LRT_9 = 14.143$ $p = 0.117$ and $LRT_{17} = 0.136$ $p = 0.287$, $LRT_{17} = 1.020$ $p = 0.312$, $LRT_{16} = 2.284$ $p = 0.319$, $LRT_{16} = 0.331$ $p = 0.848$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias. Two studies exerted very high influence over the models (see Additional file 10).

IT–HT > 30 cm There was no significant pattern in effect sizes for the simple model of IT versus HT from > 30 cm (Fig. 32). There was no heterogeneity amongst studies in this model ($Q_{15} = 12.765$ $p = 0.621$), nor in the moderated model ($QE_6 = 0.731$ $p = 0.994$). The single moderators latitude and climate zone were not signficiant (see Additional file 10). Reference SOC was significant ($LRT_{11} = 4.335$ $p = 0.037$). Study duration,

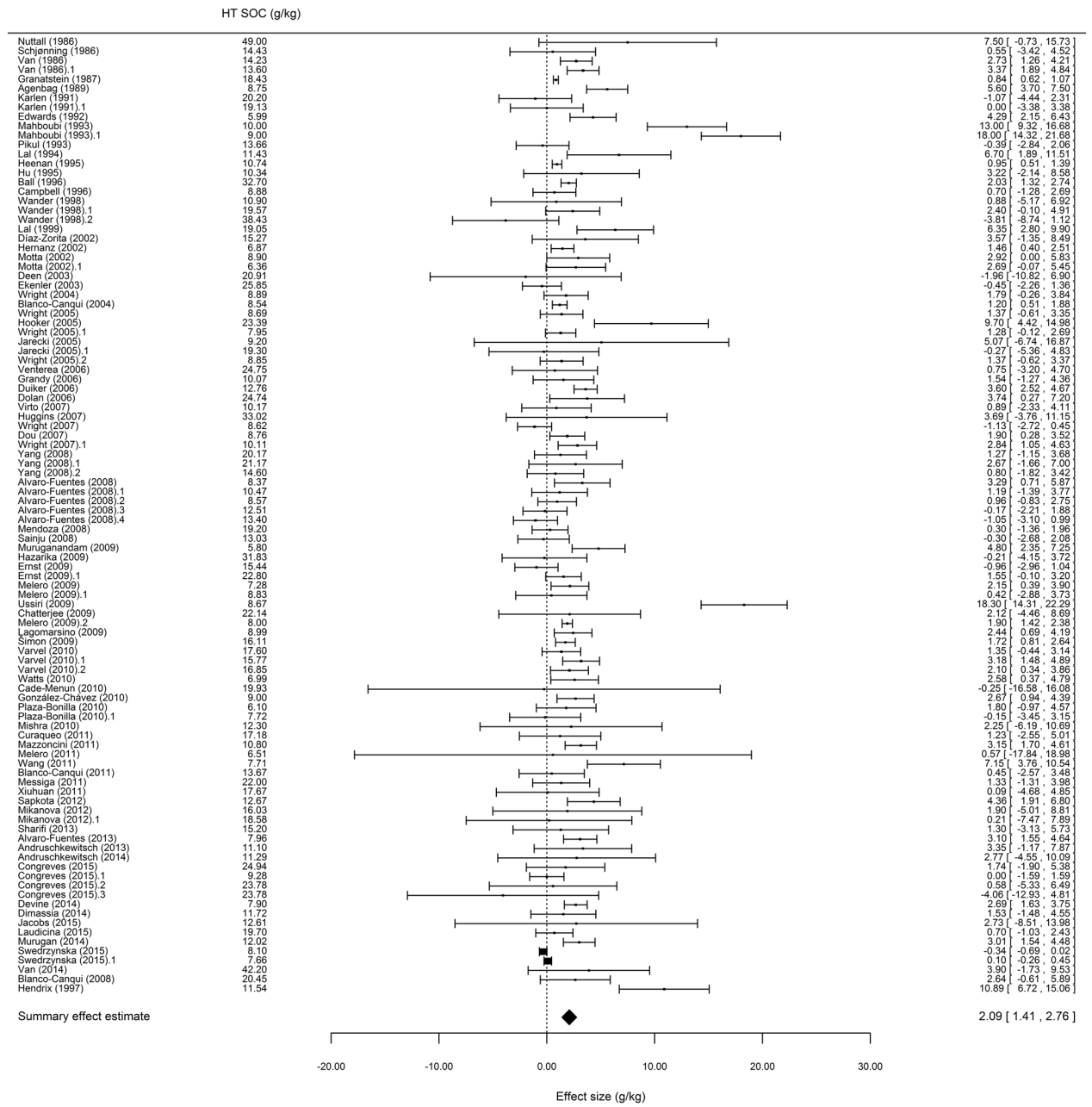

Fig. 13 Forest plot for meta-analysis of concentration data for NT–HT comparison at 0–15 cm depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

soil class, HT depth category and IT depth category were not significant ($LRT_{11} = 0.296$ $p = 0.587$, $LRT_{8} = 3.777$ $p = 0.437$, $LRT_{11} = 0.021$ $p = 0.886$, and $LRT_{10} = 1.327$ $p = 0.515$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias, with one study particularly influential in this small meta-analysis (see Additional file 10). Figure 33 shows the relationship between reference SOC and difference in SOC in IT relative to HT at > 30 cm, indicating that as reference SOC increases, the difference in SOC becomes more negative.

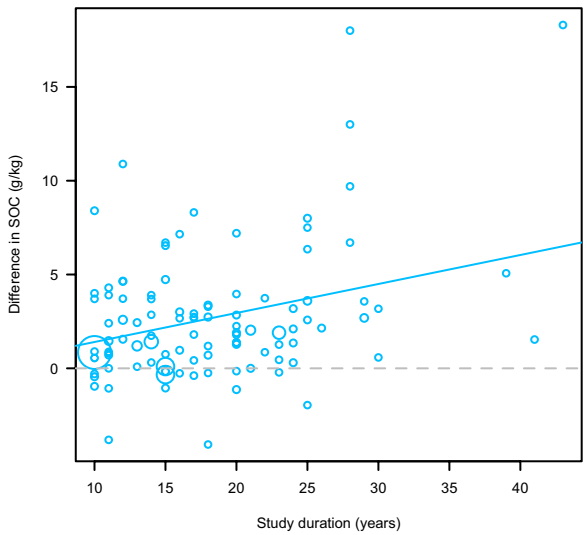

Fig. 14 Meta-regression of SOC concentration against study duration for NT–HT at 0–15 cm. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

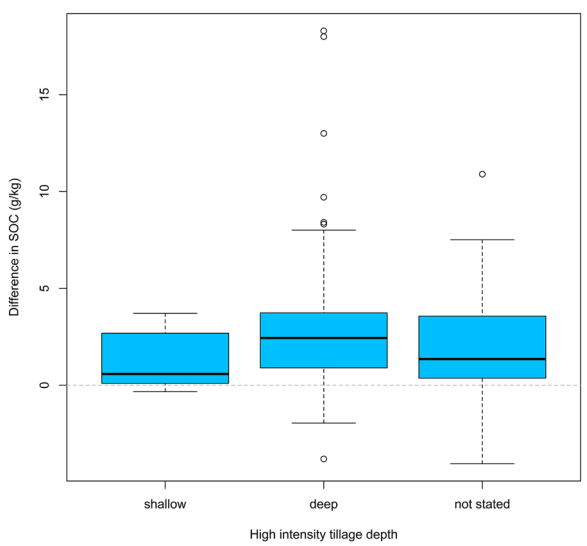

Fig. 15 Boxplot of difference in SOC concentration for NT–HT at 0–15 cm as affected by HT depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

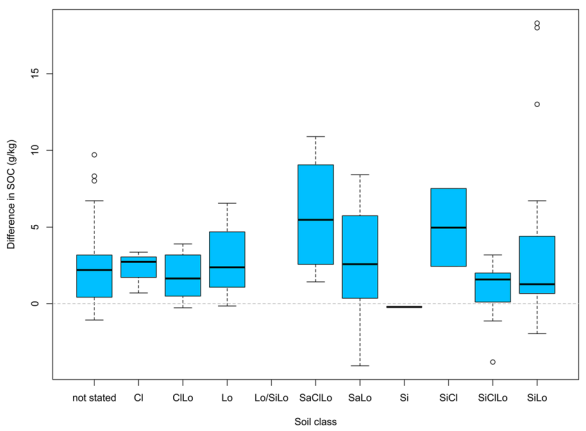

Fig. 16 Boxplot of difference in SOC concentration for NT–HT at 0–15 cm as affected by soil class. *NT* no tillage, *HT* high intensity tillage. See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

Stocks data

Figure 34 and Table 9 show the summary effect estimates for all six of the stocks data meta-analyses (basic models without moderators, as discussed above for concentration data).

NT–HT upper layer (0–30 cm) A significant positive overall effect was found for NT versus HT at 0–30 cm (Fig. 35), with significant heterogeneity present ($Q_{28} = 559.881$ $p < 0.001$). Latitude and climate zone were not significant (see Additional file 11). Soil class, reference SOC stock and HT depth category were not significant ($LRT_6 = 3.075$ $p = 0.799$, $LRT_{12} = 0.525$ $p = 0.469$, and $LRT_{11} = 2.582$ $p = 0.275$, respectively), whilst study duration was significant ($LRT_{12} = 19.583$ $p < 0.001$). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias. However, there was evidence of publication bias ($z = 2.720$ $p = 0.007$), with a greater number of less precise studies showing a positive effect than more precise studies (see Additional file 11).

Residual heterogeneity was not significantly reduced by including moderators in the model ($QE_{18} = 62.937$ $p < 0.001$). Figure 36 shows the positive relationship between study duration and difference in SOC.

NT–HT full profile (0–150 cm) No significant effect on soil C stocks was detected for NT versus HT for the full soil profile (Fig. 37), with significant heterogeneity present ($Q_{13} = 568.853$ $p < 0.001$). Climate zone could not be tested due to low sample size. Latitude, soil class, reference SOC, study duration and HT depth category were all significant, however ($LRT_8 = 6.475$ $p = 0.011$, $LRT_6 = 13.719$ $p = 0.001$, $LRT_8 = 9.699$ $p = 0.002$, $LRT_8 = 12.279$ $p < 0.001$, and $LRT_8 = 12.074$ $p < 0.001$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias (see Additional file 11).

Moderators did not reduce the residual heterogeneity in the model significantly ($QE_7 = 17.5621$ $p = 0.014$).

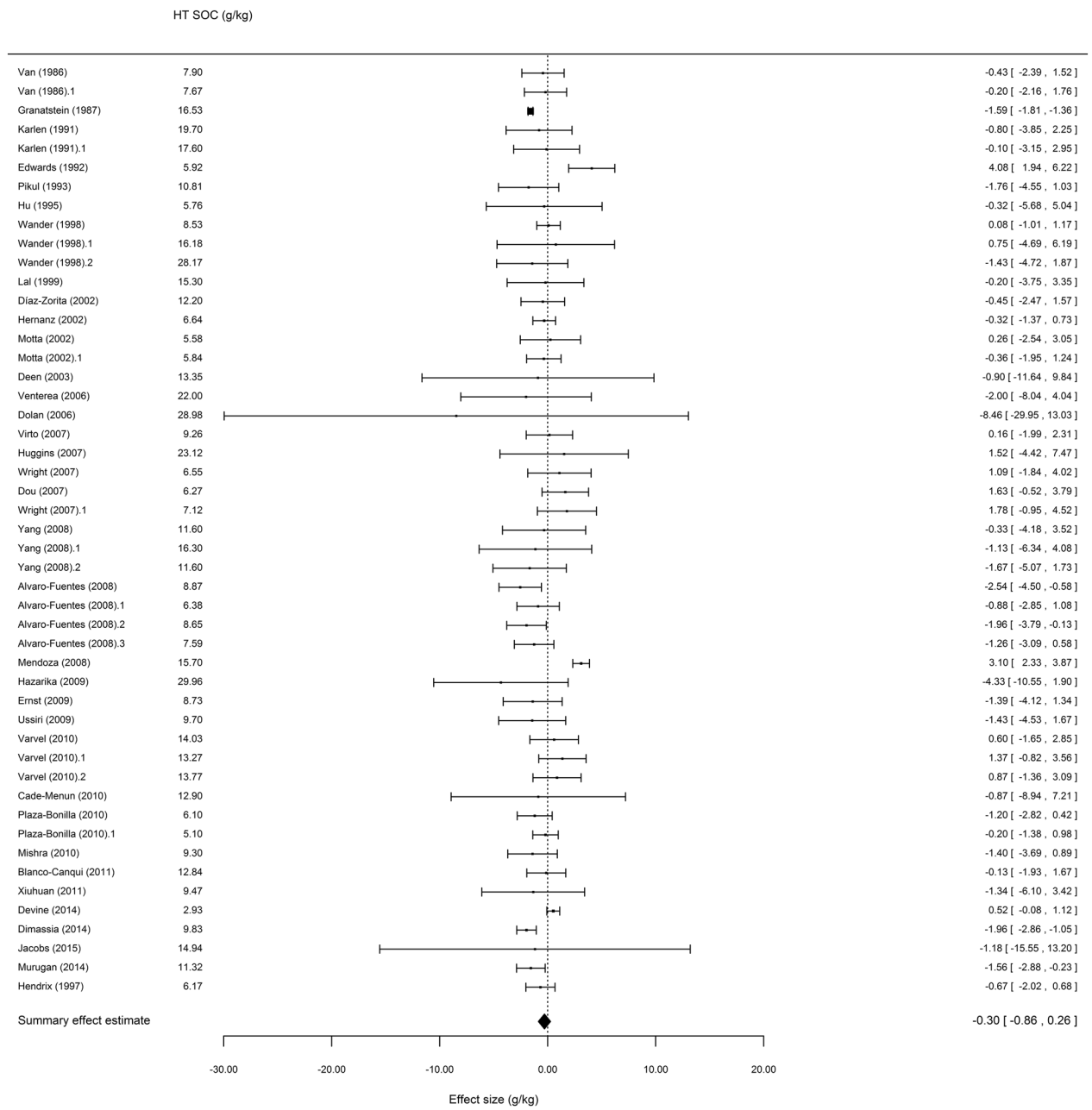

Fig. 17 Forest plot for meta-analysis of concentration data for NT–HT comparison at 15–30 cm depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

Latitude was positively correlated with difference in SOC stocks for the full profile (Fig. 38). The analysis of soil class suffered from a lack of data and low sample size, although data suggest that silty loams (SiLo) had a more positive response that the rest of the evidence base that mostly missed data (Fig. 39). The analysis of HT depth similarly suffered from a low sample size, with significance likely due to spurious differences between deep tillage studies and those missing this information (Fig. 40). The relationship between reference SOC stocks and difference in SOC stocks may be statistically significant but the effect size is very small and may not represent a biologically significant phenomenon (regression line not shown in Fig. 41). Finally, Fig. 42 suggests a

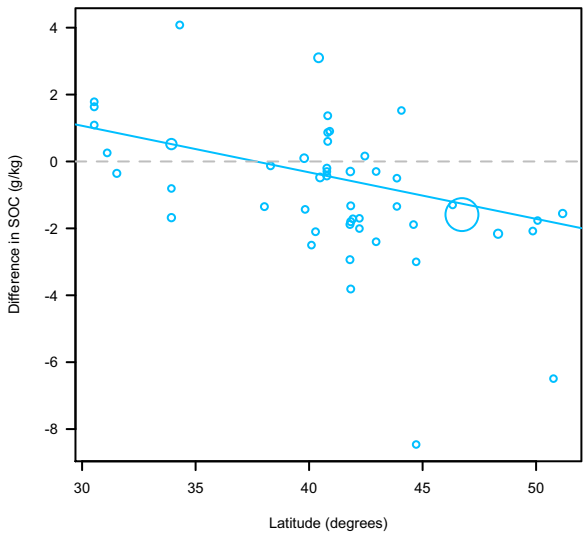

Fig. 18 Meta-regression of SOC concentration against latitude for NT–HT at 15–30 cm. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

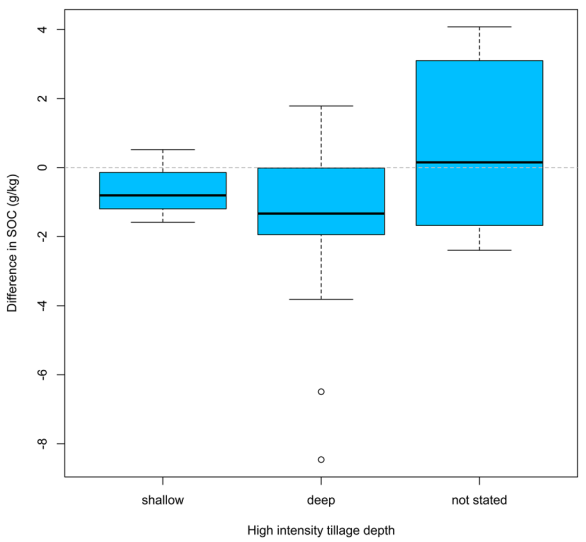

Fig. 20 Boxplot of difference in SOC concentration for NT–HT at 15–30 cm as affected by HT depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

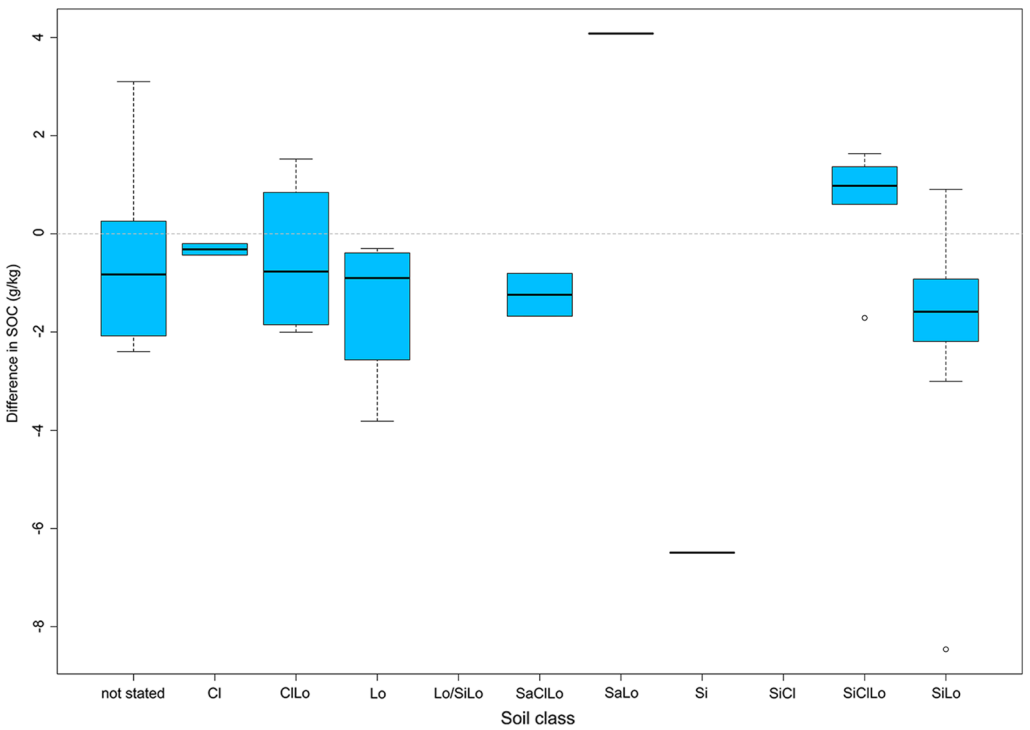

Fig. 19 Boxplots of difference in SOC concentration for NT–HT at 15–30 cm as affected by soil class. *NT* no tillage, *HT* high intensity tillage (see text for explanation). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

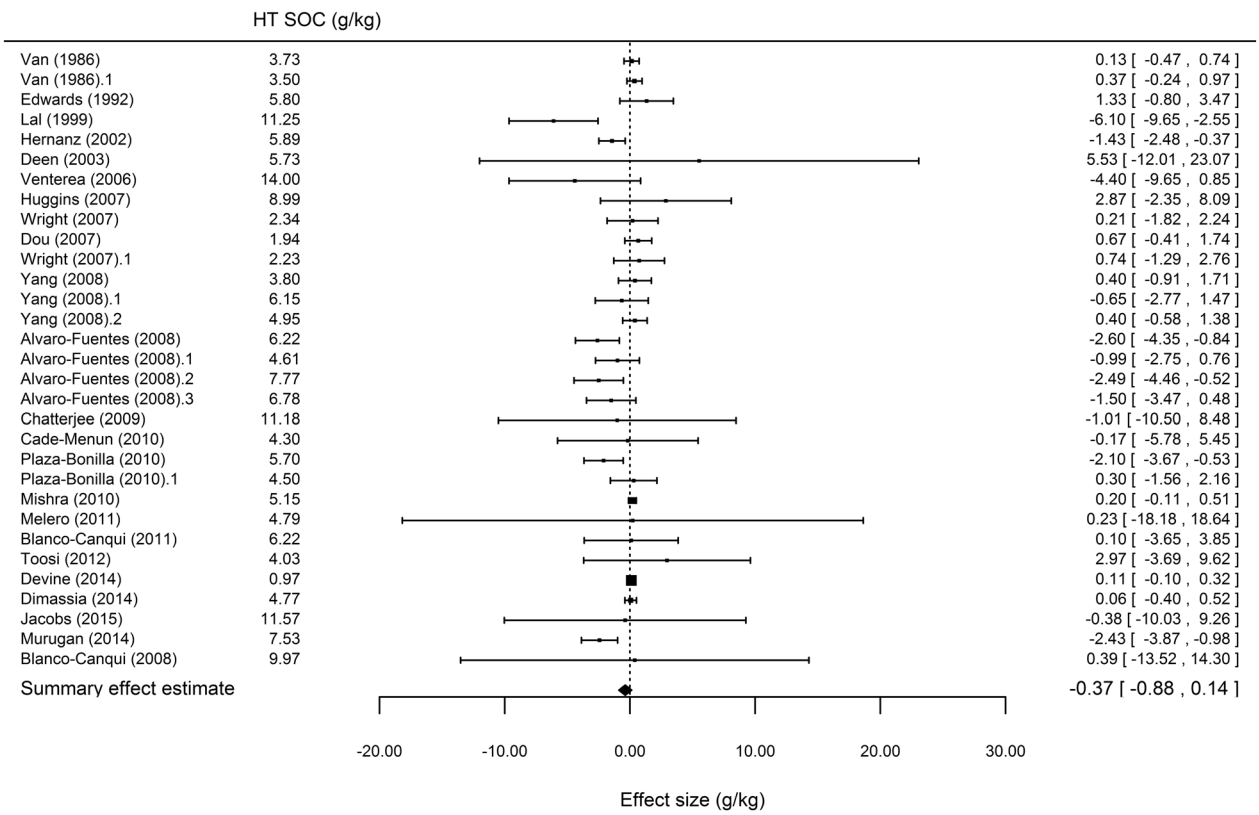

Fig. 21 Forest plot for meta-analysis of concentration data for NT–HT comparison at > 30 cm depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

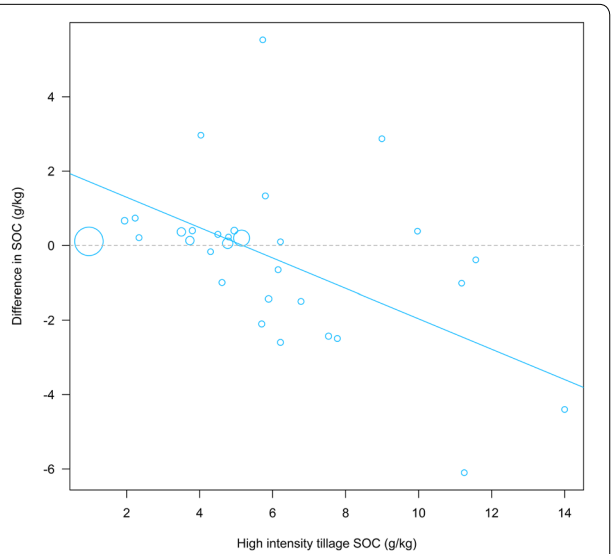

Fig. 22 Meta-regression of SOC concentration against reference SOC for NT–HT at > 30 cm. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

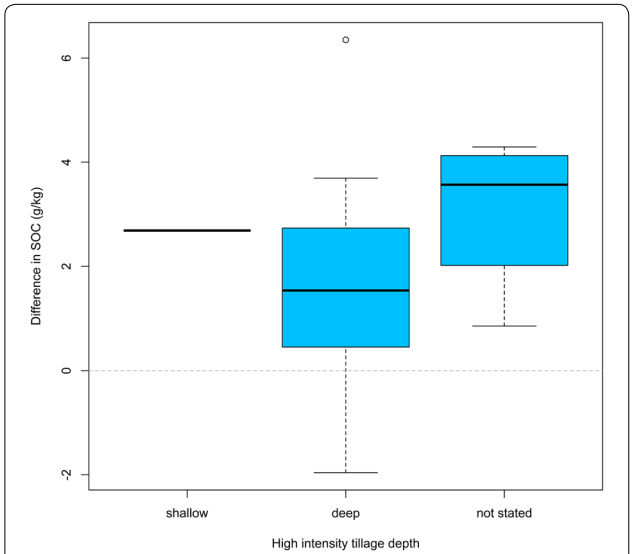

Fig. 23 Boxplot of difference in SOC concentration for NT–HT at > 30 cm as affected by HT depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

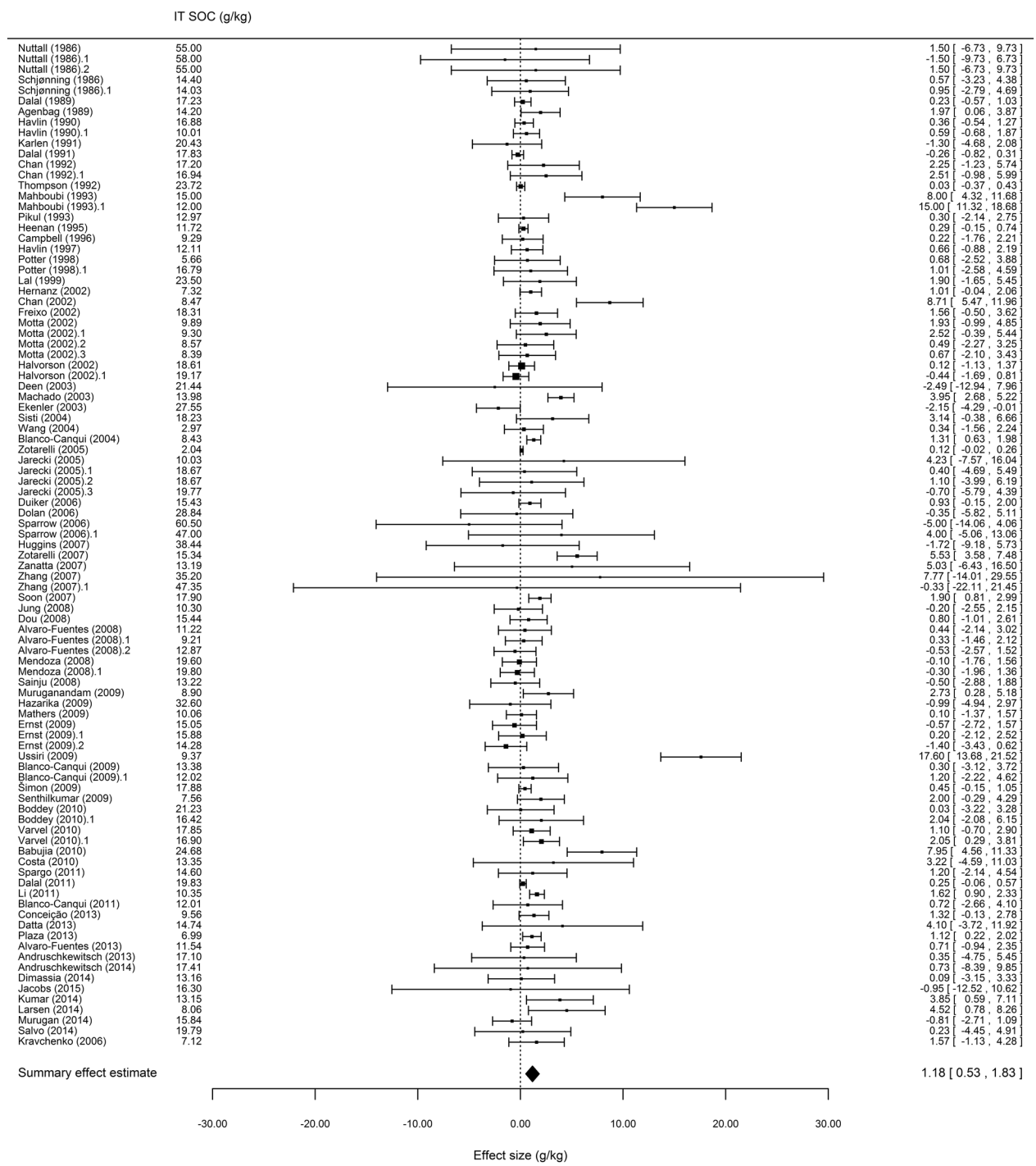

Fig. 24 Forest plot for meta-analysis of concentration data for NT–IT comparison at 0–15 cm depth. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

positive relationship between study duration and difference in SOC stocks, although sample size here is small and the regression line is thus not plotted.

NT–IT upper layer (0–30 cm) An overall significant positive effect estimate was found for NT versus IT soil C stocks for the upper profile (Fig. 43), with significant

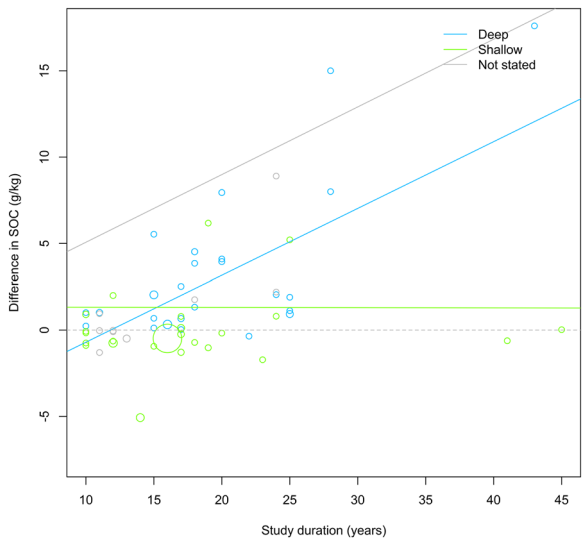

Fig. 25 Meta-regression of SOC concentration against study duration and HT depth category for NT–IT at 0–15 cm. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

heterogeneity present ($Q_{31} = 392.889$ $p < 0.001$). Latitude and climate zone were not significant (see Additional file 11), nor were any of the key moderators study duration, reference SOC stocks, soil class and IT depth category ($LRT_{13} = 3.043$ $p = 0.081$, $LRT_{13} = 2.315$ $p = 0.128$, $LRT_{7} = 3.2924$ $p = 0.857$, and $LRT_{12} = 4.652$ $p = 0.098$, respectively). The sensitivity analysis for critical appraisal category demonstrated no evidence of bias, and there was no evidence of publication bias. However, the sensitivity analysis of high reliability variability data resulted in the loss of significance, likely due to low sample size and high variability in this subset (see Additional file 11). The inclusion of moderators in the model did not remove significant heterogeneity ($QE_{20} = 160.944$ $p < 0.001$), indicating other sources of heterogeneity exist that were not accounted for.

NT–IT full profile (0–150 cm) No significant pattern was identified across the evidence base for SOC stocks in NT versus IT for the full soil profile (Fig. 44), although significant heterogeneity was present ($Q_{12} = 555.316$ $p < 0.001$). Latitude and climate zone were not significant (see Additional file 11), nor was reference SOC ($LRT_{9} = 0.528$ $p = 0.467$). Study duration, soil class and IT depth category were significant, however ($LRT_{9} = 19.816$ $p < 0.001$, $LRT_{6} = 18.327$ $p < 0.001$, and $LRT_{8} = 8.436$ $p = 0.015$, respectively). The sensitivity analysis for critical appraisal category demonstrated no evidence of bias, and there was no evidence of publication bias. However, the sensitivity analysis of high reliability variability data resulted in a significant effect estimate due to extremely low sample size (see Additional file 11).

Inclusion of moderators in the model explained the significant heterogeneity ($QE_{5} = 3.063$ $p = 0.690$). Figures 45, 46, and 47 show the relationships between difference in SOC stocks for the full soil profile and study duration, IT depth and soil class, respectively. Due to low sample size in certain subgroups (e.g. deep IT), these results should be viewed with caution (no regression lines have been plotted, accordingly). Longer studies are associated with more positive differences in SOC, and clay (Cl) and clay loam (ClLo) soils appear to show positive and negative impacts on SOC stocks for the full soil profile of a switch to NT from IT, respectively. The significant pattern in IT depth is likely driven by the large body of evidence that does not state tillage depth.

IT–HT upper layer (0–30 cm) No significant effect estimate was found for the model of SOC stocks in IT versus HT in the upper layer (Fig. 48), although significant heterogeneity was present ($Q_{28} = 285.388$ $p < 0.001$). Latitude and climate zone were not significant (see Additional file 11), nor was reference SOC ($LRT_{13} = 1.572$ $p = 0.210$). Soil class, study duration, HT depth category and IT depth category were all significant, however ($LRT_{7} = 28.893$ $p < 0.001$, $LRT_{13} = 4.633$ $p = 0.031$, $LRT_{13} = 4.946$ $p = 0.026$, $LRT_{12} = 10.857$ $p = 0.004$, respectively). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias (see Additional file 11).

The inclusion of moderators in the model accounted for the significant heterogeneity ($QE_{17} = 7.968$ $p = 0.967$). Figure 49 shows soil classes and SOC stock difference for the upper layer, suggesting that loamy sands (LoSa) and silty loams (SiLo) showed a more positive response that other soil types. Study duration was positively correlated with difference in SOC, although the power of this analysis was low due to a relatively small sample size (regression line not plotted in Fig. 50). Figure 51 suggests that a conversion from deep HT may produce a greater difference in SOC, although there was a lack of shallow HT studies for this depth. Conversion to deep IT, however, appears to result in SOC loss, whilst conversion to shallow IT has a positive effect on SOC (Fig. 52).

IT–HT full profile (0–150 cm) No significant overall summary effect was detected for IT versus HT SOC stock for the full soil profile (Fig. 53), although significant heterogeneity can be observed ($Q_{9} = 83.835$ $p < 0.001$). Latitude and climate zone were not significant (see Additional file 11), nor were reference SOC stock and IT depth category ($LRT_{7} = 0.754$ $p = 0.385$ and $LRT_{7} = 0.101$

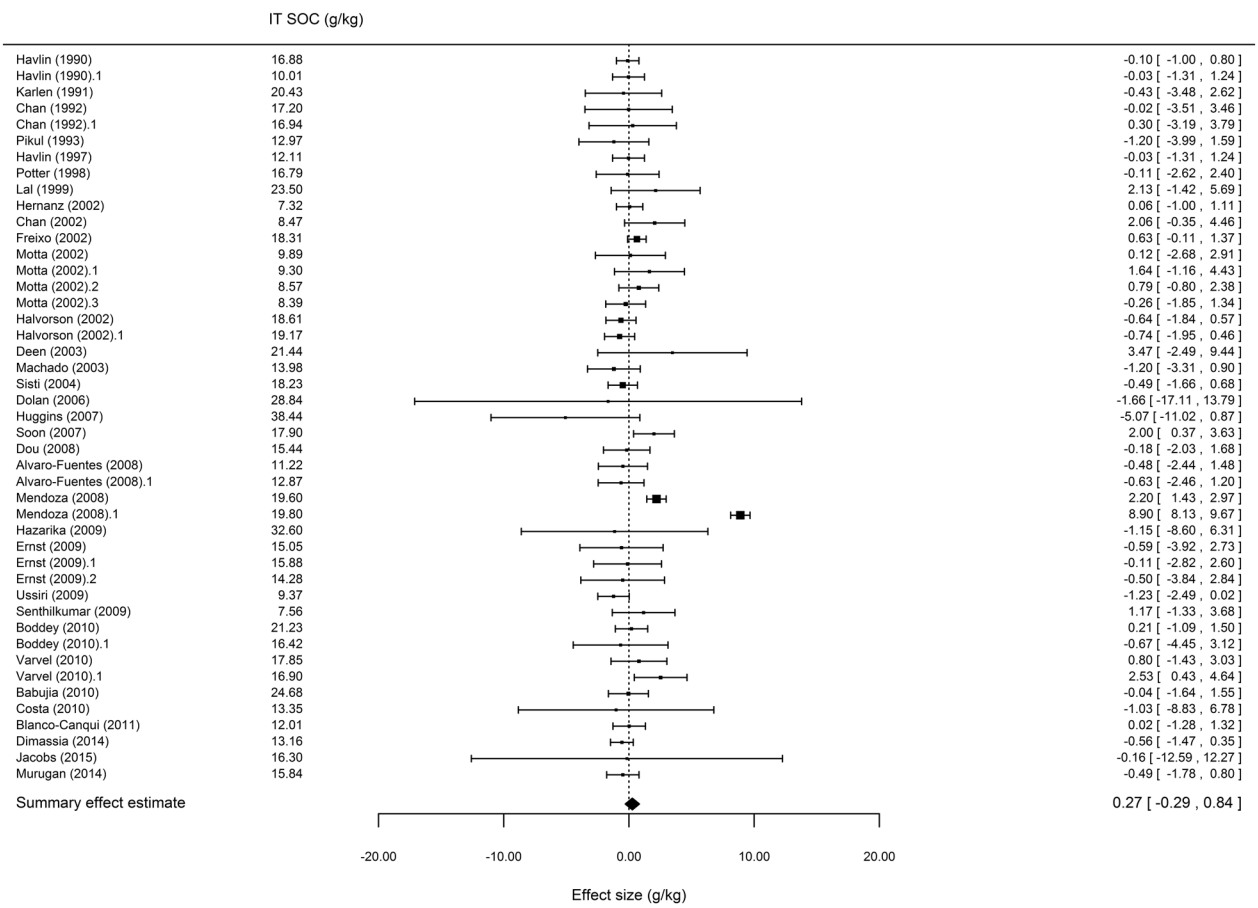

Fig. 26 Forest plot for meta-analysis of concentration data for NT–IT comparison at 15–30 cm depth. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

$p = 0.750$, respectively) (HT depth category could not be tested due to low sample size). Soil class and study duration were significant, however ($LRT_6 = 9.847$ $p = 0.002$ and $LRT_7 = 14.312$ $p < 0.001$). Sensitivity analyses for critical appraisal category and variability type demonstrated no evidence of bias, and there was no evidence of publication bias (see Additional file 11).

Residual heterogeneity in the stock data for the full soil profile was accounted for by including moderators in the model ($QE_4 = 0.363$ $p = 0.985$). Figure 54 suggests that silty loams (SiLo) may have a negative effect size, whilst other soils are generally positive ('not stated' soil types). Figure 55 suggests a positive relationship between study duration and difference in SOC, however sample size in this meta-regression is low and one study is particularly influential, suggesting that these results should perhaps be viewed with caution (regression line not plotted).

Discussion

Review findings in the context of existing knowledge

This meta-analysis showed that NT has higher SOC concentration and SOC stocks in the top layer (0–15 cm) of soil compared to HT and IT. It also showed that NT increased SOC stocks for the upper layer (0–30 cm) compared to HT. Yet C stocks for the full soil horizon (0–150 cm) were similar between all compared tillage types. The transition of tilled croplands to NT and conservation tillage has been credited with substantial potential to mitigate climate change via C storage [31, 52, 53]. Changes in C stock due to management via reduced tillage has been estimated to be around 0.4 Mg/ha per year in the US [54]. However, based on our results, the level of C stock increase under NT compared to HT was in the upper soil around 4.6 Mg/ha (0.78–8.43 Mg/ha, 95% CI) during a minimum of 10 years, while no effect was detected in the full horizon.

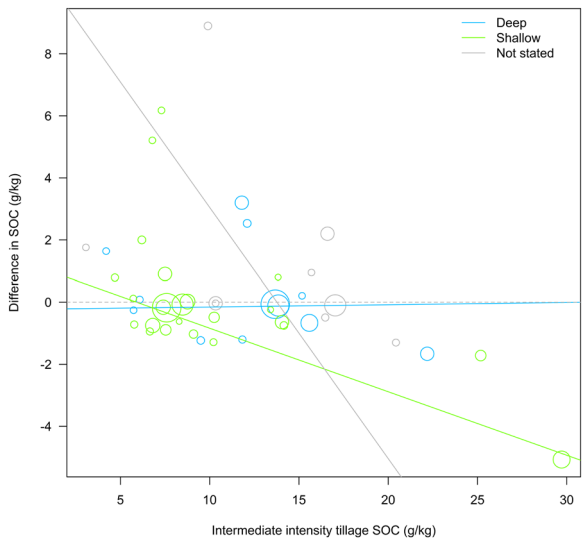

Fig. 27 Meta-regression of SOC concentration against reference SOC and IT depth category for NT–IT at 15–30 cm. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

Comparison of results across soil depths

Only 66 studies of the 351 studies (19%) in this meta-analysis sampled soil below 30 cm, and relatively few studies (32%) sampled below 15 cm. The predominance of data from the soil surface layer helps to explain the excitement for the potential for C storage in soil. Although the surface soil can rapidly accumulate SOC and microbial C with NT [29, 55, 56], the C inputs below the surface layer is less clear. Root density has been shown to be greater under NT down to 30 cm [57], and to be restricted below 15 cm compared to conventional tillage, possibly due to factors such as compaction and lower temperatures [31]. NT and conservation tillage potentially produce benefits that result from soil C accumulation in the surface soil, such as improved infiltration, water-holding capacity, erosion reduction, nutrient cycling and soil biodiversity [53]. Any effects of greenhouse gas mitigation by NT and IT can also be caused by indirect factors such as lower fossil fuel consumption in tillage and water transport, and less demand for synthetic N fertiliser with its energy demands and potential for nitrous oxide emissions [30].

Certain conditions may be more conducive to SOC accumulation under NT or IT. The meta-analysis indicates that for soils with a low starting SOC concentration, NT is more likely to increase SOC below 30 cm, as compared to HT. A higher starting SOC concentration makes for greater SOC loss at 15–30 cm with shallow IT than NT. In a C-depleted soil (e.g. a soil with 10 g SOC/kg soil) a small SOC input into the soil profile sequestered by roots and organisms will become a detectable difference, while the same addition of SOC in a soil with an initially higher SOC level (e.g. 40 g/kg) will give a relatively lower increase of SOC.

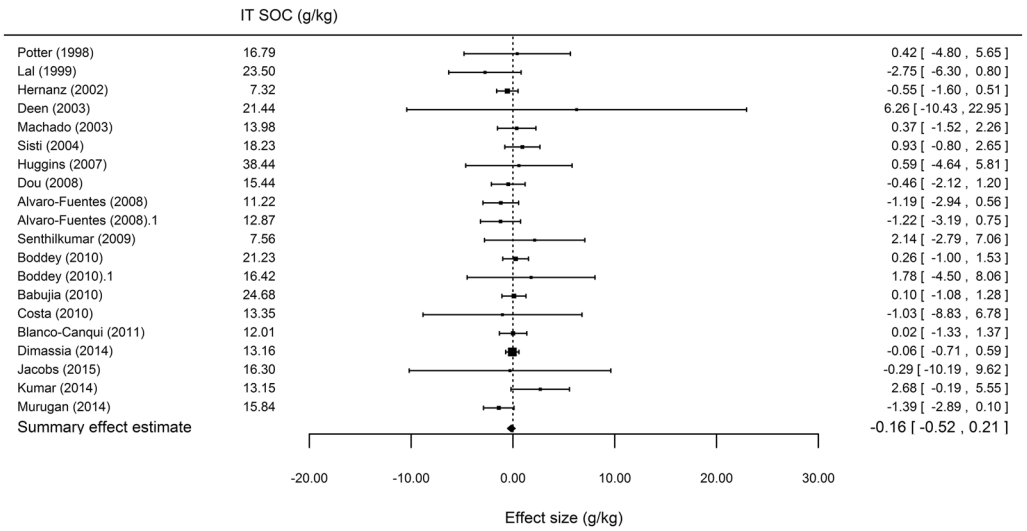

Fig. 28 Forest plot for meta-analysis of concentration data for NT–IT comparison at > 30 cm depth. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

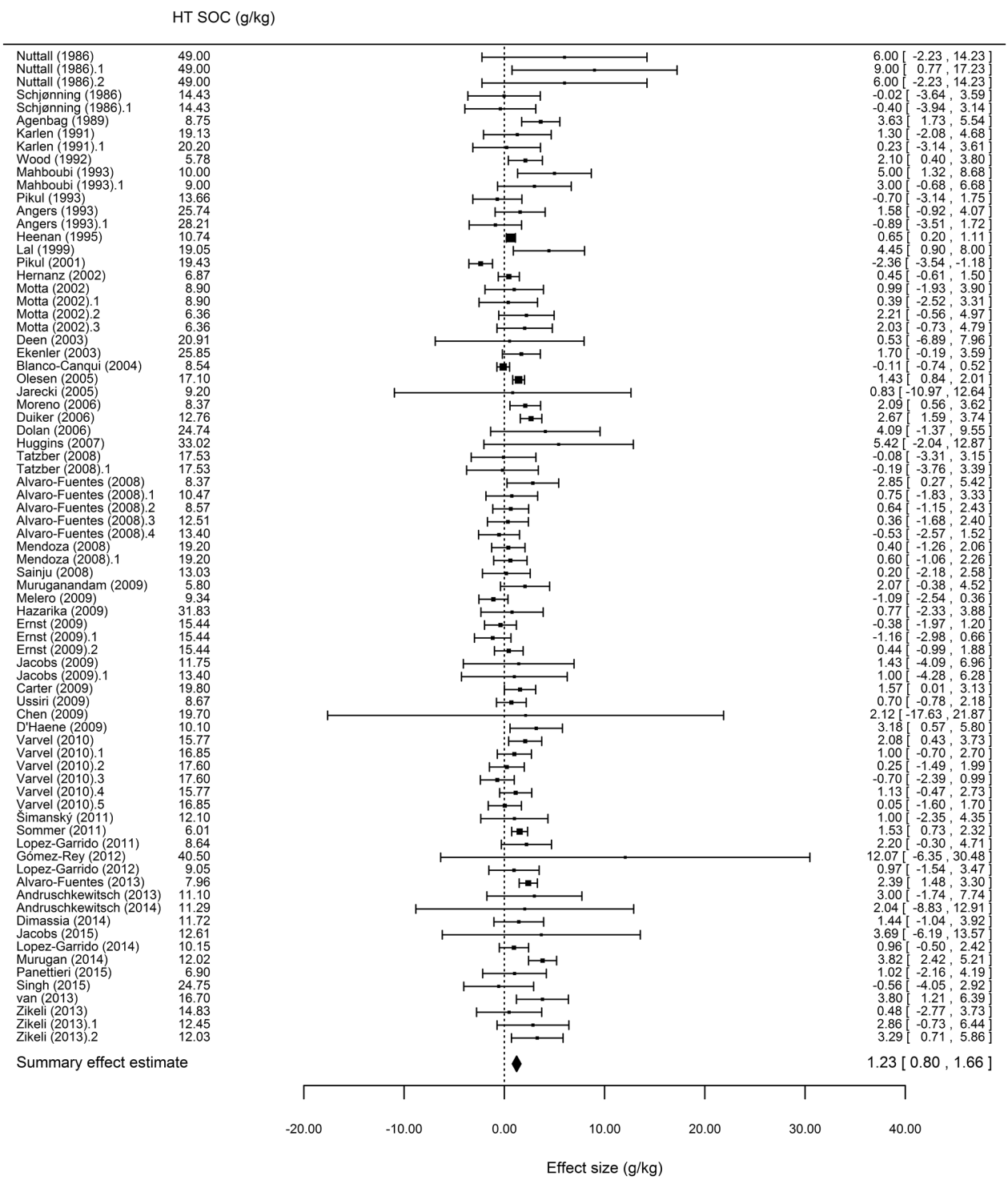

Fig. 29 Forest plot for meta-analysis of concentration data for IT–HT comparison at 0–15 cm depth. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

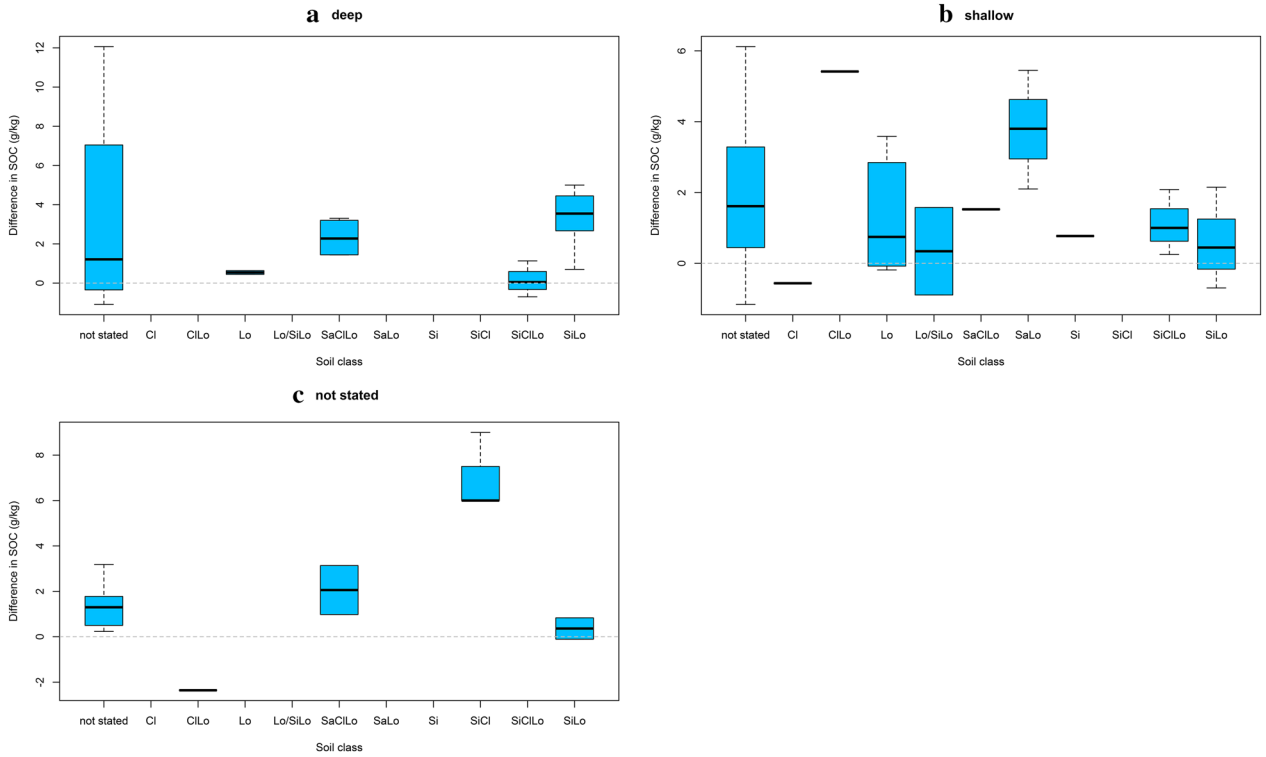

Fig. 30 Boxplots of difference in SOC concentration for IT–HT at 0–15 cm as affected by soil class. IT depth categories shown are: **a** deep, **b** shallow, and **c** not stated. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

Reasons for heterogeneity

The starting premise of this review was to include studies of more than 10 years' duration to ensure that treatment differences would be detected [33]. Analysis of relationships between study duration and SOC concentrations and stocks in the upper layers of soil confirmed that 10 years was indeed a valid minimum intervention period. For deeper soil depths, study duration was not consistently associated with SOC concentration, possibly due to greater heterogeneity among studies, or to different rates of accumulation deeper in the profile.

Soil type did not influence the effects of tillage on SOC stocks and SOC concentrations from 0 to 15 cm, however deeper down (15–30 cm) SOC concentrations had a larger increase in sandy clay loam and silty clay soils under NT compared to HT. Those soil types have, on average, a clay content of about 30 and 45%, respectively, which may help to slowdown SOC decomposition compared to coarser soils [58, 59]. This is related to the fact that clay particles can help to stabilise decomposing litter by mineral associated bonds [1, 2] and the aggregation is stronger, also promoting physical inaccessibility of SOC to the microbial community [3]. Climate zone did not affect the relationship between tillage and SOC, but as there was a limited range of sites within the boreo-temporal regions, this may not have been sufficiently variable to yield significant differences. However, site latitude was positively correlated to differences in full profile C stocks. Whether this is dependent on a lower decomposition rate at higher latitudes due to lower temperatures could be possible but the rates are also determined by interactions of a number of physical and chemical factors influencing the microbial enzymatic activities in soils [59].

A comparison of stocks and concentration data

Many of the long-term studies considered in this systematic review were set-up when climate change was not considered a significant problem or only an emerging issue. The focus was likely more oriented towards crop productivity, soil quality and environmental aspects of different management systems [60]. Within this view, SOC was considered as the most important indicator of soil quality and agronomic sustainability due to its impact on physical, chemical and biological properties [61]. In fact, half of the studies of this systematic review reported only C concentration (e.g. g/kg or %), corroborating the

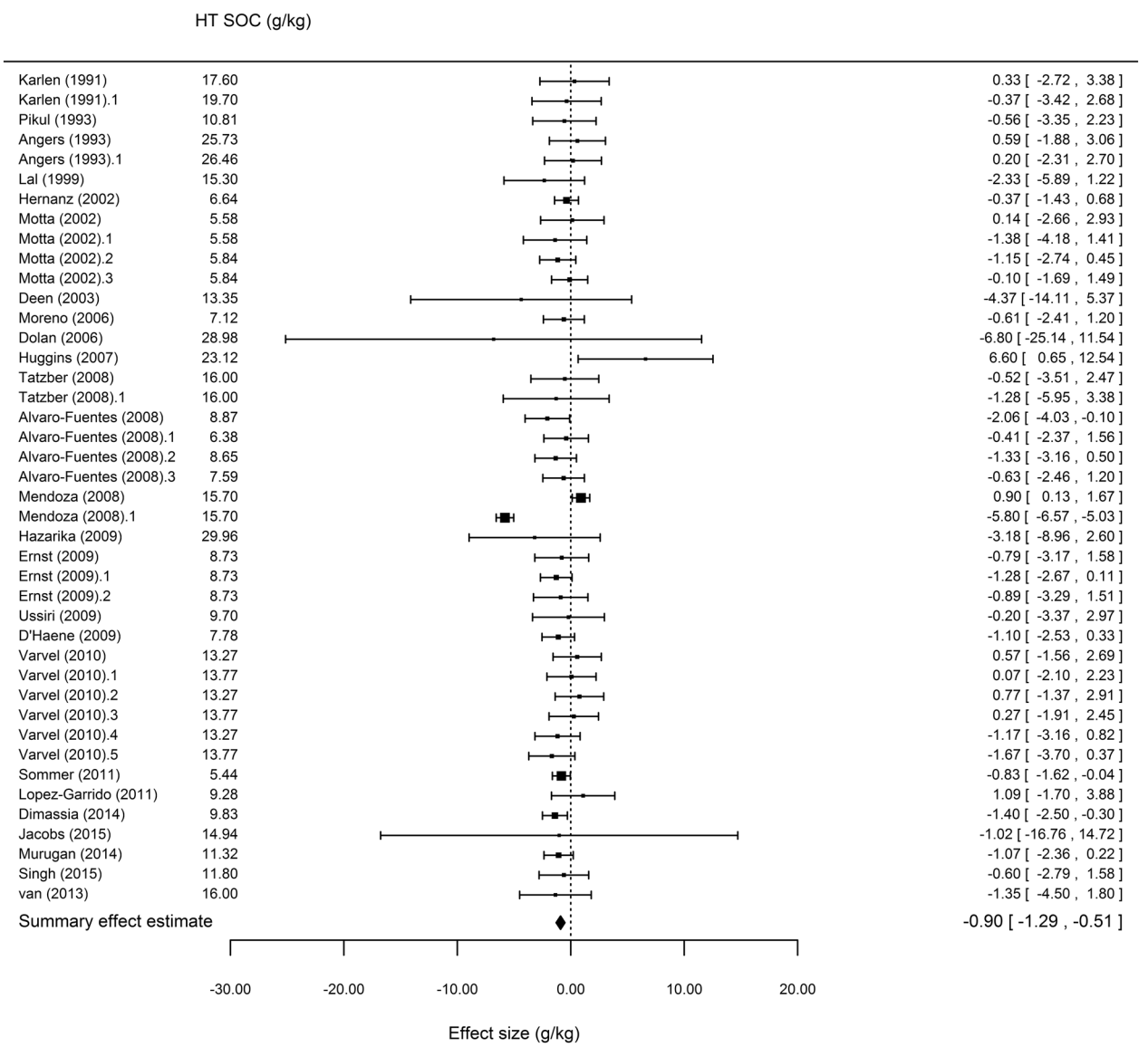

Fig. 31 Forest plot for meta-analysis of concentration data for IT–HT comparison at 15–30 cm depth. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

requirement for addressing soil quality and reducing, at the same time, the cost and time necessary to carry out the additional bulk density sampling and analysis.

However, SOC concentration alone may be less adequate if the focus is on a quantitative SOC balance, such as is necessary for assessments of carbon sequestration capacity for climate change mitigation. In particular, when the management under investigation could significantly alter soil density, as is the case for tillage interventions in general [62], bulk density becomes a fundamental parameter for accurately calculating SOC stock. Bulk density measurements undoubtedly give more transparency to the experimental results but may not guarantee the greatest accuracy, if depth is not properly considered. For example, soils with the same SOC concentration but with a different density as a result of different tillage regimes may be erroneously considered to have different SOC stock if the same depth is considered.

In much past research, most of the comparisons among treatments were made simply by multiplying SOC concentration with bulk density, considering a fixed depth. This method often introduces significant errors when

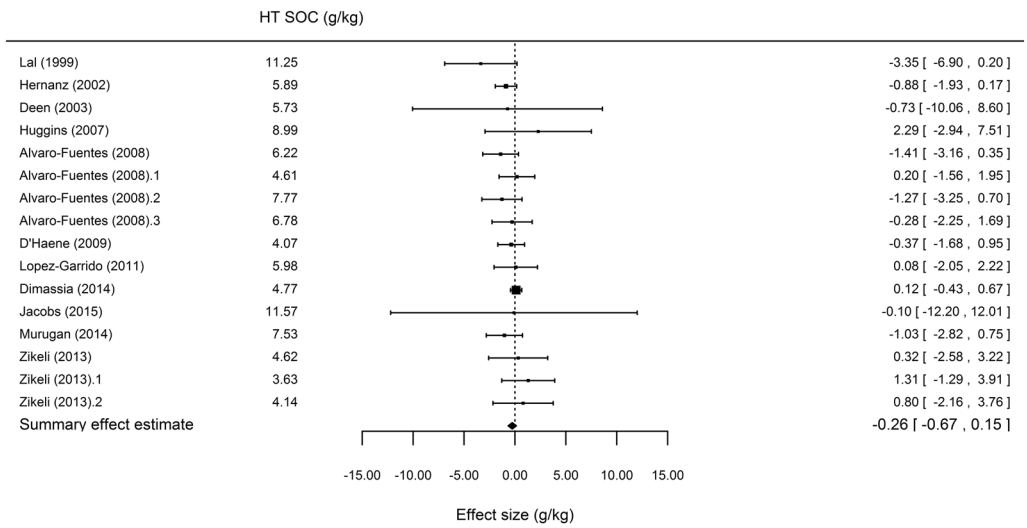

Fig. 32 Forest plot for meta-analysis of concentration data for IT–HT comparison at > 30 cm depth. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

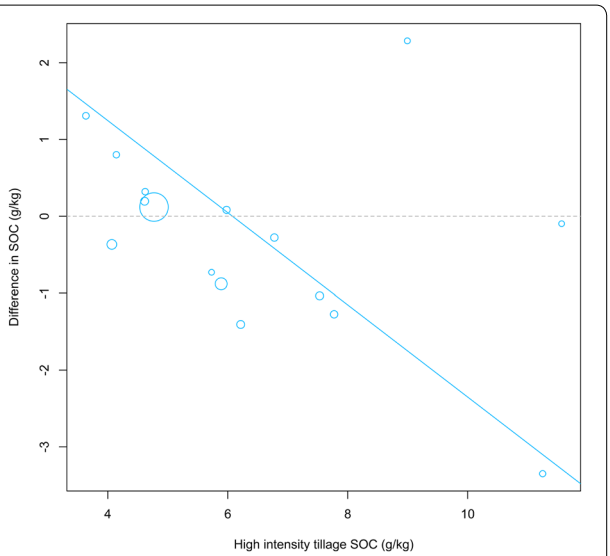

Fig. 33 Meta-regression of difference in SOC concentration for IT–HT against reference SOC at > 30 cm. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

soil bulk density differs among treatments under study, such as between tillage and no-tillage [63, 64]. In order to undertake more rigorous quantitative SOC estimations, both the bulk density measurement and calculations based on equivalent soil mass (ESM) should be reported [65, 66]. Furthermore, a similar but simpler approach based on cumulative mass could be considered, in which C density is reported for a fixed mineral mass per unit area [67]. Although the latter methods are formally more accurate than a simple comparison of concentrations to detect (and quantify) differences on SOC, they introduce further uncertainty associated with all the parameters needed for calculation; SOC, bulk density, depth and gravel content errors, coming from different sources (e.g. sampling, analysis, etc.), which propagate non-linearly [68]. This is likely the reason why the confidence intervals of SOC differences in the meta-analysis are proportionately much larger with stocks than concentrations.

Direct and indirect effects of tillage on soil functions and crop growth

Minimum or no tillage practices have also been introduced as a mitigation measure for erosion control. The experimental sites included in this systematic review were assumed to represent either stable soil conditions or a situation where eventual lateral transport of soil did not disproportionally affect experimental treatments. This assumption may be a source of bias since the mulch layer under NT conditions may have reduced erosion at alluvial positions or increased deposition at colluvial positions in the landscape compared to tilled treatments. The implications of soil erosion for carbon cycling are not straightforward [69]. Although soil erosion is a major threat to soil fertility and food security [70], it may actually lead to higher carbon retention at the landscape scale [71]. Thus, observed treatment-induced changes in SOC should not be translated directly into net transfer

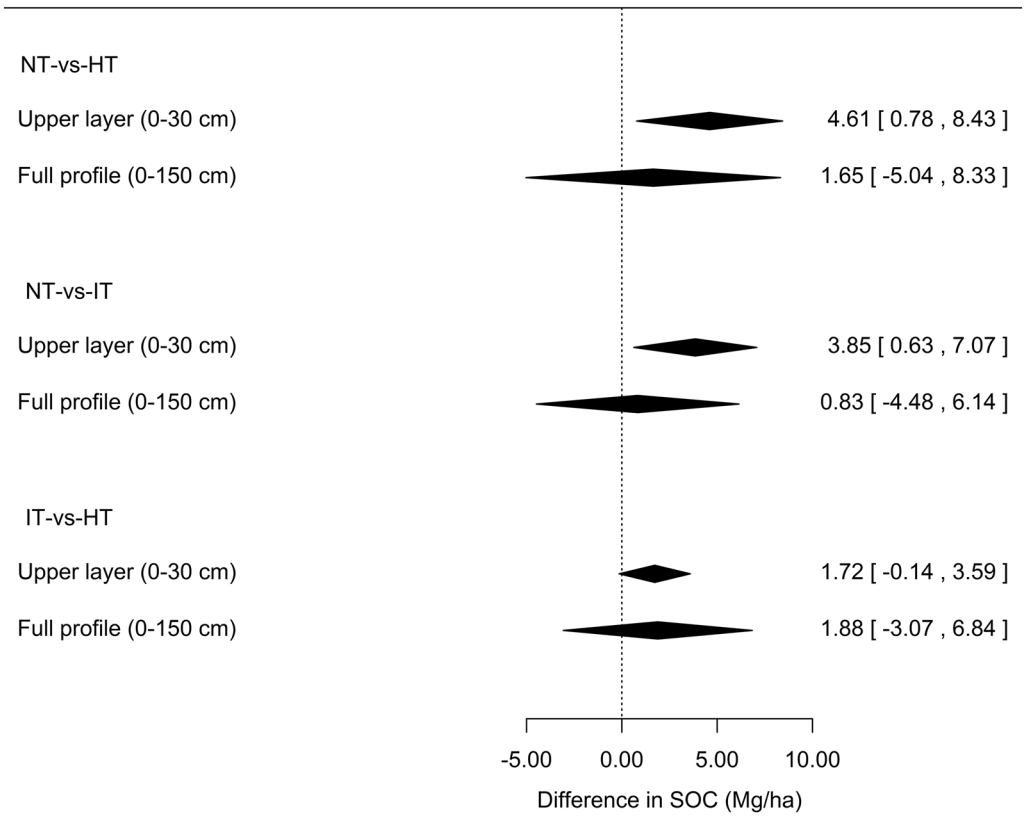

Fig. 34 Summary effect estimates (difference in SOC, Mg/ha) for stocks data meta-analyses. Three tillage comparisons are shown: NT, no tillage; *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Diamonds are centred on the summary effect estimate for each meta-analyses, with the points of the diamonds representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

Table 9 Summary effect estimates for meta-analyses of stocks data

Tillage comparison	Soil depth (cm)	Summary effect estimate (Mg/ha)	Standard error	z statistic	p value	Lower 95% confidence interval	Upper 95% confidence interval	Number of studies
NT–HT	Upper layer	4.606	1.954	2.357	0.018	0.776	8.435	29
	Full profile	1.646	3.409	0.483	0.629	− 5.035	8.327	14
NT–IT	Upper layer	3.852	1.644	2.343	0.019	0.630	7.075	32
	Full profile	0.831	2.711	0.306	0.759	− 4.483	6.144	13
IT–HT	Upper layer	1.724	0.952	1.811	0.070	− 0.142	3.589	29
	Full profile	1.883	2.529	0.745	0.457	− 3.074	6.840	10

'Upper layer' corresponds to measurements between 0 and up to 30 cm depth, whilst 'full profile' corresponds to measurements taken between 0 and between 30 and 150 cm depth

NT no tillage, *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation)

of atmospheric CO_2 to SOC, i.e., climate mitigation, at larger scales beyond single fields.

Crop yield is also affected by tillage and has, in a recent review, been shown that in order to maintain or increase yields reduced tillage needs to be combined with other management activities. Such practices include soil coverage by plants or returning residues to fields, otherwise low tillage can give lower yields [72]. To get a more holistic view of the effects of tillage on potential trade-offs between SOC accumulation and crop production we

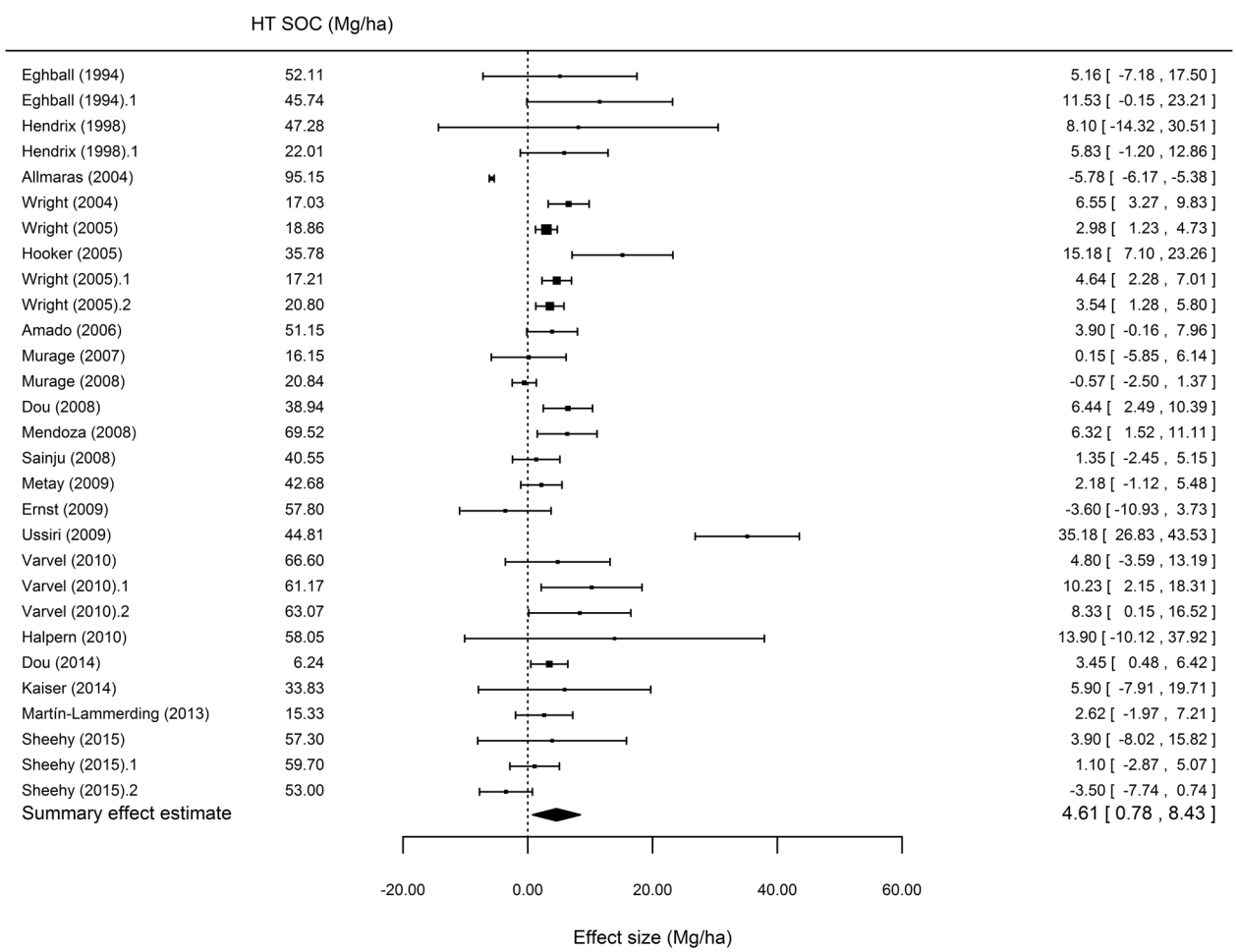

Fig. 35 Forest plot for meta-analysis of stock data for NT–HT comparison in upper layer. *NT* no tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

plan to investigate the evidence for yield effects in our database in a meta-analysis of yields.

From the perspective of climate change mitigation, any benefit of increased SOC should be considered together with components of greenhouse gas production that may differ between tillage treatments, such as emissions related to the fuel needed for field operations or the production of fertilisers and pesticides. Nitrous oxide (N_2O) is the greatest contributor to greenhouse gas emissions from crop production where the soil water content, nitrate concentrations and available carbon are the major determinants regulating emission rates. Temporary water logging due to high bulk density or insufficient drainage is considered to have a great influence on N_2O emissions in humid climates, as this will provide temporary anaerobic conditions where nitrate will be turn into N_2O by denitrifying bacteria in the soil [73]. Therefore, higher N_2O emissions are suggested to occur where bulk density values are higher, due to moister and denser soil conditions, which may eventually offset positive effects on SOC balances [22, 23]. There is no compelling evidence for changes in bulk density resulting from tillage, since some authors observe no changes whilst others find lower bulk density with increased SOC levels [74–77]. An increase in soil bulk density may offset positive effects on SOC balances, since more greenhouse gases including N_2O may be produced, for example due to anaerobic conditions [22, 23]. This potential negative climate impact may however be counteracted considerably by introducing controlled traffic farming, which will give lower bulk densities [78].

It is unclear whether observed effects of tillage treatments are mainly input or output (decomposition) driven. The increase in respiration after tillage treatment

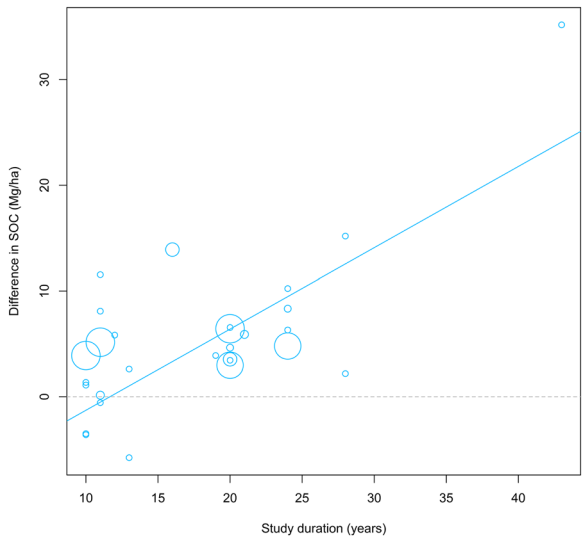

Fig. 36 Meta-regression of difference in SOC stock for NT–HT against study duration in upper layer. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

observed in numerous studies has often been ascribed to the disruption of soil aggregates, whereby occluded particulate organic material becomes available to decomposers [e.g. 79]. However, changes in soil moisture and temperature and treatment-specific distribution of crop residues have been found to be highly important [e.g. 80–83]. According to a meta-analysis conducted by Virto et al. [32] differences in SOC stocks between NT and inversion tillage were significantly and positively correlated with differences in crop yields. Thus, they concluded that the observed effect on SOC was indirect and governed mainly by the crop production response to tillage treatment. Thus, the evidence is still not conclusive whether losses of C through decomposition or yield effects are the main drivers for observed differences in SOC between tillage treatments.

Input by crop roots, their corresponding carbon allocation and the soil organism communities are considered as the major carbon sources in all soil layers [84, 85]. Soil organisms in particular are affected by tillage, for example earthworms and arbuscular mycorrhizal fungi [86, 87]. Less intensive tillage can promote the soil organism

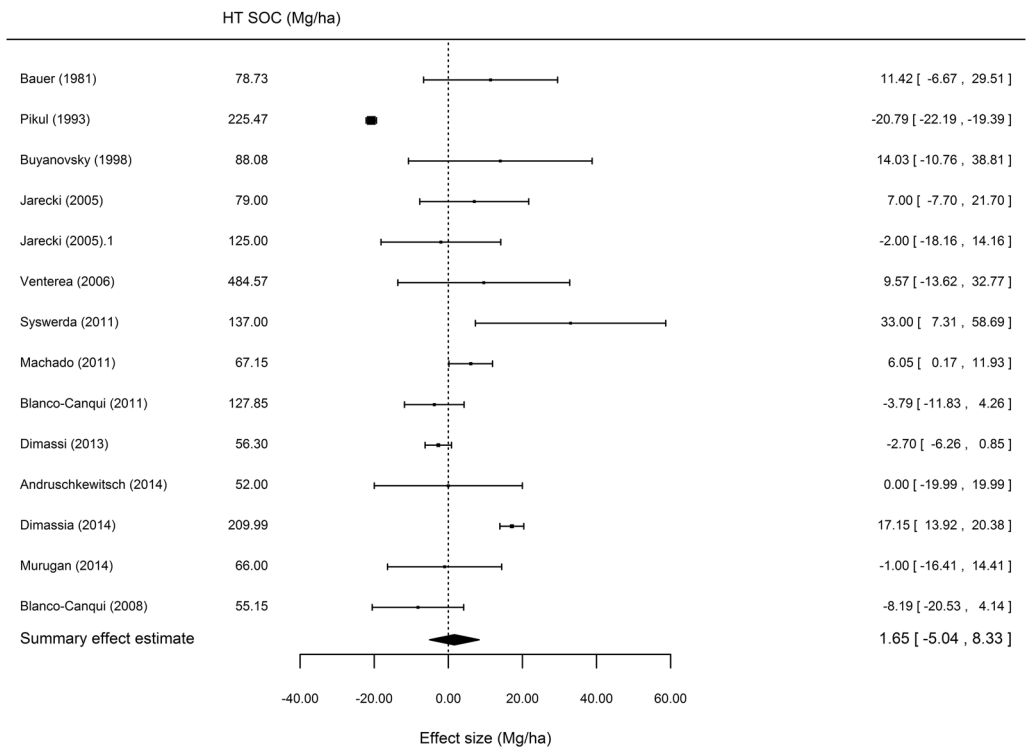

Fig. 37 Forest plot for meta-analysis of stock data for NT–HT comparison in full soil profile. *NT* no tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

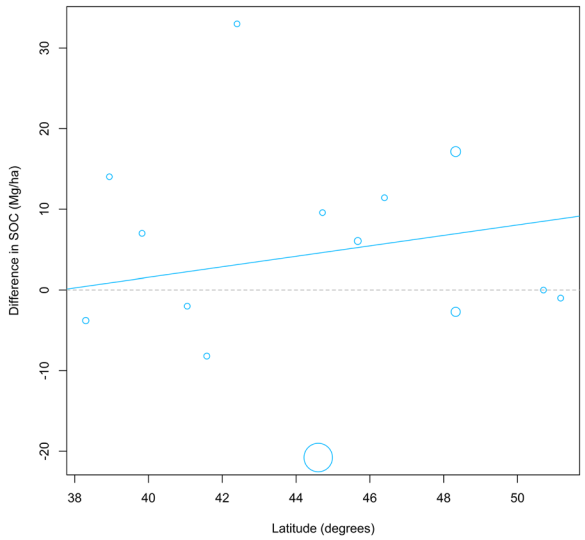

Fig. 38 Meta-regression of difference in SOC stock for NT–HT against latitude in full soil profile. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

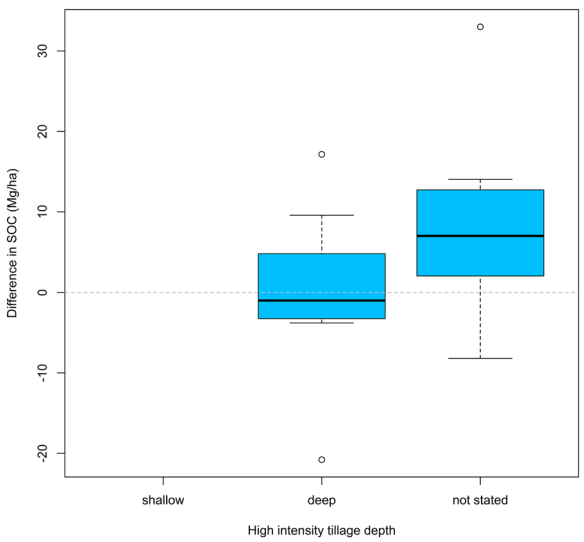

Fig. 40 Boxplot of difference in SOC stock for NT–HT in full soil profile as affect by tillage depth. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

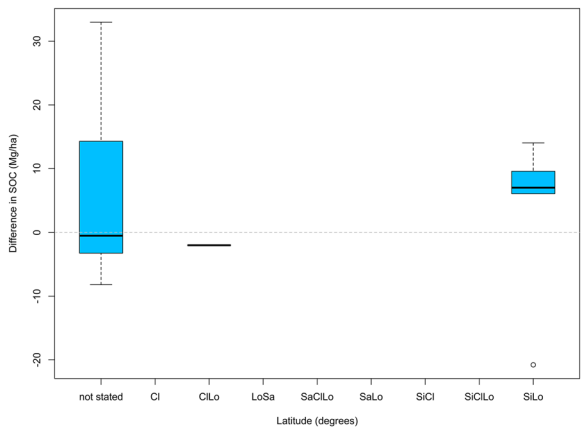

Fig. 39 Boxplot of difference in SOC stock for NT–HT in full soil profile as affected by soil class. *NT* no tillage, *HT* high intensity tillage (see text for explanation). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

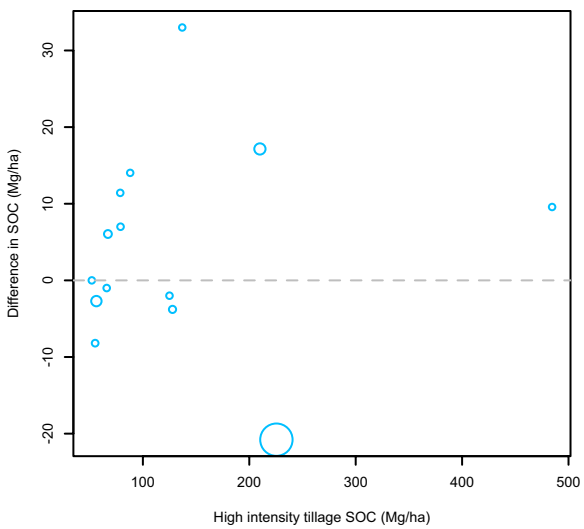

Fig. 41 Meta-regression of difference in SOC stock for NT–HT against reference SOC in full soil profile. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

communities by increasing the fungal-based parts of the soil food webs, which reduces leaching of nutrients and losses of soil carbon [88]. It has been proposed that the fungal based webs contribute more to soil C sequestration than bacterial-based soil food webs that are present at intensive management [89]. Furthermore, it has been suggested that the biomass of fungal communities also contributes substantially to the sequestration of soil C [90].

Review limitations

Limitations of the review

Our review involves a considerable number of meta-analyses, mostly consisting of a large number of studies (up to 102 studies). Some meta-analyses were based on a low sample size, however (as low as 10 studies) and a relatively low sample size for models with a complex

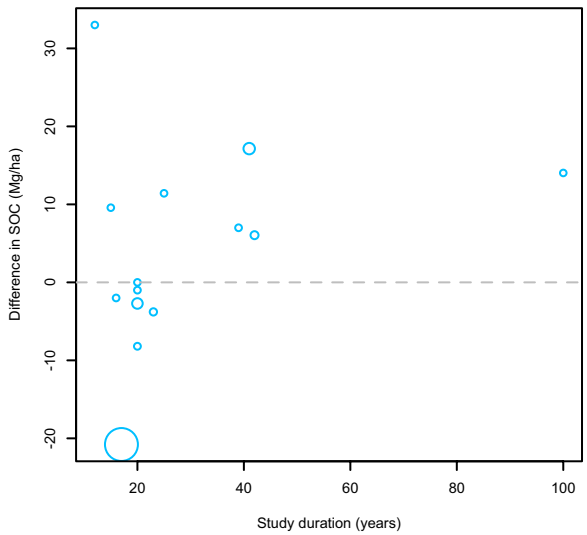

Fig. 42 Meta-regression of difference in SOC stock for NT–HT against study duration in full soil profile. *NT* no tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

structure of moderators. Relative to other meta-analyses, these tests are large [e.g. 91, 92]. Still, the robustness of some of our smaller models would be improved could studies missing data be included and as more research is published over time. Cumulative meta-analysis suggests this may not be necessary for the larger meta-analyses, however.

Whilst we have attempted to account for various moderators in our analyses, we have often run the risk of over parameterisation. We have chosen to be transparent and supply results for both basic (unmoderated) and moderated models, but the risk of over-parameterisation would be reduced in future as more research is published, particularly where information is richer, for example soil texture data, allowing a greater proportion of the evidence base to be included in complex analyses. Similarly, we have not removed outliers, but we have plotted influential studies. Since out meta-analyses are relatively large, the influence of single studies is unlikely to be unacceptably large. We appreciate that another approach could have been to remove outliers and repeat analyses, but we felt that transparency about these analyses was more appropriate than removing studies based on their influence.

It was not possible with the available resources and the volume of evidence to assess the effect of combining tillage with other interventions, such as amendments, crop rotation or fertiliser. Some 49% of the studies in the evidence base involved such factorial or combined analyses, and further investigation of these 172 studies would provide useful insights for practitioners attempting to reduce SOC loss from their soils.

Limitations of the evidence base

Due to the volume of evidence that we have encountered relating to the impacts of tillage on SOC the search update has taken 9 person months to screen, critically appraise, extract data from and integrate into the ongoing synthesis of evidence from the existing systematic map. In addition, the high publication rate of relevant research over the past 2 years (20% of the total evidence base across a 27 year history) indicates that evidence will continue to be published at this rate or higher in the coming years. Together, these facts mean that future syntheses could struggle to bring together the rapidly expanding body of evidence in an affordable, timely manner: review updates would essentially involve a similar investment of resources as many other smaller systematic reviews. Furthermore, the length of time needed to update the review could mean that an update is required by the time the review report is published. However, we can be hopeful that the analyses herein would not be significantly affected by the addition of novel research, since the cumulative meta-analysis showed that the last 2 years of evidence were not highly influential in at least one of the analyses.

A further limitation of the evidence base was missing data and meta-data. Table 10 shows some of the commonly missing information within this evidence base. The most common form of missing meta-data was soil descriptions, which hampered our analysis of this source of heterogeneity. Indeed, whilst we tried to convert soil texture classifications to a common scale using available information, certain texture classes were severely underrepresented in some analyses (e.g. silt loam in the comparison of NT relative to IT at 0–15 cm). It is common for study authors to fail to report spatial replication, study duration and study design and rates of reporting of this information in our review were in line with these rates [93]. Tillage descriptions (i.e. depth and machinery) were missing in 31% of studies, which made it difficult to investigate the impact of tillage depth and prevented any form of analysis of tillage equipment. Missing quantitative data in the form of variability measures around the mean was also a problem. Over half of the studies in the review failed to report this data. For some of these studies we were able to estimate treatment variability using an overall variability measure, which had no significant impact on our analyses (shown by sensitivity analysis). However, our meta-analyses were smaller than the available evidence, since the studies without true or estimated variability could not be included.

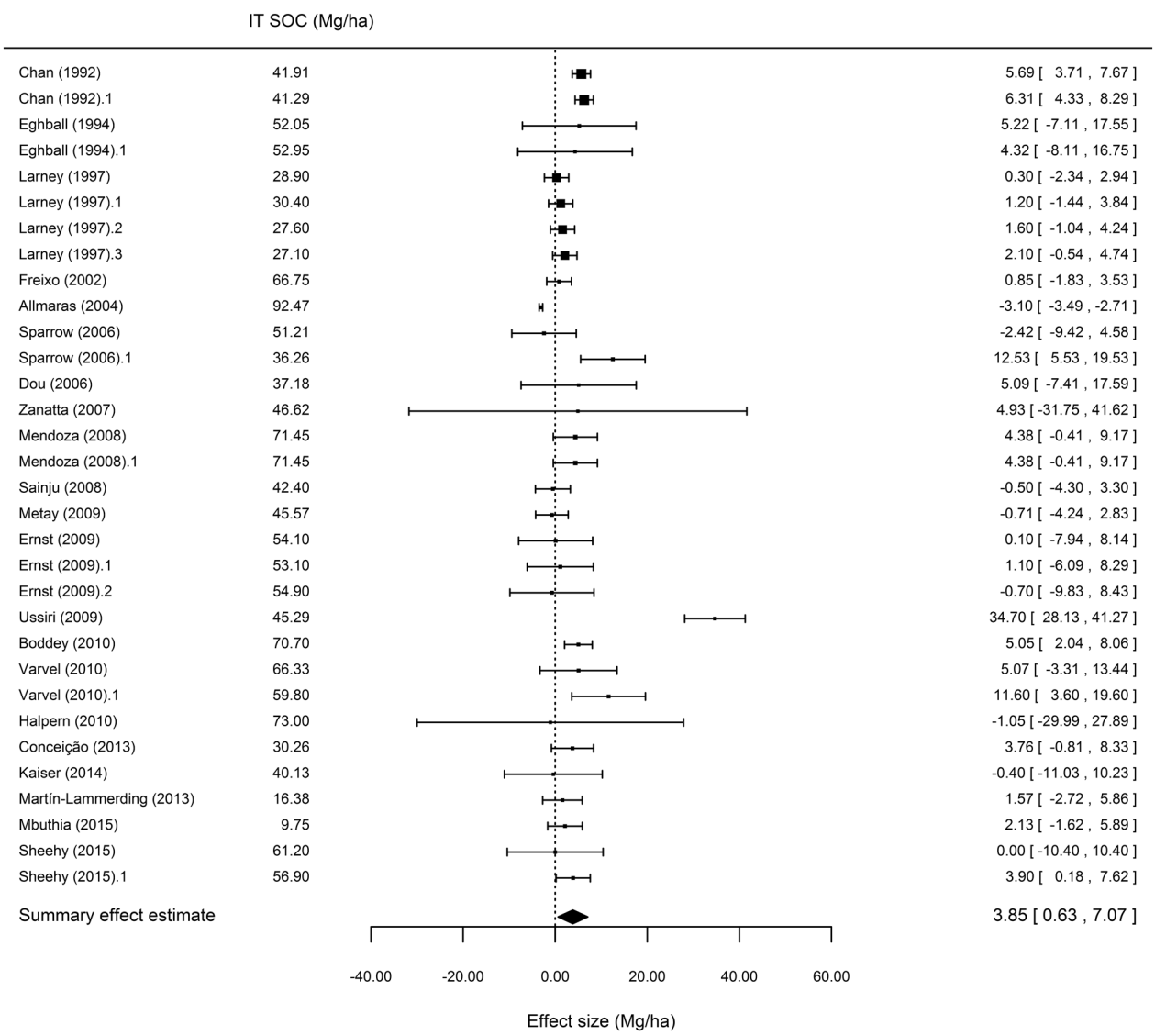

Fig. 43 Forest plot for meta-analysis of stock data for NT–IT comparison in upper layer. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

Our sensitivity analyses and assessments of publication bias, on the whole, failed to identify critical bias in the evidence base. However, there were some notable suggestions of publication bias in the concentration data meta-analyses for NT–HT at > 30 cm and IT–HT at 0–15 cm, and in the stocks data meta-analysis at 0–30 cm. All three instances were for positive trends in less precise studies, where more precise studies showed evenly distributed effect sizes. By accounting for variability in weighting our meta-analysis by inverse variance we have attempted to account for some of this publication bias. We also attempted to reduce the possibility for publication bias in the original systematic map by searching for grey literature [33]. Another factor that may limit the impact of publication bias on our review is that SOC data is often not the main outcome of interest for studies in our review: frequently they focus on other outcomes in addition, such as yield, microbial abundance, or greenhouse gas emissions. As a result, there is not such a clear link between significantly positive SOC data and perceived significance by authors, editors and peer-reviewers, possibly reducing the risk of publication bias. However, we should be aware that our effect estimates may slightly overestimate true effects at least for the three comparisons where evidence of publication bias was found.

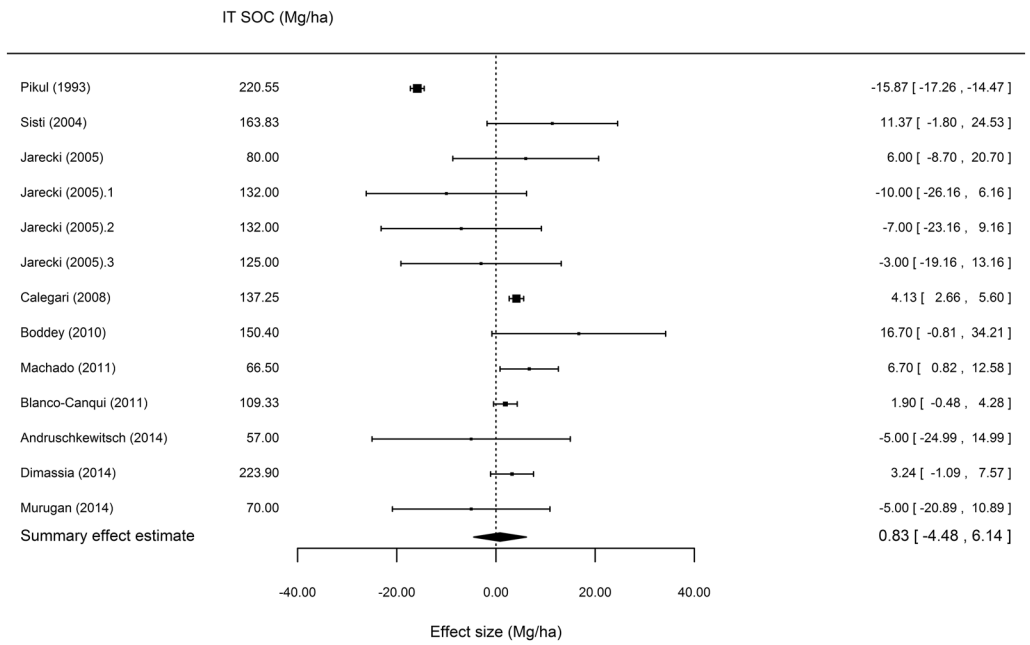

Fig. 44 Forest plot for meta-analysis of stock data for NT–IT comparison in full soil profile. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

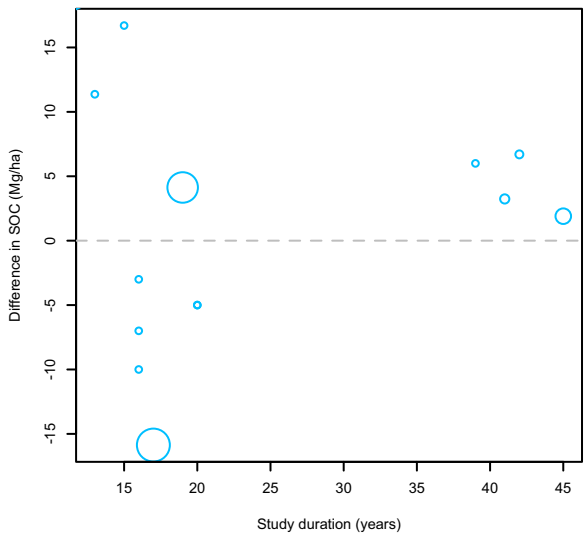

Fig. 45 Meta-regression of difference in SOC stock for NT–IT against study duration in full soil profile. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

Conclusions

Implications for policy and practice

The farming community has a strong interest in management practices not only from the perspective of agronomy but also in relation to the climate. Increasing SOC levels in the upper soil layers can reduce costs for nitrogen applications, since higher SOC level can increase the fertiliser efficiency for a given crop [94]. Among a number of management options to increase SOC for farmers, reduced tillage could provide a means to further reduce losses of SOC in the upper soil layers and contribute to economic efficiency in the long run.

The European agricultural policy that promotes conservation of soil organic matter is outlined in the guidelines for good agricultural and environmental conditions (GAEC) [95]. The policy does not currently contain measures that explicitly deal with tillage, but the results from the meta-analyses contained herein could provide evidence that NT and IT are potential means to promote SOC in the top soil, and thus could be used in formulation of GAECs concerning soils at national levels.

In the United Nations Framework Convention on Climate Change (UNFCCC) soils are considered as an important factor for mitigating C losses, and during the Paris COP meeting in 2015, there was an initiative launched that stated that if soils can annually store 0.4‰ of the global soil stocks this can be used to mitigate a large proportion of the greenhouse gases emissions to the atmosphere [96]. This will not only mitigate climate change but is also intended to provide better food security by increasing soil fertility. The FAO has also launched the Global Soil Partnership, a voluntary partnership open to governments, regional organisations, institutions and

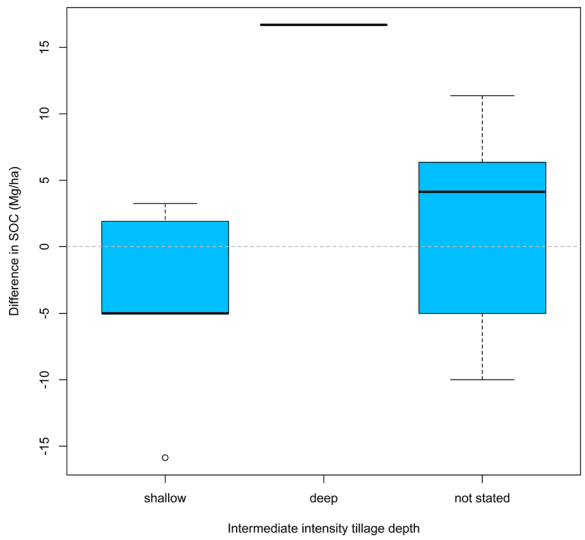

Fig. 46 Boxplot of difference in SOC stock for NT–IT in full soil profile as affected by IT depth. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

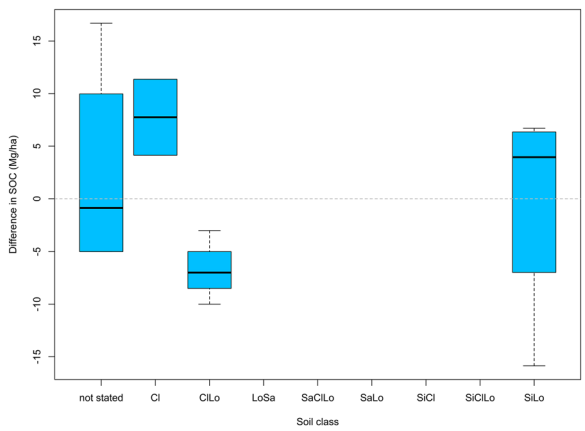

Fig. 47 Boxplot of difference in SOC stock for NT–IT at in full soil profile as affected by soil class. *NT* no tillage, *IT* intermediate intensity tillage (see text for explanation). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

other stakeholders at various levels [97]. The Partnership is guided by an intergovernmental technical panel on soils that provides scientific and technical advice on global soil issues addressing sustainable soil management across various sustainable development agendas. The evidence from this systematic review on SOC stocks from a full soil profile does not show a change due to tillage management, though the collection of evidence (and the apparent lack of data from full profiles) can hopefully be used to support further work to find solutions to increase and maintain C stocks in agricultural soils.

Implications for research

Knowledge gaps and knowledge clusters

Across the evidence considered within this systematic review a suite of other management practices was investigated. Farmers rarely make decisions based on single management practices, but rather consider their field management in a holistic way. However, the majority of the evidence base examined the effect of tillage as a standalone practice (Fig. 56). Key knowledge gaps, therefore, exist around the combined effects of tillage and amendments (such as farmyard manure application and stubble management) on SOC. Similarly, the combined effects of tillage and fertiliser were poorly studied. These represent partial knowledge gaps where further investigation and possibly primary may be warranted.

However, a modest evidence base was found relating to the combined impacts of tillage and crop rotations (Fig. 56): some 88 studies. Whilst the large variety of possible rotations may preclude meta-analysis on this number of studies, it may prove fruitful. Furthermore, a combined approach may be particularly appropriate for this topic, whereby primary research aiming to fill this knowledge gap is combined with further synthesis of existing research identified here.

Methodology

Our results provide quantitative evidence in support of the previously held view that changes in SOC cannot be detected within a 10 year timeframe [41]. This evidence should further strengthen guidance to ensure experiments are in place for longer than a decade before measurements aiming to detect SOC change are made, and researchers should ensure that investigations of SOC seek funding to cover periods of more than 10 years of study to have the necessary power to detect significant change.

Researchers may also benefit particularly from the appraisal that we have undertaken as part of this review. The key limitations to the usefulness of research studies related to missing descriptive information and missing data. Despite the following variables being vital aspects of study design and experimentation, a surprising proportion of the evidence base was deficient for one or more variables, which hampered analysis.

In particular the following meta-data were poorly documented and should be universally reported in detail to facilitate future analyses:

- Study location (i.e. specific geographical location including coordinates).

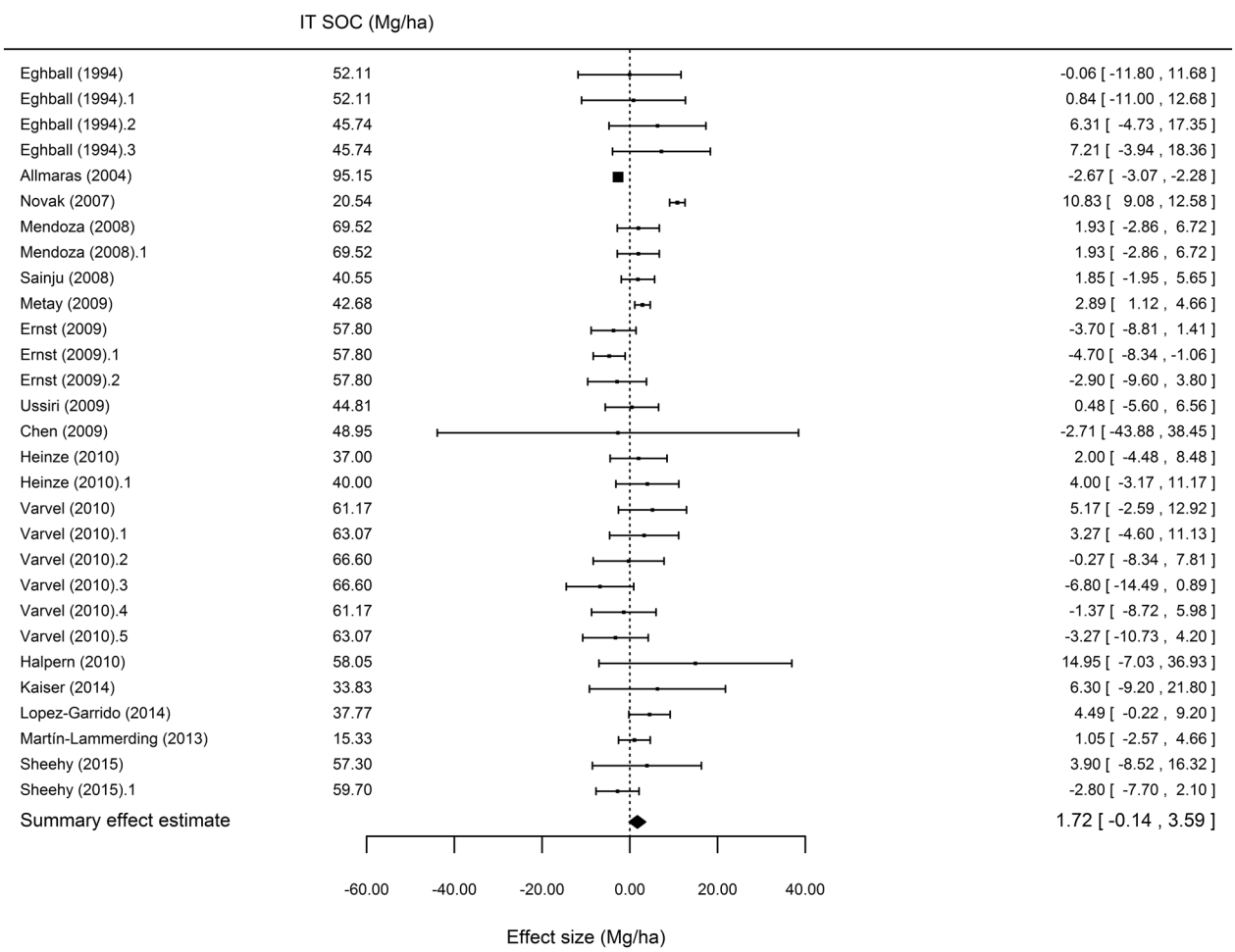

Fig. 48 Forest plot for meta-analysis of stock data for IT–HT comparison in full soil profile. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

- Experimental name or field identifier (if a frequently studied long-term experiment or if multiple long-term experiments conducted at the same site).
- Study AND experimental timing (i.e. both the period of measurement and the period over which the management practice or experiment was in place).
- Soil type (reported as clay/silt/sand or universally accepted soil texture classification).
- Detailed description of the context, including cropping regimes, fertilider rates, soil chemical and physical parameters.
- Detailed description of the study design and experimental layout, including the type and level of randomisation (i.e. how were plots randomly assigned, at what level of the experimental design was randomisation applied [treatment, block, plot, subplot]), the type of study design used, the level of true spatial replication [block, plot, subplot, split plot], the number of true spatial replicates, the number of temporal replicates and the timing of measurements, the dimension of plots).
- Detailed descriptions of the sampling design, including the depths at which soil samples were taken, the method of extraction of soil samples, the number of soil samples taken per plot/subplot.

An additional significant problem related to missing data, including:

- Individual bulk density data across all treatments and depths investigated, including measures of variability

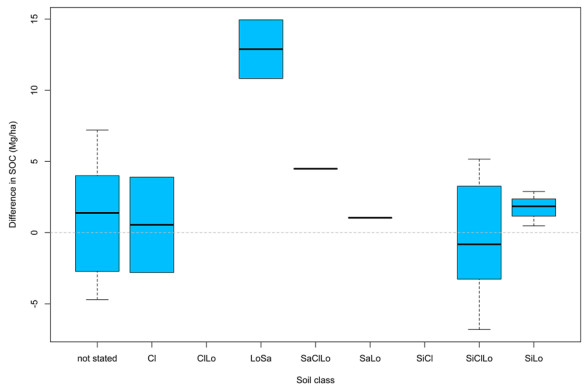

Fig. 49 Boxplot of difference in SOC stock for IT–HT in upper layer as affected by soil class. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

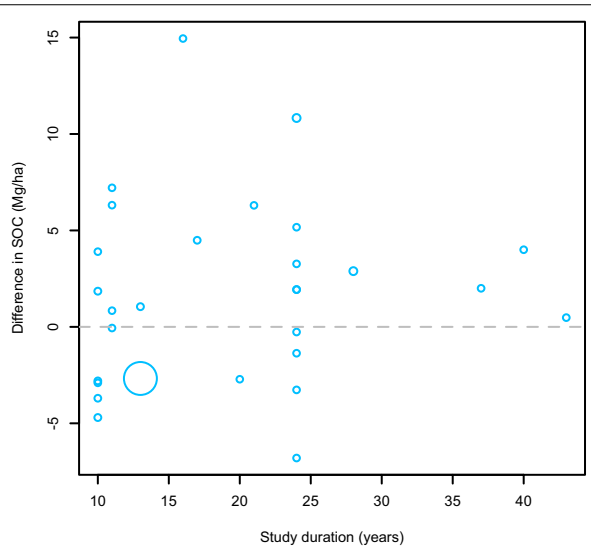

Fig. 50 Meta-regression of difference in SOC stock for IT–HT against study duration in upper layer. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

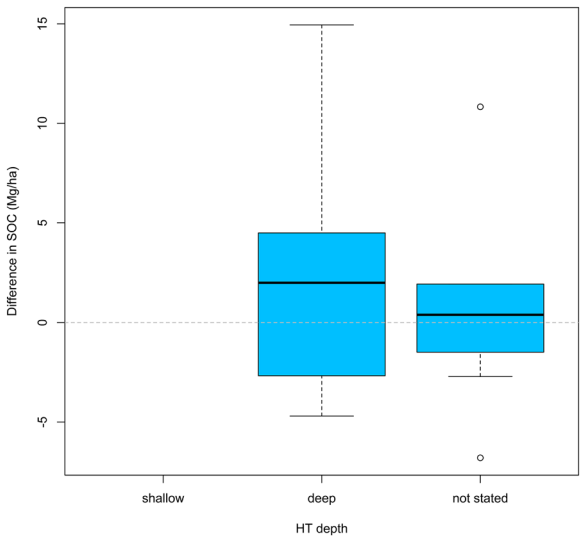

Fig. 51 Boxplot of difference in SOC stock for IT–HT in upper layer as affected by HT depth. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

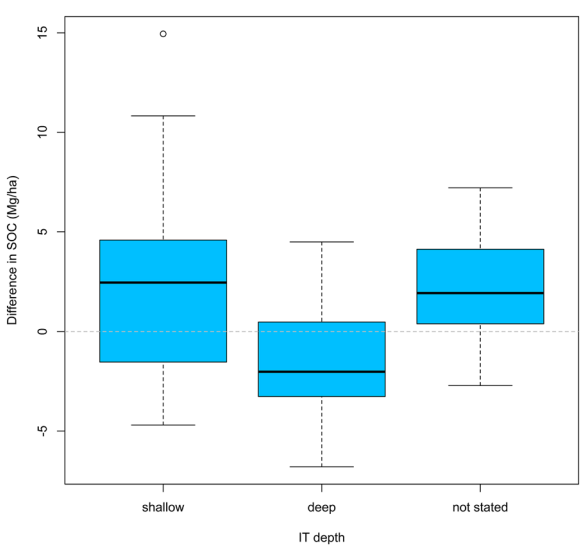

Fig. 52 Boxplot of difference in SOC stock for IT–HT in upper layer as affected by IT depth. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

where available (rather than means across sites, treatments or depth profiles).

- Measures of variability separated by treatment, soil depth and other factors considered, including other farming practices such as different crop rotations or fertiliser rates (i.e. standard deviation, 95% confidence interval, standard error).
- Sample sizes for true replicates (true replicates are those that occur at the same level as the factor of interest, e.g. if tillage treatments are applied to different fields then true replicates must occur at the field scale; subplots are pseudoreplicates).
- Long-term study data separated over time (i.e. all time points summaries using means for each time, or raw data provided).

Wherever possible all raw data should be provided, allowing synthesists to maximize the legacy and impact

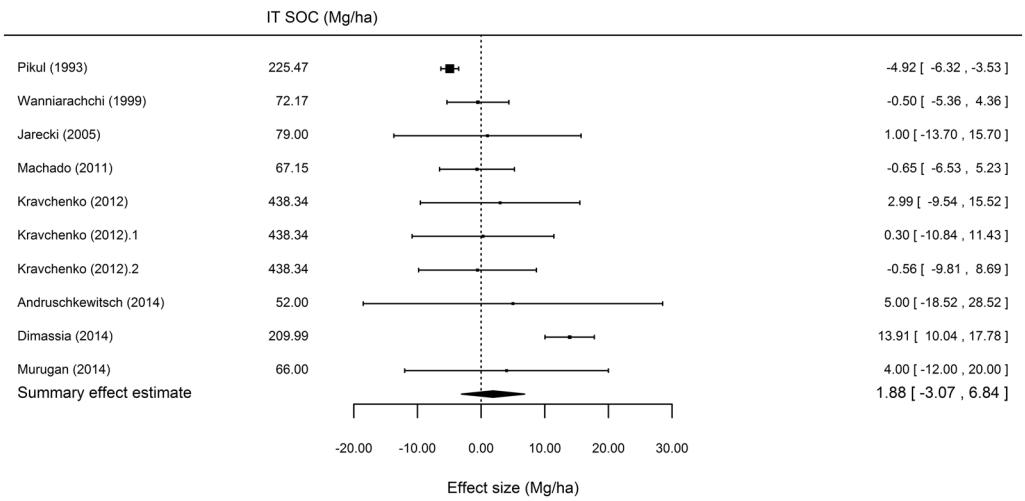

Fig. 53 Forest plot for meta-analysis of stock data for IT–HT comparison in full soil profile. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). The summary effect estimate diamond is centred on the summary effect estimate for the meta-analysis, with the points of the diamond representing the 95% confidence intervals. Numbers in the right hand column are summary effect estimates [lower 95% CI, upper 95% CI]

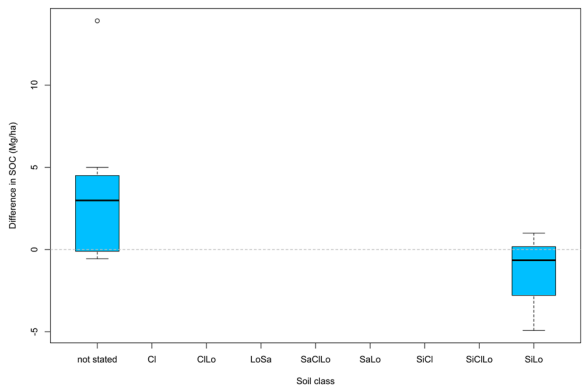

Fig. 54 Boxplot of difference in SOC stock for IT–HT in full soil profile as affected by soil class. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). See text for explanation of tillage groups and soil classes (USDA classification). Thick line, median; boxes, interquartile ranges (Q1–Q3); whiskers, non-outlier range; points, outliers

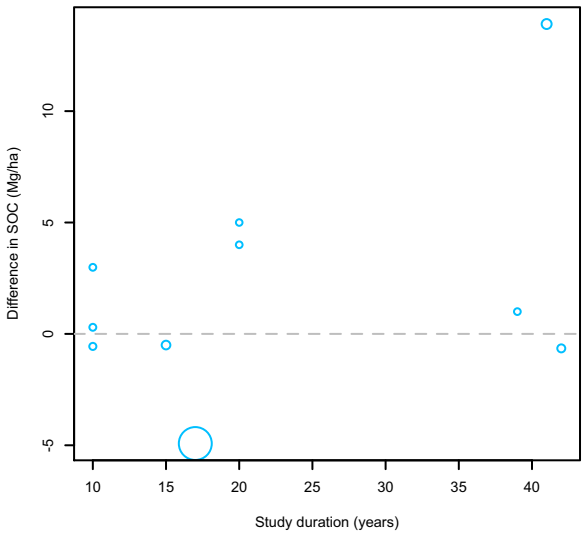

Fig. 55 Meta-regression of difference in SOC stock for IT–HT against study duration in full soil profile. *IT* intermediate intensity tillage, *HT* high intensity tillage (see text for explanation). Point size represents study weighting in the analysis (inverse variance)

of primary research. Primary research authors should see secondary synthesis in the form of systematic maps and systematic reviews as a valuable demonstration of impact of their research outputs. Such activities seek to combine research outputs to examine patterns across scales that would likely be impossible within current constraints of funding, resources and administration.

General conclusions

In this review, we compare tillage treatment effects on SOC concentrations and stocks in the upper layers of agricultural soils that have accumulated over at least a decade. This can be of importance for a number of ecosystem services, such as climate mitigation and nutrient retention. Whether observed positive changes in these measures correspond to positive absolute changes in total SOC over time has not been investigated here but will be subject to a subsequent meta-analysis for a subset of studies for which time-series measurement are available [98]. However, for mitigation of climate change, site-specific relative changes in SOC following certain management practices

Table 10 Commonly missing meta-data and data from the evidence base within this review

Missing information	No. of studies	% of evidence	Description
Soil description	37	11	Studies lacking any description of soil type
Study duration	1	0.3	Study lacks any description of how long the experiments lasted
Study timing	20	6	Studies lacking information on when the experiments started and ended
Study design	29	8	Studies lacking information about the experimental design (i.e. randomisation, blocking, split-plot, strip-plot, latin square, paired design, purposive)
Spatial replication	18	5	Studies lacking details of how many true spatial replicates were sampled
Temporal replication	3	0.9	Studies lacking information of how many samples were taken over time
Soil sampling depth	1	0.3	Studies lacking information on how many soil samples were taken and at what depth
Tillage description	108	31	Studies lacking information on the type and intensity of tillage treatments
Measure of variability	196	56	Studies failing to report a measure of variability around treatment means

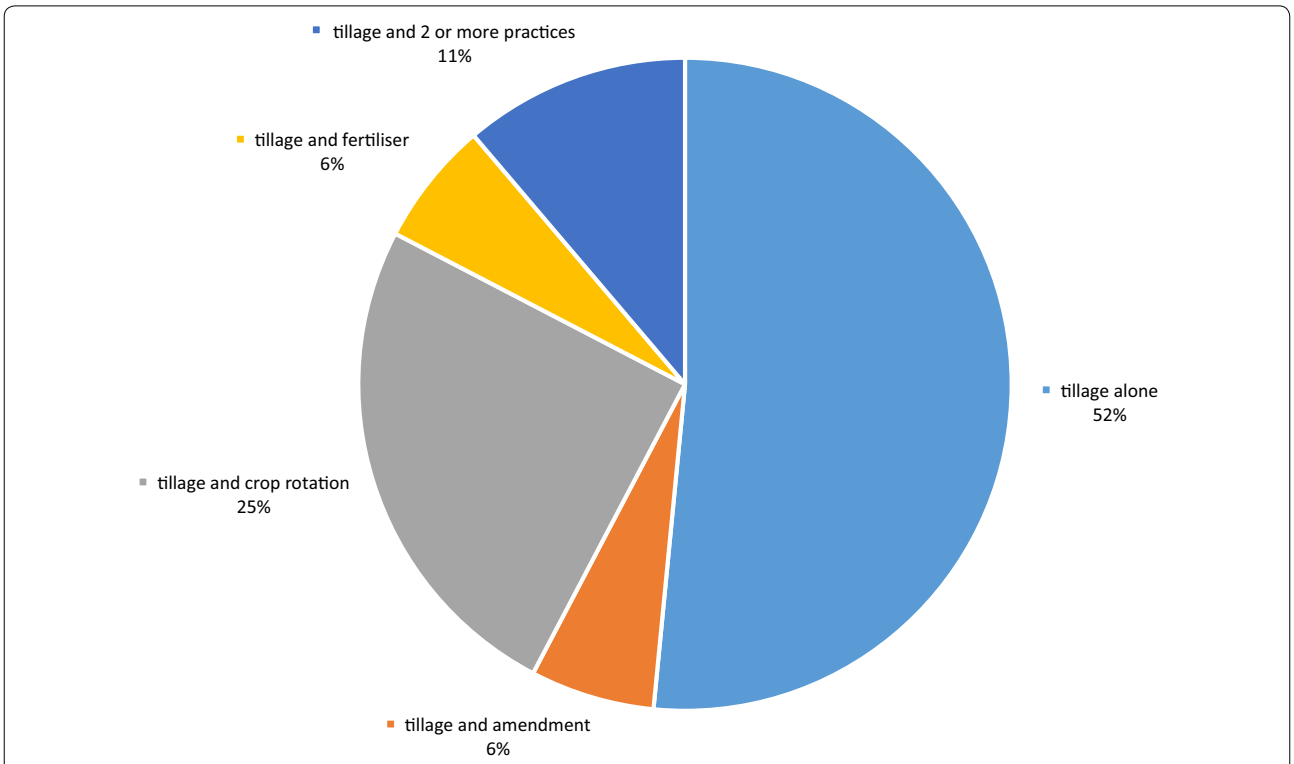

Fig. 56 Pie chart of the key farming practices investigated alongside tillage. Practices are followed by the percentage of the evidence base

are very important since absolute changes are mainly determined by initial SOC states rather than treatments imposed in a specific experiment [99]. The environmental impact of tillage needs to be considered for a number of factors influencing both farmers (crop production, future soil fertility) as well as society.

Additional files

Additional file 1. Bibliographic database search record.

Additional file 2. Unobtainable articles.

Additional file 3. Excluded and included articles at full text screening.

Additional file 4. Data extraction files (extracted study findings and effect size calculations from all included studies).

Additional file 5. Effect size calculation plan.

Additional file 6. Tillage studies systematic map database.

Additional file 7. GIS help file.

Additional file 8. Input data for concentration SOC meta-analyses.

Additional file 9. Input data for stocks SOC meta-analyses.

Additional file 10. Concentration SOC meta-analysis report.

Additional file 11. Stocks SOC meta-analysis report.

Additional file 12. Concentration SOC meta-analysis R (text file).

Additional file 13. Stocks SOC meta-analysis R (text file).

Additional file 14. Bulk density and effect size calculation checking report.

Abbreviations

C: carbon; SOC: soil organic carbon; SOM: soil organic matter; 95% CI: 95% confidence interval; CV: coefficient of variation; HT: high intensity tillage; IT: intermediate intensity tillage; LSD: least squares difference; NT: no tillage; SD: standard deviation; SE: standard error; SED: standard error of the difference.

Authors' contributions

This review is based on a draft written by NRH. NRH and HBJ performed searches, screened identified records and extracted data. NRH performed analyses. All authors assisted in editing and revising the draft. All authors read and approved the final manuscript.

Author details

[1] Mistra Council for Evidence-Based Environmental Management (EviEM), Stockholm Environment Institute, Box 24218, 10451 Stockholm, Sweden. [2] Department of Biology, Lund University, 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 of Ecology, SLU, P.O. Box 7044, 750 07 Uppsala, Sweden. [5] European Commission, Joint Research Centre (JRC), Directorate for Sustainable Resources, Land Resources Unit, Via E. Fermi, 2749, 21027 Ispra, VA, Italy. [6] Department of Agroecology, Aarhus, University, P.O. Box 50, 8830 Tjele, Denmark. [7] Department of Statistics, Lund University, 220 07 Lund, Sweden.

Acknowledgements

The authors thank two anonymous reviewers whose advice improved this protocol.

Competing interests

The authors declare that they have no competing interests. The reviewers involved in this review that are also authors of relevant articles were not included in the decisions connected to inclusion and critical appraisal of these articles.

Availability of data and materials

A list of excluded studies along with exclusion reasons is provided in additional information. An updated systematic map for tillage studies is also provided, along with extracted tables and figures and summary statistics generated from them for use in meta-analysis. R scripts for all analyses are included along with summary statistics for all models performed, irrespective of significance.

Consent for publication

Not applicable.

Ethical approval and consent to participate

Not applicable.

Funding

This review was financed by the Mistra Council for Evidence-Based Environmental Management (EviEM).

Publisher's Note

Received: 23 January 2017 Accepted: 27 October 2017
Published online: 18 December 2017

References

1. Follett R. Soil management concepts and carbon sequestration in cropland soils. Soil Tillage Res. 2001;61(1):77–92.
2. Kimble JM, Follett RF, Cole CV. The potential of US cropland to sequester carbon and mitigate the greenhouse effect. Boca Raton: CRC Press; 1998.
3. Sauerbeck D. CO_2 emissions and C sequestration by agriculture—perspectives and limitations. Nutr Cycl Agroecosyst. 2001;60(1–3):253–66.
4. Schlesinger W. Biogeochemistry: an analysis of global change. San Diego: Academic Press; 1991.
5. 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. Agric For Meteorol. 2007;142(2):216–33.
6. Kucharik CJ, Brye KR, Norman JM, Foley JA, Gower ST, Bundy LG. Measurements and modeling of carbon and nitrogen cycling in agroecosystems of southern Wisconsin: potential for SOC sequestration during the next 50 years. Ecosystems. 2001;4(3):237–58.
7. Reicosky D. Tillage-induced CO_2 emissions and carbon sequestration: effect of secondary tillage and compaction. Conservation agriculture. Berlin: Springer; 2003. p. 291–300.
8. 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 Tillage Res. 2012;122:52–60.
9. Lal R, Delgado J, Groffman P, Millar N, Dell C, Rotz A. Management to mitigate and adapt to climate change. J Soil Water Conserv. 2011;66(4):276–85.
10. 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 (63–64N). Agric Ecosyst Environ. 2010;138(3):335–42.
11. Lal R, Follett R. Soils and climate change. Soil carbon sequestration and the greenhouse effect. Madison: SSSA Special Publication; 2009. p. 57.
12. 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. Agric Ecosyst Environ. 2007;119(1):127–34.
13. 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. J Plant Nutr Soil Sci. 2011;174(5):775–84.
14. Barrios E. Soil biota, ecosystem services and land productivity. Ecol Econ. 2007;64(2):269–85.
15. Lal R, Reicosky D, Hanson J. Evolution of the plow over 10,000 years and the rationale for no-till farming. Soil Tillage Res. 2007;93(1):1–12.
16. Kern J, Johnson M. Conservation tillage impacts on national soil and atmospheric carbon levels. Soil Sci Soc Am J. 1993;57(1):200–10.
17. West TO, Post WM. Soil organic carbon sequestration rates by tillage and crop rotation. Soil Sci Soc Am J. 2002;66(6):1930–46.

18. Holland J. The environmental consequences of adopting conservation tillage in Europe: reviewing the evidence. Agric Ecosyst Environ. 2004;103(1):1–25.
19. Davies DB, Finney JB. Reduced cultivations for cereals: research, development and advisory needs under changing economic circumstances. Kenilworth: Home Grown Cereals Authority; 2002.
20. Pittelkow CM, Linquist BA, Lundy ME, Liang X, Van Groenigen KJ, Lee J, Van Gestel N, Six J, Venterea RT, Van Kessel C. When does no-till yield more? A global meta-analysis. Field Crops Res. 2015;183:156–68.
21. Van den Putte A, Govers G, Diels J, Gillijns K, Demuzere M. Assessing the effect of soil tillage on crop growth: a meta-regression analysis on European crop yields under conservation agriculture. Eur J Agron. 2010;33(3):231–41.
22. Basche AD, Miguez FE, Kaspar TC, Castellano MJ. Do cover crops increase or decrease nitrous oxide emissions? A meta-analysis. J Soil Water Conserv. 2014;69(6):471–82.
23. Rochette P, Worth DE, Lemke RL, McConkey BG, Pennock DJ, Wagner-Riddle C, Desjardins R. Estimation of N_2O emissions from agricultural soils in Canada. I. Development of a country-specific methodology. Can J Soil Sci. 2008;88(5):641–54.
24. Alvarez R. A review of nitrogen fertilizer and conservation tillage effects on soil organic carbon storage. Soil Use Manag. 2005;21(1):38–52.
25. Amini S, Asoodar MA. The effect of conservation tillage on crop yield production (The Review). N Y Sci J 2015;8(3):25–9.
26. Angers D, Eriksen-Hamel N. Full-inversion tillage and organic carbon distribution in soil profiles: a meta-analysis. Soil Sci Soc Am J. 2008;72(5):1370–4.
27. Govaerts B, Verhulst N, Castellanos-Navarrete A, Sayre KD, Dixon J, Dendooven L. Conservation agriculture and soil carbon sequestration: between myth and farmer reality. Crit Rev Plant Sci. 2009;28(3):97–122.
28. Six J, Feller C, Denef K, Ogle S, de Moraes Sa JC, Albrecht A. Soil organic matter, biota and aggregation in temperate and tropical soils-effects of no-tillage. Agronomie. 2002;22(7–8):755–75.
29. Dimassi B, Mary B, Wylleman R, Labreuche J, Couture D, Piraux F, Cohan J-P. Long-term effect of contrasted tillage and crop management on soil carbon dynamics during 41 years. Agric Ecosyst Environ. 2014;188:134–46.
30. Powlson DS, Stirling CM, Jat M, Gerard BG, Palm CA, Sanchez PA, Cassman KG. Limited potential of no-till agriculture for climate change mitigation. Nat Clim Change. 2014;4(8):678–83.
31. Baker JM, Ochsner TE, Venterea RT, Griffis TJ. Tillage and soil carbon sequestration—What do we really know? Agric Ecosyst Environ. 2007;118(1):1–5.
32. Virto I, Barré P, Burlot A, Chenu C. Carbon input differences as the main factor explaining the variability in soil organic C storage in no-tilled compared to inversion tilled agrosystems. Biogeochemistry. 2012;108(1–3):17–26.
33. Haddaway NR, Hedlund K, Jackson LE, Kätterer T, Lugato E, Thomsen IK, Jørgensen HB, Söderström B. What are the effects of agricultural management on soil organic carbon in boreo-temperate systems? Environ Evid. 2015;4(1):1.
34. Amini S, Asoodar MA. The effect of conservation tillage on crop yield production. NY Sci J. 2016;8(3):25–9.
35. Haddaway NR, Hedlund K, Jackson LE, Jørgensen HB, Kätterer T, Lugato E, Söderström B, Thomsen IK. What are the effects of agricultural management on soil organic carbon (SOC) stocks? A systematic map. Environ Evid. 2015;4(1):23.
36. Söderström B, Hedlund K, Jackson LE, Kätterer T, Lugato E, Thomsen IK, Jørgensen HB. What are the effects of agricultural management on soil organic carbon (SOC) stocks? Environ Evid. 2014;3(1):1.
37. Haddaway NR, Hedlund K, Jackson LE, Kätterer T, Lugato E, Thomsen IK, Jørgensen HB, Isberg P-E. How does tillage intensity affect soil organic carbon? A systematic review protocol. Environ Evid. 2016;5(1):1.
38. Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. PLoS ONE. 2015;10(9):e0138237.
39. Haddaway NR. The use of web-scraping software in searching for grey literature. Grey J. 2015;11(3):186–90.
40. Necpálová M, Anex R, Kravchenko AN, Abendroth LJ, Del Grosso SJ, Dick WA, Helmers MJ, Herzmann D, Lauer JG, Nafziger ED. What does it take to detect a change in soil carbon stock? A regional comparison of minimum detectable difference and experiment duration in the north central United States. J Soil Water Conserv. 2014;69(6):517–31.
41. Smith P. How long before a change in soil organic carbon can be detected? Glob Change Biol. 2004;10(11):1878–83.
42. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213.
43. Soil Survey Staff. Soil survey manual. 1993.
44. Team RC. R: A language and environment for statistical computing. Vellore: Team RC; 2016.
45. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36(3):1–48.
46. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed effects modelling for nested data. Mixed effects models and extensions in ecology with R. Berlin: Springer; 2009. p. 101–42.
47. Poeplau C, Don A. Carbon sequestration in agricultural soils via cultivation of cover crops—a meta-analysis. Agric Ecosyst Environ. 2015;200:33–41.
48. Wiesmeier M, von Lützow M, Spörlein P, Geuß U, Hangen E, Reischl A, Schilling B, Kögel-Knabner I. Land use effects on organic carbon storage in soils of Bavaria: the importance of soil types. Soil Tillage Res. 2015;146:296–302.
49. Mao D, Wang Z, Li L, Miao Z, Ma W, Song C, Ren C, Jia M. Soil organic carbon in the Sanjiang Plain of China: storage, distribution and controlling factors. Biogeosciences. 2015;12(6):1635–45.
50. Cochran WG. The combination of estimates from different experiments. Biometrics. 1954;10(1):101–29.
51. Viechtbauer W, Cheung MWL. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010;1(2):112–25.
52. Lal R. Global potential of soil carbon sequestration to mitigate the greenhouse effect. Crit Rev Plant Sci. 2003;22(2):151–84.
53. Busari MA, Kukal SS, Kaur A, Bhatt R, Dulazi AA. Conservation tillage impacts on soil, crop and the environment. Int Soil Water Conserv Res. 2015;3(2):119–29.
54. Sperow M. Estimating carbon sequestration potential on US agricultural topsoils. Soil Tillage Res. 2016;155:390–400.
55. Minoshima H, Jackson L, Cavagnaro T, Sánchez-Moreno S, Ferris H, Temple S, Goyal S, Mitchell J. Soil food webs and carbon dynamics in response to conservation tillage in California. Soil Sci Soc Am J. 2007;71(3):952–63.
56. Olson KR. Impacts of tillage, slope, and erosion on soil organic carbon retention. Soil Sci. 2010;175(11):562–7.
57. Liu X, Zhang X, Chen S, Sun H, Shao L. Subsoil compaction and irrigation regimes affect the root–shoot relation and grain yield of winter wheat. Agric Water Manag. 2015;154:59–67.
58. Abdalla K, Chivenge P, Ciais P, Chaplot V. No-tillage lessens soil CO_2 emissions the most under arid and sandy soil conditions: results from a meta-analysis. Biogeosci Discus. 2015;13:3619–33.
59. Xu X, Shi Z, Li D, Rey A, Ruan H, Craine JM, Liang J, Zhou J, Luo Y. Soil properties control decomposition of soil organic carbon: results from data-assimilation analysis. Geoderma. 2016;262:235–42.
60. 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.
61. Reeves D. The role of soil organic matter in maintaining soil quality in continuous cropping systems. Soil Tillage Res. 1997;43(1):131–67.
62. Carter M. Relative measures of soil bulk density to characterize compaction in tillage studies on fine sandy loams. Can J Soil Sci. 1990;70(3):425–33.
63. Ellert B, Janzen H, Entz T. Assessment of a method to measure temporal change in soil carbon storage. Soil Sci Soc Am J. 2002;66(5):1687–95.
64. Wuest SB. Correction of bulk density and sampling method biases using soil mass per unit area. Soil Sci Soc Am J. 2009;73(1):312–6.
65. Lee J, Hopmans JW, Rolston DE, Baer SG, Six J. Determining soil carbon stock changes: simple bulk density corrections fail. Agric Ecosyst Environ. 2009;134(3):251–6.
66. Wendt J, Hauser S. An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers. Eur J Soil Sci. 2013;64(1):58–65.
67. McBratney AB, Minasny B. Comment on "Determining soil carbon stock changes: simple bulk density corrections fail". Agric Ecosyst Environ. 2010;136(1):185–6.

68. Goidts E, Van Wesemael B, Crucifix M. Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales. Eur J Soil Sci. 2009;60(5):723–39.
69. Doetterl S, Berhe AA, Nadeu E, Wang Z, Sommer M, Fiener P. Erosion, deposition and soil carbon: a review of process-level controls, experimental tools and models to address C cycling in dynamic landscapes. Earth Sci Rev. 2016;154:102–22.
70. Amundson R, Berhe AA, Hopmans JW, Olson C, Sztein AE, Sparks DL. Soil and human security in the 21st century. Science. 2015;348(6235):1261071.
71. Sommer M, Augustin J, Kleber M. Feedbacks of soil erosion on SOC patterns and carbon dynamics in agricultural landscapes—the CarboZALF experiment. Soil Tillage Res. 2016;156:182–4.
72. Pittelkow CM, Liang X, Linquist BA, Van Groenigen KJ, Lee J, Lundy ME, van Gestel N, Six J, Venterea RT, van Kessel C. Productivity limits and potentials of the principles of conservation agriculture. Nature. 2015;517(7534):365–8.
73. Ball B. Soil structure and greenhouse gas emissions: a synthesis of 20 years of experimentation. Eur J Soil Sci. 2013;64(3):357–73.
74. Awale R, Chatterjee A, Franzen D. Tillage and N-fertilizer influences on selected organic carbon fractions in a North Dakota silty clay soil. Soil Tillage Res. 2013;134:213–22.
75. Abdullah AS. Minimum tillage and residue management increase soil water content, soil organic matter and canola seed yield and seed oil content in the semiarid areas of Northern Iraq. Soil Tillage Res. 2014;144:150–5.
76. Johnston AE, Poulton PR, Coleman K. Soil organic matter: its importance in sustainable agriculture and carbon dioxide fluxes. Adv Agron. 2009;101:1–57.
77. Dimassi B, Cohan J-P, Labreuche J, Mary B. Changes in soil carbon and nitrogen following tillage conversion in a long-term experiment in Northern France. Agric Ecosyst Environ. 2013;169:12–20.
78. Antille DL, Chamen WC, Tullberg JN, Lal R. The potential of controlled traffic farming to mitigate greenhouse gas emissions and enhance carbon sequestration in arable land: a critical review. Trans ASABE. 2015;58(3):707–31.
79. Golchin A, Oades J, Skjemstad J, Clarke P. Study of free and occluded particulate organic matter in soils by solid state 13C CP/MAS NMR spectroscopy and scanning electron microscopy. Soil Res. 1994;32(2):285–309.
80. Angers D, Bolinder M, Carter M, Gregorich E, Drury C, Liang B, Voroney R, Simard R, Donald R, Beyaert R. Impact of tillage practices on organic carbon and nitrogen storage in cool, humid soils of eastern Canada. Soil Tillage Res. 1997;41(3):191–201.
81. Kainiemi V, Arvidsson J, Kätterer T. Short-term organic matter mineralisation following different types of tillage on a Swedish clay soil. Biol Fertil Soils. 2013;49(5):495–504.
82. Kainiemi V, Arvidsson J, Kätterer T. Effects of autumn tillage and residue management on soil respiration in a long-term field experiment in Sweden. J Plant Nutr Soil Sci. 2015;178(2):189–98.
83. Kainiemi V, Kirchmann H, Kätterer T. Structural disruption of arable soils under laboratory conditions causes minor respiration increases. J Plant Nutr Soil Sci. 2015;179(1):88–93.
84. Ladygina N, Hedlund K. Plant species influence microbial diversity and carbon allocation in the rhizosphere. Soil Biol Biochem. 2010;42(2):162–8.
85. 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. Agric Ecosyst Environ. 2011;141(1):184–92.
86. Tsiafouli MA, Thébault E, Sgardelis SP, Ruiter PC, Putten WH, Birkhofer K, Hemerik L, Vries FT, Bardgett RD, Brady MV. Intensive agriculture reduces soil biodiversity across Europe. Glob Change Biol. 2015;21(2):973–85.
87. Helgason T, Daniell T, Husband R, Fitter A, Young J. Ploughing up the wood-wide web? Nature. 1998;394(6692):431.
88. de Vries FT, Thébault E, Liiri M, Birkhofer K, Tsiafouli MA, Bjørnlund L, Jørgensen HB, Brady MV, Christensen S, de Ruiter PC. Soil food web properties explain ecosystem services across European land use systems. Proc Natl Acad Sci. 2013;110(35):14296–301.
89. Six J, Frey SD, Thiet RK, Batten KM. Bacterial and fungal contributions to carbon sequestration in agroecosystems. Soil Sci Soc Am J. 2006;70(2):555–69.
90. De Vries FT, Bracht Jørgensen H, Hedlund K, Bardgett RD. Disentangling plant and soil microbial controls on carbon and nitrogen loss in grassland mesocosms. J Ecol. 2015;103(3):629–40.
91. Bernes C, Carpenter SR, Gårdmark A, Larsson P, Persson L, Skov C, Speed JD, Van Donk E. What is the influence of a reduction of planktivorous and benthivorous fish on water quality in temperate eutrophic lakes? A systematic review. Environ Evid. 2015;4(1):1.
92. Haddaway NR, Burden A, Evans CD, Healey JR, Jones DL, Dalrymple SE, Pullin AS. Evaluating effects of land management on greenhouse gas fluxes and carbon balances in boreo-temperate lowland peatland systems. Environ Evid. 2014;3(1):1.
93. Haddaway NR, Verhoeven JT. Poor methodological detail precludes experimental repeatability and hampers synthesis in ecology. Ecol Evol. 2015;5(19):4451–4.
94. Brady MV, Hedlund K, Cong R-G, Hemerik L, Hotes S, Machado S, Mattsson L, Schulz E, Thomsen IK. Valuing supporting soil ecosystem services in agriculture: a natural capital approach. Agron J. 2015;107(5):1809–21.
95. REGULATION (EU) No 1307/2013. 2013.
96. Agenda L-PA. Join the 4/1000 Initiative Soils for Food Security and Climate: UNFCCC; 2016. http://newsroom.unfccc.int/lpaa/agriculture/join-the-41000-initiative-soils-for-food-security-and-climate/. Accessed 9 Dec 2016.
97. FAO. Global Soil Partnership: Food and Agriculture Organization of the United Nations; 2016. http://www.fao.org/global-soil-partnership/en/. Accessed 9 Dec 2016.
98. Haddaway NR, Hedlund K, Jackson LE, Kätterer T, Lugato E, Thomsen IK, Jørgensen HB, Isberg P-E. Which agricultural management interventions are most influential on soil organic carbon (using time series data)? Environ Evid. 2016;5(1):1.
99. Kätterer T, Bolinder M, Berglund K, Kirchmann H. Strategies for carbon sequestration in agricultural soils in northern Europe. Acta Agric Scandinavica A-Anim Sci. 2012;62(4):181–98.