https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Quality Assessment of Meta-Analyses on Soil Organic Carbon

Julia Fohrafellner[1]*, Sophie Zechmeister-Boltenstern[2], Rajasekaran Murugan[2], Elena Valkama[3]

[1] BIOS Science Austria, Dietrichgasse 27, 1130 Vienna, Austria

[2] Institute of Soil Research (IBF), Department of Forest- and Soil Sciences, University of Natural Resources and Life Sciences, Vienna (BOKU), Gregor-Mendel-Straße 33, 1180 Vienna, Austria

[3] Natural Resources Institute Finland (Luke), Bioeconomy and environment, Sustainability Science and Indicators. Tietotie 4, 31600 Jokioinen, Finland

Corresponding author: Julia Fohrafellner (julia.fohrafellner@boku.ac.at)

Abstract. Soil organic carbon (SOC) plays a vital role in the global carbon cycle and is a potential sink for carbon dioxide. Agricultural management practices can support carbon sequestration and therefore offer potential removal strategies, whilst improving overall soil quality. Meta-analysis allows to summarize results from primary articles by calculating an overall effect size and hence to reveal the source of variation across studies. The number of meta-analyses published in the field of agriculture is continuously rising. At the same time, more and more articles refer to their synthesis work as a meta-analysis, despite applying less than rigorous methodologies. As a result, poor quality meta-analyses are published, which may lead to questionable conclusions and recommendations to scientists, policymakers and farmers.

This study aims at quantitatively analyzing 31 meta-analyses, published between the years 2005-2020, studying the effects of different management practices on SOC. We compiled a quality criteria-set, suitable for soil and agricultural sciences, by adapting existing meta-analytical guidelines from other disciplines. The set is supported by a scoring scheme, which allows a quantitative analysis. The retrieved meta-analyses were structured according to 11 management categories, such as tillage, cover crops, residue management, biochar application etc., which allowed us to assess the state-of-knowledge on these categories. Major deficiencies were found in the use of standard metrics for effect size calculation, independence of effect sizes, standard deviation extraction for each study and weighting by the inverse of variance. Only one out of 31 SOC meta-analyses, which studied the effects of no-till/reduced tillage compared to conventional tillage, was found to be of high quality. Therefore, improved meta-analyses on the effects of e.g., organic agriculture, biochar, fertilization or crop diversification on SOC are urgently needed.

We conclude that, despite the efforts over the last 15 years, the quality of meta-analyses on SOC research is still low. In order for the scientific community to provide high quality synthesis work and to make advancements in the sustainable management of agricultural soils, we need to adapt rigorous methodologies of meta-analysis as quickly as possible.

Keywords: effect size, soil management, synthesis, tillage, treatment effect, weighting

1. Introduction

1.1. Meta-analysis as a method and application in different disciplines

Meta-analysis was first defined by Glass as "the statistical analysis of a large collection of results for the purpose of integrating these findings" (1976, p.3). A newer, more precise definition by Koricheva and Gurevitch (2013) describes it as "a set of statistical methods for combining the magnitude of the outcomes (effect sizes) across different data sets addressing the same research question". It supports the structuring of the increasing amount of information (Koricheva and Gurevitch, 2014), which researchers of all fields face, and offers tools to process information with increased precision and reliability (Cooper et al., 2019b; Nakagawa and Cuthill, 2007).

Meta-analysis was developed to facilitate quantitative evidence synthesis in medical, social, and behavioral sciences (Hedges et al., 1999; Gurevitch et al., 2018). The method was first applied in ecology and evolutionary biology about 30 years ago, at a time where a need for quantitative assessment of urgent issues such as climate change or biodiversity losses arose. Since then, meta-analysis has developed within the field of ecology, establishing centers and collaborations for research synthesis (Gurevitch et al., 2018). The results of these contributions frequently provide relevant stakeholders and decision-makers with evidence-based information (Stewart, 2010).

In agricultural research, meta-analysis has only attracted a broader interest in the last decade (Fig. 1). Particularly, the use of meta-analysis as a tool to investigate the effects of agricultural management practices on relevant response variables, such as yield or soil physical or chemical parameters, is becoming increasingly prominent (Valkama et al., 2019, 2015). As these developments are rather recent, the knowledge on appropriate meta-analytical methodology is still finding its place in the research community.

Because of their close relationship, many applications of meta-analyses in ecology are also transferable to the field of agriculture and soil sciences. When looking at the possible diversity of study outcomes on certain topics in agronomy, meta-analysis can provide clarification by synthesizing conflicting evidence from primary studies. Combining results across several sites or assessing the impacts of environmental drivers, as climate change, are tasks which are processable by meta-analysis (Koricheva and Gurevitch, 2014). Nevertheless, research on agriculture and soil encounters issues, which are often specific to these fields. Firstly, changes in soil are often slower than other physiological and biogeochemical changes; e.g., changes within plant tissue. Therefore, long time experiments are needed to detect treatment effects on soil parameters or soil health indicators, like soil organic carbon (SOC). Moreover, these systems are very complex, as not only pedoclimatic conditions influence soil, but also agricultural management practices impact variables of interest. Especially the mix or combination of practices, e.g. tillage plus residue retention, makes it difficult to distinguish between sources of effects (Xiao et al., 2021). Therefore, it is crucial to define not only the treatment but also the control of the experiments precisely to allow computation of heterogeneity.

Lastly, when it comes to soil parameters and indicators, several methods are available for computation, which may cause difficulties in comparing outcomes. A good example is bulk density, which can be measured in a field experiment or estimated using pedotransfer functions in order to compute SOC stocks from concentrations. The potential uncertainty which arises by applying a pedotransfer function developed in a particular area, and which is then applied on different sites (Schillaci et al., 2021) can diminish the precision of final results.

1.2. Available guidelines and their applicability

So far, there are no collaborations or guidelines for publishing systematic reviews or meta-analyses on agricultural or soil issues. In contrast, healthcare (The Cochrane Collaboration) and social sciences (The Campbell Collaboration) established such

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

collaborative networks to develop high quality reviews already in the 1990s (Gurevitch et al., 2018; Collaboration for Environmental Evidence, 2018). These collaborations are focusing on specific disciplines and some of their tools, as trainings or the Cochrane Handbook for Systematic Reviews of Interventions, are partly applicable for agricultural and soil research (Table S1). Moreover, there are other voluntary guidelines available, which aim to support researchers in e.g., reporting or producing meta-analyses. Checklists for evaluating social science research synthesis (Cooper et al., 2019a) or evidence-based minimum item sets for reporting in systematic reviews and meta-analysis as PRISMA (Page et al., 2021) support synthesis consumers and authors. PRISMA-EcoEvo is a PRISMA extension for syntheses in ecology and evolutionary biology, which can be used for reporting, planning, registration and reviewing (O'Dea et al., 2021). Moreover, for meta-analyses in ecology, a checklist of quality criteria is available (Koricheva and Gurevitch, 2014). The Collaboration for Environmental Evidence (CEE) provides guidelines and standards for evidence synthesis in environmental management, which can be used for conducting, commissioning or using the findings of systematic reviews and systematic maps in environmental management. Further, reporting standards (ROSES), a checklist for appraisal of confidence of evidence reviews (CEESAT) and free-to-access online training courses are offered by CEE. The collaboration even brought forth "Environmental Evidence", a journal facilitating the publication of evidence synthesis in environmental management (https://environmentalevidencejournal.biomedcentral.com/). Lastly, reviews by Philibert et al. (2012), Beillouin et al. (2019) and Krupnik et al. (2019) assessed the quality of agronomic meta-analyses or compared different meta-analytical methods with the help of quality criteria. However, they are formulated rather generally.

Although all these guidelines are available, they each use different criteria which are sometimes not reported exhaustively (Koricheva and Gurevitch, 2014), making it difficult to apply them interdisciplinarily (Nakagawa and Cuthill, 2007; Lortie et al., 2015), as for the quality assessment of meta-analyses in agricultural and soil sciences. Additionally, as mentioned above, soil and agricultural scientists encounter specific issues different to ecology or medicine, when aiming to synthesize research outcomes meta-analytically. The guidelines and standards for evidence synthesis in environmental management and the CEESAT checklist by CEE clearly benefit scientists and other consumers of soil and agricultural meta-analyses, but do mainly focus on systematic reviews and maps and contain elements not necessary in meta-analysis (e.g. registration, gathering a maximum of available relevant literature or performing critical appraisal). Moreover, the guideline is exhaustive and requires inexperienced readers time and effort to understand. Many, who are not aiming to become experts in the method themselves, might not be able to find the time for such an elaborate reading.

1.3. Why we need meta-analytical guidelines in agricultural and soil research

The contribution of agriculture to the global anthropogenic greenhouse gas (GHG) emissions (Tubiello et al., 2015) and the possibilities of sequestering carbon through improved soil management in the form of SOC (Smith, 2012; Paustian et al., 2016; Smith et al., 2005) are topics that have occupied soil and agricultural researchers over the last decades. Since 2000, the number of articles published on SOC has increased yearly (Fig. 1), due to climate change pushing the scientific community to search for mitigation and adaption opportunities in numerous ways, such as through agronomic practices. Carbon sequestration in soils has gained increased resonance on the EU political agenda (EU Green Deal, Farm to Fork Strategy, EU Soil Strategy for 2030) - especially since the launch of "4 per mille initiative - Soils for Food Security and Climate" at COP21, and the publication of the global potentials of this initiative (Minasny et al., 2017).

Simultaneously, the number of meta-analyses published in the field of agriculture is continuously rising. We searched the Web of Science Core Collection for all available entries on "meta-analysis AND agriculture" since the year 2000 (Fig. 1, search conducted January 13th, 2022). Between 2000 and 2010, there was little change in the number of meta-analyses published; a steady

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

rise can only be seen since 2010. The increasing amount of available information, not only in agriculture and SOC research but across all scientific fields, is creating the need to synthesize data into a form which is easier to comprehend and allows the detection of overarching patterns (Culina et al., 2018). Unfortunately, as a consequence of the rising popularity of this method, more and more publications refer to their synthesis work as meta-analyses, despite applying less than rigorous methodologies. Many times, the term is misapplied to publications synthesizing information of primary studies, regardless of the methodologies used (Gurevitch et al., 2018). In fact, only studies using well-established statistical procedures - most importantly suitable effect-size calculation, correct weighting by the inverse of variance, analysis of possible heterogeneity and appropriate statistical models which account for the structure of the meta-analytical data - should use the term “meta-analysis” to describe their synthesis method (Vetter et al., 2013; Gurevitch et al., 2018).

The previously mentioned reviews by Philibert et al. (2012) and Krupnik et al. (2019), who analyzed the quality of meta-analyses in agronomy, found that the overall quality of meta-analyses in this field is low. Philibert et al. (2012) concluded that more than half of the publications in the searched databases mentioned meta-analyses as a method but did not carry the method out. Further issues regarding effect size metrics, weighting, and heterogeneity analysis were found. The more recent review by Krupnik et al. (2019), which analyzed meta-analyses studying the effects of conservation and organic agriculture on yield, also reported lacks in heterogeneity testing and weighting. Similarly, Beillouin et al. (2019), who studied meta-analyses on crop diversification, found issues on weighting, sensitivity analysis and database presentation. These results imply that the methodology applied in agronomical meta-analyses is variable and often not done according to standard metrics. The authors of the reviews concluded that there is a need for improvement of meta-analyses in agronomy.

Finally, it is a misconception that a high number of citations always equals quality (Aksnes et al., 2019; Leydesdorff et al., 2016). Koricheva and Gurevitch (2014) found that even in high-impact journals, cases of incorrect usage of the term “meta-analysis” can be encountered. This suggests that not only authors but also peer reviewers and journal editors do occasionally misunderstand the rules under which a meta-analysis must be conducted. O’Leary et al. (2016) analyzed the effects of journal impact factor on review quality and concluded that a high impact factor does not guarantee high quality of reviews and therefore did not recommended to use impact factor as a proxy for review quality.

All this provides reason to assume that core criteria, necessary in conducting meta-analyses, are not clear to many researchers in the field of agricultural and soil sciences. As a result, poor quality meta-analyses are published, which might report questionable conclusions and recommendations to other scientists, policymakers and farmers. Moreover, the interest in SOC sequestration and subsequent increase in related publications raises the question whether there are meta-analyses synthesising this knowledge. If so, does their quality show similar trends to agricultural meta-analyses reviewed in the past by Philibert et al. (2012), Beillouin et al. (2019) and Krupnik et al. (2019)?

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

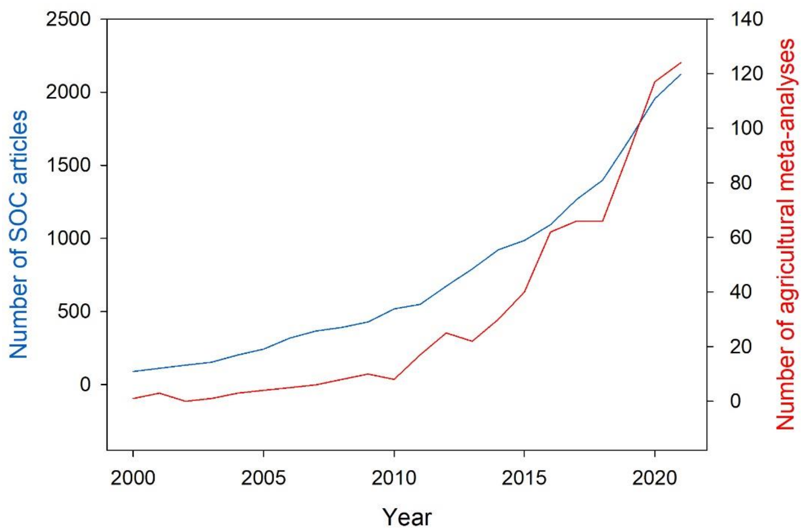

Figure 1. Number of meta-analyses in agriculture and primary research articles on soil organic published between 2000-01-01 and 2021-12-31 (search conducted on the 13.01.2022 on Web of Science Core Collection, searched in "Topic", results taken from WoS "Analyse Results" tool; Boolean search string for MA in agriculture: meta-analysis AND agriculture, carbon; Boolean search string for articles on SOC: "soil organic carbon")

1.4. Objectives

This study aims to quantitatively analyze 31 meta-analyses, studying the effects of different management practices on SOC, relevant for European cropland, published between the years 2005-2020. We compiled a quality criteria-set suitable for soil and agricultural sciences by adapting existing meta-analytical guidelines from other disciplines. The set is supported by a scoring scheme, which allows a quantitative analysis. A subsequent evaluation of the management practices studied in these SOC meta-analyses gives information on which agricultural operations require more or improved research. Finally, the aim was to demonstrate how to conduct a quick assessment of meta-analyses relevant for decision making, such as the IPCC report, by using the most critical criteria, as their presence is strictly necessary for a research synthesis that intents be termed as "meta-analysis".

2. Material and methods

2.1. Quality criteria-set

The quality criteria-set is based on the previous work of many experienced researchers with expert knowledge on meta-analysis (Table S1). The "Checklist of quality criteria for meta-analysis for research synthesis, peer reviewers and editors" by Koricheva and Gurevitch (2014) was used as a basis for the composition of the 17 quality criteria (Table 1). Their checklist is also built upon the previous efforts of other scientists who established quality criteria-sets in the fields of ecology, environmental management, conservation biology and agronomy. Other literature such as, "Introduction to Meta-Analysis" by Borenstein et al. (2009), "Handbook to Meta-analysis in Ecology and Evolution" by Koricheva, Gurevitch and Mengersen (2013), and "Handbook of Research Synthesis and Meta-Analysis" by Cooper, Hedges and Valentine (2019c) further supported the criteria construction and acted as sources for in depth explanation of those criteria, providing the reader with additional information (Table S2).

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

The 17 quality criteria were structured according to three groups: "Literature Search and Inclusion / Exclusion Criteria", "Meta-analysis", and "Results and Database Presentation". Additionally, a further division of the "quality criteria" into "sub-criteria" was conducted to provide a more detailed guidance. Each quality criterium or - if available - sub-criterium, was specified with the help of the column "Is criterion applied in meta-analysis (to which extent)", which offers the reader possible options, based on the availability of data or items within the analyzed meta-analysis. Each option ends with a numerical "Score", which indicates its quality. All individual scores can be summarized into a total score with a maximum of 28; the higher the total score, the better the overall quality of the meta-analysis. Furthermore, the quality- and sub-criteria were specified in the column "Description" to provide the reader with more detailed information. The final column offers references of relevant literature, supporting the authors' decisions on criteria formulation and scoring. In the supplementary material (Table S2) an extended version of this column can be found, where direct quotes of cited experts are provided.

Of these 17 quality criteria, we defined three as so called **"cut-off" criteria (criteria 6-8 in Table 1)**, namely "effect size", "standard deviation extracted" and "studies weighted by 1/variance". When these criteria are not fulfilled by a meta-analysis, the most essential and relevant steps in this specific synthesis method are not met. These "cut-off" criteria aim to help consumers of soil and agricultural meta-analyses to identify the defining elements of the article and judge whether it is a "true" meta-analysis or not.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Table 1. Quality criteria-set for meta-analyses in agricultural and soil research.

Group	Quality criteria	Sub-criteria	Is criterion applied in Meta-analysis (and to which extent)	Score	Description	References
Literature search and inclusion / exclusion criteria	1. Literature search	Published literature extracted from	> 4 databases	3	Several databases should be used for extracting published literature to reduce the risk of selection bias	Côté et al. 2013, p. 40
			between 2 and 4 databases	2		
			1 database	1		
			not reported	0		
		Grey literature (unpublished reports and experiments, project reports etc.) included	yes	1	Grey literature maximizes comprehensiveness and reduces risk of bias. Whether conducting a grey literature research is necessary or not is dependent on the meta-analysis itself and needs to be assessed by the authors	Borenstein et al. 2009, p. 280
			no	0		
		Keywords/search string reported	yes	1	The search string(s) used to retrieve literature from different databases should be stated	Côté et al. 2013, p. 43
			no	0		
	2. Authors checked the reference lists of other existing meta-analyses and reviews for available literature		yes	1	Ensures the inclusion of more relevant articles, as occasionally, keyword searching in databases does not provide results for all available literature	Borenstein et al. 2009, p. 278
			no	0		
	3. Inclusion and exclusion criteria reported		yes	1	Inclusion and exclusion criteria should be clearly described and decisions for exclusion of studies should be transparent	Côté et al. 2013, p. 50
			no	0		
	4. Control (C) and treatment (T) described		yes	1	Inclusion of studies on the same research topic and avoiding mixing "apples and oranges"	Stewart et al. 2013, p. 28
			no	0		
	5. Moderators and their range or groups described		yes	1	Defining moderators is essential to evaluate the source of variation across studies. Their range or groups are important to indicate the limits within which moderators were studied	Stewart et al. 2013, p. 32
			no	0		
	6. Effect size	ln(R) (log response ratio)		2	Easily interpretable by back-transformation of ln (R) to a percentage change from the control	Rosenberg et al. 2013, p. 63f
		Raw mean difference (D)		1	Not recommended for meta-analyses having a range of control levels/scales. Example: when SOC stocks are studied, initial values can have a wide range (10-100 t/ha)	Borenstein et al. 2009, p. 21ff
		Standardized mean difference (e.g., Hedges´d)		1	Difficult to interpret; |d | = 0.2 – small effects |d | = 0.5 – moderate effects |d | = 0.8 – large effects	Borenstein et al. 2009, p. 26; Rosenberg et al. 2013, p. 63f

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Meta-analysis		non-standard metrics used or not calculated		0	Without the calculation of effect sizes, a synthesis does not qualify as a meta-analysis (all following quality criteria of group "Meta-analysis" account for 0)	Koricheva and Gurevitch 2014, p. 840
	7. Standard deviation extracted		From each study	2	Standard deviations need to be extracted from all studies in order to calculate the weight for each study. They should not be estimated	Nakagawa et al. 2017, p.11
			From some studies	1		
			Not extracted	0		
	8. Studies weighted by 1/variance		For each study	2	Weighting of studies is only correct when done by the inverse of variance. Meta-analyses that extracted SD only from some studies but weighted by 1/variance did weigh "some studies" or "partly"	Koricheva and Gurevitch 2013, p. 9
			For some studies	1		
			Not weighted / reported	0		
	9. Subgroup analysis and meta-regression		yes	1	Categorical and/or continuous moderators should be assessed by Q-test	Rosenberg et al. 2000, p. 111f
			no	0		
	10. Model used	Random-effect or mixed-effect model		1	When conducting a meta-analysis in the field of agriculture, the random-effect or mixed-effect model should be chosen, as it acknowledges between-study variation	Mengersen et al. 2013, p. 94
		Fixed-effect model		0		
		No model reported		0		
	11. Software used for meta-analysis		Meta-analytical software (as MetaWin, Metafor package, etc.) or other software (as SPSS, SAS, Stata, R, etc.)	1	Used software should be stated; when using general statistical analysis software, correct model choice (weighted + random model) and implementation are necessary	Schmid et al. 2013, p. 174
			Spreadsheet (as MS Excel) or not reported	0		
	12. Independence of effect sizes		1-2 effect sizes per study/site extracted	1	Effect sizes should be independent. Golden rule: one study or site, one effect size extracted. When several combinations of treatment and control were studied, only one effect size per study/site should be extracted.	Gurevitch and Hedges 1999, p. 1147; Hungate et al. 2009, p. 2009f; Nakagawa et al. 2017, p. 3
			> 2 effect sizes per study/site extracted	0		
	13. Response variables and relevant parameters for their calculation were measured		yes	1	Response variables need to be measured in an experiment, not estimated or modelled (e.g., pedotransfer functions only provide estimates for bulk density and therefor introduce the risk of inaccurate calculation of SOC stock contents)	Xu et al. 2015, p. 1574
			no	0		
	14. Sensitivity analysis to test robustness of meta-analysis	Outliers and effect size distribution	yes	1	Presence of outliers should be tested and can be identified via effect size distribution in weighted histograms, box-plots, etc.	Rothstein et al. 2013, p. 333
			no	0		
		Detecting publication bias	yes	1	Magnitude of publication bias should be esmimated by funnel plots, Egger's regression or Fail-safe test	Borenstein et al. 2009, p. 291
			no	0		

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Results and database presentation	**15. Results presentation in figures and tables**	Summarized effect size and Confidence Intervals	yes	1	Summarized effect size and confidence intervals should be presented in a table or figure	Borenstein et al. 2009, p. 6
			no	0		
		Moderator analysis (sub-group analysis and/or meta-regression)	yes	1	Moderator (covariates) analysis should be presented in form of figures or tables	Borenstein et al. 2009, p. xxii
			no	0		
		Forest plot	yes	1	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	Lortie et al. 2013, p. 344f
			no	0		
	16. Description of meta-data		yes	1	A description of meta-data (authors and year plus experimental location, treatments, etc.) in the article or appendix should to be provided as a table	Gonçalves and Musen 2019, p. 2
			no	0		
	17. Full database including most of the following criteria for each study	Article ID and/or first author plus year	Available and includes most of the listed criteria; Article ID and/or first author plus year must be described; either effect size or mean, SD and sample size of T/C must be described	2	Each study should have a distinctive number to allow easy identification First author and year of publication should be stated	- -
		Country/location			Country and exact location of experiment	-
		Control (C)			Measure used as control	-
		Treatment (T)			Measure under investigation	-
		Moderators			Possible moderators (pedo-climatic, experimental conditions, duration of experiments, land use/crops/cropping systems)	Curtis et al. 2013 p. 53
		Means of C/T Standard deviations of C/T Sample sizes C/T			Means, SDs and sample sizes of treatment and control should be stated to allow replication of the meta-analysis	Curtis et al. 2013, p. 52
		Effect size			Calculated effect sizes should be stated	
		Article ID and/or first author plus year Means of C/T Standard deviations of C/T Sample sizes C/T Effect size	Available and includes the listed criteria (either effect size or mean, SD and sample size of T/C must be described)	1	see above in point 17	see above in point 17
			Not available or includes less information than score 1	0	If no database is provided, the meta-analysis is not transparent	Mayo-Wilson and Grant, 2019, p. 481
			Maximum reachable score per meta-analysis:	**28**		

2.2. Quality assessment of meta-analyses on SOC

2.2.1. Inclusion criteria, exclusion criteria and search strategy

First, inclusion (IC) and exclusion criteria (EC) were defined to create a framework for the literature screening (Table 2). Studies were included when they (IC1) used the term "meta-analysis" in their title, abstract or author keywords. (IC2) Land uses included were arable- or crop land, also in combination with others as e.g., agroforestry or grassland. The (IC3) assessment of the effects of one or several management practices on SOC needed to be the aim of the study. Moreover, (IC7) European experiments needed to be a part of the (global) meta-analyses, as we wanted to collect and evaluate syntheses relevant for Europe. Articles were excluded when, for example, modelling was used to obtain SOC results (EC1).

Table 2: Inclusion and exclusion criteria for the literature screening process.

	Inclusion criteria (IC)	Exclusion criteria (EC)
1.	Term meta-analysis used in title, abstract or keywords to describe study style	Systematic reviews and studies using modelling to obtain results
2.	a) Cropland/arable land needs to be part of study; b) other agricultural forms as e.g., agroforestry, paddy soils/upland soils, grassland can be part of study	a) If primary data are from one experimental site (literature not found through database search - not possible to evaluate according to our criteria-set); b) Land-use change studied; c) Cropland/arable land plus forest studied (forest not comparable to arable land)
3.	Effects of management practice on total SOC stocks or concentrations studied	Impact on SOC fractions investigated
4.	Management practice effects on SOC is central topic	Management practice effects on SOC is not a central topic
5.	Field experiments	Laboratory experiments
6.	Conducted on mineral soils	Conducted on organic soils
7.	European studies need to be part of studied experimental sites	Included only non-European experimental sites

The second step was the collection of existing meta-analyses on SOC changes due to different agricultural management practices. Therefore, the Web of Science Core Collection (timeframe 1900-2020) and Scopus (timeframe 1960-2020) databases were searched on January 5th, 2021. Due to limited human resources, only these two scientific databases were searched. The following Boolean search string was used to retrieve relevant articles: (meta-analy*) AND soil AND (agriculture OR management) AND (SOC OR OC OR "soil organic carbon" OR "organic carbon"). 552 articles were found (344 and 208 in Web of Science and Scopus, respectively) and automatic (conducted by Mendeley and JabRef software) and manual duplicate removal reduced the results by 167 articles (Fig. 2). The results were compared with the meta-analyses identified by Bolinder et al. (2020), who synthesized meta-analyses studying the effects of several management practices on SOC changes in agroecosystems. This led to the identification of one further study which was included in our evaluation. 386 articles were exported into excel and screened by title, abstract and full text according to the pre-defined inclusion- and exclusion criteria. In total, 31 meta-analyses relevant for the scope of our study were found. Figure 2 shows a flow diagram of the complete screening process. The full information of the literature

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

gathering, all 386 retrieved articles plus the screening decisions can be found in the supplementary material (Table S3 and S4, respectively). The complete reference list of the 31 meta-analyses can be found in the appendix (Table A1).

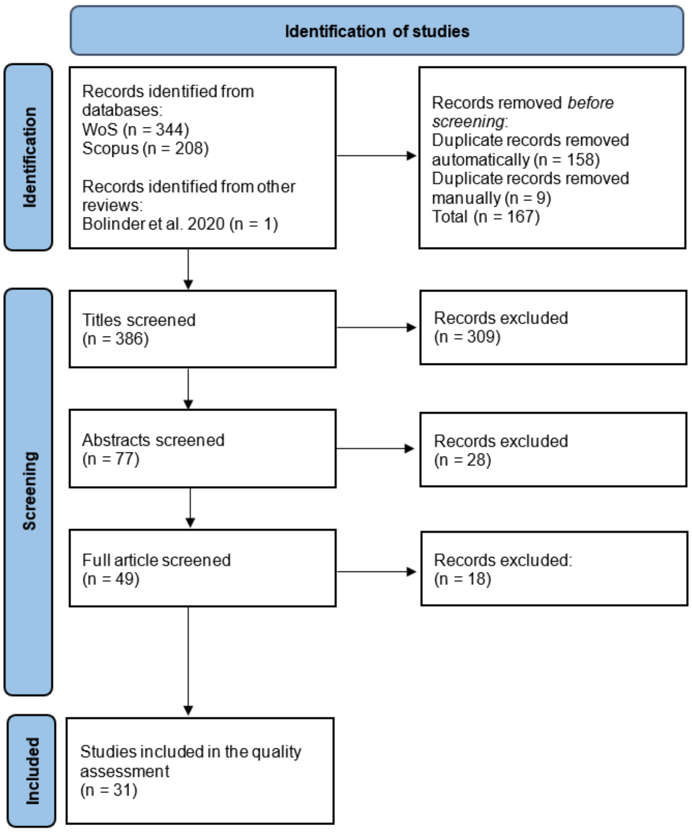

Figure 2. Flow diagram of literature search and screening. Adapted from: Page et al. (2021)

2.2.2. Quality analysis

The 31 retrieved meta-analyses were analyzed by two authors for their quality according to the quality criteria-set in Table 1. Each article was read thoroughly to ascertain whether certain criteria were fulfilled or not. Total scores for each meta-analysis were calculated, with a maximum reachable score of 28. The complete dataset containing the scores for each meta-analysis and all calculations can be found in the supplementary material (Table S2, S5). SigmaPlot version 14.5 and Microsoft Excel version 1808 were used for plotting of figures and tables and for calculations.

2.2.3. Management categories

The retrieved data also offered the possibility to analyze the “state of knowledge” on meta-analyses studying management effects on SOC. The aim was to assess how many meta-analyses were conducted on a certain management practice and whether their quality was sufficient to stop the production of new meta-analyses on the respective practices. This information will aid future research by guiding it towards knowledge needs and avoiding redundant work. We therefore grouped the meta-analyses according to the

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

management practices they studied. 11 management categories were formed and are described in Table 3. These categories aim to structure the collected SOC meta-analyses and allow a simplified investigation. As some meta-analyses studied the effects of more than one practice, they were added to all respective categories.

Table 3: Defined management categories, their included management practices and meta-analyses that studied their effects on SOC.

Nr.	Category	Description	Meta-analyses
1.	Tillage	no-till, reduced and deep tillage	Aguilera (2013), Angers (2008), Bai (2019), Cooper (2016), Feng (2020), González-Sánchez (2012), Haddaway (2017), Kopittke (2017), Li (2020), Luo (2010), Meurer (2018), Mondal (2020), Ogle (2005), Sun (2020), Virto (2012)
2.	Organic	organic practices	Aguilera (2013), Cooper (2016), García-Palacios (2018), Gattinger (2012), Kopittke (2017), Tuomisto (2012)
3.	Cover crop	cover crops used in crop rotation	Aguilera (2013), Bai (2019), González-Sánchez (2012), Jian (2020), Poeplau (2015), Sun (2020)
4.	Residue	crop residues were either left or removed from the field	Han (2016), Li (2020), Sun (2020), Xia (2018), Xu (2019)
5.	Fertilization	application of organic or mineral fertilizer	Aguilera (2013), Han (2016), Ladha (2011), Xia (2018)
6.	Amendments	application of amendments (e.g., manure)	Aguilera (2013), Chen (2018), Kopittke (2017), Maillard (2014)
7.	Biochar	application of biochar	Bai (2019), Liu (2016), Majumder (2019)
8.	Diversification	more or different crops were used in rotation	King (2018), Mathew (2020), McDaniel (2014)
9.	Combined	effect of several practices combined was studied	Aguilera (2013)
10.	High input system	system that aims in increasing carbon by e.g., irrigation, winter crops, etc. according to IPCC (1997)	Ogle (2005)
11.	Set-aside	effect of setting-aside land from crop production and planting trees or grasses	Ogle (2005)

Finally, the total number of articles per category were calculated and meta-analyses with the highest scores identified. Simultaneously, information on treatment and control, the geographical scale and soil depth were extracted. As the overall score does not give information on whether the "cut-off" criteria were fulfilled, we extracted this information as well. We presented the overall effect sizes of the meta-analyses only when both these elements were fulfilled.

Overall treatment effects on SOC are shown in percentage change from the control; when results were displayed in log response ratio (LnR), we calculated percentages with the Eq. (1):

$$\%\, change = (\mathrm{Exp}\ (\mathrm{LnR}) - 1) * 100\% \quad (1)$$

2.3. Quick assessment of meta-analyses relevant for policy making – An example

To provide readers with an example of the impacts of meta-analytical quality on policy- and decision making, we screened Chapter 2: "Land–climate interactions", of the Intergovernmental Panel on Climate Change (IPCC) "Special Report - Climate Change and Land" (Jia et al., 2019) for cited articles which used the term "meta-analysis" in the title. We chose this report by the IPCC, as their outputs are highly relevant for combating the global climate crisis and are often the basis of policy-making (IPCC, 2019), and because this exact chapter is deeply connected to the contents of this review. In total, 16 articles were retrieved and checked against the cut-off criteria of the quality criteria-set (Table S6).

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

3. Results

The investigation of meta-analyses, studying management effects on SOC published between 1990 and 2020, found that Ogle et al. (2005) published the first article on this topic. Nevertheless, the synthesis did not qualify as a formal meta-analysis, as no effect size was calculated. The first formal meta-analysis on SOC was published by Luo et al. (2010), who looked at the effect of no-till versus conventional tillage. Overall, the number of SOC meta-analyses, published between 2005 and 2020, increased over time (Fig. 3A). Scores also experienced a rise (15-year period) and related with the publication year (y= -1889.8980+0.9437*x; R^2= 0.39) (Fig. 3B) (normal distribution of scores tested with Shapiro-Wilk test; P= 0.052). If the observed rise in quality is projected into the future, without any intervention, a score of 28 will only be reached by the year 2032. As the meta-analysis by Haddaway et al. (2017) (ID= 10; score= 27) is an outlier which influences the regression result, we also calculated how the prognosis would change if we removed this meta-analysis. The new regression line (y= -1813.1622+0.9054*x; R^2= 0.45) estimates that scores of 28 will be reached in 2033.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

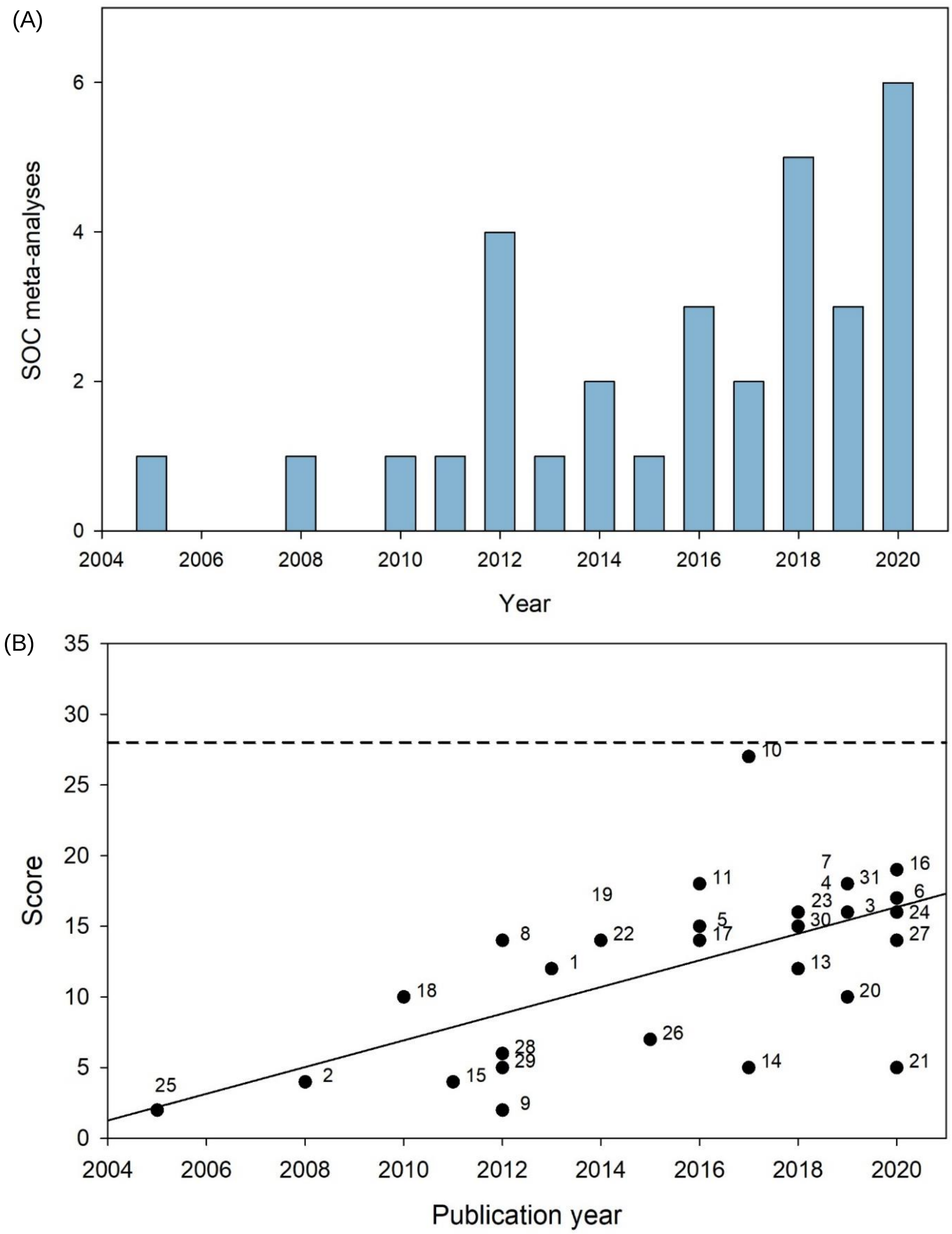

Figure 3. (A) Number of SOC meta-analyses published per year. (B) Scores of SOC meta-analyses over time (between 2005-2020) and corresponding regression line. Numbers beside dots indicate SOC meta-analysis ID (ID and linked author information in Table A1 and Table S2). Dashed line indicates maximum score 28.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

3.1. Literature search and inclusion / exclusion criteria

This group consisted of five quality criteria. The first criterium, “Literature Search”, was satisfied by more than half of the meta-analyses (Fig. 4). In nearly a quarter of the analyses, authors checked the reference lists of other existing meta-analyses and reviews for available literature. Therefore, the usefulness of this method seems to be widely underestimated. By comparing retrieved literature to other existing publications, we can not only gain confidence in our search strategy, but also encounter information which might be difficult to find otherwise (e.g., grey literature).

Inclusion and exclusion criteria, as well as a description of treatment and a control, were presented by almost all meta-analyses. Moderators were described by over half of SOC meta-analyses. Description of moderators, including their range (for continuous explanatory variables) or groups (for categorical explanatory variables) are necessary to present the way in which moderator analysis will be conducted. Results for the sub-criteria can be found in the supplementary material (Table S5).

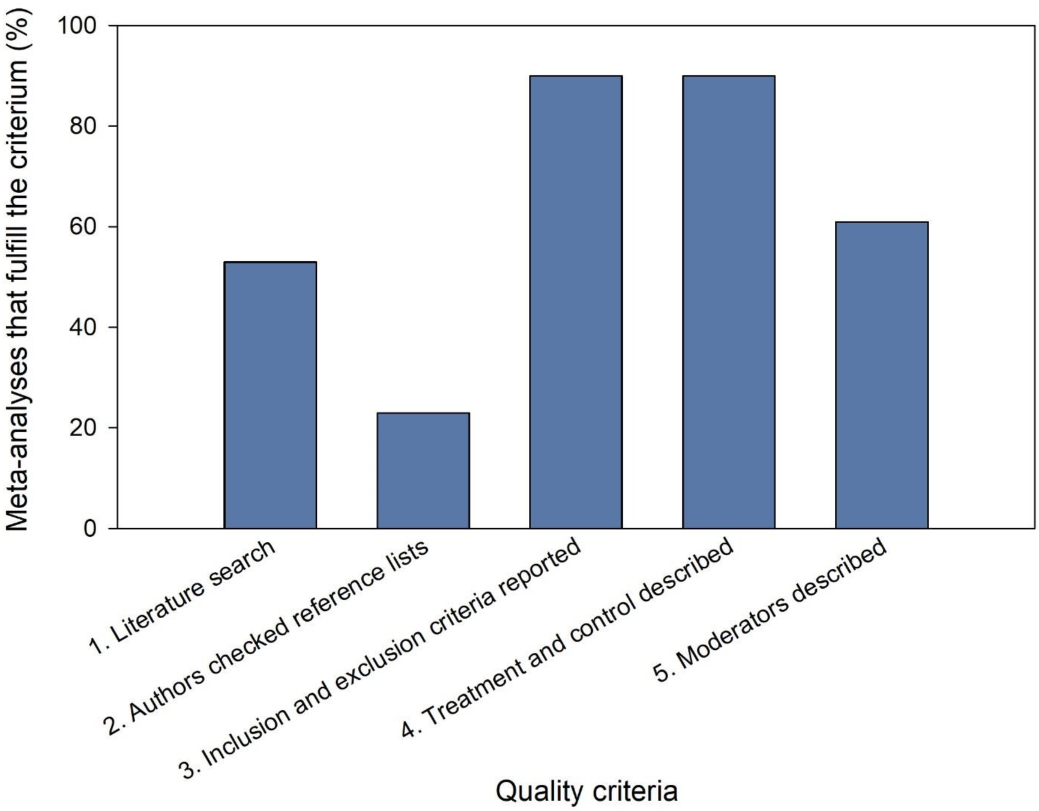

Figure 4. Compliance of meta-analyses with the criteria in group „Literature search and inclusion / exclusion criteria”.

3.2. Meta-analysis

The “Meta-analysis” group consisted of nine quality criteria, which were satisfied by the SOC meta-analyses to variant extents. Effect sizes were calculated according to standard metrics by 74% of meta-analyses (Fig. 5A). Almost half of meta-analyses used log response ratio for effect size calculation and about a third applied raw mean difference or standardized mean difference.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Standard deviation (SD) was extracted from all primary studies by 16% and partly by 42% of meta-analyses. Weighting each study by 1/variance was done by 13% of meta-analyses, whereas 19% weighted partly (Fig. 5B) (for a detailed description of the criteria for weighting, see quality criterium number eight in Table 1). Accordingly, weighting was not done in over two thirds of analyses. We classified these three criteria (effect size estimate, SD extracted and weighting by 1/variance) as "cut-off" criteria (6-8 in Table 1). When these are not fulfilled, a meta-analysis does not account as such. In our quality assessment, we acknowledged when authors partially weighted by the inverse of variance (as they only partially extracted SD) with one point for each. Nevertheless, we urge authors to extract SDs for each study and further weight them by the inverse of variance in order to conduct a high-quality meta-analysis.

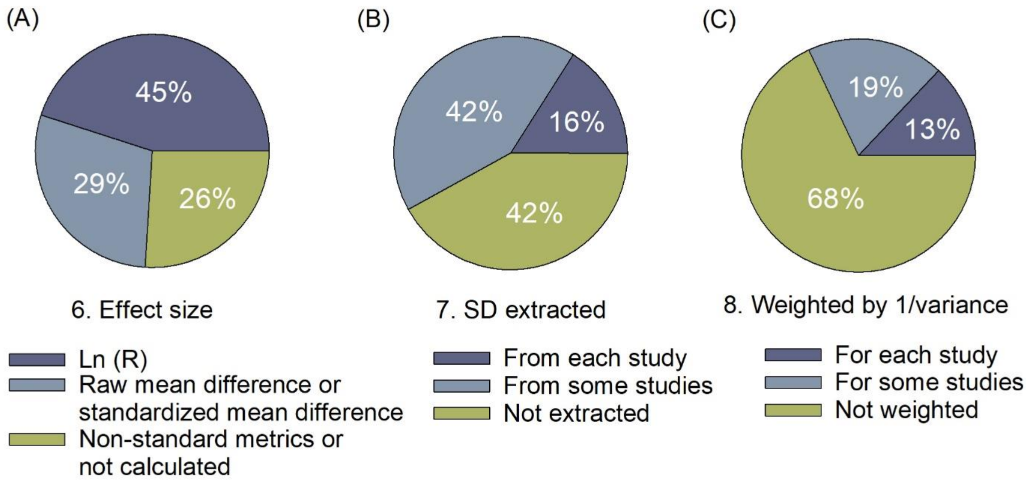

Figure 5. Compliance meta-analyses with "Cut-off" criteria in the group "Meta-analysis": (A) Ratio of effect size metrics used by the meta-analyses. (B) Ratio of meta-analyses which extracted standard deviations from each study. (C) Ratio of meta-analyses which weighted by the inverse of variance.

In Figure 6, satisfaction of criteria following the "cut-off" criteria are displayed for 1) all studied SOC meta-analyses and 2) only meta-analyses that fulfilled the "cut-off" criteria. In the following, we will describe only the results for all SOC meta-analyses. For results regarding the "cut-off" criteria, please refer to the figures. Corresponding data used for the calculation of these results can be found in the supplementary material (Table S7).

Subgroup analysis and meta-regression, which identify the source of variation between studies, were assessed by almost half of meta-analyses (Fig. 6). Models applied and software used were reported more frequently. Only about 25% of meta-analyses had no problems with non-independence of effect size, while the rest extracted several effect sizes per study. Bulk density was measured in 35% of meta-analyses, the other 65% used pedotransfer function to estimate this parameter, therefore introducing a source of uncertainty in SOC stock estimation. Lastly, sensitivity analysis of the meta-analytical results was done rarely.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

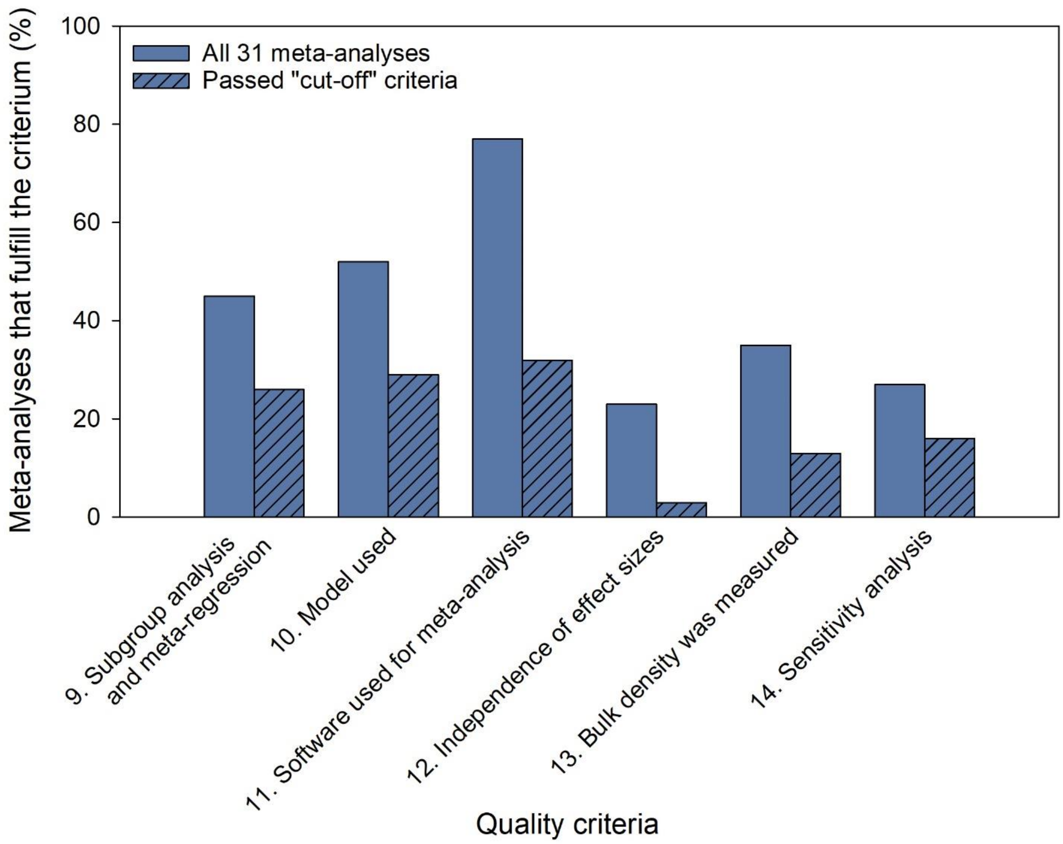

Figure 6. Compliance of meta-analyses with the criteria 9-14 in the group "Meta-analysis".

3.3. Results and database presentation

Figure 7 shows the results for the group "Results and database presentation". Almost half of the meta-analyses displayed their results in the form of figures or tables. Summarized effect sizes and confidence intervals or moderator analysis were presented graphically or in tabular form by 65% and 68% of meta-analyses respectively. Forest plots were presented by 6% of meta-analyses. Meta-data was presented in over two third of analyses, whereas a full database was made available to the readers in 13% and partly in 3% of cases (for further explanation see criterium 17 in Table 1). Information on the calculation of these results can be found in the supplementary material (Table S5).

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

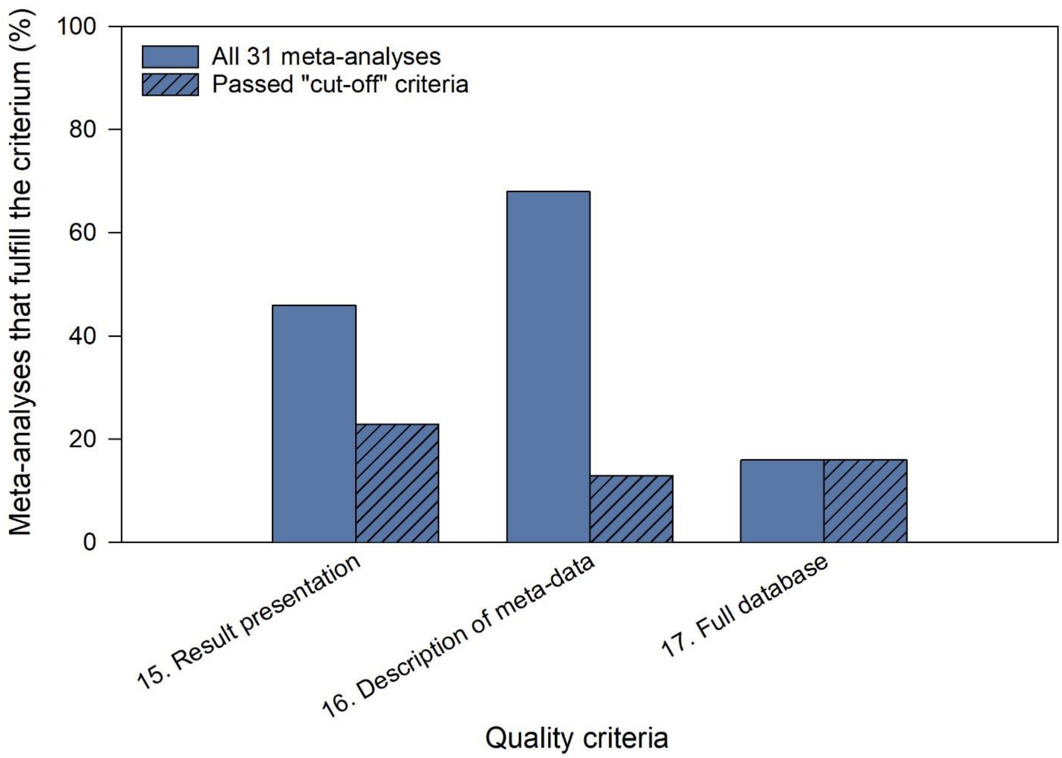

Figure 7. Compliance of meta-analyses with the criteria in the group "Results and database presentation".

3.4. Overarching findings

When looking at the overall results across the three quality criteria groups, quality varied greatly between the 31 analyses with a maximum score of 27, a minimum score of 2 and a median of 14. Haddaway et al. (2017) produced a meta-analysis of high quality which received the highest score according to our assessment. However, they used raw mean difference to calculate effect sizes, which may not be the most suitable for meta-analyses in the soil and agricultural field. In Sect. 4.2. "Meta-analysis" we will go more into detail on this issue. There were seven meta-analyses with scores up to five, the majority achieved scores between five and 15. 10 meta-analyses reached scores between 15 and 20, whereas only one reached a score above 20. Only four out of 31 meta-analyses weighted studies by the inverse of variance (Fig. 8).

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

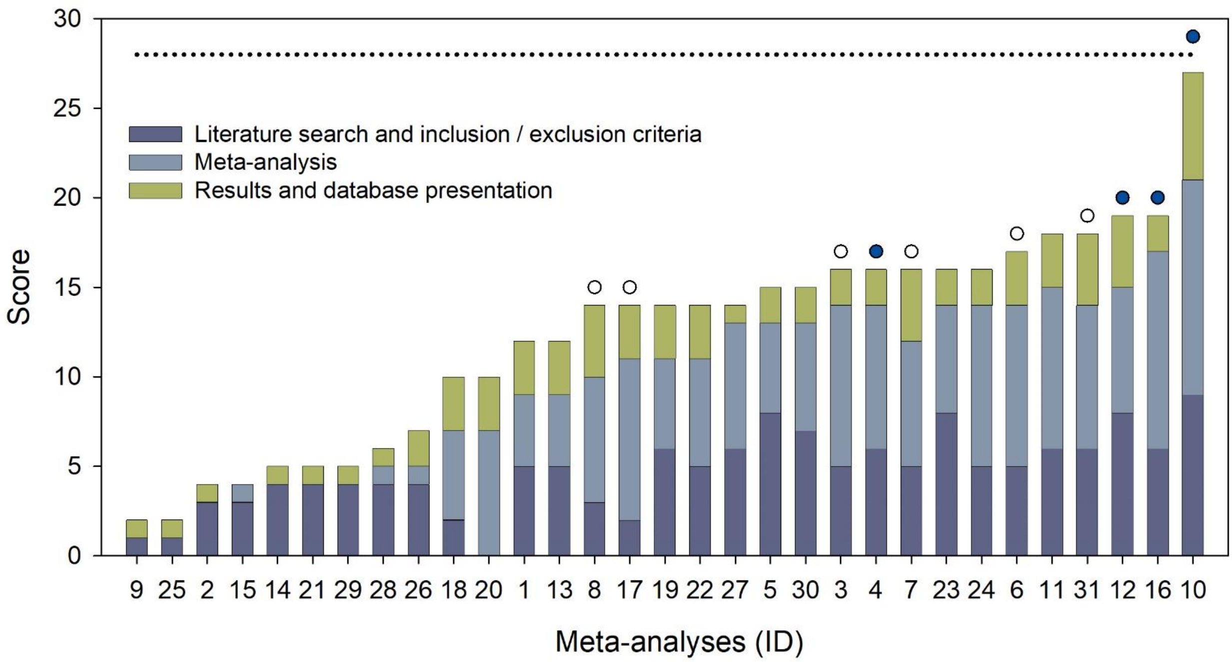

Figure 8. Scores of individual SOC meta-analyses displayed as scores per group. Sorted from lowest to highest achieved score. Meta-analysis ID and full reference information appear in Table A2. Dashed line indicates maximum reachable score of 28. Filled circles indicate meta-analyses which weighted each study by inverse variance. Open circles indicated meta-analyses which weighed some studies by inverse variance.

3.5. Analyzing management categories

Management practices studied in the meta-analyses were counted in order to assess their incidence. We found that almost half of the 31 meta-analyses studied the effects of tillage on SOC (in some cases besides other management practices) (Table 4). Other practices studied frequently were “organic agriculture” and “cover crop cultivation” (6 times each). Data on “residue”, “fertilization”, “amendments”, “biochar” and “diversification” were synthesized less often. The effects of “combined practices”, “high input” and “setting aside” on SOC were each assessed once. We found that meta-analyses, which passed the “cut-off” criteria, are available for four out of the 11 management categories (tillage, cover crop, residue, amendment). For tillage, we decided to show the three meta-analyses with the best scores (Bai et al., 2019; Li et al., 2020; Haddaway et al., 2017), as several analyses above average quality were available. Nevertheless, only Haddaway et al. (2017) fulfilled the criteria for effect size calculation, SD and weighting, whilst also achieving an overall high score and is therefore the one publication providing a high-quality meta-analysis on the effects of management practices on SOC. In the categories “organic”, “fertilization”, “biochar”, “diversification”, “combined”, “high input” and “set-aside”, no meta-analyses conducted according to the standards are currently available. In the last column of Table 4, overall effect sizes for SOC can be found. As Haddaway et al. (2017) calculated effect sizes by raw mean difference, if was not possible to transform their results from stock into percentages. For the five management categories where no meta-analysis weighed by the inverse of variance (“fertilization”, “diversification”, “combined”, “high input system” and “set-aside”), overall effect sizes for SOC change are not displayed. When looking at the retrieved data on SOC changes per management category (Table

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

4), it is apparent that the largest increases of SOC compared to the controls were achieved in the categories “organic”, “cover crop”, “amendments” and “biochar”.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

364 **Table 4. Meta-analyses with the highest score per management category, effect size and weighting used, and overall SOC responses.**

	Number of times this practice was studied in all 31 meta-analyses	Meta-analysis with the highest score	Treatment and control	Scale of meta-analysis	Score	Effect size used (standard metrics)	Weighted by 1/variance	Soil depth (cm)	SOC response
Tillage	15	Haddaway et al. (2017)	no tillage - high intensity tillage	warm temperate climate zone; global	27	Raw mean difference	yes	0-30	4.61 Mg ha^{-1}
		Li et al. (2020)	no tillage - conventional tillage	global	19	ln(R)	yes	0-30	11%
		Bai et al. (2019)	conservation tillage (NT, RT) - conventional tillage	global	17	ln(R)	partly	0-100	8%
Organic	6	García-Palacios et al. (2018)	organic farming - conventional farming	global	16	ln(R)	partly	0-90; average 0-18	27%
Cover crop	6	Jian et al. (2020)	cover crop - no cover crop	global	19	ln(R)	yes	0-30; >30	15.5% or 30.0%*
Residue	5	Li et al. (2020)	no tillage with residue retention - no tillage without residue retention	global	19	ln(R)	yes	0-30	13%
Fertilization	4	Han et al. (2016)	-	global	18	ln(R)	no	-	-
Amendments	4	Chen et al. (2018)	organic amendments plus inorganic fertilization - inorganic fertilization	global	16	ln(R)	yes	7-30; average 15-20	29% and 49%, respectively

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Biochar	3	Bai et al. (2019)	biochar - no biochar	global	17	In(R)	partly	0-10	28%
Diversification	3	McDaniel et al. (2014)	-	global	14	In(R)	no	-	-
Combined	1	Aguilera et al. (2013)	-	Mediterranean climate; global	12	In(R)	no	-	-
High input system	1	Ogle et al. (2005)	-	temperate and tropical regions; global	2	no	no	-	-
Set-aside	1	Ogle et al. (2005)	-	temperate and tropical regions; global	2	no	no	-	-

Color scheme in green, yellow and grey aims to support the visual presentation of these outcomes. Meta-analyses which did not calculate effect sizes according to standard metrics or weighted by the inverse of variance do not qualify as meta-analyses; results are therefore not presented. The maximum achievable score is 28.

*Results for when all comparisons were included or only comparisons with SD values were included, respectively.

3.6. Example of quick quality assessment of meta-analysis, relevant for policy making

Our quick analysis of the IPCC special report (Jia et al., 2019) found that of 16 articles, 50% did not qualify as "true" meta-analyses, as five did not calculate effect sizes according to standard metrics and three which did fail to extract SD and to weight by the inverse of variance. The other half of the articles did in fact conduct meta-analysis correctly. Six meta-analyses used log response ratio to calculate effect sizes, two used standardized mean difference. These eight meta-analyses extracted SD for each study and weighted by the inverse of variance. Calculations and references of all 16 analyzed articles can be found in Table S6.

4. Discussion

Previous guidelines and expert knowledge on meta-analysis from other disciplines were adapted to construct an easy-to-use criteria-set for the quantitative quality assessment of meta-analyses in soil and agricultural research. With the help of these criteria, we analyzed 31 meta-analyses, studying the effects of different management practices on SOC. Moreover, the retrieved meta-analyses were structured according to 11 categories of agricultural management practices, which allowed us to assess and analyze the state-of-knowledge on these categories. Hence, recommendations for future meta-analytical research and general improvement of applied methodology can be given. We found major deficiencies in the reporting of literature searches, application of standard metrics for effect size calculation, correct weighting by the inverse of variance, extraction of independent effect sizes and database presentation.

In the following, we will discuss the results of the quality assessment of meta-analyses on SOC with the findings of four quality assessments of meta-analyses and quantitative reviews in agronomy and ecology. We included the study by Philibert et al. (2012), focusing on agri-environment and -biodiversity, the review of Krupnik et al. (2019), looking at conservation and organic agriculture, the study by Beillouin et al. (2019), studying crop diversification and the excellent evaluation of meta-analyses in plant ecology by Koricheva and Gurevitch (2014). To simplify the discussion, not all information for the 17 quality criteria was extracted from the reviews. Instead, we selected quality criteria to be discussed according to 1) the information available in most of the reviews, which allowed a comparison of results and 2) relevance (as e.g., effect size metrics), as certain quality-criteria are more important than others.

4.1. Literature search and inclusion and exclusion criteria

The comparison of reviews for the criterium "Literature search reported" showed that our study found higher compliance (41%) with this criterium than the ones of Philibert et al. (2012) or Koricheva and Gurevitch (2014) (Table 5). Beillouin et al. (2019) reported that 46% of meta-analyses presented the search string and 86% the eligibility criteria. Krupnik et al. (2019) found that all analyzed meta-analyses presented the literature search sufficiently. This high agreement may be caused by the small study number (n=17) or the definition of less demanding criteria by the authors.

A quality criterium, which is of special significance to the soil and agricultural field, is the inclusion of grey literature. Here, exceptionally large amounts of data are available, as governmental research activities are not focused on publishing results in scientific journals. Therefore, although the inclusion of grey literature is not compulsory, it is highly encouraged (Culina et al., 2018). When conducting meta-analyses on an international or global scale, analysts will find that grey literature is often available in national languages only, which complicates and restricts its inclusion. Nevertheless, the most essential part of searching for literature, whether scientific or grey, is complete reporting. Our results show that this reporting of search strategies is often limited. Therefore, essential information to allow reproduction of the study is lacking and possible differences in outcomes between meta-analyses, studying the same effects, cannot be fully explained. If a synthesis is not replicable, it cannot be fully trusted, as mistakes in methodological proceedings are possible (Haddaway et al., 2020; Parker et al., 2016). In another review, Hungate et al. (2009) showed how important complete reporting of search and screening strategy is. Lack of transparency prompted criticism on the results of meta-analyses. Non-identical time frames over which literature was gathered, differences regarding inclusion criteria and, in our eyes most importantly, limited search methods can influence the number of articles found and taken up into a meta-analysis. This indicates the need to draw quality criteria and disseminate good practices across research fields and to improve the power of meta-analytical results.

4.2. Meta-analysis

Effect size calculation is an essential and mandatory part of meta-analysis (Koricheva and Gurevitch, 2013). Therefore, the term "meta-analysis" should only be used when data is quantitatively synthesized as described in the textbooks of Borenstein et al. (2009), Cooper, Hedges and Valentine (2019c) and Koricheva et al. (2013). The investigation regarding the compliance of our SOC meta-analyses with the criterium "Effect size calculated according to standard metrics", showed that about three quarters of meta-analyses did calculate effect sizes according to such metrics. Koricheva and Gurevitch (2014) came to similar conclusions in their review of meta-analyses in plant ecology (Table 5). Further, only about half of SOC meta-analyses used log response ratio for effect size calculation.

These findings indicate that correct calculations of effect sizes are not applied consistently in the fields of SOC and plant ecology, although they represent the most fundamental and critical part in meta-analysis. Among the several possible choices in effect size metrics, we recommend using log response ratios when creating soil and agricultural meta-analyses. They are easy to interpret, and effect sizes are not affected by different variances of control and experimental groups. Overall, they are more suitable for meta-analyses studying agricultural management effects on soil parameters as e.g., SOC, than the standardized mean difference (Hedge's d). When using the standardized mean difference, the results are more difficult to interpret compared to log response ratios, which present the treatment effects in form of percent changes from the control. Moreover, effect sizes must be normally distributed, which is almost always the case when using log response ratios or the standardized mean difference.

In Sect. 3.3 "Results and database presentation", we mentioned that, in our opinion, raw mean difference (also called unstandardized mean difference) is not recommended for calculating effect sizes in the field of soil and agricultural research. Unlike response ratio, raw mean difference does not consider variations in control levels, which are often highly variable across field experiments, particularly, on a global scale. In case of SOC studies, control levels may vary between 10 and 100 t C ha^{-1}, which makes using raw mean difference between treatment and control as an index of effect size meaningless. It may result in similar effect sizes for the relatively large as for small responses, as illustrated in Figure 9. Therefore, raw mean difference can only be applied when all experiments studied in the meta-analysis are using the same scale (Borenstein et al., 2009). Raw mean difference usually does not result in a normal distribution of effect sizes, which is a prerequisite. Although this metric is easy to use, it may be suitable for meta-analyses when controls do not present a large variation across studies. That, however, is hardly possible to achieve for the diversity of pedo-climatic conditions.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

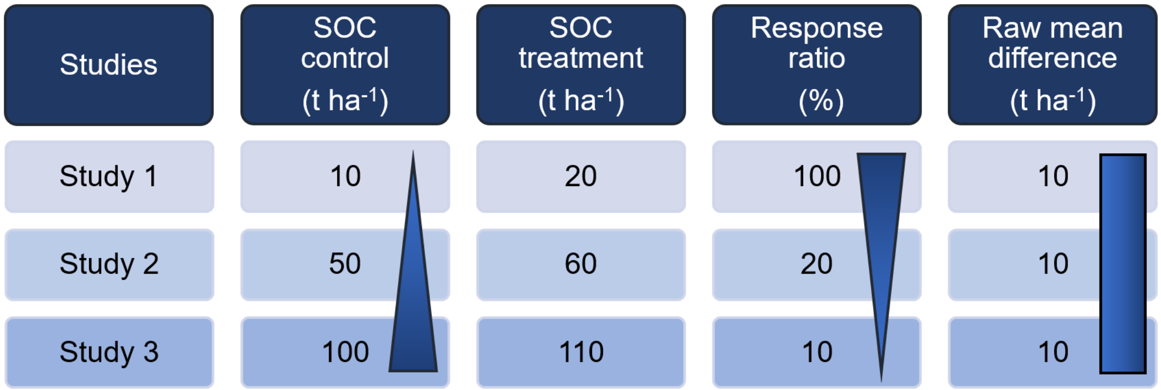

Figure 9. Example of the relationship between the SOC levels in control and effect sizes measured as response ratio or raw mean difference for three studies. Response ratio indicates increasing effect size with decreasing control level. Raw mean difference indicates equal effect sizes for all experiments and does not consider variation in control levels. Triangles indicate an increase or decrease of values; rectangle indicates constant values.

Weighting is essential, as different studies have different precision, and more precise studies with larger sample size need to be more heavily weighted in an analysis. The weighting should be done by the inverse of variance. Applying it in other ways, for example by sample size, can lead to several problems such as the introduction of unknown biases. When not weighted at all, variation within- and between-studies is not separated. Therefore, common- and random-effects models are not useable, leading to difficulties in assessing heterogeneity (Gurevitch et al., 2018). All these possible biases can adulterate the results of meta-analyses and therefore lead to false conclusions. According to findings by Hungate et al. (2009), depending on the functions used for weighting, differences in mean estimates of the effect sizes can be found. Weighting by sample size or not weighting resulted in comparable effect size estimates which often were larger than when weighted by inverse of variance. Our assessment showed that only 13% of SOC meta-analyses weighted by the inverse of variance, whereas Philibert et al. (2012) found 37% compliance. Koricheva and Gurevitch (2014) reported that three quarters of meta-analyses weighted by 1/variance. Meta-analyses studied by Krupnik et al. (2019) weighted by sample size, therefore are not correctly conducted according to our criteria-set. Beillouin et al. (2019) found that 40% of meta-analyses, studying diversification effects, weighted by 1/variance (and in some cases by sample size).

Effect sizes might show a certain amount of variability that cannot be explained by sampling errors alone, raising the question whether moderator effects may have influenced the results. A moderator is a third variable that conditions the relations between two others. Therefore, moderator analysis must be conducted to identify these effects (Lipsey, 2019). In agricultural and soil sciences, abiotic factors (climatic zone, temperature, soil pH, clay content, etc.) as well as other applied management practices can moderate the results and should subsequently be accounted for (Valkama et al., 2015). We found that subgroup analysis or meta-regression were performed by about half of analyzed SOC meta-analyses. Results by reviews of Philibert et al. (2012), Koricheva and Gurevitch (2014) and Beillouin et al. (2019) showed that meta-analyses in agri-environment, plant ecology and conservation agriculture complied almost twice as much with this criterium.

Another issue frequently found in meta-analysis is the non-independence of effect size estimates, which occurs when effect sizes are not extracted independently, but are somehow related to each other - for example observations from different soil layers, from different treatment levels, or from sites located nearby and which share the same pedo-climatic conditions. This non-independence can lead to the underestimation of standard error of the mean effect and subsequently can impact the free evaluations of the effects' statistical significance. Therefore, meta-analysts should be aware of the sources of non-independence and should select only one effect size among several related effect sizes (Gurevitch and Hedges, 1999; Nakagawa et al., 2017).

Lastly, the degree of sensitivity of meta-analytical results should be assessed. When results are sensitive to e.g., publication bias, it is indicated that these factors need specific attention (Koricheva and Gurevitch, 2014). Funnel plots can support the interpretation of statistics by visualizing bias and highlighting outliers (Borenstein et al., 2009), which should be excluded to conduct the analyses without them and see if the overall results are affected (Rothstein et al., 2013). Another possibility is the testing via the Fail-safe N. The computation of this number allows us to detect how many additional studies it would take to reduce the overall effect to a non-significant one (Rosenthals's method) or an arbitrary minimal level (Orwin's method) (Borenstein et al., 2009). Philibert et al. (2012) reported that less than 10% of meta-analyses conducted sensitivity analysis. About 30% of SOC meta-analyses fulfilled this criterium. Beillouin et al. (2019) and Krupnik et al. (2019) found that about 40% conducted sensitivity analysis, whereas Koricheva and Gurevitch (2014) found a higher agreement of their meta-analyses or reviews with this criterium.

4.3. Results and database presentation

In the group "Result and database presentation", the presentation and availability of results and full database, which give all necessary information to reproduce an analysis, were compared. Extracted data should be provided to an extent sufficient to inform readers about all subsequent synthesis work (Woodcock et al., 2014). Full datasets promote the use of the data by others and enable updating and detection of errors (Koricheva and Gurevitch, 2014). If data is not provided sufficiently enough to update studies, information must be gathered once again, causing redundant work. Of all four reviews, our findings complied least with this criterium (Table 5). Only 16% SOC meta-analyses reported databases, including all relevant information to allow recalculation of effect sizes. Overall, results were poor. Philibert et al. (2012) received similar results, Koricheva and Gurevitch (2014) and Beillouin et al. (2019) found higher correspondences, and Krupnik et al. (2019) identified the highest agreement (over 70%) with the criterium. This might be explained by the small sample size or less demanding criteria, as in our analysis of criterium "Literature Search Reported".

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Table 5. Comparison of quality assessment of meta-analyses in disciplines of soil science, agronomy and plant ecology.

Topic	Author	Number of meta-analyses or quantitative reviews under assessment	Literature search reported	Effect size calculated according to standard metrics	Weight	Subgroup analysis and meta-regression	Sensitivity analysis	Full database
Soil organic carbon	This study	31	41%	74%	13% by 1/variance, 19% partly	45% (by Q-test)	26% outlier or effect size distribution; 29% detection of publication bias	16%
Agri-environment and - biodiversity	Philibert et al. (2012)	73	22%	not reported	37% weighted by 1/variance	95%	8%	18%
Conservation and organic agriculture	Krupnik et al. (2019)	17	100%	not reported	82% weighted by replication, plot or yield sampling area	not reported	41% sensitivity analysis, 47% publication bias	71%
Crop diversification	Beillouin et al. (2019)	99	46% search string, 86 % eligibility criteria	Not reported	40% weighted 1/variance or sample size	98% explored causes of heterogeneity	40%	32%
Plant ecology	Koricheva and Gurevitch (2014)	322	32%	85%	74% weighted by 1/variance	89% explored causes of heterogeneity	61% publication bias; 25% applied some form of sensitivity analysis	31%

Data in percentages was directly extracted from the publications or calculated from total numbers.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

4.4. Management categories

The results (Table 4) show that the management category "Tillage" was studied by 15 meta-analyses, with the highest score of 27 by the review of Haddaway et al. (2017), who provided a in depth and high-quality synthesis of no-till/reduced tillage versus conventional tillage effects on SOC at a global level using raw mean difference as effect size. Therefore, we suggest that the topic is well covered for the moment and no further global meta-analysis is needed until there is a substantial number of new results. Nevertheless, high quality meta-analyses and systematic reviews studying tillage effects on SOC in specific pedoclimatic zones or continents, such as Europe, are still missing. The maximum score (16) in the organic management category was reached by the publication of García-Palacios et al. (2018), which lacked in-depth reporting of the search strategy and independency of effect sizes, used studies where pedotransfer functions were applied, did not check for outliers, only extracted SD partly, and thus weighted partly by 1/variance. Regarding the effect of cover crops on SOC, Jian et al. (2020) produced the meta-analysis which reached the highest score (19) in this category. Reporting of literature searches and effect size calculations was conducted well, but the study failed to calculate moderator effects, conduct sensitivity analysis, did not extract effect sizes independently, and included studies with pedotransfer function application. In the category "Residue", the maximum score of 19 was reached by the meta-analysis of Li et al. (2020). Literature search reporting, effect size calculation and moderator analysis was done well, but effect sizes were not extracted independently, outliers were not assessed, and a full database was not provided. Maximum scores in all other management categories did not achieve scores above 18. We therefore conclude that there is a need for further and improved meta-analyses on all management categories, except no-till/reduced tillage versus conventional tillage.

4.5. Impact of meta-analysis quality on policy making

In our quick quality assessment of meta-analyses cited in chapter 2 of the IPCC "Special Report - Climate Change and Land" (Jia et al., 2019), we found that 50% of studies (eight out of 16) which used the term "meta-analysis" in their title, were in fact no true meta-analyses, as they did not fulfil the cut-off criteria. As not even the key criteria for conducting a meta-analysis were followed by these articles, the quality of the overall study and therefore the reliability of their results is unsure. In a study by O'Leary et al. (2016), 92 reviews were assessed on their value for decision-making with the help of the Collaboration for Environmental Evidence Synthesis Assessment Tool (CEESAT) (Woodcock et al., 2014; Collaboration for Environmental Evidence, 2020), which contains elements for analysing transparency, objectivity and comprehensiveness. They found that the evidence reviews did perform poorly, with a median score of 2.5 (of possible 39). Further, many of these reviews showed low reliability in methodology, which enhances the risk that the current knowledge is not adequately reflected. They concluded, that "such reviews thus have the potential to misinform decision-making, especially if selectively used by stakeholders with particular priorities" (O'Leary et al., 2016, p.80).

Scientific literature is used increasingly for environmental management decision making (Dicks et al., 2014). Especially documents that synthesize the results of multiple studies and peer-reviewed publications are primary sources of information for respondents (Seavy and Howell, 2010). Although science is by far not the only factor which is influencing policy decisions, there have been cases in which scientific findings have had crucial impacts on policy changes (Pullin and Knight, 2012). Therefore, researchers are obligated to ensure that their evidence reviews (such as meta-analyses) accurately reflect the primary evidence base and are reliable and transparent (O'Leary et al., 2016).

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

4.6. How to fix the problem

The described limitations call for advances in meta-analyses conducted in soil and agricultural research. Firstly, to improve the overall quality, it is crucial to support education at university level and implement training for interested scientists and stakeholders. Gurevitch et al. (2018) stressed that such trainings should be part of the curriculum for higher-degree students. Furthermore, they point out that not only scientists but also editors, reviewers and science-policy practitioners would greatly benefit from knowledge on meta-analytical methodology, as it would enable them to assess the quality of meta-analyses and interpret results.

Secondly, readers of meta-analyses should check for the presence of key elements assuring transparency and replicability of the article (Lortie et al., 2015). Krupnik et al. (2019) argue that scientists and policy makers need to evaluate meta-analyses critically regarding treatment definition, data collection and analysis. Results of meta-analyses on highly politicized agronomic topics should be interpreted especially carefully. We fully agree with these claims and support the appeal to be critical when it comes to meta-analytical outcomes. The proposed quality criteria-set should aid this demanding process.

An issue that meta-analysts frequently face, is that many primary publications do not report standard deviations, which are needed to calculate variance and subsequently weight studies by the inverse of it. As a result, many studies cannot be included in the meta-analysis, thereby reducing the amount of valuable information needed to gain rigorous results. To solve this issue, a new tool named "EX-TRACT" was recently developed (Acutis et al., 2022). The easy-to-use Excel© worksheet application allows to obtain pooled error standard deviations (s_w) from ANOVA and in Multiple Comparison Tests (MCT) outcomes. By using this tool, we can double the number of studies which can be included in a meta-analysis (Acutis et al., 2022) and avoid discarding primary literature which fits our scope.

Another available and highly useful tool allows the computation of SOC stock and its SD for a single soil layer based on SOC concentration and BD (also from multiple sub-layers) (Tadiello et al., 2022). The Excel© workbook automatically computes the means of stocks and SD, saving the results in a ready-to-use database. This is especially helpful when conducting a meta-analysis. Since in original articles, SOC observations are often presented for multiple sub-layers, but not for the complete soil profile, meta-analysts tend to extract all available observations per a study, leading to a non-independence of effect sizes. With the help of this tool, it is possible to "fuse" the results from all layers into one, *independent* effect size.

The publication of protocols prior to a meta-analysis would benefit the method by allowing constructive criticism and suggestions for improvement by the scientific community (Moher et al., 2015; Brandt et al., 2013). Gurevitch et al. (2018) described that the pre-registration of planned meta-analyses, which are then peer-reviewed and published before the actual analysis is conducted, can aid the reduction of selective reporting and publication bias. Systematic review protocols for environmental sciences from the journal "Environmental Evidence" or the initiative "ROSES" are available and can be used for the construction of meta-analytical protocols.

Lastly, another viable asset in improving the quality of future meta-analyses in soil science would be the creation of a European meta-analysis hub, which focuses on 1) the development of high-quality products, 2) the assessment of quality and 3) the creation of a European database. The database should comprise all available information of former meta-analyses on soil and agricultural research, providing researchers with valuable data. With the help of this database, new meta-analyses, studying management practices relevant for the pedoclimatic zones present in Europe, could be conducted.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

This is important, as the inclusion of global experiments into an analysis can lead to over-diversification and therefore to the combination of "apples and oranges", which is not expedient.

5. Conclusions

Quality assessment of meta-analyses, especially in the complex agricultural set up, are highly warranted to harness the power of meta-analyses. We demonstrate that meta-analyses in soil and agricultural research encounter specific issues, which differ to other fields like medicine, environment or ecology. Therefore, we adapted meta-analytical guidelines from other disciplines to construct an easy-to-use criteria-set, which is suited to quantitatively assess the quality of meta-analyses in agriculture and soil sciences. With the help of these criteria, we further investigated the quality of 31 meta-analyses, studying the effects of agricultural management practices on SOC. By doing so, we aimed to present the application of the criteria-set and analyze the quality of quantitative reviews within this prominent topic. Our analysis showed that the overall quality of analyses improved over time, but only one achieved a high score. Deficits were found in literature search, statistical analyses, and data presentation. The correct weighting by 1/variance of effect sizes was found to be a challenge for many authors. In some cases, the term "meta-analysis" is still falsely used to describe quantitative syntheses of any style, independent of methodology applied. The analysis also revealed that out of 11 identified management categories studied by the meta-analyses, only the effects of no-till/reduced tillage versus conventional tillage on SOC are studied sufficiently in form of a high quality meta-analytical synthesis.

Our results indicate that the quality of meta-analyses in agricultural and soil sciences is, despite all efforts, still not satisfactory. As the information presented in summarizing research articles is frequently used by decision makers, this can also have negative impacts on evidence-based policymaking. It is high time that the agricultural and soil scientific community adapts rigorous meta-analytical methodologies and improves the quality of its output. We believe that the method is a viable and indispensable tool in quantitative synthesis of agricultural and soil research and only with combined efforts and collaborations between stakeholders across disciplines we will be able to overcome the presented challenges.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

6. Appendix

Table A1. Assessed SOC meta-analyses and their identification numbers.

Identification number (ID)	Reference of meta-analysis
1	Aguilera, E., Lassaletta, L., Gattinger, A., Gimeno, B.S., 2013. Managing soil carbon for climate change mitigation and adaptation in Mediterranean cropping systems: A meta-analysis. Agric. Ecosyst. Environ. 168, 25–36. https://doi.org/10.1016/j.agee.2013.02.003
2	Angers, D.A., Eriksen-Hamel, N.S., 2008. Full-Inversion Tillage and Organic Carbon Distribution in Soil Profiles: A Meta-Analysis. Soil Sci. Soc. Am. J. 72, 1370–1374. https://doi.org/10.2136/sssaj2007.0342
3	Bai, X., Huang, Y., Ren, W., Coyne, M., Jacinthe, P.-A., Tao, B., Hui, D., Yang, J., Matocha, C., 2019. Responses of soil carbon sequestration to climate-smart agriculture practices: A meta-analysis. Glob. Chang. Biol. 25, 2591–2606. https://doi.org/10.1111/gcb.14658
4	Chen, Y., Camps-Arbestain, M., Shen, Q., Singh, B., Cayuela, M.L., 2018. The long-term role of organic amendments in building soil nutrient fertility: a meta-analysis and review. Nutr. Cycl. AGROECOSYSTEMS 111, 103–125. https://doi.org/10.1007/s10705-017-9903-5
5	Cooper, J., Baranski, M., Stewart, G., Nobel-de Lange, M., Bàrberi, P., Fließbach, A., Peigné, J., Berner, A., Brock, C., Casagrande, M., Crowley, O., David, C., De Vliegher, A., Döring, T.F., Dupont, A., Entz, M., Grosse, M., Haase, T., Halde, C., Hammerl, V., Huiting, H., Leithold, G., Messmer, M., Schloter, M., Sukkel, W., van der Heijden, M.G.A., Willekens, K., Wittwer, R., Mäder, P., 2016. Shallow non-inversion tillage in organic farming maintains crop yields and increases soil C stocks: a meta-analysis. Agron. Sustain. Dev. 36. https://doi.org/10.1007/s13593-016-0354-1
6	Feng, Q., An, C., Chen, Z., Wang, Z., 2020. Can deep tillage enhance carbon sequestration in soils? A meta-analysis towards GHG mitigation and sustainable agricultural management. Renew. Sustain. Energy Rev. 133, 110293. https://doi.org/https://doi.org/10.1016/j.rser.2020.110293
7	García-Palacios, P., Gattinger, A., Bracht-Jørgensen, H., Brussaard, L., Carvalho, F., Castro, H., Clément, J.-C., De Deyn, G., D'Hertefeldt, T., Foulquier, A., Hedlund, K., Lavorel, S., Legay, N., Lori, M., Mäder, P., Martínez-García, L.B., da Silva, P., Muller, A., Nascimento, E., Reis, F., Symanczik, S., Paulo Sousa, J., Milla, R., 2018. Crop traits drive soil carbon sequestration under organic farming. J. Appl. Ecol. 55, 2496–2505. https://doi.org/10.1111/1365-2664.13113
8	Gattinger, A., Muller, A., Haeni, M., Skinner, C., Fliessbach, A., Buchmann, N., Mäder, P., Stolze, M., Smith, P., Scialabba, N.E.H., Niggli, U., 2012. Enhanced top soil carbon stocks under organic farming. Proc. Natl. Acad. Sci. U. S. A. 109, 18226–18231. https://doi.org/10.1073/pnas.1209429109

9 González-Sánchez, E.J., Ordóñez-Fernández, R., Carbonell-Bojollo, R., Veroz-González, O., Gil-Ribes, J.A., 2012. Meta-analysis on atmospheric carbon capture in Spain through the use of conservation agriculture. Soil Tillage Res. 122, 52–60. https://doi.org/10.1016/j.still.2012.03.001

10 Haddaway, N.R., Hedlund, K., Jackson, L.E., Katterer, T., Lugato, E., Thomsen, I.K., Jorgensen, H.B., Isberg, P.-E., 2017. How does tillage intensity affect soil organic carbon? A systematic review. Environ. Evid. 6. https://doi.org/10.1186/s13750-017-0108-9

11 Han, P., Zhang, W., Wang, G., Sun, W., Huang, Y., 2016. Changes in soil organic carbon in croplands subjected to fertilizer management: a global meta-analysis. Sci. Rep. 6. https://doi.org/10.1038/srep27199

12 Jian, J., Du, X., Reiter, M.S., Stewart, R.D., 2020. A meta-analysis of global cropland soil carbon changes due to cover cropping. Soil Biol. Biochem. 143, 107735. https://doi.org/10.1016/j.soilbio.2020.107735

13 King, A.E., Blesh, J., 2018. Crop rotations for increased soil carbon: Perenniality as a guiding principle: Perenniality. Ecol. Appl. 28, 249–261. https://doi.org/10.1002/eap.1648

14 Kopittke, P.M., Dalal, R.C., Finn, D., Menzies, N.W., 2017. Global changes in soil stocks of carbon, nitrogen, phosphorus, and sulphur as influenced by long-term agricultural production. Glob. Chang. Biol. 23, 2509–2519. https://doi.org/10.1111/gcb.13513

15 Ladha, J.K., Reddy, C.K., Padre, A.T., van Kessel, C., 2011. Role of Nitrogen Fertilization in Sustaining Organic Matter in Cultivated Soils. J. Environ. Qual. 40, 1756–1766. https://doi.org/10.2134/jeq2011.0064

16 Li, Y., Li, Z., Chang, S.X., Cui, S., Jagadamma, S., Zhang, Q., Cai, Y., 2020. Residue retention promotes soil carbon accumulation in minimum tillage systems: Implications for conservation agriculture. Sci. Total Environ. 740, 140147. https://doi.org/10.1016/j.scitotenv.2020.140147

17 Liu, S., Zhang, Y., Zong, Y., Hu, Z., Wu, S., Zhou, J., Jin, Y., Zou, J., 2016. Response of soil carbon dioxide fluxes, soil organic carbon and microbial biomass carbon to biochar amendment: a meta-analysis. Glob. Chang. Biol. BIOENERGY 8, 392–406. https://doi.org/10.1111/gcbb.12265

18 Luo, Z., Wang, E., Sun, O.J., 2010. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. Agric. Ecosyst. Environ. 139, 224–231. https://doi.org/10.1016/j.agee.2010.08.006

19 Maillard, É., Angers, D.A., 2014. Animal manure application and soil organic carbon stocks: A meta-analysis. Glob. Chang. Biol. 20, 666–679. https://doi.org/10.1111/gcb.12438

20 Majumder, S., Neogi, S., Dutta, T., Powel, M.A., Banik, P., 2019. The impact of biochar on soil carbon sequestration: Meta-analytical approach to evaluating environmental and economic advantages. J. Environ. Manage. 250. https://doi.org/10.1016/j.jenvman.2019.109466

21 Mathew, I., Shimelis, H., Mutema, M., Minasny, B., Chaplot, V., 2020. Crops for increasing soil organic carbon stocks - A global meta analysis. Geoderma 367. https://doi.org/10.1016/j.geoderma.2020.114230

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

22 McDaniel, M.D., Tiemann, L.K., Grandy, A.S., 2014. Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? A meta-analysis. Ecol. Appl. 24, 560–570. https://doi.org/10.1890/13-0616.1

23 Meurer, K.H.E., Haddaway, N.R., Bolinder, M.A., Kätterer, T., 2018. Tillage intensity affects total SOC stocks in boreo-temperate regions only in the topsoil—A systematic review using an ESM approach. Earth-Science Rev. 177, 613–622. https://doi.org/10.1016/j.earscirev.2017.12.015

24 Mondal, S., Chakraborty, D., Bandyopadhyay, K., Aggarwal, P., Rana, D.S., 2020. A global analysis of the impact of zero-tillage on soil physical condition, organic carbon content, and plant root response. L. Degrad. Dev. 31, 557–567. https://doi.org/10.1002/ldr.3470

25 Ogle, S.M., Breidt, F.J., Paustian, K., 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry. https://doi.org/10.1007/s10533-004-0360-2

26 Poeplau, C., Don, A., 2015. Carbon sequestration in agricultural soils via cultivation of cover crops - A meta-analysis. Agric. Ecosyst. Environ. 200, 33–41. https://doi.org/10.1016/j.agee.2014.10.024

27 Sun, W., Canadell, J.G., Yu, Lijun, Yu, Lingfei, Zhang, W., Smith, P., Fischer, T., Huang, Y., 2020. Climate drives global soil carbon sequestration and crop yield changes under conservation agriculture. Glob. Chang. Biol. 26, 3325–3335. https://doi.org/10.1111/gcb.15001

28 Tuomisto, H.L., Hodge, I.D., Riordan, P., Macdonald, D.W., 2012. Does organic farming reduce environmental impacts? - A meta-analysis of European research. J. Environ. Manage. 112, 309–320. https://doi.org/10.1016/j.jenvman.2012.08.018

29 Virto, I., Barre, P., Burlot, A., Chenu, C., 2012. Carbon input differences as the main factor explaining the variability in soil organic C storage in no-tilled compared to inversion tilled agrosystems. Biogeochemistry 108, 17–26. https://doi.org/10.1007/s10533-011-9600-4

30 Xia, L., Lam, S.K., Wolf, B., Kiese, R., Chen, D., Butterbach-Bahl, K., 2018. Trade-offs between soil carbon sequestration and reactive nitrogen losses under straw return in global agroecosystems. Glob. Chang. Biol. 24, 5919–5932. https://doi.org/10.1111/gcb.14466

31 Xu, H., Sieverding, H., Kwon, H., Clay, D., Stewart, C., Johnson, J.M.F., Qin, Z., Karlen, D.L., Wang, M., 2019. A global meta-analysis of soil organic carbon response to corn stover removal. Glob. Chang. Biol. BIOENERGY 11, 1215–1233. https://doi.org/10.1111/gcbb.12631

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Supplementary material

The supplementary material can be found through this link: 10.6084/m9.figshare.19447121

Table S1. List of resources for assessing the quality of, reporting or creating meta-analyses in several disciplines
Table S2. 31 soil organic carbon meta-analyses and their individual scores per (sub)-criterium
Table S3. Literature gathering
Table S4. Information on screening process
Table S5. Results for all meta-analyses per group and (sub)-criterium
Table S6. Quick quality assessment of meta-analyses in IPCC report with the help of "cut-off" criteria
Table S7. Results for criteria 9-17 from meta-analyses that passed the three cut-off criteria
Table S8. Template for quality assessment

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 862695 EJP SOIL.

Author contribution

JF and EV compiled the quality criteria-set and performed the quality assessment. SZB initiated the study and raised the funding. All authors contributed to the conceptual planning of the study, to the writing and to interpretation of results. All authors provided input on earlier versions of the manuscript and read and approved of the final review.

Declaration of Competing Interest

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

Acknowledgements

We wish to thank Prof. Marco Acutis for his thoughtful comments and Dylan Goff, MSc. for proofreading this manuscript.

7. References

Acutis, M., Tadiello, T., Perego, A., di Guardo, A., Schillaci, C., and Valkama, E.: EX-TRACT: An excel tool for the estimation of standard deviations from published articles, Environ. Model. Softw., 147, 105236, https://doi.org/10.1016/j.envsoft.2021.105236, 2022.

Aksnes, D. W., Langfeldt, L., and Wouters, P.: Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories, SAGE Open, 9, https://doi.org/10.1177/2158244019829575, 2019.

Bai, X., Huang, Y., Ren, W., Coyne, M., Jacinthe, P.-A., Tao, B., Hui, D., Yang, J., and Matocha, C.: Responses of soil carbon sequestration to climate-smart agriculture practices: A meta-analysis, Glob. Chang. Biol., 25, 2591–2606, https://doi.org/10.1111/gcb.14658, 2019.

Beillouin, D., Ben-Ari, T., and Makowski, D.: Evidence map of crop diversification strategies at the global scale, Environ. Res. Lett., https://doi.org/10.1088/1748-9326/ab4449, 10 January 2019.

Bolinder, M. A., Crotty, F., Elsen, A., Frac, M., Kismanyoky, T., Lipiec, J., Tits, M., Toth, Z., and Katterer, T.: The effect of crop residues, cover crops, manures and nitrogen fertilization on soil organic carbon changes in agroecosystems: a synthesis of reviews, Mitig. Adapt. Strateg. Glob. Chang., 25, 929–952, https://doi.org/10.1007/s11027-020-09916-3, 2020.

Borenstein, M., Hedges, L. V., Higgins, J., and Rothstein, H.: Introduction to meta-analysis, Wiley, 421 pp., 2009.

Brandt, K., Średnicka-Tober, D., Barański, M., Sanderson, R., Leifert, C., and Seal, C.: Methods for Comparing Data across Differently Designed Agronomic Studies: Examples of Different Meta-analysis Methods Used to Compare Relative Composition of Plant Foods Grown Using Organic or Conventional Production Methods and a Protocol for a Systematic Review, J. Agric. Food Chem., 61, 7173−7180, https://doi.org/10.1021/jf4008967, 2013.

Collaboration for Environmental Evidence: Guidelines and Standards for Evidence Synthesis in Environmental Management. Version 5., edited by: AS Pullin, GK Frampton, B Livoreil, and G Petrokofsky, 2018.

Collaboration for Environmental Evidence: The Collaboration for Environmental Evidence Synthesis Appraisal Tool (CEESAT). Version 2.1., https://environmentalevidence.org/ceeder/about-ceesat/, 2020.

Cooper, H., Hedges, L. V., and Valentine, J. C.: Potentials and Limitations of Research Synthesis, in: The Handbook of Research Synthesis and Meta-Analysis, edited by: Cooper, H., Hedges, L. V., and Valentine, J. C., Russell Sage Foundation, 518–525, 2019a.

Cooper, H., Hedges, L. V., and Valentine, J. C.: Research Synthesis as a Scientific Process, in: The Handbook of Research Synthesis and Meta-Analysis, edited by: Cooper, H., Hedges, L. V., and Valentine, J. C., Russell Sage Foundation, New York, 4–15, 2019b.

Cooper, H., Hedges, L. V., and Valentine, J. C.: The handbook of research synthesis and meta-analysis, 3rd ed., Russell Sage Foundation, New York, 556 pp., 2019c.

Côté, I. M., Curtis, P. S., Rothstein, H. R., and Stewart, G. B.: Gathering Data: searching Literature and selection Criteria, Princeton University Press, 37–51, 2013.

Culina, A., Crowther, T. W., Ramakers, J. J. C., Gienapp, P., and Visser, M. E.: How to do meta-analysis of open datasets, Nat. Ecol. Evol., Vol. 2, Issue 7, Nature Publishing Group, pp. 1053–1056, https://doi.org/10.1038/s41559-018-0579-2, 1 July 2018.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Curtis, P. S., Mengersen, K., Lajeunesse, M. J., Rothstein, H. R., and Stewart, G. B.: Extraction and Critical appraisal of Data, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 52–60, 2013.

Dicks, L. v., Walsh, J. C., and Sutherland, W. J.: Organising evidence for environmental management decisions: A “4S” hierarchy, Trends Ecol. Evol., Vol. 29, Issue 11, Elsevier Ltd., pp. 607–613, https://doi.org/10.1016/j.tree.2014.09.004, 1 November 2014.

García-Palacios, P., Gattinger, A., Bracht-Jørgensen, H., Brussaard, L., Carvalho, F., Castro, H., Clément, J.-C., De Deyn, G., D’Hertefeldt, T., Foulquier, A., Hedlund, K., Lavorel, S., Legay, N., Lori, M., Mäder, P., Martínez-García, L. B., da Silva, P., Muller, A., Nascimento, E., Reis, F., Symanczik, S., Paulo Sousa, J., and Milla, R.: Crop traits drive soil carbon sequestration under organic farming, J. Appl. Ecol., 55, 2496–2505, https://doi.org/10.1111/1365-2664.13113, 2018.

Glass, G. V: Primary, Secondary, and Meta-Analysis of Research, Educ. Res., 5, 3–8, https://doi.org/10.3102/0013189X005010003, 1976.

Gonçalves, R. S. and Musen, M. A.: Analysis: The variable quality of metadata about biological samples used in biomedical experiments, Sci. Data, 6, 1–15, https://doi.org/10.1038/sdata.2019.21, 2019.

Gurevitch, J. and Hedges, L. V.: Statistical Issues in Ecological Meta-Analyses, Ecology, 80, 1142, https://doi.org/10.2307/177061, 1999.

Gurevitch, J., Koricheva, J., Nakagawa, S., and Stewart, G.: Meta-analysis and the science of research synthesis, Nature, 555, 175–182, https://doi.org/10.1038/nature25753, 2018.

Haddaway, N. R., Hedlund, K., Jackson, L. E., Kätterer, T., Lugato, E., Thomsen, I. K., Jørgensen, H. B., and Isberg, P. E.: How does tillage intensity affect soil organic carbon? A systematic review, BioMed Central, 1–48 pp., https://doi.org/10.1186/s13750-017-0108-9, 2017.

Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A. S., Savilaakso, S., and Stewart, G. B.: Eight problems with literature reviews and how to fix them, Nat. Ecol. Evol., 4, 1582–1589, https://doi.org/10.1038/s41559-020-01295-x, 2020.

Hedges, L. V., Gurevitch, J., and Curtis, P. S.: The Meta-Analysis of Response Ratios in Experimental Ecology, Ecology, 80, 1150, https://doi.org/10.2307/177062, 1999.

Hungate, B. A., van Groenigen, K. J., Six, J., Jastrow, J. D., Luo, Y., de Graaff, M. A., van Kessel, C., and Osenberg, C. W.: Assessing the effect of elevated carbon dioxide on soil carbon: A comparison of four meta-analyses, Glob. Change Biol., 15, 2020–2034, https://doi.org/10.1111/j.1365-2486.2009.01866.x, 2009.

IPCC: Summary for Policymakers, in: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, edited by: Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H. O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., and Malley, J., 2019.

Jia, G., Shevliakova, E., Artaxo, P., de Noblet-Ducoudré, N., Houghton, R., House, J., Kitajima, K., Lennard, C., Popp, A., Sirin, A., Sukumar, R., and Verch, L.: Land–climate interactions, in: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, edited by: Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner,

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

H.-O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, K., Kissick, K., Belkacemi, M., and Malley, J., 2019.

Jian, J., Du, X., Reiter, M. S., and Stewart, R. D.: A meta-analysis of global cropland soil carbon changes due to cover cropping, Soil Biol. Biochem., 143, 107735, https://doi.org/10.1016/j.soilbio.2020.107735, 2020.

Koricheva, J. and Gurevitch, J.: Place of Meta-analysis among other Methods of research synthesis, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princton University Press, Princeton, 3–13, 2013.

Koricheva, J. and Gurevitch, J.: Uses and misuses of meta-analysis in plant ecology, J. Ecol., 102, 828–844, https://doi.org/10.1111/1365-2745.12224, 2014.

Koricheva, J., Gurevitch, J., and Mengersen, K. (Eds.): Handbook of Meta-analysis in Ecology and Evolution, Princton University Press, 520 pp., 2013.

Krupnik, T. J., Andersson, J. A., Rusinamhodzi, L., Corbeels, M., Shennan, C., and Gérard, B.: Does size matter? a critical review of meta-analysis in agronomy, Exp. Agric., 55, 200–229, https://doi.org/10.1017/S0014479719000012, 2019.

Leydesdorff, L., Bornmann, L., Comins, J. A., and Milojević, S.: Citations: Indicators of Quality? The Impact Fallacy, Front. Energy Res., 1, https://doi.org/10.3389/frma.2016.00001, 2016.

Lipsey, M. W.: Identifying Potentially Interesting Variables And Analysis Opportunities, in: The Handbook of Research Synthesis and Meta-Analysis, edited by: Cooper, H., Hedges, L. V., and Valentine, J. C., Russell Sage Foundation, New York, 142–151, 2019.

Li, Y., Li, Z., Chang, S. X., Cui, S., Jagadamma, S., Zhang, Q., and Cai, Y.: Residue retention promotes soil carbon accumulation in minimum tillage systems: Implications for conservation agriculture, Sci. Total Environ., 740, 140147, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.140147, 2020.

Lortie, C. J., Lau, J., and Lajeunesse, M. J.: Graphical Presentation of results, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 339–347, 2013.

Lortie, C. J., Stewart, G., Rothstein, H., and Lau, J.: How to critically read ecological meta-analyses, Res. Synth. Methods, 6, 124–133, https://doi.org/10.1002/jrsm.1109, 2015.

Luo, Z., Wang, E., and Sun, O. J.: Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments, Agric. Ecosyst. Environ., 139, 224–231, https://doi.org/https://doi.org/10.1016/j.agee.2010.08.006, 2010.

Mayo-Wilson, E. and Grant, S.: Transparent Reporting: Registrations, Protocols, and Final Reports, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Cooper, H., Hedges, L. V., and Valentine, J. C., Russell Sage Foundation, New York, 471–488, 2019.

Mengersen, K., Schmid, C. H., Jennions, M. D., and Gurevitch, J.: Statistical Models and approaches to Inference, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 89–107, 2013.

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z. S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S. Y., Mandal, B., Marchant, B. P., Martin, M., McConkey, B. G., Mulder, V. L., O'Rourke, S., Richer-de-Forges, A. C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C. C., Vågen, T. G., van Wesemael, B., and Winowiecki, L.: Soil carbon 4 per mille, Geoderma, 292, 59–86, https://doi.org/10.1016/j.geoderma.2017.01.002, 2017.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., and Group, P.-P.: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, Syst. Rev., 4, 1–9, https://doi.org/10.1186/2046-4053-4-1, 2015.

Nakagawa, S. and Cuthill, I. C.: Effect size, confidence interval and statistical significance: A practical guide for biologists, Biol. Rev., 82, 591–605, https://doi.org/10.1111/j.1469-185X.2007.00027.x, 2007.

Nakagawa, S., Noble, D. W. A., Senior, A. M., and Lagisz, M.: Meta-evaluation of meta-analysis: Ten appraisal questions for biologists, BMC Biology, 15, 1–14, https://doi.org/10.1186/s12915-017-0357-7, 2017.

O'Dea, R. E., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W. A., Parker, T. H., Gurevitch, J., Page, M. J., Stewart, G., Moher, D., and Nakagawa, S.: Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension, Biol. Rev., 96, 1695–1722, https://doi.org/10.1111/brv.12721, 2021.

Ogle, S. M., Breidt, F. J., and Paustian, K.: Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions, Biogeochemistry, https://doi.org/10.1007/s10533-004-0360-2, 2005.

O'Leary, B. C., Kvist, K., Bayliss, H. R., Derroire, G., Healey, J. R., Hughes, K., Kleinschroth, F., Sciberras, M., Woodcock, P., and Pullin, A. S.: The reliability of evidence review methodology in environmental science and conservation, Environ. Policy Gov., 64, 75–82, https://doi.org/10.1016/j.envsci.2016.06.012, 2016.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., and Moher, D.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ, 372, https://doi.org/10.1136/bmj.n71, 2021.

Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., Kelly, C. D., Gurevitch, J., and Nakagawa, S.: Transparency in Ecology and Evolution: Real Problems, Real Solutions, Trends Ecol. Evol., 31, 711–719, https://doi.org/10.1016/j.tree.2016.07.002, 2016.

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., and Smith, P.: Climate-smart soils, Nature, 532, 49-57, https://doi.org/10.1038/nature17174, 6 April 2016.

Philibert, A., Loyce, C., and Makowski, D.: Assessment of the quality of meta-analysis in agronomy, Agric. Ecosyst. Environ., 148, 72–82, https://doi.org/10.1016/j.agee.2011.12.003, 2012.

Pullin, A. S. and Knight, T. M.: Science informing Policy - A health warning for the environment, Environ. Evid., 1, https://doi.org/10.1186/2047-2382-1-15, 19 December 2012.

Rosenberg, M. S., Gurevitch, J., and Adams, D.: MetaWin Statistical Software for Meta-Analysis, Sinnauer Associates, 2000.

Rosenberg, M. S., Rothstein, H. R., and Gurevitch, J.: Effect sizes: Conventional Choices and Calculations, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 61–71, 2013.

Rothstein, H., Lortie, C. J., Stewart, G. B., Koricheva, J., and Gurevitch, J.: Quality standards for research syntheses, Princton University Press, 323-338, 2013.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Schillaci, C., Perego, A., Valkama, E., Märker, M., Saia, S., Veronesi, F., Lipani, A., Lombardo, L., Tadiello, T., Gamper, H. A., Tedone, L., Moss, C., Pareja-Serrano, E., Amato, G., Kühl, K., Dămătîrcă, C., Cogato, A., Mzid, N., Eeswaran, R., Rabelo, M., Sperandio, G., Bosino, A., Bufalini, M., Tunçay, T., Ding, J., Fiorentini, M., Tiscornia, G., Conradt, S., Botta, M., and Acutis, M.: New pedotransfer approaches to predict soil bulk density using WoSIS soil data and environmental covariates in Mediterranean agro-ecosystems, Sci. Total Environ., 780, https://doi.org/10.1016/j.scitotenv.2021.146609, 2021.

Schmid, C. H., Stewart, G. B., Rothstein, H. R., Lajeunesse, M. J., and Gurevitch, J.: Software for statistical Meta-analysis, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 147–194, 2013.

Seavy, N. E. and Howell, C. A.: How can we improve information delivery to support conservation and restoration decisions?, Biodivers. Conserv., 19, 1261–1267, https://doi.org/10.1007/s10531-009-9752-x, 2010.

Smith, P.: Soils and climate change, Curr. Opin. Environ. Sustain., 4, 539-544, https://doi.org/10.1016/j.cosust.2012.06.005, 2012.

Smith, P., Andrén, O., Karlsson, T., Perälä, P., Regina, K., Rounsevell, M., and van Wesemael, B.: Carbon sequestration potential in European croplands has been overestimated, Glob. Change Biol., 11, 2153–2163, https://doi.org/10.1111/j.1365-2486.2005.01052.x, 2005.

Stewart, G.: Meta-analysis in applied ecology, Biol. Lett., 6, 78–81, https://doi.org/10.1098/rsbl.2009.0546, 2010.

Stewart, G. B., Côté, I. M., Rothstein, H. R., and Curtis, P. S.: First steps in Beginning a Meta-analysis, in: Handbook of Meta-analysis in Ecology and Evolution, edited by: Koricheva, J., Gurevitch, J., and Mengersen, K., Princeton University Press, Princeton, 27–36, 2013.

Tadiello, T., Perego, A., Valkama, E., Schillaci, C., and Acutis, M.: Computation of total soil organic carbon stock and its standard deviation from layered soils, MethodsX, 9, https://doi.org/10.1016/j.mex.2022.101662, 2022.

Tubiello, F. N., Salvatore, M., Ferrara, A. F., House, J., Federici, S., Rossi, S., Biancalani, R., Condor Golec, R. D., Jacobs, H., Flammini, A., Prosperi, P., Cardenas-Galindo, P., Schmidhuber, J., Sanz Sanchez, M. J., Srivastava, N., and Smith, P.: The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012, Glob. Change Biol., 21, 2655–2660, https://doi.org/10.1111/gcb.12865, 2015.

Valkama, E., Virkajärvi, P., Uusitalo, R., Ylivainio, K., and Turtola, E.: Meta-analysis of grass ley response to phosphorus fertilization in Finland, Grass Forage Sci., 71, 36–53, https://doi.org/10.1111/gfs.12156, 2015.

Valkama, E., Usva, K., Saarinen, M., and Uusi-Kämppä, J.: A Meta-Analysis on Nitrogen Retention by Buffer Zones, J. Environ. Qual., 48, 270–279, https://doi.org/10.2134/jeq2018.03.0120, 2019.

Vetter, D., Rucker, G., and Storch, I.: Meta-analysis: A need for well-defined usage in ecology and conservation biology, Ecosphere, 4, 1–24, https://doi.org/10.1890/ES13-00062.1, 2013.

Woodcock, P., Pullin, A. S., and Kaiser, M. J.: Evaluating and improving the reliability of evidence syntheses in conservation and environmental science: A methodology, Biol. Conserv., 176, 54–62, https://doi.org/10.1016/j.biocon.2014.04.020, 2014.

Xiao, L., Kuhn, N. J., Zhao, R., and Cao, L.: Net effects of conservation agriculture principles on sustainable land use: A synthesis, Glob. Change Biol., 27, 6321–6330, https://doi.org/10.1111/gcb.15906, 2021.

https://doi.org/10.5194/egusphere-2022-488
Preprint. Discussion started: 15 August 2022

Xu, L., He, N. P., Yu, G. R., Wen, D., Gao, Y., and He, H. L.: Differences in pedotransfer functions of bulk density lead to high uncertainty in soil organic carbon estimation at regional scales: Evidence from Chinese terrestrial ecosystems, J. Geophys. Res. Biogeosci., 120, 1567–1575, https://doi.org/10.1002/2015JG002929, 2015.