Check for updates

Building a solid foundation: advancing evidence synthesis in agri-food systems science

Pierre Ellssel[1], Georg Küstner[2],
Magdalena Kaczorowska-Dolowy[3], Eduardo Vázquez[4],
Claudia Di Bene[5], Honghong Li[6,7], Diego Brizuela-Torres[8,9],
Elansurya Elangovan Vennila[10], José Luis Vicente-Vicente[11,12]
and Daniel Itzamna Avila-Ortega[6,7]*

[1]Department of Crop Sciences, Institute of Agronomy, BOKU University, Vienna, Austria, [2]Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Würzburg, Germany, [3]Agriculture and Environment, Harper Adams University, Newport, United Kingdom, [4]Departamento de Producción Agraria, ETSIAAB, Universidad Politécnica de Madrid (UPM), Madrid, Spain, [5]CREA - Research Centre for Agriculture and Environment, Rome, Italy, [6]Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden, [7]Global Economic Dynamics and the Biosphere, The Royal Swedish Academy of Sciences, Stockholm, Sweden, [8]Department of Conservation Biology and Social Ecological Systems, Helmholtz Centre for Environmental Research GmbH – UFZ, Leipzig, Germany, [9]German Center for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher, Leipzig, Germany, [10]Department of Agri- Food Engineering and Biotechnology (DEAB), Barcelona School of Agri-food and Biosystems Engineering (EEABB), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, [11]Spanish National Research Council (CSIC), Institute of Economics, Geography and Demography, Madrid, Spain, [12]Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany

Enhancing the reliability of literature reviews and evidence synthesis is crucial for advancing the transformation of agriculture and food (agri-food) systems as well as for informed decisions and policy making. In this perspective, we argue that evidence syntheses in the field of agri-food systems research often suffer from a suite of methodological limitations that substantially increase the risk of bias, i.e., publication and selection bias, resulting in unreliable and potentially flawed conclusions and, consequently, poor decisions (e.g., policy direction, investment, research foci). We assessed 926 articles from the Collaboration for Environmental Evidence Database of Evidence Reviews (CEEDER) and recent examples from agri-food systems research to support our reasoning. The analysis of articles from CEEDER (*n* = 926) specifically indicates poor quality (Red) in measures to minimize subjectivity during critical appraisal (98% of all reviews), application of the eligibility criteria (97%), cross-checking of extracted data by more than one reviewer (97%), critical appraisal of studies (88%), establishment of an *a priori* method/protocol (86%), and transparent reporting of eligibility decisions (65%). Additionally, deficiencies (Amber) were found in most articles (>50%) regarding the investigation and discussion of variability in study findings (89%), comprehensiveness of the search (78%), definition of eligibility criteria (72%), search approach (64%), reporting of extracted data for each study (59%), consideration and discussion of the limitations of the synthesis (56%), documentation of data extraction (54%) and regarding the statistical approach (52%). To enhance the quality of evidence synthesis in agri-food science, review authors should use tried-and-tested methodologies and publish peer-reviewed *a priori* protocols. Training in evidence synthesis methods should be scaled, with universities playing a crucial role. It is the shared duty of research authors, training providers, supervisors, reviewers, and editors to ensure that rigorous and robust evidence syntheses are made available to decision-makers. We argue that all these actors should be cognizant of these common mistakes

to avoid publishing unreliable syntheses. Only by thinking as a community can we ensure that reliable evidence is provided to support appropriate decision-making in agri-food systems science.

1 Why is rigourous evidence synthesis necessary in agri-food systems?

Agri-food systems are vital for primary production, food distribution, household consumption, food and nutrition security (Willett et al., 2019). They also substantially contribute to global greenhouse gas (GHG) emissions, alteration of carbon, nitrogen and phosphorus cycles, land-use change, biodiversity loss, and intensive freshwater and groundwater use and contamination (IPBES, 2019; IPCC, 2019; Richardson et al., 2023).

Agri-food systems are projected to face growing challenges, including impacts from climate change, increasing populations, and vulnerability to shocks and stresses. The need to transform agri-food systems has been widely recognized (HLPE, 2019; IPCC, 2019; Willett et al., 2019; SAPEA, 2020; Ruggeri Laderchi et al., 2024), requiring reliable scientific evidence to inform and guide policy and enhance public awareness. This evidence must be obtained through a systematic process of evidence synthesis to ensure transparency and replicability (GCSA, 2019).

Evidence synthesis methods, such as systematic reviews, provide a more rigorous methodology, aiming for transparency, procedural objectivity, repeatability, and minimising bias throughout the review process, in comparison to traditional, non-systematic literature reviews (Pullin and Knight, 2001; Gough et al., 2017; CEE, 2022). Nevertheless, there are numerous flaws in published reviews, including those intended to be systematic. Such flaws include redundancy and poor methodological reliability, including lack of comprehensiveness, transparency, and objectivity, and hence high susceptibility to bias (e.g., selection and publication bias) (Uttley et al., 2023). These issues have been observed across research fields, such as medicine and health care (Ioannidis, 2016), education and international development (Haddaway et al., 2017), and environmental science and conservation (O'Leary et al., 2016; Haddaway et al., 2020).

In this perspective, we aim to point out methodological problems in evidence synthesis in agri-food systems research. To support our reasoning, we assessed and discussed articles from the Collaboration for Environmental Evidence Database of Evidence Reviews (CEEDER) and recent examples from agri-food systems research using a published critical appraisal tool (CEESAT) to identify and assess their risks of bias and limitations. We propose practical mitigation measures to ensure rigorous and robust evidence syntheses. Additionally, we highlight a diverse range of Open Access resources to support authors in writing evidence syntheses.

2 State of the evidence base in agri-food systems research

The CEE Database of Evidence Reviews (CEEDER) is the only database that comprehensively collects and assesses evidence reviews, specifically in the context of environmental policy and practice (CEE, 2024a), including evidence reviews from agri-food systems science. Hence, it is a valuable resource for assessing the current state of evidence in the field. Indexed reviews are evaluated against a set of methodological standards using the CEE Synthesis Assessment Tool (CEESAT) (Woodcock et al., 2014; CEE, 2024b). The database allows users to find and access reviews and evidence overviews along with robust assessments of their rigor. Additionally, it can aid in identifying 'synthesis gaps' (i.e., unaddressed review questions) to avoid redundancy in research efforts (Konno et al., 2020b). CEESAT encompasses seven review components, comprising 16 elements in total (Table 1; Figure 1). These elements are assessed and rated by at least two reviewers, earning one of four designations: Red (serious deficiencies), Amber (deficiencies), Green (acceptable standard) and Gold (meets the standards) (Konno et al., 2020b). The reviewers are selected from a pool of volunteers who have received training in applying CEESAT. The reviewers rate the evidence reviews using the appraisal tool (CEESAT), and the criteria defined herein serve as an indicator for assessing the reliability of evidence reviews (Woodcock et al., 2014).

We aim to provide evidence on the quality of evidence syntheses in agri-food systems science. To assess the agri-food systems evidence syntheses, we searched the CEEDER database for 'evidence reviews' on 27.06.2024 using the following search string, which was iteratively built using key terms of agri-food systems (see Supplementary material): "(farm) OR (agriculture) OR (agricultural) OR (crop) OR (field) OR (arable) OR (meadow) OR (livestock) OR (food) OR (pasture) OR (animal husbandry) OR (yield) OR (nutrition) OR (value chain) OR (consumption) OR (waste) OR (rural)." This search yielded 1,143 hits from 2018 to 2024, accounting for approximately 55% of the total database and 66% of evidence reviews in the database. After removing duplicates ($n = 2$) and articles not directly related to agri-food systems ($n = 215$), such as those dealing with conservation, forestry, marine biology, 926 articles (54% of evidence reviews) remained for analysis. All articles were cross-checked by a second team member.
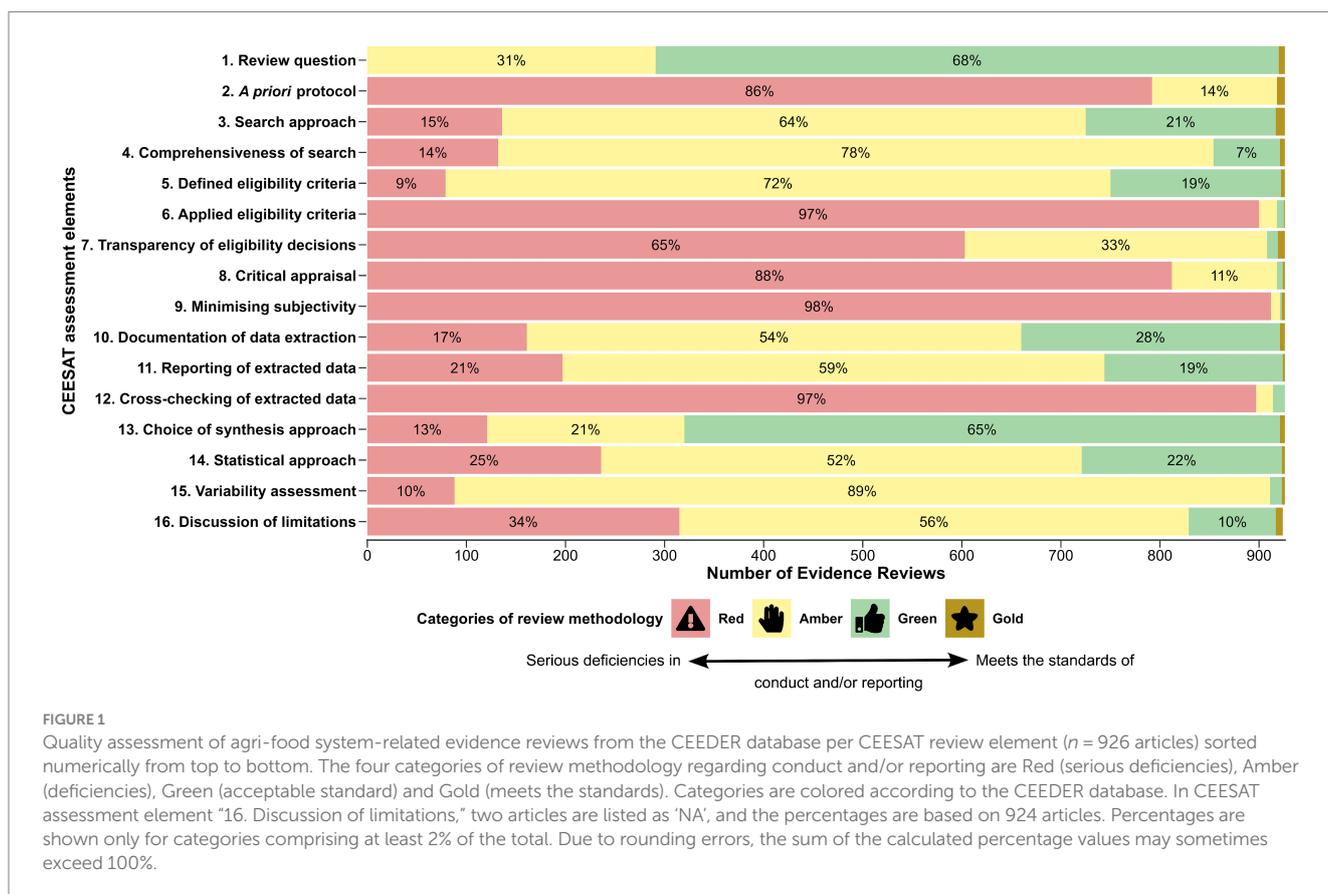
On average, 43% of the 16 CEESAT elements were categorised as Red, 40% as Amber, 17% as Green, and only about 1% as Gold.

The review elements in which most articles were classified as poor quality (Red) were related to minimising subjectivity during critical appraisal (98% of all reviews), application of the eligibility criteria

TABLE 1 The seven review components and respective review elements of the Collaboration for Environmental Evidence Synthesis Assessment Tool (CEESAT).

Review components	Review elements
Review question	1. Are the elements of the review question clear?
Method/Protocol	2. Is there an *a priori* method/protocol document?
Searching for studies	3. Is the approach to searching clearly defined, systematic and transparent?
	4. Is the search comprehensive?
Including studies	5. Are eligibility criteria clearly defined?
	6. Are eligibility criteria consistently applied to all potentially relevant articles and studies found during the search?
	7. Are eligibility decisions transparently reported?
Critical appraisal	8. Does the review critically appraise each study?
	9. During critical appraisal was an effort made to minimize subjectivity?
Data extraction	10. Is the method of data extraction fully documented?
	11. Are the extracted data reported for each study?
	12. Were extracted data cross-checked by more than one reviewer?
Data synthesis	13. Is the choice of synthesis approach appropriate?
	14. Is a statistical estimate of pooled effect (or similar) provided together with measure of variance and heterogeneity among studies?
	15. Is variability in the study findings investigated and discussed?
	16. Have the authors considered limitations in the synthesis?

Source: Konno et al. (2020b) and CEE (2022).

FIGURE 1
Quality assessment of agri-food system-related evidence reviews from the CEEDER database per CEESAT review element (*n* = 926 articles) sorted numerically from top to bottom. The four categories of review methodology regarding conduct and/or reporting are Red (serious deficiencies), Amber (deficiencies), Green (acceptable standard) and Gold (meets the standards). Categories are colored according to the CEEDER database. In CEESAT assessment element "16. Discussion of limitations," two articles are listed as 'NA', and the percentages are based on 924 articles. Percentages are shown only for categories comprising at least 2% of the total. Due to rounding errors, the sum of the calculated percentage values may sometimes exceed 100%.

(97%), cross-checking of extracted data by more than one reviewer (97%), critical appraisal of studies (88%), establishment of an *a priori* method/protocol (86%) and transparent reporting of eligibility decisions (65%) (Figure 1).

Most articles (>50%) had deficiencies (Amber) regarding the investigation and discussion of variability in study findings (89%), comprehensiveness of the search (78%), definition of eligibility criteria (72%), search approach (64%), reporting of extracted data for each

study (59%), consideration and discussion of the limitations of the synthesis (56%), documentation of data extraction (54%) and regarding the statistical approach (52%).

Only in the case of review elements regarding the clarity of the review question and the choice of the synthesis approach were most articles rated as acceptable (Green), at 68 and 65%, respectively. Among the other 14 review elements, the percentage of acceptable (Green) evidence reviews was always less than 30%. Less than 10 articles per review element (≤1%) were rated as high quality (Gold) (Figure 1).

3 Problems in and solutions for evidence syntheses

The following sections outline issues in evidence synthesis, citing examples from our assessment of the CEEDER database in the second section and other recent examples from the published literature on agri-food systems research, and propose solutions to mitigate these problems.

3.1 Review question

Description: A well-crafted review question is essential for any evidence synthesis because it establishes the core focus and system boundaries for the review process (Pullin et al., 2009). The key elements must be clearly defined.

Example: Zahra et al. (2021) do not state a review question or hypothesis but rather vaguely state a problem and an objective: "[…] this study was designed to determine the nexus between biochar and compost […]." The lack of clarity in the question means that the review's system boundaries are unclear.

Solution: Authors aiming for a systematic review should separate and specify their question's key elements using a framework such as the PI/ECO framework for comparative intervention/exposure questions (Population, Intervention/ Exposure, Comparator, Outcome) (see, e.g., Higgins et al., 2019; CEE, 2022).

3.2 Method/protocol

Description: Defining an *a priori* protocol that outlines the methods is critical when conducting evidence syntheses to improve transparency, comprehensiveness, reproducibility and to avoid mission creep and selective study choices (bias). It also helps to ensure focus and efficiency throughout the review process, precluding major deviations from the review objectives, search strategy, inclusion, and appraisal criteria. An *a priori* protocol reduces bias by explicitly specifying procedures that will be followed at each step, reducing the opportunity for subjectivity and transparently outlining the planned methods in detail (CEE, 2022).

Example: Viana et al. (2022) and Palomo-Campesino et al. (2018) lack a protocol and details in their methods description, for example, specifics about searched databases within the Web of Science platform

and the handling of consistency and subjectivity among reviewers. As in many cases, the authors argue that their review is systematic because of the use of PRISMA. However, this is a common misconception and demonstrates that the authors are confusing reporting guidance for conduct guidance. The review by Delitte et al. (2021) is particularly problematic because it lacks both method description and protocol. The lack of internationally accepted best practices and minimum standards for conduct and reporting raises the risk of bias, and insufficient details impede the assessment of the review's validity.

Solution: Authors aiming for evidence synthesis should design an *a priori* protocol outlining planned methods for searching, screening, data extraction, critical appraisal, and synthesis. Deviations from the protocol during study conduct should be described and justified. Organizations such as Cochrane, Campbell Collaboration and Collaboration for Environmental Evidence (CEE) provide standards and guidance for conducting high-quality evidence synthesis. In addition, authors can apply internationally recognized review reporting standards, such as ROSES (Reporting standards for Systematic Evidence Syntheses) (Haddaway et al., 2018b) or PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) (Page et al., 2021). They assist authors in ensuring that they accurately document all their actions in a clear language, whereas review protocols guide review conduct. Ideally, authors can avoid shortcomings by having their protocols peer-reviewed and published beforehand in suitable journals (e.g., Environmental Evidence), obtaining feedback from methodological experts during the peer-review process. However, authors facing monetary and/or time constraints may consider registering protocols using PROCEED's free service (JKI and CEE, 2023). Protocols are not peer-reviewed but undergo an editorial process with guidance and possible revisions before acceptance (Pullin, 2023).

3.3 Searching for studies

Description: Searching in evidence syntheses requires a detailed approach involving tried-and-tested search strings that carefully balance sensitivity and specificity. Validation against benchmark papers is essential to ensure the comprehensiveness of the searches. Reviewers should ensure that they search across multiple, relevant, and diverse sources for evidence, including bibliographic databases (e.g., Scopus, LILACS), theses and grey literature repositories and other relevant organizational websites. Neglecting grey literature may introduce publication bias, especially for negative or non-significant results that are typically underrepresented in peer-reviewed journals (Franco et al., 2014). Commonly used search facilities, such as Google Scholar, may be deemed inappropriate because they may introduce publication bias, among other limitations (Møller and Jennions, 2001; Aguillo, 2012; Bramer et al., 2016; Kugley et al., 2017; Gusenbauer and Haddaway, 2020). The exclusion of non-English languages can lead to bias (Konno et al., 2020a). Reviews that examine regionally specific topics must consider local language evidence - both in terms of search design and study inclusion.

Example: Dagunga et al. (2023) may have overlooked valuable sources by omitting grey literature, particularly in "agroecology"

research, where it can serve as an important repository, as highlighted by D'Annolfo et al. (2014) and Hussain et al. (2014). Malhi et al. (2021) exclusively relied on Google Scholar for their search, an unrepeatable resource from which results cannot be exported in full.

Solution: Including or consulting information specialists or librarians in the review process is an important means of avoiding common mistakes when constructing a search strategy (Rethlefsen et al., 2015; Meert et al., 2017). To enhance comprehensiveness and representativeness, a suite of different search methods should be employed, including: multiple databases, forward and backward citation chasing, web searches, manual checking of non-indexed journals (Gusenbauer, 2024), and calls for submission of relevant studies. Having a peer-reviewed *a priori* protocol enhances the validity of the search strategy by facilitating constructive feedback.

3.4 Including studies

Description: Systematic eligibility screening or study selection requires *a priori* and well-defined criteria that are consistently applied to all studies identified in the search process. This is to minimize errors or selection bias, which occurs when the included articles fail to be representative of the evidence base (McDonagh et al., 2013). Consistency of screening decisions and resolving disagreements among the review team are essential to reduce subjectivity in study inclusion. The methods used for consistent screening decisions must be described and reported transparently.

Example: Dagunga et al. (2023) do neither report on consistency checking nor define what a "smallholder farmer" is. It is not clear which studies are included, those farmers who have a certain smallholding size, or those that use the label "smallholder." The main terms "agroecology" and "resilience," as well as their measurements, are only defined in the results section as part of the analysis. As central definitions and inclusion criteria for studies (e.g., type of measurement and empirical vs. non-empirical studies) should be defined *a priori*, this indicates that the review might not be entirely systematic. Ricciardi et al. (2021) aimed at comparing farms with similar management systems but lack details on eligibility criteria for such management systems, how many screeners were involved and how they dealt with consistency. Authors may have conducted a noble study, but unreported numbers and excluded articles during screening (Muneret et al., 2018; Jat et al., 2020) or the absence of a list of included studies (Bai et al., 2022), limit transparency and trust in the results and conclusions.

Solutions: The review question should be used as the basis for defining eligibility criteria. Evidence reviews applying PICO/PECO-type research questions may use these as key elements [Population(s), Intervention(s), Comparator(s), Outcome(s)] for defining the eligibility criteria. Additional entries, such as study design, language, geography, and study design should also be clarified (Bernes et al., 2018; Macura et al., 2019). Involving multiple reviewers and piloting the screening process for disagreements further enhances consistency (Waffenschmidt et al., 2019). Before full-text screening begins, disagreements must be resolved, screening decisions kept consistent, and criteria clearly defined for operational clarity. This process is crucial at multiple stages of the review. Transparency and replicability

are ensured by providing details on the eligibility criteria application, the complete list of screened articles, included studies, and reasons for article exclusion at the full-text stage.

3.5 Critical appraisal and minimising subjectivity

Description: Critical appraisal aims to assess the internal validity (quality) and external validity (generalizability) of individual studies included in a review. This process is important because research studies vary considerably in terms of their reliability, for example, study design (CEE, 2022), and overlap with review aims. Evidence syntheses should be based on reliable primary research, but moderate-quality research can be included; however, all studies should be weighted accordingly and tested for robustness (CEE, 2022). The validity assessment process could be influenced by implicit bias (i.e., the reviewer's perspective) (Konno et al., 2024) and should hence not depend on a single reviewer. It should be based on tried-and-tested frameworks for critical appraisal, for example, tools for assessing risk of bias in randomized control trials, such as RoB2.0. It is essential to provide detailed information regarding the decision-making process during critical appraisal.

Example: Abdalla et al. (2018) and Liu et al. (2018) omitted critical appraisal or failed to report a validity assessment; hence, their review's reliability cannot be assessed by the reader. Only three systematic reviews in our CEEDER assessment are rated "high-quality" (Gold) for their critical appraisal (Bernes et al., 2018; Eales et al., 2018; Meurer et al., 2018). All three studies clearly presented a critical appraisal process and validity assessment.

Solution: Critical appraisal should be planned and tested *a priori* to ensure consistency among the review team. The critical appraisal criteria should reflect what the review team deemed to be critical variables influencing the reliability of the study, as for example in Eales et al. (2018). The study validity rating should avoid complex scoring systems, but instead use a simple rating, such as low, high, and unclear validity, to avoid introducing errors (Higgins et al., 2011). Various established and verified critical appraisal tools are available (e.g., Frampton et al., 2017).

3.6 Data extraction

Description: Data extraction and coding in evidence synthesis involves systematically retrieving relevant information from included articles in a standardized manner, including qualitative and/or quantitative data. This process requires a description of the data extraction method and coding of the studies. This includes variables that are extracted as data and others that are coded (metadata). Additionally, it should be outlined how the process was tested for repeatability and consistency between reviewers (Frampton et al., 2017). The documentation of each extraction step and the approach to acquiring missing or unclear data are crucial for transparency (CEE, 2022).

Example: Adil et al. (2022) partly report how they extracted data but do not report specific steps, criteria, or procedures followed

during the data extraction process, nor how the data extraction was validated to minimize errors and bias in the results. Additionally, the extracted data are not presented anywhere, despite a data availability statement stating the opposite. In comparison, Blouin et al. (2019) provide a more detailed account of their data extraction process. However, more explicit information on the validation process for the extracted data and the approach taken for handling missing or insufficient data are lacking.

Solution: Good data extraction practices for enhanced objectivity, reproducibility, and transparency involve: (i) fully documenting the type of data to be extracted and extraction method; (ii) employing a data extraction form pretested by two reviewers; (iii) including appendices with full extracted data/meta-data from each primary study; (iv) specifying pre-analysis calculations and data transformations, e.g., calculation of effect sizes; (v) ensuring consistency through multiple reviewers' testing for human error, and in the best case, two independent reviewers extracting data from each study (CEE, 2022).

3.7 Data synthesis

Description: Data synthesis combines and analyses data from studies included in the evidence synthesis to answer the review question, draw conclusions, or identify patterns. This involves the synthesis of study characteristics through narrative, quantitative (e.g., meta-analysis), qualitative (e.g., thematic synthesis) or mixed methods synthesis (Sandelowski et al., 2010, 2012; Pearson et al., 2015; Higgins et al., 2019; CEE, 2022). Vote counting as a form of quantitative synthesis should not be used, as it is misleading and neglects each study's validity, statistical power, precision and magnitude (Friedman, 2001).

Example: Dittmer et al. (2023) and Cozim-Melges et al. (2024) used vote counting to show that agroecology and diversified farming practices positively affect climate change adaptation and enhance biodiversity. The identified interventions' effect size, magnitude, and statistical power are missing. Dhaliwal et al. (2019) opted for a narrative synthesis. Nonetheless, the effect of agricultural addendums in a comparative way is missing, given the lack of a quantitative synthesis of the included studies. On the other hand, Tuttle and Donahue (2022) clearly state that their evidence synthesis is a narrative synthesis. This is accompanied by tables and visualizations that summarize the information retrieved from the included articles.

Solution: The synthesis of data should follow tried-and-tested methods. In quantitative synthesis, vote counting should never be used instead of meta-analysis. In cases where meta-analysis is not feasible (e.g., due to data heterogeneity), a narrative synthesis should be employed to describe the evidence base. All included studies' characteristics, outcomes, and validity should be listed and tabulated. The chosen methods/models for synthesis should be justified, assumptions clarified (e.g., for missing data), and decisions made explicit (e.g., choice of effect measure). Following the call for 'Open synthesis' (Haddaway, 2018a), input data for quantitative synthesis and analytical code should be provided where not legally restricted or data summaries where restrictions exist. Authors should transparently outline potential limitations and uncertainties of the analysis approach, detailing their potential impact on the overall conclusions.

4 Conclusion

4.1 Improving awareness of the need for methodological rigor in evidence synthesis at all levels

We recognize the challenging nature of systematic reviews, which demand substantial resources, expertise, and time. Moreover, the pressure in academic research to frequently produce and publish (Van Dalen, 2021; Becker and Lukka, 2023) can drive researchers to cut corners to meet deadlines, thus reducing methodological rigor and validity. In resource-constrained scenarios, researchers may employ a 'rapid review' or 'scoping review' methodology, incorporating systematic review elements. However, the risk of bias can increase (CEE, 2022; Sabiston et al., 2022). Pursuing methodological rigor in evidence synthesis must be a shared responsibility among authors, peer reviewers, and editors to enhance the quality, reliability, and credibility of evidence synthesis in agri-food science. Despite challenges such as lack of awareness, limited training opportunities, and time constraints, it is critical to foster commitment within the research community to adhere to best practices. This requires willingness to learn, appreciation of the efforts involved in rigorous reviews, and collaboration among stakeholders. To avoid unreliable syntheses, all actors in the research community should be cognizant of common mistakes.

4.2 Scaled training for researchers in evidence synthesis methods

Many researchers, as well as postgraduate and PhD students, have not been trained in best-practice evidence synthesis. Providing comprehensive and accessible training to researchers can address the current gaps in awareness, skills, and best practices. Several organizations, such as Collaboration for Environmental Evidence and Cochrane, offer training opportunities (CEE, 2024c; Cochrane, 2024). However, fostering a proactive approach requires the active involvement of universities and other research institutions in incorporating such training programs into their curricula and continuing education.

4.3 Peer review and publication of *a priori* protocols

Peer review and publication of review protocols are regarded as best practices and can enhance the quality of evidence synthesis thanks to the critical evaluation of the review's design, methods, and planned analyses. Thus, redundancies and wasted research efforts can be avoided before conducting the study. Overall, this practice promotes transparency, reduces bias, and enhances credibility.

4.4 Better gatekeeping at publication to catch substandard methodology

The lenient acceptance of papers claiming to be "systematic" is turning "systematic" into a meaningless buzzword devoid of its original rigor and purpose. To enhance gatekeeping at publication, editors and publishers must establish robust peer-review processes that

ensure that evidence syntheses meet established methodological and reporting standards. Journals can provide authors with clear guidelines that require transparency and reporting in sufficient detail, as well as adherence to recognized protocols (e.g., PRISMA, ROSES). Additionally, editors can enlist expert reviewers with knowledge of evidence synthesis methodology to thoroughly assess submitted papers.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

PE: Data curation, Formal analysis, Validation, Writing – original draft, Writing – review & editing. GK: Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. MK-D: Writing – original draft, Writing – review & editing. EV: Writing – original draft, Writing – review & editing. CB: Writing – original draft, Writing – review & editing. HL: Writing – original draft, Writing – review & editing. DB-T: Writing – original draft, Writing – review & editing. EE: Writing – original draft, Writing – review & editing. JV-V: Conceptualization, Writing – original draft, Writing – review & editing. DIA-O: Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsufs.2024.1410205/full#supplementary-material

SUPPLEMENTARY TABLE S1
Number of articles per iterative agri-food systems key term combination.

SUPPLEMENTARY TABLE S2
Lists of included and excluded articles.

References

Abdalla, M., Hastings, A., Chadwick, D. R., Jones, D. L., Evans, C. D., Jones, M. B., et al. (2018). Critical review of the impacts of grazing intensity on soil organic carbon storage and other soil quality indicators in extensively managed grasslands. *Agric. Ecosyst. Environ.* 253, 62–81. doi: 10.1016/j.agee.2017.10.023

Adil, M., Zhang, S., Wang, J., Shah, A. N., Tanveer, M., and Fiaz, S. (2022). Effects of fallow management practices on soil water, crop yield and water use efficiency in winter wheat monoculture system: a Meta-analysis. *Front. Plant Sci.* 13:825309. doi: 10.3389/fpls.2022.825309

Aguillo, I. F. (2012). Is Google scholar useful for bibliometrics? A webometric analysis. *Scientometrics* 91, 343–351. doi: 10.1007/s11192-011-0582-8

Bai, S. H., Omidvar, N., Gallart, M., Kämper, W., Tahmasbian, I., Farrar, M. B., et al. (2022). Combined effects of biochar and fertilizer applications on yield: a review and meta-analysis. *Sci. Total Environ.* 808:152073. doi: 10.1016/j.scitotenv.2021.152073

Becker, A., and Lukka, K. (2023). Instrumentalism and the publish-or-perish regime. *Crit. Perspect. Account.* 94:102436. doi: 10.1016/j.cpa.2022.102436

Bernes, C., Macura, B., Jonsson, B. G., Junninen, K., Müller, J., Sandström, J., et al. (2018). Manipulating ungulate herbivory in temperate and boreal forests: effects on vegetation and invertebrates. A systematic review. *Environ. Evid.* 7:13. doi: 10.1186/s13750-018-0125-3

Blouin, M., Barrere, J., Meyer, N., Lartigue, S., Barot, S., and Mathieu, J. (2019). Vermicompost significantly affects plant growth. A meta-analysis. *Agron. Sustain. Dev.* 39:34. doi: 10.1007/s13593-019-0579-x

Bramer, W. M., Giustini, D., and Kramer, B. M. R. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google scholar: a prospective study. *Syst. Rev.* 5:39. doi: 10.1186/s13643-016-0215-7

CEE (2022). Guidelines and Standards for Evidence Synthesis in Environmental Management. Version 5.1. Collaboration for Environmental Evidence (CEE). Available at: www.environmentalevidence.org/information-for-authors (Accessed January 6, 2024).

CEE (2024a). About the CEEDER database. Collaboration for Environmental Evidence (CEE). Available at: https://environmentalevidence.org/ceeder-search/ (Accessed March 6, 2024).

CEE (2024b). What is CEESAT? About CEESAT. Collaboration for environmental evidence (CEE). Available at: https://environmentalevidence.org/ceeder/about-ceesat/ (Accessed March 6, 2024).

CEE (2024c). Training content. Collaboration for Environmental Evidence (CEE). Available at: https://environmentalevidence.org/training-2/ (Accessed March 6, 2024).

Cochrane (2024). Cochrane Interactive Learning. Available at: https://training.cochrane.org/interactivelearning (Accessed March 6, 2024).

Cozim-Melges, F., Ripoll-Bosch, R., Veen, G. F., Oggiano, P., Bianchi, F. J. J. A., Van Der Putten, W. H., et al. (2024). Farming practices to enhance biodiversity across biomes: a systematic review. *NPJ Biodivers* 3:1. doi: 10.1038/s44185-023-00034-2

D'Annolfo, R., Graeub, B., and Gemmill-Herren, B. (2014). Agroecological socio-economics: Agroecology's contribution to farm incomes, labour and other socio-economic dimensions of food systems. Proceedings of the FAO International Symposium, (Rome, Italy). 19, 332–347.

Dagunga, G., Ayamga, M., Laube, W., Ansah, I. G. K., Kornher, L., and Kotu, B. H. (2023). Agroecology and resilience of smallholder food security: a systematic review. *Front. Sustain. Food Syst.* 7:1267630. doi: 10.3389/fsufs.2023.1267630

Delitte, M., Caulier, S., Bragard, C., and Desoignies, N. (2021). Plant microbiota beyond farming practices: a review. *Front. Sustain. Food Syst.* 5:624203. doi: 10.3389/fsufs.2021.624203

Dhaliwal, S. S., Naresh, R. K., Mandal, A., Walia, M. K., Gupta, R. K., Singh, R., et al. (2019). Effect of manures and fertilizers on soil physical properties, build-up of macro and micronutrients and uptake in soil under different cropping systems: a review. *J. Plant Nutr.* 42, 2873–2900. doi: 10.1080/01904167.2019.1659337

Dittmer, K. M., Rose, S., Snapp, S. S., Kebede, Y., Brickman, S., Shelton, S., et al. (2023). Agroecology can promote climate change adaptation outcomes without compromising yield in smallholder systems. *Environ. Manag.* 72, 333–342. doi: 10.1007/s00267-023-01816-x

Eales, J., Haddaway, N. R., Bernes, C., Cooke, S. J., Jonsson, B. G., Kouki, J., et al. (2018). What is the effect of prescribed burning in temperate and boreal forest on biodiversity, beyond pyrophilous and saproxylic species? A systematic review. *Environ. Evid.* 7:19. doi: 10.1186/s13750-018-0131-5

Frampton, G. K., Livoreil, B., and Petrokofsky, G. (2017). Eligibility screening in evidence synthesis of environmental management topics. *Environ. Evid.* 6, 1–13. doi: 10.1186/s13750-017-0102-2

Franco, A., Malhotra, N., and Simonovits, G. (2014). Publication bias in the social sciences: unlocking the file drawer. *Science* 345, 1502–1505. doi: 10.1126/science.1255484

Friedman, L. (2001). Why vote-count reviews don't count. *Biol. Psychiatry* 49, 161–162. doi: 10.1016/S0006-3223(00)01075-1

GCSA (2019). Scientific advice to European policy in a complex world. LU: Group of Chief Scientific Advisors.

Gough, D., Thomas, J., and Oliver, S. (2017). An introduction to systematic reviews: Sage Publications Ltd., 352.

Gusenbauer, M. (2024). Beyond Google scholar, Scopus, and web of science: an evaluation of the backward and forward citation coverage of 59 databases' citation indices. *Res. Synth. Methods*. doi: 10.1002/jrsm.1729

Gusenbauer, M., and Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google scholar, PubMed, and 26 other resources. *Res. Synth. Methods* 11, 181–217. doi: 10.1002/jrsm.1378

Haddaway, N. R. (2018a). Open synthesis: on the need for evidence synthesis to embrace Open Science. *Environ. Evid.* 7, 1–5. doi: 10.1186/s13750-018-0140-4

Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., et al. (2020). Eight problems with literature reviews and how to fix them. *Nat. Ecol. Evol.* 4, 1582–1589. doi: 10.1038/s41559-020-01295-x

Haddaway, N. R., Land, M., and Macura, B. (2017). A little learning is a dangerous thing: a call for better understanding of the term systematic review. *Environ. Int.* 99, 356–360. doi: 10.1016/j.envint.2016.12.020

Haddaway, N. R., Macura, B., Whaley, P., and Pullin, A. S. (2018b). ROSES RepOrting standards for systematic evidence syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environ. Evid.* 7:7. doi: 10.1186/s13750-018-0121-7

Higgins, J. P. T., Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343:d5928. doi: 10.1136/bmj.d5928

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., et al. (2019). Cochrane Handbook for Systematic Reviews of Interventions. *2nd* Edn: Wiley-Blackwell. doi: 10.1002/9781119536604

HLPE (2019). Agroecological and other innovative approaches for sustainable agriculture and food systems that enhance food security and nutrition. Rome: High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security.

Hussain, S., Miller, D., Gemmill-Herren, B., and Bogdanski, A. (2014). Agroecology and economics of ecosystems and biodiversity: the devil is in the detail., in proceedings of the FAO international symposium. Italy: Rome.

Ioannidis, J. P. A. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and Meta-analyses. *Milbank Q.* 94, 485–514. doi: 10.1111/1468-0009.12210

IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Paris, France: Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.

IPCC (2019). "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. eds. P. R. Shukla, J. Skea, E. C. Buendia, V. Masson-Delmonte, H.-O. Pörtner and D. C. Robertset al. (Cambridge University Press). Available at: https://www.ipcc.ch/srccl/

Jat, M. L., Chakraborty, D., Ladha, J. K., Rana, D. S., Gathala, M. K., McDonald, A., et al. (2020). Conservation agriculture for sustainable intensification in South Asia. *Nat. Sustain.* 3, 336–343. doi: 10.1038/s41893-020-0500-2

JKI, and CEE (2023). PROCEED. *PROCEED 2024 v0.9.1.* Available at: https://www.proceedevidence.info/ (Accessed March 6, 2024).

Konno, K., Gibbons, J., Lewis, R., and Pullin, R. (2024). Potential types of bias when estimating causal effects in environmental research and how to interpret them. *Environ Evid*. vol. 13:6373.

Konno, K., Akasaka, M., Koshida, C., Katayama, N., Osada, N., Spake, R., et al. (2020a). Ignoring non-English-language studies may bias ecological meta-analyses. *Ecol. Evol.* 10, 6373–6384. doi: 10.1002/ece3.6368

Konno, K., Cheng, S. H., Eales, J., Frampton, G., Kohl, C., Livoreil, B., et al. (2020b). The CEEDER database of evidence reviews: an open-access evidence service for researchers and decision-makers. *Environ. Sci. Pol.* 114, 256–262. doi: 10.1016/j.envsci.2020.08.021

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. K., Hammerstrøm, K., et al. (2017). Searching for studies: a guide to information retrieval for Campbell systematic reviews. *Campbell Syst. Rev.* 13, 1–73. doi: 10.4073/cmg.2016.1

Liu, X., Yang, T., Wang, Q., Huang, F., and Li, L. (2018). Dynamics of soil carbon and nitrogen stocks after afforestation in arid and semi-arid regions: a meta-analysis. *Sci. Total Environ.* 618, 1658–1664. doi: 10.1016/j.scitotenv.2017.10.009

Macura, B., Byström, P., Airoldi, L., Eriksson, B. K., Rudstam, L., and Støttrup, J. G. (2019). Impact of structural habitat modifications in coastal temperate systems on fish recruitment: a systematic review. *Environ. Evid.* 8:14. doi: 10.1186/s13750-019-0157-3

Malhi, G. S., Kaur, M., and Kaushik, P. (2021). Impact of climate change on agriculture and its mitigation strategies: a review. *Sustain. For.* 13:1318. doi: 10.3390/su13031318

McDonagh, M., Peterson, K., Raina, P., Chang, S., and Shekelle, P. (2013). Avoiding Bias in selecting studies. Methods guide for comparative effectiveness reviews. Rockville, MD: Agency for Healthcare Research and Quality.

Meert, D., Torabi, N., and Costella, J. (2017). Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. *JMLA*:104. doi: 10.5195/jmla.2016.139

Meurer, K. H. E., Haddaway, N. R., Bolinder, M. A., and 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 Sci. Rev.* 177, 613–622. doi: 10.1016/j.earscirev.2017.12.015

Møller, A. P., and Jennions, M. D. (2001). Testing and adjusting for publication bias. *Trends Ecol. Evol.* 16, 580–586. doi: 10.1016/S0169-5347(01)02235-2

Muneret, L., Mitchell, M., Seufert, V., Aviron, S., Djoudi, E. A., Pétillon, J., et al. (2018). Evidence that organic farming promotes pest control. *Nat. Sustain.* 1, 361–368. doi: 10.1038/s41893-018-0102-4

O'Leary, B. C., Kvist, K., Bayliss, H. R., Derroire, G., Healey, J. R., Hughes, K., et al. (2016). The reliability of evidence review methodology in environmental science and conservation. *Environ. Sci. Pol.* 64, 75–82. doi: 10.1016/j.envsci.2016.06.012

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71

Palomo-Campesino, S., González, J., and García-Llorente, M. (2018). Exploring the connections between Agroecological practices and ecosystem services: a systematic literature review. *Sustain. For.* 10:4339. doi: 10.3390/su10124339

Pearson, A., White, H., Bath-Hextall, F., Salmond, S., Apostolo, J., and Kirkpatrick, P. (2015). A mixed-methods approach to systematic reviews. *Int. J. Evid. Based Healthc.* 13, 121–131. doi: 10.1097/XEB.0000000000000052

Pullin, A. (2023). Introducing PROCEED: a new service for fast registration and publication of protocols for environmental evidence syntheses, including rapid reviews. *Environ. Evid.* 12:s13750-022-00295–7. doi: 10.1186/s13750-022-00295-7

Pullin, A. S., and Knight, T. M. (2001). Effectiveness in conservation practice: pointers from medicine and public health. *Conserv. Biol.* 15, 50–54. doi: 10.1111/j.1523-1739.2001.99499.x

Pullin, A. S., Knight, T. M., and Watkinson, A. R. (2009). Linking reductionist science and holistic policy using systematic reviews: unpacking environmental policy questions to construct an evidence-based framework. *J. Appl. Ecol.* 46, 970–975. doi: 10.1111/j.1365-2664.2009.01704.x

Rethlefsen, M. L., Farrell, A. M., Osterhaus Trzasko, L. C., and Brigham, T. J. (2015). Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J. Clin. Epidemiol.* 68, 617–626. doi: 10.1016/j.jclinepi.2014.11.025

Ricciardi, V., Mehrabi, Z., Wittman, H., James, D., and Ramankutty, N. (2021). Higher yields and more biodiversity on smaller farms. *Nat. Sustain.* 4, 651–657. doi: 10.1038/s41893-021-00699-2

Richardson, K., Steffen, W., Lucht, W., Bendtsen, J., Cornell, S. E., Donges, J. F., et al. (2023). Earth beyond six of nine planetary boundaries. *Sci. Adv.* 9:eadh2458. doi: 10.1126/sciadv.adh2458

Ruggeri Laderchi, C., Lotze-Campen, H., DeClerck, F., Bodirsky, B. L., Collignon, Q., Crawford, M. S., et al. (2024). The economics of the food system transformation. *Food*

System Economics Commission (FSEC), Global Policy Report. Available at: https://foodsystemeconomics.org/wp-content/uploads/FSEC-Global_Policy_Report.pdf

Sabiston, C. M., Vani, M., De Jonge, M., and Nesbitt, A. (2022). Scoping reviews and rapid reviews. *Int. Rev. Sport Exerc. Psychol.* 15, 91–119. doi: 10.1080/1750984X.2021.1964095

Sandelowski, M. J., Voils, C. I., and Barroso, J. (2010). Defining and designing mixed research synthesis studies. *Research in the schools: a nationally refereed journal sponsored by the Mid-South Educational Research Association and the University of Alabama.* 13:29. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2809982/

Sandelowski, M., Voils, C. I., Leeman, J., and Crandell, J. L. (2012). Mapping the mixed methods-mixed research synthesis terrain. *J. Mixed Methods Res.* 6, 317–331. doi: 10.1177/1558689811427913

SAPEA (2020). A sustainable food system for the European Union. *1st* Edn. DE: SAPEA.

Tuttle, L. J., and Donahue, M. J. (2022). Effects of sediment exposure on corals: a systematic review of experimental studies. *Environ. Evid.* 11:4. doi: 10.1186/s13750-022-00256-0

Uttley, L., Quintana, D. S., Montgomery, P., Carroll, C., Page, M. J., Falzon, L., et al. (2023). The problems with systematic reviews: a living systematic review. *J. Clin. Epidemiol.* 156, 30–41. doi: 10.1016/j.jclinepi.2023.01.011

Van Dalen, H. P. (2021). How the publish-or-perish principle divides a science: the case of economists. *Scientometrics* 126, 1675–1694. doi: 10.1007/s11192-020-03786-x

Viana, C. M., Freire, D., Abrantes, P., Rocha, J., and Pereira, P. (2022). Agricultural land systems importance for supporting food security and sustainable development goals: a systematic review. *Sci. Total Environ.* 806:150718. doi: 10.1016/j.scitotenv.2021.150718

Waffenschmidt, S., Knelangen, M., Sieben, W., Bühn, S., and Pieper, D. (2019). Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. *BMC Med. Res. Methodol.* 19, 132–139. doi: 10.1186/s12874-019-0782-0

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., et al. (2019). Food in the Anthropocene: the EAT–lancet commission on healthy diets from sustainable food systems. *Lancet* 393, 447–492. doi: 10.1016/S0140-6736(18)31788-4

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

Zahra, M. B., Fayyaz, B., Aftab, Z.-E.-H., and Haider, M. S. (2021). Mitigation of degraded soils by using biochar and compost: a systematic review. *J. Soil Sci. Plant Nutr.* 21, 2718–2738. doi: 10.1007/s42729-021-00558-1