Ecosystem Services 81 (2026) 101905

Contents lists available at ScienceDirect

Ecosystem Services

journal homepage: www.elsevier.com/locate/ecoser

Review Paper

Potential of temperate agroforestry systems to deliver ecosystem services: An evidence map

Andrea Schievano [a,*], Camille Rubeaud [b], Margret Köthke [c], Beatrice Landoni [d], Damien Beillouin [e], Marta Pérez-Soba [f], Simona Bosco [a], Rui Catarino [a], Giovanni Tamburini [g], David Makowski [h], Sonja Kay [b]

[a] European Commission, Joint Research Centre (JRC), Ispra (VA), Italy
[b] Agroscope, Zürich, Switzerland
[c] Thünen Institute of Forestry, Hamburg, Germany
[d] Department of Bioscience, University of Milano, Milano, Italy
[e] CIRAD, UPR HortSys, Univ Montpellier, Montpellier, France
[f] Wageningen University and Research, Droevendaalsesteeg 4, 6708, PB, Wageningen, the Netherlands
[g] Department of Soil, Plant and Food Sciences (DiSSPA – Entomology and Zoology), University of Bari Aldo Moro, Bari, Italy
[h] University Paris-Saclay, INRAe, AgroParisTech, Unit MIA, PS, Palaiseau, France

ARTICLE INFO

Keywords:
Agroforestry
Ecosystem services
Temperate pedoclimates
Evidence map
Meta analysis

ABSTRACT

Agroforestry is increasingly promoted as a nature-based strategy towards a more sustainable agriculture, yet quantitative evidence of its benefits, especially in temperate regions, remains fragmented. Here, we synthesized 42 meta-analyses to map the potential of temperate agroforestry to deliver ecosystem services, assess their methodological quality, and identify knowledge gaps using the CICES classification framework. We find strong and consistent evidence for a positive effect on regulating ecosystem services: over 65% of assessments reported significantly positive outcomes for soil carbon sequestration and soil quality indicators. Effects on provisioning services are more variable, with around 50% of assessments for cultivated plant production reporting negative outcomes, a finding that likely reflects the narrow metrics (e.g. single-crop yield) used, rather than true system-level deficits, as more integrative indicators (such as the land equivalent ratio) are rarely reported. Mapping results onto CICES highlights that only 8 of 38 ecosystem service groups are covered by meta-analytical evidence, with cultural services entirely absent. Methodological quality was heterogeneous, with frequent shortcomings in publication bias assessment, data sharing and transparent reporting of study selection. We also provide a georeferenced database of 1495 primary studies to facilitate future syntheses and a global map of study locations. This synthesis provides a robust evidence base for policymakers, pinpointing both established benefits and critical research gaps needed to fully leverage agroforestry in temperate regions.

1. Introduction

Agroforestry – the deliberate integration of trees into cropping and livestock systems - is increasingly recognized as a multifunctional land-use strategy to address multiple global challenges simultaneously, including climate change mitigation, biodiversity conservation, soil health regeneration, and sustainable food production (Beillouin et al., 2021; Castle et al., 2022; Nair et al., 2021; Tamburini et al., 2020). Agroforestry encompasses a broad range of practices and systems, including silvoarable, silvopastoral, and agrosilvopastoral systems. These systems are highly diverse, with varied structural and botanical compositions, and each has distinct environmental and productivity implications (Castle et al., 2022; Kuyah et al., 2019). The integration of trees with crops or livestock has often been shown to significantly enhance biodiversity, carbon sequestration, soil health, and overall system resilience, particularly in subtropical and tropical regions (Ngaba et al., 2024). The effectiveness of agroforestry practices or systems (hereafter simply called "agroforestry") in delivering ecosystem services varies considerably depending on the type of system, tree species, spatial arrangement, management intensity and the pedo-climatic

* Corresponding author.
E-mail address: schievanoa@gmail.com (A. Schievano).

https://doi.org/10.1016/j.ecoser.2026.101905
Received 17 October 2025; Received in revised form 18 June 2026; Accepted 19 August 2026
Available online 2 September 2026

conditions (Kim et al., 2016; Shi et al., 2018). This highlights the need for region-specific research to optimize and assess agroforestry outcomes.

However, most available evidence, as synthesized by published meta-analyses (MAs) pertains to agroforestry in tropical and subtropical latitudes (tropical, arid or semi-arid pedo-climates), primarily in the southern hemisphere, whereas information remains comparatively scarce for temperate latitudes (including pedo-climates: temperate, mediterranean, boreal, continental, polar, etc.) (Torralba et al., 2016). In the EU, agroforestry is particularly prevalent in Mediterranean regions, as nature-based solutions for climate adaptation, biodiversity conservation, and rural development. For example, the Montado and Dehesa systems of Portugal and Spain that combine holm oaks or cork oaks with grazing (sheep, pigs) and cereal cultivation. Therefore, continuous update of the available evidence, especially in these contexts, is of primary importance. However, comprehensive evidence on the impacts of agroforestry across sustainability outcomes in these regions is still lacking.

Previous evidence syntheses have, in parallel, mapped the effects of agroforestry on various outcomes, including biodiversity, soil properties, productivity (Beillouin et al., 2019; Mathieu et al., 2025; Tamburini et al., 2020). Other evidence maps (Castle et al., 2022; Köthke et al., 2022) revealed that linear boundary plantings, such as hedgerows and windbreaks, have been extensively studied for their impacts on biodiversity, soil and water quality, and carbon sequestration. However, gaps remain in our understanding of agroforestry's effects on a broad range of sustainability outcomes, particularly for northern-hemisphere pedo-climates. Importantly, keeping evidence syntheses current is a key challenge. For instance, the most recent map of meta-analyses was published in 2025 (Mathieu et al., 2025); however, this analysis is based on a literature search performed in 2021. Given the rapid evolution of synthesis literature in this field (see Fig. 1), keeping evidence maps up to date and fostering collaboration across research teams is therefore essential.

It also remains unclear how robust the available empirical evidence is, to explain the potential contributions of agroforestry to the provision of ecosystem services, especially in temperate regions. According to some systematic literature maps (Mathieu et al., 2025; Tamburini et al., 2020), agroforestry in temperate regions provides numerous ecosystem services, including carbon sequestration, soil fertility enhancement, and biodiversity conservation. For instance, European studies highlight the benefits of wood pastures in Mediterranean, Atlantic, and Continental landscapes, focusing on habitat provision, climate regulation, and aesthetic value (Fagerholm et al., 2016). However, most reviews do not provide transparent linkages between empirical metrics (used in field experiments) and ecosystem services, as defined by the CICES

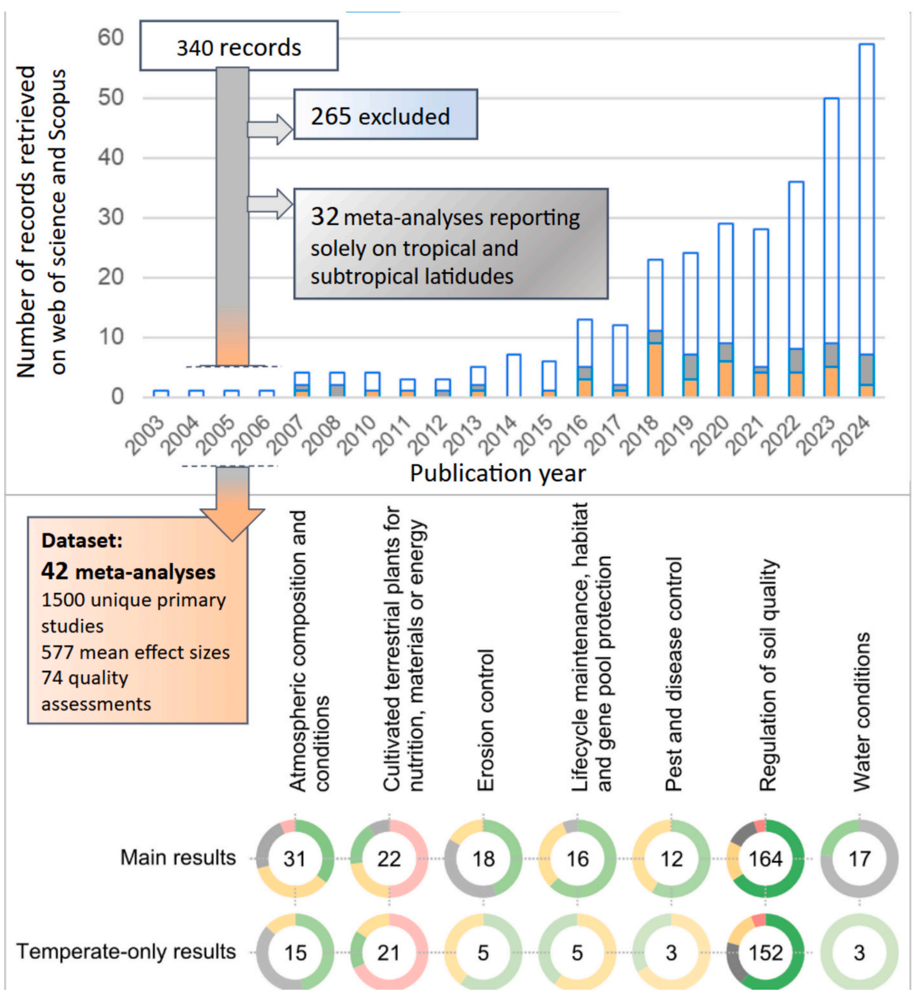

Fig. 1. Chronological distribution of the systematic review of meta-analyses (MAs), schematic PRISMA diagram and overview of the main results. Results (i.e. mean effect sizes extracted from the select MAs) were extracted for main assessments, as well as for temperate-only subgroups. Metrics used in empirical studies were linked across classes of Ecosystem services (CICES classification). Donut plots show the share of results showing significant positive (green) or negative effects (red), non-significant effect (yellow) or non-statistically-tested results (grey), as presented by the selected MAs. The numbers report the total count of effect sizes reported by the selected MAs. The full PRISMA statement diagram (Page et al., 2021) is reported in Supplementary Fig. 1. Detailed classification of specific metrics belonging to each class of outcomes is available in Supplementary Table 4.

classification (Sauer et al., 2021; Tamburini et al., 2020).

Another issue regards the quality of the MAs published on agroforestry. The increasing number of MAs, reporting on agroforestry and more in general in agricultural and environmental sciences, underscores their value for informing policy and practice by quantitatively synthesizing numerous experimental results (Schievano et al., 2024). Ideally, MAs should adhere to a rigorous, transparent systematic review process—encompassing clear scoping, comprehensive literature searches, unbiased study selection, standardized data extraction, appropriate statistical analysis, and thorough bias assessment—to ensure the validity of their conclusions. However, previous assessments in the field of agricultural sustainability (including on agroforestry) have revealed significant shortcomings (Beillouin et al., 2018; Fohrafellner et al., 2023; Schievano et al., 2026). For instance, crucial steps like publication bias assessment are often overlooked (reported in only 16–40% of MAs in some reviews), and replicability is frequently compromised by limited data sharing (18–35% providing datasets) and inadequately documented search and selection strategies (Beillouin et al., 2018; Fohrafellner et al., 2023). This lack of transparency contravenes FAIR (Findability, Accessibility, Interoperability, and Reusability) data principles, hindering the scientific community's ability to re-analyse data, integrate new studies, and support a robust, cumulative evidence base and reduce bias.

This work was performed in the context of a wider synthesis of MAs reporting on a broad range of sustainable agricultural practices, performed by the European Commission, Joint Research Centre. A first version of this map was already available within the "JRC-Farming Practices Evidence Library" (European Commission. Joint Research Centre., 2025; Schievano et al., 2024), based on a comprehensive dataset including a collection of around 570 MAs (version 2023) reporting on the effects of farming practices. However, the agroforestry data in the JRC dataset were collected in early 2020 and have become outdated due to the rapid advancement of research in this field. This paper aims at mapping the potential of agroforestry to deliver ecosystem services, as well as at identifying the knowledge and quality gaps. While drawing from global evidence, the review has a specific focus on agroforestry within temperate pedoclimatic zones. Specifically, the objectives were:

1. To update a map of the published MAs reporting on the effects of agroforestry at a global scale and specifically including articles regarding temperate pedo-climates;
2. To characterize how existing meta-analytical evidence is distributed across different agroforestry practices/systems and sustainability outcomes, as quantified by empirical metrics;
3. To link empirical metrics to the Common International Classification of Ecosystem Services version 5.2 and pinpointing knowledge gaps across the whole range of ecosystem services (CICES).
4. To assess the methodological quality of the included MAs and examining potential trends in quality over time or across sustainability outcome types.
5. To recompile the underlying primary studies references, determine the extent to which they overlap between different MAs and deliver a unique list of references for future further literature analyses, including a map the geographical location of the experiments.

By addressing these objectives, this study aims at providing policymakers, researchers, and practitioners with robust, synthesized, and current evidence based on agroforestry knowledge in temperate pedoclimates. This information is intended to highlight areas of strong consensus, identify critical research gaps, and ultimately support the development of effective agroforestry for sustainable land management, biodiversity conservation, climate change adaptation and mitigation, and food security.

2. Methods

We build on previous evidence syntheses on agroforestry based on published meta-analyses (MAs) (Beillouin et al., 2019; Castle et al., 2022; Köthke et al., 2022; Mathieu et al., 2025; Tamburini et al., 2020; Terasaki Hart et al., 2023) and the current version of the "JRC-Farming Practices Evidence Library" (European Commission. Joint Research Centre., 2025; Schievano et al., 2024).

2.1. Search strategy and inclusion/exclusion criteria

Our systematic review of MAs is performed in accordance with a methodological framework developed in the context of the JRC Evidence Library (Schievano et al., 2025). In brief, we followed the PRISMA statement guidelines (Page et al., 2021) and the Cochrane Handbook, to comprehensively identify MAs on agroforestry published before June 2024 in Web of Science and Scopus. We developed specific search strings, combining keywords related to agroforestry practices and meta-analytical methods (Supplementary Table 1). To update the first search run in 2020 (European Commission. Joint Research Centre., 2025), we performed three successive searches between 2023 and 2024 with refined keywords, (Fig. 1). We also incorporated MAs identified by previous systematic maps (Beillouin et al., 2019; Castle et al., 2022; Köthke et al., 2022; Tamburini et al., 2020). Four reviewers independently screened titles and abstracts, followed by full-text assessment based on predefined inclusion criteria (a complete list of the selection (exclusion/inclusion) criteria is reported in Supplementary Table 2). We included peer-reviewed MAs that reported quantitative results on the effects of agroforestry compared to tree-less agriculture or, in few cases, silvopastoral systems compared to monocultural timber plantations (Feliciano et al., 2018; Pent, 2020). During the data extraction process, we included MAs that provide evidence on agroforestry in temperate regions, including specific pedo-climates and biomes, such as temperate, boreal, continental, and Mediterranean. We excluded MAs focusing solely on either tropical or arid/semi-arid pedo-climates from the final data extraction (typically these are datasets that include only agroforestry studies performed in tropical or subtropical arid zones). As some global MAs provide data from temperate and other pedo-climates, we mapped the following in the present manuscript: 1) main results (global MAs including at least some temperate studies) and 2) results (e.g. of subgroup analysis) specifically including temperate studies.

2.2. Data extraction and classification

Data extraction was conducted using a standardized spreadsheet, developed in the context of the JRC-Farming Practices Evidence Library dataset (Schievano et al., 2025, 2024), to capture study characteristics, methodological details, quality and reported outcomes. In accordance with the methodology used in the JRC-Farming practices Evidence Library (European Commission. Joint Research Centre., 2025), the effect sizes were classified as significantly positive, significantly negative, non-significant, or lacking formal statistical testing. Extracted metadata included systematic review methods, characteristics of original experiments (intervention, comparator, outcome metrics, population variables), key results and conclusions, and quantitative effect size data (means, confidence intervals, sample sizes, effect size type, and statistical models).

Several MAs had global coverage, i.e. experimental sites were spread across many different pedo-climatic conditions and geographical locations. Other MAs were more specifically focused on geographical areas or pedo-climatic zones (however, according to our selection criteria, at least including temperate pedo-climates). We extracted the main assessments (i.e. mean effect sizes), as calculated by the authors of each MA, i.e. reflecting the overall original population (sometimes grand means pooling together different types of agroforestry practices and systems). We also extracted effect sizes for subgroups defined by specific agroforestry practice types, pedo-climatic conditions, or geographic zones.

To facilitate synthesis across studies, we classified intervention-comparator pairs and outcome metrics into harmonized categories: the main classes of agroforestry were categorised according to the JRC classification of farming practices (Angileri et al., 2024) and the EURAF Agroforestry Typology (Worms and Lawson, 2024), as follows: i) silvopastoral systems, ii) silvoarable systems, iii) agrosilvopastoral systems, iv) landscape woody features and v) others. Each class may include several specific practices. For instance, "Landscape woody features" would include e.g. hedgerows, buffer strips. We report an updated classification in Table 1.

In total, 80 empirical metrics (e.g. soil organic carbon stock, soil nitrogen concentration, soil sediment loss, forage yield, etc.) were grouped into 16 "impact categories" (e.g. increase carbon sequestration, increase soil nutrients, decrease soil erosion, increase crop yield, respectively), which contribute to 6 independent "sustainability outcomes" (e.g. carbon sequestration, soil health, agricultural productivity) (see supplementary table 4 for a full list and assortment of sustainability outcomes, impact categories and metrics). This classification of outcomes is currently adopted by the JRC-Farming practices Evidence Library (European Commission. Joint Research Centre., 2025) and matches the main thematic areas published in the Food System Sustainability Model (European Commission. Joint Research Centre., 2024), developed for the EU Food System Monitoring Framework (European Commission. Joint Research Centre., 2024).

Subsequently, we matched the empirical metrics retrieved from the selected MAs (as defined by the authors of each meta-analysis) to the most recent version of the ecosystem services classification (CICES 5.2). We used the most disaggregated category (CICES classes of four digits) to match the metrics (the full list of metrics matched to the CICES classes can be found in Supplementary Table 4).

2.3. Quality assessment and primary literature overlap

We evaluated the quality of included MAs using 16 criteria (Supplementary Table 5), that cover aspects of the systematic review process, statistical analysis, and risk of bias, in accordance with the JRC-farming practices dataset (Schievano et al., 2024). For MAs reporting metrics belonging to multiple impact categories, we assessed separately the quality per each impact category. The total number of criteria met indicated the quality score following the standardized quality assessment method by Schievano et al. (2026).

To evaluate the overlap of studies included in MAs, we extracted all primary-study references from the included MAs. Missing bibliographic metadata were retrieved using a custom R-based pipeline that accessed the Crossref metadata database via the rcrossref package (Chamberlain et al., 2025). For entries with missing DOIs but known titles, we implemented a looped query system that matched candidate records using a Jaro-Winkler string similarity algorithm (threshold ≥0.95). The threshold of ≥0.95 was selected to prioritize precision over recall, minimizing the risk of incorrectly merging records from distinct publications; entries below this threshold were either manually reviewed or retained as unmatched. For references lacking both title and DOI, full citation strings (APA, Chicago, or AYJ styles) were parsed using regular expressions to extract key metadata fields such as author, year, journal, volume, and pagination. These were then used to reconstruct missing information through targeted queries to Crossref. All records were subjected to internal consistency checks and harmonized into a standardized format. We then quantified the overlap of primary studies across MAs and tracked the accumulation of unique studies over time.

GPS coordinates of experimental sites were systematically extracted from the methods sections of primary studies using a combination of automated text mining and pattern-matching algorithms designed to recognize diverse coordinate formats, including decimal degrees, degrees-minutes-seconds, and other common notations. Cases where no precise coordinates could be determined were resolved through the manual annotation interface; when locality information permitted only country-level attribution, records were assigned to country centroids. To ensure data completeness and accuracy, missing or ambiguous locations were further curated manually through an interactive Shiny-based annotation interface, leveraging contextual information from titles, abstracts, and external databases (see Supplementary materials). This multi-tiered approach allowed for comprehensive harmonization and high-confidence georeferencing across the assembled dataset.

2.4. Statistical tests

We applied a Bayesian ordinal regression framework to quantify and compare the probabilities of observing Negative, Neutral, and Positive effects across the "Temperate-only" and "Full" datasets. Models were implemented in R using the brms package (Bürkner, 2017), with effect direction treated as an ordered factor (Negative < Neutral < Positive). Two model families were used depending on the number of distinct effect categories observed within each CICES group.

For groups with three observed categories, a Bayesian cumulative logit model was fitted:

$$\mathrm{logit}\left[P\left(Y_{ij} \le k\right)\right] = \tau_k - \left(\beta \times x_i + u_i\right),\ k = 1, 2$$

where Y_{ij} is the ordered effect for observation j in paper i, taking values 1 (Negative), 2 (Neutral), and 3 (Positive); τ_k are threshold parameters; k = 1, 2 indexes the two cumulative thresholds separating the three ordered categories; x_i is the number of primary studies included in the meta-analysis (continuous covariate); β is the regression coefficient associated with x_i, a positive value indicating a shift towards more positive effect categories; and $u_i \sim \mathrm{Normal}(0, \sigma^2)$ is a paper-level random intercept, identified by DOI. Thus, P(Negative) = $P(Y_{ij} \le 1)$, P (Neutral) = $P(Y_{ij} \le 2) - P(Y_{ij} \le 1)$, and P(Positive) = $1 - P(Y_{ij} \le 2)$. For groups with only two observed categories (Positive and Negative), a Bayesian binary logistic model was fitted using the same covariate x_i and paper-level random intercept structure. Groups with fewer than 10 observations were excluded.

A weakly informative prior Normal (0,10) was placed on β; default brms priors were used for threshold parameters and σ. All models were run with 4 chains, 2000 iterations (500 warmup), and adapt_delta = 0.95. The key quantity of interest — P(Positive) − P(Negative) — was computed from posterior predicted probabilities at the median number of primary studies, marginalizing over random effects. Differences between the Full and Temperate-only datasets were estimated by fitting models to each dataset separately and computing draw-level differences in this quantity across groups. Results are reported as posterior mean, 95% credible interval, and the posterior probability that P(Positive) > P (Negative), and visualized with density plots (see Supplementary materials).

Table 1
- Classification of agroforestry systems and practices used in this study, as adapted from the JRC classification of farming practices (Angileri et al., 2024) and the EURAF Agroforestry Typology (Worms and Lawson, 2024).

Main agroforestry systems	Specific agroforestry practices
Silvoarable	Alley cropping
	Alley coppice
Silvopastoral	Wood pasture & Orchard grazing
Landscape woody features	Hedges/wooded strips
	Buffer strips
	Trees in line (Windbreaks and shelterbelts)
	Isolated trees
Agrosilvopastoral	Settlement agroforestry (Homegarden)
Other Systems	Shaded perennials
	Improved fallow
	Woodlot

3. Results

A total of 340 records were identified, and 42 meta-analyses (MAs) meeting our inclusion criteria were selected (Fig. 1), published between 2007 and June 2024. Of these, 29 MAs report results at a global scale (including results on temperate systems) and 13 MAs focus specifically on agroforestry systems in temperate pedo-climates. A total of 32 MAs were identified that focus specifically on agroforestry studies located in tropical/subtropical or arid/semi-arid regions, which did not undergo data extraction (Supplementary Table 3). The full PRISMA statement diagram (Page et al., 2021) on the review workflow is reported in Supplementary Fig. 1 and the full critical appraisal in Supplementary Table 3.

3.1. Overview of available assessments

Fig. 1 provides a synthetic overview of the main effects of agroforestry, as reported by MAs across the Groups of Ecosystem services (3-digits of the CICES classification). In parallel, we report the distribution of results across empirical metrics (see Supplementary Fig. 2), as named in the original MAs and as classified in the JRC-Farming practices Evidence Library (European Commission. Joint Research Centre., 2025).

Both the main results (i.e. mean effect sizes estimated using the full dataset of each MA) and the temperate-only results (i.e. mean effect sizes estimated using only subgroups of experiments located in temperate pedo-climates) cover seven classes of ES. By far, the majority of assessments (i.e. estimated effect sizes) were published on “Regulation of soil quality” (main: $n = 164$; temperate: $n = 152$). The large majority of these assessments were reported for the empirical metrics (see Supplementary Table 4): Organic carbon stock (soil), Organic carbon sequestration rate (soil), Organic carbon content (soil); fewer results were available for other metrics such as Soil water, Soil phosphorous/nitrogen and Soil biological quality (e.g. Taxonomic parameters (microbes)). High shares of these assessments (i.e. around 65% for both temperate-only and main results) reported significant positive effects.

The ES group “Atmospheric composition and conditions” (main: n = 31; temperate: n = 15) showed a moderate share of positives (35–50%), while many results lacked statistical analysis (Fig. 1). These results were mainly assessed through metrics related to organic carbon sequestration in trees biomass and to soil GHG emissions (Supplementary Table 6). A small share (5%) of negative effects were reported for specific metrics related to soil GHG emissions (such as N2O and CH4, see Supplementary Fig. 2).

Over 65% of positive effects (and zero negative effects) were reported for ES groups of “Erosion control” (main: $n = 18$; temperate: $n = 5$), “Pest and disease control” (main: $n = 12$; temperate: $n = 6$), “Lifecycle maintenance” (main: $n = 16$; temperate: $n = 5$) and “water conditions” (main: $n = 17$; temperate: $n = 3$).

For the ES group “Cultivated terrestrial plants for nutrition, materials or energy” (main: n = 22; temperate: n = 21), around 50% of negative effects were reported, with lower shares of positive effects (8% and 25% for main and temperate-only results, respectively). These results were mainly related to the metrics “Crop yield (arable crops)” and “Net primary productivity (grass)”, which hardly reflect the overall productivity of an agroforestry system, rather than of fractions of land dedicated to either single arable crops or pasture. Very few assessments reported results for more holistic assessments of the agricultural productivity of the whole agroforestry system (e.g. “Land equivalent ratio (LER)”, which includes the assessment of the intercropped trees and arable crops or livestock production.

We quantified the likelihood that interventions produce beneficial versus detrimental outcomes across ecosystem services. Using the main results (global dataset), interventions on Atmospheric composition and conditions were somewhat more likely to be beneficial than harmful, with a mean probability difference of 0.21 (95% CI –0.53 to 0.83, posterior probability 0.77), indicating a 21-percentage-point higher chance of positive effects. For Cultivated terrestrial plants for nutrition, materials or energy, the global dataset slightly favored negative outcomes (mean – 0.37, 95% CI –0.82 to 0.27, posterior probability 0.094). Metrics, however, were mainly focusing on specific crops, rather than the whole-system's productivity. In contrast, interventions targeting Regulation of soil quality were strongly likely to be beneficial, with a 69-percentage-point higher probability of positive effects (95% CI 0.54–0.83, posterior probability 1).

Restricting the analysis to the temperate subset yielded broadly consistent patterns. Cultivated terrestrial plants showed no detectable difference between positive and negative outcomes (mean 0.016, 95% CI –0.34 to 0.41, posterior probability 0.53), whereas Regulation of soil quality interventions remained strongly positive (mean 0.48, 95% CI –0.03 to 0.83, posterior probability 0.97). Comparisons between the global and temperate datasets revealed only minor differences, indicating that focusing on temperate regions does not substantially alter predicted effect probabilities. Overall, soil quality interventions consistently show a high likelihood of positive outcomes, while results for other services are more variable and context-dependent.

3.2. Available assessments on classes of agroforestry

While most meta-analyses (MAs) presented aggregated results, pooling effect sizes from various agroforestry systems (referred to as “unspecified agroforestry types”), Figs. 2 and 3 provide visual summaries of the evidence base disaggregated by agroforestry class (following Table 1) and CICES Ecosystem Service (ES) Groups. “Unspecified agroforestry type” is the most frequently represented category overall, particularly within the temperate climate results (Fig. 3), covering approximately one third of the findings. Silvoarable and silvopastoral systems are also prominent, with a high number of assessments in both global (Fig. 2) and temperate-specific contexts (Fig. 3). Conversely, “Landscape woody features” consistently show fewer results across both climate groups.

A consistent pattern across all agroforestry classes is a positive effect on the ES Group “Regulation of soil quality.” For instance, silvopastoral systems contribute positively to “Soil nutrients” and “Soil physico-chemical quality” (Fig. 2). Fig. 2 highlights a concentration of meta-analytical evidence on the ES Group “Atmospheric composition and conditions,” primarily driven by carbon sequestration, which stands out with the highest number of synthesized results across nearly all agroforestry types, particularly for silvopastoral and silvoarable systems. Other ES Groups receiving significant attention include aspects of “Regulation of soil quality,” notably “Soil nutrients” and “Soil physico-chemical quality,” as well as “Lifecycle maintenance” (biodiversity) and “Cultivated terrestrial plants for nutrition, materials or energy” (specifically grassland production under silvopastoral systems). Conversely, ES Groups like “Atmospheric composition and conditions” (direct GHG emissions and biomass carbon stocks, distinct from soil carbon stocks), “Pollination,” and overall farming system productivity appear less frequently assessed.

Regarding the nature of the findings (Fig. 2), “Atmospheric composition and conditions” predominantly shows positive effects across agroforestry types. However, for silvopastoral systems, a substantial proportion of these results are non-significant, and a considerable portion of carbon sequestration results lack formal statistical testing. Similarly, positive impacts are frequently reported for “Soil nutrients,” “Soil physico-chemical quality,” “Soil biological quality,” and “Lifecycle maintenance,” suggesting a general consensus in the MAs on benefits in these areas.

However, outcomes related to “Cultivated terrestrial plants for nutrition, materials or energy” show more variability. For instance, grassland production under silvopastoral systems yields a mix of significantly positive, negative, and non-significant results, while crop yield under silvoarable systems includes both significantly negative and non-significant findings. Furthermore, the prevalence of non-significant

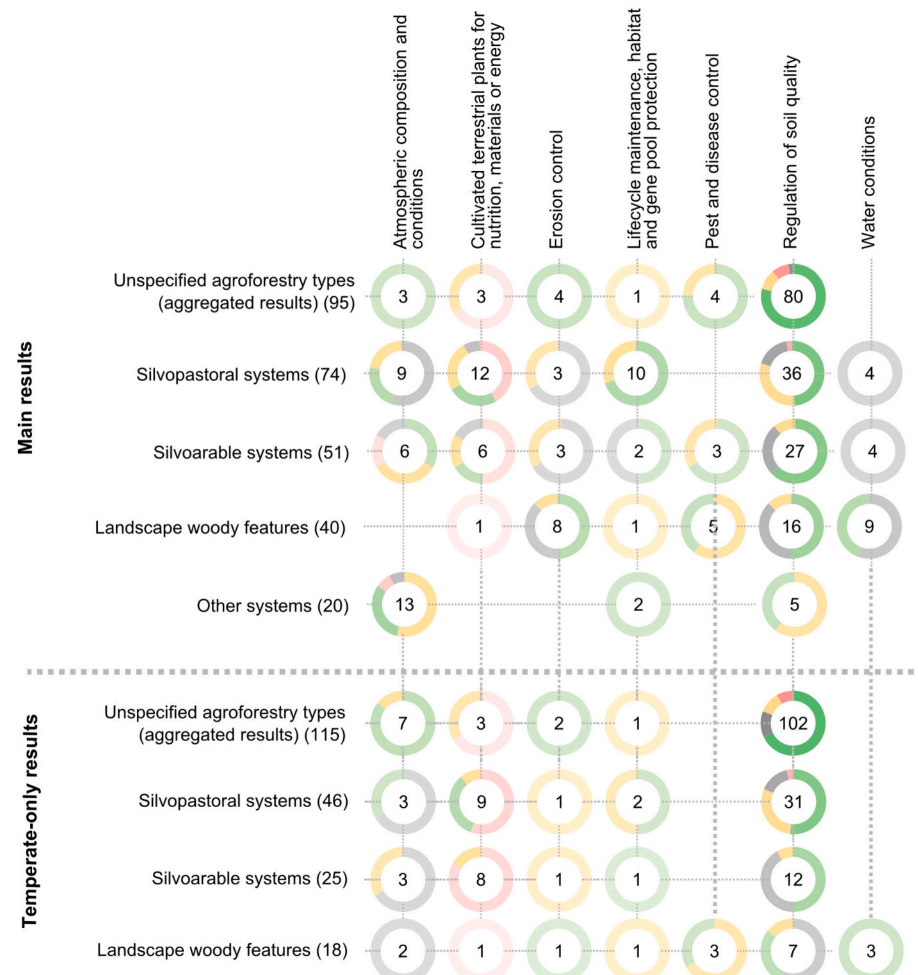

Fig. 2. – Meta-analytical evidence regarding the effects of main classes of agroforestry systems, reporting the main effect sizes estimated by the selected meta-analyses. Results (i.e. mean effect sizes extracted from the select MAs) were extracted for main assessments, as well as for temperate-only subgroups. The effects are classified for Ecosystems services (ES) groups (i.e. 3-digits of the CICES classification). The chart illustrates the distribution of results showing significant positive (green) or negative effects (red), non-significant effect (yellow) or non-statistically-tested results (grey), for the selected combination of practices and ES. The numbers represent the count of effect sizes reported by the selected MAs. When the same MA reports results at different aggregation levels, we did not double count these results in the same donut graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

findings in ES Groups like "Erosion control" and "Pest and disease control," alongside the significant proportion of results lacking statistical testing across multiple ES Groups (e.g., "Global warming potential," "Nutrient leaching"), underscores the need for cautious interpretation and highlights areas where the synthesized evidence remains inconclusive or requires more rigorous assessment.

3.3. Knowledge Gaps across CICES classification

Results reported by MAs were associated with the CICES classifications, namely with 2 (out of 6) sections and 3 (out of 19) divisions (represented graphically in Fig. 4) and with 8 (out of 38) groups and 10 (out of 99) classes (see detailed report in Supplementary Table 6). This means that 66% of sections, 84% of divisions, 85% of groups and 91% of classes are not covered by any result reported in MAs. While many "Abiotic/Geophysical" services may not be directly affected by the implementation of agroforestry systems, "Biotic/Biophysical" services are also relatively uncovered, with only 4% coverage rate for "Provisioning" services, 32% for "Regulation & maintenance" services and no coverage at all for "Cultural" services.

The division with the highest coverage rate (57%) is "2.3 - Regulation of physical, chemical, biological conditions", with 8 out of 14 classes covered by evidence (Fig. 4). However, most results concern soil-related services, while very important classes remain either uncovered (e.g. for "Control of wind erosion rates", "Regulation runoff and base flows", "Seed dispersal", "Maintaining or regulating refuge habitats", "Regulation of temperature and humidity, including ventilation and transpiration at local scales", "Wind protection", "Fire protection") or underrepresented (e.g. "water conditions", see Figs. 1,2,3). This highlights a particular research void concerning agroforestry's potential key role in enhancing biodiversity, improving water quality/management and hydrological cycles and driving resilience, which are especially pertinent given increasing concerns about biodiversity losses, water quality, fire risks and extreme weather events.

Provisioning services remain relatively uncovered by evidence in many relevant classes (4% coverage rate for "biotic" services, Fig. 4), such as "Fibres and other materials from cultivated plants, fungi, algae and bacteria for direct use or processing (excluding genetic materials)" (1.1.1.2), "Animals reared for nutritional purposes" (1.1.3.1). The available results within provisioning services are narrowly focused on crop-yield or grassland/forage-productivity metrics, typically at plot or field scales for a limited number of main crops (see supplementary Figures), neglecting the broader spectrum of products and services that agroforestry systems may provide (e.g. timber, fruit, essential oils, etc.).

Furthermore, cultural services (both Biotic and Abiotic) have been entirely overlooked in published meta-analyses on agroforestry. This

Fig. 3. –Meta-analytical evidence regarding the effects of specific classes of agroforestry systems. Results (i.e. mean effect sizes extracted from the select MAs) were extracted for main assessments, as well as for temperate-only subgroups. The effects are classified for Ecosystems services (ES) groups (i.e. 3-digits of the CICES classification). The chart illustrates the distribution of results showing significant positive (green) or negative effects (red), non-significant effect (yellow) or non-statistically-tested results (grey), for the selected combination of practices and sustainability outcomes. The numbers represent the count of effect sizes reported by the selected MAs. When the same MA reports results at different aggregation levels, we did not double count these results in the same donut graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

represents a critical void, as agroforestry systems often hold significant social, aesthetic, recreational, and spiritual value, which are crucial for holistic sustainability assessments and important for enhancing the level of uptake and economic success of agroforestry systems. The absence of evidence in this domain underscores a need for research methodologies that can effectively capture these less tangible, yet profoundly important, benefits.

3.4. Quality assessment of the meta-analyses

The methodological quality of MAs was assessed for 74 combinations of practice-impact (Fig. 5A). Our evaluation of the quality revealed that the majority of papers reported on the procedure of data search and selection and on funding sources. While statistical methods are described in most of the assessments, a correct reporting of MA-statistics, such as reporting and application of individual effect sizes, weighting of studies and assessment of the heterogeneity of results is neglected in around half of the studies (Fig. 5A). Only 40 assessments reported consequently the details of the selection process according to the PRISMA statement. Importantly, substantial progress should be made in the communication of a retrievable list of included primary studies (e.g. including DOIs), as around 1 out of three MAs do not

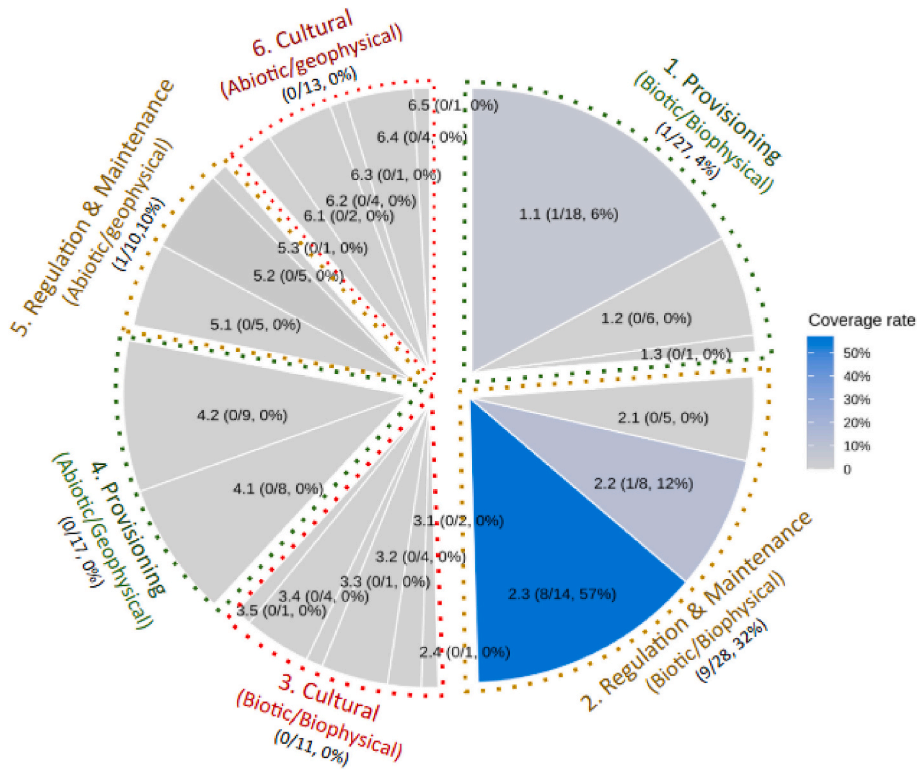

Fig. 4. – Graphical glance of how current evidence reported in meta-analyses cover the full list of CICES classes (4-digits code, e.g. 1.1.1.1 - Cultivated terrestrial plants grown for nutritional purposes). Sectors of the pie-chart represent CICES Divisions (indicated with the corresponding 2-digits code, e.g. 1.1 - Biomass), and grouped by dotted lines into CICES Sections (1-digit code, e.g. 1 - Provisioning (biotic/biophysical). For each Section or Division, the figures between brackets report the coverage rate (as ratio and as percentage and represented with blue color scale) of classes covered by evidence. Knowledge gaps across the CICES classification are shown as pie-chart sectors in grey color. Supplementary Table 6 reports the list of specific CICES sections, divisions, groups and classes covered by results of meta-analyses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

publish a consistent list. Even more, the primary dataset is not available for nearly 2 out of 3 MAs. The assessment of the publication bias and the reporting of individual effect sizes were the least fulfilled criteria (with 30 and 22 out of 74 assessments).

Fig. 5B shows that only 6 of the considered assessments (8%) fulfilled all 16 quality criteria, and around 80% met at least 8 criteria. All assessments fulfill at least 4 out of 16 criteria, while about 93% fulfill at least 7 criteria. The curve shows a steep decline after meeting a quality threshold of 10 out of 16 criteria. These visualizations reveal that while most MAs meet basic reporting standards, there is significant room for improvement in more rigorous criteria such as individual effect sizes, dataset availability, and publication bias analysis.

Between 2015 and 2024, the number of MAs reporting on agroforestry increased consistently (Fig. 5C). This increase over time has, however, been accompanied by a non-homogeneous increase in quality level (Fig. 5C). While the average quality of the MAs increased over time, some MAs with lower quality (less than 50% of criteria met) were also published in the most recent years.

No statistically significant differences in quality scores were detected across impact categories by the application of ANOVA and Kruskal-Wallis test ($p > 0.05$) (Supplementary Fig. 6 A). It should be noted, though, that the number of assessments per impact category is rather low, leading to relatively high standard deviations.

3.5. Primary literature map

In total, we collected 2066 references, of which 161 could not be processed (lack of sufficient information, such as DOI or title) and 1905 were retrieved using our workflow querying the Crossref database (https://www.crossref.org). Of these, 1495 were unique primary papers, synthesized by the 42 selected MAs. The full list of references is available in the JRC - Farming practices data collection (European Commission - Joint Research Center, 2025).

Fig. 6 A shows the overlaps of primary literature papers across the 42 selected MAs. Most primary literature appears in only one meta-analysis (over 1200 of the 1495 primary studies, and over time new primary studies are continuously added to the evidence base. Particular overlap in synthesis occurred in the years 2019–2021 with several MAs published with limited increase of unique primary studies added to the cumulative curve (Fig. 6 A).

The extraction of geographical coordinates (from direct mention in the original primary paper text or inferred from reported place names such as cities or villages) yielded 625 out of 1495 studies (Fig. 6B). Interestingly, 435 were located in temperate pedo-climatic zones and 190 were located in other pedo-climates. The distribution of temperate studies is mostly concentrated in North America (USA mainly), Europe and China. Few temperate studies are located in the southern hemisphere Countries (e.g. Chile, Australia, New Zealand).

4. Discussion

This synthesis aimed to provide a comprehensive overview of the current state of meta-analytical evidence on the potential of agroforestry to enhance ecosystem services, identifying well-researched areas as well as critical knowledge gaps requiring further investigation, with a specific focus on temperate pedo-climates. Our findings offer a robust evidence map for policymakers and researchers, guiding future efforts to strengthen the understanding and implementation of agroforestry. Our analysis reveals that temperate agroforestry systems demonstrate significant potential to enhance several key ecosystem services, particularly for soil health and carbon sequestration.

However, the consistency of positive effects is not universal.

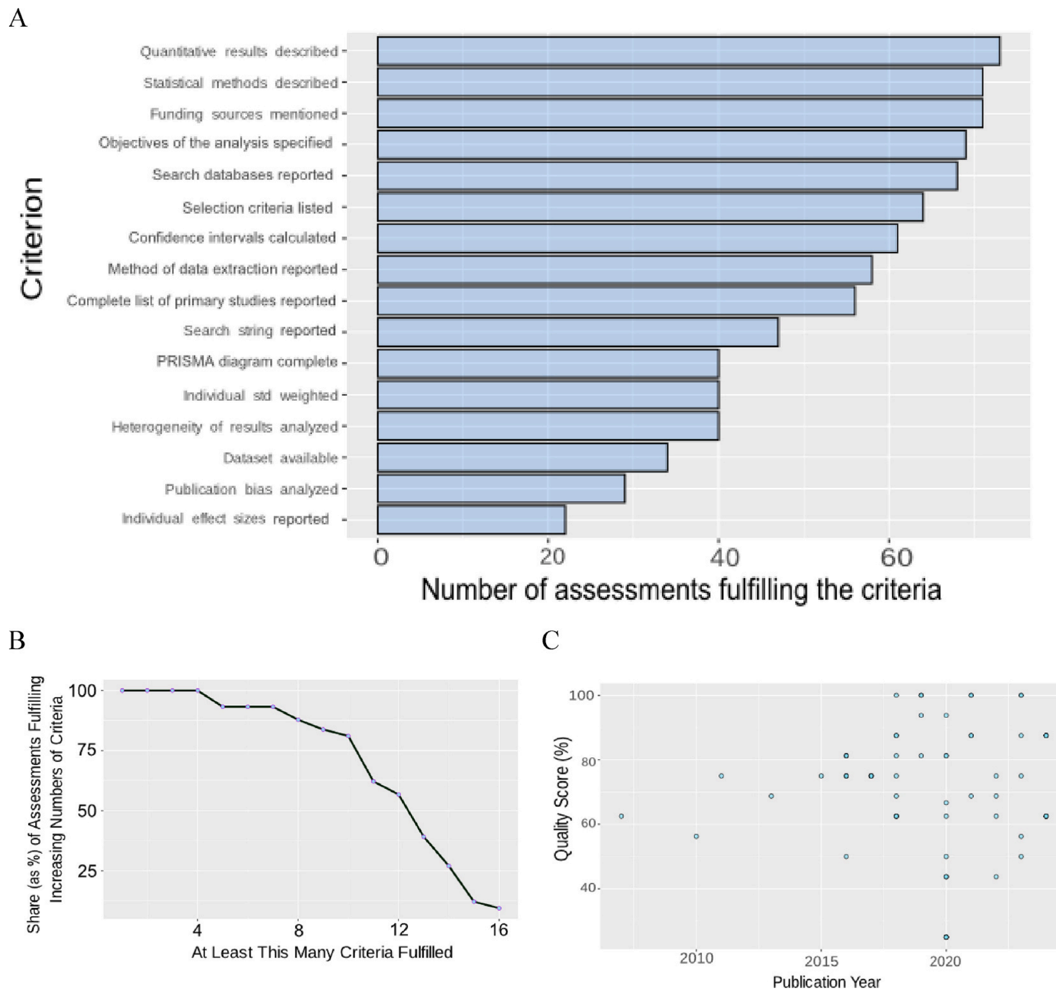

Fig. 5. – Distribution of quality assessments performed on each selected meta-analysis reporting on one or more impact categories (n = 74). A) Number of assessments fulfilling each of the 16 quality criteria; B) Share of assessments fulfilling an increasing number of criteria; and C) distribution of quality scores (% of fulfilled criteria) over time (publication year).

Furthermore, a considerable portion of findings across various ES or impact categories, especially for temperate-only assessments, were either non-significant or lacked formal statistical testing (grey segments in Fig. 1, 2, and 3). This indicates that while many studies exist, their conclusions are not always robustly established, or the effects are highly context-dependent. For the smallest subgroups ($n \leq 5$), results should be treated as exploratory and interpreted alongside the wide credible intervals reported.

Provisioning services through agricultural productivity outcomes exhibit relatively high variability. Grassland/forage production under silvopastoral systems, for instance, shows a mix of significantly positive, negative, and non-significant results, while single-crops yield under silvoarable systems includes both significantly negative and non-significant findings reported in the MAs (Fig. 2). Importantly, the indicators used to quantify provisioning services linked to agricultural production are often limited to crop or grass yields, overlooking the broader system productivity that includes shrubs/trees-products yields (e.g. tree-forages, fruits, berries, timber, biomass, etc.) or overall Land Equivalent Ratio (LER) (Dupraz et al., 2018; Pent, 2020). Many MAs do not explicitly define how agroforestry system boundaries are considered for productivity metrics (e.g., yield per hectare of cropped area vs. total agroforestry area), hindering consistent interpretation of "negative" impacts on specific crop yields. This highlights a significant knowledge gap in assessing whole agroforestry system productivity in the existing meta-analytical literature.

Agroforestry systems are distinguished by their multidimensional character, incorporating diverse components such as trees, crops, and animals. However, prevailing research methodologies frequently focus on one-dimensional effects, thereby overlooking the potential for comprehensive, multidimensional impact assessment. A key question is whether existing research methodologies can adequately capture the complexity, diversity and long-term nature of agricultural systems. This issue was also raised by Douxchamps et al. (2017). The authors reviewed models focusing on monitoring and evaluating the climate resilience of agriculture, and found that the models failed to assess 'stability', 'transformation' and the topic's overall complexity.

Are existing meta-analyses methodologically sound?

Our assessment of methodological quality across 74 combinations of practice-impact revealed mixed results, with significant room for improvement, echoing findings by Beillouin et al. (2018), Fohrafellner et al. (2023) and Schievano et al. (2026). While most MAs adequately reported on data search and selection procedures and funding sources, crucial aspects of rigorous systematic review were often neglected. Only 40 out of 74 assessments consistently reported selection details according to PRISMA statement guidelines (Page et al., 2021).

A major concern is the reporting and application of MA statistics, including the assessment of heterogeneity and robust publication bias analyses. Furthermore, accessibility and transparency of underlying data remain a challenge. Approximately one-third of MAs failed to provide a retrievable list of included primary studies (e.g., with DOIs),

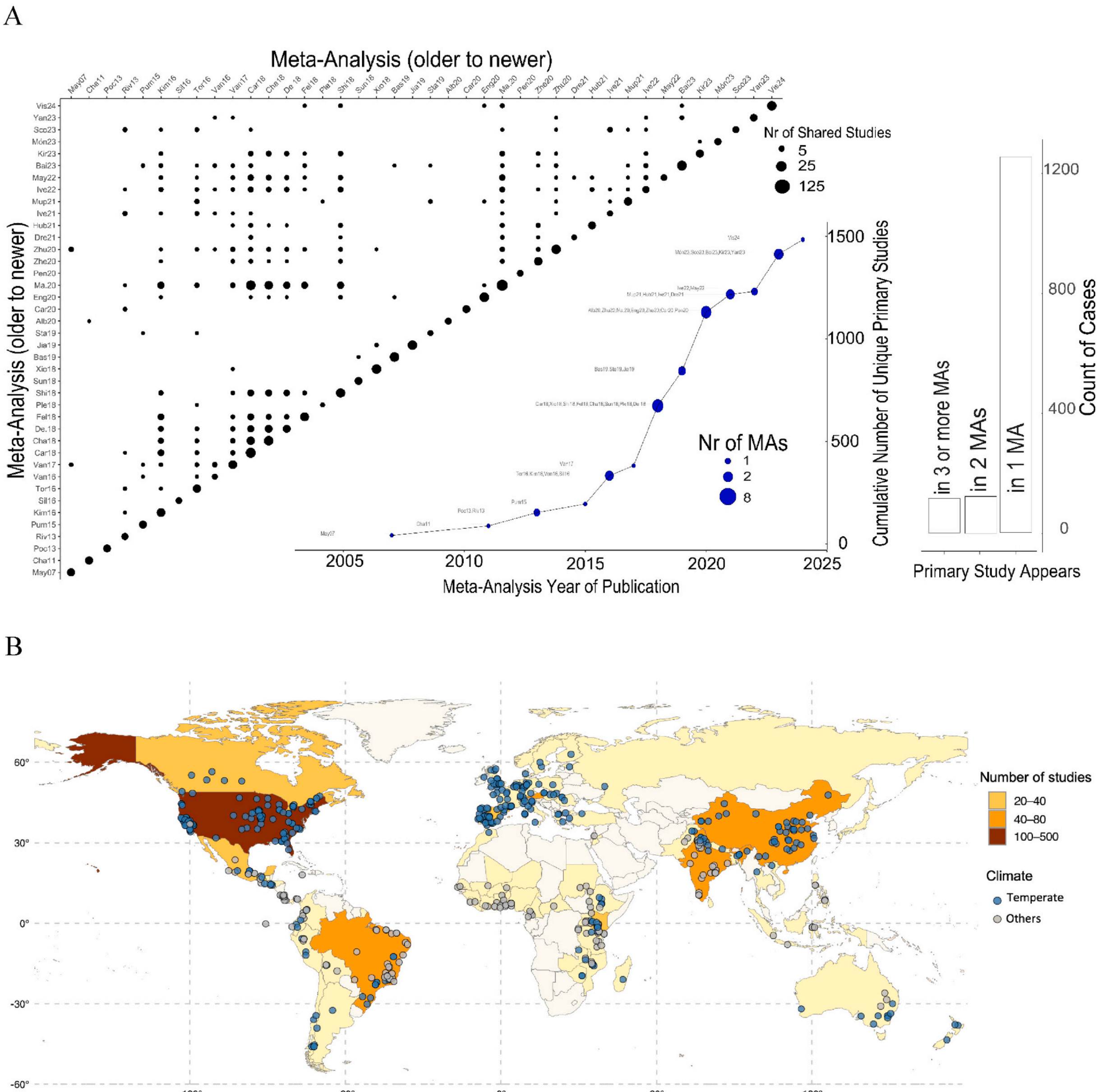

Fig 6. - Primary literature studies map (whole dataset obtained from the selected 42 meta-analyses, including temperate and non-temperate studies). (A) Accumulation over time and overlap of primary studies across the 42 meta-analyses; The size of the dots represent the number of shared primary literature papers, i.e. in common for each pair of MAs. (B) Global map showing GPS coordinates extracted from primary studies. Points on the map are marked using different colours and counted distinguishing between studies performed in temperate and non-temperate pedo-climatic conditions. Studies are counted per Country (represented with color scale). Map lines delineate study areas and do not necessarily depict accepted national boundaries.

and nearly two-thirds did not make their primary datasets available. This lack of transparency contravenes FAIR data principles and compromises the replicability and reusability of these valuable syntheses (Fig. 5A).

While the number of MAs on agroforestry has consistently increased between 2015 and 2024 (Fig. 1), the improvement in quality has been uneven. The average quality has risen over time, but some MAs with very low quality (less than 50% of criteria met) were still published in recent years (Fig. 5C). This suggests a need for stricter adherence to methodological standards in evidence synthesis, such as the establishment of a standardized quality assessment protocol, as also highlighted by Schievano et al. (2024), to ensure robust and credible conclusions that can effectively inform policy and practice.

What critical knowledge gaps limit our understanding of agroforestry in temperate regions?

Our evidence mapping exercise revealed several significant and pressing knowledge gaps that limit a comprehensive understanding of agroforestry's potential in temperate regions. First, it should be noted that some primary studies georeferenced within temperate pedo-climatic zones are located at high-altitude tropical sites (e.g. southern Brazil, Madagascar). Although temperature regimes at such sites may resemble temperate conditions, differences in seasonality, soil characteristics, and species assemblages mean that results from these locations should be interpreted with caution when generalizing to continental temperate agroforestry contexts.

Certain types of agroforestry systems (especially the most diversified) are underrepresented. While silvopastoral systems (e.g. orchard grazing) and silvoarable systems (e.g. alley cropping) are well-studied (Castle et al., 2022), there is a notable scarcity of meta-analytical evidence for agrosilvopastoral systems (e.g. successional agroforestry) and improved fallows. Hedge systems and other types of woody features also receive comparatively limited academic attention, despite their recognized potential (Fig. 2, Fig. 3). This imbalance restricts our ability to provide a holistic assessment of agroforestry's benefits across its diverse forms.

In addition, the dominance of the 'unspecified agroforestry type' category in the evidence base itself constitutes a finding: it indicates that existing MAs rarely provide disaggregated, system-specific conclusions, limiting the statistical power available for type-specific inference and underscoring the need for more targeted primary research and synthesis.

A critical void exists in the assessment of socio-cultural and cultural ecosystem services (ESS) provided by agroforestry. As highlighted in Fig. 4, cultural services (both Biotic and Abiotic) are entirely overlooked in the published meta-analyses. This aligns with findings by Köthke et al. (2022) and represents a significant limitation, as agroforestry systems often hold profound social, aesthetic, recreational, and spiritual value that are crucial for holistic sustainability assessments and public acceptance.

Despite the increasing concerns about water availability, quality, and extreme weather events, our study revealed a significant lack of meta-analytical information regarding the impact of agroforestry on water balances and hydrological cycles. The CICES classification clearly shows "water conditions" as underrepresented, and critical classes like "Control of wind erosion rates," "Regulation of runoff and base flows," and "Regulation of temperature and humidity" remain either uncovered or underrepresented (Fig. 4).

Overlap in primary literature and geographical coverage.

Our review found some overlap (Fig. 6 A) in the primary literature across different meta-analyses. The overlap was particularly evident for widely studied practices such as alley cropping. While new primary studies are constantly being added to upcoming meta-analyses, indicating an ongoing expansion of the evidence base, this overlap suggests a need for more diverse primary research to explore novel aspects of agroforestry impacts. A certain degree of overlap is unavoidable, and therefore needs to be assessed when comparing and combining different MAs. MAs focusing on narrower indicators, however, often show lower overlap, suggesting a broad and not yet sufficiently explored research field where individual MAs specialize in different aspects or regions. The geographical distribution of primary studies included in MAs reveals an uneven distribution (Fig. 6B). While our review specifically focused on MAs including temperate regions, the global map highlights that much of the research is still concentrated in North America (mainly USA), Europe, and China. This uneven distribution limits our understanding of agroforestry's potential across diverse temperate climatic conditions and management contexts, particularly in other temperate zones.

Implications for temperate pedo-climates.

The current evidence map, while comprehensive for meta-analytical syntheses, underscores the unique challenges and opportunities for agroforestry in temperate pedo-climates. The generally positive evidence for soil health and carbon sequestration provides a strong foundation for promoting agroforestry in these regions as a climate change mitigation and adaptation strategy. However, the identified knowledge gaps – particularly the underrepresentation of certain practices, the scarcity of long-term studies, and the neglect of socio-cultural and water-related impacts – are critical for optimizing agroforestry adoption and policy formulation in temperate zones. Policy recommendations must consider not just ecological benefits but also the economic viability (profitability studies are lacking for high-income countries, Castle et al., 2022) and social acceptance of these systems.

Recommendations for future research and policies.

This study's strength lies in its comprehensive review of 42 meta-analyses, following a rigorous protocol aligned with the PRISMA statement and Cochrane Handbook guidelines. It provides a detailed classification of outcomes against the CICES framework and includes a quality assessment of the MAs themselves. A key limitation of this study is that it relies exclusively on meta-analytical evidence. However, we provide here (see Supplementary materials) the list of references of all primary studies that were selected by these MAs. We strongly encourage the scientific community to engage in the reconstruction of the literature base, ideally using a common and standardized dataset. In line with FAIR principles, the authors of these 42 MAs should agree to share the primary data and make an effort in merging their structured datasets.

Future research should pursue two main objectives: first, filling the knowledge gaps on hedgerows, diversified agroforestry systems and silvopastoral systems; second, developing methodologies that can effectively encapsulate the multifaceted nature of agroforestry, thereby ensuring that its comprehensive value is duly recognized.

The overall positive picture of agroforestry emerging from this synthesis can support policy-makers and practitioners in promoting these systems. However, the knowledge gaps identified should encourage policymakers to develop new explicit support systems and thus motivate practitioners to promote more and new agroforestry systems. This will expand our knowledge and support a more resilient and sustainable agriculture.

CRediT authorship contribution statement

Andrea Schievano: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Camille Rubeaud:** Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Margret Köthke:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Beatrice Landoni:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Damien Beillouin:** Writing – review & editing, Visualization, Validation, Software, Formal analysis. **Marta Pérez-Soba:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Simona Bosco:** Writing – review & editing, Methodology. **Rui Catarino:** Writing – review & editing, Software, Methodology, Formal analysis. **Giovanni Tamburini:** Funding acquisition, Data curation. **David Makowski:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Sonja Kay:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Funding acquisition, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Google AI studio in order to re-phrase and improve readability of sentences. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility.

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.

Acknowledgement

The present paper is based on the JRC-Farming-Practices dataset (2023 version), which has been compiled in the frame of an Administrative Arrangement entitled "Integrated modelling platform for agro-economic commodity and policy analysis (IMAP4AGRI)" between DG Agriculture and Rural Development (DG AGRI) and DG Joint Research Centre (DG JRC), providing the funding of the work. David Makowski was partly funded by the project CLAND (French Research Agency, 16-ANR-0003). Giovanni Tamburini was partly funded by the Italian

Agritech National Research Center (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). Sonja Kay and Camille Rubeaud were supported by the DigitAF project (Grant Agreement N° 101059794), co-funded by the European Commission, European Research Agency, within the Horizon Europe programme, Cluster 6: "Food, bioeconomy, natural resources, agriculture and environment".

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoser.2026.101905.

Data availability

Original data are available at: https://data.jrc.ec.europa.eu/collection/id-00399

References

Angileri, V., Fernandez, I.G., Weiss, F., 2024. A classification scheme based on farming practices. Publications Office of the European Union. https://doi.org/10.2760/33560.

Beillouin, D., Ben-ari, T., Makoswki, D., 2018. Assessing the Quality and Results of meta-Analyses on Crop Diversification Protocol for Systematic Review and Evidence Map.

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

Beillouin, D., Ben-Ari, T., Malézieux, E., Seufert, V., Makowski, D., 2021. Positive but variable effects of crop diversification on biodiversity and ecosystem services. Glob. Chang. Biol. 27, 4697–4710. https://doi.org/10.1111/gcb.15747.

Bürkner, P.-C., 2017. Brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw. 80, 1–28. https://doi.org/10.18637/jss.v080.i01.

Castle, S.E., Miller, D.C., Merten, N., Ordonez, P.J., Baylis, K., 2022. Evidence for the impacts of agroforestry on ecosystem services and human well-being in high-income countries: a systematic map. Environ. Evid. 11, 1–27. https://doi.org/10.1186/s13750-022-00260-4.

Chamberlain, S., Zhu, H., Jahn, N., Boettiger, C., Ram, K., 2025. Rcrossref: client for various "CrossRef" "APIs". . R package version 1.2.009.

Douxchamps, S., Debevec, L., Giordano, M., Barron, J., 2017. Monitoring and evaluation of climate resilience for agricultural development – a review of currently available tools. World Dev. Perspect. 5, 10–23. https://doi.org/10.1016/j.wdp.2017.02.001.

Dupraz, C., Lawson, G.J., Lamersdorf, N., Papanastasis, V.P., Rosati, A., Ruiz-Mirazo, J., 2018. Temperate agroforestry: the European way. In: temperate agroforestry systems. CABI books, pp. 98–152. https://doi.org/10.1079/9781780644851.0098.

European Commission - Joint Research Center, 2025. JRC-Farming-Practices data collection – An evidence library of the effects of Farming Practices on the environment and the climate.

European Commission. Joint Research Centre, 2024. EU Food System Monitoring Framework: From Concepts to Indicators. Publications Office, LU.

European Commission. Joint Research Centre, 2025. The JRC Farming Practices Evidence Library: This Library Synthesizes a Large Amount of Scientific Evidence to Assess the Effects of Farming Practices on Sustainability Outcomes, Mainly Regarding the Environment, the Climate, and Agricultural Productivity. Publications Office, LU.

Fagerholm, N., Torralba, M., Burgess, P.J., Plieninger, T., 2016. A systematic map of ecosystem services assessments around European agroforestry. Ecol. Indic. 62, 47–65. https://doi.org/10.1016/j.ecolind.2015.11.016.

Feliciano, D., Ledo, A., Hillier, J., Nayak, D.R., 2018. Which agroforestry options give the greatest soil and above ground carbon benefits in different world regions? Agric. Ecosyst. Environ. 254, 117–129. https://doi.org/10.1016/j.agee.2017.11.032.

Fohrafellner, J., Zechmeister-Boltenstern, S., Murugan, R., Valkama, E., 2023. Quality assessment of meta-analyses on soil organic carbon. SOIL 9, 117–140. https://doi.org/10.5194/soil-9-117-2023.

Kim, D.-G., Kirschbaum, M.U.F., Beedy, T.L., 2016. Carbon sequestration and net emissions of CH4 and N2O under agroforestry: synthesizing available data and suggestions for future studies. Agric. Ecosyst. Environ. 226, 65–78. https://doi.org/10.1016/j.agee.2016.04.011.

Köthke, M., Ahimbisibwe, V., Lippe, M., 2022. The evidence base on the environmental, economic and social outcomes of agroforestry is patchy—an evidence review map. Front. Environ. Sci. 10. https://doi.org/10.3389/fenvs.2022.925477.

Kuyah, S., Whitney, C.W., Jonsson, M., Sileshi, G.W., Öborn, I., Muthuri, C.W., Luedeling, E., 2019. Agroforestry delivers a win-win solution for ecosystem services in sub-Saharan Africa. A meta-analysis. Agron. Sustain. Dev. 39, 1–18. doi:https://doi.org/10.1007/s13593-019-0589-8.

Mathieu, A., Martin-Guay, M.-O., Rivest, D., 2025. Enhancement of agroecosystem multifunctionality by agroforestry: a global quantitative summary. Glob. Chang. Biol. 31, e70234. https://doi.org/10.1111/gcb.70234.

Nair, P.K.R., Kumar, B.M., Nair, V.D., 2021. An introduction to agroforestry: four decades of scientific developments. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-75358-0.

Ngaba, M.J.Y., Mgelwa, A.S., Gurmesa, G.A., Uwiragiye, Y., Zhu, F., Qiu, Q., Fang, Y., Hu, B., Rennenberg, H., 2024. Meta-analysis unveils differential effects of agroforestry on soil properties in different zonobiomes. Plant Soil 496, 589–607. https://doi.org/10.1007/s11104-023-06385-w.

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., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, n71. https://doi.org/10.1136/bmj.n71.

Pent, G.J., 2020. Over-yielding in temperate silvopastures: a meta-analysis. Agrofor. Syst. 94, 1741–1758. https://doi.org/10.1007/s10457-020-00494-6.

Sauer, T.J., Dold, C., Ashworth, A.J., Nieman, C.C., Hernandez-Ramirez, G., Philipp, D., Gennadiev, A.N., Chendev, Y.G., 2021. Agroforestry practices for soil conservation and resilient agriculture. In: Udawatta, R.P., Jose, S. (Eds.), Agroforestry and Ecosystem Services. Springer International Publishing, Cham, pp. 19–48. https://doi.org/10.1007/978-3-030-80060-4_2.

Schievano, A., Pérez-Soba, M., Bosco, S., Montero-Castaño, A., Catarino, R., Chen, M., Tamburini, G., Landoni, B., Mantegazza, O., Guerrero, I., Bielza, M., Assouline, M., Koeble, R., Dentener, F., Van Der Velde, M., Rega, C., Furlan, A., Paracchini, M.L., Weiss, F., Angileri, V., Terres, J.-M., Makowski, D., 2024. Evidence library of meta-analytical literature assessing the sustainability of agriculture – a dataset. Sci. Data 11, 979. https://doi.org/10.1038/s41597-024-03682-6.

Schievano, A., Bosco, S., Perez-Soba, M., Catarino, R., Montero, C.A., Chen, M., Tamburini, G., Landoni, B., Mantegazza, O., Rega, C., Guerrero, I., Bielza, M., Terres, J.-M., Makowski, D., 2025. Umbrella-Review of meta-Analyses: A Methodological Framework to Support Evidence-Based Policymaking. JRC Publ. Repos. https://doi.org/10.2760/4592550.

Schievano, A., Bosco, S., Pérez-Soba, M., Montero-Castaño, A., Catarino, R., Chen, M., Tamburini, G., Landoni, B., Mantegazza, O., Rega, C., Guerrero, I., Bielza, M., Van der Velde, M., Paracchini, M.L., See, L., Terres, J.-M., Makowski, D., 2026. The quality of meta-analyses in agricultural sustainability has been increasing, but weaknesses persist. Npj Sustain. Agric. 4, 48. https://doi.org/10.1038/s44264-026-00148-7.

Shi, L., Feng, W., Xu, J., Kuzyakov, Y., 2018. Agroforestry systems: Meta-analysis of soil carbon stocks, sequestration processes, and future potentials. Land Degrad. Dev. 29, 3886–3897. https://doi.org/10.1002/ldr.3136.

Tamburini, G., Bommarco, R., Wanger, T.C., Kremen, C., van der Heijden, M.G.A., Liebman, M., Hallin, S., 2020. Agricultural diversification promotes multiple ecosystem services without compromising yield. Sci. Adv. 6, eaba1715. https://doi.org/10.1126/sciadv.aba1715.

Terasaki Hart, D.E., Yeo, S., Almaraz, M., Beillouin, D., Cardinael, R., Garcia, E., Kay, S., Lovell, S.T., Rosenstock, T.S., Sprenkle-Hyppolite, S., Stolle, F., Suber, M., Thapa, B., Wood, S., Cook-Patton, S.C., 2023. Priority science can accelerate agroforestry as a natural climate solution. Nat. Clim. Chang. 13, 1179–1190. https://doi.org/10.1038/s41558-023-01810-5.

Torralba, M., Fagerholm, N., Burgess, P.J., Moreno, G., Plieninger, T., 2016. Do European agroforestry systems enhance biodiversity and ecosystem services? A meta-analysis. Agric. Ecosyst. Environ. 230, 150–161. https://doi.org/10.1016/j.agee.2016.06.002.

Worms, P., Lawson, G., 2024. Agroforestry and the green Deal. . Zenodo. https://doi.org/10.5281/zenodo.14604732.