The implementation of network meta-analysis in Ecology; a case study using crop yield data

https://github.com/maxanochirim/network-meta-analysis-using-crop.yield.data.git

Maximus Anochirim[a], Matthew Grainger[b], Gavin Stewart[c], Elina Takola[a]

[a]Department of Computational Landscape Ecology, UFZ – Helmholtz Centre for Environmental Research, Permoserstrasse 15, Leipzig, 04318, Germany

[b]Norwegian Institute for Nature Research, Trondheim, Norway – NINA

[c]Evidence Synthesis Lab, School of Natural and Environmental Science, University of Newcastle, Newcastle-upon-Tyne, UK

ORCiD

MA: 0000-0002-1366-5488

MG: 0000-0001-8426-6495

GS: 0000-0001-5684-1544

ET: 0000-0003-1268-5513

ABSTRACT

Network meta-analysis (NMA) is a method commonly used in medical research that allows for the comparison of multiple interventions in a single, coherent analysis. In this study, we explore how NMA can be applied in ecological studies – specifically, in comparing the effectiveness of multiple interventions in field experiments. Our study aims to provide a general and non-technical introduction of network meta-analysis to ecologists, particularly on key assumptions and methods. Using an example, we demonstrate how NMA can serve as a tool to compare the effectiveness of different interventions used in enhancing crop yield. We conducted a systematic review and extracted data from meta-analytical studies that explored the response of yield to an intervention. Each study structured data as a pairwise comparison between an intervention and a control. Using yield as the measure of effectiveness, we evaluated four interventions – Liming, Straw return, Super Absorbent Polymers, and considered no intervention as the control. All measurements came from field experiments, and we analyzed data from 3733 independent studies that were included in these meta-analytical reviews. The results of our analysis demonstrate the potential of NMA as a valuable statistical method in ecological research, providing more precise comparisons of multiple interventions. However, we emphasize the need for careful consideration of important assumptions such as transitivity and consistency when implementing NMA in ecological studies. This study offers a novel approach to synthesizing ecological data, contributing to improved decision-making in agriculture, ecology, and environmental sciences.

Keywords: Network meta-analysis, Multi-treatment analysis, Mixed-treatment comparison, Evidence synthesis, Systematic reviews, Policy, Decision making, Crop yield, Interventions, Agriculture, Ecology, Environmental science

1. INTRODUCTION

1.1. What is Network Meta-Analysis (NMA)

The advent of meta-analysis in ecology in the early nineties (Arnqvist & Wooster, 1995; Fernandez- Duque & Valeggia, 1994; Jarvinen, 1991) provided researchers with a statistical method that helped to systematically synthesize empirical findings (Bilotta et al., 2014; H. M. Cooper et al., 2019). In contrast to narrative reviews and other qualitative methods that were previously used to summarize study findings (Gurevitch et al., 2018; Slavin, 1995), the quantitative synthesis of empirical findings allowed for data from multiple studies to be drawn together. Additionally, and since the former did not require a clear formulation of research questions nor a quantitative synthesis of the data (Vetter et al., 2013), meta-analysis proved extremely useful and was considered a powerful statistical method for summarizing research findings. Furthermore, the quantitative synthesis of results across studies offers better precision as well as a higher statistical power as it helps in identifying sources of variation in study outcomes (Gurevitch et al., 2018). Subsequently, the use of meta-analysis (which is a part of the larger field of research synthesis) became widely accepted (**Figure S1**) and adapted across many scientific disciplines despite its initial skepticism (Cadotte et al., 2012; Gurevitch et al., 2018; Hillebrand & Cardinale, 2010; Hunt, 1997; Whittaker, 2010). It offered many advantages including transparency, replicability, generalizability, and the ability to better quantify effect size by analyzing the result from multiple independent studies (Conn et al., 2012; Hernandez et al., 2020).

In ecological and environmental contexts, the application of these methods has not only helped in the identification of research gaps (Thomsen et al., 2012), but has also allowed for the estimation of the direct effects of major environmental drivers – including climate change, invasive species, and habitat fragmentation (Aguilar et al., 2006; Clewley et al., 2012; Jactel et al., 2012). Additionally, these approaches have allowed for comparisons of these effects across different scales, taxa, and ecosystems (Powell et al., 2011; Rodríguez-Castañeda, 2013), thus helping to enhance the reliability and generalization of ecological conclusions (Cadotte et al., 2012; Maki et al., 2018; Nakagawa et al., 2023). Likewise in the field of plant ecology, the quantitative synthesis of research results has had a tremendous impact, proving extremely useful in objectively summarizing findings from numerous studies (Koricheva & Gurevitch, 2014). This has led to the development of standardized protocols, usage, and organizations such as the Collaboration for Environmental Evidence – CEE that guide both scientists and policymakers in providing the best evidence for high quality environmental decision-making processes (Koricheva & Gurevitch, 2014; Nakagawa et al., 2023; Pullin & Stewart, 2007; Vetter et al., 2013), a move similar to already established standards like The Cochrane Collaboration in the field of medical sciences.

Despite all of this, meta-analysis still has its limitations. For example, when one would like to make inferences about all the indicated treatments for the same condition and sample characteristics, the standard meta-analytic methodology experiences a significant drawback. Furthermore, since meta-analysis is focused on comparing only two interventions at the same time, the use of this method implies that all studies must have a common treatment. Likewise, the growing multiplicity of available scientific literature on a topic, particularly in ecology produces a plethora of information/interventions which could unfortunately complicate decision-making processes (Dias & Caldwell, 2019; Koricheva et al., 2013; Roberts et al., 2006). Additionally, ecological systems are oftentimes highly interconnected, with multiple drivers of change acting simultaneously. The limiting ability of meta-analytical methods to adequately mirror the complexity of natural ecosystems allows room for the desire of a statistical method that can optimally address questions about how different environmental factors/interventions interact and compare in their effects. There is therefore growing a new perspective to make inferences about competing treatments for the same condition, one such as the network meta-analysis.

First coined by Lumley (2002), the methodology of network meta-analysis, (sometimes called multi-treatment analysis or mixed-treatment comparison, and hereafter referred to as NMA) focuses on combining both the direct and indirect information across a network of randomized studies to infer about the relative effectiveness of multiple interventions (Dias, 2018; Lumley, 2002; Madden et al., 2016). In a scenario with multiple interventions, as is the case with most ecological data, it is clear that not all possible pairwise comparisons will have been separately/directly carried out. The method of NMA solves this issue by combining in a single coherent analysis, the result of those studies/interventions where a direct comparison has been made (direct evidence) together with those studies/interventions where there is no pairwise comparison (indirect evidence). The term 'indirect' is used because it relies on evidence against the common comparator or other all relevant comparators of interest, and not on 'direct' head-to-head evidence. Thus, with NMA, all interventions can be compared with one another, including comparisons not evaluated within any of the primary studies further strengthening inferences concerning the relative efficacy of treatments (Lu & Ades, 2004). This type of method incorporates all available evidence into a general statistical framework for the comparison of all available treatments thus resolving the limitations of traditional pairwise meta-analyses" (Tu, 2014).

With NMA, it becomes possible to statistically analyze the result from multiple independent studies with different treatments (Higgins & Whitehead, 1996; Salanti, 2012; Salanti & Schmid, 2012); thus providing a more effective comparative review method and subsequently forming the basis for coherent, evidence- based treatment decisions. A major advantage of NMA is that correlations of estimated treatment effects are automatically considered when an appropriate model is used thus improving the precision for the estimated effect sizes. It facilitates simultaneous inference regarding all treatments in order to select the best possible treatment for example (Lu & Ades, 2004; **Figure 1**). Additionally, NMA makes it possible for the comparisons of interventions that are not possible in a single study because all treatments of interest may not be included in any given study. Other advantages of a network meta-analysis include; (i) the preservation of randomization: Because a network meta-analysis does not directly compare treatment arms across studies but rather the relative differences between treatment comparisons, it preserves randomization (i.e., takes into consideration the fact that samples were randomized to interventions within studies but not across studies), (ii) the maximization of all available evidence. (iii) the production of treatment ranking which is useful in decision-making processes or in revealing the most effective components of complex interventions.

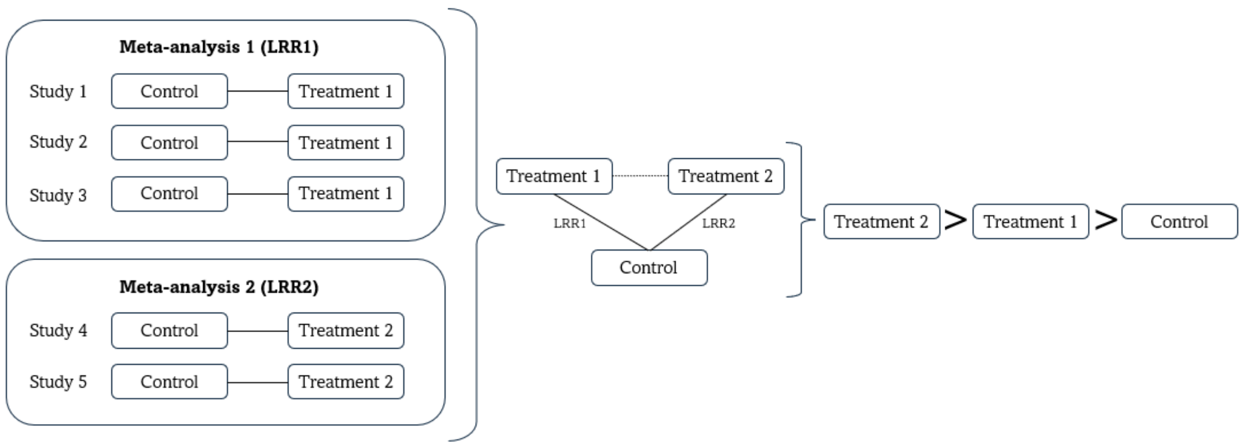

Figure 1. A hypothetical scenario where NMA combines all available evidence including those that have not previously been compared to each other.

By focusing on the use of network meta-analytic approaches in ecological studies, we argue that this relatively new approach and subsequent improvements will aid both researchers and policymakers in their ability to promptly and effectively craft policies that address pressing environmental issues. The development of user-friendly software, protocols, and guidelines such as the PRISMA-NMA extension (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Hutton et al. (2015); http://www.prisma-statement.org/nma) is evidence that the method of NMA is growing in acceptance and popularity. Already, it is being embraced by national and international policy-making bodies as a tool to best answer complex policy-relevant questions (Salanti & Schmid, 2012).

When conducting a network meta-analysis, it is however important to note that analytical methods are more complex and can be quite challenging (Tu, 2014). In addition, confusion about how to choose models, fit them to the data, or interpret the results can be expected for users who are not yet familiar with the different statistical models (Dias, 2018; Madden et al., 2016).

1.2. **Research Objectives**

While studies comparing more than two competing interventions are quite common in several research fields including medicine, pharmacology, social psychology, and education, the implementation of the NMA methodology is not yet common practice in ecological studies. A search for published literature on NMA using the keyword “network meta-analysis” in the literature database ISI Web of Science revealed that the field of ecology contributed less than 1% out of the approximately 27 thousand entries (**Figure S2**). Our main research objective therefore was to investigate the feasibility of applying the methodology of indirect comparisons (network meta-analysis) in the scientific field of ecology as a way of further improving the precision of results that could be obtained from ecological studies. We aim to provide a general and non-technical introduction of network meta-analysis to ecologists, and the general research community in the field of environmental management.

We showed an example of how NMA works by developing a research question that seeks to find out if we could implement a multi-treatment meta-analysis to uncover the effects of different interventions used in the yield production of croplands. To the best of our knowledge, we believe this is the first detailed guideline on the application of NMA in ecology, particularly in how this method can be utilized to assess the efficacy of interventions used in improving crop yield.

2. METHODS

2.1. Basic Concepts and Assumptions of a Network Meta-Analysis

2.1.1. Concepts and terminology

The methodological advantage behind a network meta-analysis is a very simple one: "indirect comparisons". Assuming we have two independent studies (study 1: directly comparing treatment A versus C [μ^{D}_{AC}], and study 2 directly comparing treatment B versus C [μ^{D}_{BC}] but no studies directly comparing A versus B [μ^{D}_{AB}]), we can indirectly compare treatment A to treatment B via the common comparator C. This is done by statistically combining the information from all A versus C (AC) and B versus C (BC) studies represented here as: $\mu^{I}_{AB} = \mu^{D}_{AC} - \mu^{D}_{BC}$ (**Figure 2**). With a network meta-analysis, it is also possible to further improve the precision of treatment estimates assuming a scenario exists where both the direct [μ^{D}_{AB}] and indirect estimates [μ^{I}_{AB}] are available for the same comparison. In this situation, a 'mixed' effect size [μ^{M}_{AB}] is calculated by taking the weighted average of μ^{D}_{AB} and μ^{I}_{AB} (Bucher et al., 1997).

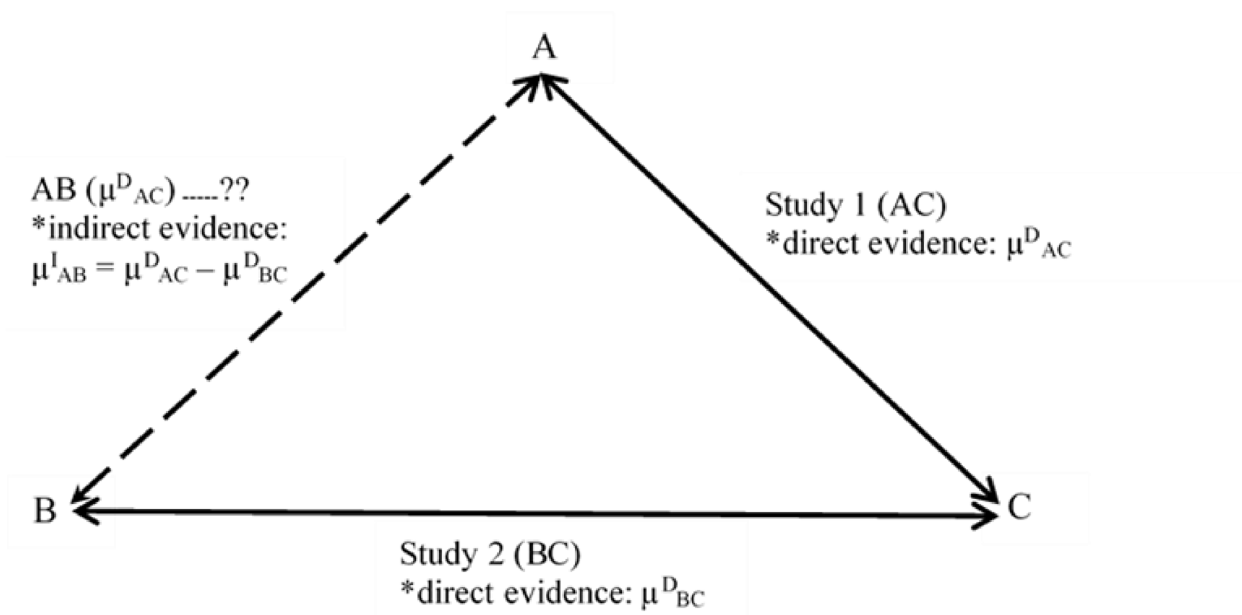

Figure 2. An example of a network of three treatments (ABC) compared in two studies (solid black lines), where an indirect comparison can be made (dashed grey line).

2.1.2. **Statistical assumptions**

Since indirect and mixed comparisons are generally considered observational in nature (Catalá-López et al., 2014), two major assumptions need to be considered and where possible, met, before a network meta-analysis study can be considered as valid.

1. <u>Assumption of transitivity:</u> The main assumptions that underpin the validity of indirect and mixed comparisons is that there are no significant differences between the studies making different comparisons other than the treatments that are being compared (Cipriani et al., 2013). Since the aim of a network meta-analytical study is to compare two treatments via a third one, it assumes that indirect comparison validly estimates unobserved head-to-head comparison (Salanti, 2012). With our earlier example comparing treatment A versus treatment B via treatment C, our common comparator which is C is regarded as 'transitive'. This is because it allows a valid comparison of the treatment to which it is linked (Salanti, 2012). With transitivity, similarity is not required for all characteristics of studies and samples across the evidence base i.e., valid indirect comparisons can be obtained even when studies are dissimilar in characteristics which are not effect modifiers (Dias, 2018). In technical terms, what this means is that indirect comparisons can still be made between a study conducted in pot experiments and another study that was conducted in field conditions provided that the study samples were not shown to have a modifying effect on the result from any of the study. Additionally, this assumption requires that treatments/interventions should be comparable among themselves in practice. For example, assuming we were comparing two treatments A and B for plant yield. If treatment A needs a precondition before it can be implemented (e.g. only as a second-line treatment or perhaps only to samples with certain conditions) and Treatment B doesn't need these preconditions, then the law of transitivity is violated because in practice, the observational samples would be different for each treatment (we would not be

able to randomly assign the samples to just any treatment in a fresh study). Furthermore, samples that are included in a network should be able to be randomized (sample randomization) to any of the treatments/interventions. i.e. (in principle), one should be able to apply any of the treatments/interventions randomly to all participants in the network. The assumption of transitivity could be violated if interventions have different indications.

Questions to be asked regarding the assumption of transitivity:

i. In the planned (ecological) study, can the two treatments that want to be compared indirectly form a common node?
ii. Are the missing treatments ‘missing at random’ or is it directly associated with the true relative effectiveness of the interventions (intervention effectiveness bias)?
iii. Is the choice of the comparator random? If the choice of the comparison is associated, directly or indirectly, with the relative effectiveness of the interventions, then the assumption of transitivity is violated.

2. <u>Assumption of consistency</u>: Consistency is the extension of transitivity over a loop of evidence (Cipriani et al., 2013). With consistency, the major assumption is that the direct and indirect estimates/sources of evidence agree i.e., both the direct and indirect evidence are estimating the same underlying treatment effect. This assumption can be measured/evaluated statistically with the use of a simple z-test, often called the Bucher method (Bucher et al., 1997). It can be evaluated only when there is direct and indirect evidence existing in the evidence network for a particular comparison of interventions (Dias et al., 2010; Dias & Caldwell, 2019; Higgins et al., 2012). When the direct comparisons of means are different from indirect comparisons, then the network is said to be inconsistent (Cipriani et al., 2013; Lu & Ades, 2004, 2006). The assumption of consistency is a prerequisite in calculating a valid mixed estimate. A significant *f*-test for the design and treatment interaction is an indication of inconsistency.

Questions to be asked regarding the assumption of consistency:

i. Do the ecological studies use comparable methodologies (e.g., measurement of species abundance, or habitat quality) to ensure consistency in effect estimates?

ii. Is there significant variability in ecological contexts that might lead to inconsistent results between direct and indirect comparisons?

iii. Do studies reporting direct comparisons conflict with indirect evidence due to context-specific factors?

3. <u>Exchangeability assumption</u>: This assumes that the distribution of effect modifiers (variables that influence treatment effects) is balanced across studies. i.e., that two sets of studies e.g. AC and BC, do not differ with respect to the distribution of effect modifiers. For example: if 'site' was an effect modifier in all AC studies (irrespective of their distribution – heterogeneous studies), then in order to make a valid indirect comparison of AB; 'site' should also be an effect modifier (distributed in a similar proportion) in all BC studies - AC and BC studies should therefore cover the entire spectrum of the observed effect modifier. It is therefore important to identify apriori possible effect modifiers and compare their distributions across comparisons when synthesizing evidence from many comparisons. Adjustments can be used to improve transitivity (through network meta-regression or a subgroup analysis) if an imbalanced distribution of effect modifiers is identified. Adjustment should take place only for study or sample characteristics that are categorized as effect modifiers.

Questions to be asked regarding the assumption of exchangeability:

i. Are key ecological variables (e.g., temperature, precipitation, soil) influencing the treatment effects similarly distributed across the studies in the network?

ii. Are the interventions or treatments implemented in similar ecological settings, or are there systematic differences?

iii. Are there outliers among the studies (e.g., studies conducted in extreme environments or with highly specialized species) that might bias the overall effect estimate?

2.1.3. Statistical approaches to fitting a network meta-analysis

Both Bayesian and frequentist approaches can be used in fitting a network meta-analysis model (Dias & Caldwell, 2019; Hong et al., 2013). Bayesian methods for NMA require selecting a prior probability distribution that describes the range and probability of plausible values for the parameters of interest (e.g., treatment effect). Using Bayes theorem, this is then combined with a likelihood statement that provides information on the collected data (Dias & Caldwell, 2019). On the other hand, the frequentist approach to network meta-analysis does not require prior knowledge or beliefs. Relying solely on the data collected, the frequentist approach calculates probabilities and estimates based on how likely the observed data would be under different assumptions about the parameter of interest (e.g., treatment effect) (Rücker, 2012). The results from both analyses however are quite similar (Dias & Caldwell, 2019; Hong et al., 2013) with the main difference being the way results are presented. Results from a frequentist approach are presented as estimated relative effects and a corresponding 95% Confidence Interval (CI), while results from a Bayesian NMA analysis are presented as summaries of the effect (typically, mean or median) and a 95% credible interval (CrI) (Dias & Caldwell, 2019). Both approaches can be implemented in commonly used statistical software such as R. Bayesian approaches to network meta-analysis are conducted with packages such as; “gemtc”, “bnma”, “pcnetmeta”, “multinma”, “nmaINLA”, “bayesmeta”, “BUGSnet” while the package “netmeta” developed by Balduzzi et al. (2023) is used for frequentist methods.

2.2. An example of network meta-analysis in ecology: a case study using crop yield data

2.2.1. Data collection

In this systematic review and network analysis, we searched the database used in Takola et al. (unpublished) to identify published studies of interventions applied in the production of various crop yields. We compiled a dataset of meta-analytical studies from around the globe that had investigated the response of yield production to an agricultural intervention (**Figure 3**). These strategies – ranging from; straw return, liming, super absorbent polymers, alternative fertilization options, substitution of mineral fertilizers with manure nitrogen, removal of topsoil, etc. are among some of the most common agricultural management strategies for balancing yield increase (a full list of interventions can be found in **Table S1**). Our initial dataset (from Takola et al., unpublished) consisted of 11 meta-analytical studies, with 13 interventions and 8814 yield data (**Figure 4**). Study information extracted include (i) study identifiers (e.g. title, abstract, authors, publication year); (ii) study characteristics (e.g. study design, study location); (iii) participant characteristics (e.g. sample size, standard deviation); (iv) intervention details; and (v) outcome data (e.g. effect sizes of intervention and control). Our primary outcome of interest was yield production measured in kilogram per hectare (kg/ha).

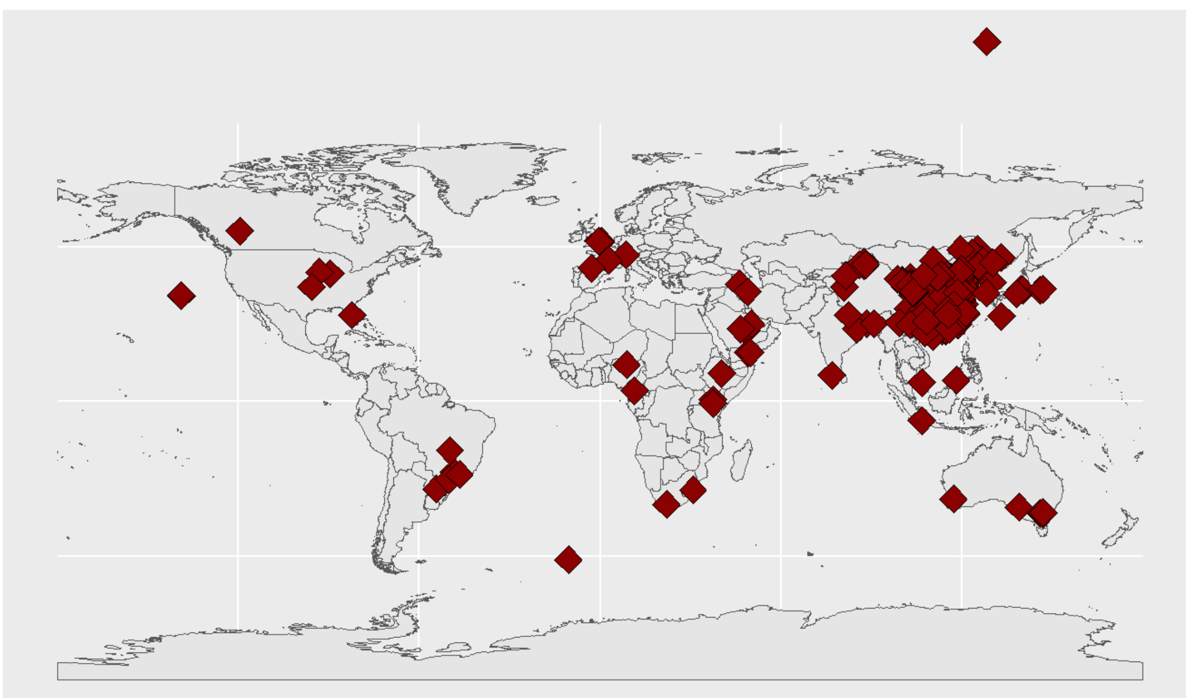

Figure 3. Locations of sites included in the network meta-analysis. (see Appendix for a complete list of references)

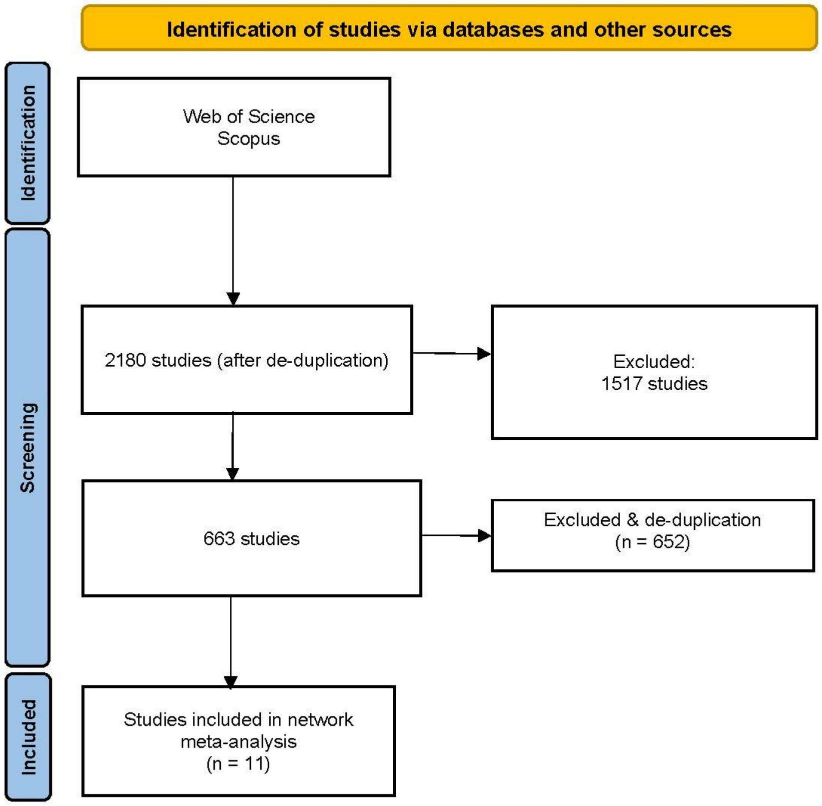

Figure 4. A PRISMA flow diagram showing the systematic review process (‘identification’, ‘screening’, ‘eligibility’, and ‘included’) of selecting publications relevant to our network meta-analysis.

2.2.2. **Inclusion criteria**

To help ensure consistency in our analysis, we defined a set of criteria to screen out irrelevant studies and only included studies that (i) were randomized controlled trials (ii) were paired i.e., compared an active intervention with a control (iii) had no missing data e.g. SD, sample size (n), effect sizes, etc. (iv) used interventions not applied as a second-line treatment.

After applying our selection criteria (**Table 1**), we extracted data from 3 meta-analytical studies of interest (**Table 2**). These studies comprising of 3 paired interventions and 3733 yield data met all inclusion criteria and were then used for the network meta-analysis.

Table 1. Summary of eligibility criteria

	Inclusion criteria	Exclusion criteria

Population	Studies with no missing data. E.g. SD, sample size (n).	No SD
	Studies that were open-access	
	Studies that had supplementary datasets available	
Intervention	Interventions for improving yield	
Study design	Randomized controlled trials	
Comparison	Paired. i.e., Intervention vs Control group (no intervention)	Interventions not paired
Outcome	Yield values	

Table 2. Interventions of Interest

No.	Intervention	Dataset
1.	Straw return	D331
2.	Liming	D973
3.	Super Absorbent Polymer (SAP)	D652
4.	No Intervention (Control)	D331, D973, D652

2.2.3. **Data analysis**

To make a comparison of these selected interventions on yield production, we conducted a network meta-analysis in R (version 4.3.2) using the *netmeta* package. This approach adopts frequentist methods and calculates point estimates and their corresponding confidence intervals based on weighted least squares regression (Rücker, 2012). We calculated the effect of the different agricultural interventions on yield production. As a summary measure of effect size, we estimated standardized mean differences (SMD) of each intervention relative to the control group using pairwise and a random effects meta-analytical model. In choosing a random-effects model for our network meta-analysis, we assume that heterogeneity exists among studies and that all effect sizes did not come from one population i.e., each study has different overall means and they don't have one true overall mean. By choosing a random-effects model, our analysis further accounts for variability both within and between studies. The alternative to this is the 'fixed-effect' model commonly referred to as the 'common-effect' model which assumes that all effect sizes (from different studies) come from one population (Nakagawa et al., 2023). From the generated estimates and confidence intervals, probability scores (*P*-scores) were calculated and these were used to hierarchically rank each intervention according to their effects on yield using methods developed by Rücker & Schwarzer, 2015. Estimates, confidence intervals, and *P*-scores then allowed us to construct, forest plots and league tables which are useful in visualizing the comparisons.

Our data input was based on a wide-arm format with 3733 rows (where each row corresponds to a single study with multiple or double comparisons) and it was subsequently transformed during the analytical process into the standard contrast-based format using the auxiliary *pairwise* function of the *netmeta* package. Our dataset has a continuous outcome wherein all variables containing information on group sample sizes (argument n), means (mean), and standard deviations (SD) are provided – hence the reason why we chose the SMD as an effect size. SMD is used as a summary statistic when studies assess the same outcome but use different measurement methods (*9.2.3.2 The Standardized Mean Difference*, n.d.). In such cases, it becomes necessary to standardize the results of the studies to a common scale before combining them. SMD (often times referred to as Hedges' g or Cohen's d) is considered a comparative measure because it is typically used when comparing two groups (Nakagawa et al., 2023). In the *netmeta* package, the method by Crippa & Orsini, (2016) is used to guarantee consistent SMDs and standard errors for multi-arm studies (Balduzzi et al., 2023).

$$\text{SMD} = \frac{Difference\ in\ mean\ outcome\ between\ groups}{Standard\ deviation\ of\ outcome\ among\ participants}$$

Similarly, a workflow detailing the steps involved in performing a network meta-analysis, along with the related functions of *netmeta* in R is provided in **Figure S3**. All codes used in our study are also provided in the supplement section.

3. RESULTS

3.1. Data summary

This dataset contains cleaned data extracted from a systematic review assessing the effect of different interventions on yield production in different crop farms (maize, wheat, cotton, oat, etc.). They are a combination of three pairwise meta-analyses comparing the effects of "Straw return“, “Super Absorbent Polymer“, and "Liming“ respectively, with a Control (no treatment). The primary outcome was “Yield production”. Relative treatment effects were expressed as SMD. Data on this outcome were available for 3733 (total number of samples) from 3733 single pairwise comparison studies; all of which are two-arm studies comparing an active treatment against the absence of a treatment (control), thus providing information for two means (y1, y2), two standard deviations (sd1, sd2), two group sample sizes (n1, n2), along with two treatment labels (t1, t2).

3.2. **Summary of Networks**

The number of treatments of interest in our network (also called nodes or vertices) is 4 (n=4) and d which is the number of designs is 3. Additionally, each study contributes a number of pairwise comparisons (m) and the total sum of all pairwise comparisons across studies in our network is 3732. K (which is the number of independent studies) simultaneously corresponds to m and this is because there are only two-arm studies in our network. Assuming there was at the least, one study evaluating more than two treatments, then m will be greater than k.

3.3. **Net graph**

In the network graph (**Figure 5**), we get a graphical presentation of the network structure with each treatment represented as a point (node) in the plane. It shows a network of interventions compared in a yield production study. Furthermore, treatments are connected by a line (edge) if at least one direct pairwise comparison exists with the thickness of the edges being proportional to the number of studies directly comparing treatments. There are 3 edges in the plot, suggesting that 3 of the 6 pairwise comparisons had direct evidence, while the remaining 3 (SAP versus Straw return, SAP versus Liming, and Liming versus Straw return) had only indirect evidence. Our network graph also visualizes the number of studies contributing to each pairwise comparison. From the net graph, we immediately see from the line width that the comparison of Super Absorbent Polymer (SAP) versus Control has the largest number of studies (1424). Furthermore, all studies were two-armed and only had direct comparisons.

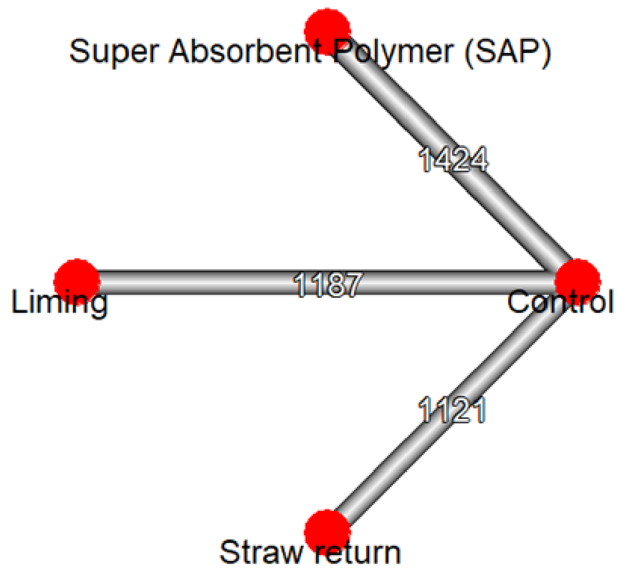

Figure 5. Network graph without crossings for yield production network meta-analysis. Line widths are proportional to the number of studies directly comparing treatments

3.4. Forest plot

Forest plots provide a graphical display of the observed effect, confidence interval, and often times the weight of each study (Harrer et al., 2021). It can also be used as a way to better visualize the uncertainty in our network. With “Control” as the comparison group and from the forest plot that was produced, we visually see that the intervention “Super Absorbent Polymer” works better in comparison with the other interventions in improving crop yield, while the intervention “Straw return” does not have a strong impact on yield (**Figure 6 and 7**).

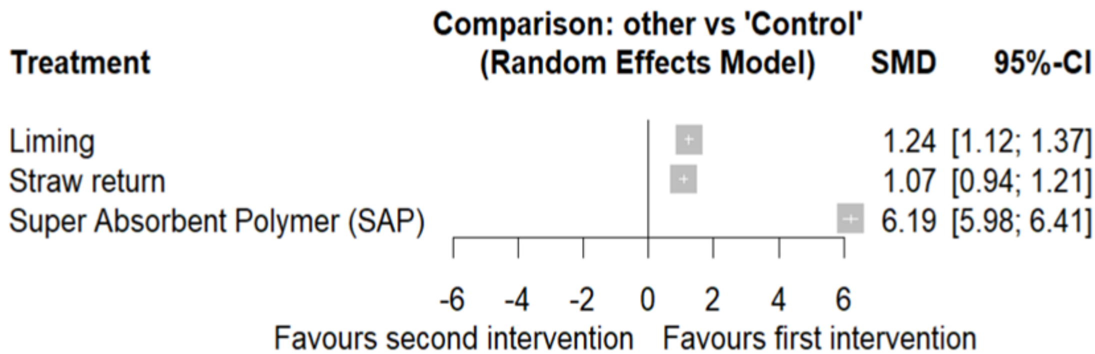

Figure 6. Forest plot for yield production network meta-analysis; with control as reference.

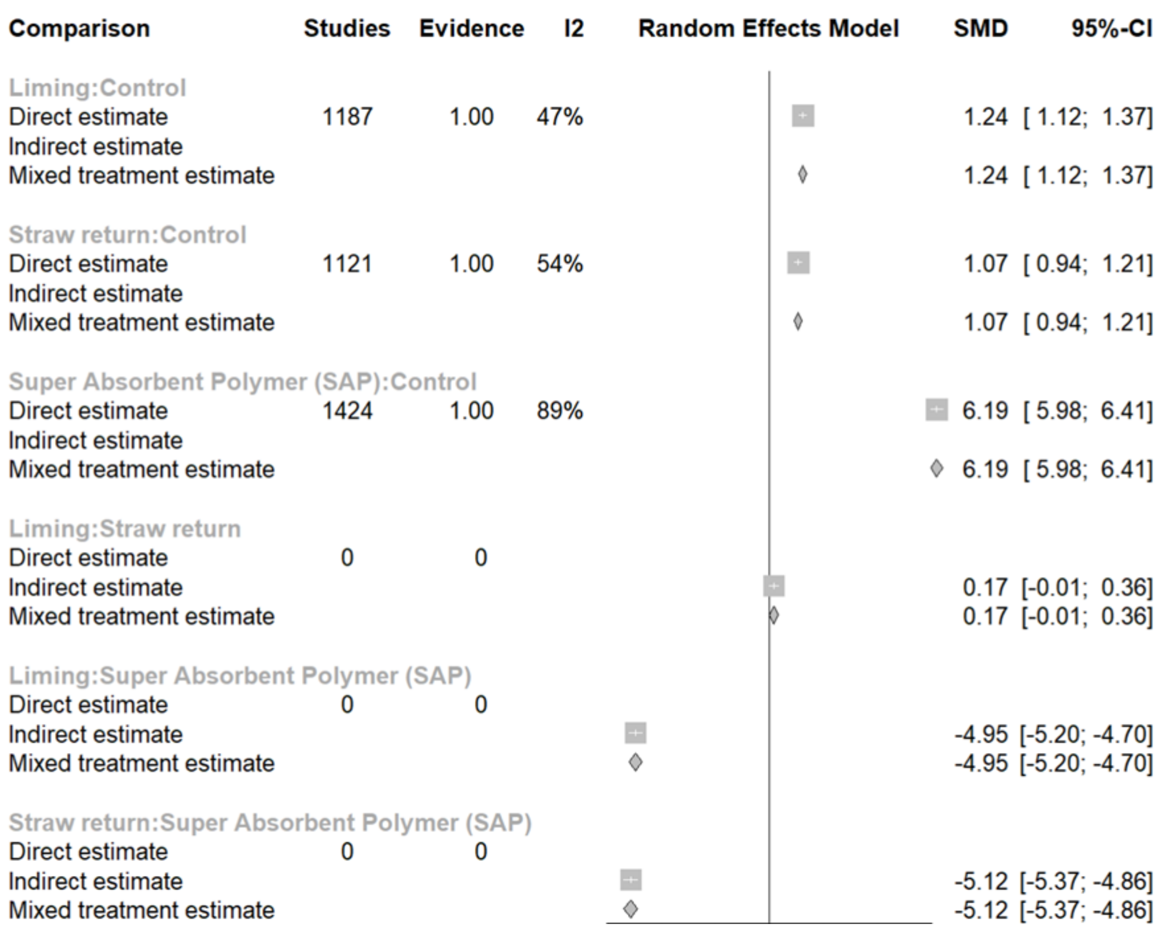

Figure 7. Forest plot for yield production network meta-analysis (active intervention versus all other treatments).

3.5. Treatment estimates (SMD) – Assessment of interventions on yield

The network estimates for the random effects models are provided, using Control as the reference group. We produced network estimates, lower and upper confidence limits for all observed and unobserved pairwise comparisons. The results in **Table 4** below summarizes the standardized mean differences (SMDs) of Yield (in kg/ha) for each treatment compared to the "Control" along with their 95% Confidence Intervals.

Table 4. Summary of standardized mean differences of yield

	SMD	95%-CI	z	p-value
Control	.	.	.	.
Liming	1.24	[1.12; 1.37]	18.89	< 0.0001
Straw return	1.07	[0.94; 1.21]	15.88	< 0.0001
Super Absorbent Polymer (SAP)	6.19	[5.98; 6.41]	55.72	0

The random effects NMA shows strong evidence that all treatments have a significant positive effect on yield production compared to the control, as indicated by p-values < 0.0001. Liming (SMD = 1.24, 95% CI [1.12, 1.37]) had a moderate positive effect compared to the control, with a statistically significant result ($p < 0.0001$). The intervention - liming, increases crop yield by 1.24 kg/ha, compared with an absence of it. In addition, we can also say that there is a 95% probability that this increase is between 1.12 and 1.37 kg/ha. Straw return (SMD = 1.07, 95% CI [0.94, 1.21]) also showed a moderate positive effect relative to the control, and the result was statistically significant ($p < 0.0001$). SAP showed the largest effect size (SMD = 6.19), indicating a very strong impact compared to the other treatments.

3.6. Assessing heterogeneity (within designs) and inconsistency (between designs)

$$tau^2 = 4.1014; tau = 2.0252; I^2 = 78.2\% [77.6\%; 78.9\%]$$

I^2 measures the proportion of total variation in effect estimates that is due to heterogeneity rather than chance. An I^2 of 78.2% suggests substantial heterogeneity, indicating that the treatment effects vary considerably across studies.

Table 5. Quantifying heterogeneity and inconsistency

	Q	d. f.	p-value
Total	17123.15	3729	0
Within designs	17123.15	3729	0
Between designs	0.00	0	--

The results of the Q statistic tests for heterogeneity among studies (Total Q) show a p-value of < 0.0001, thus indicating significant heterogeneity. Additionally, the significant p-value of the Q statistic examining heterogeneity within the groups of studies that share the same design (within designs) indicates substantial variability even within these groups. The tests of inconsistency (Between designs) typically assess differences between groups of studies with different designs. The results of our study suggest that all heterogeneity is captured within designs since there are no degrees of freedom provided for this test.

Based on the results of the tests for heterogeneity and inconsistency **(Table 5)**, we can deduce that there is significant variation among studies (this is supported by the high value of I^2 and the significance of the Q value. Given this, we performed the following recommended steps as recommended by N. J. Cooper et al., 2009.

- **Investigating sources of Heterogeneity:** With the use of (i) Subgroup Analysis, there is the possibility of conducting analyses within subgroups of studies to explore whether certain study characteristics would account for the observed heterogeneity, and (ii) the use of meta-regression techniques to identify factors that might explain the variability in effect sizes across studies. In this study however, we did not explore the causes of heterogeneity due to limitations of time, resource, and available information on important covariates at a global scale (such as field size, management history, management intensity, etc.).

3.7. Hierarchy/Ranking of competing treatments

Table 6 shows treatment rankings and the probability that each intervention is the 'best' or 'worst' in improving crop yield. Here ranks are reported for effectiveness, such that rank 1 means that the intervention is most effective. Control/No Intervention has a rank of 4 (*P-score* = 1.0000). That is, on average, the absence of no intervention was ranked approximately fourth out of the four available treatments (i.e., worst) for improving crop yield. Conversely, Super Absorbent Polymer (SAP) was ranked first out of all four treatments and had a higher probability of being the most effective treatment to improve crop yield. The area under the cumulative ranking curve (SUCRA), a summarization method that gives an index of the overall performance of the treatments (Salanti et al., 2011) can then be derived from the ranking probability of each treatment. SUCRA ranges from 0 to 1 and the higher the SUCRA value, the greater the likelihood that a treatment is better than the other treatments in the network. On the other hand, the closer the SUCRA value to 0, the greater the likelihood that a treatment is worse than the other treatments.

Table 6. Treatment ranking based on P-scores and SUCRA values

Intervention	**P-score**	**SUCRA values**

Control	1.0000	1.0000
Straw return	0.6548	0.6543
Liming	0.3452	0.3457
Super Absorbent Polymer (SAP)	0.0000	0.0000

- *SUCRA values based on 1000 simulations*

4. DISCUSSION

As experimental evidence in Ecology increases, so does the need for methodologies and statistical models to analyze them. Meta-analyses are powerful tools that synthesize evidence and help decision-making (Bilotta et al., 2014). Currently, pairwise meta-analyses are the most commonly used statistical technique to quantitatively synthesize research findings and to compare treatments. However, they have limitations as they can only compare two treatment options at a time. Network meta-analyses are a tool to compare multiple treatments because they combine in a single coherent analysis, studies where a direct comparison has been made (direct evidence) together with studies where there is no pairwise comparison (indirect evidence). In addition, they provide a helpful framework to present a comprehensive, and reproducible synthesis of the evidence (Bilotta et al., 2014). However, the uptake of this method by ecological research has been rather slow.

In this study, we present a network meta-analysis and implement it on data from agricultural field experiments. We provide a general introduction to network meta-analysis (including its advantages over the traditional pairwise meta-analysis), showcase the steps involved in conducting/correctly reporting one, and discuss the major assumptions that guide a standard NMA, such as consistency and transitivity. In this NMA, we were able to compare different interventions used in improving crop yield. Doing so also helped us make comparisons of interventions that had not been previously addressed in any individual primary study (indirect evidence). The dataset we used contains experiments with the following treatments: Super Absorbent Polymer, straw return, liming, and a control group (no intervention). Our results show that the most effective treatment for yield enhancement is the Super Absorbent Polymer. The hierarchy of the treatments was as follows: Super Absorbent Polymer, liming, straw return, and control.

The benefits of Super Absorbent Polymer on crop yield have already been summarized in previous reviews by Zheng et al. (2023). When compared to a control group (no intervention), the addition of SAP significantly increased ($p < 0.01$) crop yields by 12.8% (CIs: 12.1 - 13.4%). The effect sizes of crop yield under liming in comparison to control treatments (no liming) showed that liming similarly had a positive influence on crop yield (Enesi et al., 2023; Li et al., 2019; Liao et al., 2021; average of 12.9%). Likewise, Wang et al. (2021) reported an annual increase of 5.83% in the yield of agricultural products like corn when an optimal scheme of straw return was implemented relative to straw removal. Assuming we were to look at the result of our NMA individually (i.e. Treatment vs Control), these pairwise meta-analyses comparisons confirm our results that all interventions significantly contribute to improving crop yield. Additionally, they could offer some insight into the validity of the treatment rankings obtained during our network meta-analysis. To the best of our knowledge, the global network meta-analysis study published by Herrmann et al. (2022) on the promotion of crop growth, yield, and quality by bioeffectors is the first study that utilizes NMA methods on agricultural field experiments. We however recommend that subsequent studies that plan on implementing this method follow reporting protocols and already established guidelines such as the one suggested by Hutton et al. (2015) when conducting or reporting the results of an NMA.

Network meta-analysis (NMA) can be a valuable tool for ecologists when synthesizing evidence from studies that compare multiple interventions or management strategies in controlled experimental settings. For example, lattice square agricultural experiments, where different combinations of crops, fertilizers, or farming methods are systematically tested across multiple plots, provide structured data suitable for NMA. Similarly, controlled field experiments designed to evaluate different ecological restoration techniques, pest management strategies, or habitat interventions often involve overlapping comparisons that align well with NMA assumptions, such as transitivity and consistency. On the other hand, the applicability of NMA is more limited in monitoring or impact studies, as these often lack the controlled settings and standardized comparisons needed for a robust analysis. Monitoring studies frequently involve highly variable contexts, such as natural ecosystems with diverse species interactions or long-term impacts influenced by unmeasured confounders (e.g., climate variability, historical land use). These complexities make it challenging to meet the assumptions of exchangeability and consistency, reducing the reliability of indirect comparisons in such cases.

A common criticism of the implementation of network meta-analyses in ecological contexts is that experiments are not randomized because the researcher is not typically blind to the control and treatment groups. Randomization is essential in medical studies because it helps to make causal inferences between the treatment and the effect. Although the ecologist cannot be blind to field treatments, we argue that causal inference is facilitated from the variation of the contexts in which each treatment is applied (i.e. temporal and spatial contexts). Regarding the assumption of transitivity which essentially seeks to ensure that studies are comparable to each other with respect to any potential effect-modifying characteristics, we emphasize the need to carefully assess the quality of studies to ensure they are comparable before they are included in a network meta-analysis. Violating this crucial assumption can lead to inconsistencies in a network, which can lead to inaccuracies in the result of the analysis.

In conclusion, the use of NMA has continued to grow over the last decades with expansions into different research fields. We showcase with our study, how this method can be implemented on data from agro-ecological experiments, but there is a plethora of other contexts in which this method can be applied: for example, (i) in the estimation of the effects of environmental drivers – including climate change, invasive species, and habitat fragmentation, (ii) in the comparison of interventions used in nature conservation or forest management (iii) long-term experiments in grassland ecosystems (such as The Jena Experiment). More importantly, the resulting hierarchy of treatments, based on their effectiveness, is a very valuable and important tool to inform decision-making. For example, policy-makers can evaluate specific conservation measures as well as their interactions, to effectively design protected areas. Overall, network meta-analysis is a novel tool at the disposal of ecologists, in their effort to find the best nature-based solutions.

List of abbreviations

NMA – Network Meta-analysis

CEE – Collaboration for Environmental Evidence

PRISMA NMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

ISI – Institute for Scientific Information

CI – Confidence Interval

CrI – Credible Interval

SD – Standard Deviation

SAP – Super Absorbent Polymer

PICO – Population, Intervention, Comparison and Outcome

SMD – Standardized Mean Differences

SUCRA – Surface Under The Cumulative Ranking Curve

DECLARATIONS

Ethics approval and consent to participate

- Not applicable

Consent for publication

- Not applicable

Availability of data and materials

- The datasets generated and analysed during the current study are publicly available online at the following repository - https://github.com/Helmholtz-UFZ/network-meta-analysis-using-crop.yield.data

Competing interests

- The author(s) declare that they have no competing interests

Funding

- Not applicable

Acknowledgements

- Not applicable

REFERENCES

9.2.3.2 The standardized mean difference. (n.d.). Retrieved 21 August 2024, from https://handbook-5-1.cochrane.org/chapter_9/9_2_3_2_the_standardized_mean_difference.htm

Aguilar, R., Ashworth, L., Galetto, L., & Aizen, M. A. (2006). Plant reproductive susceptibility to habitat fragmentation: Review and synthesis through a meta-analysis. *Ecology Letters*, *9*(8), 968–980. https://doi.org/10.1111/j.1461-0248.2006.00927.x

Arnqvist, G., & Wooster, D. (1995). Meta-analysis: Synthesizing research findings in ecology and evolution. *Trends in Ecology & Evolution*, *10*(6), 236–240. https://doi.org/10.1016/S0169-5347(00)89073-4

Balduzzi, S., Rücker, G., Nikolakopoulou, A., Papakonstantinou, T., Salanti, G., Efthimiou, O., & Schwarzer, G. (2023). **netmeta**: An *R* Package for Network Meta-Analysis Using Frequentist Methods. *Journal of Statistical Software*, *106*(2). https://doi.org/10.18637/jss.v106.i02

Bilotta, G. S., Milner, A. M., & Boyd, I. (2014). On the use of systematic reviews to inform environmental policies. *Environmental Science & Policy*, *42*, 67–77. https://doi.org/10.1016/j.envsci.2014.05.010

Bucher, H. C., Guyatt, G. H., Griffith, L. E., & Walter, S. D. (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology*, *50*(6), 683–691. https://doi.org/10.1016/S0895-4356 (97)00049-8

Cadotte, M. W., Mehrkens, L. R., & Menge, D. N. L. (2012). Gauging the impact of meta-analysis on ecology. *Evolutionary Ecology*, *26*(5), 1153–1167. https://doi.org/10.1007/s10682-012-9585-z

Catalá-López, F., Tobías, A., Cameron, C., Moher, D., & Hutton, B. (2014). Network meta-analysis for comparing treatment effects of multiple interventions: An introduction. *Rheumatology International*, *34*(11), 1489–1496. https://doi.org/10.1007/s00296-014-2994-2

Cipriani, A., P.T. Higgins, J., R. Geddes, J., & Salanti, G. (2013). Conceptual and Technical Challenges in Network Meta-analysis. *Annals of Internal Medicine*. https://www.acpjournals.org/doi/10.7326/0003-4819-159-2-201307160-00008

Clewley, G. D., Eschen, R., Shaw, R. H., & Wright, D. J. (2012). The effectiveness of classical biological control of invasive plants. *Journal of Applied Ecology*, *49*(6), 1287–1295. https://doi.org/10.1111/j.1365-2664.2012.02209.x

Conn, V. S., Ruppar, T. M., Phillips, L. J., & Chase, J.-A. D. (2012). Using meta-analyses for comparative effectiveness research. *Nursing Outlook*, *60*(4), 182–190. https://doi.org/10.1016/j.outlook.2012.04.004

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *Handbook of research synthesis and meta-analysis* (3rd Edition). Russell Sage Foundation.

Cooper, N. J., Sutton, A. J., Morris, D., Ades, A. E., & Welton, N. J. (2009). Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non- rheumatic atrial fibrillation. *Statistics in Medicine*, *28*(14), 1861–1881. https://doi.org/10.1002/sim.3594

Crippa, A., & Orsini, N. (2016). Dose-response meta-analysis of differences in means. *BMC Medical Research Methodology*, *16*(1), 91. https://doi.org/10.1186/s12874-016-0189-0

Dias, S. (Ed.). (2018). *Network meta-analysis for decision making*. Wiley.

Dias, S., & Caldwell, D. M. (2019). Network meta-analysis explained. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *104*(1), F8–F12. https://doi.org/10.1136/archdischild-2018-315224

Enesi, R. O., Dyck, M., Chang, S., Thilakarathna, M. S., Fan, X., Strelkov, S., & Gorim, L. Y. (2023). Liming remediates soil acidity and improves crop yield and profitability—A meta-analysis. Frontiers in Agronomy, 5. https://doi.org/10.3389/fagro.2023.1194896

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*(7695), 175–182. https://doi.org/10.1038/nature25753

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Chapter 12 Network Meta-Analysis | Doing Meta-Analysis in R*. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/netwma.html#netw-which-model

Hernandez, A. V., Marti, K. M., & Roman, Y. M. (2020). Meta-Analysis. *Chest*, *158*(1), S97–S102. https://doi.org/10.1016/j.chest.2020.03.003

Herrmann, M. N., Wang, Y., Hartung, J., Hartmann, T., Zhang, W., Nkebiwe, P. M., Chen, X., Müller, T., & Yang, H. (2022). A Global Network Meta-Analysis of the Promotion of Crop Growth, Yield, and Quality by Bioeffectors. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.816438

Higgins, J. P. T., & Whitehead, A. (1996). Borrowing Strength from External Trials in a Meta-Analysis. *Statistics in Medicine*, *15*(24), 2733–2749. https://doi.org/10.1002/ (SICI) 1097-0258(19961230)15:24<2733: AID-SIM562>3.0.CO; 2-0

Higgins, J. P. T., Jackson, D., Barrett, J. K., Lu, G., Ades, A. E., & White, I. R. (2012). Consistency and inconsistency in network meta- analysis: Concepts and models for multi- arm studies. Research Synthesis Methods, 3(2), 98–110. https://doi.org/10.1002/jrsm.1044

Hillebrand, H., & Cardinale, B. J. (2010). A critique for meta- analyses and the productivity–diversity relationship. *Ecology*, *91*(9), 2545–2549. https://doi.org/10.1890/09-0070.1

Hong, H., Carlin, B. P., Shamliyan, T. A., Wyman, J. F., Ramakrishnan, R., Sainfort, F., & Kane, R. L. (2013). Comparing Bayesian and Frequentist Approaches for Multiple Outcome Mixed Treatment Comparisons. Medical Decision Making, 33(5), 702–714. https://doi.org/10.1177/0272989X13481110

Hunt, M. M. (1997). *How science takes stock: The story of meta-analysis*. Russell Sage Foundation.

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P. A., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catalá-López, F., Gøtzsche, P. C., Dickersin, K., Boutron, I., Altman, D. G., & Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. Annals of Internal Medicine, 162(11), 777–784. https://doi.org/10.7326/M14-2385

Jactel, H., Petit, J., Desprez-Loustau, M.-L., Delzon, S., Piou, D., Battisti, A., & Koricheva, J. (2012). Drought effects on damage by forest insects and pathogens: A meta-analysis. *Global Change Biology*, *18*(1), 267–276. https://doi.org/10.1111/j.1365-2486.2011.02512.x

Jarvinen, A. (1991). A meta- analytic study of the effects of female age on laying- date and clutch- size in the Great Tit Parus major and the Pied Flycatcher Ficedula hypoleuca. Ibis, 133(1), 62–67. https://doi.org/10.1111/j.1474-919X.1991.tb04811.x

Koricheva, J., & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, *102*(4), 828–844. https://doi.org/10.1111/1365-2745.12224

Koricheva, J., Gurevitch, J., & Mengersen, K. (2013). *Handbook of Meta-analysis in Ecology and Evolution*. Princeton University Press.

Li, Y., Cui, S., Chang, S. X., & Zhang, Q. (2019). Liming effects on soil pH and crop yield depend on lime material type, application method and rate, and crop species: A global meta-analysis. Journal of Soils and Sediments, 19(3), 1393–1406. https://doi.org/10.1007/s11368-018-2120-2

Liao, P., Huang, S., Zeng, Y., Shao, H., Zhang, J., & van Groenigen, K. J. (2021). Liming increases yield and reduces grain cadmium concentration in rice paddies: A meta-analysis. Plant and Soil, 465(1), 157–169. https://doi.org/10.1007/s11104-021-05004-w

Lu, G., & Ades, A. E. (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine*, *23*(20), 3105–3124. https://doi.org/10.1002/sim.1875

Lu, G., & Ades, A. E. (2006). Assessing Evidence Inconsistency in Mixed Treatment Comparisons. *Journal of the American Statistical Association*, *101*(474), 447–459. https://doi.org/10.1198/016214505000001302

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine*, *21*(16), 2313–2324. https://doi.org/10.1002/sim.1201

Madden, L. V., Piepho, H.-P., & Paul, P. A. (2016). Statistical Models and Methods for Network Meta-Analysis. *Phytopathology®*, *106*(8), 792–806. https://doi.org/10.1094/PHYTO-12-15-0342-RVW

Maki, A., Cohen, M. A., & Vandenbergh, M. P. (2018). Using Meta-Analysis in the Social Sciences to Improve Environmental Policy. In W. Leal Filho, R. W. Marans, & J. Callewaert (Eds.), *Handbook of Sustainability and Social Science Research* (pp. 27–43). Springer International Publishing. https://doi.org/10.1007/978-3-319-67122-2_2

Nakagawa, S., Yang, Y., Macartney, E. L., Spake, R., & Lagisz, M. (2023). Quantitative evidence synthesis: A practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence*, *12*(1), 8. https://doi.org/10.1186/s13750-023-00301-6

Powell, K. I., Chase, J. M., & Knight, T. M. (2011). A synthesis of plant invasion effects on biodiversity across spatial scales. *American Journal of Botany*, *98*(3), 539–548. https://doi.org/10.3732/ajb.1000402

Pullin, A., & Stewart, G. (2007). Guidelines for Systematic Review in Environmental Management. *Conservation Biology : The Journal of the Society for Conservation Biology*, *20*, 1647–1656. https://doi.org/10.1111/j.1523-1739.2006.00485.x

Roberts, P. D., Stewart, G. B., & Pullin, A. S. (2006). Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation*, *132*(4), 409–423. https://doi.org/10.1016/j.biocon.2006.04.034

Rodríguez-Castañeda, G. (2013). The world and its shades of green: A meta-analysis on trophic cascades across temperature and precipitation gradients. *Global Ecology and Biogeography*, *22*(1), 118–130. https://doi.org/10.1111/j.1466-8238.2012.00795.x

Rücker, G. (2012). Network meta- analysis, electrical networks and graph theory. *Research Synthesis Methods*, *3*(4), 312–324. https://doi.org/10.1002/jrsm.1058

Rücker, G., & Schwarzer, G. (2015). Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Medical Research Methodology*, *15*(1), 58. https://doi.org/10.1186/s12874-015-0060-8

Salanti, G., Ades, A. E., & Ioannidis, J. P. A. (2011). Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: An overview and tutorial. Journal of Clinical Epidemiology, 64(2), 163–171. https://doi.org/10.1016/j.jclinepi.2010.03.016

Salanti, G. (2012). Indirect and mixed- treatment comparison, network, or multiple- treatments meta- analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods*, *3*(2), 80–97. https://doi.org/10.1002/jrsm.1037

Salanti, G., & Schmid, C. H. (2012). Research Synthesis Methods special issue on network meta-analysis: Introduction from the editors. *Research Synthesis Methods*, *3*(2), 69–70. https://doi.org/10.1002/jrsm.1050

Slavin, R. E. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*, *48*(1), 9–18. https://doi.org/10.1016/0895-4356 (94)00097-A

Thomsen, M. S., Wernberg, T., Engelen, A. H., Tuya, F., Vanderklift, M. A., Holmer, M., McGlathery, K. J., Arenas, F., Kotta, J., & Silliman, B. R. (2012). A Meta-Analysis of Seaweed Impacts on Seagrasses: Generalities and Knowledge Gaps. *PLOS ONE*, *7*(1), e28595. https://doi.org/10.1371/journal.pone.0028595

Tu, Y.-K. (2014). Use of Generalized Linear Mixed Models for Network Meta-analysis. *Medical Decision Making*, *34*(7), 911–918. https://doi.org/10.1177/0272989X14545789

Vetter, D., Rücker, G., & Storch, I. (2013). Meta-analysis: A need for well-defined usage in ecology and conservation biology. *Ecosphere*, *4*(6), art74. https://doi.org/10.1890/ES13-00062.1

Wang, S., Huang, X., Zhang, Y., Yin, C., & Richel, A. (2021). The effect of corn straw return on corn production in Northeast China: An integrated regional evaluation with meta-analysis and system dynamics. Resources, Conservation and Recycling, 167, 105402. https://doi.org/10.1016/j.resconrec.2021.105402

Whittaker, R. J. (2010). Meta-analyses and mega-mistakes: Calling time on meta-analysis of the species richness–productivity relationship. *Ecology*, *91*(9), 2522–2533. https://doi.org/10.1890/08-0968.1

Zheng, H., Mei, P., Wang, W., Yin, Y., Li, H., Zheng, M., Ou, X., & Cui, Z. (2023). Effects of super absorbent polymer on crop yield, water productivity and soil properties: A global meta-analysis. Agricultural Water Management, 282, 108290. https://doi.org/10.1016/j.agwat.2023.108290

SUPPLEMENT SECTION

1. Figures

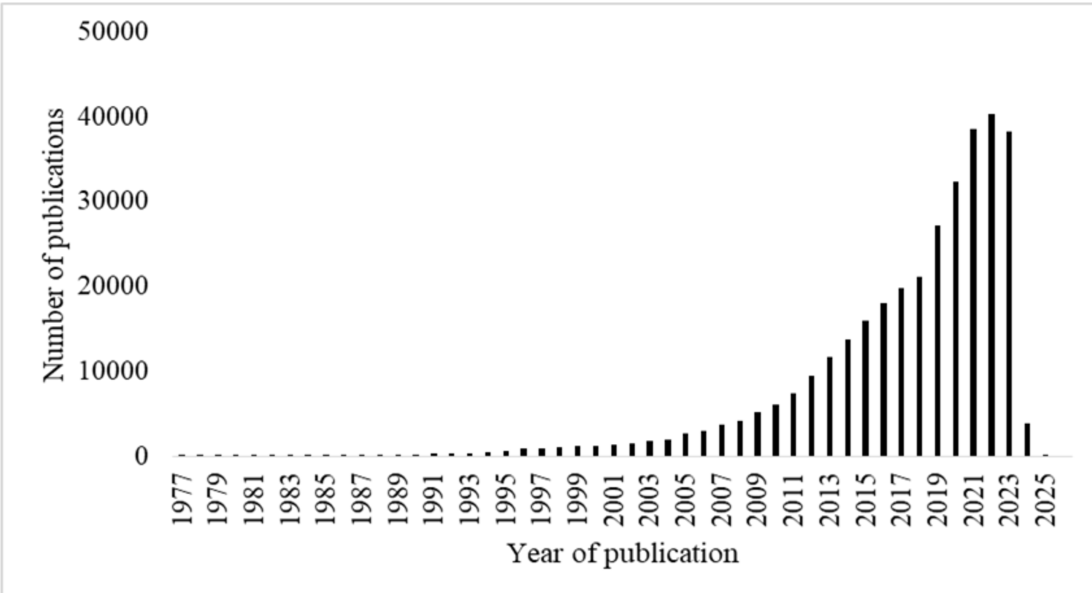

Figure S1. Number of publications on meta-analysis (search web of science using string “meta-analysis” until March 2024). The number of papers using meta-analytical methods has increased exponentially over the years.

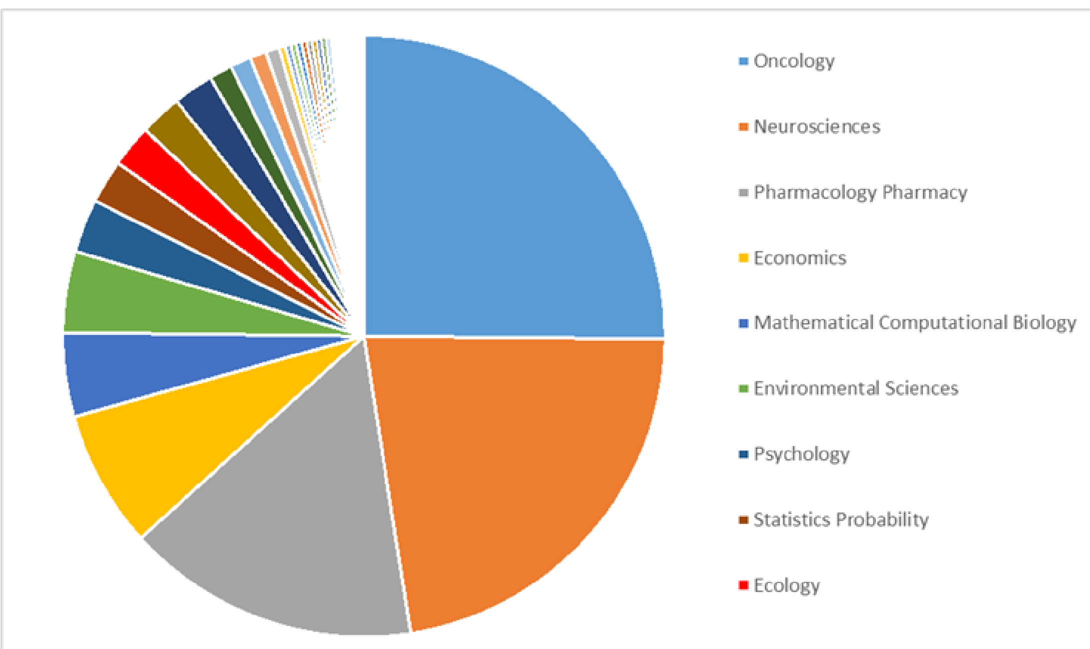

Figure S2. Pie chart showing the percentages of scientific contributions/articles on network meta-analysis according to research area (search web of science using string “network meta-analysis” until March 2024)

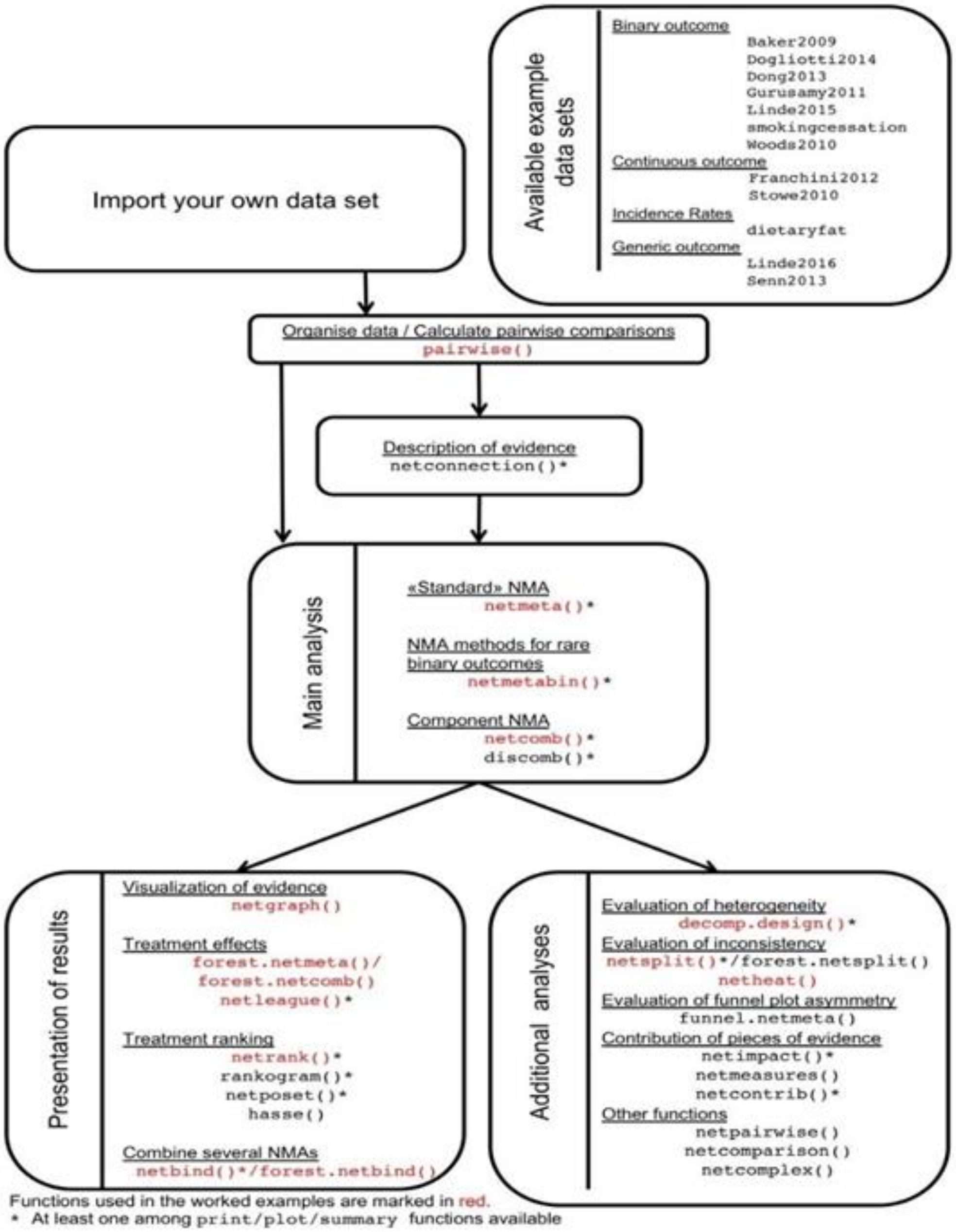

Figure S3. Workflow to perform a (component) network meta-analysis with the R package *netmeta*. Adapted from (Balduzzi et al., 2023)

2. Tables

Table S1. Summary of retrieved datasets with accompanying interventions

Dataset ID	Title of the meta-analytical study paper	Intervention	Citation
D331	Agricultural management strategies for balancing yield increase, carbon sequestration, and emission reduction after straw return for three major grain crops in China: A meta-analysis	Straw return	Liu, D., Song, C., Xin, Z., Fang, C., Liu, Z., & Xu, Y. (2023). Agricultural management strategies for balancing yield increase, carbon sequestration, and emission reduction after straw return for three major grain crops in China: A meta-analysis. Journal of Environmental Management, 340, 117965. https://doi.org/10.1016/j.jenvman.2023.117965

D973	Potential benefits of liming to acid soils on climate change mitigation and food security	Liming	Wang, Y., Yao, Z., Zhan, Y., Zheng, X., Zhou, M., Yan, G., Wang, L., Werner, C., & Butterbach-Bahl, K. (2021). Potential benefits of liming to acid soils on climate change mitigation and food security. Global Change Biology, 27(12), 2807–2821. https://doi.org/10.1111/gcb.15607
D1120	The adaptive capacity of maize-based conservation agriculture systems to climate stress in tropical and subtropical environments: A meta-regression of yields	Straw addition	Steward, P. R., Dougill, A. J., Thierfelder, C., Pittelkow, C. M., Stringer, L. C., Kudzala, M., & Shackelford, G. E. (2018). The adaptive capacity of maize-based conservation agriculture systems to climate stress in tropical and subtropical environments: A meta-regression of yields. Agriculture, Ecosystems & Environment, 251, 194–202. https://doi.org/10.1016/j.agee.2017.09.019

D921A & D921B	Integrated biochar solutions can achieve carbon-neutral staple crop production	Integrated biochar solutions, Straw addition	Xia, L., Cao, L., Yang, Y., Ti, C., Liu, Y., Smith, P., van Groenigen, K. J., Lehmann, J., Lal, R., Butterbach-Bahl, K., Kiese, R., Zhuang, M., Lu, X., & Yan, X. (2023). Integrated biochar solutions can achieve carbon-neutral staple crop production. Nature Food, 4(3), 236–246. https://doi.org/10.1038/s43016-023-00694-0
D309	Improving yield and nitrogen use efficiency through alternative fertilization options for rice in China: A meta-analysis.	Green manure, Organic fertilizer, Secondary and micronutrient fertilizer, Slow release fertilizer, Straw return	Ding, W., Xu, X., He, P., Ullah, S., Zhang, J., Cui, Z., & Zhou, W. (2018). Improving yield and nitrogen use efficiency through alternative fertilization options for rice in China: A meta-analysis. Field Crops Research, 227, 11–18. https://doi.org/10.1016/j.fcr.2018.08.001

D669	Effects of the Ratio of Substituting Mineral Fertilizers with Manure Nitrogen on Soil Properties and Vegetable Yields in China: A Meta-Analysis	Substituting mineral fertilizers with manure nitrogen	Wang, S., Lv, R., Yin, X., Feng, P., & Hu, K. (2023). Effects of the Ratio of Substituting Mineral Fertilizers with Manure Nitrogen on Soil Properties and Vegetable Yields in China: A Meta-Analysis. Plants, 12(4), Article 4. https://doi.org/10.3390/plants12040964
D473	A global meta-analysis of cover crop response on soil carbon storage within a corn production system	Cover crops	Joshi, D. R., Sieverding, H. L., Xu, H., Kwon, H., Wang, M., Clay, S. A., Johnson, J. M., Thapa, R., Westhoff, S., & Clay, D. E. (2023). A global meta-analysis of cover crop response on soil carbon storage within a corn production system. Agronomy Journal, 115(4), 1543–1556. https://doi.org/10.1002/agj2.21340

D652	Effects of super absorbent polymer on crop yield, water productivity and soil properties: A global meta-analysis	Super absorbent polymer (SAP)	Zheng, H., Mei, P., Wang, W., Yin, Y., Li, H., Zheng, M., Ou, X., & Cui, Z. (2023). Effects of super absorbent polymer on crop yield, water productivity and soil properties: A global meta-analysis. Agricultural Water Management, 282, 108290. https://doi.org/10.1016/j.agwat.2023.108290
D906	Assessment of drainage nitrogen losses on a yield-scaled basis	Drainage nitrogen losses	Zhao, X., Christianson, L. E., Harmel, D., & Pittelkow, C. M. (2016). Assessment of drainage nitrogen losses on a yield-scaled basis. Field Crops Research, 199, 156–166. https://doi.org/10.1016/j.fcr.2016.07.015
D352	Effect of soil erosion depth on crop yield based on topsoil removal method: a meta- analysis	Topsoil removal experiment	Zhang, L., Huang, Y., Rong, L., Duan, X., Zhang, R., Li, Y., & Guan, J. (2021). Effect of soil erosion depth on crop yield based on topsoil removal method: A meta-analysis. Agronomy for Sustainable Development, 41(5), 63. https://doi.org/10.1007/s13593-021-00718-8