University of Groningen

Systematisch èn snel wetenschappelijk literatuuronderzoek t.b.v. beleid

Sijtsma, Frans; Vreeling, Liselotte; Angelstam, Per ; El- Hacen, Hacen; van der Heide, Martijn; Horlings, Ina; Piersma, Theunis; Tittonell, Pablo

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Sijtsma, F., Vreeling, L., Angelstam, P., El- Hacen, H., van der Heide, M., Horlings, I., Piersma, T., & Tittonell, P. (2021). *Systematisch èn snel wetenschappelijk literatuuronderzoek t.b.v. beleid: Een verkenning van het gebruik van ASReview rond de effectiviteit van beleidsinstrumenten die duurzame landbouw stimuleren*. University of Groningen.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 21-01-2023

university of groningen

Systematic and fast scientific literature review for policy purposes

An exploration of using ASReview software about the effectiveness of policy instruments promoting sustainable agriculture

University of Groningen
Groningen, December 2021

By Dr. Frans J. Sijtsma and MSc. Liselotte H. Vreeling

In collaboration with Prof. Per Angelstam, Hacen El-hacen, PhD, Prof. Martijn van der Heide, Prof. Ina Horlings, Prof. Theunis Piersma & Prof. Pablo Tittonell.

In cooperation with the Netherlands Environmental Assessment Agency (Planbureau voor de Leefomgeving).

Systematic and fast scientific literature review to the benefit of policy

An exploration of using ASReview software about the effectiveness of policy instruments promoting sustainable agriculture.

Groningen, December 2021.

Dr. Frans J. Sijtsma en Liselotte H. Vreeling, MSc.
University of Groningen
In collaboration with Prof. Per Angelstam, Hacen El-hacen, PhD, Prof. Martijn van der Heide, Prof. Ina Horlings, Prof. Theunis Piersma & Prof. Pablo Tittonell.
In cooperation with the Netherlands Environmental Assessment Agency (Planbureau voor de Leefomgeving).

Summary

In the context of Article 3.1 of the compatibility law 2016 (Comptabiliteitswet) (CW 2016), which aims to bring about more scientific underpinning of policy, the House of Representatives in 2019 examined the extent to which policy proposals address the efficiency and effectiveness of the policy instruments to be deployed (Sneller & Snels, 2019). They concluded that while they often address **how** these instruments should contribute to the objective, they often do not address **to what extent**.

Making such statements about efficiency and effectiveness requires an understanding of scientific evidence. As a result, the Netherlands Environmental Assessment Agency (PBL) anticipates that they will receive more frequent questions from the government about what science says about the effectiveness of proposed policies. PBL already has a strong scientific orientation in its work and recognizes the great importance of a request for scientific foundation. But is also foresees that answering this type of question about scientific evidence for policy proposals can be difficult in practice. This is due to the short period in which these questions must generally be answered on the one hand and the time required for (systematic) literature review on the other. Conducting a (systematic) literature search using artificial intelligence (AI) could offer a solution to this problem and that is what this report is about.

When using AI, the literature search is supported by a so-called learning algorithm, which learns as the process progresses to better and better assess what is relevant literature for the researcher. However, little experience has been gained with the use of such software in policy-oriented research. PBL therefore asked the UG to investigate how the open-source AI software ASReview could offer a solution in efficiently meeting the demand for scientific insights from the government. At PBL's request, the UG researched three substantive questions around the effectiveness of policy instruments for sustainable agriculture and used these to test the process of the AI supported literature review. In this research, the UG worked with a combination of AI-supported literature screening and sounding boards of academics who fed the search at the beginning and interpreted the results substantively at the end. Such a sounding board enables the distillation of substantive lessons in a relatively short time.

The findings are:

1. It is possible to conduct a quick and good systematic literature search by combining AI-supported literature search with a sounding board of scientific experts. The search for scientific insights that could potentially provide substantiation yielded a database of 40.000 potentially relevant papers, from which a diverse set of 100 relevant papers were selected using AI. Using the sounding board of experts, an even smaller set of 12 papers was created from this list that were deemed most urgent for policymakers to study.
2. Because the AI software ASReview puts content first and hides reputation of authors and journals during screening, objectivity and breadth is stimulated;
3. However, ASReview does present other challenges that, if not taken into account, can compromise the objectivity of research in other ways;
    i) One of the main risks of using ASReview is what we call ‘trap formation’, especially when there is a short time frame for screening. This means that you

end up on a particular 'track' of articles on a particular (sub)topic, which means that other also relevant articles are not found. The broader the query and the less time, the more this poses a risk to the efficiency and reliability of the screening. By using certain settings for the screening, this can be taken into account to a certain extent.

ii) The amount of efficiency gains that can be achieved with ASReview depends on the breadth and multidisciplinarity of the research question. The topic of sustainable agriculture has many facets, both in terms of instruments and outcomes. This makes it more difficult for the program to quickly learn what is most relevant, so screening will take more time;

iii) AI is not a fully automated process. Use of the program requires skill from the researcher to drive the algorithm and expert knowledge to start the process and expert knowledge to interpret the results.

4. The content of this research has resulted in two sets of scientific articles: a Top 100 and a Top 12.
5. The first end result of the Systematic Review is a list (Top 100) of relevant articles ranked by their "scientific recognition." Here, because of the requirement for speed in the process, scientific recognition is simply operationalized as a combination of the number of citations of the article per year and the impact factor of the journal (as a measure of the seriousness of the blind review process) in which they were published. While there are caveats to this method of ranking, it does provide a way to make a large amount of knowledge manageable within a relatively short period of time, and in doing so, gives policymakers and researchers a foothold that they can study "the most important first".
6. The second end result is a selection from the top 100 by the scientific experts: which ones are most important for policy? The six multi-disciplinary scholars each selected three papers that they felt were most important for policymakers and researchers to take to task: combined, this yielded a set of 12 articles. The Top 12 is a manageable set of articles that was selected quickly, and can be studied in content in a short period of time while being largely systematic.
7. In conclusion. What does the process tested here offer compared to what we might call "the standard quick search for scientific evidence" of a PBL staff member? This search often consists of manually consulting Google Scholar and/or individually contacting a scientific expert. Compared to manual searches via Scholar, searching the literature with ASReview offers the opportunity to systematically review a vast amount of studies for relevance. After each selection by the researcher, the entire database is reordered. Ultimately, it has been shown, this leads to a greater diversity of studies than a manual search process. Compared to contacting experts individually, the added value of utilizing a group of experts in this study is not only their diversity of expertise, but also that they all reflect on the same scientific dataset and choose from it (quickly and rationally) the most relevant ones for policy. This is a much more systematic process than asking for their scientific views separately.

Table of Contents

Summary 3
1. Introduction 7
1.1 Rationale 7
1.2 Research goal 7
1.3 The research team 9
1.4 The research approach 9
2. Methodology 11
2.1 What is ASReview? 11
2.2 What do you need to use ASReview? 12
2.3. What we have used for this research 12
3. Results 16
3.1 Process-oriented results 16
3.1.1. Broad research question (1) 16
3.1.2.1. Narrower research question (2) 17
3.1.2.2. Narrower research question (2), 95% max + 5% uncertainty 17
3.1.3. Narrowest research question (3), 95% max + 5% uncertain 18
3.2. Content Results 20
3.2.1 A list with relevance-ranked papers 20
3.2.2 The expert result: A selection by experts of the 12 most important papers 21
4. Recommendations for and comments on ASReview 24
4.1 General 24
4.2 Search strategy 24
4.3 Abstract-screening with the help of AI 25
4.4 Objectivity and transparency 25
5. Conclusions 25
6. References 28
7. Appendices 30
Appendix A – Search query 30
Appendix B – The 42 'key articles' supplied by the sounding boards 31
Appendix C – Quick Guide ASReview 34
Appendix D – The top 100 papers 37
Appendix E – The instructions for the experts in the sounding board groups. 41
Appendix F – List of 15 articles found for the broadest research question 43
Appendix G – 35 relevant articles for the smallest (3) research question 44
Appendix H – Abstracts of Top 12 47

1. Introduction

1.1 Rationale

The Netherlands environmental assessment agency (PBL) asked the UG to conduct a rapid systematic literature review (SR) in approximately three months to test the use of the open-source AI software ASReview in the context of policy-oriented research. PBL wants to gain insight into the scientific knowledge on the effectiveness of various policy instruments that can be used to make agriculture and agricultural land use more sustainable. This insight into the scientific knowledge is particularly important for providing policymakers with the best possible substantiated policy advice in the short term.

Unfortunately, the process of systematically reviewing scientific literature is usually time-consuming, making it difficult in practice to provide sound advice within a short period of time. As a result, as the meeting held at the start of this project revealed, many PBL researchers do not systematically search for literature, but often use articles they are already familiar with, and search for new literature in online databases such as Google Scholar. This sometimes leaves them with the impression that they are not basing their studies on the most relevant or highest quality literature, which could compromise the reliability and objectivity of their advice. In addition, it is not very easy to conduct systematic research in the context of the greatly increased number of scientific publications.[1] All this means that, especially in the case of multidisciplinary subjects, it is not easy to obtain a complete and up-to-date overview of relevant studies, which complicates the goal of providing policymakers with sound advice in an efficient manner.

PBL hopes to achieve the combination of efficiency and depth through the use of the software ASReview. ASReview (https://asreview.nl/) is an "open source machine learning framework for efficient and transparent systematic reviews" (Van de Schoot et al., 2021, p. 125) developed by researchers at Utrecht University, intended to contribute to the efficiency, but also to the transparency and objectivity of SRs. However, the software is relatively new and there is little experience with its use for policy-oriented research.

1.2 Research goal

This research by the UG deals with the usability of the open-source AI software ASReview for doing systematic literature reviews applied to the field of sustainability-oriented agricultural policies. It therefore has a dual purpose. In terms of content, this research aims to provide insight into the effectiveness of instruments to stimulate more sustainable agriculture. Process-wise, the study aims to gain experience with ASReview-supported systematic literature review on such a topic.

[1] For example, the Google Scholar search "effectiveness agricultural policy instruments" returns just under 400,000 hits, including 12,000 in the year 2021.

The process-oriented research question is:

> To what extent and in what way is it possible to use the Artificial Intelligence software ASReview to substantially shorten the process of systematic review of scientific literature?

Given the exploratory nature of the study, the substantive research question was refined at two points based on the experience gained with ASReview. The research question that has been answered most fully is:

> What is known in the scientific literature about the effectiveness of subsidies that are or can be used in Western countries to make agriculture and agricultural land use more sustainable?

Around this fully answered question, we experimented with two other substantive questions: one that is broader (not just "subsidies" but "the palette of policy instruments") and one that is more specific (not "sustainable agriculture" but "reduction of greenhouse gas and nitrogen emissions to air in agriculture").

The initial approach was to create an Evidence Gap Map (EGM) based on the literature found. In such a map, the evidence is categorized by instrument and outcome, in order to be able to see at a glance which instruments and outcomes there is more or less scientific evidence about (Lopez-Avila et al., 2017). While ASReview can indeed speed up doing literature reviews, the process prior to title-and-abstract-screening in ASReview in particular still takes a fair amount of time. So much so that during the course of the project, it quickly became clear that it was not feasible to complete a full SR, including full-text screening and content synthesis, within the relatively short duration of this project. As a result, it is not possible to make confident statements about the amount of scientific evidence for each of the tools/outcomes; and thus the possibility of an Evidence Gap Map is dropped. At the same time, we see a possible disadvantage to EGMs: although they also present the substantive scientific insights, they are not the main topic in the Map. This is actually a shortcoming, because that is precisely why the review takes place; policy makers want to strengthen their substantiation by listening to as many and as broadly as possible scientific insights.

There are two important features of EGMs that we adopt as the end result of our review, but in a modified form. EGMs present the most relevant literature in an orderly format and they usually also give an indication of the degree of reliability of the available studies. We take these two aspects, overview and reliability, as inspiration for the alternative form of synthesis.

Since it was not feasible within the time frame to synthesize the results in an EGM, we searched for another form of synthesis in which the content of available scientific evidence can be learned relatively easily. As a substantive outcome, this study presents a two-fold selection of scientific papers : 1) ranking of relevant studies and 2) selection of the most important articles by experts.

The first list is a ranked list of papers found in ASReview (Top 100), ranked by the reading urgency for policy makers. The list was created as an MS Excel file and contains the titles, abstracts and digital links to the scientific papers. The ranking was made based on objective indicators in terms of recognition within science (citations weighted for year and impact factorpr of the journals). The second is a list of articles made by experts. Experts have made a selection of the most important papers from the list of papers. The combination of ranking and selection allows the enrichment of knowledge by science to take place efficiently and effectively. The policy maker with very little time focuses on the selection of articles by experts or on the top of the ranked list. The policy maker with more time can delve into the entire list of abstracts and links to articles. In doing so, the ranking in terms of scientific recognition always provides guidance that the extra time spent is not arbitrary, but when working from top to bottom, one first takes note of those insights that have been taken most seriously scientifically: most cited and most seriously reviewed.

1.3 The research team

This research was carried out by a core team of two researchers in collaboration with two sounding board groups: a sounding board of scientists from the Sustainable Landscapes research group at the UG and a sounding board of policy researchers from PBL. These contributed to the study at regular intervals. This is explained in the next section of this chapter.

The core team of researchers consisted of: Frans Sijtsma and Liselotte Vreeling.
The academic sounding board was formed by a selection of Sustainable Landscapes fellows from the UG: Professor Ina Horlings, Professor Per Angelstam, Professor Martijn van der Heide, Professor Pablo Tittonell, and Professor Theunis Piersma / PhD. Hacen El Hacen.

The sounding board group of PBL researchers consisted of: Jetske Bouma and Jarry Porsius, together with Jan van Dam, Petra van Egmond, Stefan van der Esch, Sonja Kruitwagen, Gusta Renes, Stefan Troost, Rob Weterings.

1.4 The research approach

The research team divided the project into several steps. This is visualized in Figure 1. Starting with the research question, we went through 10 steps to get to the substantive result. As mentioned earlier, the research process was one of advancing understanding. This means that the research question was adjusted twice based on the experience gained with ASReview. A total of five screenings based on three research questions were conducted in ASReview (Step 5).

The sounding boards described above contributed to the study at four points:

- Sounding board of PBL researchers: a starting workshop and a meeting around the draft final report
- Both sounding boards: when delivering the so-called key articles on the subject, as a starting point for drawing up the search strategy and to be used as a starting point in ASReview (early October);

- Scientific sounding board: when assessing/ranking the final set of (potentially) relevant literature based on their own expertise.

In addition, Jetske Bouma and Jarry Porsius provided feedback and tips individually and together several times.

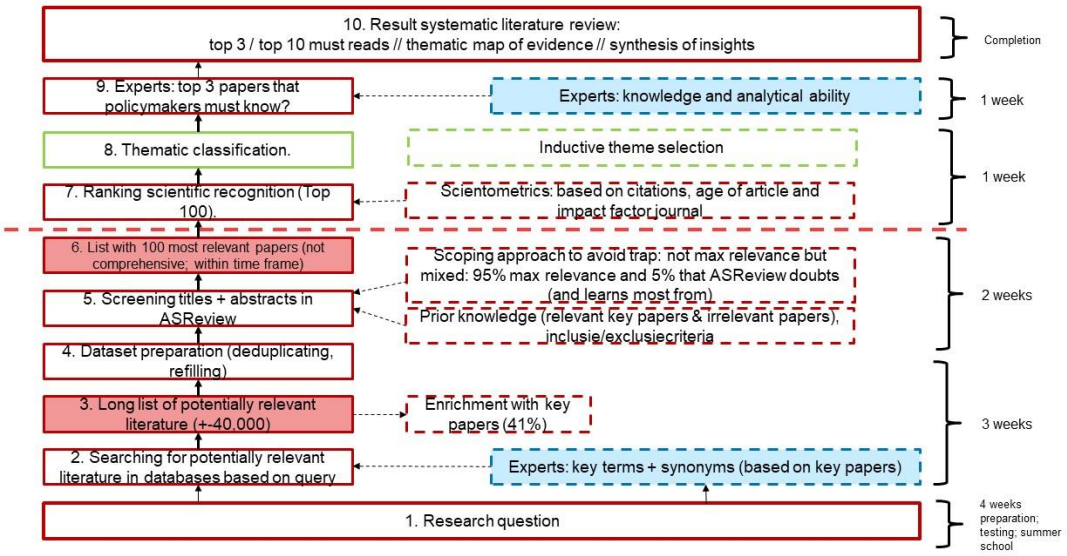


Figure 1 – visualization of the research process and time indication per step.

2. Methodology

2.1 What is ASReview?

According to Abrami et al. (2010, p.373), there is a "growing need not just for comprehensive and systematic reviews of evidence, but also for high-quality brief reviews. ASReview is open-source software developed by researchers at Utrecht University. The program is designed to use active machine learning to make the screening of titles and abstracts for an SR more efficient and accurate. The combination of AI and a researcher holds the promise of being able to screen (the titles and abstracts of) a large number of publications more efficiently, transparently and, moreover, objectively. This allows for a much quicker arrival at a limited set of articles to include in full-text screening compared to a manual SR (Van de Schoot et al., 2021). Especially for an institute like PBL, which wants to be able to provide policy makers with sound advice in a short period of time, these promises make the program interesting to explore further. At the same time, AI-assisted screening of literature is a new development, with which there is still little experience in the context of policy-oriented research.

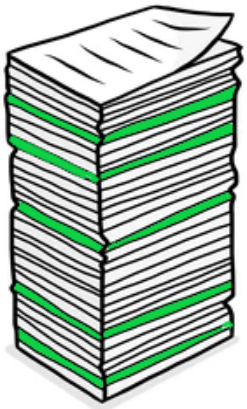


Figure 1 – Visualization of manual SR

How does this software work? Briefly, what ASReview does is calculate which article in the dataset has the highest probability of relevance based on the choices the researcher has already made (which articles are relevant or irrelevant). Each time during the screening process, the researcher is shown the title and abstract of the article that currently has the highest relevance probability. The researcher screens these and then makes the decision whether this article is indeed relevant or still irrelevant. Based on this information, ASReview recalculates which article has the highest probability of relevance, and the researcher is shown the title and abstract again, etc. The screening process thus consists of an interaction between AI and researcher: AI calculates the relevance probability and determines which article the researcher will see, but the researcher decides whether it is actually relevant. Unlike so-called 'black box' algorithms, in which the algorithm itself decides what is and is not relevant, this is therefore a transparent and interactive process in which the researcher ultimately makes that choice (Van de Schoot et al., 2021).

Figure 2 - Visualization of the ordering made by ASReview, based on relevance- probability

Thus, unlike a manual SR, in which the researcher screens each article "on the stack" piece by piece (Figure 2), it is as if the stack is ordered after each article screened: the literature with a high(er) relevance probability always ends up on top, while the literature with a low(er) probability ends up at the bottom of the stack (Figure 3). Moreover, with each choice made (relevant or irrelevant), the algorithm becomes "smarter"; it learns to better and better assess which articles are relevant and which are not. In this way, without having to screen the entire stack of potentially relevant literature himself, the researcher arrives at a set of articles to screen full-text. Depending on the research question, the size of the dataset and the starting articles chosen, the use of ASReview can save up to 95% of the time that would be needed to perform a manual SR (Van de Schoot et al., 2021).

In addition to these time savings, ASReview has several other advantages. First, the software is open-source, which means that anyone can install and use the program, and all codes and manuals are public. Second, the program only shows the title, DOI and abstract. This contributes to the objectivity of the SR. When a researcher herself is well known in the field, if an article was written by a particular author and/or published in a particular journal, she may be inclined to quickly believe that "this article would probably be relevant. Because this information is not seen in ASReview, each article is reviewed purely on the basis of the title and abstract. The quality control of these articles is done in the step after ASReview, where the full-texts are screened.

The screening of titles and abstracts in ASReview continues until a certain "stopping point" is reached. When this point is reached cannot be predicted in advance but is determined by the researcher(s). In general, the process is stopped when the number of articles offered is increasingly marked as irrelevant. A brief guide to using ASReview is provided in Appendix C.

2.2 What do you need to use ASReview?

In preparation for screening in ASReview, the following is required:

1. A clearly stated, delineated research question;
2. Clear inclusion and exclusion criteria;
3. A number of "key" articles on the research topic to screen for keywords (item 4a) and to use as a starting point (item 5);
4. A prepared dataset in .ris or .csv format;
    a. A search strategy using keywords;
    b. Searching one or more online databases based on the search strategy;
    c. Exporting the dataset to a reference manager (e.g. EndNote);
    d. Deduplicate and prepare the dataset;
5. 1 to 5 relevant and 1 to 5 irrelevant articles from the dataset as a starting point for learning ("prior knowledge")

2.3. What we have used for this research

2.3.1. A clearly stated, delineated research question

ASReview is best (i.e., most reliable and efficient) used with a narrower research question. This is because the program works on the basis of a learning algorithm. This means that the program constantly learns what you are looking for, based on the choices you make. Herein lies, especially when there is little time available, also one of the biggest challenges in using the program: trap formation. If the research question is broad, i.e. if many different (sub-)topics are relevant, then the algorithm can ensure that you only get and stay on one track, because you cannot indicate that you are also interested in other things. When the track on this subject is finished, articles on other relevant subjects can be found. It is just not possible to say in advance when such a track will be finished and when articles on the next subject will emerge. This reduces the efficiency gain of using ASReview compared to a manual SR in such a case.

Given the breadth of the topic and the limited time in this project, broader and narrower research questions were used to test how trap formation works and how best to deal with it.

The three research questions screened for are:

1. What is known in the scientific literature about the effectiveness of the palette of instruments that are or can be used in Western countries to stimulate sustainable agriculture and sustainable agricultural land use?
2. What is known in the scientific literature about the effectiveness of subsidies to stimulate sustainable agriculture and agricultural land use in Western countries?
3. What is known in the scientific literature about the effectiveness of financial resources to reduce greenhouse gas and nitrogen emissions to the air in the agricultural sector?

2.3.2. Clear inclusion and exclusion criteria

As with a manual SR, in order to systematically assess the literature as relevant or irrelevant, a number of inclusion and exclusion criteria must be established based on the research question. Because of the experimental approach, a total of three research questions were used. The inclusion and exclusion criteria were adjusted each time to fit the relevant research question. The inclusion and exclusion criteria used in this project are:

Inclusion criteria:

- Evaluates/assesses policy instrument
- Sustainable agriculture and land use;
- In western countries;
- Grants, AES, finance, compensation;[2]
- Reduction of greenhouse gas emissions and nitrogen to air.[3]

Exclusion criteria:

- Greenhouse horticulture, biofuel, fisheries, aquaculture;
- Deals broadly with sustainable agriculture, but not specifically with (the effectiveness of) instruments and/or outcomes;
- Does not evaluate, but advocates or proposes model or analysis framework;
- "How does X affect the effectiveness of subsidy";
- Non-western country, tropical country, developing country;
- Health-related articles (e.g., obesity);
- Women's empowerment and development assistance.

2.3.1. A number of key articles

To be able to carry out the next two steps, a number of articles are needed that are already certain to be relevant to the research. Preferably, leading ('key') articles on the subject are used for this. These articles are then scanned for keywords and synonyms of these words. Moreover, they can be used as a starting point for the learning process (see 2.3.5).

[2] This criterion is used only for the second and third research questions.
[3] This criterion is used only for the third research question.

This study used key articles supplied by the two expert sounding boards. A total of 42 articles were provided. These were screened for relevance, keywords and synonyms of these keywords. Some of these articles were found to be irrelevant to the topic during the screening process after all. For example, Ferraro (2011) and Miteva et al. (2012) were not about agriculture, but about nature conservation. Second, it took a relatively long time for all articles to be delivered. Based on the keywords from these articles, the search strategy (2.3.4.1) was created. An overview of these articles can be found in Appendix B.

2.3.2. A prepared dataset in .ris or .csv-format

For the reliability of the research, it is important that the compilation of the dataset is done precisely. After all, if the query used does not contain all possible (combinations of) keywords, you could miss out on relevant literature.

What is known in the scientific literature about the effectiveness of the palette of instruments that are or can be used in western countries to make agriculture and agricultural land use more sustainable?

Domain	**Determinant (instrument)**	**Outcome**
Policy	Effect	Sustain*
Govern*	Assess*	“Environment* friendly”
Institut*	Evaluat*	Climate
CAP	Eviden*	Biodiversity
“Common Agricultural Policy”	Perform*	Conservation
Regulation	Implement*	Agricultur*
Regulatory		Agri
Law		Farm
Legislation		Ecosystem
Instrument		“agro-ecosystem”
Measure		Agroecosystem
Intervention		“ecosystem service”
arrangement		Cropland
		Grassland
		Livestock
		Cattle
		cereal

Table 1 - The DDO used to compose the dataset

2.3.2.1. Keyword search strategy

First, a search strategy must be drawn up (DDO/PICOS, see Table 1), containing the key words of the research question and synonyms for these key words. In this study, we searched for these synonyms by scanning the key articles provided by the two focus groups (step 2.3.3).

2.3.2.2. Create a search query for an online database

Based on the above DDO, a search query was prepared (see Appendix A). This was used to search the online database Scopus for literature on October 26, 2021 and found 40,223 potentially relevant articles.

2.3.2.3. Exporting as .ris in a reference manager such as EndNote

After this, the found literature must be exported. This is a time-consuming and labor-intensive task that is also prone to errors. In Scopus, a maximum of 2,000 articles can be exported at once, and for each "batch" it must be checked whether all the necessary information is exported (in any case author, DOI, title, journal, page numbers, keywords and abstract). I did this by exporting the articles by year (and further broken down by research field if necessary). If you make a mistake in exporting, it will affect the reliability of the study. Therefore, it is important to report this properly.

2.3.2.4. Deduplicate and complete missing information

For this study, literature was only searched in Scopus. There can always be duplicates in the dataset, especially when the dataset is compiled from multiple online databases. For irrelevant articles this matters less than for relevant articles, but it is nevertheless advisable to deduplicate the dataset (this also applies to manual SR, by the way). If a relevant article is duplicated (or perhaps even more so) in the dataset, the researcher will see it three times in a row in ASReview and thus label it as relevant three times in a row. This causes ASReview to overestimate the value of the terms in these articles; that is, ASReview does not see that it is exactly the same article that is labeled as relevant three times in a row, but only sees that the terms in it are labeled as relevant three times in a row. Second, ASReview needs at least a title, abstract and keywords from each article. Thus, it is important to clean up the dataset before screening in ASReview can begin. Reference managers such as EndNote offer the ability to deduplicate datasets and fill in missing information. However, this is time-consuming. For example, EndNote can search for missing information for ±250 articles per hour.

2.3.3. 1 to 5 relevant and 1 to 5 irrelevant articles from the dataset as a starting point for learning ("prior knowledge")

In the case of this study, the dataset included articles on obesity or emancipation, and articles focused on agriculture in tropical areas or developing countries; these were known in advance to be irrelevant to this study. Thus, in order to quickly track down these irrelevant articles, it pays to have an idea of approximately what is in the dataset. When this is not the case, ASReview also provides an option to search random articles. Since there is probably much more irrelevant literature in the dataset, it is easy to add irrelevant literature this way. You only have less influence as a researcher on what you add as irrelevant literature. Therefore, we recommend selecting the irrelevant literature based on the inclusion and exclusion criteria.

The choice of 1-5 articles that are known in advance to be relevant is more difficult, especially with a broad research question like this one. These should be articles of which it is certain in advance that they are relevant to the research question and articles that do not deal with a too specific subject, as this could push the algorithm into a certain 'trap' while other subjects are also relevant. In this study we used the key articles provided by experts as prior knowledge. These were screened for title, abstract and broad content to determine which were appropriate.

3. Results

3.1 Process-oriented results

Due to the exploratory nature of the study, we always screened in ASReview based on insights obtained on a new research question or otherwise. In total, we screened five times with three research questions. Initially, the principal investigator undertook the screening. However, the second researcher was available to provide input or feedback when in doubt.

3.1.1. Broad research question (1)

What is known in the scientific literature about the effectiveness of the palette of instruments that are or can be used in western countries to make agriculture and agricultural land use more sustainable?

In the exploratory phase of the project, in September 2021, a trial was run based on the above research question and the initial inclusion and exclusion criteria (found in 2.3.3). This trial was done to provide input for the kick-off meeting. The search strategy, dataset and priors were pragmatically designed and chosen for this trial. That is, the dataset, consisting of 17,453[4] articles, was most likely incomplete, as not all keywords and possible synonyms were included in the search strategy. A total of one working day was screened. Of the 63 articles screened, 15 were identified as (potentially) relevant.

As a starting point for this first trial, two articles were specified as relevant. These were found by searching the dataset for "sustainable agriculture" and "instrument" and then screening the abstracts of the articles:

- Catarino, Gaba & Bretagnolle (2019). Experimental and empirical evidence shows that reducing weed control in winter cereal fields is a viable strategy for farmers;
- McElwee *et al.* (2020). The impact of interventions in the global land and agri-food sectors on Nature's Contributions to People and the UN Sustainable Development Goals.

Soon, the researcher noticed that the priors specified and the first few choices have a major impact on how the algorithm learns. In this case, because the two articles used as a starting point were about "weed control" (a specific part of sustainable agriculture) and "sustainable development goals" (a buzzword), the screening quickly found itself in a trap. It is not possible to communicate to the program that other topics (in this case: other tools and outcomes) are also relevant. Therefore, the test was run again, with a larger number and less specific items as a starting point.

For the second trial, five articles were selected as relevant starting points:

- Howden *et al.* (2007). Adapting agriculture to climate change.
- Pretty (2008). Agricultural sustainability: Concepts, principles and evidence.

[4] Found via Scopus on 20-09-2021 with the search string: (TITLE-ABS-KEY (policy OR govern* OR institut* OR measure OR intervention OR instrument)) AND (TITLE-ABS-KEY (agricultur* OR farm* AND sustainab*)) AND (TITLE-ABS-KEY (assessment OR assess* OR effect* OR evidence)).

- Springmann *et al.* (2018). Options for keeping the food system within environmental limits.
- Torralba *et al.* (2016). Do European agroforestry systems enhance biodiversity and ecosystem services? A meta-analysis.
- Van Passel *et al.* (2007). Measuring farm sustainability and explaining differences in sustainable efficiency.

A total of one working day was spent screening for this research question. During this time, 63 articles were screened, of which 15 were marked as relevant. Although care was taken to ensure that the priors were more general in their selection, the screening still seemed to remain primarily on one track: that of agroforestry, and particularly in the context of Mediterranean countries. See Appendix F for the fifteen articles selected.

3.1.2.1. Narrower research question (2)

What is known in the scientific literature about the effectiveness of subsidies to stimulate sustainable agriculture and agricultural land use in Western countries?

After the kick-off meeting, a new dataset consisting of +- 40,000 articles was prepared (see 2.3.4). This dataset was used for screening for this research question and the next research question. As a starting point for screening for this question, five articles were chosen as relevant based on their title, abstract and broadly on their content:

- Batáry *et al.* (2015). The role of agri-environment schemes in conservation and environmental management.
- Daniel & Perraud (2008). The multifunctionality of agriculture and contractual policies. A comparative analysis of France and the Netherlands.
- Migliorelli & Dessertine (2017). Time for new financing instruments? A market-oriented framework to finance environmentally friendly practices in EU agriculture.
- Piñeiro *et al*. (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes.
- Pretty (2008). Agricultural sustainability: Concepts, principles and evidence.

Approximately sixteen hours of screening was conducted (128 articles in total; of which 91 were included). Despite screening only for grants, there was still a trap. Through the topics of agri-environmental schemes (AES) and biodiversity, the program quickly arrived at birds. This in itself is not irrelevant, but again very specific. After screening about 50 articles, the 'bird pile' seemed to be running out and other articles on AES and biodiversity were also shown.

In addition, it was noticed that there were quite a few articles (on a limited topic) being included; a signal that something was not quite right. Therefore, the starting articles were slightly modified to more general articles about the effectiveness of subsidies and a different query strategy was applied. The results of this are discussed in the next section.

3.1.2.2. Narrower research question (2), 95% max + 5% uncertainty

What is known in the scientific literature about the effectiveness of subsidies to stimulate sustainable agriculture and agricultural land use in Western countries?

As a starting point for this screening, five articles from those provided by the expert sounding boards were chosen as relevant. Three of these are the same as the articles used for the screening in Section II; two are additional:

- Batáry *et al.* (2015). The role of agri-environment schemes in conservation and environmental management.
- Parra-López *et al*. (2009). An integrated approach for ex-ante evaluation of public policies for sustainable agriculture at landscape level.
- Piñeiro *et al.* (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes.
- Plieninger *et al.* (2012). Mainstreaming ecosystem services through reformed European agricultural policies.
- Pretty (2008). Agricultural sustainability: Concepts, principles and evidence.

The 95% maximum and 5% uncertainty setting means that 95% of the time the researcher will see the article that has the highest relevance probability, and 5% of the time the article that ASReview is currently most uncertain about (i.e., the article whose P is closest to 0.5). In this way, you can partially overcome the trap effect, since you will occasionally see an article that is not on top of the pile. Moreover, the algorithm can learn the most from the choice made about the articles about which it has the most doubts. Compared to the screening described in the previous section (II), this setting yielded a greater diversity of relevant literature. For example, although most of the articles were still about AES and less about other forms of agricultural subsidies, studies on birds emerged as well as other forms of biodiversity.

The 'stopping point' was chosen pragmatically for this screening. In view of the remaining duration of the project, it was decided to devote a maximum of five working days to screening. In this time 400 articles (1% of the dataset) were screened, of which 100 were included. As visible Figure 4, this was a point where there were already gaps in the relevance screening: a longer time with irrelevance of the articles offered by the AI as most relevant. These 100 included articles were used for the content synthesis of this study, found in Chapter 3.2. Thus, the stopping point was chosen not only on the basis of the data, but more importantly on the basis of time. Given the statistics, it is quite possible that there were still relevant articles 'in the pile' (see Figure 7). Therefore, it cannot be said with certainty that all (potentially) relevant articles were extracted from the dataset. This is partly due to trap formation and partly due to the chosen stopping point. It should be emphasized here that both of these things are related to the breadth of the research question: the broader the question, the harder it is for the algorithm to learn what is relevant, and the more time it takes to get to a point where it can be said with some certainty that (as good as) all relevant literature has been found.

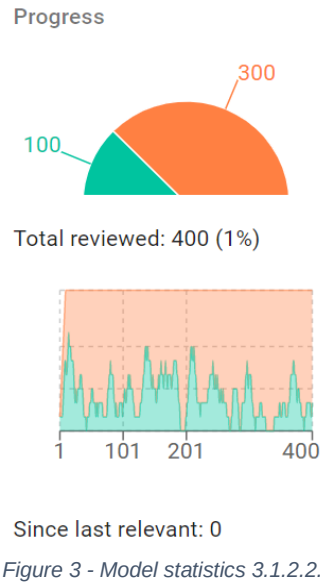


Figure 3 - Model statistics 3.1.2.2.

3.1.3. Narrowest research question (3), 95% max + 5% uncertain

What is known in the scientific literature about the effectiveness of financial resources to reduce greenhouse gas and nitrogen emissions to the air in the agricultural sector?

After consulting with PBL on the above findings, it was decided to screen one last time with an even more narrowly defined research question, focusing only on greenhouse gas emissions in the agricultural sector. Due to the short time available to investigate this question, it was decided not to compile another dataset. Therefore, we were limited to the existing dataset, also in terms of selecting the priors. To find these, we searched both via Google Scholar and in the dataset itself (using the terms 'agriculture' 'sustainable' 'ammonia' 'emission' and later greenhouse gas).

Finding suitable priors proved not to be so easy: because the dataset was not specifically compiled for this research question and there were also no articles on greenhouse gas emissions in the articles supplied by experts, there was relatively little literature on this subject in the dataset. It is therefore highly advisable to work with the same research question from start to finish: this way, one can both start with a smaller dataset and search more specifically for suitable priors to use as a starting point, which benefits the quality of the screening in ASReview. As priors for this research question, the research team chose the following articles from the existing dataset:

- Aneja *et al.* (2009). Ammonia assessment from agriculture: U.S. status and needs.
- Biffi *et al.* (2021). Aligning agri-environmental subsidies and environmental needs: A comparative analysis between the US and EU.
- Dace *et al.* (2015). Searching for solutions to mitigate greenhouse gas emissions by agricultural policy decisions - Application of system dynamics modeling for the case of Latvia.
- Jongebreur *et al.* (2001). Prevention and control of losses of gaseous nitrogen compounds in livestock operations: a review.
- Wei *et al.* (2018). Greenhouse gas and ammonia emissions and mitigation options from livestock production in peri-urban agriculture: Beijing – A case study.

Screening went a lot smoother with this research question, because you can tell very quickly when an article does not meet the inclusion criteria. In less than two working days, 220 articles were screened, of which 35 were included (see Appendix G). Screening stopped after eleven consecutive irrelevant articles were offered by the program. Nevertheless, there are some issues with the set of 35 articles that were included. Although they were labeled as relevant during the selection, looking at the final set still raises doubts as to whether these articles adequately address the core of the research question. This may be because we had to rely on the existing dataset of 40,223 articles, which was not compiled based on this research question, but also because there is simply little scientific literature on this topic. The model statistics in Figure 8 show that after approximately 110 screened articles the number of included articles decreased considerably; an indication that there were few (potentially) relevant articles left 'on the pile' at the time the screening was stopped.

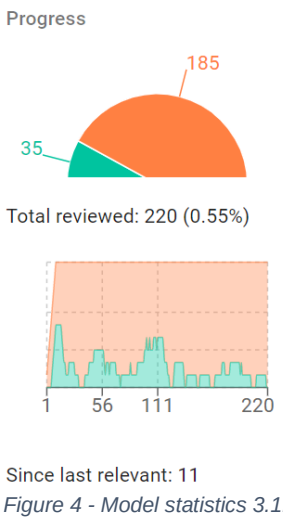


Figure 4 - Model statistics 3.1.3

3.2. Content Results

3.2.1 A list with relevance-ranked papers

Based on the screening performed in the previous section (3.1.2.2), a ranked, interactive dataset consisting of 100 articles was compiled. For this purpose, see Appendix D.

At the top of this list is a 2006 paper by Kleijn et al: 'Mixed biodiversity benefits of agri-environment schemes in five European countries'. A paper in the journal Ecology Letters (impact factor 9.5, with almost 600 citations. In this paper the biodiversity effects based on species density of Agri-Environmental Schemes in five different European countries are analyzed. The paper compares agricultural plots under nature management and conventionally managed plots. The paper finds some positive effect for biodiversity, but only for fairly common species. Much less effect is found for rare or threatened species. The abstract says the following:

> *Agri-environment schemes are an increasingly important tool for the maintenance and restoration of farmland biodiversity in Europe but their ecological effects are poorly known. Scheme design is partly based on non-ecological considerations and poses important restrictions on evaluation studies. We describe a robust approach to evaluate agri-environment schemes and use it to evaluate the biodiversity effects of agri-environment schemes in five European countries. We compared species density of vascular plants, birds, bees, grasshoppers and crickets, and spiders on 202 paired fields, one with an agri-environment scheme, the other conventionally managed. In all countries, agri-environment schemes had marginal to moderately positive effects on biodiversity. However, uncommon species benefited in only two of five countries and species listed in Red Data Books rarely benefited from agri-environment schemes. Scheme objectives may need to differentiate between biodiversity of common species that can be enhanced with relatively simple modifications in farming practices and diversity or abundance of endangered species which require more elaborate conservation measures*.

The second article in the list is from 2013, made by Scheper et al. "Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss - a meta-analysis," with, incidentally, David Kleijn as co-author. The article shows that the success of policies aimed at pollination strongly depends on the landscape context in which the farmer finds himself. The abstract reads as follows:

> *In Europe, agri-environmental schemes (AES) have been introduced in response to concerns about farmland biodiversity declines. Yet, as AES have delivered variable results, a better understanding of what determines their success or failure is urgently needed. Focusing on pollinating insects, we quantitatively reviewed how environmental factors affect the effectiveness of AES. Our results suggest that the ecological contrast in floral resources created by schemes drives the response of pollinators to AES but that this response is moderated by landscape context and farmland type, with more positive responses in croplands (vs. grasslands) located in simple (vs. cleared or complex) landscapes. These findings inform us how to promote pollinators and associated pollination services in species-poor landscapes. They do not, however, present viable strategies to mitigate loss of threatened or endangered*

species. This indicates that the objectives and design of AES should distinguish more clearly between biodiversity conservation and delivery of ecosystem services.

The third paper is a paper from the journal Nature and it is again a paper by Kleijn and co-authors but this time focusing exclusively on Dutch agriculture. The paper is already from 2001. The title of the paper is resolutely negative about the contribution of policy to biodiversity: 'Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes'. As in the later 2006 paper, the approach is based on a pairwise comparison of plots under nature-oriented agriculture with simply managed plots. The paper finds confounding effects (little effect or more effect on regularly managed plots and calls (in 2001) for much more serious evaluation of this policy. The abstract is:

Roughly 20% of the European Union's farmland is under some form of agri-environment scheme to counteract the negative impacts of modern agriculture on the environment. The associated costs represent about 4% (1.7 billion euros) of the European Union's total expenditure on the Common Agricultural Policy and are expected to rise to 10% in the near future. Although agri-environment schemes have been implemented in various countries for well over a decade, to date no reliable, sufficiently replicated studies have been performed to test whether such measures have the presumed positive effects on biodiversity. Here we present the results of a study evaluating the contribution of agri-environment schemes to the protection of biodiversity in intensively used Dutch agricultural landscapes. We surveyed plants, birds, hover flies and bees on 78 paired fields that either had agri-environment schemes in the form of management agreements or were managed conventionally. Management agreements were not effective in protecting the species richness of the investigated species groups: no positive effects on plant and bird species diversity were found. The four most common wader species were observed even less frequently on fields with management agreements. By contrast, hover flies and bees showed modest increases in species richness on fields with management agreements. Our results indicate that there is a pressing need for a scientifically sound evaluation of agri-environment schemes.

3.2.2 The expert result: A selection by experts of the 12 most important papers

The experts were asked to make a selection of three papers from the ranked list (Top 100) that are most important to policy makers concerned with the question "What is known in the scientific literature about the effectiveness of subsidies that are or can be used in Western countries to make agriculture and agricultural land use more sustainable? Obviously, not every expert selected the same three papers. The experts also used different criteria and methods for themselves.

Below is an impression of the criteria used (by various experts):

- Recently (last 5 years) published;
- To what extent they take into account the existing literature;

- The strength of the evidence regarding effectiveness and the rigor of the study design;
- Geographic scope;
- Multiple disciplines represented and in particular encompassing both ecology and social and economic aspects;
- Open access;
- Immediately usable (and readable) in Dutch agricultural policy.

A total of 12 papers were selected (see Table 2 on page 23). This is, according to our scholars, the selection of most important papers for policy makers to read. Below, we first show that table, ranking it in the Top 100 presented earlier. The main purpose of our quick and systematic review is the transfer of knowledge from science. Therefore, in addition to the table, we present the 12 abstracts of the selected papers, so that the contents can be immediately learned (see Appendix H). The references preceding the abstract also include a so-called DOI, a digital object identifier, which links to the real article (access depends on the institute from which one accesses the website).

One can ask many questions about this selection. What if we had asked the experts for 5 instead of three? What if there had been other experts on the sounding board? Would that have produced different papers? Possibly. At the same time, it is good to emphasize how many scientific voices shine through in this selection and to compare that to the "the standard quick search for scientific support" of a PBL employee we mentioned earlier.[5]

In this selection of 12 papers, a great many systematically collected "voices of science" resonate together. At the base are the 40,000 papers and their tens of thousands of authors from Scopus. With the intelligence of the algorithm and the knowledge of the priors and inclusion and exclusion criteria, these have been objectively tested for relevance as far as possible to produce a manageable set of 100 papers that reflect the state of scientific knowledge with respect to the question posed. From these, the combined but also diverse expertise and wisdom of six different scholars in the Sustainable Landscapes group selected these 'twelve voices' as the most important for policy to read. May the wisdom and insight of these twelve papers inspire policy!

[5] The manual search via Google Scholar or consulting an expert individually.

Papers selected by experts from the ranked list (Top 100)			
#	Year	# Top 100	Reference
1	2006	1	Kleijn, D., Baquero, R. A., Clough, Y., Díaz, M., De Esteban, J., Fernández, F. ... & Yela, J. L. (2006). Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology letters*, *9*(3), 243-254.
2	2013	2	Scheper, J., Holzschuh, A., Kuussaari, M., Potts, S. G., Rundlöf, M., Smith, H. G. & Kleijn, D. (2013). Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss–a meta-analysis. *Ecology letters*, *16*(7), 912-920.
3	2020	4	Piñeiro, V., Arias, J., Dürr, J., Elverdin, P., Ibáñez, A. M., Kinengyere, A. ... & Torero, M. (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes. *Nature Sustainability*, *3*(10), 809-820.
4	2021	5	Biffi, S., Traldi, R., Crezee, B., Beckmann, M., Egli, L., Schmidt, D. E. ... & Ziv, G. (2021). Aligning agri-environmental subsidies and environmental needs: a comparative analysis between the US and EU. *Environmental Research Letters*, *16*(5), 054067.
5	2015	7	Batáry, P., Dicks, L. V., Kleijn, D. & Sutherland, W. J. (2015). The role of agri-environment schemes in conservation and environmental management. *Conservation Biology*, *29*(4), 1006-1016.
6	2021	23	Brown, C., Kovács, E., Herzon, I., Villamayor-Tomas, S., Albizua, A., Galanaki, A. ... & Zinngrebe, Y. (2021). Simplistic understandings of farmer motivations could undermine the environmental potential of the Common Agricultural Policy. *Land Use Policy*, *101*, 105136.
7	2004	24	Kleijn, D., Berendse, F., Smit, R., Gilissen, N., Smit, J., Brak, B. & Groeneveld, R. (2004). Ecological effectiveness of agri-environment schemes in different agricultural landscapes in the Netherlands. *Conservation biology*, *18*(3), 775-786.
8	2015	25	Van Dijk, W. F., Lokhorst, A. M., Berendse, F. & de Snoo, G. R. (2015). Collective agri-environment schemes: How can regional environmental cooperatives enhance farmers' intentions for agri-environment schemes? *Land Use Policy*, *42*, 759-766.
9	2015	31	Merckx, T. & Pereira, H. M. (2015). Reshaping agri-environmental subsidies: From marginal farming to large-scale rewilding. *Basic and Applied Ecology*, *16*(2), 95-103.
10	2019	32	Zingg, S., Ritschard, E., Arlettaz, R. & Humbert, J. Y. (2019). Increasing the proportion and quality of land under agri-environment schemes promotes birds and butterflies at the landscape scale. *Biological conservation*, *231*, 39-48.
11	2014	62	Mouysset, L. (2014). Agricultural public policy: Green or sustainable? *Ecological Economics*, *102*, 15-23.
12	2013	100	Sanders, M. E., Nieuwenhuizen, W., Dirkx, G. H. P., Schrijver, R. A. M. & Smidt, R. A. (2013). Bedrijfsvoering zit in de weg: Landbouw slechts beperkt inzetbaar voor natuur-en landschapsbehoud. *Landschap: tijdschrift voor landschapsecologie en milieukunde*, *30*(2), 57-66.

Table 2 – The 12 most important papers for policymakers to read (selected by scientific experts)

4. Recommendations for and comments on ASReview

4.1 General

1. ASReview can certainly contribute to doing (systematic) literature reviews efficiently and transparently. Especially if the aim is to conduct a quick and systematic research, ASReview works better (i.e. more efficiently and reliably) with a narrow research question. Despite the fact that during the screening process care was taken not to get caught in a "trap", we mainly ended up with articles on the effectiveness of subsidies on biodiversity, and we found few articles on emissions of greenhouse gases such as CO2 and nitrogen. It should be noted, however, that this is related to the limited time available: if there had been more time to screen, we might have found more articles on other subjects.
2. A research question on this topic that allows ASReview to be used most efficiently delineates both the tool and the outcome. This has to do with the learning process of the algorithm. The clearer and more delineated the research question, the better the algorithm can be trained and the more efficiently ASReview can be used. In the case of this project, the initial research question was actually very broad, and many different tools as well as outcomes are relevant. This makes it more difficult for the algorithm to predict the probability of relevance of an article. Not only does this reduce the amount of time saved compared to a manual SR, as more articles need to be screened, but it also introduces another problem for doing systematic research, namely, trap formation.
3. Despite the fact that ASReview contributes to objectivity and efficiency of SRs, because the researcher is not influenced by information about authors or journal, and because not all articles in the stack need to be screened, it is still possible that some relevant articles will not be found. These may be, for example, "atypical" articles, whose title, abstract, and keywords contain terms other than those the algorithm knows are relevant, thus preventing the researcher from seeing them. Thus, there is always a trade-off between efficiency and completeness: because not all articles are screened, as in a manual SR, it cannot be guaranteed that 100% of the relevant articles will be found. Moreover, it is questionable whether this can be guaranteed in manual SRs.

4.2 Search strategy

4. The trap formation, which can be an obstacle especially for rapid screenings, can be avoided to some extent by the chosen query strategy. During the research process, a new query strategy within ASReview was also discussed: Clustering. This strategy could help with a broad research question and prevent trap formation. Because of the short time frame and from the knowledge that the researcher would first have to invest time in mastering this way, it was decided not to apply this strategy. In order to still cushion the trap somewhat, we used the "95% max + 5% uncertainty" setting. However, in ASReview you have no way of knowing which articles you will see that fall into this category; although sometimes you can reasonably sense it, for example when you are shown an article that is about an area in Africa.

5. It is not possible to switch search strategies halfway through the screening process. You can only choose the starting point and the settings (max, 95 max + 5 uncertainty, 95 max + 5 random, clustering and random) beforehand, and you are stuck with this throughout the entire screening process. I would have liked to be able to switch after screening x-number of articles, and thus be able to screen in a different way based on the same algorithm that is getting smarter all the time.

4.3 Abstract-screening with the help of AI

6. The choices you make at the beginning are very important in how the algorithm learns. For example, if (one of) your starting point articles is about a specific part of sustainable agriculture, such as bird populations or soil management, this will steer the algorithm in a certain direction.
7. ASReview does not (yet) have a "pause button. For each article that the researcher sees, a choice must be made between relevant or irrelevant. I found this difficult in some cases, especially with the broader research questions.
8. In contrast to more demarcated research fields such as medical science, in which terminology, (patient) population, intervention, comparison, outcome and type of research are clearly demarcated, this is not the case for this subject. This makes it more difficult, first, to assemble the dataset, and second, to screen efficiently with ASReview, and because there are more synonyms that ASReview does not immediately know to mean the same thing.
9. During the screening process, consideration should be given to trap formation. This means that once you are on the track of e.g. a specific outcome of policy, biodiversity on agricultural land, you can hardly deviate from it. As a result, articles on other outcomes, such as emission reduction, end up at the bottom of the pile, even though they are relevant to the study.
10. Algorithms can learn a lot, but they are not perfect. They can't "read comprehension" like humans; they can only convert words to numerical values, and based on the researcher's choices, determine the extent to which a value contributes to whether or not the article is relevant. But an algorithm does not know, for example, that **good** and **not good** are each other's opposite.

4.4 Objectivity and transparency

11. To ensure the objectivity of the screening, you judge the relevance of an article in ASReview by title and abstract only. However, due to the way of exporting in Scopus, some articles still have information at the end of the abstract about the year and in some cases also about the journal and/or author(s). This information can influence the researcher's choice and thus reduce the reliability of the study. It is recommended that this 'bug' be taken out in future versions of the software.
12. The researcher only gets to see (depending on your settings at least in 95% of the cases) the article that is currently on top of the pile. ASReview also shows the article that has the highest probability of relevance at that moment, but does not show exactly how big that probability is and on what basis it was calculated. In other words, the formula used to calculate the relevance probability is not transparent during the screening. Nor can the researcher see which article is in second, third or fourth place in the pile. This information is only available after the dataset is exported.

5. Conclusions

Procedural Conclusions

The goal of this project was to explore the usefulness of the open-source AI software ASReview for doing faster and reliable SRs aimed at policy advice. We did this in the context

of insights into the effectiveness of policy instruments to promote sustainable agriculture and sustainable agricultural land use in Western countries. We worked with three research questions of different breadth, which were screened a total of five times in ASReview. In this way, we got a good idea of what ASReview can, but more importantly what it cannot contribute to rapid policy-oriented systematic literature review.

ASReview can certainly contribute to the efficiency and objectivity of a rapid (systematic) literature review. However, the program does have some obstacles that need to be taken into account. For example, its usefulness and efficiency depends very much on the breadth of the research question. Also, in our experience, a lot of time goes into the pre-screening phase of titles and abstracts with ASReview, i.e. finding key articles and keywords, preparing the search strategy and dataset, although this is also the case with manual SRs. So while AI-assisted literature searches can take a lot of work out of the hands of the researcher, the time needed to properly prepare the abstract screening should certainly be taken into account.

Therefore, with regard to using ASReview to provide short-term, evidence-based advice for policy, our recommendation is to determine at the outset what the core of the problem is, in order to delineate the research question as much as possible. The narrower the research question, the more the use of ASReview can speed up the screening process, and the less the pitfalls of ASReview, such as trap formation, pose a problem for the reliability of the research.

In a process that aims to combine speed and depth, we believe it is necessary to use a broad set of scientific experts to secure the results. Unlike consulting experts solely on the basis of their own reputation and expertise, here the experts are asked to feed the broadly searching process of the systematic review with key papers (and key concepts) at the beginning, and (optionally) to make a concise selection at the end.

Content-wise conclusions

In a rapid and in-depth systematic review, one can present the final result in two policy-oriented ways. The first is to present a ranked list based on objective indicators of scientific recognition. This list makes the abstract of the papers easily accessible via a click-through Excel file. In this study, in terms of ranking the papers, we chose the number of citations by others (per year) and the scientific standing of the journal in which the papers were published. This makes the knowledge easily manageable for policy makers: the large set of 100 articles can be accessed in an orderly fashion according to scientific recognition. Because of the arrangement, every ten minutes is well spent because one can be assured of the greatest possible scientific relevance of what one is reading at any given moment.
For the second and most compact way of presenting the scientific knowledge, the experts are also asked to prioritize the raw list of results at the end. In this case, that resulted in a list of 12 articles.

Several "voices of science" come together in this most concise selection. At the base are the 40,000 papers and their tens of thousands of authors from Scopus. With the intelligence of the ASReview algorithm and the knowledge of the priors and inclusion and exclusion criteria, these have been tested for relevance as objectively as possible to produce a manageable set of 100 papers that reflect the state of scientific knowledge with respect to the question

posed. The combined expertise and wisdom of six different scientists from the Sustainable Landscapes group selected 'twelve voices' from these. May the wisdom and insight of these twelve papers inspire policy!

Compared to 'the standard quick search for scientific evidence' by a PBL staff member, screening literature with ASReview has a number of advantages. The standard search often consists of manually consulting Google Scholar and/or contacting a scientific expert individually. Compared to manual searches via Scholar, screening literature with ASReview provides the opportunity to systematically review a vast amount of studies for relevance while simultaneously reducing the time spent screening non-relevant literature. After each selection by the researcher, the entire database is reordered. Ultimately, this has been shown to result in a greater diversity of studies than a manual search process.

6. References

Abrami, P.C., Borokhovski, E., Bernard, R.M. … Surkes, M.A. (2010). Issues in conducting and disseminating brief reviews of evidence. *Evidence & Policy: A Journal of Research, Debate and Practice,* 6(3), 371-389. Doi: 10.1332/174426410X524866

Aneja, V.P., Blunden, J., James, K., … Cole, S. (2008). Ammonia assessment from agriculture: U.S. status and needs. *Journal of Environmental Quality,* 37(2), 515-520.

Batáry, P., Dicks, L.V., Kleijn, D., & Sutherland, W.J. (2015). The role of agri-environment schemes in conservation and environmental management. *Conservation Biology, 29*(4), 1006-1016. Doi:10.1111/cobi.12536

Biffi, S., Traldi, R., Crezee, B., Beckmann, M., Egli, L., Epp Schmidt, D., ... Ziv, G. (2021). Aligning agri-environmental subsidies and environmental needs: A comparative analysis between the US and EU. *Environmental Research Letters, 16*(5). doi:10.1088/1748-9326/abfa4e

Catarino, R., Gaba, S., & Bretagnolle, V. (2019). Experimental and empirical evidence shows that reducing weed control in winter cereal fields is a viable strategy for farmers. *Scientific Reports, 9*(1). doi:10.1038/s41598-019-45315-8

Dace, E., Muizniece, I., Blumberga, A., & Kaczala, F. (2015). Searching for solutions to mitigate greenhouse gas emissions by agricultural policy decisions - Application of system dynamics modeling for the case of Latvia. *Science of the Total Environment, 527-528*, 80-90. doi:10.1016/j.scitotenv.2015.04.088

Daniel, F. J., & Perraud, D. (2009). The multifunctionality of agriculture and contractual policies. A comparative analysis of France and the Netherlands. *Journal of Environmental Management, 90*(SUPPL. 2), S132-S138. doi:10.1016/j.jenvman.2008.11.015

Howden, S.M., Soussana, J., Tubiello, F.N., Chhetri, N., Dunlop, M. & Meinke, H. (2007). Adapting agriculture to climate change. *PNAS, 104*(40), 19691-19696. Doi: 10.1073/pnas.0701890104

Jongebreur, A.A., & Monteny, G.J. (2001). Prevention and control of losses of gaseous nitrogen compounds in livestock operations: a review. *The Scientific World Journal, 1 Suppl 2*, 844-851. doi: 10.1100/tsw.2001.339

Kleijn, D., Baquero, R. A., Clough, Y., Díaz, M., De Esteban, J., Fernández, F., ... Yela, J. L. (2006). Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology letters*, *9*(3), 243-254.

Kleijn, D., Berendse, F., Smit, R., & Gilissen, N. (2001). Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes. *Nature*, *413*(6857), 723-725.

Lopez-Avila, D., Husain, S., Bhatia, R., Nath, M. & Vinaygyam, R. (2017). *Agricultural innovation – An evidence gap map.* Report 12. New Delhi: International Initiative for Impact Evaluation (3ie).

McElwee, P., Calvin, K., Campbell, D., Cherubini, F., Grassi, G., Korotkov, V., ... Smith, P. (2020). The impact of interventions in the global land and agri-food sectors on Nature's Contributions to People and the UN Sustainable Development Goals. *Global Change Biology, 26*(9), 4691-4721. doi:10.1111/gcb.15219

Migliorelli, M., & Dessertine, P. (2018). Time for new financing instruments? A market-oriented framework to finance environmentally friendly practices in EU agriculture. *Journal of Sustainable Finance and Investment, 8*(1), 1-25. doi:10.1080/20430795.2017.1376270

Parra-López, C., Groot, J.C.J., Carmona-Torres, C., & Rossing, W.A.H. (2009). An integrated approach for ex-ante evaluation of public policies for sustainable agriculture at landscape level. *Land Use Policy, 26*(4), 1020-1030. doi:10.1016/j.landusepol.2008.12.006

Piñeiro, V., Arias, J., Dürr, J., Elverdin, P., Ibáñez, A. M., Kinengyere, A., ... Torero, M. (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes. *Nature Sustainability, 3*(10), 809-820. doi:10.1038/s41893-020-00617-y

Plieninger, T., Schleyer, C., Schaich, H., Ohnesorge, B., Gerdes, H., Hernández-Morcillo, M., & Bieling, C. (2012). Mainstreaming ecosystem services through reformed European agricultural policies. *Conservation Letters, 5*(4), 281-288. doi:10.1111/j.1755-263X.2012.00240.x

Pretty, J. (2008). Agricultural sustainability: Concepts, principles and evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 363(1491), 447-465.

Scheper, J., Holzschuh, A., Kuussaari, M., Potts, S. G., Rundlöf, M., Smith, H. G., & Kleijn, D. (2013). Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss–a meta-analysis. *Ecology letters*, *16*(7), 912-920.

Sneller, J. & Snels, B. (2019). *Onderbouwing van beleid – Het belang van artikel 3.1 van de Comptabiliteitswet 2016 om de regering te controleren.* Den Haag: Tweede Kamer der Staten-Generaal.

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B.L. . . . Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature, 562,* 519-525. Doi: 10.1038/s41586-018-0594-0

Torralba, M., Fagerholm, N., Burgess, P.J., Moreno, G. & Plieninger, T. (2016). Do European agroforestry systems enhance biodiversity and ecosystem services? A meta-analysis. *Agriculture, Ecosystems & Environment, 230,* 150-161. Doi: 10.1016/j.agee.2016.06.002

Van de Schoot, R., Bruin, J. de, Schram, R., Zahedi, P., Boer, J. de, . . . Oberski, D.L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, 3, 125-133. Doi: 10.1038/s42256-020-00287-7

Van Passel, S., Nevens, F., Mathijs, E. & Huylenbroek, G. van (2007). Measuring farm sustainability and explaining differences in sustainable efficiency. *Ecological Economics, 62*(1), 149-161. Doi: 10.1016/j.ecolecon.2006.06.008

Wei, S., Bai, Z. H., Chadwick, D., Hou, Y., Qin, W., Zhao, Z. Q., . . . Ma, L. (2018). Greenhouse gas and ammonia emissions and mitigation options from livestock production in peri-urban agriculture: Beijing – A case study. *Journal of Cleaner Production, 178*, 515-525. Doi: 10.1016/j.jclepro.2017.12.257

7. Appendices

Appendix A – Search query

TITLE-ABS-KEY (policy OR govern* OR institut* OR cap OR "Common Agricultural Policy" OR regulation OR regulatory OR law OR legislation OR instrument OR measure OR intervention OR arrangement) AND TITLE-ABS-KEY (effect* OR assess* OR evaluat* OR eviden* OR perform* OR implement*) AND TITLE-ABS-KEY (sustain* OR "environment* friendly" OR climate OR biodiversity OR conservation) AND TITLE-ABS-KEY (agricultur* OR agri OR farm* OR ecosystem OR "agro-ecosytem" OR agroecosystem OR "ecosystem service" OR cropland OR grassland OR livestock OR cattle OR cereal) AND TITLE-ABS-KEY (practice OR "land use" OR manag*) AND (LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015) OR LIMIT-TO (PUBYEAR , 2014) OR LIMIT-TO (PUBYEAR , 2013) OR LIMIT-TO (PUBYEAR , 2012) OR LIMIT-TO (PUBYEAR , 2011) OR LIMIT-TO (PUBYEAR , 2010) OR LIMIT-TO (PUBYEAR , 2009) OR LIMIT-TO (PUBYEAR , 2008) OR LIMIT-TO (PUBYEAR , 2007) OR LIMIT-TO (PUBYEAR , 2006) OR LIMIT-TO (PUBYEAR , 2005) OR LIMIT-TO (PUBYEAR , 2004) OR LIMIT-TO (PUBYEAR , 2003) OR LIMIT-TO (PUBYEAR , 2002) OR LIMIT-TO (PUBYEAR , 2001) OR LIMIT-TO (PUBYEAR , 2000)) AND (LIMIT-TO (AFFILCOUNTRY , "United States") OR LIMIT-TO (AFFILCOUNTRY , "United Kingdom") OR LIMIT-TO (AFFILCOUNTRY , "Australia") OR LIMIT-TO (AFFILCOUNTRY , "Germany") OR LIMIT-TO (AFFILCOUNTRY , "Italy") OR LIMIT-TO (AFFILCOUNTRY , "Canada") OR LIMIT-TO (AFFILCOUNTRY , "Spain") OR LIMIT-TO (AFFILCOUNTRY , "France") OR LIMIT-TO (AFFILCOUNTRY , "Netherlands") OR LIMIT-TO (AFFILCOUNTRY , "Sweden") OR LIMIT-TO (AFFILCOUNTRY , "Switzerland") OR LIMIT-TO (AFFILCOUNTRY , "Japan") OR LIMIT-TO (AFFILCOUNTRY , "Belgium") OR LIMIT-TO (AFFILCOUNTRY , "Portugal") OR LIMIT-TO (AFFILCOUNTRY , "Austria") OR LIMIT-TO (AFFILCOUNTRY , "Denmark") OR LIMIT-TO (AFFILCOUNTRY , "Norway") OR LIMIT-TO (AFFILCOUNTRY , "New Zealand") OR LIMIT-TO (AFFILCOUNTRY , "Finland") OR LIMIT-TO (AFFILCOUNTRY , "Greece") OR LIMIT-TO (AFFILCOUNTRY , "Poland") OR LIMIT-TO (AFFILCOUNTRY , "Russian Federation") OR LIMIT-TO (AFFILCOUNTRY , "Czech Republic") OR LIMIT-TO (AFFILCOUNTRY , "Ireland") OR LIMIT-TO (AFFILCOUNTRY , "Turkey") OR LIMIT-TO (AFFILCOUNTRY , "Romania") OR LIMIT-TO (AFFILCOUNTRY , "Hungary") OR LIMIT-TO (AFFILCOUNTRY , "Slovakia") OR LIMIT-TO (AFFILCOUNTRY , "Estonia") OR LIMIT-TO (AFFILCOUNTRY , "Slovenia") OR LIMIT-TO (AFFILCOUNTRY , "Lithuania") OR LIMIT-TO (AFFILCOUNTRY , "Ukraine") OR LIMIT-TO (AFFILCOUNTRY , "Croatia") OR LIMIT-TO (AFFILCOUNTRY , "Bulgaria") OR LIMIT-TO (AFFILCOUNTRY , "Serbia") OR LIMIT-TO (AFFILCOUNTRY , "Latvia") OR LIMIT-TO (AFFILCOUNTRY , "Cyprus") OR LIMIT-TO (AFFILCOUNTRY , "Iceland") OR LIMIT-TO (AFFILCOUNTRY , "Luxembourg") OR LIMIT-TO (AFFILCOUNTRY , "Malta") OR LIMIT-TO (AFFILCOUNTRY , "Albania") OR LIMIT-TO (AFFILCOUNTRY , "Georgia") OR LIMIT-TO (AFFILCOUNTRY , "North Macedonia") OR LIMIT-TO (AFFILCOUNTRY , "Montenegro") OR LIMIT-TO (AFFILCOUNTRY , "Azerbaijan") OR LIMIT-TO (AFFILCOUNTRY , "Monaco") OR LIMIT-TO (AFFILCOUNTRY , "Belarus") OR LIMIT-TO (AFFILCOUNTRY , "Bosnia and Herzegovina") OR LIMIT-TO (AFFILCOUNTRY , "Armenia") OR LIMIT-TO (AFFILCOUNTRY , "Undefined"))

Appendix B – The 42 'key articles' supplied by the sounding boards

Delivered by PBL sounding board

Beers, C. van, Bergh, J.C.J.M. van den, Moor, A. de & Oosterhuis, F.H. (2003) Milieu-effecten van indirecte subsidies. *Econ Stat Berichten, 88,* 129-131.

Conway, A.G. (1991). Plenary paper 1: A role for economic instruments in reconciling agricultural and environmental policy in accordance with the Polluter Pays Principle. European Review of Agricultural Economics, 18 (3-4) 467–484. Doi: 10.1093/erae/18.3-4.467

Czyzewski, B. Matuszczak, A., Grzelak, A., Guth, M. & Majchrzak, A. (2021). Environmental sustainable value in agriculture revisited: How does Common Agricultural Policy contribute to eco-efficiency? *Sustainability Science*, *16,* 137-152.

Díaz, M. & Concepción, E.D. (2016). Enhancing the effectiveness of CAP Greening as a Conservation Tool: a Plea for Regional Targeting Considering Landscape Constraints. *Current Landcape Ecology Reports, 1,* 168-177.

Engel, S. (2016), The devil in the detail: a practical guide on designing payments for environmental services. *International Review of Environmental and Resource Economics,* 9 (1-2), 131- 177

Ferraro, P.J. (2011). The future of payments for environmental services. *Conservation Biology*, 25(6), 1134-1138.

Gottschalk, T. et al. (2007). Impact of agricultural subsidies on biodiversity at the landscape level. *Landscape Ecology*, 22, 643-656.

Hermann, D., Sauthoff, S. & Musshoff, O. (2017). Ex-ante evaluation of policy measures to enhance carbon sequestration in agricultural soils. *Ecological Economics, 140*, 241-250. Doi: 10.1016/j.ecolecon.2017.05.018

Khanna, M., Isik, M. & Zilberman, D. (2005). Cost-effectiveness of alternative green payment policies for conservation technology adoption with heterogeneous land quality. *Agricultural Economics*, 27(2), 157-174

Miteva, D. A., Pattanayak, S. K., & Ferraro, P. J. (2012). Evaluation of biodiversity policy instruments: what works and what doesn't? *Oxford Review of Economic Policy,* 28 (1), 69-92.

Parra-López, C., Groot, J.C.J., Carmona-Torres, C., & Rossing, W.A.H. (2009). An integrated approach for ex-ante evaluation of public policies for sustainable agriculture at landscape level. *Land Use Policy, 26*(4), 1020-1030. doi: 10.1016/ j.landusepol.2008.12.006

Pretty, J., Brett, C., Gee, D., Hine, R., Mason, C., Morison, J., Rayment, M., Van Der Bijl, G., & Dobbs, T. (2001). Policy challenges and priorities for internalizing the externalities of modern agriculture. *Journal of Environmental Planning and Management,* 44(2), 263-283.

Slangen, L.H.G. (1992). Policies for nature and landscape conservation in Dutch agriculture: An evaluation of objectives, means, effects, and programme costs. *European Review of Agricultural Economics,* 19(3), 331-350.

Vásáry, M., Osztrogonácz, I, Dobó, E., Buzás, R. & Vásáry, V. (2007). Sustainable Agriculture: With or Without Subsidies? *Cereal Research Communications, 35*(2), 1285-1288.

Delivered by academic sounding board

Ansell, D., Freudenberger, D., Munro, N. & Gibbons, P. (2016). The cost-effectiveness of agri-environment schemes for biodiversity conservation: a quantitative review. *Agriculture, Ecosystems & Environment, 225,* 184-191.

Batáry, P., Dicks, L. V., Kleijn, D., & Sutherland, W. J. (2015). The role of agri-environment schemes in conservation and environmental management. *Conservation Biology, 29*(4), 1006-1016. doi:10.1111/cobi.12536

Boetzl, F.A., Krauss, J., Heinze, J., Hoffmann, H., Juffa, J. … Steffan-Dewenter, I. (2021). A multitaxa assessment of the effectiveness of agri-environmental schemes for biodiversity management. *PNAS, 118* (10).

Brzezina, N., Biely, K., Helfgott, A., Kopainsky, B., Vervoort, J., & Mathijs, E. (2017). Development of organic farming in europe at the crossroads: Looking for the way forward through system archetypes lenses. *Sustainability (Switzerland), 9*(5). doi:10.3390/su9050821

Czyzewski, B. & Smedzik-Ambrozy, K. (2017). The regional structure of the CAP subsidies and the factor productivity in agriculture in the EU 28. *Agricultural Economies Czech, 63,* 149-163.

Elts, J. & Lohmus, A. (2012). What do we lack in agri-environment schemes? The case of farmland birds in Estonia. *Agriculture, Ecosystems & Environment, 156* (1), 89-93. Doi: 10.1016/j.agee.2012.04.023

Hodgson, J.G., Montserrat-Martí, G., Tallowin, J., Thompson, K., Díaz, J. … Zak, M.R. (2005). How much will it cost to save grassland diversity? *Biological Conservation, 122* (2), 263-273.

Kleijn, D., Berendse, F., Smit, R., & Gilissen, N. (2001). Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes. *Nature, 413*(6857), 723-725. Doi: 10.1038/35099540

Laborde, D., Mamun, A., Martin, W., Piñeiro, V. & Vos, R. (2021). Agricultural subsidies and global greenhouse gas emissions*. Nature Communications* 12, 2601. Doi: 10.1038/s41467-021-22703-1

Lohr, L. & Salomonsson, L. (2000). Conversion subsidies for organic production: results from Sweden and lessons for the United States. *Agricultural Economics, 22* (2), 133-146. Doi: 10.1016/S0169-5150(99)00045-6

Mujtar, V. el, Muñoz, N. Prack McCormick, B., Pulleman, M. & Tittonell, P. (2019). Role and management of soil biodiversity for food security and nutrition; where do we stand? *Global Food Security, 20*, 132-144.

Noordwijk, M. van & Brussaard, L. (2014). Minimizing the ecological footprint of food: closing yield and efficiency gaps simultaneously? *Current Opinion in Environmental Sustainability, 8,* 62-70.

Pe'er, G., Bonn, A., Bruelheide, H., Dieker, P., Eisenhauer, N. … Lakner, S. (2020). Action needed for the EU Common Agricultural Policy to address sustainability challenges. *People and Nature, 2*(2), 305-316.

Pe'er, G., Dicks, L.V., Visconti, P., Arlettaz, R., Báldi, A. … Scott, A.V. (2014). EU Agricultural reform fails on biodiversity: Extra steps by member states are needed to protect farmed and grassland ecosystems. *Science, 344* (6188), 1090-1093.

Piñeiro, V., Arias, J., Dürr, J., Elverdin, P., Ibáñez, A. M., Kinengyere, A., ... Torero, M. (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes. *Nature Sustainability, 3*(10), 809-820. Doi: 10.1038/s41893-020-00617-y

Plieninger, T., Schleyer, C., Schaich, H., Ohnesorge, B., Gerdes, H., Hernández-Morcillo, M., & Bieling, C. (2012). Mainstreaming ecosystem services through reformed European agricultural policies. *Conservation Letters, 5*(4), 281-288. doi:10.1111/j.1755-263X.2012.00240.x

Runhaar, H.A.C. (2017). Governing the transformation towards 'nature-inclusive' agriculture: insights from the Netherlands. *International Journal of Agricultural Sustainability, 15*(4), 340-349. Doi: 10.1080/14735903.2017.1312096

Runhaar, H.A.C., Melman, T. C. P., Boonstra, F. G., Erisman, J. W., Horlings, L. G., de Snoo, G. R., ... Arts, B. J. M. (2017). Promoting nature conservation by Dutch farmers: a governance perspective†. *International Journal of Agricultural Sustainability, 15*(3), 264-281. doi: 10.1080/14735903.2016.1232015 Rose, D. C.,

Sutherland, W. J., Barnes, A. P., Borthwick, F., Ffoulkes, C., Hall, C., . . . Dicks, L. V. (2019). Integrated farm management for sustainable agriculture: Lessons for knowledge exchange and policy. *Land Use Policy, 81*, 834-842. doi:10.1016/j.landusepol.2018.11.001

Scown, M. W., Brady, M. V., & Nicholas, K. A. (2020). Billions in Misspent EU Agricultural Subsidies Could Support the Sustainable Development Goals. *One Earth, 3*(2), 237-250. doi:10.1016/j.oneear.2020.07.011

Strijker, D., Sijtsma, F. J., & Wiersma, D. (2000). Evaluation of nature conservation: An application to the Dutch ecological network. *Environmental and Resource Economics, 16*(4), 363-378. doi:10.1023/A:1008344604392

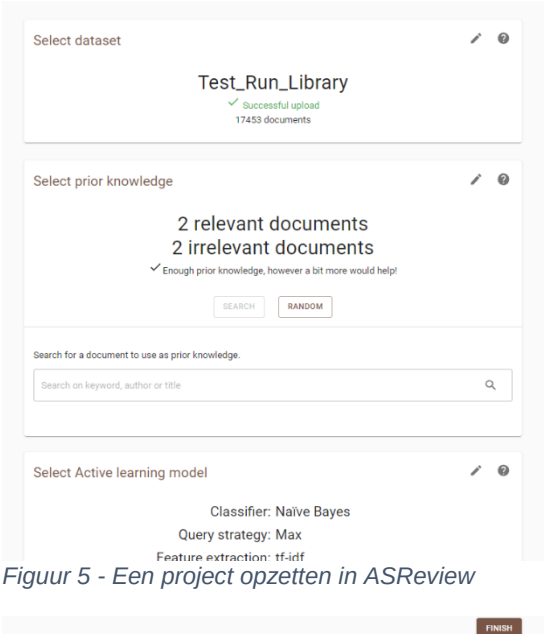
Tittonell, P. (2014). Ecological intensification of agriculture - sustainable by nature. *Current Opinion in Environmental Sustainability, 8,* 53-61.

Tittonell, P. (2020). Assessing resilience and adaptability in agroecological transitions. *Agricultural Systems, 184*. doi:10.1016/j.agsy.2020.102862

Appendix C – Quick Guide ASReview

1. Setting-up a project

The first step is to upload the prepared dataset in .csv or .ris format. Next, prior knowledge must be specified: 1-5 relevant and 1-5 irrelevant articles. The more prior knowledge specified, the faster the algorithm learns what is relevant to the researcher.



Figuur 5 - Een project opzetten in ASReview

Finally, the researcher must determine the settings of the model:

- A classifier; how the model is created. The default setting is Naïve Bayes. This is the default because research has shown that this setting works best for many different research questions (i.e., the model learns most quickly what the researcher is looking for) (ASReview Summer school, 2021).
- - The query strategy; which article the researcher will see. The default setting is 'max', which means that the researcher is always shown the article that has the highest probability of relevance at that moment. There are also other settings:
    - Random - like a manual SR;
    - Clustering - ASReview clusters the data based on the amount of words they have in common
    - Mixed (95% max, 5% random) - the researcher is shown the article with the highest relevance probability 95% of the time, and a random article 5% of the time;
    - Mixed (95% max, 5% uncertain) - the researcher is shown the article with the highest relevance probability 95% of the time, and 5% of the time the article whose relevance probability the algorithm has the most trouble determining (i.e. the article whose relevance probability is closest to 0.5).
- Feature extraction; how the model "translates" the text into terms it can understand and how the model determines the value of each of these terms. The default setting is tf-idf (Term Frequency - Inverse Document Frequency). This means that words that occur more frequently in the text are given a higher score, but words that occur frequently throughout the data set are given a lower score. This allows controlling for the fact that some words are more common in general (such as 'and' and 'the').

2. Screening titles and abstracts in ASReview

After the above steps are completed, the screening of the titles and abstracts can begin. Which article the researcher will see depends on the settings chosen in the above step. Assuming that the query strategy 'Max' has been chosen, the researcher will each time be shown the title and abstract of the article that has the relevance probability at that moment according to ASReview (see Figure 7). Because each choice affects the algorithm and thus the 'order of the stack', these must be carefully considered.

For each article, a choice must be made between relevant and irrelevant; in other words, a choice must be made. In some cases this can be difficult, especially with a broad research question like this. There is a possibility to undo choices made, but this causes the algorithm to be modified. This can interfere with the learning process of the algorithm.

Figure 7 – An example of an abstract to be screened in ASReview

During the screening, a statistics dashboard is displayed on the right-hand side of the screen (see Figure 8). In addition to the project and the size of the dataset, this also contains information about the progress of the screening and the model fit. The semicircle diagram shows how many of the screened articles were labelled as (ir)relevant and the graph underneath shows how well the model can estimate the relevance. The better the model is able to do this, the more articles in a row are labelled as relevant and the higher the green line. Finally, it shows how long ago an article was last marked as relevant. In this way, as a researcher, you have insight into your progress during screening, how many articles have already been included and how "well" the algorithm can assess relevance.

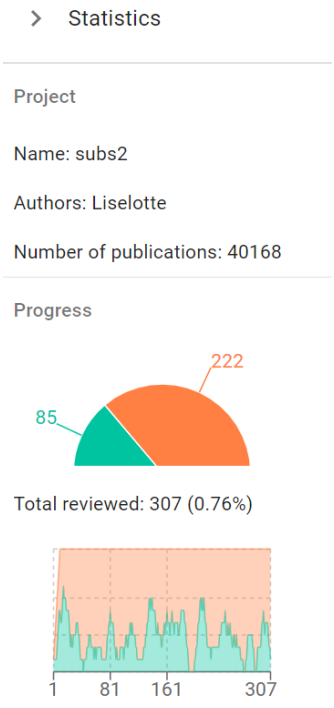


Figure 8 – Screening statistics in ASReview (example)

3. Determine the stop point

These statistics also help determine the stopping point. When fewer and fewer articles are included over time, or when it has

been a long time since an article has been labeled relevant. The stopping point can be determined based on:

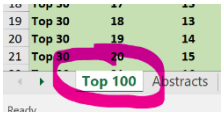
- Pragmatic; after a certain amount of time or a certain number/percentage of articles have been screened, the researcher decides that she has found enough literature. This does mean that the literature review is no longer fully 'systematic', as there may still be relevant articles 'in the pile';
- After x-number of consecutive irrelevant articles; when the researcher has consecutively labeled, for example, 50 articles as irrelevant, it can be said that (almost) all relevant articles have been found. After all, the article that according to the program has the highest chance of being relevant has already been found irrelevant by the researcher 50 times in a row. However, it should be noted that there is always a chance that relevant literature has been left 'on the pile'.

4. Closing a project in ASReview

After the screening process has stopped, the labeled dataset can be downloaded. After this, the included articles can be reviewed full-text for relevance, and then the synthesis of the SR can be prepared.

Appendix D – The top 100 papers

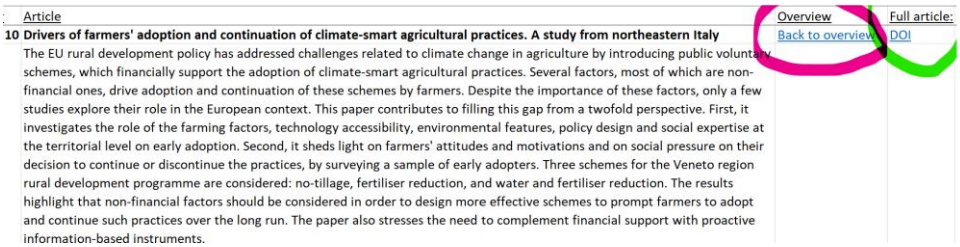
This appendix is primarily available as an Excel file. In the 'Top 100' worksheet of the Excel file, as below, one can see the titles of the 100 papers in summary form (including authors, year, and journal)



To the left of the titles (in the Abstracts column), one can also click directly through to the Abstracts on a separate worksheet.

Top	Nr	Final Rank	Abstract	Title
Top 3	1	1	1	Mixed biodiversity benefits of agri-environment schemes in five European countries
Top 3	2	2	2	Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss - a meta-analysis
Top 3	3	2	3	Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes
Top 10	4	3	4	A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes
Top 10	5	4	5	Aligning agri-environmental subsidies and environmental needs: A comparative analysis between the US and EU

In the Abstracts worksheets, one can read the entire abstract. To the right of the abstract is a link (see pink circle below) that takes the user directly back to where one left the Top 100 worksheet. Even more to the right of the Abstract is the so-called DOI (Digital Object Identifier) column that leads the user directly to the original article on the Internet. Access depends on the rights of the user or the institution from which the DOI is accessed.



Article	Overview	Full article:
10 **Drivers of farmers' adoption and continuation of climate-smart agricultural practices. A study from northeastern Italy**	Back to overview	DOI

The EU rural development policy has addressed challenges related to climate change in agriculture by introducing public voluntary schemes, which financially support the adoption of climate-smart agricultural practices. Several factors, most of which are non-financial ones, drive adoption and continuation of these schemes by farmers. Despite the importance of these factors, only a few studies explore their role in the European context. This paper contributes to filling this gap from a twofold perspective. First, it investigates the role of the farming factors, technology accessibility, environmental features, policy design and social expertise at the territorial level on early adoption. Second, it sheds light on farmers' attitudes and motivations and on social pressure on their decision to continue or discontinue the practices, by surveying a sample of early adopters. Three schemes for the Veneto region rural development programme are considered: no-tillage, fertiliser reduction, and water and fertiliser reduction. The results highlight that non-financial factors should be considered in order to design more effective schemes to prompt farmers to adopt and continue such practices over the long run. The paper also stresses the need to complement financial support with proactive information-based instruments.

The ranking in the Excel and therefore in the list below shows the 'scientific recognition' of an article. This is calculated based on the ranking of citation score per year (50%) and the ranking of the impact factor of the journal (50%).

Top 3	**1**	**1**	1	**Mixed biodiversity benefits of agri-environment schemes in five European countries**
Top 3	**2**	**2**	2	**Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss - a meta-analysis**
Top 3	**3**	**2**	3	**Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes**
Top 10	**4**	**3**	4	**A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes**
Top 10	**5**	**4**	5	**Aligning agri-environmental subsidies and environmental needs: A comparative analysis between the US and EU**
Top 10	**6**	**5**	6	**Agricultural sustainability: Concepts, principles and evidence**
Top 10	**7**	**6**	7	**The role of agri-environment schemes in conservation and environmental management**
Top 10	**8**	**6**	8	**Mixed effects of organic farming and landscape complexity on farmland biodiversity and biological control potential across Europe**
Top 10	**9**	**7**	9	**The future of agri-environment schemes: Biodiversity gains and ecosystem service delivery?**
Top 10	**10**	**7**	10	**Drivers of farmers' adoption and continuation of climate-smart agricultural practices. A study from northeastern Italy**

Top 30	**11**	**8**	11	**Will agri-environment schemes deliver substantial biodiversity gain, and if not why not?**
Top 30	**12**	**9**	12	**Assessing agricultural eco-efficiency in Italian Regions**
Top 30	**13**	**10**	13	**Exploring the knowledge landscape of ecosystem services assessments in Mediterranean agroecosystems: Insights for future research**
Top 30	**14**	**11**	14	**The role of agri-environment schemes and farm management practices in reversing the decline of farmland birds in England**
Top 30	**15**	**12**	15	**FORUM: Landscape-scale conservation: Collaborative agri-environment schemes could benefit both biodiversity and ecosystem services, but will farmers be willing to participate?**
Top 30	**16**	**13**	16	**Local and landscape effects of organic farming on butterfly species richness and abundance**
Top 30	**17**	**13**	17	**Mainstreaming ecosystem services through reformed European agricultural policies**
Top 30	**18**	**13**	18	**Farmers value on-farm ecosystem services as important, but what are the impediments to participation in PES schemes?**
Top 30	**19**	**14**	19	**Experimental evidence that the effectiveness of conservation biological control depends on landscape complexity**
Top 30	**20**	**15**	20	**Landscape-moderated biodiversity effects of ground herb cover in olive groves: Implications for regional biodiversity conservation**
Top 30	**21**	**16**	21	**Creating culturally sustainable agri-environmental schemes**
Top 30	**22**	**17**	22	**Social and ecological drivers of success in agri-environment schemes: The roles of farmers and environmental context**
Top 30	**23**	**18**	23	**Simplistic understandings of farmer motivations could undermine the environmental potential of the common agricultural policy**
Top 30	**24**	**19**	24	**Ecological effectiveness of agri-environment schemes in different agricultural landscapes in The Netherlands**
Top 30	**25**	**20**	25	**Collective agri-environment schemes: How can regional environmental cooperatives enhance farmers' intentions for agri-environment schemes?**
Top 30	**26**	**20**	26	**Current use of impact models for agri-environment schemes and potential for improvements of policy design and assessment**
Top 30	**27**	**21**	27	**Divergent farmer and scientist perceptions of agricultural biodiversity, ecosystem services and decision-making**
Top 30	**28**	**21**	28	**Impact of landscape improvement by agri-environment scheme options on densities of characteristic farmland bird species and brown hare (Lepus europaeus)**
Top 30	**29**	**21**	29	**Plant species richness decreased in semi-natural grasslands in the Biosphere Reserve Wienerwald, Austria, over the past two decades, despite agri-environmental measures**
Top 30	**30**	**22**	30	**What can management option uptake tell us about ecosystem services delivery through agri-environment schemes?**
Top 50	**31**	**23**	31	**Reshaping agri-environmental subsidies: From marginal farming to large-scale rewilding**
Top 50	**32**	**23**	32	**Increasing the proportion and quality of land under agri-environment schemes promotes birds and butterflies at the landscape scale**
Top 50	**33**	**23**	33	**Assessing the feasibility of carbon payments and Payments for Ecosystem Services to reduce livestock grazing pressure on saltmarshes**
Top 50	**34**	**24**	34	**Effects of a coordinated farmland bird conservation project on farmers' intentions to implement nature conservation practices – Evidence from the Swedish Volunteer & Farmer Alliance**
Top 50	**35**	**25**	35	**Investigating farmers' preferences for alternative PES schemes for carbon sequestration in UK agroecosystems**
Top 50	**36**	**26**	36	**Bird diversity relates to agri-environment schemes at local and landscape level in intensive farmland**
Top 50	**37**	**27**	37	**Financial imperative or conservation concern? EU farmers' motivations for participation in voluntary agri-environmental schemes**
Top 50	**38**	**27**	38	**Management of agricultural soils for greenhouse gas mitigation: Learning from a case study in NE Spain**
Top 50	**39**	**28**	39	**Farmers and nature conservation: What is known about attitudes, context factors and actions affecting conservation?**

Top 50	40	29	40	Strategies for the sustainable management of phosphorus
Top 50	41	30	41	Mixed effectiveness of French agri-environment schemes for nationwide farmland bird conservation
Top 50	42	31	42	An integrated approach for ex-ante evaluation of public policies for sustainable agriculture at landscape level
Top 50	43	31	43	A fuzzy cognitive mapping approach for the assessment of public-goods governance in agricultural landscapes
Top 50	44	32	44	Little and late: How reduced hedgerow cutting can benefit Lepidoptera
Top 50	45	33	45	A survey exploring private farm advisor perspectives of agri-environment schemes: The case of England's Environmental Stewardship programme
Top 50	46	34	46	Farmer participation in agri-environmental schemes: Towards conservation-oriented thinking?
Top 50	47	35	47	Biological conservation in dynamic agricultural landscapes: Effectiveness of public policies and trade-offs with agricultural production
Top 50	48	35	48	Implementation of landscape planning and nature conservation in the agricultural landscape - A case study from saxony
Top 50	49	35	49	Multi-scale effects of agri-environment schemes on carabid beetles in intensive farmland
Top 50	50	36	50	'Greening' green infrastructure. Good italian practices for enhancing green infrastructure through the common agricultural policy
Top 100	51	36	51	Mammals, agri-environment schemes and set-aside - What are the putative benefits?
Top 100	52	36	52	Assessing sustainability in agricultural landscapes: A review of approaches1,2
Top 100	53	36	53	Soil degradation, farming practices, institutions and policy responses: An analytical framework
Top 100	54	37	54	A landscape perspective on sustainability of agricultural systems
Top 100	55	37	55	Plant diversity partitioning in Mediterranean croplands: Effects of farming intensity, field edge, and landscape context
Top 100	56	38	56	Exploring cooperative place-based approaches to restorative agriculture
Top 100	57	38	57	The role of network bridging organisations in compensation payments for agri-environmental services under the EU Common Agricultural Policy
Top 100	58	39	58	Ecological effectiveness of French grassland agri-environment schemes for farmland bird communities
Top 100	59	40	59	Evaluating the effectiveness of conservation measures for European grassland-breeding waders
Top 100	60	41	60	Multiple benefits of carbon-friendly agricultural practices: Empirical assessment of conservation tillage
Top 100	61	42	61	Is it worth protecting groundwater from diffuse pollution with agri-environmental schemes? A hydro-economic modeling approach
Top 100	62	43	62	Agricultural public policy: Green or sustainable?
Top 100	63	44	63	Restoration of plant diversity on ditch banks: Seed and site limitation in response to agri-environment schemes
Top 100	64	45	64	Present agri-environment measures in Europe are not sufficient for the conservation of a highly sensitive bird species, the Corncrake Crex crex
Top 100	65	46	65	Influence of land sharing and land sparing strategies on patterns of vegetation and terrestrial vertebrate richness and occurrence in Australian endangered eucalypt woodlands
Top 100	66	47	66	Farmer awareness and implementation of sustainable agriculture practices in different types of farms in Poland
Top 100	67	47	67	Towards sustainable land use: Public demand for plant diversity in agricultural landscapes of central Germany
Top 100	68	48	68	Effect of agri-environment measure for the aquatic warbler on bird biodiversity in the extensively managed landscape of Biebrza Marshes (Poland)
Top 100	69	49	69	Historical development, state and perspectives of environmental management in Bulgarian agriculture
Top 100	70	50	70	Turning Farmers into Conservationists? Progress and Prospects
Top 100	71	51	71	Agri-environmental schemes and the European agricultural landscapes: The role of indicators as valuing tools for evaluation

Top 100	72	51	72	The effect of agri-environment schemes on amphibian diversity and abundance
Top 100	73	52	73	Management of high nature value farmland in the republic of ireland: 25 years evolving toward locally adapted results-orientated solutions and payments
Top 100	74	52	74	Stakeholder perceptions of the effectiveness and efficiency of agri-environment schemes in enhancing pollinators on farmland
Top 100	75	53	75	Assessing ecologically sustainable agricultural land-use in the Central Pyrénées at the field and landscape level
Top 100	76	54	76	Impacts of management at a local and landscape scale on pollinators in semi-natural grasslands
Top 100	77	54	77	Sustainable landscape, soil and crop management practices enhance biodiversity and yield in conventional cereal systems
Top 100	78	55	78	Effects of agri-environmental schemes on farmland birds: Do food availability measurements improve patterns obtained from simple habitat models?
Top 100	79	55	79	Agri-environmental schemes for biodiversity and environmental protection: How were are not yet “hitting the right keys”
Top 100	80	56	80	Towards Sustainable Agriculture? The EU framework and local adaptation in Sweden and Poland
Top 100	81	57	81	The importance of spatial and temporal scale for agri-environment scheme delivery
Top 100	82	58	82	Assessing pollinators' use of floral resource subsidies in agri-environment schemes: An illustration using Phacelia tanacetifolia and honeybees
Top 100	83	58	83	Governing agri-environmental schemes: Lessons to be learned from the new institutional-economics approach
Top 100	84	59	84	Effects of agri-environment management for cirl buntings on other biodiversity
Top 100	85	60	85	Mixing public and private agri-environment schemes: Effects on farmers participation in quebec, canada
Top 100	86	61	86	Evaluation of a dairy agri-environmental programme for restoring woody green infrastructure
Top 100	87	62	87	Alternative agricultural policy scenarios, sector modelling and indicators: A sustainability assessment
Top 100	88	62	88	Radar remote sensing as a novel tool to assess the performance of an agri-environment scheme in coastal grasslands
Top 100	89	63	89	Are the conservation requirements of pseudo-steppe birds adequately covered by spanish agri-environmental schemes? An ex-ante assessment
Top 100	90	64	90	System dynamics model to design effective policy strategies aiming at fostering the adoption of conservation agriculture practices in sicily
Top 100	91	65	91	What influences farmers 'acceptance of agrienvironment schemes? An ex-post application of the Theory of Planned Behaviour'
Top 100	92	65	92	Policy Integration for Sustainable Agricultural Landscapes: Taking Stock of UK Policy and Practice
Top 100	93	66	93	Cost-effective Biodiversity Conservation:Procurement Auctions and Payment-by-Results
Top 100	94	66	94	Evaluation of agricultural ecosystem services in fallowing land based on farmers' participation and model simulation
Top 100	95	66	95	Greenhouse gas mitigation strategies and opportunities for agriculture
Top 100	96	67	96	Valoración indirecta de un programa agroambiental enfocado a la conservación de la avifauna esteparia en el norte de España
Top 100	97	68	97	Balancing food production and biodiversity conservation in arable landscapes: Lessons from the Farm4Bio experiment
Top 100	98	69	98	Environmentalizmus a pol'nohospodárska prax
Top 100	99	69	99	Operation pollinator: Positive action for pollinators and improved biodiversity on farm
Top 100	100	70	100	Bedrijfsvoering zit in de weg: Landbouw slechts beperkt inzetbaar voor natuur- en landschapsbehoud

Appendix E – The instructions for the experts in the sounding board groups.

Highlighting important content from 100 top papers

In the attached Excel-file 'Database sustainable agriculture' you will find the Top 100 articles that we have compiled based on screening abstracts in ASReview. The purpose of our screening is to find a fast and systematic way of informing policy makers in sustainable agriculture about important insights in science.

Our research question which was at the basis of the selection of these 100 articles was:

What does the scientific literature say about the effectiveness of subsidies (subsidy/income support/compensation) that aim to stimulate sustainable agricultural practices and agricultural land use in Western countries?

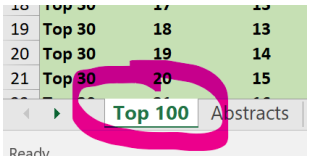
The 100 articles are ranked according to 'scientific recognition' for which we used both citations per year and the impact factors of the journals.

We would like now to dig a bit deeper into the content of what these papers say, and what you think is important for policy makers to know. Therefore we would like to know from you, based on your expertise:

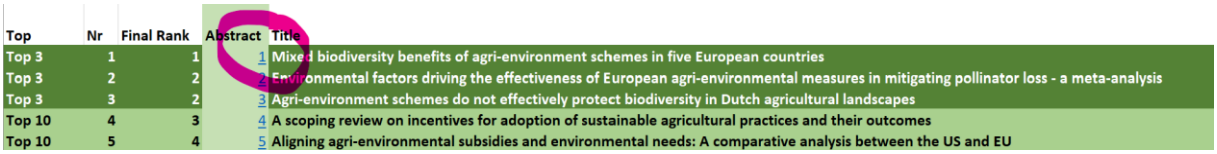
Which three of these hundred articles do you consider to be the most relevant or 'key' for policy makers, and why?

Please use the form on page 2 to fill in your answers

Tips: the way our database is structured and can be handled is as follows.
In the worksheet 'Top 100' you find the rank and titles of the articles (and to the right the authors, journal and category). This makes it easy to quickly browse through all hundred.

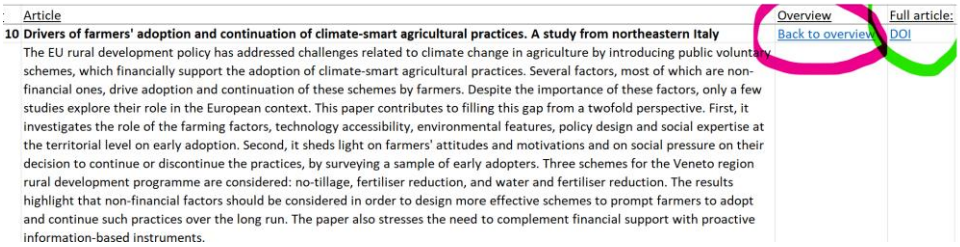


To the left of the title, in the Abstract column is a number that you can click on.



Clicking this number brings you to the Abstract-worksheet at the exact place of the abstract of paper number 1. Here you can read the abstract.

To the right of the abstract (see pink circle below) is a link that brings you back to where you left the Top 100 worksheet.
One cell more to the right (green circle below) you can find the DOI that leads you directly to the full text of the article. Access obviously depends on your institution.



Article	Overview	Full article:
10 **Drivers of farmers' adoption and continuation of climate-smart agricultural practices. A study from northeastern Italy**	Back to overview	DOI

The EU rural development policy has addressed challenges related to climate change in agriculture by introducing public voluntary schemes, which financially support the adoption of climate-smart agricultural practices. Several factors, most of which are non-financial ones, drive adoption and continuation of these schemes by farmers. Despite the importance of these factors, only a few studies explore their role in the European context. This paper contributes to filling this gap from a twofold perspective. First, it investigates the role of the farming factors, technology accessibility, environmental features, policy design and social expertise at the territorial level on early adoption. Second, it sheds light on farmers' attitudes and motivations and on social pressure on their decision to continue or discontinue the practices, by surveying a sample of early adopters. Three schemes for the Veneto region rural development programme are considered: no-tillage, fertiliser reduction, and water and fertiliser reduction. The results highlight that non-financial factors should be considered in order to design more effective schemes to prompt farmers to adopt and continue such practices over the long run. The paper also stresses the need to complement financial support with proactive information-based instruments.

Your name

Title article #
What are the most important insights provided by this article?

Why do you consider this article to be relevant or 'key' for policy makers?

Additional comments (optional)

Title article #2
What are the most important insights provided by this article?

Why do you consider this article to be relevant or 'key' for policy makers?

Additional comments (optional)

Title article #3
What are the most important insights provided by this article?

Why do you consider this article to be relevant or 'key' for policy makers?

Additional comments (optional)

Appendix F – List of 15 articles found for the broadest research question (1)

Collucia, B., Valente, D., Fusco, G. De Leo, F. & Porrini, D. (2020). Assessing agricultural eco-efficiency in Italian Regions*. Ecological Indicators, 116.*

Conijn, J.G., Bindraban, P.S., Schröder, J.J. & Jonschaap, R.E.E. (2018). Can our global food system meet food demand within planetary boundaries? Agriculture, Ecosystems and Environment, *255,* 244-256.

Fagerholm, N., Torralba, M., Burgess, P.J. & Plieninger, T. (2016). A systematic map of ecosystem services assessments around European agroforestry. Ecological Indicators, 62, 47-65.

Howden, S.M., Soussana, J.F., Tubiello, F.N., Chhetri, N., Dunlop, M. & Meinke, H. (2007). Adapting agriculture to climate change. Proceedings of the National Academy of Sciences of the United States of America, 104(50), 19691-19696.

Jordon, M.W., Willis, K.J., Harvey, W.J., Petrokofsky, L. & Petrokofsky, G. (2020). Implications of temperate agroforestry on sheep and cattle productivity, environmental impacts and enterprise economics. A systematic evidence map. *Forests, 11*(12), 1-29.

Keating, B.A., Carberry, P.S., Bindraban, P.S., Asseng, S., Meinke, H. & Dixon, J. (2010). Eco-efficient agriculture: Concepts, Challenges, and opportunities. *Crop Science, 50,* 109-199.

Lee, H., Lautenbach, S., Nieto, A.P.G., Bondeau, A., Cramer, W., Geijzendorffer, I.R. (2019).The impact of conservation farming practices on Mediterranean agro-ecosystem services provisioning—a meta-analysis*. Regional Environmental Change, 19*(8), 2187-2202.

Nieto-Romero, M., Oteros-Rozas, E., González, J.A. & Martín-López, B. (2014). Exploring the knowledge landscape of ecosystem services assessments in Mediterranean agroecosystems: Insights for future research. *Environmental Science and Policy*, 37, 121-133.

Palomo-Campesino, S., González, J.A., García-Llorente, M. (2018). Exploring the connections between agroecological practices and ecosystem services: A systematic literature review. *Sustainability (Switzerland), 10*(12)

Pretty, J. (2008). Agricultural sustainability: Concepts, principles and evidence. Philosophical Transactions of the Royal Society B: Biological Sciences, 363 (1491), 447-465.

Santiago-Freijanes, J.J., Pisanelli, A., Rois-Diaz, M., Aldrey-Vasquez, J.A. … Mosquera-Losada, M.R. (2018) Agroforestry development in Europe: Policy issues. *Land Use Policy, 76,* 144-156.

Smith, J., Pearce, B.D. & Wolfe, M.S. (2013). Reconciling productivity with protection of the environment: Is temperate agroforestry the answer? *Renewable Agriculture and Food Systems, 28*(1), 80-92.

Springmann, M., Clark, M., Mason-Dçroz, D. … Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature, 562*, 519-525.

Torralba, M., Fagerholm, N., Burgess, P.J., Moreno, G. & Plieninger, T. (2016). Do European agroforestry systems enhance biodiversity and ecosystem services? A meta-analysis. *Agriculture, Ecosystems and Environment, 230,* 150-161,

Van Passel, S., Nevens, F., Mathijs, E. & Van Huylenbroeck, G. (2007). Measuring farm sustainability and explaining differences in sustainable efficiency. *Ecological Economics, 62*(1), 149-161.

Appendix G – 35 relevant articles for the narrowest research question (3)

Aneja, V. P., Blunden, J., James, K., Schlesinger, W. H., Knighton, R., Gilliam, W. . . . Cole, S. (2008). Ammonia assessment from agriculture: U.S. status and needs. *Journal of Environmental Quality, 37*(2), 515-520. doi:10.2134/jeq2007.0002in

Aneja, V. P., Schlesinger, W. H. & Erisman, J. W. (2009). Effects of agriculture upon the air quality and climate: Research, policy, and regulations. *Environmental Science and Technology, 43*(12), 4234-4240. doi:10.1021/es8024403

Biffi, S., Traldi, R., Crezee, B., Beckmann, M., Egli, L., Epp Schmidt, D. . . . Ziv, G. (2021). Aligning agri-environmental subsidies and environmental needs: A comparative analysis between the US and EU. *Environmental Research Letters, 16*(5). doi:10.1088/1748-9326/abfa4e

Blumberga, A., Timma, L., Lauka, D., Dace, E., Barisa, A. & Blumberga, D. (2015). Achieving sustainability in non-ETS sectors using system dynamics modelling practice. *Chemical Engineering Transactions, 45*, 871-876. doi:10.3303/CET1545146

Dace, E., Muizniece, I., Blumberga, A. & Kaczala, F. (2015). Searching for solutions to mitigate greenhouse gas emissions by agricultural policy decisions - Application of system dynamics modeling for the case of Latvia. *Science of the Total Environment, 527-528*, 80-90. doi:10.1016/j.scitotenv.2015.04.088

De Pinto, A., Li, M., Haruna, A., Hyman, G. G., Martinez, M. A. L., Creamer, B. . . . Martinez, J. D. (2016). Low Emission Development Strategies in Agriculture. An Agriculture, Forestry, and Other Land Uses (AFOLU) Perspective. *World Development, 87*, 180-203. doi:10.1016/j.worlddev.2016.06.013

Durandeau, S., Gabrielle, B., Godard, C., Jayet, P. A. & Le Bas, C. (2009). Coupling biophysical and micro-economic models to assess the effect of mitigation measures on greenhouse gas emissions from agriculture. *Climatic Change, 98*(1-2), 51-73. doi:10.1007/s10584-009-9653-8

Fellmann, T., Witzke, P., Weiss, F., Van Doorslaer, B., Drabik, D., Huck, I. . . . Leip, A. (2018). Major challenges of integrating agriculture into climate change mitigation policy frameworks. *Mitigation and Adaptation Strategies for Global Change, 23*(3), 451-468. doi:10.1007/s11027-017-9743-2

Garnett, T. (2009). Livestock-related greenhouse gas emissions: impacts and options for policy makers. *Environmental Science and Policy, 12*(4), 491-503. doi:10.1016/j.envsci.2009.01.006

Gerber, P., Key, N., Portet, F., & Steinfeld, H. (2010). Policy options in addressing livestock's contribution to climate change. *Animal, 4*(3), 393-406. doi:10.1017/S1751731110000133

Haden, V. R., Dempsey, M., Wheeler, S., Salas, W. & Jackson, L. E. (2013). Use of local greenhouse gas inventories to prioritise opportunities for climate action planning and voluntary mitigation by agricultural stakeholders in California. *Journal of Environmental Planning and Management, 56*(4), 553-571. doi:10.1080/09640568.2012.689616

Havlík, P., Valin, H., Herrero, M., Obersteiner, M., Schmid, E., Rufino, M. C. . . . Notenbaert, A. (2014). Climate change mitigation through livestock system transitions. *Proceedings of the National Academy of Sciences of the United States of America, 111*(10), 3709-3714. doi:10.1073/pnas.1308044111

Henseler, M. & Dechow, R. (2014). Simulation of regional nitrous oxide emissions from German agricultural mineral soils: A linkage between an agro-economic model and

an empirical emission model. *Agricultural Systems, 124*, 70-82. doi:10.1016/j.agsy.2013.10.005

Jongebreur, A. A. & Monteny, G. J. (2001). Prevention and control of losses of gaseous nitrogen compounds in livestock operations: a review. *TheScientificWorldJournal, 1 Suppl 2*, 844-851. doi:10.1100/tsw.2001.339

Klavs, G. & Rekis, J. (2016). Introduction of Energy and Climate Mitigation Policy Issues in Energy - Environment Model of Latvia. *Latvian Journal of Physics and Technical Sciences, 53*(6), 12-20. doi:10.1515/lpts-2016-0039

Kwon, H., Liu, X., Xu, H. & Wang, M. (2021). Greenhouse gas mitigation strategies and opportunities for agriculture. *Agronomy Journal*. doi:10.1002/agj2.20844

Leifeld, J. & Fuhrer, J. (2005). Greenhouse gas emissions from Swiss agriculture since 1990: Implications for environmental policies to mitigate global warming. *Environmental Science and Policy, 8*(4), 410-417. doi:10.1016/j.envsci.2005.04.001

Liu, L., Zhang, X., Xu, W., Liu, X., Li, Y., Wei, J. . . . Wu, X. (2020). Challenges for Global Sustainable Nitrogen Management in Agricultural Systems. *Journal of Agricultural and Food Chemistry, 68*(11), 3354-3361. doi:10.1021/acs.jafc.0c00273

Liu, Z. & Liu, Y. (2018). Mitigation of greenhouse gas emissions from animal production. *Greenhouse Gases: Science and Technology, 8*(4), 627-638. doi:10.1002/ghg.1785

Ogle, S. M., McCarl, B. A., Baker, J., Del Grosso, S. J., Adler, P. R., Paustian, K. & Parton, W. J. (2016). Managing the nitrogen cycle to reduce greenhouse gas emissions from crop production and biofuel expansion. *Mitigation and Adaptation Strategies for Global Change, 21*(8), 1197-1212. doi:10.1007/s11027-015-9645-0

Poppe, K., Van Duinen, L. & De Koeijer, T. (2021). Reduction of Greenhouse Gases from Peat Soils in Dutch Agriculture. *EuroChoices, 20*(2), 38-45. doi:10.1111/1746-692X.12326

Saathoff, W., Von Haaren, C., Dechow, R. & Lovett, A. (2013). Farm-level assessment of CO2 and N2O emissions in Lower Saxony and comparison of implementation potentials for mitigation measures in Germany and England. *Regional Environmental Change, 13*(4), 825-841. doi:10.1007/s10113-012-0364-8

Singh, H., Northup, B. K., Baath, G. S., Gowda, P. P. & Kakani, V. G. (2019). Greenhouse mitigation strategies for agronomic and grazing lands of the US Southern Great Plains. *Mitigation and Adaptation Strategies for Global Change*. doi:10.1007/s11027-019-09894-1

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P. . . . Towprayoon, S. (2007). Policy and technological constraints to implementation of greenhouse gas mitigation options in agriculture. *Agriculture, Ecosystems and Environment, 118*(1-4), 6-28. doi:10.1016/j.agee.2006.06.006

Snyder, C. S., Davidson, E. A., Smith, P. & Venterea, R. T. (2014). Agriculture: Sustainable crop and animal production to help mitigate nitrous oxide emissions. *Current Opinion in Environmental Sustainability, 9-10*, 46-54. doi:10.1016/j.cosust.2014.07.005

Solazzo, R., Donati, M., Tomasi, L. & Arfini, F. (2016). How effective is greening policy in reducing GHG emissions from agriculture? Evidence from Italy. *Science of the Total Environment, 573*, 1115-1124. doi:10.1016/j.scitotenv.2016.08.066

Starmans, D. A. J., & Van der Hoek, K. W. (2007). *Ammonia: The case of the Netherlands*. Wageningen: Wageningen Academic Publishers.

Torres, M., & Pinho, P. (2011). Encouraging low carbon policies through a Local Emissions Trading Scheme (LETS). *Cities, 28*(6), 576-582. doi:10.1016/j.cities.2011.06.005

Wagner, S., Angenendt, E., Beletskaya, O. & Zeddies, J. (2017). Assessing ammonia emission abatement measures in agriculture: Farmers' costs and society's benefits – A case study for Lower Saxony, Germany. *Agricultural Systems, 157*, 70-80. doi:10.1016/j.agsy.2017.06.008

Warner, D., Tzilivakis, J., Green, A. & Lewis, K. (2017). Prioritising agri-environment options for greenhouse gas mitigation. *International Journal of Climate Change Strategies and Management, 9*(1), 104-122. doi:10.1108/IJCCSM-04-2015-0048

Wei, S., Bai, Z. H., Chadwick, D., Hou, Y., Qin, W., Zhao, Z. Q. . . . Ma, L. (2018). Greenhouse gas and ammonia emissions and mitigation options from livestock production in peri-urban agriculture: Beijing – A case study. *Journal of Cleaner Production, 178*, 515-525. doi:10.1016/j.jclepro.2017.12.257

Wirsenius, S., Hedenus, F. & Mohlin, K. (2011). Greenhouse gas taxes on animal food products: Rationale, tax scheme and climate mitigation effects. *Climatic Change, 108*(1), 159-184. doi:10.1007/s10584-010-9971-x

York, L., Heffernan, C. & Rymer, C. (2018). A systematic review of policy approaches to dairy sector greenhouse gas (GHG) emission reduction. *Journal of Cleaner Production, 172*, 2216-2224. doi:10.1016/j.jclepro.2017.11.190

Appendix H – Abstracts of Top 12

1 Kleijn, D., Baquero, R. A., Clough, Y., Díaz, M., De Esteban, J., Fernández, F., ... & Yela, J. L. (2006). Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology letters*, *9*(3), 243-254. DOI

Agri-environment schemes are an increasingly important tool for the maintenance and restoration of farmland biodiversity in Europe but their ecological effects are poorly known. Scheme design is partly based on non-ecological considerations and poses important restrictions on evaluation studies. We describe a robust approach to evaluate agri-environment schemes and use it to evaluate the biodiversity effects of agri-environment schemes in five European countries. We compared species density of vascular plants, birds, bees, grasshoppers and crickets, and spiders on 202 paired fields, one with an agri-environment scheme, the other conventionally managed. In all countries, agri-environment schemes had marginal to moderately positive effects on biodiversity. However, uncommon species benefited in only two of five countries and species listed in Red Data Books rarely benefited from agri-environment schemes. Scheme objectives may need to differentiate between biodiversity of common species that can be enhanced with relatively simple modifications in farming practices and diversity or abundance of endangered species which require more elaborate conservation measures.

2 Scheper, J., Holzschuh, A., Kuussaari, M., Potts, S. G., Rundlöf, M., Smith, H. G., & Kleijn, D. (2013). Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss–a meta-analysis. *Ecology letters*, *16*(7), 912-920. DOI

In Europe, agri-environmental schemes (AES) have been introduced in response to concerns about farmland biodiversity declines. Yet, as AES have delivered variable results, a better understanding of what determines their success or failure is urgently needed. Focusing on pollinating insects, we quantitatively reviewed how environmental factors affect the effectiveness of AES. Our results suggest that the ecological contrast in floral resources created by schemes drives the response of pollinators to AES but that this response is moderated by landscape context and farmland type, with more positive responses in croplands (vs. grasslands) located in simple (vs. cleared or complex) landscapes. These findings inform us how to promote pollinators and associated pollination services in species-poor landscapes. They do not, however, present viable strategies to mitigate loss of threatened or endangered species. This indicates that the objectives and design of AES should distinguish more clearly between biodiversity conservation and delivery of ecosystem services.

4 Piñeiro, V., Arias, J., Dürr, J., Elverdin, P., Ibáñez, A. M., Kinengyere, A., ... & Torero, M. (2020). A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes. *Nature Sustainability*, *3*(10), 809-820. DOI

The increasing pressure on agricultural production systems to achieve global food security and prevent environmental degradation necessitates a transition towards more sustainable practices. The purpose of this scoping review is to understand how the incentives offered to farmers motivate the adoption of sustainable agricultural practices and, ultimately, how and whether they result in measurable outcomes. To this end, this scoping review examines the evidence of nearly 18,000 papers on whether incentive-based programmes lead to the

adoption of sustainable practices and their effect on environmental, economic and productivity outcomes. We find that independent of the incentive type, programmes linked to short-term economic benefit have a higher adoption rate than those aimed solely at providing an ecological service. In the long run, one of the strongest motivations for farmers to adopt sustainable practices is perceived benefits for either their farms, the environment or both. Beyond this, the importance of technical assistance and extension services in promoting sustainable practices emerges strongly from this scoping review. Finally, we find that policy instruments are more effective if their design considers the characteristics of the target population, and the associated trade-offs between economic, environmental and social outcomes.

5 Biffi, S., Traldi, R., Crezee, B., Beckmann, M., Egli, L., Schmidt, D. E., ... Ziv, G. (2021). Aligning agri-environmental subsidies and environmental needs: a comparative analysis between the US and EU. *Environmental Research Letters*, *16*(5), 054067. DOI

The global recognition of modern agricultural practices' impact on the environment has fuelled policy responses to ameliorate environmental degradation in agricultural landscapes. In the US and the EU, agri-environmental subsidies (AES) promote widespread adoption of sustainable practices by compensating farmers who voluntarily implement them on working farmland. Previous studies, however, have suggested limitations of their spatial targeting, with funds not allocated towards areas of the greatest environmental need. We analysed AES in the US and EU - specifically through the Environmental Quality Incentives Program (EQIP) and selected measures of the European Agricultural Fund for Rural Development (EAFRD) - to identify if AES are going where they are most needed to achieve environmental goals, using a set of environmental need indicators, socio-economic variables moderating allocation patterns, and contextual variables describing agricultural systems. Using linear mixed models and linear models we explored the associations among AES allocation and these predictors at different scales. We found that higher AES spending was associated with areas of low soil organic carbon and high greenhouse gas emissions both in the US and EU, and nitrogen surplus in the EU. More so than successes, however, clear mismatches of funding and environmental need emerged - AES allocation did not successfully target areas of highest water stress, biodiversity loss, soil erosion, and nutrient runoff. Socio-economic and agricultural context variables may explain some of these mismatches; we show that AES were allocated to areas with higher proportions of female producers in the EU but not in the US, where funds were directed towards areas with less tenant farmers. Moreover, we suggest that the potential for AES to remediate environmental issues may be curtailed by limited participation in intensive agricultural landscapes. These findings can help inform refinements to EQIP and EAFRD allocation mechanisms and identify opportunities for improving future targeting of AES spending.

7 Batáry, P., Dicks, L. V., Kleijn, D., & Sutherland, W. J. (2015). The role of agri-environment schemes in conservation and environmental management. *Conservation Biology*, *29*(4), 1006-1016. DOI

Over half of the European landscape is under agricultural management and has been for millennia. Many species and ecosystems of conservation concern in Europe depend on agricultural management and are showing ongoing declines. Agri-environment schemes (AES) are designed partly to address this. They are a major source of nature conservation funding within the European Union (EU) and the highest conservation expenditure in Europe.

We reviewed the structure of current AES across Europe. Since a 2003 review questioned the overall effectiveness of AES for biodiversity, there has been a plethora of case studies and meta-analyses examining their effectiveness. Most syntheses demonstrate general increases in farmland biodiversity in response to AES, with the size of the effect depending on the structure and management of the surrounding landscape. This is important in the light of successive EU enlargement and ongoing reforms of AES. We examined the change in effect size over time by merging the data sets of 3 recent meta-analyses and found that schemes implemented after revision of the EU's agri-environmental programs in 2007 were not more effective than schemes implemented before revision. Furthermore, schemes aimed at areas out of production (such as field margins and hedgerows) are more effective at enhancing species richness than those aimed at productive areas (such as arable crops or grasslands). Outstanding research questions include whether AES enhance ecosystem services, whether they are more effective in agriculturally marginal areas than in intensively farmed areas, whether they are more or less cost-effective for farmland biodiversity than protected areas, and how much their effectiveness is influenced by farmer training and advice? The general lesson from the European experience is that AES can be effective for conserving wildlife on farmland, but they are expensive and need to be carefully designed and targeted.

23 Brown, C., Kovács, E., Herzon, I., Villamayor-Tomas, S., Albizua, A., Galanaki, A., ... & Zinngrebe, Y. (2021). Simplistic understandings of farmer motivations could undermine the environmental potential of the Common Agricultural Policy. *Land Use Policy*, *101*, 105136. DOI

The European Union Common Agricultural Policy (CAP) has failed to achieve its aim of preserving European farmland biodiversity, despite massive investment in subsidies to incentivise environmentally-beneficial farming practices. This failure calls into question the design of the subsidy schemes, which are intended to either function as a safety net and make farming profitable or compensate farmers for costs and loss of income while undertaking environmental management. In this study, we assess whether the design of environmental payments in the CAP reflects current knowledge about farmers' decision-making as found in the research literature. We do so on the basis of a comprehensive literature review on farmers' uptake of agri-environmental management practices over the past 10 years and interviews specifically focused on Ecological Focus Areas with policy-makers, advisors and farmers in seven European countries. We find that economic and structural factors are the most commonly-identified determinants of farmers' adoption of environmental management practices in the literature and in interviews. However, the literature suggests that these are complemented by – and partially dependent on – a broad range of social, attitudinal and other contextual factors that are not recognised in interview responses or, potentially, in policy design. The relatively simplistic conceptualisation of farmer behaviour that underlies some aspects of policy design may hamper the effectiveness of environmental payments in the CAP by over-emphasising economic considerations, potentially corroding farmer attitudes to policy and environmental objectives. We conclude that an urgent redesign of agricultural subsidies is needed to better align them with the economic, social and environmental factors affecting farmer decision-making in a complex production climate, and therefore to maximise potential environmental benefits.

24 Kleijn, D., Berendse, F., Smit, R., Gilissen, N., Smit, J., Brak, B., & Groeneveld, R. (2004). Ecological effectiveness of agri-environment schemes in different agricultural landscapes in the Netherlands. *Conservation biology*, *18*(3), 775-786. DOI

Agri-environment schemes are an instrument used by western European countries to counteract the negative effects of contemporary agriculture on biodiversity, but not much is known about their effectiveness. We investigated the ecological effects of Dutch agri-environment schemes aimed at promoting botanical diversity or meadow birds, and we tested whether the effectiveness of the schemes depends on landscape type or structure. In three different types of landscape, we surveyed plants, birds, bees, and hover flies on 78 paired fields that either had agri-environment schemes or were managed conventionally, and we collected data on a range of different environmental variables. Neither plant species richness nor abundance of meadow birds was higher on fields with agri-environment schemes. Landscape type had a significant effect on both species groups, but the effects of the schemes were independent of landscape type. Neither the diversity of plants nor the abundance of birds was related to any of the environmental variables. Agri-environment schemes designed to promote plant species richness or bird abundance did have positive side-effects because they enhanced the species richness of bees and hover flies, irrespective of the type of landscape. Furthermore, landscape type, groundwater level (hover flies), and area of wooded edges (bees) significantly affected both species groups. The failure of the schemes to promote the target species may be related to the high intensity of land use in The Netherlands. Simple conservation measures taken by farmers may not be sufficient to counteract the impact of factors that are often controlled at the landscape level (e.g., hydrology). Similar studies in other countries are needed to place the results of our study into a European context.

25 Van Dijk, W. F., Lokhorst, A. M., Berendse, F., & de Snoo, G. R. (2015). Collective agri-environment schemes: How can regional environmental cooperatives enhance farmers' intentions for agri-environment schemes?. *Land Use Policy*, *42*, 759-766. DOI

The effectiveness of agri-environment schemes (AES) in enhancing biodiversity on farmland and creating a long-lasting change in farmers' motivation towards a more environmental-friendly practice is still strongly debated. Applying a regional approach has been advocated widely to make AES more ecologically and socially sustainable. In the Netherlands, some AES are performed collectively by large regional groups of farmers called Environmental Cooperatives (EC). We hypothesise that these cooperatives enhance farmers' intention to participate by facilitating the application of AES, but also by generating group pressure. In the study at hand, we used an extended version of the Theory of Planned Behaviour (TPB) to investigate which factors are associated with farmers' intention to participate in two kinds of collective AES (ditch bank management and the protection of meadow birds). Our results demonstrate that attitude and perceived personal ability to participate in these AES are associated with the intention of farmers to participate in ditch bank management. However, for the protection of meadow birds, social pressure, self-identity and facilitation by the EC also relate to the intention of farmers. We conclude that the facilitation undertaken by ECs positively relates to farmers' intention to participate in collective AES.

31 Merckx, T., & Pereira, H. M. (2015). Reshaping agri-environmental subsidies: From marginal farming to large-scale rewilding. *Basic and Applied Ecology*, *16*(2), 95-103. DOI

Despite continued discussion about market distortions and environmental impacts, agricultural subsidies continue to be a key component of European Union policy. About 10% of the agro-forestry subsidies are targeted at supporting agri-environment schemes, and at supporting farming in Less Favoured Areas (LFA) such as mountain regions. One of the main justifications for these agri-environmental subsidies towards marginally productive land is that they promote the conservation of biodiversity by maintaining low-intensity farming practices. Here, we critically examine this assumption and argue instead for a two-tier approach to Europe's agri-environmental policy based on inherent land fertility and spatial scale: (i) at a local, single-farm scale, fertile agricultural land should preferentially be intensively but sustainably farmed with a focus on high yields, (ii) while simultaneously and at larger, regional scales, less-productive land, and especially protected areas, may be ecologically restored into 'wild' and resiliently functioning ecosystems. As such, agri-environmental subsidies towards fertile land should support the implementation of measures that benefit biodiversity while allowing, and even helping, the achievement of high agricultural yields. In contrast, agri-environmental and LFA subsidies towards marginal land and protected areas should also promote rewilding and the management of natural succession. In order for this approach to be successful, a higher proportion of the Common Agricultural Policy subsidies would need to be allocated to environmental goals.

32 Zingg, S., Ritschard, E., Arlettaz, R., & Humbert, J. Y. (2019). Increasing the proportion and quality of land under agri-environment schemes promotes birds and butterflies at the landscape scale. *Biological conservation*, *231*, 39-48. DOI

The intensification of agricultural practices that Western nations have experienced after World War II has led to an alarming decline in farmland biodiversity. With the aim of stopping and even reversing this decline, agri-environment schemes (AES) have been implemented in many European countries since the 1990s. In Switzerland, farmers are required to manage at least 7% of their land in the form of biodiversity promotion areas (BPA), which are extensively managed, wildlife-friendly farmland habitats such as hay meadows and traditional orchards. We investigated how the occurrence and characteristics of these BPA influence birds and butterflies in the Swiss lowlands. Butterfly species richness and abundance increased by 22% and 60%, respectively, when the proportion of BPA in the landscape increased from 5% to 15%. Likewise, bird species richness increased, but to a lesser extent, with the proportion of BPA in the landscape. For birds, the proportion of BPA characterized by a high ecological quality played a role in promoting both priority-farmland and red-listed species. For both taxonomic groups, the amount and quality of BPA habitats contributed more to species richness than their spatial configuration, connectivity included. This study shows that AES measures implemented at the field scale have positive effects on mobile species that are noticeable at the landscape scale, and that the fraction of AES in the cultivated landscape matters more than their spatial configuration, which has strong implications for designing multi-functional agro-ecosystems.

62 Mouysset, L. (2014). Agricultural public policy: Green or sustainable?. *Ecological Economics*, *102*, 15-23. DOI

The future of agriculture constitutes a major challenge to the achievement of sustainable development. There are new perspectives on greening (focusing on ecological objectives) and sustainability (combining both ecological and social goals). Academic papers mainly study the ecological efficiency of agricultural public policies, while real public policies, such as the European Common Agricultural Policy, examine both ecological and social considerations. The objective of this paper is to consider economic, social and ecological objectives within the design of agricultural public policies. Using a bio-economic model applied to France, we compare different optimal public strategies. We show that, when the biodiversity objectives are either very limited or very demanding, grassland subsidies are the best instruments from both green and sustainable points of view. However for medium objectives, reducing crops subsidies is the cheapest way to green the CAP, while subsidies on grasslands are the only strategy from a sustainability perspective. Our work highlights new trade-offs related to policy implementation, such as social acceptance or technical difficulties, and the spatial equity of performance among regions.

100 Sanders, M. E., Nieuwenhuizen, W., Dirkx, G. H. P., Schrijver, R. A. M., & Smidt, R. A. (2013). Bedrijfsvoering zit in de weg: Landbouw slechts beperkt inzetbaar voor natuur-en landschapsbehoud. *Landschap: tijdschrift voor landschapsecologie en milieukunde*, *30*(2), 57-66. DOI

The Dutch government wants to realize nature and landscape targets on agricultural land. However, there is serious criticism on the effects of agri-environmental schemes. It turns out, almost all farmers choose only easy management types while more demanding management types with better prospects for nature are less popular. We noticed that the demanding types on a large portion of the farm area, need another agricultural management that is not financially stimulated by the government. Hence, agricultural schemes are not the solution. We think that a transition to sustainable agriculture does not yet support nature and landscape, but has the potential to do so.