LIFT

Low-Input Farming and Territories – Integrating knowledge for improving ecosystem based farming

Research and Innovation action: H2020 – 770747
Call: H2020-SFS-2016-2017
Type of action: Research and Innovation Action (RIA)
Work programme topic: SFS-29-2017
Duration of the project: 01 May 2018 – 30 April 2022

Farm environmental performance depending on the degree of ecological approaches

Kato Van Ruymbeke[1], Hervé Dakpo[2], Laure Latruffe[2], Liesbet Vranken[1] *

[1] KU Leuven (Belgium), [2] INRAE (France)

* Deliverable leader – Contact: liesbet.vranken@kuleuven.be

DELIVERABLE D3.3

Workpackage N°3

Due date: M38

Actual delivery date: 30/06/2021

Dissemination level: Public

About the LIFT research project

Ecological approaches to farming practices are gaining interest across Europe. As this interest grows there is a pressing need to assess the potential contributions these practices may make, the contexts in which they function and their attractiveness to farmers as potential adopters. In particular, ecological agriculture must be assessed against the aim of promoting the improved performance and sustainability of farms, rural environment, rural societies and economies, together.

The overall goal of LIFT is to identify the potential benefits of the adoption of ecological farming in the European Union (EU) and to understand how socio-economic and policy factors impact the adoption, performance and sustainability of ecological farming at various scales, from the level of the single farm to that of a territory.

To meet this goal, LIFT will assess the determinants of adoption of ecological approaches, and evaluate the performance and overall sustainability of these approaches in comparison to more conventional agriculture across a range of farm systems and geographic scales. LIFT will also develop new private arrangements and policy instruments that could improve the adoption and subsequent performance and sustainability of the rural nexus. For this, LIFT will suggest an innovative framework for multi-scale sustainability assessment aimed at identifying critical paths toward the adoption of ecological approaches to enhance public goods and ecosystem services delivery. This will be achieved through the integration of transdisciplinary scientific knowledge and stakeholder expertise to co-develop innovative decision-support tools.

The project will inform and support EU priorities relating to agriculture and the environment in order to promote the performance and sustainability of the combined rural system. At least 30 case studies will be performed in order to reflect the enormous variety in the socio-economic and bio-physical conditions for agriculture across the EU.

Project consortium

No.	Participant organisation name	Country
1	**INRAE - Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement**	**FR**
2	VetAgro Sup – Institut d'enseignement supérieur et de recherche en alimentation, santé animale, sciences agronomiques et de l'environnement	FR
3	SRUC – Scotland's Rural College	UK
4	Teagasc – Agriculture and Food Development Authority	IE
5	KU Leuven – Katholieke Universiteit Leuven	BE
6	SLU – Sveriges Lantbruksuniversitet	SE
7	UNIBO – Alma Mater Studiorum – Universita di Bologna	IT
8	BOKU – Universitaet fuer Bodenkultur Wien	AT
9	UBO – Rheinische Friedrich-Wilhelms – Universitat Bonn	DE
10	JRC – Joint Research Centre – European Commission	BE
11	IAE-AR – Institute of Agricultural Economics	RO
12	MTA KRTK – Magyar Tudományos Akadémia Közgazdaság – és Regionális Tudományi Kutatóközpont	HU
13	IRWiR PAN – Instytut Rozwoju Wsi i Rolnictwa Polskiej Akademii Nauk	PL
14	DEMETER – Hellinikos Georgikos Organismos – DIMITRA	GR
15	UNIKENT – University of Kent	UK
16	IT – INRAE Transfert S.A.	FR
17	ECOZEPT Deutschland	DE

Table of contents

List of acronyms and abbreviations

AECM	agri-environmental and climatic measures
AES	agri-environmental scheme
ATT	average treatment effect on the treated
BT	Benefit transfer
C	carbon
CAP	Common Agricultural Policy
CDI	crop diversity index
CICES	Common International Classification of Ecosystem Services
CI	Composite indicator
DEA	Data Envelopment Analysis
DCM	Direct Covariate Matching
ES	Ecosystem service
FADN	Farm Accountancy Data Network
GHG	greenhouse gases
GIS	Global Information System
GS	grass strip
GWP	Global Warming Potential
ha	hectare
HI	Herfindahl index
HR	hedgerow
IPM	Integrated Pest Management
LCA	Life Cycle Analysis
LFA	Less Favoured Area
N	nitrogen
NFS	Ireland's National Farm Survey
MAC	marginal abatement cost
P	phosphorus
PLN	Polish zloty
PSM	Propensity Score Matching
REA	rapid evidence assessment
SI	Sub-indicator

SOC Soil Organic Carbon

SFA Stochastic Frontier Analysis

TSG traditional speciality guaranteed

UAA Utilised Agricultural Area

1 Summary

This deliverable (D3.3 of the LIFT project) presents the results of a series of analyses carried out to evaluate the environmental performance of (ecological) farm management practices at the farm level. We use secondary data collected through a variety of approaches in an effort to evaluate environmental performance across various dimensions, from a qualitative description, through a quantitative assessment to an empirical analysis.

The deliverable presents summaries from three distinct analyses that were carried out within Task 3.4 of the LIFT project. The analyses proceed in a pyramid-approach fashion, in which the most broad analysis is presented first, and all subsequent analyses presented increase in nuance and complexity. Prior to carrying out these analyses in this task, a rapid evidence assessment (REA) was performed. Evidence collected through the REA was compiled in a database and formed the basis for the subsequent qualitative and quantitative assessments of environmental performance of farm management practices.

The pyramid first presents results from the qualitative and quantitative evaluation of the impact of (ecological) farm management practices on the supply of ecosystem services (ES) in European agriculture. Quantitative, semi-quantitative and qualitative data was collected through the REA. Semi-quantitative data was then used to compose sub-indicators (SIs), presenting a dimensionless quantification of the impact of 26 management practices across 17 ES. Second, a monetary valuation of six ES was performed using a benefit transfer. The benefit transfer was carried out for two management practices in the Belgian Flemish case study Hageland-Haspengouw, namely for grass strips and hedgerows.

Finally, we present summarised results from the various empirical analyses on farm environmental performance that have been carried out across the LIFT project. Results from four main approaches are summarised: (i) farm bio-economic model enabling the computation of environmental indicators for a specific farm; (ii) computation of environmental indicators and comparing them across types of ecological farms; (iii) farm efficiency, including in the model both economic and environmental inputs or outputs on the farm; and (iv) environmental damage computed at the farm or the plot level. Due to the scope of the empirical analyses, these were not performed based on the evidence compiled through the REA mentioned above. More extensive descriptions of the empirical analyses of farm environmental performance can be found in LIFT Deliverable 3.1 (Niedermayr et al., 2021a) describing the empirical analyses on farm economic performance (Task 3.2 of the LIFT project).

The formulation of this deliverable and the work presented within allows us to obtain a comprehensive overview of the impact of various farm management practices on the supply of a number of ES. We see that methodologies differ in the scope and depth of results they are able to capture. As the analytical methodologies become more complex, the number of management practices and ES that can be considered within the analysis decreases. Simultaneously we see that as the number of management practices/ES evaluated decreases, the nuance in the results increases. As such, by providing an overview of results from a variety of methodologies in this deliverable, we are able to present both broad picture and nuanced results on the environmental performance of ecological farm management practices.

2 Introduction

Agroecosystems are arguably among the most important ecosystems to sustain human wellbeing. Not only do we rely on these systems for the provisioning of food and energy materials, we also derived many secondary benefits from them such as recreation, regulation of natural hazards and carbon sequestration. Historically, however, these systems have been primarily managed to sustain food production and other provisioning services (Sandhu et al., 2010; Swinton et al., 2007, 2006), with preservation of secondary benefits remaining largely on the backburner. That is not to say both primary and secondary benefits cannot be maintained simultaneously. Through well-planned and regulated farm management practices (both conventional and ecological[1]), we can manage agroecosystems to find a balance between meeting demands for productive output and maximising environmental performance to ensure long-term sustainability (Bateman et al., 2009; Pretty, 2008; Pretty and Bharucha, 2014; Wezel et al., 2017).

When evaluating the environmental performance of farm management practices, most studies adopt the ecosystem service (ES) concept (Turner and Daily, 2008). ES can be defined as the direct or indirect contribution of ecosystems to human well-being, capturing human-nature interactions (Haines-Young and Potschin, 2018). The Common International Classification of Ecosystem Services (CICES) categorises ES into three broad categories, i) regulating and maintaining services, which help maintain proper functioning of ecosystems (e.g. biodiversity), ii) provisioning services, which supply productive output that can be directly exploited (e.g. crop production), and iii) cultural services, which influence people's mental and physical wellbeing through non-material characteristics of an ecosystem (CICES, 2018). ES also contain a spatial component, with many ES emerging only if minimum scale threshold of specific service-providing processes/functions are met (Andersson et al., 2015). Relevant scale thresholds vary between ES from global (e.g. global climate mitigation) to plot level (e.g. pest control) (ibid). The human dimension of the human-nature interactions captured by the ES concept is also spatially explicit, with demand for certain services often driven by socio-cultural and/or geographic conditions (Potschin and Haines-Young, 2011).

A notable number of the ES produced by agroecosystems have seen steady degradation with recent rapid population growth and sub- and exurban expansion (Kirchner et al., 2015; Kroeger and Casey, 2007). Therefore it is worthwhile to make an assessment of ES to increase awareness amongst farmers, policy makers and the wider public to ensure their consideration when establishing farm management policies. Assessing ES can be achieved through a qualitative and/or quantitative approach (Busch et al., 2012). Qualitative assessment of ES involves the qualitative description of the state of ES stocks/flows, and is commonly approached through an evidence synthesis in which evidence from primary studies is compiled and described (Busch et al., 2012). Quantitative assessments involve the quantification of (changes in) ES stocks/flows, either in biophysical or monetary units.

The biophysical quantification of ES involves the direct or indirect measure of the state of an ecosystem. The former is achieved through surveys and questionnaires, in-field observations, and monitoring, while the latter requires further interpretation or combination of biophysical measures with other

[1] Ecological practices are understood in LIFT as low-input practices and/or practices that are environmentally friendly. The originality of LIFT in this view is not to focus on a specific type of ecological approaches, but to cover the whole continuum of farming approaches, from the most conventional to the most ecological, including the widest range of ecological approaches. This comprises the existing nomenclatures such as organic farming, low-input farming, agroecological farming, etc. It also encompasses approaches that are not yet part of a nomenclature, but that can be identified with various criteria such as management practices, on-farm diversification etc. Thus, conventional practices mean non-ecological practices.

environmental data (Burkhard and Maes, 2017). Both direct and indirect measurements of ES are expressed in biophysical units which vary depending on the ES assessed. As not all ES can be expressed in biophysical units, biophysical quantification is not appropriate for all ES. Such quantification is particularly challenging for cultural services, which are often intangible and not readily quantified.

Monetary quantification of ES is achieved through economic assessments, either marker-based or non-market based (Busch et al., 2012; Costanza et al., 2017; de Groot et al., 2002; Gómez-Baggethun et al., 2010; Kroeger and Casey, 2007). In such assessments, the stock/flow of an ES is evaluated and quantified through attributing monetary value to the (changes in) stocks/flows. Expressing ES in monetary units is valuable as this facilitates their consideration amongst land managers and policy makers. Furthermore, the commensurable nature of monetary units allows for trade-offs and synergies between services to be more readily evaluated, further promoting their integration into rural development schemes (Busch et al., 2012; Chan et al., 2012). Despite this, monetary valuation of ES is subject to critique within the literature. Firstly, such an approach assumes the various dimensions of services are commensurable, implying the value of services may be reduced to a single dimension (Kroeger and Casey, 2007; Rawluk et al., 2019). Secondly, the universal nature of money suggests that ES value can be considered in isolation from individuals' relation and history with place, which is not always the case (Rawluk et al., 2019). And third, there are fears that monetary assessments of ES open the way for the privatisation and commodification of nature (Chan et al., 2012).

In this deliverable we adopt a pyramid approach to evaluate the environmental performance of farm management practices, both conventional and ecological, through an assessment across various ES. This assessment consists of three broad dimensions. First, the provisioning of ES from farm management practices is qualitatively assessed based on a rapid evidence assessment. This is then followed by a quantitative assessment through the composition of sub- and composite indicators, measured potential ES provision and overall environmental performance of management practices respectively. Second, a quantitative monetary assessment is performed on a handful of ES through a benefit transfer. Lastly, environmental farm performance, quantitative performance (productivity) or financial performance (profitability) of the various farm management practices is evaluated.

Results from each dimension of the task described above are summarised in the body of this deliverable. Sections 3, 4 and 5 describe the methodology adopted for, and the main results obtain by, the qualitative and quantitative assessment, the quantitative monetary assessment and the farm performance analysis respectively. A more extensive description of the methodologies and results of the respective dimensions are available upon request from the authors.

3 Qualitative and quantitative ES assessment

Qualitative and quantitative assessments of the ability of (ecological) farm management practices to supply ES are assessed through a systematic evidence synthesis. Specifically, we perform a REA to explore these linkages between practices and ES across various farming systems throughout Europe. Evidence derived from the REA is then used to compose sub-indicators, evaluating potential supply of ES from management practices, as well as composite indicators, evaluating overall environmental performance of a single practice. Firstly we compile a database of synthesised results through an REA of the literature, delineating the impacts of 26 farm management practices on 17 ES across farming systems across Europe. Observations from this database are then used to quantify the impact of management practices on ES through the calculation of sub-indicators (SIs). Finally, SIs are aggregated into composite indicators (CIs) quantifying overall environmental performance of the management practices in three case study areas across Belgium and England.

The research objectives of this exercise are framed within the Ecosystem Service Cascade model as proposed by (Potschin and Haines-Young, 2011). The Cascade model presents a 'production chain' in which ecological structures and processes are linked with human wellbeing through a five-tiered process. Biophysical structures and processes underpin ecosystem functions, which in turn give rise to ES. ES then provide benefits through their (direct or indirect) exploitation by society, resulting in welfare gains. These benefits in turn are valued based on the socio-economic, cultural and spatiotemporal context of the system (ibid). The Cascade model can be interpreted as a social-ecological system, in which humans are considered a part of – rather than separate from – nature (Folke, 2007), and to which a supply/demand relationship may be applied. The ecological dimension of the Cascade model is commensurable with the concept of supply, whereby ecosystems supply ES. At the same time, the contextual characteristics specified in the social dimension of the model may be considered the underlying functions determining demand for particular ES. When supply of and demand for an ES spatially and temporally overlap, said ES is delivered. Without this spatial and temporal overlap, ES may be supplied by the ecological dimension and may be demanded by the social dimension, but there will be no delivery.

While the complete description of results found in appendix pertain to two spatial scales, farm and territorial level, in this deliverable we focus only on the farm level results. Spatial scales are defined across two dimensions, depending on whether the supply of or demand for an ES is being assessed. Specifically, the supply of ES (evaluated through the SIs) warrants delineating farm-level based on the geographical scales at which a management practice is implemented. Spatial delineation of the demand for ES (capture by CIs), on the other hand, is defined by end-users. Here, farm level end-users are defined as the farmers themselves.

Through this exercise we provide what is to our knowledge, a first attempt at incorporating such two methodologies to gain a comprehensive overview of environmental performance of farm management practices in Europe. We address two main research questions: i) how do various farm management practices (both conventional and ecological) impact the delivery of ES in agroecosystems across Europe?, and ii) what is the overall environmental performance of (conventional and ecological) farm management practices? We aim for this approach to be considered as a framework which can be applied in various case study areas throughout Europe. Similarly to Rigby et al. (2001), we do not claim that the indicators presented in this paper are decisive of either farm management practice impacts on ES nor of environmental performance of said practices overall. Rather we assert that the proposed indicators are valuable in that they provide a first attempt at summarising the multitude of evidence available in the literature in a concise, intuitive and transparent manner. In this way we hope to gain a better understanding of the current state of affairs in the literature, identify where (and which) evidence is missing, and open up a discussion on how to go about utilising the information we already have and filling the remaining information gaps.

3.1 Methodology

3.2 Methodology

3.2.1 Rapid evidence assessment

Input provided by eight LIFT partners was combined with results from LIFT deliverable D1.1 (Rega et al., 2018) to select the 26 relevant management practices included in this assessment. Based on these selected practices, a search string from which articles were derived in WebofScience was composed through an iterative process. This process consisted of formulating a search string for the individual management practices, combining these into a composite search string, and then evaluating the search string results against the inclusion of a pre-defined set of reference articles. The comprehensive search string (Table A2), the full list of management practices (Table A1) as well as the reference articles included in this assessment can be found in Annex 1.

Inclusion/exclusion criteria (described in Table A3) were used to determine the relevance of articles derived from the final search string. Using the inclusion/exclusion criteria, 647 articles were selected for inclusion based on title and abstract level screening. Reviewers consisting of experts from the LIFT project consortium, extracted meta-analytic data of the articles such as type of review, location of study, management practices adopted, specific practice(s) implemented, spatial scale considered, and ES assessed. A targeted selection of these 647 articles was conducted for full text reading. Targeted sampling consisted of, where possible, selecting five articles (of which one a meta-analysis) per management practice. This resulted in a total of 105 articles that were included in the final REA. At full text screening 10 more articles were excluded based on exclusion criteria, resulting in a final corpus of 95 articles.

For each article quantitative, semi-quantitative and qualitative data for the link between management practices and ES supply was extracted into an excel database. Semi-quantitative data was expressed as either a positive, inconclusive, or a negative relationship between the management practices evaluated and the ES assessed. As semi-quantitative data was extracted for all articles, but quantitative data was not, it was opted to use only the former for SI calculation. As such, observations are henceforth defined as semi-quantitative observations reflecting the potential supply of an ES from a management practice, which were coded as 1 (negative), 2 (inconclusive) and 3 (positive). An evaluation of overall article quality was also performed by each reviewer based on a set of 21 quality criteria adapted from Beillouin et al. (2019) and PRISMA (2015). The list of the 21 quality criteria can be found in Table A4 in Annex 1. Figures A1 through A3 in Annex 1 illustrate the template that was used to extract data from the articles reviewed.

3.2.2 Sub-indicator composition

Using semi-quantitative observations derived from the REA, SIs were calculated using a weighted arithmetic mean to reflect the potential supply of an ES from a single management practice. A measure of article quality, as well as a measure of the number of articles from which observations were derived were incorporated into SI calculations to increase transparency. Figure 1 illustrates the SI calculation process visually.

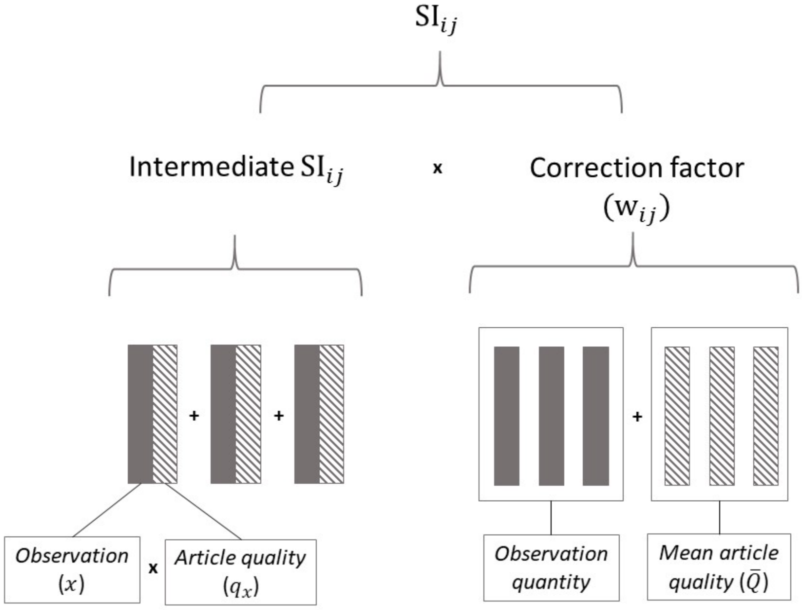

Figure 1. Visual representation of the SI calculation process. The intermediate SI_{ij} (the sum product across multiple observations (x) and their respective article quality score (q_x)) is multiplied by the correction factor (w_{ij}) to obtain SI_{ij} for each management practice i linked to ecosystem service j. The correction factor is composed of a measure of the quantity of observations and the average article quality ($\bar{Q}$) across all articles included in SI_{ij}.

SIs were composed following a three-step approach. First, observations were weighted against average article quality using a weighted arithmetic mean. Second, average article quality and number of articles were incorporated into a single correction factor per SI. Finally, the weighting factor was incorporated with the weighted observations into a single SI for each management practice i and ES j.

Reported SIs are supplemented with the respective correction factor (w), consensus value (c), number of observations (no) and number of articles (na). These supplementary measures ensure transparency and facilitate the interpretation of the SI. SIs range from -1 to +1; from a negative supply (i.e. the management practice has a negative impact on supply) to a positive supply. All supplementary measures are specific to an SI. The correction factor (w) ranges between 0-1 and illustrates the trade-off between the quantity and quality of the evidence incorporated in the SI calculation. Average article quality is corrected downwards or upwards relative to the distance between the observed evidence quantity (no) and the mean evidence quantity per SI at the farm ($\overline{na}_{farm}$) and territorial ($\overline{na}_{terr}$) level respectively. The consensus value reports the level of agreement amongst observations, and is reported on a scale of 0 (no consensus) to 1 (complete consensus). Number of observations and number of articles are reported for the sake of increasing transparency.

3.2.3 Composite indicator composition

Each management practice may influence more than one ES. In order to assess the environmental performance of a management practice overall, the relevant SIs were aggregated into a single composite indicator (CI) per management practice. To limit the degree of compensability allowed between ES in the aggregation, it was opted to implement a weighted geometric aggregation. To capture spatial variation in environmental performance as a result of variation in demands for ES, CIs were calculated for three different case study areas. Regional difference in ES demand was accounted for by deriving distinct weights for each case study area at both spatial scales.

ES weights, obtained through stakeholder engagement, reflect the relative importance of ES within three case study areas across western Europe: one area, Hageland-Haspengouw (HH), located in Flanders, Belgium, and two areas, North Kent (NK) and High Weald (HW) located in England, UK.

A comprehensive overview of the mathematical derivation of both SIs and CIs is available upon request. Sensitivity analyses were performed in which key assumptions of SI and CI composition described in Annex 2 were relaxed in order to evaluate how these affect the proposed indicators.

3.3 Results

SIs were calculated to quantify the potential supply of 17 ES from a set of 26 farm management practices. A total of 193 SIs were calculated, of which 133 SIs were calculated at farm level. A significant difference was found in the number of SIs calculated for each of the three CICES ES categories (CICES, 2018); significantly more SIs were calculated for regulating and maintaining ES compared to provisioning and cultural ES (P<0.001). The most frequently evaluated ES in the literature at the farm level was production (consisting of both crop and livestock production), followed by decontamination and fixing processes, and disease and pest control. The most frequently evaluated management practices at farm level were conservation tillage, crop rotation, mulching and the use of organic fertilisers. No significant differences were observed between the weighting factors of the SIs and the ES classes.

Agri-environmental schemes (AES) had the highest positive impact on ES supply for its impact on pollination services ($SI = 0.80, w = 0.80, c = 1, no = 4, na = 1$). This SI was calculated based on four observations derived from a single article. The directionality of the SI reflects the potential of a management practice to supply a particular ES. The correction factor (w) contains a trade-off which favours quality over quantity of evidence. Considering the complete consensus amongst observations, and the fact that the number of observations included is only slightly below the average ($\overline{na}_{farm} = 6$), the magnitude of the SI is primarily driven by the high quality of evidence. The highest negative impact was observed for the impact of the use of chemical pesticide inputs on soil formation and composition (SI=-0.73, w=0.73, c=1, no=1, na=1). Considering the low number of observations, the magnitude of the SI in this case reflects the quality and quantity of evidence in the literature identifying the negative directionality.

Using the SIs we are also able to draw conclusions on the state of the literature. Here, we find that the literature could benefit greatly from an increase in high quality research at territorial level as well as for cultural ES.

The highest positive overall environmental performance, quantified by the CIs, was observed for mulching across all three case study areas (CI_{HH}=0.23, CI_{HW}=0.24, and CI_{NK}=0.28). The highest negative overall environmental performance was observed for low agrochemical pesticide across all three case study areas (CI_{HH}= -0.13, CI_{HW}= -0.11 and CI_{NK}=-0.20). CI magnitudes (both high and low) are driven firstly by underlying SI magnitudes - which in turn are driven by a combination of evidence quality, evidence quantity and consensus amongst observations – and secondly by the weights reflecting demand for particular ES in the case study areas. The weights primarily cause the variation observed in CIs between case study areas. While the highest and lowest performing management practices are commensurable between case study areas, the magnitudes of the CIs vary considerably. This is caused by the differing weights attributed to the 17 ES included in the assessment.

4 Quantitative monetary assessment

The monetary quantification of PG and ES can be achieved through two main approaches; market and non-market based assessments (de Groot et al., 2002; Kroeger and Casey, 2007; Gomez-Baggethun et al., 2010; Busch et al., 2012; Costanza et al., 2017). Market-based assessments consider the ability of direct outputs from goods and services to be sold on the market, and are represented by the exchange value of a particular good or service in trade (de Groot et al., 2002). Non-marketed assessments are performed when markets for services are absent. Therefore, a monetary value must be attributed through proxy measures (Potschin and Haines-Young, 2011). The two most commonly applied methodologies in such assessments include revealed preference and stated preference analyses (Costanza et al., 2017). The former infers value from individuals' observed choices in a real-world setting, while the latter uses individuals' responses to hypothetical situations to calculate monetary value through willingness to pay/accept (Costanza et al., 2017; de Groot et al., 2002; Gómez-Baggethun et al., 2010).

Provisioning services are most readily monetised as these describe the material or energy outputs derived from ecosystems (Power, 2010). Markets for (derivatives of) provisioning services also frequently exist, facilitating estimation of monetary values. Regulating and maintaining services, though not as readily monetised as provisioning services (because no markets exist), can be attributed a monetary value based on non-marketed, proxy measures. The monetary assessment of cultural services, on the other hand, is more complicated since cultural worth is highly individual and place-dependent (Rawluk et al., 2019). Nonetheless, through contingent valuation, a non-market based assessment approach, cultural services may also be attributed monetary quantification (e.g. Scarpa et al., 2015).

The above-mentioned assessments to attribute monetary value to ES are all related to collecting primary data. However, when primary data is infeasible – too costly or time-consuming – a benefit transfer (BT) approach provides an appropriate alternative to quantifying monetary values (Johnston et al., 2015). BT extrapolates economic estimates from a study site, derived through one of the above mentioned approaches, to a similar policy site for which estimates are unknown (ibid). The accuracy and relevance of a BT hinges on the similarity between the two sites. The greater the similarity between the sites, in terms of site characteristics, valuation context as well as socio-demographic characteristics, the greater the quality of the BT output. If study and policy sites are ill-matched in context, potential generalisation errors increase greatly (Johnston et al., 2015). Measurement error may also undermine BT output by carrying over underlying errors from the study site to estimates made at the policy site (Bateman et al., 2009; Johnston et al., 2015). Particularly when dealing with ES, this error may be significant (Johnston et al., 2015). Therefore, careful attention must be paid to not only the approximation of the policy site to the study site, but also the quality of the data reported for the study site.

Here, a BT was used to derive monetary values for a subset of the ES quantified in section 3. For pragmatic purposes, we adopt a unit value approach in which a single number or a set of numbers from pre-existing studies are transferred to the policy site.

4.1 Methodology

4.1.1 Benefit Transfer

Using the database derived from the previously carried out REA (see section 3), we selected a relevant study to use as policy site for the BT. Specifically, we used data derived from Van Vooren et al. (2018) as this study provided detailed quantitative data relating management practices to particular ES, as well as a transparent methodology. Using ES, Van Vooren et al. (2018) evaluate the quantitative performance of grass strips (GS) and hedgerows (HR) in an arable agricultural setting in two case study areas (Polders and Heuvelland) in Flanders, Belgium. Table 1 provides and overview of the ES and their respective indicators that were monitored by Van Vooren et al. (2018).

Table 1. List of ecosystem services and respective indicators monitored by Van Vooren et al. (2018).

Ecosystem service	Indicator
Food production	Yield (ton ha^{-1})
Climate regulation	Organic carbon stock (ton ha^{-1})
Maintenance of chemical water quality	Mineral N content (kg ha^{-1})
Pest control	Species diversity/density

To attribute monetary values to these ES, we use the Nature Value Explorer tool (https://www.natuurwaardeverkenner.be) developed by Broekx et al. (2013). This tool is freely available online, and has been previously used in the literature for a similar purpose (Lerouge et al., 2016). The Nature Value Explorer uses GIS geodatabases to attribute monetary values to land use changes by comparing different scenarios of intervention against a baseline. Through the use of GIS technology, the tool may be used to identify areas within a given parcel where a particular farm management practice is implemented. The impact of this practice on ES delivery within the parcel is then compared against a baseline, which is defined by the current land use practices. This provides the user with a great deal of freedom to manipulate scenarios to best match their expectations.

Using the CICES version 4.3, the Nature Value Explorer tool expresses the value of ES within one of the seven considered ecosystems in terms of a qualitative score of 1 to 10, quantitative physical units dependent on the service, as well as monetary units (€). The quantification of biomass production within the tool is based on the area in ha harvested in the study area. Monetary value for grassland biomass, which is primarily used as fodder within the agricultural firm, is estimated indirectly from the standard gross margins (market prices) of the livestock grown with the fodder.

Soil carbon stocks within the tool are based on estimates from Meersmans et al. (2008), who performed a regression approach to estimate the spatial distribution of soil organic carbon in Flanders, Belgium. Incorporating information on soil characteristics and vegetation type, potential maximum carbon content is estimated. Expressing these estimated in monetary values is achieved through avoided abatement costs of mitigation measured for climate change. The tool uses 30€/tonne CO_2 (100€/tonne C) as an average lower value, and 100€/tonne CO_2 (366€/tonne C) as an average higher value for 2020.

Avoided N and P leaching to water (water quality) is used as a proxy measure for soil mineral N content, as this is associated with the retention of nutrients in soils. N and P contents of soils are derived from the carbon content. Similarly to the approach adopted for quantifying soil carbon stocks, soil mineral N is monetarily valued through the avoided marginal abatement cost associated with the removal of an additional kg of N or P. Costs were calculated for the Flemish river basin management under the European Water Framework Directive. The highest marginal cost calculated to reach water quality objectives are 74€/kg N and 800€/kg P. Considering most measures have impact on both N and P it was deemed impossible by the creators to individually link avoided costs to separate pollutants. In an effort to avoid double-counting, it was therefore opted to estimate the value of nutrient retention across both pollutants simultaneously. The above mentioned marginal cost was used as a high estimate, while an average low estimate was derived from the literature (5€/kg N and 80€/kg P).

4.1.2 Sample selection

Using GIS (QGIS V3.16.6), we attempted to match the modelled scenarios as accurately as possible with the sample parcels used in Van Vooren et al. (2018). To this end, scenarios illustrating the monetary value of GS were modelled for the case study area of the Polders region, while HR scenarios were modelled for Hageland-Haspengouw. Though Van Vooren et al. (2018) quantify HR impact on ES in the case study area of Heuvelland, data restrictions did not allow for this area to be modelled using the Nature Value Explorer. As such, Hageland-Haspengouw was selected as an alternative study area due to its similarity in production type and landscape characteristics.

Based on the specifications of sample parcels in Van Vooren et al. (2018), a set of criteria were identified for those plots to be used for GS and HR scenario modelling. In accordance with the below-mentioned criteria, 16 parcels (eight for GS and eight for HR) were selected for this BT modelling exercise. Parcels selected for GS modelling had to meet all of the following criteria:

1. Parcels must be located in the Polders region (defined by the municipalities Sint-Laureins, Assenede and Kaprijke);
2. Parcels must be bordering a waterbody or water stream;
3. Parcels must be sown with either potato, sugar beets, or winter wheat; and
4. Parcels must have a minimum surface area of 1 ha.

Parcels selected for HR modelling had to meet all of the following criteria:

1. Parcels must be located in Hageland-Haspengouw;
2. Parcels must be sown with either grass, legumes, of winter wheat; and
3. Parcels must have a minimum surface area of 1 ha.

Using the Nature Value Explorer, scenarios were then modelled on each of the 16 selected parcels. Eight of the 16 parcels were modelled with GS adjacent to a ditch and/or waterbody, and with a minimum width of 12 meters. The remaining eight parcels were modelled with HR between 1-2 meters in width along one border, made up of a mixture of deciduous tree species.

Monetary estimates were averaged across the eight parcels per intervention to obtain one single, average monetary value for each ES. Not all ES described by Van Vooren et al. (2018) could be modelled. However, the Nature Value Explorer allowed for some additional ES to be modelled instead. A list of the ES for which monetary value was derived using the Nature Value Explorer can be found in Table 2.

4.2 Results

Table 2 describes the resulting monetary value of a set of six ES under GS and HR management respectively. Results are benchmarked against the current, extensive land use, therefore representing the monetary value of conversion from current arable agriculture practices to either GS or HR expressed in € year^{-1}. We see that, based on the performance across all six considered ES, the conversion of arable land to HR along one parcel margin in Hageland-Haspengouw accrues a monetary value of 41.50 € year^{-1}. Similarly, GS implemented along parcel margins in the Polders region accrues a cost of -186.38 € year^{-1}. As the interventions (GS and HR) are modelled in two different case study regions, a comparison between the two is not possible. Instead, results for each should be considered in isolation and are meant to serve as an example of the monetary value that can be estimated for each ES derived from GS and HR.

Table 2. Low and high monetary value estimate of a set of six ES provided by grass strips and hedgerows (€ year^{-1}) management practices adopted in arable agricultural parcels across Flanders, Belgium.

Ecosystem service	Grass strips (GS)		Hedgerows (HR)	
	Low	High	Low	High
Food production	-373.63	-706.63	-58	-107
Wood production	-2.88	-2.88	1.25	1.25
Air quality	-25.75	-25.75	13.75	13.75
Water: Increased infiltration	0.38	0.38	-1.13	-1.13
C storage (soils)	36.38	132.50	12.75	45.75
C storage (biomass)	-2.63	-9.38	0.75	2.50
Recreation & tourism	167.88	167.88	63.75	63.75
Total	-186.38	-430	41.50	27.25

When looking at the estimated monetary value for **food production**, we see that conversion of arable agriculture to both GS and HR accrue an economic cost rather than providing an economic benefit (indicated by the negative estimates). The cost accrued by GS is of a rather high magnitude, especially when compared to the cost of HR in terms of food production. This may be explained by the larger portion of arable land that is converted to non-productive land when implementing GS compared to HR. As specified by Van Vooren et al. (2018), GS were required to have a minimum width of 12 meters while HR were only minimally 1m in width. To this end, the average area covered by GS per modelled parcel in this assessment was 0.37 ha, which was 4.84% of the average considered parcel area. HR covered on average 0.04 ha per parcel, which, due to the overall smaller parcel size in Hageland-Haspengouw, equated to 1.1% of the average parcel area.

Non-monetary quantification of ES benefits as reported by Van Vooren et al. (2018) also indicate a loss in crop yield under both intervention types. Here, the loss in crop yield as a result of HR implementation is not only attributed to loss of productive area, but also to a shading-effect, which postulates that crop yields directly adjacent to the HR may be reduced due to a reduction in solar radiation reaching the crops (Montgomery et al., 2020; Van Vooren et al., 2017). This shading-effect seems to be further illustrated in Van Vooren et al. (2017), where crop production was found to increase with an increase in distance/height ratio (the ratio of the distance from the HR to the height of the HR). Van Vooren et al. (2017) further illustrate that the overall impact of HR on crop yield may be further nuanced, with a loss of productive area in combination with the shading-effect resulting in a reduction of crop yield, while a simultaneous shelter effect results in an increase in crop yield. Without considering the impact of the loss of productive area, the authors conclude that the overall impact of HR on crop yield is positive. Unfortunately, this nuance is not possible within the current analysis due to input restrictions in the Nature Value Explorer tool used.

Though **wood production** is not a service considered at the study site, results from the Nature Value Explorer enable its inclusion in this analysis. As is to be expected, the monetary value of wood production from GS implementation in the Polders regions is negative, while this is positive, albeit only marginally so, for HR implemented in Hageland-Haspengouw. The negative monetary value for GS is likely caused by the specification of the reference scenario, i.e. the current land use. It is possible that certain areas that have been converted to GS in the modelling process are registered in the input geodatabases as contributing to wood production in the baseline, which would therefore result in a loss in this productive output with the implementation of GS.

While the monetary value of wood production from HR is positive, this estimate is expected to be highly dependent on the management scheme adopted by the farmer. The estimation model used by Broekx et al. (2013) estimates wood production for only a number of tree species using the average market prices paid for wood harvested in public forests. However, maintenance of HR does not occur along the same principles as agroforestry systems. Indeed, farmers may receive subsidies for the maintenance of HR following various strict pruning and harvesting guidelines (VLM Vlaamse Land Maatschappij, 2017). This nuance is not captured in the Nature Value Explorer, and the BT estimates for wood production should therefore be considered carefully.

A monetary value of 13.75 € year^{-1} was estimated for the impact of HR on **air quality**. On the other hand, the impact of GS on air quality was estimated to accrue a cost of -25.75 € year^{-1}. The positive monetary value estimated for the implementation of HR compared to the negative estimated value for the implementation of GS is not unexpected. While there is not much evidence pertaining particularly to agricultural landscapes, trees, shrubs and other (semi-)permanent woody species have been found to effectively reduce air pollination and increase air quality in urban environments and along roads (e.g. Gromke et al., 2016). The cost associated with the implementation of GS, on the other hand, is likely once again related to both the local characteristics of the surrounding environment (e.g. distance to nearby roads) combined with the types of vegetation that have made way for GS in the baseline scenario.

The monetary value of **increased infiltration rate** related to GS implementation along parcel edges in the Polders region is estimated at 0.38 € year^{-1}. For HR implemented in Hageland-Haspengouw, the monetary value is estimated as a cost of -1.13 € year^{-1}. Though Van Vooren et al. (2018) do not consider infiltration in their assessment, results from the quantification exercise in section 3 as well as evidence from the literature (Holden et al., 2019; Montgomery et al., 2020; Stiles, 2016) indicate that the quantitative, non-monetary value of infiltration resulting from HR implementation is positive. Though we would assume a positive relationship between the non-monetary and monetary value of ES derived from farm management practices, this discrepancy in our results may be caused by the impact of geospatial characteristics on the value of specific ES. Further, as the GS cover a greater area than the HR, these would likely increase infiltration capacity of the soils as a greater portion of the parcel is under non-productive land use compared to with HR.

Unfortunately, the impact of GS and HR on water quality could not be estimated using the Nature Value Explorer tool.

The monetary value of C storage was estimated for both soil and biomass storage. We see that for both GS and HR, the monetary value of **C storage in soils** is higher than in biomass, with an estimated value for the former of 36.38 € year^{-1} for GS and 12.75 € year^{-1} for HR. The discrepancy in monetary value between HR and GS is in accordance with the results of the non-monetary quantification reported at the study site as well as in the literature more broadly (Holden et al., 2019; Montgomery et al., 2020; Van Vooren et al., 2018, 2017). At the study site, the yearly increase in soil organic carbon (SOC) stock as a result of the conversion of a portion of arable land to GS was estimated at 1.60 ton ha^{-1} year^{-1}. According to the estimates made in this analysis, this increase equates to a monetary value of 36.38 € year^{-1}.

Estimates of the monetary value of C storage in biomass for GS are negative, indicating that the conversion of arable land into non-productive GS along parcel edges in the Polders region accrues a cost of -2.63 € year^{-1} in terms of C storage in biomass. At the same time, the monetary value for the same ES from HR implemented in Hageland-Haspengouw is positive, with a value of 0.75 € year^{-1}. Similarly to wood production, HR maintenance has an impact on the amount of C stored in the biomass, with C storage in biomass decreasing with increased trimming and pruning (Axe et al., 2017). Though incorporating specific maintenance practices into the BT carried out here is beyond the scope of this evaluation, readers should be aware of the interaction between the maintenance regimes and ES delivery/value when considering the results reported here.

Monetary value of ES derived from management practices such as GS and HR are likely to change over time, and therefore any conclusions to inform policy-making decisions should be based on a multi-temporal estimation rather than the snapshot estimation that is reported here.

Though **cultural services** are not evaluated by Van Vooren et al. (2018), the Nature Value Explorer enabled us to estimate the monetary value of recreation and tourism benefits derived from GS and HR. The estimated monetary values of these interventions are 167.88 € year^{-1} and 63.75 € year^{-1} respectively. While we are not able to compare these estimated outputs with result from the study site, we are able to refer to the estimation models used by Broekx et al. (2013) in the Nature Value Explorer to explain the discrepancy between the two interventions. Here, we see that monetary value is estimated primarily based on spatial specifications, considering the characteristics of the surrounding area as a major contributor to monetary value.

This highlights one of the key distinctions that needs to be made when comparing monetary estimates between the two interventions. Interventions were modelled in different case study regions, each with differing geographical characteristics. As the Nature Value Explorer is a spatially-explicit tool, using geospatial input data from various different geodatabases, the results must be interpreted as such. Therefore, making comparisons between the two management practices requires modelling the interventions within the same case study area. This would be a next step to adopt in this analysis.

5 Empirical analyses of farm environmental performance

Farm environmental performance, whether in quantitative terms (productivity) or in financial terms (profitability), reveals how a farm performs using various environmental indicators. The latter would theoretically be pressure (impact) indicators, but generally the data available do not allow the measurement of such indicators. Therefore, in general, in empirical assessments with farm-level data, proxies are used and are often indicators of the extent of input use (costs in monetary values, or quantities), alone or in combination with economic output indicator (in the efficiency analyses). Several studies have been conducted in the LIFT project on this issue. They are summarised below. Some papers are provided in more details in LIFT Deliverable 3.1 (Niedermayr et al., 2021a) since they also assess jointly farm economic performance.

5.1 Methodology

Farm environmental performance has been investigated empirically in LIFT with four approaches: (i) farm bio-economic model enabling the computation of environmental indicators for a specific farm; (ii) computation of environmental indicators and comparing them across types of ecological farms; (iii) farm efficiency, including in the model both economic and environmental netputs on the farm; (iv) environmental damage computed at the farm or the plot level. The different studies carried out have been included in 13 academic papers (some already published) which are described below.

5.1.1 Farm bio-economic models

- Heinrichs et al. (2021) on legumes in France and Germany

Heinrichs et al. (2021) use the bio-economic programming farm-scale model FarmDyn (Britz et al., 2016), to quantify agronomic, economic and environmental impacts of increasing legume production (peas, faba beans and alfalfa) in France and in Germany. It therefore considers that when the farm produces legume it is more ecological, as increasing the legume production implies moving the farm towards low input production. More precisely, the study investigates the interaction of policy measures that, on the one hand, aim at promoting legume production (with voluntary coupled support in the frame of the CAP greening) and, on the other hand, potentially constrain their production by regulating N supply (with Nitrate Directive).

Two farms are modelled: one French and one German intensively managed dairy farm located respectively in Pays-de-la-Loire in France and North Rhine-Westphalia in Germany, both regions being characterised by intensive livestock production under temperate climate. Data on yields and on input and output prices for legumes and other crops are extracted from public statistics and professional agricultural press. In the baseline scenario both farms are conventional since there is no voluntary coupled production for legumes and the national implementation of the Nitrate Directive. In further scenarios, a voluntary coupled support is introduced in both farms, then the German Nitrate Directive is introduced.

The environmental indicators computed in the study are input indicators, namely protein self-sufficiency and input quantity of mineral fertiliser and manure; global warming potential (GWP) of the farm; and indicator of N leaching.

- Heinrichs and Britz (2021) on AES in Germany

Heinrichs and Britz (2021) study the economic and environmental impacts of selected agri-environmental and climatic measures (AECMs) in two conventional farms, one arable crop and the other dairy, in North Rhine-Westphalia in Germany. The baseline scenario excludes the opportunity to participate at AECMs, and both farms comply with the greening regulation and the German Nitrate Directive. In the AECM scenario, AECMs exist and can be uptaken, with the corresponding subsidies which relate to catch crops, flower strips and the implementation of diversified crop rotation.

The environmental indicators computed in the study are ratios of input use, N and P balances, pressure indicators (e.g. N-Leaching), life cycle analysis (LCA) indicators, and biodiversity indicators.

- Kilcline et al. (2021) on sheep farms in Ireland

Kilcline et al. (2021) develop a nationally representative bio-economic model to compare farm level economic and environmental performance of Irish sheep farms using Teagasc National Farm Survey (NFS) data (Ireland's contribution to EU Farm Accountancy Data Network-FADN). Farm level Carbon Footprints are calculated in terms of the Carbon Dioxide (CO_2) equivalent per kg of live weight equivalent of sheep produced using a "cradle-to-farm-gate" LCA approach.

The environmental performance of distinct sheep farming systems operating at different levels of production intensity and input use is presented and compared. The data used consist of 3,235 sheep farm-year observations over the 6 year period 2010 to 2015. The average farm size is 50 ha for hill farms and 42 ha for lowland farms.

5.1.2 Environmental indicators

- Lascano Galarza et al. (2021) and Niedermayr et al. (2021b) on FADN data in Italy and Austria

Lascano Galarza et al. (2021) compute two environmental indicators for 2,117 Italian farms during 10 years (2004-2013) using EU FADN data: the cost of crop protection products per ha of UAA, and the cost of chemical fertilisers per ha of UAA. The sample of farms include all production types of farming. The authors then regress the indicators on explanatory variables, including whether the farm produces under organic farming and whether the farm is participating in an AES.

Niedermayr et al. (2021b) use FADN data for specialist dairy farms in Austria for the years 2014 and 2015. The farms in the sample have an UAA of 31 ha on average and produce 106,400€ of output. The environmental indicators computed by the authors are intensities of inputs: veterinary expenses (in € per cow), fertiliser costs (in € per ha), crop protection costs (in € per ha) and concentrate feed costs (in € per ha). The LIFT typology protocol developed in LIFT Deliverable 1.4 Rega et al. (2021) is used to assign farms to various ecological types: organic farms, integrated farms, farms that are organic and integrated, non-organic farms. In order to compare environmental indicators between farms that are similar in terms of structural characteristics, direct covariate matching (DCM) is used, where matching is performed upon several covariates at the same time. In this case, the covariates are: farm size measured in terms of standard output; site conditions proxied by LFA payments per livestock unit; the share of permanent grassland; and a dummy for the year 2014.

- Jendrzejewski and Zawalińska (2021) and Ayouba et al. (2021) on data from the LIFT large-scale farmer survey in Poland and France

Jendrzejewski and Zawalińska (2021) for two Polish regions and Ayouba et al. (2021) for three French regions using the 2018 data from the LIFT large-scale farmer survey. The questionnaire has been specifically designed (see Deliverable 2.2, Tzouramani et al. (2019) so as to collect original farm data on practices and various other information, for the year 2018. Jendrzejewski and Zawalińska (2021) and Ayouba et al. (2021) compute performance indicators, and use propensity score matching (PSM) to compare them between two types of farms: ecological farms and non-ecological farms. Both types of farms are defined in different ways as explained below, and the comparison of performance is done for each pair of types:

- Certified organic: the ecological farms are certified organic farms, the non-ecological farms are not (T_org is a dummy variable taking the value 1 if Q9_sq02_A=1, 0 otherwise).
- No chemical pest management: the ecological type is farms not using pest management chemical products, and the non-ecological type is farms using them (T_pest is a dummy variable taking the value 1 if Q20A_1=0, 0 if Q20A_1=1).
- Conservation tillage or no tillage: the ecological type is farms using conservation tillage or not tillage, and the non-ecological are farms not using this type of tillage (T_till is a dummy variable taking the value 1 if Q26C_2=1 OR Q26C_3=1; 0 otherwise).
- Antibiotics for treatment only: the ecological type is farms using antibiotics for treatment only, and the non-ecological type is farms using antibiotics for both treatment and prevention (T_antib is a dummy variable taking the value 1 if Q35A_2=1; 0 otherwise).

Environmental performance indicators (for the year 2018) include three input use indicators and one output indicator. The input use indicators are (the first two ones were used for Polish farms, while the rest was used for French farms): use of chemical fertilisers (in PLN) per ha of UAA; use of fuel (PLN) per ha of UAA; percentage of UAA on which chemical products are applied; duration that concentrates are given (number of months); while the output indicator is the percentage of UAA covered with hedgerows.

The comparison of these environmental indicators between both farm types (ecological and non-ecological farms) is made with PSM, where similar farms in terms of structural and social characteristics (the so-called covariates) are compared. The covariates used in both countries are the age of the farmer, the experience of the farmer, the UAA, the share of owned land in UAA, and soil quality. Additional covariates are used for the French sample: gender of the farmer, education of the farmer, regional dummy and dummy for farm localisation in less favoured area (LFA).

The Polish sample includes 100 farms distributed over several farming types, in particular mixed crop and livestock farms, field crop farms, dairy farms and beef cattle farms. On average their UAA is 15 ha. They are located in Lubelskie and in Podlaski regions. Lubelskie is located in the south-eastern part of Poland, and is known as the agricultural area called a food granary of Poland with 70% of the region area in UAA. By contrast, Podlaskie in north-eastern Poland, counts vast forests and numerous lakes, and presents harsh climatic conditions for agriculture. The French sample used includes 162 farms specialised in dairy farming or beef cattle farming, mainly located in mountainous region Puy-de-Dôme in central France, where dairy and cattle beef farms produce high-value cheese. Other farms in the French sample are located in western France, in Ille-et-Vilaine where dairy farms are relatively intensive and do not produce milk dairy products, and in Sarthe, where dairy and beef cattle farms have a less specialised orientation and produce also field crops. Farms in the French sample operate on average 115 ha of UAA and breed 111 livestock units for dairy and beef cattle.

5.1.3 Farm technical-economic and environmental efficiency

Four studies use frontiers methods to compute farm environmental performance together with technical-economic efficiency.

- Huang et al. (2021) on arable crop farms in Sweden

The authors used an unbalanced panel of data of 209 arable crop farms (specialist cereals, oilseeds and protein crops, general field cropping, and mixed cropping) from the Swedish FADN for the period 2009-2016 in two regions of Sweden (North and South). Farms in the sample operate 110 ha of crop land and produce €2,260 of agricultural output on average.

The authors apply the concept of eco-efficiency (OECD, 1998; WBCSD, 1992) with a directional distance function Färe et al. (2005). In addition to the standard agricultural output and inputs, they include an environmentally-favourable input, namely a crop diversity index (CDI) lagged, and an undesirable output namely the Herfindahl index (HI) which is an index of crop diversity loss. The novelty in this study is that the dynamic effects of crop diversity are accounted for in the production function, with the underlying assumption that crop diversity plays a dynamic role as input (crop diversity in previous years contributes to the production process in future years) and is a by-product in the production process. On average, the CDI of the sample is 0.79 and the HI is 0.21.

- Sintori et al. (2021) on olive farms in Greece

Similar to the above study, Sintori et al. (2021) use the concept of eco-efficiency. Their application is to the sample of Crete (Greece) olive farms from the LIFT large-scale farmer survey whose data are from 2018. The farms are located in the regions of Heraklion and Lasithi, located in the eastern part of the island. These two regions account for 15% of the total olive trees and 17% of the total olive production of the country. The sample used include 65 olive farms, operating on average 4.9 ha and generating on average €13,340 of agricultural output.

The authors use Data Envelopment Analysis (DEA) (Kuosmanen and Kortelainen (2005)) to compute eco-efficiency score with the standard agricultural output as the DEA output, and only environmental inputs are used in the inputs. The four environmental inputs include two environmentally-damaging inputs namely water consumption (total amount of irrigation water used per hectare) and fuel consumption (total amount of fuel consumed to perform everyday tasks per hectare). The two other inputs are environmentally-favourable inputs, which are derived from the LIFT typology protocol in Deliverable D1.4 Rega et al. (2021): a composite indicator of soil management that includes tillage practices and soil cover practices, and a composite indicator that includes fertilisation practices and pest management practices; the higher the indicators, the more environmentally-friendly the practices are. The inverse of the two indicators are used as inputs.

After computing the efficiency scores, the authors use a truncated regression to investigate the drivers of efficiency.

- Toma et al. (2021) on sheep and cattle farms in Scotland

Toma et al. (2021) apply DEA on FADN data for 89 cattle or sheep farms in Scotland observed each year between 2011 and 2015. The inputs used in the DEA model are total assets, intermediate consumption, paid labour and unpaid labour, while the output is an environmental output, that includes output from renewable energy, tourism, forestry.

The LIFT typology protocol in Rega et al. (2021) is used to compare the efficiency scores of high input vs low input farms on the one hand, and of high integration vs low integration farms on the other hand. In a second stage the authors estimated the effect of various explanatory variables on the efficiency scores with a regression.

- Niedermayr et al. (2021c) on dairy farms in Austria

Niedermayr et al. (2021c) use data from the LIFT large-scale farmer survey for the year 2018 in regions Steyr-Kirchdorf and Salzburg und Umgebung, both situated at the northern edge of the Alps and northern alpine foothills and characterised by a large share of farms specialised in grazing livestock.

The authors define ecological and non-ecological farms in two ways: organic farms vs. non-organic farms; and farms certified within the EU quality scheme as traditional speciality guaranteed (TSG) under the name haymilk farms vs. non haymilk farms. Many farms producing silage-free milk are certified as haymilk TSG farms in Austria. In this system, grass can be cut less often, when hay is produced, compared to silage, which can be beneficial for biodiversity on grassland. Haymilk farms can be both organic or non-organic. In the LIFT large-scale farmer survey Austrian sample, non-organic farms (n = 35) make up the biggest group, followed by organic haymilk farms (n = 20), organic farms (n = 16) and non-organic haymilk farms (n = 10). Farms in the sample are 36 ha large and produce €124,200 of output. An animal welfare index is calculated based on four animal welfare indicators: veterinary expenses as a proxy for animal health; stable size; seasonal pasture and general outdoor access. The animal welfare index is 0.48 on average for the whole sample, and increases along the degree of ecological farming, namely from non-organic farms (0.37), non-organic haymilk farms (0.40), organic farms (0.59), and organic haymilk farms (0.60).

Efficiency accounting for both agricultural and environmental outputs is computed with DEA. Two DEA models are used, both with the same standard inputs, but with different outputs. One model includes two outputs, namely animal welfare and total output excluding subsidies, and one model has one single output, namely total output including agri-environmental and organic subsidies. In a second-stage, the authors estimate the impact of various drivers on farms' efficiency.

5.1.4 Environmental damage

- Dakpo and Femenia (2021) for wheat farms in France

Dakpo and Femenia (2021) examine pesticide efficiency using stochastic frontier analysis (SFA). Pesticides are not considered as standard input, but under the damage control specification (exponential and logistic). The authors develop an input-specific efficiency measure and overcome the potential endogeneity issue by relying on a dual approach. Following the framework developed in Chambers and Lichtenberg (1994), they derive pesticide demand and then adjust this function to account for inefficiency. Using a maximum likelihood approach, the obtained model is estimated on a large sample of about 2,000 French wheat farmers in the Meuse (north-east France) region over the period 1998 to 2014, whose data are obtained from local farm accountancy offices.

- Dakpo et al. (2021) for wheat plots in France

Dakpo et al. (2021) evaluate nitrogen excess marginal abatement cost (MAC) in French wheat production. The authors consider a novel framework for modelling pollution-generating technologies based on multi-equation (namely the by-production as in Murty et al. (2012)). The basic principle of this approach is to consider that the overall technology is the intersection of two sub-technologies—one for the production of good outputs and the other for the generation of bad outputs. The estimation of the MAC is based on the potential trade-offs involving nitrogen excess. Trade-offs are estimated using a quantile approach to account for the potential inefficiency of producers. An extension of the stochastic DEA has been developed to account for the generation of bad outputs. The application is to a sample of plots cultivated with wheat in France in 2017 and managed by the French Ministry of Agriculture (database "Pratiques Culturales"). According to the authors, the plot level is the level where the environment is at stake for nitrogen pollution. In a specific farm, there may not be nitrogen excess when considering the farm as a whole. Still, there may be high excess on some plots, implying potential nitrogen losses to the environment that would need to be addressed.

5.2 Results

5.2.1 Farm bio-economic models

- Heinrichs et al. (2021) on legumes in France and Germany

Since legumes provide nitrogen, the increase in legume production has an input-saving effect by decreasing the use of purchased feed, increasing the farms protein self-sufficiency, and the reduced use of N fertiliser. Under the French Nitrate Directive, N leaching decreases almost continuously to reach a maximal decrease of 16%, whereas, under the German Nitrate Directive, it decreases only by 5%. This gap is due to the spreading of manure on grain legumes, provoking their over-fertilisation and thus, additional N leaching. The GWP decreases by 5% under the French Nitrate Directive and by 2% under the German Nitrate Directive. Additional details can be found in section 5.4 of LIFT Deliverable D3.1.

- Heinrichs and Britz (2021) on AES in Germany

According to Heinrichs and Britz (2021), in the baseline scenario without the possibility to participate at AECMs, both crop and dairy farms comply with the greening regulations. After implementing AECMs as voluntary policy measure, both farms introduce 10% of flower strips to the farm. Thereby, flower strips substitute against rape seed production. The introduction of AECMs results in a considerable decrease in input requirements on both farms. The conversion to a more extensive crop mix due to the integration of flower stripes combined with the reduced application of mineral fertilisers results in an improvement of pressure indicators. The LCA analysis indicates a decrease of the GWP by 11%. Using SMART index and SALCA methods, the authors show that biodiversity increases. Additional details can be found in section 5.3 of LIFT Deliverable D3.1.

- Kilcline et al. (2021) on sheep farms in Ireland

Kilcline et al. (2021) estimate that the average carbon footprint per kg of live weight produced to be 13% lower on lowland farms in comparison to hill farms. Taking into account the carbon sequestration value of grassland reduces the carbon footprints on farms. In line with O'Brien et al. (2015), the carbon sequestration rate had a relatively larger impact on reducing emissions for more extensive farms including hill farms. The breakdown of emissions shows that animal activities represent the largest source, with Tier I estimates of enteric fermentation and manure management comprising 64% and 6% of total emissions respectively. Other emissions include those emissions from soils (14%) and total emissions associated with feed production (16%). Additional details can be found in section 5.1 of LIFT Deliverable D3.1.

5.2.2 Environmental indicators

- Lascano Galarza et al. (2021) and Niedermayr et al. (2021b) on FADN data in Italy and Austria

Lascano Galarza et al. (2021), with their regression for Italian FADN farms, show that, conform to intuition, the cost of crop protection products per ha is significantly lower for farms that fully produce under certified organic farming than other farms, the difference being about €120 per ha. The cost for farms producing both under certified organic farming and non-organic approaches is not significantly different than other farms. As regard the cost of chemical fertilisers per ha of UAA, the results of the regression on this indicator indicate that producing under organic farming, whether on the whole farm or partly, has no impact on this cost. Both regressions show that being specialist horticulture farm has the stronger positive impact on both types of costs, but that participating in an AES reduces both types of costs.

Niedermayr et al. (2021b), using matching on four environmental indicators compute average treatment effects on the treated (ATTs). They show that most ATTs are significant for the pairwise comparisons of the four ecological types. Each type is compared to another type, along an increasing degree of uptake of ecological practices: non-organic, integrated, organic, organic and integrated together. Six comparisons are performed. In the pairwise comparisons, the less ecological farming system is always defined as the control group and the more ecological farming system the treated group. A negative ATT thus indicates that the environmental indicator decreases when switching to a more ecological farming system. The ATTs for all four environmental indicators are most frequently negative (at worst, not significant) when moving from less ecological (e.g. organic) to more ecological (e.g. organic and integrated together). Details can be found in LIFT Deliverable 3.1 where the study of Niedermayr et al. (2021b) is described in section 4.3.

- Jendrzejewski and Zawalińska (2021) and Ayouba et al. (2021) on data from the LIFT large-scale farmer survey in Poland and France

Jendrzejewski and Zawalińska (2021) in Poland and Ayouba et al. (2021) in France apply PSM on the LIFT large-scale farmer survey data. PSM allows calculating an ATT, which measures the change in performance when an ecological farm switched from non-ecological to ecological (i.e. the treatment). The ATTs are calculated for the Polish sample and for the French sample for each typology (ecological vs non-ecological). Most ATTs are not significant. As regard the significant values, in the Polish sample, conform to intuition farms with non-chemical pest management practices use significantly a lower cost of chemical fertilisers per ha of UAA than their non-ecological counterparts. As for certified organic farms, they use more fuel per ha of UAA than non-organic farms, which is consistent with the fact that under organic practices the resort to mechanic pest management is higher.

In France, the percentage of UAA on which chemical products are applied is lower for ecological farms than non-ecological farms when the former are identified as applying no tillage or conservation tillage or using antibiotics for treatment only. This suggests that farms use a combination of several ecological practices, here relating to pest management, soil management and livestock management. This is also revealed by the significant negative ATT for the duration that concentrates are given in the case where ecological farms are identified as farms not using pest management chemical products, or in combination with using antibiotics for treatment only.

As for the environmental output on the farms, namely the hedgerows measured as the percentage of UAA covered with hedgerows, it is higher (the ATT is significantly positive) for farms using antibiotics for treatment only, or in combination with not using pest management chemical products), here again suggesting a bundle of different ecological practices used on farms.

5.2.3 Farm technical-economic and environmental efficiency

- Huang et al. (2021) on arable crop farms in Sweden

The average estimated eco-efficiency score for the Swedish arable crop farms is 0.876, revealing that the farms could improve their efficiency by 12.4% on average in terms of expanding agricultural products revenue and reducing HI given unchanged inputs. The average eco-efficiency in Southern Sweden and in Northern Sweden is similar. In terms of ecological types of farms, the average scores are close: certified organic farms have an average eco-efficiency score of 0.876, while the figure for mixed organic-non-organic farms or farms in conversion to organic farming is 0.886 and for non-organic farms it is 0.876.

Results from the estimated frontier reveal a reverse U shape relationship between eco-efficiency and lagged CDI. In addition, the elasticity of eco-inefficiency with respect to this CDI indicates that a 1%

decrease of the CDI would decrease the eco-efficiency by 6.3% on average. This is not in line with previous research (e.g. Di Falco and Chavas, 2006), where crop diversity is positively related with production in lagged effects. A time dummy was included in the efficiency frontier and results indicate that eco-efficiency was higher after 2013 than before. This may be due to the 2013 CAP reform incorporating crop diversification as a mandatory greening component for direct payments.

- Sintori et al. (2021) on olive farms in Greece

The average eco-efficiency for the whole sample of 65 farms is relatively low, 0.39, indicating room for improvement, in line with a similar study on olive farms in Spain (Andalusia) who also identified wide-range eco-inefficiency (Gómez-Limón et al., 2012). Looking at the average eco-efficiency per ecological type of farms, identified with the LIFT typology protocol of Deliverable D1.4 (Rega et al., 2021), reveals that conservation farms (0.36) and organic farms (0.42) have the lowest average eco-efficiency scores. Conservation farms perform well in terms of the composite indicator for soil management, and of water and fuel consumption, but poorly in terms of the composite indicator for fertilisation and pest management.

Significant results from the second-stage regression indicate that high education increases eco-efficiency, as well as the percentage of income from farming proxying the involvement of farmers in farming activity. Farmers were asked to rank their objectives with a Likert scale and two objectives have a significant impact on eco-efficiency scores: highly ranking the objective of producing high quality products increases eco-efficiency while highly ranking the objective of protecting the environment for future generations decreases it. This suggests that the actual reason for implementing environmentally-friendly farming practices may be farmers' focus on quality products and not on environmental concerns per se. Finally, market orientation, defined as the revenue from sales related to the sum of revenue from sales and subsidies, has a positive impact on eco-efficiency, suggesting that the lowest the subsidies received in relative value, the higher the eco-efficiency. This is in line with studies focusing on technical efficiency of olive farms in Greece and the Mediterranean region (Lambarraa et al., 2007; Zhu et al., 2011).

- Toma et al. (2021) on sheep and cattle farms in Scotland

The DEA results indicate a dispersed distribution of the efficiency scores, indicating significant potential for efficiency improvements. The efficiency scores of high input vs low input farms on the one hand, and of high integration vs low integration farms on the other hand, are not significantly different. Results from the second-stage regression indicate that subsidies, included as explanatory variables, have a significant positive impact on technical-efficiency.

- Niedermayr et al. (2021c) on dairy farms in Austria

After running the two DEA models differing in terms of output on the whole sample, the authors compare the average efficiency scores across the four ecological types (non-organic, non-organic haymilk, organic, organic haymilk). The results show that there is no significant difference across types: in the model where the two outputs are farm output and animal welfare index, the average scores range from 0.83 to 0.89 along the ecological types; in the model where the single output includes both farm output and subsidies, the average scores range from 0.73 to 0.81. By contrast when no environmental output is included and the model's outputs are the milk quantity produced and the other farm output in value, there is a significant difference across ecological types, with haymilk types being the least efficient and non-haymilk types being the most efficient in particular the non-organic type.

The estimation of the drivers of efficiency in a second-stage reveals that for both models, male farmers are less efficient than female farmers, and farmers with higher education are more efficient. In addition, age has no significant impact on efficiency, while the share of household income from farming has a positive impact, and milk price and rental price as well. And specifically for the model including farm output and the animal welfare index in the outputs, the amount of AES and organic subsidies per ha of UAA decreases efficiency, while the share of dairy cows in total cattle increase it. Detailed results can be found in LIFT Deliverable 3.1 in section 4.2.

5.2.4 Environmental damage

- Dakpo and Femenia (2021) for wheat farms in France

Dakpo and Femenia (2021) find that overall, pesticides could be reduced by 5% to 35% without impacting wheat yield if all farmers were fully efficient in using pesticides. These figures are lower compared to the ones reported in the study by Lechenet et al. (2017) - between 37% and 60%. Moreover, the results obtained with the exponential and logistic specifications are virtually very similar, which shows their robustness. In addition, while investigating the effect of crop diversification on pest control, the authors find a positive effect that is significant only for lower level of pesticides use. More details can be found in Deliverable D3.1 in section 4.1.

- Dakpo et al. (2021) for wheat plots in France

Results show that, for the sample of 153 wheat plots, the average shadow price for excess nitrogen is €8 per kg of N excess (when the wheat price is assumed to be €154 per ton). Moreover, equating the MAC with the wheat revenue suggests that the constraint on organic nitrogen should not be above 110 kg nitrogen per ha for farmers to keep a positive profit on the plots.

5.3 Discussion

We provided an overview of the 13 studies carried out in LIFT to investigate farm-level environmental performance, from a quantitative point of view (productivity) whether positive (e.g. biodiversity, animal welfare, hedgerows, technical-economic efficiency) or negative (e.g. N leaching) and from a financial point of view (profitability) namely costs of chemical inputs, shadow price of environmental pressure, profit. The variety of studies shows that there is no single methodology for measuring farm environmental performance. Methods used include computing input use indicators, measuring pressure indicators, estimating trade-offs between farm output and environmental (negative or positive) output, or calculating efficiency indicators that account for both economic and environmental dimensions. Often the environmental performance indicators are input use indicators. Only with specific data (e.g. bio-economic model, agronomic database such as the French database "Pratiques Culturales") can environmental pressures indicators be computed. The challenge is to observe environmental outputs jointly produced by farms with economic outputs, and one is forced to consider proxies based on inputs and practices that limit the range of methodologies used to assess environmental performance. Further research is needed large-scale collection or estimation of environmental outputs (e.g. carbon sequestration, GHG emissions, biodiversity intensity…).

From a methodological point of view, it could be noted that including subsidies as input or as output in efficiency models is problematic. Even though some subsidies like AES are considered proxy for environmental outputs, due to their nature, it is difficult to see how they can be rigorously (in the neoclassic framework) defined as output.

Finally, in the case of the LIFT project where one objective is to compare performance between different ecological types of farms and assess performance gaps, it is difficult to disentangle environmental performance indicators and proxies used to characterised ecological farm types, since the former are often input use indicators, hence relating to practices.

Despite these limits, the studies described in this section 5 show that findings are ambiguous as regard performance gaps: farms that are more ecological may show a higher or a lower environmental performance, depending on the ecological typology, the indicator of environmental performance, the main production of the farm (type of farming), among others. Similarly, the sources of farm environmental performance depend very much on the context. In particular, the effect of subsidies on efficiency has been found to be positive, negative or not significant depending on the case study, and corroborating the conclusions of Minviel and Latruffe (2017).

6 Conclusion

In this deliverable we present the results of the work carried out within Task 3.4 under WP3 of LIFT. Results include a qualitative, quantitative, and monetary assessment of the environmental performance of (ecological) farm management practices using the ES concept, as well as an empirical analysis of farm environmental performance depending on the practices. The qualitative assessment indicates that there is a great deal of information already available in the literature on the impact of farm management practices on ES supply in European agricultural systems. Various methodologies exist in the literature, ranging from purely qualitative descriptions (literature review), through biophysical quantitative assessments to purely monetary assessments. We notice that not all ES can be assessed using all possible methodologies. For example, cultural services are most often evaluated using monetary assessments such as stated preference techniques, while regulating and maintaining services are most often quantified in purely biophysical units.

Building on this work in the literature, we present a framework that allows for a wide range of ES to be assessed in the qualitative, non-monetary way through the composition of SIs. Though a REA, evidence derived from the literature was integrated with an assessment of article quality (through the use of standardised quality criteria), to obtain a single SI quantifying the impact of a single management practice on a single ES in the context of European agriculture. SIs thus allow us to draw conclusions on the potential supply of an ES from a given management practice, as well as on the overall quality of the research being done. Though carried out for a wide variety of ES, the SIs and CIs could not be calculated for all considered ES.

Following the non-monetary assessment, we carried out a quantitative, monetary assessment of a further smaller subset of ES. Here we focussed on two management practices, grass strips and hedgerows, and the monetary value of a set of six ES within the context of two Belgian agriculture case studies. Though the derived monetary estimates are specific to the case study areas for which they were derived, and should thus not be used outside of this context, they give an indication of the value of ES derived from ecological farm management practices in Europe.

Lastly, empirical analyses allowed us to draw conclusions of the farm environmental performance depending on management practices (rather ecological or conventional). Results from multiple studies are presented, each adopting one of four presented methodologies. Methods used include computing input use indicators, measuring pressure indicators, estimating trade-offs between farm output and environmental (negative or positive) output, or calculating efficiency indicators that account for both economic and environmental dimensions. The variety of studies shows that there is no single methodology most suitable for evaluating farm environmental performance.

Through the pyramid-approach adopted to present the environmental performance assessments carried out in this deliverable, we are able to demonstrate the impact of various management practices across a variety of ES using a wide range of methodologies. This assessment allows for a comparison to be made between various management practices within each level of the pyramid, and also allows for the performance of a single management practice to be assessed across various dimensions when moving down the pyramid. However, while each level of the pyramid serves as a continuation of the work carried out in the previous level(s), the empirical analyses described in section 5 and the results from the biophysical and monetary quantifications described in sections 3 and 4 are not methodologically linked. That is to say, ideally the empirical analyses carried out in section 5 are based on the results derived from the biophysical and monetary quantification, which is not the case in this work. The reason for this is that the information necessary for the empirical analyses are not available in secondary data, are too complex to collect on a large-scale basis, and because the work carried on the study region of Hageland-Haspengouw in Belgium provides a useful methodology but would require more research to be able to extend it to the scale necessary for incorporation in empirical analysis.

Nonetheless, results presented in this deliverable demonstrate that there are a wide variety of methodologies in existence, each with its own strengths and weaknesses. We demonstrate that there is no one methodology that is best for evaluating farm environmental performance of (ecological) farm management practices, but rather that the choice of methodology depends greatly on the available data, the intended scope and depth of the analysis, as well as on more logistical constraints.

Acknowledgements

We would like to thank everyone who contributed to the work presented in this deliverable. Particularly we thank Katarzyna Bańkowska, Joana Ferreira, Vasileios Gkisakis, Monserrath Lascano Galarza, Gordana Manevska-Tasevska, Peter Matthews, Kewan Mertens, Andreas Niedermayr, Lena Schaller, Peter Walder for their efforts in screening and extracting data from articles during the rapid evidence assessment. We thank the authors of the papers listed in section 5: Kassoum Ayouba, Andrew Barnes, Wolfgang Britz, Riccardo D'Alberto, Yann Desjeux, Laura Eckart, Fabienne Femenia, Helena Hansson, Julia Heinrichs, Wei Huang, Penelope Gouta, Philippe Jeanneaux, Błażej Jendrzejewski, Yan Jin, Julia Jouan, Jochen Kantelhardt, Kevin Kilcline, Vassilia Konstantidelli, Monserrath Lascano Galarza, Gordana Manevska-Tasevska, Andreas Niedermayr, Cathal O'Donoghue, Christoph Pahmeyer, Meri Raggi, Cesar Revoredo-Giha, Mary Ryan, Lena Schaller, Alexandra Sintori, Bethan Thompson, Luiza Toma, Irene Tzouramani, Davide Viaggi, Peter Walder, Matteo Zavalloni, Katarzyna Zawalińska. We thank Laure Latruffe as LIFT project coordinator and Jochen Kantelhardt as WP leader for their input in assembling the deliverable and ensuring a coherent output. Lastly, we thank the reviewers, Lionel Vedrine and Katarzyna Zawalińska for their much valued feedback.

References

Andersson, E., McPhearson, T., Kremer, P., Gomez-Baggethun, E., Haase, D., Tuvendal, M., Wurster, D., 2015. Scale and context dependence of ecosystem service providing units. Ecosyst. Serv. 12, 157–164. https://doi.org/10.1016/j.ecoser.2014.08.001

Axe, M.S., Grange, I.D., Conway, J.S., 2017. Carbon storage in hedge biomass—A case study of actively managed hedges in England. Agric. Ecosyst. Environ. 250, 81–88. https://doi.org/10.1016/j.agee.2017.08.008

Ayouba, K., Latruffe, L., Jeanneaux, P., Desjeux, Y., Dakpo, K.H., 2021. Environmentally-friendly practices in dairy and beef cattle farming in France: results from a survey. Unpublished LIFT working paper.

Bateman, I., Brouwer, R., Cranford, M., 2009. Valuing Environmental Impacts: Practical Guidelines for the Use of Value Transfer in Policy and Project Appraisal. Value Transf. Guidel. … 44, 0–23.

Beillouin, D., Ben-ari, T., Makoswki, D., 2018. Assessing the quality and results of meta-analyses on crop diversification Protocol for systematic review and evidence map.

Beillouin, D., Ben-Ari, T., Makowski, D., 2019. A dataset of meta-analyses on crop diversification at the global scale. Data Br. 24. https://doi.org/10.1016/j.dib.2019.103898

Britz, W., Lengers, B., Kuhn, T., Schäfer, D., 2016. A highly detailed template model for dynamic optimization of farms-FARMDYN.

Broekx, S., Liekens, I., Peelaerts, W., De Nocker, L., Landuyt, D., Staes, J., Meire, P., Schaafsma, M., Van Reeth, W., Van den Kerckhove, O., Cerulus, T., 2013. A web application to support the quantification and valuation of ecosystem services. Environ. Impact Assess. Rev. 40, 65–74. https://doi.org/10.1016/j.eiar.2013.01.003

Burkhard, B., Maes, J., 2017. Mapping Ecosystem Services project Mapping and assessment of ecosystem services to improve resource management and human wellbeing in data-scarce peri-urban ecosystems View project.

Busch, M., La Notte, A., Laporte, V., Erhard, M., 2012. Potentials of quantitative and qualitative approaches to assessing ecosystem services. Ecol. Indic. 21, 89–103. https://doi.org/10.1016/j.ecolind.2011.11.010

Chambers, R.G., Lichtenberg, E., 1994. Simple Econometrics of Pesticide Productivity. Am. J. Agric. Econ. 76, 407–417. https://doi.org/10.2307/1243653

Chan, K.M.A., Satterfield, T., Goldstein, J., 2012. Rethinking ecosystem services to better address and navigate cultural values. Ecol. Econ. 74, 8–18. https://doi.org/10.1016/j.ecolecon.2011.11.011

CICES, 2018. Revision Highlights [WWW Document]. URL https://cices.eu/revision-highlights/ (accessed 12.23.20).

Costanza, R., de Groot, R., Braat, L., Kubiszewski, I., Fioramonti, L., Sutton, P., Farber, S., Grasso, M., 2017. Twenty years of ecosystem services: How far have we come and how far do we still need to go? Ecosyst. Serv. 28, 1–16. https://doi.org/10.1016/j.ecoser.2017.09.008

Dakpo, H., Desjeux, Y., Latruffe, L., 2021. Shadow price of excess nitrogen with multi-technology modelling: the case of wheat plots in France., in: XVIth EAAE Congress. Virtual.

Dakpo, H., Femenia, F., 2021. Pesticides efficiency of French wheat producers under a stochastic frontier framework. Unpublished LIFT working paper.

de Groot, R.S., Matthew, A.W., Roelof, M.J.B., 2002. A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecol. Econ. 41, 393–408.

Di Falco, S., Chavas, J.-P., 2006. Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture. . Eur. Rev. Agric. Econ. 33, 289–314.

Färe, R., Grosskopf, S., Noh, D.-W.W., Weber, W., 2005. Characteristics of a polluting technology: theory and practice. J. Econom. 126, 469–492.

Folke, C., 2007. Social-ecological systems and adaptive governance of the commons. Ecol. Res. 22, 14–15. https://doi.org/10.1007/s11284-006-0074-0

Gómez-Baggethun, E., de Groot, R., Lomas, P.L., Montes, C., 2010. The history of ecosystem services in economic theory and practice: From early notions to markets and payment schemes. Ecol. Econ. 69, 1209–1218. https://doi.org/10.1016/j.ecolecon.2009.11.007

Gómez-Limón, J.A., Picazo-Tadeo, A.J., Reig-Martínez, E., 2012. Eco-efficiency assessment of olive farms in Andalusia. Land use policy 29, 395–406. https://doi.org/10.1016/j.landusepol.2011.08.004

Gromke, C., Jamarkattel, N., Ruck, B., 2016. Influence of roadside hedgerows on air quality in urban street canyons. Atmos. Environ. 139, 75–86. https://doi.org/10.1016/j.atmosenv.2016.05.014

Haines-Young, R., Potschin, M., 2018. CICES V5. 1. Guidance on the Application of the Revised Structure. Fabis Consult. 53.

Heinrichs, J., Britz, W., 2021. The impact of agri-environmental and climate measures on sustainable farm performance – a German case study analysis. Unpublished LIFT working paper.

Heinrichs, J., Jouan, J., Pahmeyer, C., Britz, W., 2021. Integrated assessment of legume production challenged by European policy interaction: A case-study approach from French and German dairy farms. Unpublished LIFT working paper. Q Open 1.

Holden, J., Grayson, R.P., Berdeni, D., Bird, S., Chapman, P.J., Edmondson, J.L., Firbank, L.G., Helgason, T., Hodson, M.E., Hunt, S.F.P., Jones, D.T., Lappage, M.G., Marshall-Harries, E., Nelson, M., Prendergast-Miller, M., Shaw, H., Wade, R.N., Leake, J.R., 2019. The role of hedgerows in soil functioning within agricultural landscapes. Agric. Ecosyst. Environ. 273, 1–12. https://doi.org/10.1016/j.agee.2018.11.027

Huang, W., Manevska-Tasevska, G., Hansson, H., 2021. Integrating a crop diversity index to eco-efficiency measurement for cropland farms in Sweden, in: XVIth EAAE Congress. Virtual.

Jendrzejewski, B., Zawalińska, K., 2021. Indicators of techno-economic, environmental and employment performance based on the LIFT large-scale farmer survey: t-tests and propensity score matching approach. Unpublished LIFT working paper.

Johnston, R.J., Rolfe, J., Rosenberger, R.S., Brouwer, R., 2015. The Economics of Non-Market Good and Resources: Benefit Transfer of Environmental and Resource Values - A Guide for Researchers and Practitioners.

Kilcline, K., Jin, Y., Ryan, M., O'Donoghue, 2021. Development of a bioeconomic model of pasture-based livestock farms. Unpublished LIFT working paper.

Kirchner, M., Schmidt, J., Kindermann, G., Kulmer, V., Mitter, H., Prettenthaler, F., Rüdisser, J., Schauppenlehner, T., Schönhart, M., Strauss, F., Tappeiner, U., Tasser, E., Schmid, E., 2015. Ecosystem services and economic development in Austrian agricultural landscapes - The impact of policy and climate change scenarios on trade-offs and synergies. Ecol. Econ. 109, 161–174. https://doi.org/10.1016/j.ecolecon.2014.11.005

Kroeger, T., Casey, F., 2007. An assessment of market-based approaches to providing ecosystem services on agricultural lands. Ecol. Econ. 64, 321–332. https://doi.org/10.1016/j.ecolecon.2007.07.021

Kuosmanen, T., Kortelainen, M., 2005. Measuring Eco-Efficiency of Production with Data Envelopment Analysis. J. Ind. Ecol. 9, 59–72.

Lambarraa, F., Serra, T., Gil, J.M., 2007. Technical efficiency analysis and decomposition of productivity growth of Spanish olive farms. Spanish J. Agric. Res. 5, 259–270.

Lascano Galarza, M., Zavalloni, M., D'Alberto, R., Raggi, M., Viaggi, D., 2021. Discussione dei risultati preliminary sulla performance a livello di azienda agricola. Unpublished LIFT working paper.

Lechenet, M., Dessaint, F., Py, G., Makowski, D., Munier-Jolain, N., 2017. Reducing pesticide use while preserving crop productivity and profitability on arable farms. Nat. Plants 3, 17008.

Lerouge, F., Sannen, K., Gulinck, H., Vranken, L., 2016. Revisiting production and ecosystem services on the farm scale for evaluating land use alternatives. Environ. Sci. Policy 57, 50–59. https://doi.org/10.1016/j.envsci.2015.11.015

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143, 1–13. https://doi.org/10.1016/j.geoderma.2007.08.025

Minviel, J.J., Latruffe, L., 2017. Effect of public subsidies on farm technical efficiency: a meta-analysis of empirical results. Appl. Econ. 49, 213–226.

Montgomery, I., Caruso, T., Reid, N., 2020. Hedgerows as Ecosystems: Service Delivery, Management, and Restoration. Annu. Rev. Ecol. Evol. Syst. 51, 81–102. https://doi.org/10.1146/annurev-ecolsys-012120-100346

Murty, S., Russell, R.R., Levkoff, S.B., 2012. On modeling pollution-generating technologies. J. Environ. Econ. Manage. 64, 117–135.

Niedermayr, A., Kantelhardt, J., Eckart, L., Kohrs, M., Schaller, L., Walder, P., Ayouba, K., Bakucs, Z., Baráth, L., Barnes, A., Britz, W., Dakpo, K.H., D'Alberto, R., Desjeux, Y., Féménia, F., Fertő, I., Gouta, P., Hansson, H., Hein-richs, J., Huang, W., Jeanneaux, P., Jin, Y., Jouan, J., Kilcline, K., Konstantidelli, V., Krupin, V., Kuhn, T., Lascano Galarza, M.X., Latruffe, L., Letort, E., Manevska-Tasevska, G., O'Donoghue, C., Pahmeyer, C., Raggi, M., Ridier, A., Ryan, M., Sintori, A., Thompson, B., Toma, L., Tzouramani, I., Viaggi, D., Zavalloni, M., Zawalińska, K., 2021a. Farm technical-economic performance depending on the degree of ecological approaches. LIFT Deliverable 3.1.

Niedermayr, A., Walder, P., Schaller, L., Eckart, L., Kantelhardt, J., 2021b. Technical economic and environmental farm performance of dairy farms in Austrian case study regions Steyr-Kirchdorf and Salzburg und Umgebung. Unpublished LIFT working paper.

Niedermayr, A., Walder, P., Schaller, L., Kantelhardt, J., 2021c. Technical economic and environmental performance of Austrian dairy farms. Unpublished LIFT working paper.

O'Brien, D., Hennessy, T., Moran, B., Shalloo, L., 2015. Relating the carbon footprint of milk from Irish dairy farms to economic performance. J. Dairy Sci. 98, 7394–7407.

OECD, 1998. Eco-efficiency. Paris.

Potschin, M.B., Haines-Young, R.H., 2011. Ecosystem services: Exploring a geographical perspective. Prog. Phys. Geogr. 35, 575–594. https://doi.org/10.1177/0309133311423172

Power, A.G., 2010. Ecosystem services and agriculture: Tradeoffs and synergies. Philos. Trans. R. Soc. B Biol. Sci. 365, 2959–2971. https://doi.org/10.1098/rstb.2010.0143

Pretty, J., 2008. Agricultural sustainability: Concepts, principles and evidence. Philos. Trans. R. Soc. B Biol. Sci. 363, 447–465. https://doi.org/10.1098/rstb.2007.2163

Pretty, J., Bharucha, Z.P., 2014. Sustainable intensification in agricultural systems. Ann. Bot. 114, 1571–1596. https://doi.org/10.1093/aob/mcu205

PRISMA, 2015. PRISMA: TRANSPARENT REPORTING of SYSTEMATIC REVIEWS and META-ANALYSES [WWW Document]. URL http://www.prisma-statement.org/ (accessed 3.3.21).

Rawluk, A., Ford, R., Anderson, N., Williams, K., 2019. Exploring multiple dimensions of values and valuing: a conceptual framework for mapping and translating values for social-ecological research and practice. Sustain. Sci. 14, 1187–1200. https://doi.org/10.1007/s11625-018-0639-1

Rega, C., Paracchini, M.L., Mccraken, D., Saba, A., Zavalloni, M., Raggi, M., Viaggi, D., Britz, W., Frappier, L., 2018. LIFT-Deliverable D1.1 Review of the definitions of the existing ecological approaches.

Rega, C., Thompson, B., D'Alberto, R., Niedermayr, A., Kantelhardt, J., Gouta, P., Konstantidelli, V., Tzouramani, I., Desjeux, Y., Latruffe, L., Billaudet, L., Paracchini, M.L., 2021. Deliverable 1.4. LIFT farm typology developed, tested and revised, and recommendations on data needs.

Rigby, D., Woodhouse, P., Young, T., Burton, M., 2001. Constructing a farm level indicator of sustainable agricultural practice. Ecol. Econ. 39, 463–478. https://doi.org/10.1016/S0921-8009(01)00245-2

Sandhu, H.S., Wratten, S.D., Cullen, R., 2010. Organic agriculture and ecosystem services. Environ. Sci. Policy 13, 1–7. https://doi.org/10.1016/j.envsci.2009.11.002

Scarpa, R., Campbell, D., Hutchinson, W.G., 2015. Benefit Estimates for Landscape Improvements: Sequential Bayesian Design and Respondents' Rationality in a Choice Experiment. Land Econ. 83, 617–634. https://doi.org/10.3368/le.83.4.617

Sintori, A., Gouta, P., Konstantidelli, V., Tzouramani, I., 2021. Estimating eco-efficiency of the olive farms in Crete, Greece. Unpublished LIFT working paper.

Stiles, W., 2016. The Potential of tree and Hedgerow Planting to Reduce the Frequency and Impact of Flood Events in the UK [WWW Document]. Welsh Gov. URL https://businesswales.gov.wales/farmingconnect/news-and-events/technical-articles/potential-tree-and-hedgerow-planting-reduce-frequency-and-impact-flood-events-uk (accessed 6.14.21).

Swinton, S.M., Lupi, F., Robertson, G.P., Hamilton, S.K., 2007. Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. Ecol. Econ. 64, 245–252. https://doi.org/10.1016/j.ecolecon.2007.09.020

Swinton, S.M., Lupi, F., Robertson, G.P., Landis, D.A., 2006. Ecosystem services from agriculture: Looking beyond the usual suspects. Am. J. Agric. Econ. 88, 1160–1166. https://doi.org/10.1111/j.1467-8276.2006.00927.x

Toma, L., Revoredo-Giha, C., Thompson, B., Barnes, A., 2021. An exploration of the labour and environmental efficiency of Scottish cattle and sheep farms. Unpublished LIFT working paper.

Turner, R.K., Daily, G.C., 2008. The ecosystem services framework and natural capital conservation. Environ. Resour. Econ. 39, 25–35. https://doi.org/10.1007/s10640-007-9176-6

Tzouramani, I., Latruffe, L., Konstantidelli, V., Desjeux, Y., Bailey, A., Bardounioti, M., Barnes, A., Bigot, G., Dakpo, H., Davidova, S., Duval, J., Eichhorn, T., Gerner, L., Henderson, S., Hostiou, N., Jeanneaux, P., Kantelhardt, J., Larmet, V., Legras, S., Niedermayr, A., Paracchini, M.L., Polge, E., Rega, C., Schaller, L., Solomou, A., Thompson, B., Toma, L., Treguer, S., Tzanopoulos, J., Vedrine, L., Walder, P., 2019. Deliverable 2.2. LIFT Large-Scale Farmer Survey Questionnaire.

Van Vooren, L., Reubens, B., Ampoorter, E., Broekx, S., Pardon, P., Van Waes, C., Verheyen, K., 2018. Monitoring the Impact of Hedgerows and Grass Strips on the Performance of Multiple Ecosystem Service Indicators. Environ. Manage. 62, 241–259. https://doi.org/10.1007/s00267-018-1043-4

Van Vooren, L., Reubens, B., Broekx, S., De Frenne, P., Nelissen, V., Pardon, P., Verheyen, K., 2017. Ecosystem service delivery of agri-environment measures: A synthesis for hedgerows and grass strips on arable land. Agric. Ecosyst. Environ. 244, 32–51. https://doi.org/10.1016/j.agee.2017.04.015

VLM Vlaamse Land Maatschappij, 2017. Beheerovereenkomsten voor kleine landschapselementen. Brussels.

WBCSD, 1992. See Under Cross Cutting Themes: Eco-efficiency, International Food Policy Research Institute (IFPRI). Washington DC.

Wezel, A., Nicholls, C.I., Altieri, M.A., Vazquez, L., 2017. Agroecological Principles for the Conversion of Farming Systems. Agroecol. Pract. Sustain. Agric. 1–18. https://doi.org/10.1142/9781786343062_0001

Zhu, X., Karagiannis, G., Oude Lansink, A., 2011. The impact of direct income transfers of CAP on Greek olive farms’ performance: using a non-monotonic inefficiency effects model. J. Agric. Econ. 62, 630–638.

Annex 1

Table A1. Final list of management practices included in the REA, adapted from Rega et al. (2018).

Management practice (practice clusters)	Individual practices within the clusters	
Use of chemical fertiliser inputs	- Use of inorganic fertiliser / chemical fertiliser - Agrochemical input – fertilisers - Mineral fertiliser	
Low fertiliser input	- Low nutrient input - Reduced fertiliser application - Low-solubility mineral fertilisers	
Biological N fixation	- Biological nitrogen fixation - Legume-cereal rotations	- Legumes - Pulse crops
Use of organic fertilisers (incl. manure)	- Manure fertiliser - Farmyard manure - Organic manure	
Use of chemical pesticide inputs	- Use of inorganic pesticide inputs - Herbicide input	- Insecticide input - Agrochemical input - pesticides
Biological pest control	- Bio-control - Biological pest control - Natural pest control	- Plant extract bio-control - Diversionary strategy
Use of organic pesticides	- Biological insecticide - Amendments	- Copper - Sulphur
Low agrochemical pesticide input	- Reduced herbicide application - Reduced insecticide use - Low pesticide input - Seed selection	- Crop variety improvements - Varietal diversity - Local variety - Insect-resistant crops
Alternative weed management strategies	- Fumigation - Mechanical weeding - Push-pull system	- Manual weeding - Integrated pest management (IPM)
Cover crops	- Catch crop - Clover	
Conservation tillage	- Strategic tillage - Reduced soil cultivation - Minimum tillage - Shallow tillage - No tillage	- Occasional tillage - Ridge till - Asynchronous tilling - Direct sowing
Crop rotation	- Crop sequence - Dryland rotation	- Irrigated rotation - Diversification of crop rotation
Crop residue management	- Crop sequence - Dryland rotation	- Irrigated rotation - Diversification of crop rotation
Mulching	- Organic mulching - mulching	

Sustainable water management	- Deficit irrigation - Reduced irrigation - Drainage	- No irrigation - Flooding
Agroforestry	- Agroforestry	
Extensive livestock systems	- Transhumance - Silvopasture	
Crop livestock integration	- Animal circulation - Crop-livestock integration - Grassland-livestock integration	
Semi-natural habitats	- Diversified field edges - Conservation buffers - Border planting - Ecological compensation areas - Ecological focus area - (Agro) ecological infra-structure (management) - Grassy buffer strips	- Semi-natural habitat - Wildlife plots - Hedgerows - Insectary strips - Living fences - Noncrop plantings - Beneficial fauna
Spatial heterogeneity	- Diversification - Farm heterogeneity	- Spatial diversity - Patch intensification
Agri-environmental schemes	- Agri-environmental schemes	
Sustainable grazing	- Grass ley - Ley farming - Perennial leys with leg-umes - Improved pastures - Grassland mixtures - Grazing	- Grazing on crop residues - Low density of livestock - Low stocking rates - Use of fallow - Rotational grazing
Selection of breeds (genetic diver-sity, traditional/local breeds)	- Breed selection - Genetic diversity - Local breed	
Low mechanisation	- No mechanisation - Low mechanisation	- Manual cuts - Blade mowing machine cuts
Precision farming	- Precision farming - Precision livestock farming	
Intercropping	- Alley intercropping - Intercropping - Multiple intercropped species	- Relay intercropping - Polyculture

Table A2. Comprehensive search string and search string components used in the REA.

Ecological farming practices key words	
1	(((("chemical input$" OR fertili* OR nutrient$ OR nitrogen OR "low input$" OR fertili* OR nutrient$ OR nitrogen OR "low input$" OR "seed selection" OR "crop variet*" OR "variety improvement" OR "varietal diversity" OR "local variet*" OR "insect resist*" OR "water management" OR precision OR "organic farming" OR "organic agrosystems") NEAR/2 (farm* OR agr*))) OR (mechani$ation NEAR (farm* OR agr*)) OR ("crop rotation$" OR "crop residue" OR (diversification NEAR/2 (crop OR farm* OR agr*)) OR (spatial NEAR/2 (heterogeneity OR diversity) OR habitat NEAR/2 (heterogeneity OR diversity)) AND (farm* OR agr*)) OR ((pest$ NEAR management OR bio$control OR pest NEAR control OR (("diversionary strategy" OR push-pull) AND pest$)) AND (farm* OR agr* OR crop$)) OR ("breed selection" OR "local breed$" OR breed NEAR "genetic diversity") OR (((buffer OR flower OR grass) NEAR/1 strip) OR "field margin" OR hedgerow) OR (integrat* NEAR/0 (farm* OR agr* OR "crop-livestock$")) OR (agroforestry) OR ("cover crop$" OR "catch crop$") OR (till* AND (farm* OR agr*)) OR ((ley$ NEAR/1 (grass OR farming OR perennial) OR "improved pasture$" OR "grassland mixture$" OR grazing NEAR ("stocking rate$" OR density OR pressure OR intensity OR management) OR fallow OR "rotational$grazing")))
Article type key words	
2	(review OR "meta analysis" OR meta-analysis OR "rapid evidence assessment" NOT "systematic map*")
Geographic qualifier keywords	
3	(Global* OR Temperate OR Mediterranean OR Euro* OR Eurasia OR Albania* OR Andorra* OR Armenia* OR Austria OR Austrian OR Azerbaijan* OR Belarus* OR Belgium OR Bosnia* OR Herzegovina* OR Bulgaria* OR Croatia* OR Cyprus OR Czech OR Dannish OR Denmark OR Escadinavia* OR Estonia* OR Finnish OR Finland OR French OR France OR Georgia* OR German* OR Greek OR Greece OR Hungar* OR Iberia* OR Iceland* OR Irish OR Ireland OR Ital* OR Kazakhstan* OR Kosov* OR Latvia* OR Liechtenstein* OR Lithuania* OR Luxembourg* OR Macedonia* OR Malt* OR Moldova* OR Monegasque OR Monaco* OR Montenegr* OR Dutch OR Holland OR Netherlands OR Norwegian OR Norway OR Polish OR Poland OR Portug* OR Romania* OR Russia* OR Sammarinense "San Marino" OR Serbia* OR Slovak* OR Slovenia* OR Spanish OR Spain OR Swed* OR Swiss OR Switzerland OR Turk* OR Ukrain* OR British OR Britain OR English OR England OR Scottish OR Scotland OR Welsh OR Wales OR "United Kingdom" OR UK OR Vatican)
Topical qualifier key words	
4	(Tropic* OR smallholder OR small$holder OR marine OR aquaculture)
Final combination of search groups in search string	
5	1 **AND** 2 **AND** 3 **NOT** 4

List of 4 reference articles used in REA

- Palomo-Campesino, S., González, J.A. and García-Llorente, M., 2018. Exploring the connections between agroecological practices and ecosystem services: A systematic literature review. Sustainability, 10(12), p.4339.
- Laura, V.V., Bert, R., Steven, B., Victoria, N., Paul, P. and Kris, V., 2017. Ecosystem service delivery of agri-environment measures: a synthesis for hedgerows and grass strips on arable land. Agriculture, Ecosystems & Environment, 244, pp.32-51.
- Snyder, C.S., Bruulsema, T.W., Jensen, T.L. and Fixen, P.E., 2009. Review of greenhouse gas emissions from crop production systems and fertilizer management effects. Agriculture, Ecosystems & Environment, 133(3-4), pp.247-266.
- Lee, H., Lautenbach, S., Nieto, A.P.G., Bondeau, A., Cramer, W. and Geijzendorffer, I.R., 2019. The impact of conservation farming practices on Mediterranean agro-ecosystem services provisioning—a meta-analysis. Regional Environmental Change, pp.1-16.

Table A3. PICO (Population, Intervention, Comparator, Outcome) used as inclusion/exclusion criteria in the REA.

PICO	COMPONENT	OBJECTIVE
POPULATION	Quantitative or qualitative literature review studies	Robustly inform environmental assessment indicators using pre-existing literature reviews, quick scoping reviews, rapid evidence assessments, meta-analyses, systematic reviews and reviews of reviews; quantitative and qualitative data was selected to be input into indicators
POPULATION	European agricultural land	Use the most locally-relevant data on practices and their effects
INTERVENTION/EXPOSURE	Ecological farming practices	Cover the variety of practices to be included in the environmental assessment
COMPARATOR	Conventional agricultural practices	Compare conventional approaches to agriculture with more ecological approaches (embedded within the literature reviews)
OUTCOME	Ecosystem services/public good (ES and PG) provision	Measure, through the use of indicators, proxies or qualitative data, the impact of adoption of farming practices on ES and PG provision

Table A4. Quality criteria used in the REA to assess quality of articles, adapted from Beillouin et al. (2019) and PRISMA (2015).

Expected section in article	Criterion number	Criterion name	Criterion description
Results	RS-01	Research question(s) and objectives are explicitly stated	An explicit statement of the questions and objectives being addressed is provided, with reference to their key elements (i.e. population or participants, concepts, interventions, comparators, outcomes, context and study designs) or other relevant key elements used to conceptualise the review questions and/or objectives.
Methods	SS-01	Literature databases are mentioned	All databases used in the search are listed.
Introduction	SS-05	Additional literature search is performed	At least one of the 5 following items is met: 1) the literature search is performed in the Google Scholar database; 2) additional studies were searched through a local scientific journal search engine; 3) additional studies were identified through expert knowledge; 4) additional studies were identified in any other way; 5) additional studies were collected from the reference list of published articles.
Methods	SS-04	Date(s) of the search(es) are mentioned	Date(s) of the search(es) for each information source are given. Or, at least, the date of the most recent search executed is given.
Methods	SS-02	Keywords used in the search are given	The keywords used in the database search(es) are given.
Methods	SS-03	Full boolean search string is provided	The full electronic search string for at least 1 database is presented, including booleans and any limits used, such that it could be repeated.
Methods	RS-02	Inclusion and exclusion criteria are mentioned	The study characteristics (e.g. population, interventions, outcomes) and reporting characteristics (e.g. years considered, language, publication status) used as criteria for eligibility, and respective rationale for inclusion/exclusion, are given.
Methods	SS-06	Steps of the screening process are reported	The screening process is transparent, i.e. the following 3 items are given, ideally using a flow diagram: 1) the number of studies originally retrieved from the literature search (including any grey literature searched); 2) the number of selected (included) studies at the end of the

			screening process; 3) the number of studies excluded at each step of the screening process, with reasons for exclusions at each stage.
Results	SS-08	Characteristics of the included studies are given	For each study, the characteristics for which data were extracted (e.g. study length, population, location, specific interventions) are provided.
Results	SS-09	List of excluded studies is provided	The study includes the final reference list of excluded individual studies.
Results	DS-03	Sofware tools used in the review are mentioned	The names of the software/package(s) used to perform statistical analyses (e.g. Metafor in R), or any other software tools used in the review, are given.
Results	DS-02	Methods of handling and summarising the data are described.	The methods of handling data and combining results of studies are described, and statistical models used identified. If meta-analyses have been done, the model used to estimate average effect size should be clear and reproducible (e.g. is other information on the statistical model presented (e.g. equations, parameters, etc.)?
Methods	DS-01	Method of data extraction is described	Methods for extracting data from figures and charts, such as the use of digitising software (e.g. Webplot Digitilizer, Datathief, GetData_Graph_Digitizer) or obtaining the source data from authors, are described.
Results	SS-07	List of included studies is given	The study includes the final reference list of the selected individual studies used in the review.
Methods	BI-01	A protocol is published prior to publication of the review	An a priori review protocol exists and can be accessed (e.g. paper, web address), and, if available, provides a registration number.

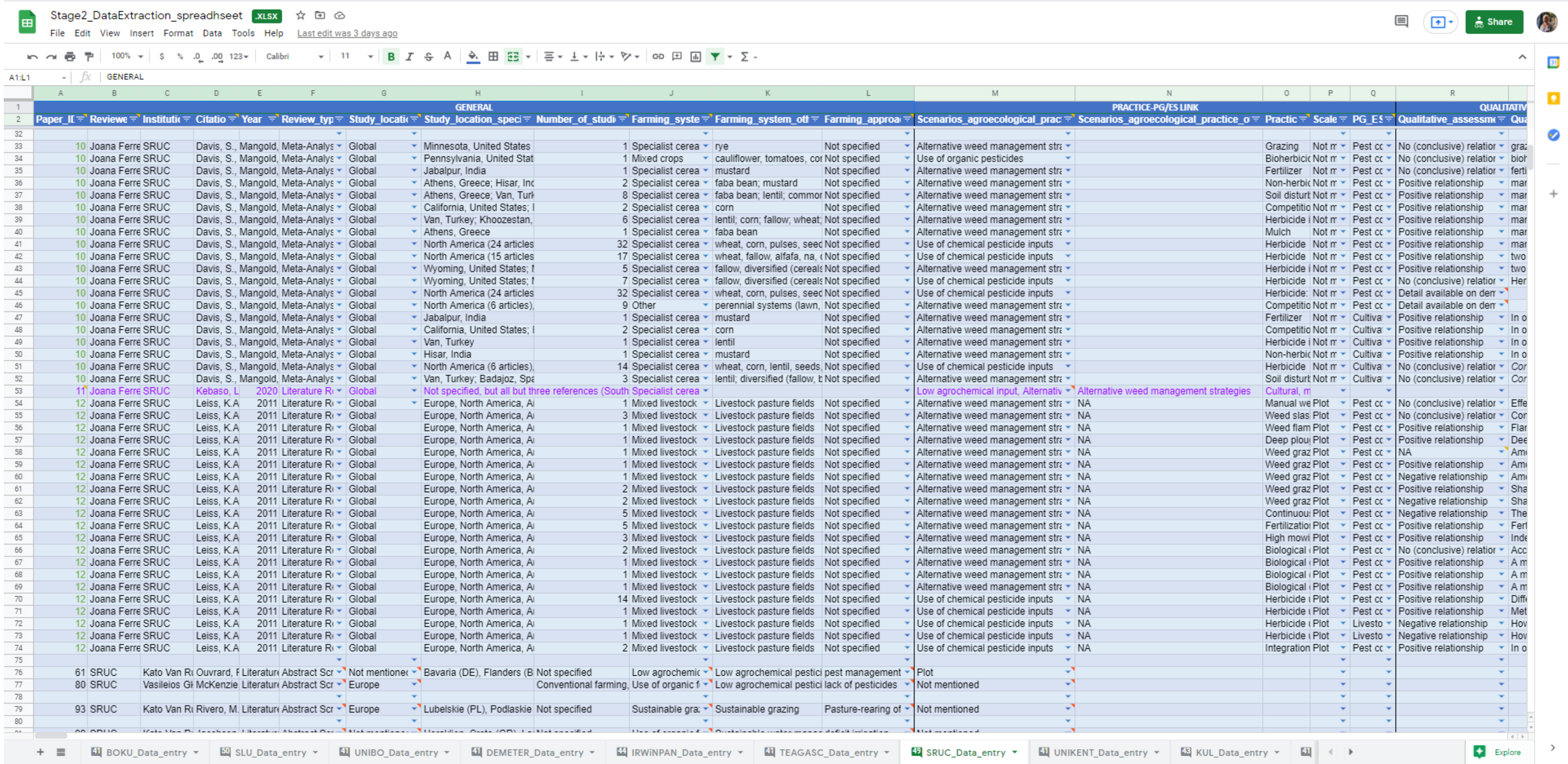

Figure A1. Screenshot of the template used to extract meta-analytic (columns D-Q) data from articles included in the REA.

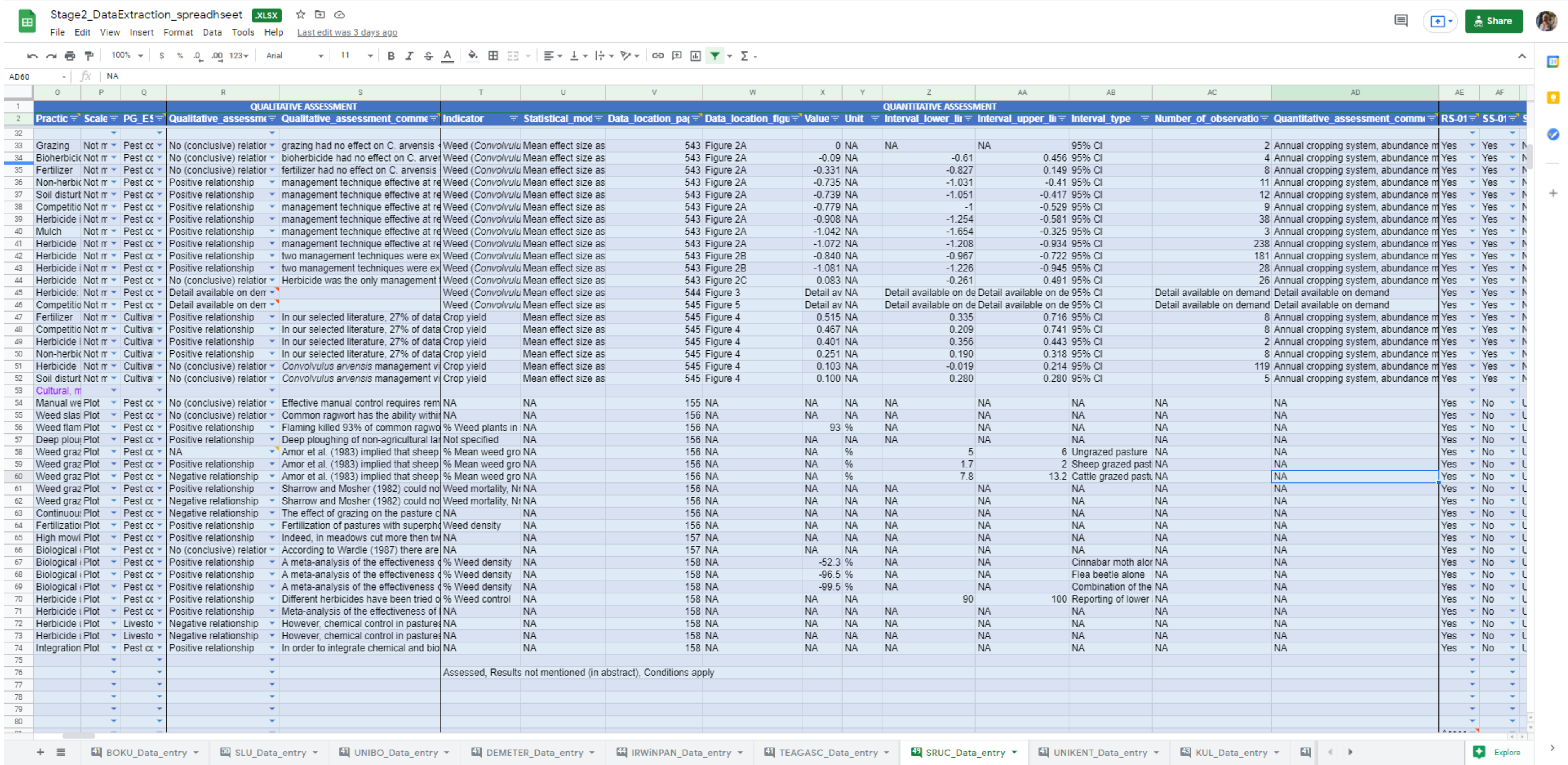

Figure A2. Screenshot of the template used to extract qualitative (column S), semi-qualitative (column R) and quantitative (columns T-AD) data from the articles included in the REA.

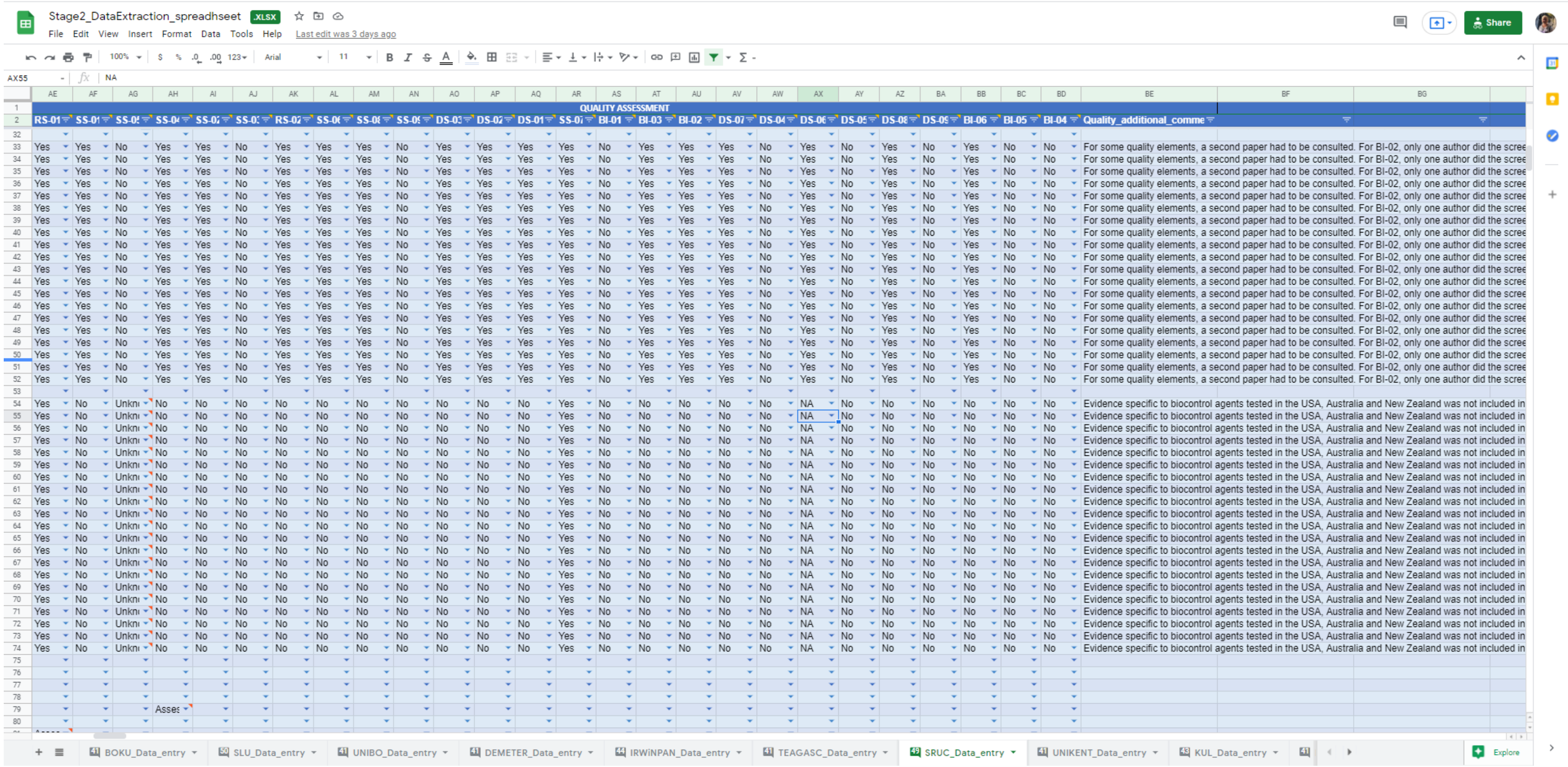

Figure A3. Screenshot of the template used to extract quality (columns AG-BE) data from the articles included in the REA using the 26 quality criteria adapted from Beillouin et al. (2018) and PRISMA (2015).

Annex 2

In this paper we calculate SIs for 26 management practices at farm and territorial level, holding key assumptions regarding the trade-off between quantity and quality of evidence, the definition of an observation, as well as the suitability of the consensus measure as a measure of variance. We perform sensitivity analyses in which we relax the above-mentioned assumption and look at how this affects our SIs.

As observations were derived from articles which synthesised results from a variety of primary articles, and because many of the ES against which management practices were evaluated were quite broadly defined during data collection, we allowed for multiple observations to be derived for the same management practice-ES link derived from the same article. This way we were able to ensure that enough variation is captured by the SI. Though output from the REA was thoroughly cleaned prior to SI composition, we performed a sensitivity analysis to assess for double counting. To do this, we performed a separate calculation of the SIs, this time allowing for only one observation per management practice-ES link from a single article to be included. A comparison between the two sets of SIs found no significant differences, and a ranking exercise demonstrated that the highest and lowest ranked SIs did not change between sets.

Further, we compare the use of consensus (c) against the use of variance to evaluate the degree to which observations within a single SI take the same value. We found a strong correlation between the two measures, $r(861) = -0.9506, p < 0.001$. Therefore, considering the consensus is more suited to reflect agreement amongst ordinal data, we opt to maintain the use of this measure in any further reporting.

Finally, and perhaps most importantly, we test the assumption of the increased importance of evidence quality over quantity made in the correction factor (w_{ij}). We do this by calculating SIs for each trade-off factor r ranging from $r = 0.1$ to $r = 0.9$, increasing r by 0.1 with each iteration. An ANOVA found a significant difference between the SIs composed using r values ranging from $r = 0.1$ to $r = 0.9$ at both farm ($p < 0.001$) and territorial ($p < 0.001$) level. This demonstrates that by changing the assumptions regarding the trade-off between quality and quantity of evidence to increasingly favour quantity, the magnitude of the SIs change. However, a ranking exercise demonstrated that the order of the SIs ranked from highest to lowest magnitude at both farm and territorial level does not change with increasing r values. Therefore we maintain the assumption made regarding the trade-off between quality and quantity of evidence in favour of quality, positing that when considering secondary literature as a data source, evidence quality more accurately captures confidence reflected within the correction factor.