Assessing sustainability at farm-level: Lessons learned from a comparison of tools in practice

Article *in* Ecological Indicators · July 2016

DOI: 10.1016/j.ecolind.2016.01.047

5 authors, including:


Evelien De Olde
Aarhus University


Frank Oudshoorn
SEGES


Claus G Sørensen
Aarhus University

Ecological Indicators 66 (2016) 391–404

Contents lists available at ScienceDirect

Ecological Indicators

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

Assessing sustainability at farm-level: Lessons learned from a comparison of tools in practice

Evelien M. de Olde[a,c,*], Frank W. Oudshoorn[a,b], Claus A.G. Sørensen[a], Eddie A.M. Bokkers[c], Imke J.M. de Boer[c]

[a] Aarhus University, Department of Engineering, Inge Lehmanns Gade 10, 8000 Aarhus, Denmark
[b] SEGES, Agro Food Park 15, 8200 Aarhus N., Denmark
[c] Animal Production Systems Group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 6 July 2015
Received in revised form 19 January 2016
Accepted 25 January 2016
Available online 19 February 2016

Keywords:
Agricultural production
Farm level
Livestock
Relevance
Sustainability assessment tools

ABSTRACT

In the past decades a wide variety of tools have been developed to assess the sustainability performance of farms. Although multiple studies have compared tools on a theoretical basis, little attention has been paid to the comparing tools in practice. This research compared indicator-based sustainability assessment tools to gain insight in practical requirements, procedures and complexity involved in applying sustainability assessment tools. In addition, the relevance of the tools, as perceived by farmers, was evaluated. An overview of 48 indicator-based sustainability assessment tools was developed to, subsequently, select tools that address the environmental, social and economic dimension of sustainability, are issued in a scientific publication and suitable for assessing the sustainability performance of livestock and arable farms in Denmark. Only four tools (RISE, SAFA, PG and IDEA) complied with the selection criteria and were used to assess the sustainability performance of five Danish farms. The tools vary widely in their scoring and aggregation method, time requirement and data input. The farmers perceived RISE as the most relevant tool to gain insight in the sustainability performance of their farm. The findings emphasize the importance of context specificity, user-friendliness, complexity of the tool, language use, and a match between value judgements of tool developers and farmers. Even though RISE was considered as the most relevant tool, the farmers expressed a hesitation to apply the outcomes of the four tools in their decision making and management. Furthermore, they identified limitations in their options to improve their sustainability performance. Additional efforts are needed to support farmers in using the outcomes in their decision making. The outcomes of sustainability assessment tools should therefore be considered as a starting point for discussion, reflection and learning.

1. Introduction

Agricultural production significantly contributes to, for example, climate change, water pollution, and loss of biodiversity, and increasingly competes for natural resources, such as land and phosphorus (Steinfeld et al., 2006). Moreover, social concerns arise about the impact of agricultural production on public health and animal welfare, and diminishing farm profitability (Bos et al., 2009). The urgency of sustainable development of agricultural production, therefore, is increasingly acknowledged (Pretty, 2008; Tilman et al., 2002; Wiskerke, 2009). To enable a transition towards more sustainable production, a wide range of tools have been developed to gain insight in the sustainability performance of agricultural systems (Binder et al., 2010; Schader et al., 2014). Indicator-based sustainability assessment tools vary widely in their scope (geographical and sector), target group (e.g. farmers or policy makers), selection of indicators, aggregation and weighing method, and time requirement for execution (Binder et al., 2010; Marchand et al., 2014; Schader et al., 2014). Although many stress the importance of integrating environmental, economic and social themes in sustainability assessment tools, environmental themes and tools generally receive more attention (Binder et al., 2010; Finkbeiner et al., 2010; Lebacq et al., 2013; Marta-Costa and Silva, 2013; Schader et al., 2014).

* Corresponding author at: Aarhus University, Department of Engineering, Inge Lehmanns Gade 10, 8000 Aarhus, Denmark. Tel.: +45 23360087.
E-mail address: evol@eng.au.dk (E.M. de Olde).

http://dx.doi.org/10.1016/j.ecolind.2016.01.047
1470-160X/

1.1. Hierarchical structure in sustainability assessment tools

Indicator-based sustainability assessment tools are generally structured following three or four hierarchical levels (Fig. 1). A wide diversity of terminology, however, is used to define the various levels (Fig. 1) (Bausch et al., 2014; Bélanger et al., 2012; De Boer and Cornelissen, 2002; Guerci et al., 2013; Haas et al., 2000; van Calker et al., 2007; Van Cauwenbergh et al., 2007). This paper follows the structure suggested in the SAFA guidelines (FAO, 2013). A dimension is a pillar of sustainability and is the highest and most general level in the structure of a tool. On the intermediate level, universal sustainability goals are translated into themes and, in some cases, made more explicit in subthemes. Finally, indicators are measurable variables to evaluate the sustainability performance for the (sub) theme (FAO, 2013). The indicator value can be derived in different ways, e.g. through measurement, expert opinion or model estimates (Van Cauwenbergh et al., 2007). To evaluate the indicator value, a desired level for each indicator is described by means of a reference value (Acosta-Alba and van der Werf, 2011; Van Cauwenbergh et al., 2007). Reference values can be absolute or relative values. Absolute values can be divided into target values identifying a desirable condition (e.g. legal norm), and threshold values defining a minimum or maximum acceptable level (e.g. political interpretation of scientific findings) (Van Cauwenbergh et al., 2007). Relative reference values compare indicator values with an initial value, regional or sample average or desirable trend (Lebacq et al., 2013; Van Cauwenbergh et al., 2007).

1.2. The adoption of sustainability assessment tools in practice

Sustainability assessment tools can provide support to on-farm decision making and hereby may have a significant impact on a sustainable development of farms (Le Gal et al., 2011; Marchand et al., 2014). So far, however, the actual adoption of sustainability assessment tools by agricultural practice is relatively limited (Binder et al., 2010; Triste et al., 2014). In the development of a sustainability tool, tool developers make value judgements and assumptions, for example, on what is sustainability, what is a sustainable level of production, which indicators to select, and how to measure, weigh and aggregate the indicators (Gasparatos, 2010). A mismatch between these value judgements and assumptions of tool developers and its users (i.e. farmers and advisors) can result from insufficient involvement of these users during the development of a tool, and is considered as a reason for the limited adoption of sustainability assessment tools in farming practice (De Mey et al., 2011; Gasparatos, 2010; Triste et al., 2014; Van Meensel et al., 2012). Furthermore, data availability and quality, time and budget requirements as well as factors related to unfamiliar terminology, user-friendliness, and tool accessibility influence the farmers' perception of the tool's relevance and, consequently, the adoption of tools (Lynch et al., 2000; Marchand et al., 2014; Van Meensel et al., 2012).

Farmers' adoption of sustainability assessment tools and their outcomes is a key issue when considering to use sustainability assessment tools to contribute to the sustainable development of farms (Triste et al., 2014). Literature on the adoption of tools by farmers emphasizes the importance of the perceived relevance of the tool which is determined by a combination of factors mentioned above (Van Meensel et al., 2012). Relevance can be defined as: '*Something is relevant to a task if it increases the likelihood of accomplishing the goal which is implied by the task*' (Hjørland and Christensen, 2002). As stated by McCown (2002) farmers cease to care about tools when they can't see sufficient value for action resulting from the output.

The aim of this study was to compare sustainability assessment tools in practice and discuss the relevance as perceived by farmers. The importance of such an end-user validation of sustainability indicators and methods was raised by Bockstaller and Girardin (2003). By applying multiple tools on farms, insights are obtained in the practical and operational requirements, procedures and the complexity involved in applying sustainability assessment tools in practice. This adds another dimension to existing studies focused on comparing tools on a theoretical basis. An overview of existing tools was developed to, subsequently, select tools that address the environmental, social and economic dimension of sustainability, focus at farm level, are issued in a scientific publication and suitable for assessing the sustainability performance of livestock and arable farms in North-West Europe. The tools were applied on five Danish farms as a case, and compared using the framework of Marchand et al. (2014), adapted from Binder et al. (2010).

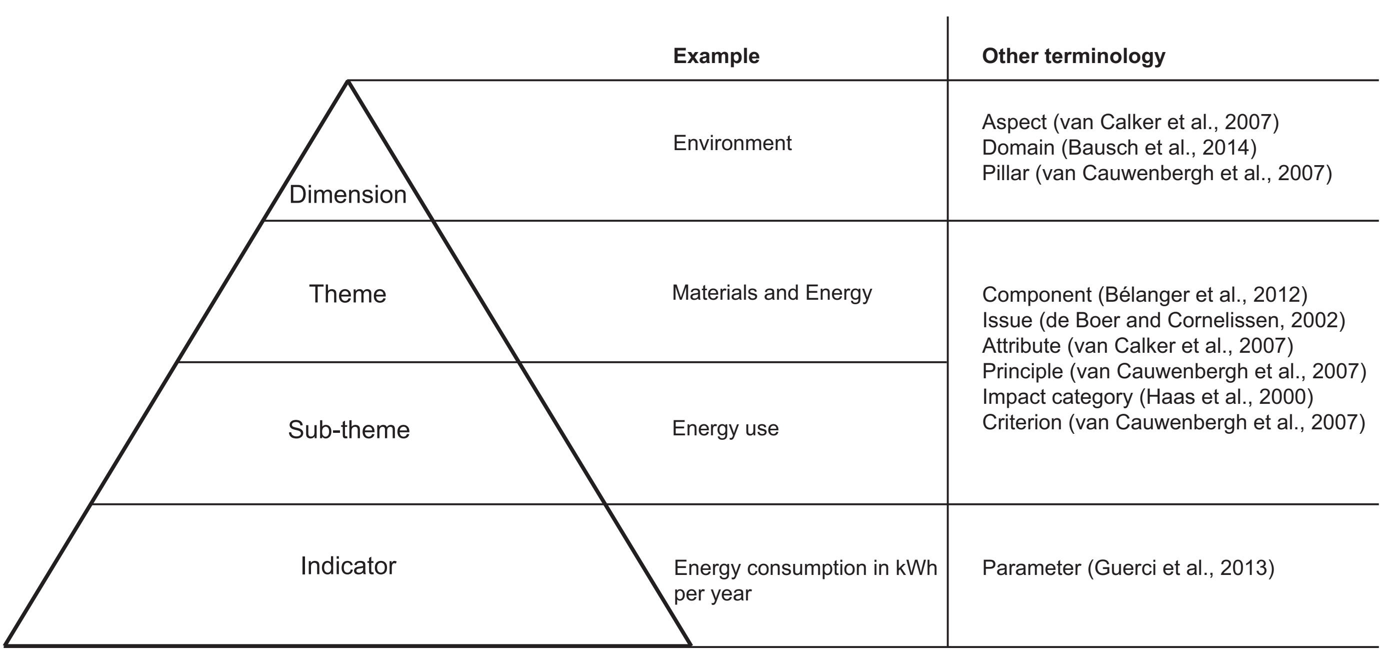

Fig. 1. Hierarchical levels in sustainability assessment according to SAFA as used in the present study, and terminology used in other sustainability assessment studies.

2. Methods

2.1. Overview of sustainability assessment tools

Different terms are used in literature to describe sustainability assessments such as methods, methodological approaches, frameworks, and tools (Marchand et al., 2014; Schader et al., 2014; Schindler et al., 2015). In this paper we focus on those sustainability assessments that have been developed into tools aimed at ex post assessments of the sustainability performance of farms using indicators, so-called indicator-based sustainability assessment tools. An overview of available tools was established using the search engine Scopus and the snowball method. The following list of reviews and scientific papers on sustainability assessments were identified through a literature study (Acosta-Alba and van der Werf, 2011; Binder et al., 2010; FAO, 2013; Galan et al., 2007; Halberg et al., 2005; Marchand et al., 2012; Marta-Costa and Silva, 2013; Schader et al., 2014; Van Cauwenbergh et al., 2007; van der Werf et al., 2007; Van Passel and Meul, 2012). In addition, several recently developed tools that were presented at scientific conferences were added to the list and the developers were contacted to discuss the characteristics of the tools. This resulted in a comprehensive, yet possibly not exhaustive, list of 48 relevant tools (Appendix A).

To select sustainability assessment tools at farm level the following selection criteria were used:

1. To fit the scope of this study the tool has to be aimed at assessing the sustainability performance at farm level using an indicator-based sustainability assessment.
2. To ensure scientific rigour, the tool has to be published in a peer-reviewed scientific journal and/or peer-reviewed scientific report. The publication has to be written by the tool developers and focus on the tool. Thereby, tools that are only mentioned in reviews, without any other scientific publication, are excluded.
3. Economic, environmental and social sustainability indicators should be included to allow an integrated assessment of the farm.
4. The tool has to be suitable for the assessment of livestock and arable farms since this research focuses on Denmark where livestock production generally is combined with arable production.
5. The tool has to be suitable for North-West Europe and applied in more than one country to allow contextualization.
6. The tool should be available in English and/or Danish to allow sufficient understanding of the tool and application in the Danish context.

Tool developers were contacted to check the actual status of the tool and to assure it fitted the selection criteria. Based on the selection criteria, four sustainability farm assessment tools remained for the comparison: RISE (Häni et al., 2003), SAFA (FAO, 2013), Public Goods (PG) (Gerrard et al., 2012), and IDEA (Zahm et al., 2008) (Table 1). Although the (PG) tool has an emphasis on public goods instead of sustainability, the tool developers do consider it as a suitable tool for sustainability assessment and it complied to the selection criteria (Leach et al., 2013). RISE, PG and IDEA are tailored specifically for farm assessments whereas SAFA applies a wider scope by also extending through supply chains in agriculture, forestry and fisheries. To align the tools with the terminology described in the introduction, the aggregation of indicators in RISE, PG and IDEA is considered as sub-themes, instead of what they define indicators.

2.2. Comparison framework

To compare the tools, we applied the framework of Marchand et al. (2014), adapted from Binder et al. (2010) (Table 2). This framework was selected as it has been described in-detail and used to compare a wide range of characteristics of sustainability assessment tools (Marchand et al., 2014). The chronological order of the framework allows the reader to follow the steps included in the sustainability assessment tools. Binder et al. (2010) distinguishes normative, systemic and procedural aspects in sustainability assessment processes. In the normative aspects, the sustainability concept, goal, scoring method and tool function are described (Marchand et al., 2014). The systemic aspects address the ability of the tool to translate the complexity of a system (Binder et al., 2010). As Binder et al. (2010) state, a tool should be as simple as possible while addressing the complexity of the system, and covering relations among indicators. Finally, Marchand et al. (2014) extended the framework of Binder et al. (2010) by incorporating critical success factors for the implementation of sustainability assessment tools as developed by De Mey et al. (2011). The procedural aspects include, amongst others, user-friendliness, data availability and effectiveness. The description of the characteristic 'effectiveness' given by Marchand et al. (2014), however, focuses on relevance. We, therefore, changed the characteristic's name effectiveness into relevance accordingly. This concept is considered more appropriate as it evaluates the contribution to reaching a goal (i.e. sustainable development of agriculture), instead of effectiveness which can be defined as producing a desired result. To compare multiple aspects of user-friendliness, we divided the aspect user-friendliness into four aspects: (1) understanding the tool, (2) working with the tool, (3) usability for the farmer and (4) time requirement. The time and effort to get to know and learn to work with the tool as an assessor (first author) is included as an aspect of the user-friendliness of the tool. Altogether, the normative, systemic and procedural aspects provide a comprehensive insight in the tools.

Table 1
General characteristics of the tools that complied with the six selection criteria.

Tool	Full name	Target group	Publication	Origin	Start	Used version	No. assess.	Countries
RISE	Response Inducing Sustainability Evaluation	Farmers	Häni et al. (2003)	Switzerland (Bern University of Applied Sciences	1999	2.0 from 2011	>2300	>51
SAFA	Sustainability Assessment of Food and Agriculture Systems	Food and agricultural enterprises, organizations, governments	FAO (2013)	Multiple countries and institutes	2009	3.0 from 2013	>8600	>30
PG	Public Goods Tool	Farmers, policy-makers	Gerrard et al. (2012)	UK (Organic Research Centre)	2010	1.0 from 2011	>140	>9
IDEA	Indicateurs de Durabilité des Exploitations Agricoles	Farmers, policy-makers, education	Zahm et al. (2008)	France (multiple institutes)	1996	3.0 from 2008	>1500	>5

Table 2
Framework to compare sustainability assessment tools by Marchand et al. (2014), adapted from Binder et al. (2010).

	Characteristic	Description
Normative aspects		
	Sustainability concept	The concept of sustainability adopted
	Goal setting	How goals were set for the sustainability assessments
	Scoring and aggregation method	The method for indicator assessment, weighing and aggregation
	Tool function	The function, or purpose, of the tool
Systemic aspects		
	Simplicity	Is simplicity of the system representation a goal of the tool
	Sufficiency (complexity)	Is sufficiency of the system representation a goal of the tool
	Indicator interaction	Is interaction between indicators addressed in the tool
Procedural aspects		
Preparatory phase		Preparation requirements
Phase of indicator selection		Possibility of indicator selection
Measurement phase (quantification of indicators)	Data correctness	The user's perception of the correctness of the data provided
	Data availability	The availability of the required data
	User-friendliness	The user's perception of the user-friendliness of the tool
	Compatibility	The extent to which the tool is compatible with existing data systems
Assessment phase (aggregation of indicators)	Transparency	Transparency of the tool's calculations, weighing and aggregation
Applicability of assessment results and follow up	Output accuracy	The user's perception of the accuracy; the proximity of the results to the true value
	Complexity	The complexity of the tool procedures, presentation and interpretation of the results
	Communication aid	Ability to use the tool as a communication aid to discuss sustainability
	Relevance (effectiveness)	The extent to which the tool is perceived by the users as relevant to use and implement

The comparison of the normative and systemic aspects of the tools is based on information gathered from the tools' publications and manuals. The tools were applied on Danish farms to compare the procedural aspects of the tools. The experiences of the first author, who carried out the assessments, and the experiences of the farmers, collected through a questionnaire, provided the input for the comparison of the procedural aspects. These findings are described in Appendix C (Table C.3). Although this research puts emphasis on the application of the tools (procedural aspects), in-depth understanding of the tools requires a complete overview using the entire framework of Marchand et al. (2014).

2.3. Farm assessments

The manuals of the four selected tools were consulted: RISE (Grenz et al., 2012), SAFA (FAO, 2013), PG (Gerrard et al., 2012) and IDEA (Vilain, 2008). To ensure a good understanding and fluent application of the assessment tools, the assessor (first author) first tested each tool in a pilot study on a Danish farm. This pilot study indicated that data for RISE, IDEA and PG is accessible through farm data and interviews with farmers. SAFA, however, distinguishes different levels of data collecting, depending on their required quality, ranging from high (e.g. data collected for SAFA) to low (e.g. assessor's estimates, based on available farm data and farmer interview). Farm interviews and existing farm data were used as an input for SAFA to ensure a similar assessment procedure for each tool and each farm. No additional measurements (e.g. air emission analysis) or stakeholder interviews were carried out, consequently, some of the data input for SAFA is based on estimates with limited contextualization. To ensure a similar approach in SAFA, compared to the other tools, the assessment focused on the farm level and on data of 1 year. No sub-themes and indicators were excluded from the analysis. RISE involves an elaborate contextualization process in which a large set of regional data is entered in the database of the software. Also for the other tools, Danish reference values were collected to ensure a comparable approach of the tools. For the contextualization of SAFA, PG and IDEA information such as Danish wage levels, prices for agricultural products, income levels, red list species, rare livestock breeds and plant species, was gathered. Where possible, questions of PG and IDEA were specified to the Danish situation by giving examples from Denmark (e.g. Danish energy audits, agri-environmental schemes, rare breeds).

To compare the sustainability assessment tools in practice, the network of an agricultural consultancy was used to select five farmers (not the pilot farm) based on their interest in sustainability and willingness to cooperate. Selected farms involved two dairy farms and three pig farms, with 230 up to 550 hectares of land for crop production (Appendix D). The four tools selected were applied on these five Danish farms, resulting in a total of 20 assessments. Farm assessments were carried out in January and February 2015. To enable inclusion of the most recent annual financial report, assessments covered farm data of 2013. To prepare the assessments, not only this financial report, but also the fertilizer, pesticide and crop rotation plan of 2013 were obtained through the agricultural consultant.

Each assessment tool was scheduled on a different day and the order of the assessments was randomized across farms. All assessments were carried out by the same assessor (first author). Farmers' comments on the tools were registered as well. The time needed for the preparation, the assessment, and reporting was monitored. Preparation time includes time allocated to prepare the assessment in advance using farm data (including farm accounts, feeding compositions, crop rotation plan, fertilization plan) provided by the agricultural consultant and the farmer. These farm data provided answers to part of the questions, which were checked with the farmer during the interview. The assessment time includes the on-farm time in which the remaining questions were answered by the farmer during the interview. The reporting time includes time spent on filling in additional data, data quality control, calculations and developing the report. In case additional information had to be looked up in the farm management system, this data was processed after the interview. All the assessments were checked for their completeness and coherency before calculating the results and making the report.

2.4. Farmer's evaluation of the tools

Results of the assessments were discussed during a fifth meeting on-farm. The results of each tool were discussed for 20–30 min whereupon the farmer's perception of each tool was assessed using a questionnaire of ten questions (Appendix B). This procedure was chosen to prevent confusion regarding the characteristics of – and experiences with – the tools during the farmers' evaluation. The farmers were asked to respond to each question using the Likert scale scoring from 1 to 5, (1 means strongly disagree, and 5 strongly agree). The individual scores are presented in Appendix C (Table C.3). This approach was taken to allow for comparison of the responses. This questionnaire was developed following the framework of Marchand et al. (2014) and included the procedural phases the farmers were involved and had insight in: the measurement and application phase. The questionnaire focused on the procedural aspects that can play a role in the relevance of the tool as perceived by the farmer. The last five questions evaluated the perceived relevance more in detail by focusing on the contribution of sustainability assessment tools to support the farmer in the sustainable development of their farm (Appendix B). Additional comments of the farmers were collected by the first author and quoted in the paper, where relevant, to elaborate on the findings of the questionnaire. At the end of this meeting, farmers were asked to rank tools according to the relevance for their decision-making regarding sustainability.

3. Results

Results follow the structure of the framework of Marchand et al. (2014), as explained in Table 2. Detailed answers for the normative, systemic and procedural aspects can be found in Appendix C.

3.1. Normative aspects

The sustainability concept adopted by RISE, IDEA and SAFA is similar and follows the concept of sustainable development as introduced by Brundtland (1987) (Appendix C, Table C.1). The three tools provide a definition of sustainable agriculture that is economically viable, environmentally sound and socially just. The PG tool does not adopt a sustainability concept, but uses public goods (services and goods provided by agriculture) as a lens to assess farm performance. The goal setting of PG and RISE occurred through stakeholder consultation as well as through experts and literature review. For SAFA and IDEA, the goals were defined through a top-down procedure following literature and experts.

Scoring methodologies of tools differ, even though all tools apply a similar structure and 'weight-and-sum' aggregation on theme and/or sub-theme level. RISE uses a complex set of calculations to determine the score of each sub-theme. In addition, references to regional data (e.g. regional crop yields) and so-called master data (e.g. toxicity of pesticides) are included. SAFA provides an extensive manual to evaluate the performance of each indicator through several measurements. In addition, performance-, practice-, and target-based indicators are distinguished, and connected to a hierarchy in weighting. In PG and IDEA, indicator scores are derived directly from answers given by the farmer and farm data. The function of the tool is similar for all the tools studied; to assess the sustainability performance of farms. An important secondary function of the tools is to stimulate learning and dissemination of sustainability, and for PG the dissemination of public goods provided by agriculture.

3.2. Systemic aspect

The systemic aspects within an assessment tool give insight in the translation of the complexity of the system into indicators and addresses the simplicity, sufficiency and indicator interaction (Binder et al., 2010). The simplicity of system representation is mentioned only in the PG manual as an explicit goal (Appendix C, Table C.2). For the other tools, simplicity is not mentioned, but is an implicit goal, as they aim to develop an understandable tool that can contribute to awareness and education on sustainable agriculture.

Sufficiency of a tool is defined as the provision of a complete representation of the elements of an agricultural system (Binder et al., 2010). RISE, SAFA, and IDEA explicitly mention the importance of an integrative approach to farm sustainability. In PG, sufficiency is an implicit goal as it aims to embrace a wide scope of public goods delivered by agriculture. Finally, the possibility to identify interactions between indicators is not explicitly addressed in any of the tools; they all consider the indicators independently.

3.3. Procedural aspects

3.3.1. Preparatory phase

In the preparation of an assessment, SAFA requires a description of the assessment boundaries and supply chain stakeholders. In this study, the preparation time needed for SAFA was 15 min on average as a similar assessment boundary, scope and time frame were applied (Table 3). In addition, SAFA does not require direct input of quantitative farm data in the software. The preparatory phase of the other tools is focused on planning the visit and informing the farmer and advisor about the required data. Data available beforehand (e.g. data on crop production, nutrient balance, energy consumption and economics) was used to prepare the assessments. Consequently, the preparation time of the other tools was higher (Table 3).

3.3.2. Phase of indicator selection

Indicator selection is only possible in SAFA; the other tools provide a standard set of indicators.

3.3.3. Measurement phase

The correctness of the farm data and interview answers used in the tool was perceived positively by the farmers (Appendix C, Table C.3, Question 1). Nevertheless, they also stated that the qualitative questions in the tools allowed them to influence their assessment results and possibly affect the correctness. Furthermore, the assessor considered it difficult to determine the level of austerity to be applied in SAFA: *'Although the underlying data and answers were provided, the decision whether to allocate, for example, a score of 3 or 4 was in some cases difficult.'* In SAFA, the translation of qualitative information into quantitative scores is made by the assessor and puts him or her in a critical position.

Table 3
Minimum and maximum time requirements.

Min. and max. time requirements	RISE	SAFA	PG	IDEA
Preparation time	105–180 min	10–25 min	30–60 min	60–75 min
Assessment time	120–165 min	105–140 min	75–120 min	45–90 min
Calculation and reporting time	105–180 min	15 min	30–60 min	45–60 min
Total assessment time	330–525 min	125–185 min	135–240 min	150–225 min

All farmers and the assessor perceived data availability as high (Appendix C, Table C.3, Question 2). The assessor experienced that data required for RISE, PG and IDEA was relatively easy to access due to the availability of farm accounts, farm management systems and farmers' knowledge and overview of farm operations. In SAFA, however, data-availability was scored as moderate as different types of data need to be collected for the assessor to make an evaluation, and high quality data requires additional analysis and interviews with stakeholders.

Learning to work with the tools as an assessor, i.e. understand the questions and the software, requires time. Working with RISE requires training given by the tool developers to understand the intention of the questions asked, gain insight in the underlying calculations that are used to aggregate indicators into sub-themes and themes, and to get familiar with the software. Furthermore, reference data for the region, in this case Denmark, has to be put in the database. The assessor, therefore, considered that the time required learning RISE was higher compared with other tools. SAFA requires a moderate time investment, as it includes extensive manuals to be studied to understand all indicators and their way of measurement. IDEA and PG require a relative limited time investment, as manuals are more concise, and questions and calculations are relatively easy.

Four out of five farmers, as well as the assessor, indicated that RISE and IDEA provided the clearest questions, whereas SAFA includes long questions with complex wording (Appendix C, Table C.3, Question 3). During the assessments, the assessor noticed that several questions related to nature conservation in PG included unfamiliar terminology to farmers.

Time needed for on-farm assessment varied widely among the tools (Table 3). When adding the preparation, assessment and reporting time, RISE requires the largest time investment compared with the other tools (Table 3).

The assessor perceived RISE, PG and IDEA as compatible with existing data systems, since all data required for these assessments can be retrieved from the farm management system, economic accounts and the farm interview. Compared to the other tools where data needs to be entered into the tool, data for SAFA needs to be collected and is evaluated by the assessor. The compatibility of SAFA depends on the desired data quality; for high data quality, additional analysis and stakeholder interviews are required. Therefore, the assessor perceived this tool as partly compatible.

3.3.4. Assessment phase

Transparency, defined as the ease to understand how tool results are computed, was evaluated by the assessor. Transparency is relatively high in PG and IDEA due to their simple aggregation, calculation and scoring method. In SAFA, the aggregation and calculation method is transparent, whereas the scoring method is less transparent due to translation of qualitative information into scores by the assessor. Transparency of RISE is lower compared with other tools; although all calculations are provided, their complexity makes it less transparent to understand how results are computed.

3.3.5. Applicability of assessment results and follow-up

The output accuracy, defined by the proximity of the measurement result to the true value, was evaluated by the farmers and assessor. The assessor perceived the output accuracy of RISE as high due to involvement of experts and use of scientific literature in tool development, and the quantitative approach using farm data. Although the development of SAFA involved a high number of experts, the scoring procedure is rather qualitative and depends on the assessor, and, therefore, could result in less accurate outcomes. Also in IDEA and PG, a wide range of experts were involved, but the more qualitative approach of the tools may lower the accuracy. Furthermore, noticeable assumptions and value judgements on what is considered sustainable, in PG and IDEA, e.g. being organic certified contributes to the farm sustainability, negatively affected the farmers' perception of the accuracy of these tools.

Results of SAFA were considered by the farmers as too positive. One of the farmers indicated that he considered his farm's economic vulnerability as problematic, whereas SAFA presented a very positive score on this topic. Another farmer came to a similar conclusion and added: '*Perhaps that is the Danish context; the basic level is rather green already*.'

The assessor perceived the complexity of RISE and SAFA as high compared with PG and IDEA. Complexity for RISE resulted from the complex set of calculations and the high data input in RISE. For SAFA it resulted from the extensive manuals and approaches to measure each indicator. PG and IDEA provide more simple calculation methods with lower data input. The farmers evaluated the complexity of understanding the tool, and considered RISE as relatively easy to understand, and SAFA as least easy (Appendix C, Table C.3, Question 5). The assessor observed that the wording used in RISE is more recognizable for farmers compared with the more abstract wording applied in the themes of SAFA (e.g. rule of law) and IDEA (e.g. organization of space).

All tools present results of the assessment in a polygon on theme level, and tables on sub-theme or indicator level. RISE, SAFA and PG use a similar style and colour in their visual presentation, comprising a categorization from red (negative sustainability performance), orange, to green (positive sustainability performance). IDEA uses a more basic style of polygon and tables which is perceived as a less strong visual aid, as one farmer said: '*IDEA is compared to RISE less pedagogic, it is less easy to read the score due to the absence of colours as evaluation.*'

3.3.6. Relevance

Farmers recognized to a high degree the results of RISE as a reflection of the strengths and weaknesses of their farm. The results of IDEA were least recognized by all farmers (Appendix C, Table C.3, Question 6). The relevance of the tools to gain insight in the level of sustainability of the farm scored highest for RISE, followed by SAFA and PG, and lowest for IDEA (Question 7). The farmers were more skeptical about applying the tool outcomes in the decision making and management of the farm. RISE scored relatively highest followed by SAFA, and PG and IDEA (Question 8). Moreover, the farmers indicated that the tools provided limited new knowledge and insights to them (Question 9). The farmers were asked whether they saw aspects in the farm to improve or change, based on the tool outcomes. The scores were moderate, as the farmers indicated that they felt restricted in their opportunities to improve their sustainability due to the complexity of the system they are part of (Question 10).

Overall, farmers considered RISE as the most relevant tool to get insight into their sustainability performance. One of the farmers stated: '*The results are specified in the separate sub-themes; it enables you to see more specific results.*' In the discussions with farmers about the relevance of tools, the assessor observed that farmers strongly emphasized the relevance to the Danish context. They want to know how they perform compared to their colleagues, and on themes relevant in the Danish context.

The relevance of the PG tool was scored as second best by the farmers. Nevertheless, the farmers raised the possible influence of their answers on the questions and the underlying value judgements of the tool developers. Questions on landscape, heritage and nature conservation, for example, were considered as strongly oriented on the British context. Similarly, several themes and questions in IDEA were experienced as typically French and can be due to differences in the thematic scope of the tools. Also a limited visual presentation, in IDEA, and complex language use, in SAFA, had its influence on the farmers' perception of the relevance. Moreover, certain assumptions made by the tools, like organic is more sustainable, were strongly criticized by the farmers and referred to in their evaluation of the tools' relevance.

4. Discussion

The comparison of the four sustainability assessment tools indicated differences in various characteristics of the tools including their scoring and aggregation method, time requirements, transparency, output accuracy, and complexity. In the development of sustainability assessment tools aspects as time requirement, output accuracy, and complexity need to be balanced in relation to the user and tool function (Marchand et al., 2014). Marchand et al. (2014) present two extremes within sustainability assessment tools: full sustainability assessments (FSA) and rapid sustainability assessments (RSA). RSA tools require a limited time investment, with a high transparency of the tool, yet, the output-accuracy is considered lower. FSA tools would require a higher time investment, up to weeks, with a more scientifically underpinned output, however, with a low transparency, higher complexity and accuracy. Based on the comparison, RISE could be positioned in between RSA and FSA, whereas IDEA and PG would tend more towards a RSA tool. The position of SAFA in this continuum depends on the approach taken and desired data quality. Both extremes of FSA and RSA, and the four tools situated within, however, present strengths and weaknesses (Marchand et al., 2014).

Despite a higher time investment, and lower transparency, farmers considered RISE as the most relevant tool. Compared to the other tools they perceived the results as more precise, and the outcomes as understandable and recognizable. In addition, the more subjective character of IDEA and PG, noticeable from value judgement and assumptions of the tool developers (e.g. the assumption that organic farming is more sustainable) negatively affected the perceived relevance. Results of this research show that when the value judgements embedded in a tool do not reflect those of the farmers, results may become irrelevant for the farmer and the knowledge produced is not considered useful for them (Gasparatos and Scolobig, 2012; Vatn, 2005). This stresses the importance of the selection of an appropriate tool in which a fit between the value judgements of the tool developers and the users (i.e. farmers) is found (Gasparatos, 2010; Marchand et al., 2014).

Farmers in this study raised the importance of a context-specific (e.g. Denmark-specific) approach to farm level sustainability assessments. They are interested in comparing their results to their colleagues, using a tool that is based on the regional context. A context-specific approach allows inclusion of context specific characteristics, like regional sustainability challenges and norms (Gasso et al., 2015). Context-specific approaches will provide outcomes that are more likely to fit to the context in which the farmer is operating, and as such may stimulate farmers in taking action to improve the sustainability of their farm. At the same time, a context specific approach can reduce the possibility to compare the sustainable performance of different sectors and countries and risks neglecting global sustainability issues (Mascarenhas et al., 2010). This highlights the dilemma between generic and specific approaches (Gasso et al., 2015). The specificity level of the assessment should, therefore, be aligned with the assessment purpose (Gasso et al., 2015). Consequently, different sustainability assessment tools are needed to reflect the various contexts and value judgements (de Ridder et al., 2007; Gasso et al., 2015; Schader et al., 2014).

By assessing the sustainability performance of farms, tool developers aim to support farmers in their decision making towards a more sustainable development of the farm. Sustainability assessment tools can have different functions including monitoring, policy advice and certification, as described by (Schader et al., 2014). The tools manuals used in this study mention stimulating dissemination and learning as an additional function of sustainability assessment tools. Although the assessment of sustainability performance may result in discussions regarding reference values, selection of indicators and value-loaded assumptions, the tools succeed in fulfilling their secondary aim. Through sustainability assessments, farmers are triggered to discuss a wide range of sustainability themes and reflect upon their practices. The assessment outcomes function hereby as a starting point for discussion, reflection and learning.

In the evaluation of the relevance of the tools, the farmers were rather skeptical about using the knowledge produced by the tools in the decision making and management of their farm. Moreover, the farmers experienced restricted possibilities to improve or change their sustainability performance due to the complexity of the agricultural system they are part of. This emphasizes the need for additional support to utilize and implement the knowledge developed in a sustainability assessment (Binder et al., 2010). Support from advisors in constructing a farm development plan, or discussion groups of farmers can stimulate learning and reflection on the results (Marchand et al., 2014). Constructing a farm development plan, supported by outcomes of the tools, may contribute substantially to the relevance and adoption of the tools (De Mey et al., 2011). Moreover, development plans can facilitate integration of different knowledge types (i.e. practical, theoretical, transdisciplinary) and contextual information important for farm-level decision making (Darnhofer, 2010).

From a list of 48 sustainability assessment tools, only four tools complied with our set of selection criteria. This procedure may have excluded valuable tools that, for example, focus only on one dimension of sustainability, on one agricultural sector, or were developed for one specific country. None of the tools was excluded for the language criterion. Nevertheless, the selection of the four tools allowed assessments of multiple dimensions of sustainability, in the Danish context. RISE required a large input of Danish reference data; this input of data is relatively limited for the other tools. The high degree of contextualization of RISE may have positively influenced the farmers' perception of the relevance of RISE. Besides differences in the degree of contextualization allowed in the design of a tool, differences in thematic scope may have influenced the farmers' perception of relevance. Themes selected in each tool may be considered more relevant in one context to the other and may have played a role in the perceived relevance of the tools. Additional efforts to allow contextualization of tools, reduce complex language use, and improve user-friendliness may add to the perceived relevance and adoption of tools (Gasso et al., 2015; Van Meensel et al., 2012). At the same time, the adoption of tools is not primarily dependent on farmers but also depends on policy structures, research projects, and awareness of the users. Involving other stakeholders in the development of sustainability assessment tools is often suggested to enhance dialogue and understanding on sustainable development (Bell and Morse, 2008; Schindler et al., 2015; Triste et al., 2014).

Sustainability is an evolving concept; implying that sustainability assessment tools will evolve as well. The authors recognize that some of the comments in this paper might already be addressed in upcoming versions of the tools. One of the limitations of the study entails the high investment of time required to learn the tools and carry out twenty assessments. As a consequence, the high time investment per tool allowed for a limited number of farmers to be included in this study. However, the more an assessor is using a tool, the better it is trained to use it and to work with it efficiently. This research compared four sustainability assessment tools, and applied the tools in practice to evaluate the relevance as perceived by farmers, as an example of end-user validation. In addition, the framework of Marchand et al. (2014) and Binder et al. (2010) demonstrated to be useful to compare tools and can contribute to the selection of relevant tools. Additional research efforts are needed to evaluate the design validation and output validation of the tools and discuss their ability to draw evidence based conclusions on the sustainability performance of farms. Moreover, research efforts could focus on developing approaches to use sustainability assessment tools as a starting point for developing strategic farm choices, and monitor their long term performance.

5. Conclusion

Four out of 48 identified sustainability assessment tools met the criteria to be able to assess Danish farms for their sustainability performance. The four tools (RISE, SAFA, PG and IDEA) vary widely in their scoring and aggregation method, time investment, and data requirements. The farmers perceived RISE as the most relevant tool to gain insight in the sustainability performance of their farm, because this tool is based on the input of quantitative farm data and uses a context specific approach using regional data. Other factors contributing to the perceived relevance of the four tools were user-friendliness, complexity of the tool, language use and value judgements of tool developers and users (i.e. farmers). Alignment of the value judgements embedded in tools is essential for the acceptance of the assessment outcomes and the application in improving the sustainability of the farm.

Even though RISE was considered as the most relevant tool, the farmers expressed hesitation to apply the outcomes of the four tools in their decision making and management. Furthermore, they identified limitations in the options to improve their sustainability performance. Additional efforts are needed to support farmers in using the outcomes in their decision making on the local level. The outcomes of sustainability assessment tools should therefore be considered as a starting point for discussion, reflection and learning.

Acknowledgements

This research has received funding from the European Community under the Seventh Framework Programme for the project Autograssmilk under grant agreement SME-2012-2-314879. We thank the agricultural consultancy JYSK Landbrugsrådgivning for their support in finding five farmers. We are very grateful for the willingness of the farmers to go through four different sustainability assessments and discuss the results and their perception. Thanks also to the developers of the four sustainability tools for providing assistance for using the tools, and to SEGES, in particular Anke Stubsgaard, for developing a Danish version of RISE and support during the test phase. Finally, we would like to thank the anonymous reviewers for providing valuable suggestions to improve the paper.

Appendix A. Overview of tools

No.	Tool	Full name	1. Assessment level	2. Peer reviewed	Reference	3. Economic, environmental and social	4. Sector	5. Suitable NW Europe/ multiple countries	6. Language
1	AEMBAC	European Analytical Framework for the Development of Local Agri-Environmental Programmes	Landscape	Yes	Bastian et al. (2007), Simoncini (2009)	Environmental	Universal	Yes/Yes	English
2	AESIS	Agro-Environmental Sustainability Information System	Farm	Yes	Pacini et al. (2009), Pacini et al. (2011)	Environmental	Universal	Yes/No	Unknown
3	Agro-Eco-Index		Farm	Yes	Viglizzo et al. (2006)	Environmental	Universal	No/No	Spanish
4	ANCA	Annual Nutrient Cycle Assessment	Farm	Yes	Aarts et al. (2015)	Environmental	Dairy	Yes/No	Dutch
5	APOIA-NOVO RURAL		Farm	Yes	Rodrigues et al. (2010)	Economic, Environmental, Social	Universal	No/Yes	Portuguese
6	ARBRE	Arbre de l'Exploitation Agricole Durable	Farm	No	Pervanchon (2004)	Economic, Environmental, Social	Universal	Yes/No	French
7	AUI	Agrarumweltindikatoren	Farm	No	www.blw.admin.ch	Environmental	Universal	Yes/No	German
8	Avibio	AVIculture BIOlogique	Farm, chain	Yes	Pottiez et al. (2012)	Economic, Environmental, Social	Poultry	Yes/No	Unknown
9	BROA	Biodiversity Risk and Opportunity Assessment	Landscape	No	www.batbiodiversity.org/broa	Environmental	Universal	Yes/Yes	English
10	COSA	Committee On Sustainability Assessment	Farm	Yes	COSA (2013)	Economic, Environmental, Social	Coffee and cacao	No/Yes	Multiple
11	Coteur et al. (2014)		Farm	Yes	Coteur et al. (2014)	Economic, Environmental, Social	Fruit, arable, greenhouse production	Yes/No	Dutch
12	DairySAT	Dairy Self-Assessment Tool	Farm	No	www.dairysat.com.au	Environmental	Dairy	Yes/No	English
13	Dantsis et al. (2010)		Farm, regional	Yes	Dantsis et al. (2010)	Economic, Environmental, Social	Plant production	Yes/No	Unknown
14	DIAGE	DIAgnostic Global d'Exploitation	Farm	No	www.cooperation-agricole.asso.fr/sites/saf/guide/fiches/methodes_evaluation_systeme_individuelles/diage.aspx	Environmental	Universal	Yes/No	French

No.	Tool	Full name	1. Assessment level	2. Peer reviewed	Reference	3. Economic, environmental and social	4. Sector	5. Suitable NW Europe/ multiple countries	6. Language
15	DIALECTE	DIAgnostic Liant Environnement et Contrat Territorial d'Exploitation	Farm	No	http://dialecte.solagro.org/	Environmental	Universal	Yes/No	French
16	DIALOGUE	Diagnostic agri-environnemental global d'exploitation	Farm	No	www.solagro.org/site/im_user/014plaquette_dialogue.pdf	Environmental	Universal	Yes/No	French
17	DLG	DLG – Zertifikat Nachhaltige Landwirtschaft	Farm	No	www.nachhaltige-landwirtschaft.info	Economic, Environmental, Social	Universal	Yes/No	German
18	DSI	Dairyman Sustainability Index	Farm	Yes	Elsaesser et al. (2015)	Economic, Environmental, Social	Dairy	Yes/Yes	English
19	DSR	Driving Force State Response	Regional	No	OECD (2001)	Economic, Environmental, Social	Universal	Yes/Yes	English
20	EF	Ecological Footprint	Farm, product, local, regional	Yes	Wackernagel et al. (1999)	Environmental	Universal	Yes/Yes	English
21	EMA	Environmental management for agriculture	Farm	Yes	Lewis and Bardon (1998)	Environmental	Universal	Yes/No	English
22	EP	Ecopoints	Farm	No	www.oekopunkte.at	Environmental	Universal	Yes/No	German
23	FARMSMART		Farm	Yes	Tzilivakis and Lewis (2004)	Economic, Environmental, Social	Universal	Yes/No	English
24	Field Print Calculator		Farm	No	www.fieldtomarket.org	Economic, Environmental, Social	Arable farming	Yes/No	English
25	GA	Green Accounts for Farms	Farm	No	www.landbrugsinfo.dk/miljoe/natur-og-arealforvaltning/tilskudsordninger/groenne-regnskaber	Environmental	Universal	Yes/Yes	Danish
26	IDEA	Indicateur de Durabilité des Exploitations Agricoles	Farm	Yes	Zahm et al. (2008)	Economic, Environmental, Social	Universal	Yes/Yes	French, English
27	IFSC	Illinois Farm Sustainability Calculator	Farm	No	www.ideals.illinois.edu/handle/2142/13458	Environmental, Economic	Universal	No/No	English
28	INDIGO®		Farm	Yes	Thiollet-Scholtus and Bockstaller (2014)	Environmental	Crop production; viticulture	Yes/No	French
29	ISAP	Indicator of Sustainable Agricultural Practice	Farm	Yes	Rigby et al. (2001)	Economic, Environmental, Social	Horticulture	Yes/No	English
30	KSNL	Kriteriensystem Nachhaltige Landwirtschaft	Farm	No	Breitschuh (2009)	Economic, Environmental, Social	Universal	Yes/No	German
31	LCA	Life Cycle Assessment	Product	Yes	e.g. Haas et al. (2000)	Environmental	Universal	Yes/Yes	Multiple
32	MESMIS	Framework for Assessing the Sustainability of Natural Resource Management Systems	Farm, local	Yes	López-Ridaura et al. (2002), Speelman et al. (2007)	Economic, Environmental, Social	Smallholder	No/No	English, Spanish
33	MMF	Multiscale Methodological Framework	Farm, local, regional	Yes	Lopéz-Ridaura et al. (2005)	Economic, Environmental, Social	Smallholder	No/No	Unknown
34	MOTIFS	Monitoring Tool for Integrated Farm Sustainability	Farm	Yes	Meul et al. (2008)	Economic, Environmental, Social	Dairy	Yes/No	Dutch
35	PG	Public Goods Tool	Farm	Yes	Gerrard et al. (2012)	Economic, Environmental, Social	Universal	Yes/Yes	English
36	RAD	Diagnostic de durabilité of Réseau de l'Agriculture Durable	Farm	Yes	Le Rohellec and Mouchet (2008)	Economic, Environmental, Social	Dairy	Yes/No	French
37	REPRO	Reproduction of Soil Fertility	Farm	No	www.nachhaltige-landbewirtschaftung.de/repro/	Environmental	Universal	Yes/No	German
38	RISE	Response-Inducing Sustainability Evaluation 2.0	Farm	Yes	Häni et al. (2003)	Economic, Environmental, Social	Universal	Yes/Yes	Multiple
39	SAFA	Sustainability Assessment of Food and Agriculture systems	Farm, chain	Yes	FAO (2013)	Economic, Environmental, Social	Universal	Yes/Yes	English

No.	Tool	Full name	1. Assessment level	2. Peer reviewed	Reference	3. Economic, environmental and social	4. Sector	5. Suitable NW Europe/ multiple countries	6. Language
40	SAFE	Sustainability Assessment of Farming and the Environment	Farm, landscape, regional	Yes	Van Cauwenbergh et al. (2007)	Economic, Environmental, Social	Universal	Yes/No	English
41	SAI – SPA	Farmer Self Assessment 2.0	Farm	No	www.standardsmap.org/fsa	Economic, Environmental, Social	Universal	Yes/Yes	English, Spanish, French
42	SALCA	Swiss Agricultural Life Cycle Assessment	Farm, product, system	Yes	Nemecek et al. (2011)	Environmental	Universal	Yes/No	Unknown
43	SeeBalance®		Product	Yes	Saling et al. (2005)	Economic, Environmental, Social	Universal	Yes/Yes	English
44	SLCA	Social Life Cycle Assessment	Product	Yes	Benoît and Mazijn (2009)	Social	Universal	Yes/Yes	English
45	SMART	Sustainability Monitoring and Assessment RouTine	Farm	No	www.fibl.org/en/themes/smart-en.html	Economic, Environmental, Social	Universal	Yes/Yes	Unknown
46	Soil&More Flower	Sustainability Flower Quick Assessment	Farm	No	www.soilandmorefoundation.org/projects/sustainability-flower	Economic, Environmental, Social	Universal	Yes/Yes	English
47	Sustainability Dashboard		Farm	No	www.triplehelix.com.au/documents/FarmSustainabilityDashboard.pdf	Economic, Environmental, Social	Universal	Yes/No	English
48	van Calker et al. (2006)		Farm	Yes	van Calker et al. (2006)	Economic, Environmental, Social	Dairy	Yes/No	Unknown

Appendix B. Questionnaire for the farmers to evaluate the tools

Phase	Characteristic	Question
Measurement phase	Data correctness	1. Do you consider the data you've provided for the assessment correct?
	Data availability	2. Was the data easily accessible?
	User-friendliness	3. Does the tool present the questions in an easy and understandable way?
Applicability of assessment results and follow up	Output accuracy	4. Do you consider the results of the tool precise?
	Complexity	5. Does the tool present the outcomes in an easy and understandable way?
	Relevance	6. Do you recognize the results of the tool?
		7. Do you consider this as a relevant tool to gain insight in the level of sustainability of your farm
		8. Can you apply the tool outcomes in the decision making and management of the farm?
		9. Does the tool provide new knowledge and insights to you?
		10. Based on the tool outcomes, do you see aspects in the farm to improve or change?

Appendix C. Comparison of sustainability assessment tools

Table C.1
Normative aspects of sustainability assessment tools.

Characteristic	RISE	SAFA	PG	IDEA
Sustainability concept	Sustainable development (Brundtland); sustainable development in agriculture is environmentally non-degrading, technically appropriate, economically viable and socially acceptable) (FAO, 1989)	Sustainable development in agriculture is environmentally non-degrading, technically appropriate, economically viable and socially acceptable (FAO, 1989)	Farming provides public goods beyond production of food only (Cooper et al., 2009)	Sustainable development (Brundtland); sustainable agriculture is agriculture that is ecologically sound, economically viable, socially just and humane (Francis et al., 1990).
Goal setting	Stakeholder involvement as well as top-down	Top-down approach	Stakeholder involvement as well as top-down	Top-down approach
Scoring and aggregation method	RISE covers 10 themes divided into 50 sub-themes. The scores of the sub-themes ranges between 0 and 100 and is based on an aggregation of indicators. The online software calculates the scores based on a farm interview, data from farm accounts and references to regional and master data.	SAFA includes 21 themes, 58 sub-themes and 116 indicators. The score of each indicator is evaluated on a scale from 1 to 5. To determine this score as an assessor, SAFA indicates ways to measure the indicator. The scores of the indicators are aggregated to the sub-theme and theme level. SAFA distinguishes performance-, practice- and target-based indicators with differences in weighting.	PG covers 11 themes with 33 sub-themes. The scores of the sub-themes are calculated using various indicators scored between 1 and 5. The scores are derived through a farm interview and include a few comparisons to regional averages.	IDEA includes 10 themes divided into 42 sub-themes. A farm interview forms the basis for this tool. For each indicator a certain amount of points can be obtained. In addition, IDEA has defined a maximum value for each sub-theme and theme in order to limit the possibility to compensate for low scores.
Tool function	Holistic assessment of the sustainability of agricultural production at farm level, and stimulate discussion.	Provide a holistic assessment of food and agricultural systems on four domains of sustainability.	Assess public goods provided by a farm, and provide a learning tool.	Provide a tool that could assess and raise awareness for sustainability on-farm.

Table C.2
Systemic aspects of sustainability assessment tools.

Characteristic	RISE	SAFA	PG	IDEA
Simplicity	Implicit goal, as it aims to support farmers and advisors in sustainable development	Implicit goal, as it aims to develop a shared and understandable way to communicate about sustainability in agriculture	Explicit goal mentioned in Gerrard et al. (2012)	Implicit goal from the commissioner as it should be accessible to many.
Sufficiency (complexity)	Explicit goal as the tool aims to provide a holistic approach	Explicit goal to address the complexity and relationships of all dimensions of sustainability	Implicit goal as the tool aims to cover a wide range of public goods	Explicit goal as the tool aims to provide a holistic approach
Indicator interaction	Not explicit	Not explicit	Not explicit	Not explicit

Table C.3
Procedural aspects of sustainability assessment tools. The comparison of the procedural aspects is based on the experiences of the assessor and the results of the questionnaire. The results of the questionnaire are presented in Likert scores given by the farmers for each question (farmer 1-2-3-4-5).

Phase	Characteristic	RISE	SAFA	PG	IDEA
Preparatory phase		Plan farm visit and discuss required data, start filling in farm data in software	Plan visit, describe the assessed entities, boundaries and supply chain	Plan farm visit	Plan farm visit and prepare necessary documentation
	Preparation time	02:10	00:15	00:50	01:10
Phase of indicator selection		Not included: a standard set of indicators is used	Included: contextualization of the tool involves listing relevant sub-themes and indicators	Not included: a standard set of indicators is used	Not included: a standard set of indicators is used
Measurement phase	Data correctness	High, due to data input from accountancy and management system	Moderate, due to influence of farmers' perception and the austerity of the assessor	Moderate, due to influence of farmers' perception	Moderate, due to influence of farmers' perception
	1. Do you consider the data you've provided for the assessment correct?	4-4-4-5-4	5-4-4-5-4	4-4-4-5-3	4-5-4-5-4
	Data availability	Good, data is available from farm accounts, management system and through interviews	Moderate, depending on the desired level of data quality. Part of the data can be retrieved from farm accounts, analysis and through interviews. Yet, high quality data requires additional analyses.	Good, data is available from farm accounts, management system and through interviews	Good, data is available from farm accounts, management system and through interviews
	2. Is the data easily accessible?	5-4-4-5-4	5-4-4-5-4	4-4-4-5-3	5-4-4-5-4
	User-friendliness				
	Understanding the tool	High time investment due to required training, and time needed to get familiar with the questions and to understand the calculations. In addition, regional reference data has to be put in the database.	Moderate time investment due comprehensive assessment guidelines and time needed to get familiar with the questions	Limited time investment to understand the tool.	Limited time investment to understand the tool and manual
	Working with the tool	The tool is supported with user-friendly, online software available in multiple languages. The software combines the regional data, master data and farm data to compute the scores.	Supported with free, simple software. The software function is limited as it only includes the rating and not additional information collected in the assessment.	Simple, using Excel. Yet, some tables are very large, this reduces the user-friendliness.	Simple, using the score forms provided by the tool. Either in Excel or on paper. Some tables are, however, only in French.
	3. Does the tool present the questions in an easy and understandable way?	5-4-4-3-4	3-2-4-4-4	3-4-3-4-4	5-4-4-4-4
	Assessment time	02:20	02:00	01:30	01:10
	Calculation and reporting time	02:30	00:15	00:40	00:50
	Total assessment time	07:00	02:30	03:00	03:10
	Compatibility	Good compatible with farm accounts	Partly compatible, some questions require other types of data, dependent on desired data quality	Good compatible with farm accounts	Good compatible with farm accounts

Table C.3 (*Continued*)

Phase	Characteristic	RISE	SAFA	PG	IDEA
Assessment phase	Transparency	Moderate transparency; although insight in underlying calculations is given, the its complexity reduces the transparency	Moderate transparency; the scores are the assessor's interpretation of the measurements	High transparency in scoring method and documentation	High transparency in scoring method and documentation
Applicability of assessment results and follow up	Output accuracy	High, due to input of quantitative farm data, and involvement of experts and scientific reference in the development.	Moderate, although experts and scientific literature were involved in the development of the tool, a more qualitative approach gives more influence to the farmer and assessor.	Moderate, although experts and scientific literature were involved in the development of the tool, a more qualitative approach gives more influence to the farmer and assessor.	Moderate, although experts and scientific literature were involved in the development of the tool, a more qualitative approach gives more influence to the farmer and assessor.
	4. Do you consider the results of the tool precise?	4-4-4-5-4	2-3-4-5-4	4-4-3-4-3	2-4-3-4-4
	Complexity	High complexity due to complex calculations, high data requirements and time investment.	High complexity due to extensive guidelines on collecting information for each indicator	Limited complexity resulting from relative easy format and limited data requirements	Limited complexity resulting from relative easy format and limited data requirements
	5. Does the tool present the outcomes in an easy and understandable way?	5-4-4-4-4	2-2-3-4-4	3-4-3-4-3	3-3-4-4-3
	Communication aid	High. The results are shown in a polygon and tables with colours (green to red). The division into sub-themes tables allows in-depth discussion of the results	Moderate. The results are presented in a coloured polygon and tables. The tool is visually strong but the wording is sometimes abstract.	High. The results are shown in Excel using a coloured polygon on theme level and a chart on sub-theme level.	Moderate. The results are shown in multiple graphs including a polygon on theme level. The tables and polygon are less clear compared to the other tools due to missing colours and some unknown terminology for farmers.
	Relevance				
	6. Do you recognize the results of the tool?	4-4-4-5-4	4-2-4-5-4	4-3-4-5-3	3-3-3-3-3
	7. Do you consider this as a relevant tool to gain insight in the level of sustainability of your farm?	4-4-3-5-5	3-2-3-4-4	3-3-3-4-3	2-3-2-4-3
	8. Can you apply the tool outcomes in the decision making and management of the farm?	3-3-3-5-4	2-2-3-4-3	2-2-3-3-3	2-2-3-3-2
	9. Does the tool provide new knowledge and insights to you?	3-3-3-3-4	2-2-3-3-4	2-2-3-3-4	3-2-3-3-3
	10. Based on the tool outcomes, do you see aspects in the farm to improve or change?	4-3-3-3-4	2-2-3-3-3	2-2-3-3-4	2-2-3-3-2
	11. Ranking tools	1-1-1-1-1	3-4-3-4-2	2-2-2-2-3	4-3-4-3-4

Appendix D. Characteristics of the farms

	Hectares	Dairy cows	Sows	Slaughter pigs	Veal calves
Farm 1	230	370	–	–	–
Farm 2	250	350	–	–	–
Farm 3	325	–	550	5000	–
Farm 4	550	–	–	30,000	1500
Farm 5	250	–	–	30,000	–

References

Aarts, H., De Haan, M., Schröder, J., Holster, H., De Boer, J., Reijs, J., Oenema, J., Hilhorst, G., Sebek, L., Verhoeven, F., 2015. Quantifying the Environmental Performance of Individual Dairy Farms – The Annual Nutrient Cycling Assessment. European Grassland Federation, Wageningen, the Netherlands.

Acosta-Alba, I., van der Werf, H.M.G., 2011. The use of reference values in indicator-based methods for the environmental assessment of agricultural systems. Sustainability 3, 424–442.

Bastian, O., Corti, C., Lebboroni, M., 2007. Determining environmental minimum requirements for functions provided by agro-ecosystems. Agron. Sustain. Dev. 27, 279–291.

Bausch, J.C., Bojórquez-Tapia, L., Eakin, H., 2014. Agro-environmental sustainability assessment using multicriteria decision analysis and system analysis. Sustain. Sci., 1–17.

Bélanger, V., Vanasse, A., Parent, D., Allard, G., Pellerin, D., 2012. Development of agri-environmental indicators to assess dairy farm sustainability in Quebec, Eastern Canada. Ecol. Indic. 23, 421–430.

Bell, S., Morse, S., 2008. Sustainability indicators: measuring the immeasurable? Earthscan.

Benoît, C., Mazijn, B., 2009. Guidelines for Social Life Cycle Assessment of Products. UNEP, Paris, France.

Binder, C.R., Feola, G., Steinberger, J.K., 2010. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. Environ. Impact Assess. Rev. 30, 71–81.

Bockstaller, C., Girardin, P., 2003. How to validate environmental indicators. Agric. Syst. 76, 639–653.

Bos, A.P., Groot Koerkamp, P.W.G., Gosselink, J.M.J., Bokma, S., 2009. Reflexive interactive design and its application in a project on sustainable dairy husbandry systems. Outlook Agric. 38, 137–145.

Breitschuh, T., 2009. Folgenabschätzung einer zunehmenden Bereitstellung von Bioenergieträgern auf die Nachhaltigkeit landwirtschaftlicher Unternehmen, bewertet mit dem Kriteriensystem Nachhaltige Landwirtschaft (KSNL). Umweltbundesamt.

Brundtland, G.H., 1987. Report of the World Commission on Environment and Development: Our Common Future. United Nations, Oslo.

COSA., 2013. The COSA Measuring Sustainability Report: Coffee and Cocoa in 12 Countries. The Committee on Sustainability Assessment, Philadelphia, PA.

Cooper, T., Hart, K., Baldock, D., 2009. The provision of public goods through agriculture in the European Union. Report prepared for DG Agriculture and Rural Development. Institute for European Environmental Policy, London, Contract No 30-CE-023309/00-28.

Coteur, I., Marchand, F., Debruyne, L., Bijttebier, J., Triste, L., Lauwers, L.,2014. Development and evaluation of an on-demand sustainability tool in Flanders. In: 11th European IFSA Symposium. IFSA Europe, Berlin, Germany.

Dantsis, T., Douma, C., Giourga, C., Loumou, A., Polychronaki, E.A., 2010. A methodological approach to assess and compare the sustainability level of agricultural plant production systems. Ecol. Indic. 10, 256–263.

Darnhofer, I., 2010. Strategies of family farms to strengthen their resilience. EPG 20, 212–222.

De Boer, I.J.M., Cornelissen, A.M.G., 2002. A method using sustainability indicators to compare conventional and animal-friendly egg production systems. Poultry Sci. 81, 173–181.

De Mey, K., D'Haene, K., Marchand, F., Meul, M., Lauwers, L., 2011. Learning through stakeholder involvement in the implementation of MOTIFS: an integrated assessment model for sustainable farming in Flanders. Int. J. Agric. Sustain. 9, 350–363.

de Ridder, W., Turnpenny, J., Nilsson, M., Von Raggamby, A., 2007. A framework for tool selection and use in integrated assessment for sustainable development. J. Environ. Assess. Policy Manage. 9, 423–441.

Elsaesser, M., Jilg, T., Herrmann, K., Boonen, J., Debruyne, L., Laidlaw, A.S., Aarts, F., 2015. Quantifying sustainability of dairy farms with the DAIRYMAN sustainability-index, European Grassland Federation. In: Grassland Science in Europe, Wageningen, the Netherlands, pp. 367–376.

FAO, 1989. Sustainable development and natural resources management. In: Conference, T.-F. (Ed.), Paper C 89/2 - Sup. 2. Food and Agriculture Organization, Rome.

FAO, 2013. Sustainability Assessment of Food and Agriculture Systems (SAFA): Guidelines, Version 3.0. Food and Agricultural Organization of the United Nations.

Finkbeiner, M., Schau, E.M., Lehmann, A., Traverso, M., 2010. Towards life cycle sustainability assessment. Sustainability 2, 3309–3322.

Francis, C.A., Flora, C.B., King, L.D., 1990. Sustainable agriculture in temperate zones. John Wiley & Sons, New York and Chichester.

Galan, M.B., Peschard, D., Boizard, H., 2007. ISO 14 001 at the farm level: analysis of five methods for evaluating the environmental impact of agricultural practices. J. Environ. Manage. 82, 341–352.

Gasparatos, A., 2010. Embedded value systems in sustainability assessment tools and their implications. J. Environ. Manage. 91, 1613–1622.

Gasparatos, A., Scolobig, A., 2012. Choosing the most appropriate sustainability assessment tool. Ecol. Econ. 80, 1–7.

Gasso, V., Oudshoorn, F.W., de Olde, E., Sørensen, C.A.G., 2015. Generic sustainability assessment themes and the role of context: the case of Danish maize for German biogas. Ecol. Indic. 49, 143–153.

Gerrard, C., Smith, L.G., Pearce, B., Padel, S., Hitchings, R., Measures, M., 2012. Public Goods and Farming, Farming for Food and Water Security, 10. Sustainable Agriculture Reviews, No. 8380. Springer, Dordrecht/Heidelberg/New York/London, pp. 1–22.

Grenz, J., Schoch, M., Stämpfli, A., Thalmann, C., 2012. RISE 2.0 Field Manual. School of Agricultural, Forest and Food Sciences (Bern University of Applied Sciences), Zollikofen, Switzerland.

Guerci, M., Knudsen, M.T., Bava, L., Zucali, M., Schönbach, P., Kristensen, T., 2013. Parameters affecting the environmental impact of a range of dairy farming systems in Denmark, Germany and Italy. J. Clean. Prod. 54, 133–141.

Haas, G., Wetterich, F., Geier, U., 2000. Life cycle assessment framework in agriculture on the farm level. Int. J. Life Cycle Assess. 5, 345–348.

Halberg, N., Van Der Werf, H.M.G., Basset-Mens, C., Dalgaard, R., De Boer, I.J.M., 2005. Environmental assessment tools for the evaluation and improvement of European livestock production systems. Livest. Prod. Sci. 96, 33–50.

Häni, F., Braga, F., Stämpfli, A., Keller, T., Fischer, M., Porsche, H., 2003. RISE, a tool for holistic sustainability assessment at the farm level. Int. Food Agribus. Manage. Rev. 6, 78–90.

Hjørland, B., Christensen, F.S., 2002. Work tasks and socio-cognitive relevance: A specific example. J. Am. Soc. Inf. Sci. Tec. 53, 960–965.

Le Rohellec, C., Mouchet, C., 2008. Efficacité économique de systemes laitiers herbagers en agriculture durable (RAD): une comparaison avec le RICA. Fourrages 193, 107–113.

Leach, K., Gerrard, C.L., Padel, S., 2013. Rapid Sustainability Assessment of Organic and Low-input Farming Across Europe and Identification of Research Needs Deliverable 1.1, No. EU-SOLID Project. ORC, Hamstead Marshall, Newbury.

Lebacq, T., Baret, P.V., Stilmant, D., 2013. Sustainability indicators for livestock farming. A review. Agron. Sustain. Dev. 33, 311–327.

Le Gal, P.Y., Dugué, P., Faure, G., Novak, S., 2011. How does research address the design of innovative agricultural production systems at the farm level? A review. Agr. Syst. 104, 714–728.

Lewis, K., Bardon, K., 1998. A computer-based informal environmental management system for agriculture. Environ. Model. Softw. 13, 123–137.

López-Ridaura, S., Masera, O., Astier, M., 2002. Evaluating the sustainability of complex socio-environmental systems. The MESMIS framework. Ecol. Indic. 2, 135–148.

Lopéz-Ridaura, S., Keulen, H.V., Ittersum M.K.v., Leffelaar, P.A., 2005. Multiscale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. Environ. Dev. Sustain. 7, 51–69.

Lynch, T., Gregor, S., Midmore, D., 2000. Intelligent support systems in agriculture: how can we do better? Anim. Prod. Sci. 40, 609–620.

Marchand, F., Debruyne, L., Lauwers, L., 2012. A Comparison of Complex Expert-based Assessment versus Quickscan Assessment. IFSA, Aarhus, Denmark.

Marchand, F., Debruyne, L., Triste, L., Gerrard, C., Padel, S., Lauwers, L., 2014. Key characteristics for tool choice in indicator-based sustainability assessment at farm level. Ecol. Soc., 19.

Marta-Costa, A.A., Silva, E., 2013. Approaches for sustainable farming systems assessment. In: Marta-Costa, A.A., Soares da Silva, E.L.D.G. (Eds.), Methods and Procedures for Building Sustainable Farming Systems. Springer, pp. 21–29.

Mascarenhas, A., Coelho, P., Subtil, E., Ramos, T.B., 2010. The role of common local indicators in regional sustainability assessment. Ecol. Indic. 10, 646–656.

McCown, R.L., 2002. Changing systems for supporting farmers' decisions: problems, paradigms, and prospects. Agric. Syst. 74, 179–220.

Meul, M., Passel, S.V., Nevens, F., Dessein, J., Rogge, E., Mulier, A., Hauwermeiren, A.V., 2008. MOTIFS: a monitoring tool for integrated farm sustainability. Agron. Sustain. Dev. 28, 321–332.

Nemecek, T., Dubois, D., Huguenin-Elie, O., Gaillard, G., 2011. Life cycle assessment of Swiss farming systems: I. Integrated and organic farming. Agric. Syst. 104, 217–232.

OECD, 2001. Environmental Indicators for Agriculture. Methods and Results. Organisation for Economic Co-operation and Development, Paris.

Pacini, C., Lazzerini, G., Migliorini, P., Vazzana, C., 2009. An indicator-based framework to evaluate sustainability of farming systems: review of applications in Tuscany. Ital. J. Agron. 4, 23–40.

Pacini, G.C., Lazzerini, G., Vazzana, C., 2011. AESIS: a support tool for the evaluation of sustainability of agroecosystems. Example of applications to organic and integrated farming systems in Tuscany, Italy. Ital. J. Agron. 6, e3.

Pervanchon, F., 2004. L'arbre de l'exploitation agricole durable. Construire en groupe son projet d'agriculture durable. Travaux et Innovations, Paris, France, pp. 5–8.

Pottiez, E., Lescoat, P., Bouvarel, I., 2012. AVIBIO: a method to assess the sustainability of the organic poultry industry. In: 10th European International Farming Systems Association (IFSA) Symposium, IFSA Europe Group, Aarhus, Denmark.

Pretty, J., 2008. Agricultural sustainability: concepts, principles and evidence. Philos. Trans. R. Soc. B: Biol. Sci. 363, 447–465.

Rigby, D., Woodhouse, P., Young, T., Burton, M., 2001. Constructing a farm level indicator of sustainable agricultural practice. Ecol. Econ. 39, 463–478.

Rodrigues, G.S., Rodrigues, I.A., Buschinelli C.C.D.A., de Barros, I., 2010. Integrated farm sustainability assessment for the environmental management of rural activities. Environ. Impact Assess. Rev. 30, 229–239.

Saling, P., Maisch, R., Silvani, M., König, N., 2005. Assessing the environmental-hazard potential for life cycle assessment, eco-efficiency and SEEbalance. Int. J. Life Cycle Assess. 10, 364–371.

Schader, C., Grenz, J., Meier, M.S., Stolze, M., 2014. Scope and precision of sustainability assessment approaches to food systems. Ecol. Soc. 19.

Schindler, J., Graef, F., König, H., 2015. Methods to assess farming sustainability in developing countries. A review. Agron. Sustain. Dev. 35, 1043–1057.

Simoncini, R., 2009. Developing an integrated approach to enhance the delivering of environmental goods and services by agro-ecosystems. Reg. Environ. Change 9, 153–167.
Speelman, E.N., López-Ridaura, S., Colomer, N.A., Astier, M., Masera, O.R., 2007. Ten years of sustainability evaluation using the MESMIS framework: lessons learned from its application in 28 Latin American case studies. Int. J. Sustain. Dev. World Ecol. 14, 345–361.
Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., Haan, C.d., 2006. Livestock's long shadow: environmental issues and options. In: Food and Agriculture Organization of the United Nations (FAO).
Thiollet-Scholtus, M., Bockstaller, C., 2014. Using indicators to assess the environmental impacts of wine growing activity: The INDIGO® method. Eur. J. Agron. 62, 13–25.
Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. Nature 418, 671–677.
Triste, L., Marchand, F., Debruyne, L., Meul, M., Lauwers, L., 2014. Reflection on the development process of a sustainability assessment tool: learning from a Flemish case. Ecol. Soc., 19.
Tzilivakis, J., Lewis, K.A., 2004. The development and use of farm-level indicators in England. Sustain. Dev. 12, 107–120.
van Calker, K.J., Berentsen, P.B.M., de Boer, I.J.M., Giesen, G.W.J., Huirne, R.B.M., 2007. Modelling worker physical health and societal sustainability at farm level: An application to conventional and organic dairy farming. Agr. Syst. 94, 205–219.
van Calker, K.J., Berentsen, P.B.M., Romero, C., Giesen, G.W.J., Huirne, R.B.M., 2006. Development and application of a multi-attribute sustainability function for Dutch dairy farming systems. Ecol. Econ. 57, 640–658.
Van Cauwenbergh, N., Biala, K., Bielders, C., Brouckaert, V., Franchois, L., Garcia Cidad, V., Hermy, M., Mathijs, E., Muys, B., Reijnders, J., Sauvenier, X., Valckx, J., Vanclooster, M., Van der Veken, B., Wauters, E., Peeters, A., 2007. SAFE-A hierarchical framework for assessing the sustainability of agricultural systems. Agric. Ecosyst. Environ. 120, 229–242.
van der Werf, H.M.G., Tzilivakis, J., Lewis, K., Basset-Mens, C., 2007. Environmental impacts of farm scenarios according to five assessment methods. Agric. Ecosyst. Environ. 118, 327–338.
Van Meensel, J., Lauwers, L., Kempen, I., Dessein, J., Van Huylenbroeck, G., 2012. Effect of a participatory approach on the successful development of agricultural decision support systems: the case of Pigs2win. Decis. Support Syst. 54, 164–172.
Van Passel, S., Meul, M., 2012. Multilevel and multi-user sustainability assessment of farming systems. Environ. Impact Assess. Rev. 32, 170–180.
Vatn, A., 2005. Rationality, institutions and environmental policy. Ecol. Econ. 55, 203–217.
Viglizzo, E.F., Frank, F., Bernardos, J., Buschiazzo, D.E., Cabo, S., 2006. A rapid method for assessing the environmental performance of commercial farms in the Pampas of Argentina. Environ. Monit. Assess. 117, 109–134.
Vilain, L., 2008. La méthode IDEA: indicateurs de durabilité des exploitations agricoles. Educagri éditions.
Wackernagel, M., Onisto, L., Bello, P., Linares, A.C., Falfán, I.S.L., García, J.M., Guerrero, A.I.S., Guerrero, M.G.S., 1999. National natural capital accounting with the ecological footprint concept. Ecol. Econ. 29, 375–390.
Wiskerke, J.S.C., 2009. On places lost and places regained: reflections on the alternative food geography and sustainable regional development. Int. Plan. Stud. 14, 369–387.
Zahm, F., Viaux, P., Vilain, L., Girardin, P., Mouchet, C., 2008. Assessing farm sustainability with the IDEA method – from the concept of agriculture sustainability to case studies on farms. Sustain. Dev. 16, 271–281.