OCL
Oilseeds & fats Crops and Lipids

Available online at:
www.ocl-journal.org

RESEARCH ARTICLE – DOSSIER   OPEN ACCESS

DOSSIER

CRITÈRES DE DURABILITÉ
SUSTAINABILITY CRITERIA

Issues in evaluating sustainability of farming systems with indicators

Christian Bockstaller[1,2,*], Pauline Feschet[1,2] and Frédérique Angevin[3]

[1] INRA, UMR 1121 Agronomie et Environnement, INRA-Université de Lorraine, BP20507, 68021 Colmar Cedex, France
[2] Université de Lorraine, UMR 1121, Agronomie et Environnement, BP20507, 68021 Colmar Cedex, France
[3] INRA, UAR 1240 Eco-Innov, avenue Lucien Brétignières, 78850 Thiverval-Grignon, France

Abstract – The growing concern about side-effects of policies focusing on economic growth or even technological innovations, as well as agriculture intensification leads more and more stakeholders to pay attention to the questions of monitoring and evaluation of agricultural practices. This step of evaluation is now essential in policy decision, in research and design of innovative solutions, in NGOs' development projects, as well as in improvement process in ISO certification. The aim of this article is to review steps in the evaluation of sustainability in agriculture, starting in a first section with the necessity to develop a conceptual indicator framework to precise evaluators' own vision of sustainability. In a second section, we address the necessity to answer preliminary questions that will guide the selection of a set of indicators or an assessment method. In a third section, after discussing the way to categorize indicators, we provide an overview of available indicators for two sustainability themes of the environmental dimension regarding respectively nitrogen management and biodiversity. In a fourth section, we highlight the diversity of evaluation methods of sustainability through six examples in France. Finally we conclude the article with a general discussion on questions that remain to address.

Keywords: Multi-criteria assessment / framework / aggregation / nitrogen management / biodiversity

Résumé – Choix d'un ensemble d'indicateurs pour évaluer la durabilité des pratiques agricoles. Le souci croissant au sujet des effets secondaires de politiques centrées sur la croissance économique ou des innovations technologiques mais aussi de l'intensification de l'agriculture conduit de plus en plus d'acteurs à s'intéresser à la question de l'évaluation et le suivi des pratiques agricoles. Cette étape de l'évaluation est devenue primordiale dans la décision publique, dans la recherche de solutions innovantes, dans des projets développement d'ONG ou encore dans les démarches de progrès en certification ISO. L'objectif de cet article est de passer en revue les étapes d'une démarche d'évaluation de la durabilité en agriculture, en partant de la nécessité de développer un cadre conceptuel des indicateurs afin de préciser la vision de la durabilité de l'évaluateur. Dans une seconde partie, nous mettons en évidence la nécessité de répondre à des questions préliminaires qui vont guider le choix d'un ensemble d'indicateurs et/ou d'une méthode d'évaluation. Dans une troisième, après avoir discuté de la possibilité d'établir une typologie d'indicateurs, nous présentons une vue d'ensemble des indicateurs disponibles pour deux thématiques de la dimension environnementale, la gestion de l'azote et celle de la biodiversité. Dans une quatrième, nous illustrons la diversité des méthodes d'évaluation de la durabilité en agriculture au travers de six exemples français. Enfin, nous terminons l'article par une discussion sur les questions à approfondir.

Mots clés : Évaluation multicritère / cadre conceptuel / agrégation / gestion de l'azote / biodiversité

* Correspondence: `christian.bockstaller@colmar.inra.fr`

1 Introduction

Following the Rio conference in 1992, the environmental issue and, more generally, the question of sustainability became a concern in the developed countries and at the planet level. Though its ability to federate, this concept failed to meet a consensus on its implementation until now (Robinson, 2004) so that Lacousmes (2005) spoke about a "driving illusion". However, the growing concern about side-effects of policies focusing on economic growth or technological innovations, as well as agriculture intensification led more and more stakeholders to pay attention to the questions of monitoring and evaluation. This step of evaluation has become now essential in policy decision, in research and design of innovative solutions, in the NGOs' development projects, as well as in improvement process in ISO certification (López-Ridaura *et al.*, 2005; Niemeijer and de Groot, 2008). Thus, there is a general agreement on the need of developing sustainability indicators that have to be organized in a conceptual framework to form an evaluation method. The use of indicators can be easily explained by the impossibility to measure directly environmental impacts in routine outside of research context, or by difficulties when addressing complex systems or concepts such as biodiversity and sustainability (Gras *et al.*, 1989; Maurizi and Verrel, 2002). This has fostered a great development of studies on indicators, especially in the agricultural sector (Riley, 2001; Rosnoblet *et al.*, 2006).

The aim of this article is to review steps in the evaluation of sustainability in agriculture, starting in Section 1 with the necessity to develop a conceptual indicator framework to precise evaluators' own vision of sustainability. In Section 2, we address the necessity to answer preliminary questions that will guide the selection of a set of indicators or an assessment method. In Section 3, after discussing the way to categorize indicators, we provide an overview of available indicators for two sustainability themes of the environmental dimension regarding respectively nitrogen management and biodiversity. In Section 4, we highlight the diversity of evaluation methods of sustainability through six examples in France. Finally we conclude the article with a general discussion on questions that remain to address.

2 Different sustainability frameworks

Sustainability is by nature a multidimensional issue which addresses hence a set of criteria which can be simply organized in a list or in a more complex framework (Ledoux *et al.*, 2005). In any case, a general conceptual indicator framework is a prerequisite to any indicator selection to avoid a unconsidered and even biased assessment of sustainability (Alkan Olsson *et al.*, 2009). Hansen (1996) distinguished between sustainability as an approach of agriculture and sustainability as a property of agriculture. He separated the former between (i) an alternative ideology and (ii) a set of strategies, and the latter between (iii) an ability to fulfill goals and (iv) an ability to continue. If the first approach based on an ideology remains general and vague, the three others have been translated in operational principles that inspired different evaluation frameworks (Smith and McDonald, 1998). Defining sustainability as a set of strategies

or practices led to numerous assessment methods implementing a scoring system of farmers' practices like the IDEA methods for the environmental pillar of sustainability (Zahm *et al.*, 2008) or the indicator set of Rigby *et al.* (2001). The goal-based approach of sustainability is also very common and encompasses a framework based on a set of general goals often divided in more operational goals (Bockstaller *et al.*, 1997). An example of an environmental general goal can be "preserving water quality" that can be translated into several operational goals like "reducing nitrate leaching" or "decreasing pesticide transfer to ground or surface water". In some cases, the goals can be quantified (*e.g.* "reducing the nitrate leaching by $X\%$"). In many other systems, goals are expressed more vaguely in form of themes and sub-themes (*e.g.* Alkan Olsson *et al.*, 2009; Ledoux *et al.*, 2005). In Life-Cycle Analysis focusing mainly on the environmental dimension, goals refer to environmental impacts derived from the cause-effect chain (Payraudeau and van der Werf, 2005). More recently, Life-Cycle Analysis was applied to the social dimension (Falque *et al.*, 2013; Feschet, 2014). The last approach of sustainability referring to a propriety of agriculture to continue was extended by several authors to a set of systemic properties. Bossel (2000) proposed six basic axes linked to systemic properties: existence, effectiveness, freedom of action, security, adaptability, and coexistence across totally different systems like cultural and social systems, ecosystems. López-Ridaura *et al.* (2005) collected a long list of attributes or properties of sustainability to develop an assessment method of sustainability of small peasant farms in Mexico. Finally he focused on five main attributes: productivity (ability to produce a combination of outputs), stability (to reproduce the former), reliability (ability to remain at an equilibrium state in normal conditions), resilience (ability to recover a normal stage following a perturbation) and adaptability or flexibility (ability to function in new conditions). Although this last approach is interesting in its genericity and in avoiding long list of indicators, its implementation raises problem when a property should be translated into concrete indicators (Alkan Olsson *et al.*, 2009).

Most of sustainability frameworks are structured across three sustainability dimensions or pillars: economic, environmental and social. Within each dimension, a list of items, goals, themes, etc. are defined. In some cases, those are organized in a hierarchical way which leads the aggregation step of indicators like for the MASC (Craheix *et al.*, 2012) and DEXiPM models (Pelzer *et al.*, 2012), or in the method of van Asselt *et al.* (2014). The property-based approach proposes generic properties across the three dimensions of sustainability. A similar attempt was presented by Alkan Olsson *et al.*, 2009) for a goal-based indicator framework (GOF) that classifies sustainability themes between *ultimate goal*, *process to achieve* (goal) and *means* (Alkan Olsson *et al.*, 2009). This classification of themes across sustainability dimensions considers the action chain. A policy is motivated by one or several *ultimate goals* (*e.g.* human health, viability, see Bossel (2000)) requiring some *process to achieve* it (*e.g.* balance of environmental function, improvement of economic performance) and *means* (*e.g.* protecting environmental compartment, increasing financial capital), across the three sustainability dimensions.
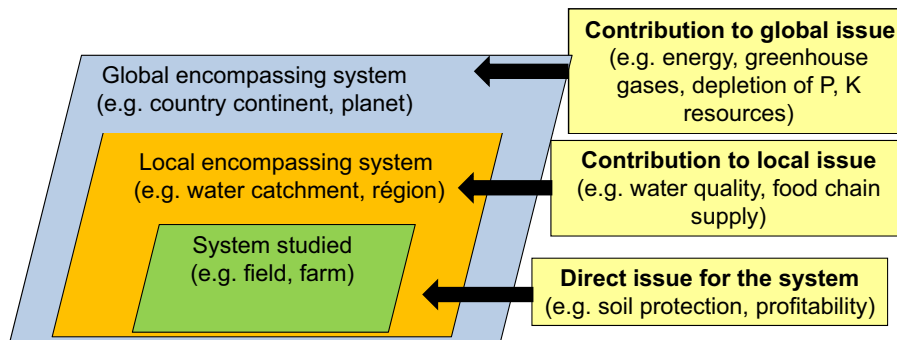
Fig. 1. Definition of system and related sustainability issues.

3 Preliminary choices

The clarification of a sustainability framework that structures the indicator selection has to be completed by several preliminary choices and assumptions (Bockstaller *et al.*, 2008). An initial diagnosis is required to identify the actual issues regarding sustainability for a given system, to put in evidence stakeholders and participants implicated, processes involved, degree of severity of impacts etc. (answering therefore the question: *why to evaluate?*).

The identification of the end-users (*to evaluate for whom?*) and the definition of the practical objectives (*to evaluate for what?*) of the indicator, were pointed out as an essential step by several authors (Brooks and Bubb, 2014; Girardin *et al.*, 1999; Mitchell *et al.*, 1995). This preliminary steps will serve as a basis to design or select indicators that meet end-users needs and requirements. Different users groups can be identified like, for example, scientists, advisors, farmers, decision makers, or consumers. The group of people doing the calculations and the group of people using the results should be differentiated. In many cases, although farmers are the targeted user group, they are actually not direct users of the evaluation method but they are end-users of the results (Cerf and Meynard, 2006). It should be noticed that the position of users and end-users of results is in many cases not only driven by scientific considerations (Bouleau, 2012; Gudmundsson, 2003). Indeed, they may interest in the selection of indicators or their results to defend their interests.

An indicator can be developed for various objectives. Those can be ordered in three main usages: (i) to gain knowledge about a system, *e.g. ex post* evaluation of an action at the end or during its implementation, monitoring purpose with an alert role, or checking the respect of regulation. (ii) for decision support: *ex ante* evaluation of actions in a planning phase to select the "best" system (Sadok *et al.*, 2008), decision support in real time to drive the system, (iii) communication which implies a reduced number of indicators easy to understand (Mitchell *et al.*, 1995).

The design of a sustainability framework allows to define the issue of concerns or criteria to precise the content of evaluation (*to evaluate what?*). Those can be presented in form of a set of strategies, of goals, themes, systemic properties, etc. as presented in the previous section. The definition of the system boundaries is another important step directly linked to the previous one (Van Cauwenbergh *et al.*, 2007). It includes the calculation scales, spatial and temporal (*to evaluate where? When?*), the organizational level, which will be influenced by the user needs, the issues of concerns, etc. In Life Cycle Analysis approaches (LCA), users are compelled to define the system boundaries (Brentrup *et al.*, 2004). It can be the product, the farm including or not upstream such as production of inputs and offstream activities such as waste management. In many other evaluation methods, this definition seems to be neither explicit nor unified between indicators. For example, in several methods indirect energy cost due to fertilizer, pesticide, machine production are included in the energy calculation like in LCA although for the rest of issues like water quality or emissions of pollutant to air, such approach is not implemented and only direct impact at field or farm level are covered. Regarding spatial and time scales one should pay attention to the resolution of calculation and the level at which basic calculations are carried out. Farm and year are typical resolution for environmental indicators. This should not be confused with the extent, *i.e.* the whole area, (*e.g.* the region), or time span, (*e.g.* the crop rotation), covered by the indicators calculation (Purtauf *et al.*, 2005).

Another aspect to consider is the differentiation between the system itself and the encompassing systems that can be separated between the local and the global ones (Fig. 1). This refers respectively to "on-site" (on the system) and "off-site" issues (outside the system), (Smith and McDonald, 1998). "On-site" issues (*e.g.* soil quality of farm fields) are linked to the sustainability of the studied system (*e.g.* farm) while "off-site" issues concern the encompassing system, (*e.g.* region where the farm is located), or the society as a whole (Alkan Olsson *et al.*, 2009). The former refers to sustainability of agriculture itself, the latter may be considered as the contribution of agricultural systems to sustainable development. For the latter, local and global issues can be distinguished. In any case, in the perspective of sustainability, a balanced choice between direct issues of the system, and contribution to local and global issues should be done (Alkan Olsson *et al.*, 2009).

Last but not least, means and resources should be assessed before selecting a method. It includes budget, time, data availability, etc. to be sure that they meet the requirements of a selected method.

4 Different types of indicators

Different typologies were proposed to categorize the broad variety of indicators proposed in the last decades. At the
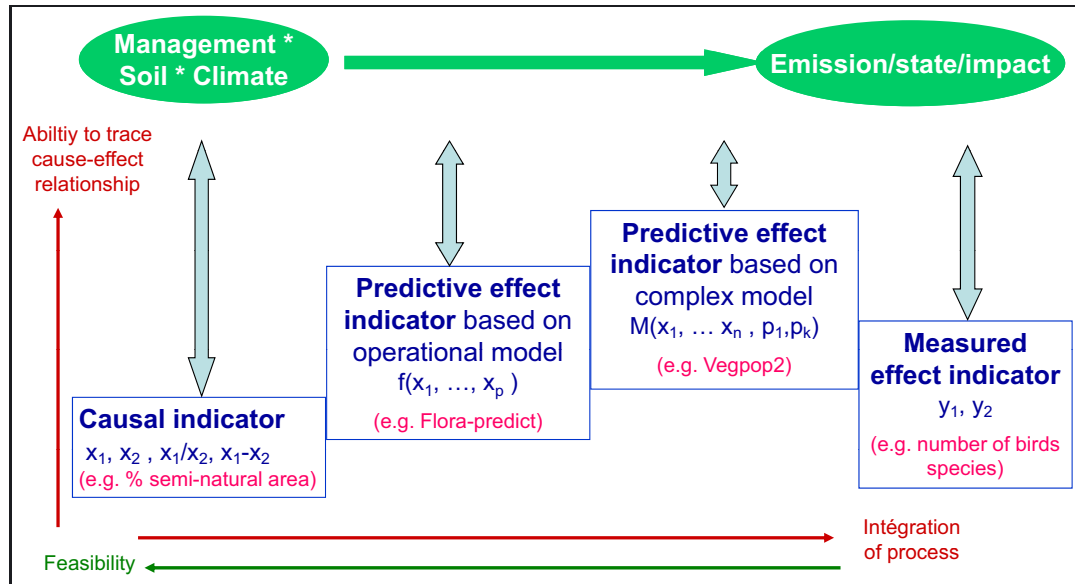
Fig. 2. Typology of indicators according to their nature (adapted from Bockstaller *et al.*, 2011).

international level the well-known Pressure/State/Response (PSR) and Driving-force/Pressure/State/Impact/Response (DPSIR) frameworks were inspired by the cause-effect chain. They were developed to ascertain the relevance of environmental indicators for human activities and their consequences at national level. These frameworks are extended by some authors to lower scales (*e.g.* Maurizi and Verrel, 2002) in spite of criticisms formulated by different authors (Niemeijer and de Groot, 2008). One major drawback is the impression of linearity between pressure, state and impact given by the framework, whereas the reality is more complex and closer to a causal network than to a chain. Another flaw is the ambiguity of the item, for instance pressure. Behind, you can find several types of indicators (see below). For example, pressure encompasses emission indicators which can be measured (*e.g.* nitrogen content at bottom of the root zone, measured by ceramic cup) or model output (from field leaching model), as well as simple indicators based on information from farmers' management data (*e.g.* amount of nitrogen input).

Several authors made the difference between (i) *means-based indicators* (van der Werf and Petit, 2002), or *action-oriented indicators* (Braband *et al.*, 2003) using information on farmers' practices or other causal variables; and (ii) *effect-based indicators* or *result-oriented indicators*, based on an assessment of the effect at different stages of the cause-effect chain. Concerning biodiversity, some authors also distinguished between indirect (means-based) and direct (effect-based) indicators. However, observing that these classes still cover indicators totally different like measurement or model output, Bockstaller *et al.* (2011) and, Feschet and Lairez (2015) proposed another typology taking into account the nature and structure of the indicators. Here we propose an adaptation of the former based on the following classes:

(i) *Causal indicator* based on a causal variable or a simple combination of variables of same nature (sum, product, ratio). Causal indicators can be based on management
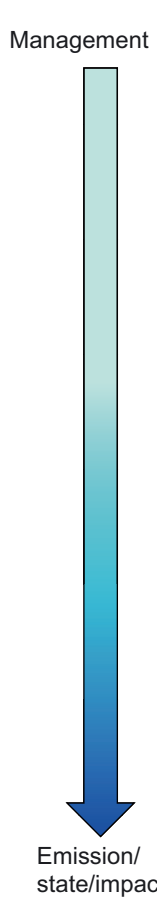
variables or environmental variables (soil, climate, etc.). For the former, the term "means" is relevant but not for the latter;

(ii) *Predictive effect indicators* based on model output that can be operational (with a reduced and available number of input variables) or complex (from the research point of view, without considering the number and availability of input data);

(iii) Measured effect indicators based on field assessment or observation.

Effect indicators may refer more precisely to different stages on the cause-effect chain: emission, state or impact (Bockstaller *et al.*, 2008). As shown on Figure 2, those types do not show same qualities. Indicators belong to the groups of causal indicators and result in most cases in a poor predictive quality whereas measurement indicators may provide more precise information about the state or the impact, without providing information on the causes. Predictive indicators are useful for *ex ante* assessment and to relate effect to cause (Bockstaller *et al.*, 2008). In any case, all those types have their utility, causal indicator to highlight changes in management or environment sensitivity, measured effect indicators for monitoring, predictive indicators for analyzing cause-effect relations in order to improve the system.

5 Example of indicators and of evaluation methods

In this section we illustrate the diversity of indicators for two major environmental issues: impact of nitrogen management and biodiversity. We also describe different evaluation methods based on a set of indictors to highlight their variability.

Management

Category	Number	Scale	Example
Input management	11	Farm, *Year*	*Storage capacitiy*[a]
Good management	13	Farm, Field *Year*	*Number of fertilizations*[a]
Fertilized area	4	Farm, Water catchment *Year*	*% reference area fertilized with organic fertilizer*[a]
Nitrogen amount	20	Farm, Field, Cropping system, *Year*	*Nitrogen rate*[a, b]
Nitrogen balance	31	Farm, Field *Pluriannual*	*Farm-gate balance*[a]
Deviation from the recommended nitrogen rate	5	Farm, Field *Year*	Deviation from recommendation[a]
Productivity / Efficience	7	Farm, Field *Year*	Nitrogen use efficiency[b]
Soil cover management	20	Field *Year*	*Bare soil during drainage period*[a]
Sensitivity of environementt o transfer	15	Field, Water catchment *Pluriannual*	*DRASTIC*[b]
Nitrogen balance x Coefficient	3	Field, Cropping system *Pluriannual*	*N-Index Tier-1*[b]
Qualitative operational model	5	Field, Cropping system *Year, Pluriannual*	*MERLIN*[a, c]
Quantitative operational model	17	Field, Cropping system *Year, Pluriannual*	*IN INDIGO*[d]
Measurement of soil nitrogen	6	Field *Year, Pluriannual*	*Soil mineral nitrogen*[b]
Measurement of water quality	8	Water catchment, Field *Pluriannual*	*Nitrate concentration*[a]

Emission/ state/impact

DOSSIER

Fig. 3. Typology of nitrogen indicators (Schneller *et al.*, 2013a). References for the examples: [a] CORPEN, 2006, [b] Buczko and Kuchenbuch, 2010, [c] Aveline *et al.*, 2009, [d] Bockstaller *et al.*, 2008.

5.1 Nitrogen indicators

The typology presented in Figure 3 results from the analysis of a database of 1464 environmental indicators issued from 112 methods or reviews on indicators. The database was created by Rosnoblet *et al.* (2006) and extended and analyzed by Schneller *et al.* (2013b). The whole variability presented in Figure 3 ranges from causal indicators based on management data, like the manure storage capacity, to measured effect indicators of water quality with in-between effect predictive indicators based on operational models. The category "nitrogen balance" is the most popular with more than 30 proposals. Those could be differentiated between farm-gate, soil surface and soil system budgets (Oenema *et al.*, 2003). Although nitrogen balance is recognized as the most commonly used indicator to assess nitrogen management (Langeveld *et al.*, 2007), its predictive quality of nitrogen losses, especially nitrate leaching remains questionable, especially in the case of calculations with annual data in situations with low surplus (Oenema *et al.*, 2005).

5.2 Biodiversity indicators

Since the 80s, a large number of direct measured indictors for biodiversity has been proposed in the literature and extensively discussed by some authors (*e.g.* Lindenmayer and Likens, 2011). Indicators based on species diversity and/or abundance among a given taxon or several taxa (*e.g.* birds, plants, carabid beetles, etc.) are the most commonly used at different scales, from field to national level. Many proposals also exist for causal or indirect indicators. Among the 91 indicators listed for agriculture by Delbaere (2003) more than half belongs to this type. When considering the general model explaining biodiversity in farmland (Le Roux *et al.*, 2008) they can be in two groups: (i) indicators related to management of farmland like the percentage of semi-natural area and, (ii) indicators addressing cropping practices, which can be expressed in amount of inputs per area unit or in percentage of area disturbed by fertilizer, pesticides, irrigation, tillage. Both groups can be expressed at different scales.

Contrary to the two previous types of indicators, examples of predictive effect indicators according to the typology in Figure 2 are less numerous. Table 1 provides an overview of initiatives based on an operational model in arable farming. Most of the methods belong to a multi-criteria assessment method. SALCAbd (Jeanneret *et al.*, 2014) was developed to complete the SALCA method based on Life Cycle Analysis although it only deals with direct effects at field and not upstream or off-stream indirect impacts (see Sect. 3). Outputs are in form of a probability of presence or a decreased of number of species,

Table 1. Examples of predictive biodiversity indicators with their main characteristics (adapted from Bockstaller *et al.*, 2011).

Name reference	Taxonomic group (species)	Species number	Expression of result	Aggregation function	Input variable or component	Scale
VEM (Sanderson *et al.*, 1995)	Plants	534	Probability of presence between 0 and 1	Calculation of suitability index from the British National Vegetation Classification. A probability of presence is derived from this index	*E.g.* grazing (yes/no) Slurry application (yes/no) Mineral fertilization (kg N/ha)	Fields, farm, landscape
Keichinger (2001)	Pheasant Partridge Field hare Wild rabbit	1 1 1 1	Score between 0 (maximum impact) and 10 (no impact)	Decision tree with fuzzy subsets for global indicators and components	Soil cover Crop diversity Machine use Pesticide risk Irrigation Semi-natural area	Farm, landscape
Meyer-Aurich *et al.* (2003)	Amphibian Partridge	Not explicit 1	Disturbance impact scored between 0 (none) and 1 (high)	Continuous function	*E.g.* number of tillage perturbation amount of nitrogen number of herbicides	Field, farm, landscape
SALCAbd (Jeanneret *et al.*, 2014)	Plants Mammals Birds Amphibians Snails and slugs Spiders Carabid beetles Orthoptera Bees and bumblebees Butterflies	Not explicit	Score between 1 (negative impact) and 5 (positive impact)	Each cropping practices is scored and an average value is calculated	All cropping practices	Field, farm
Butler *et al.* (2009)	Birds Bumblebees Butterflies Mammals Broadleaf plants	63 14 23 44 190	Risk score between 0 (none) and 3 to 6 (high)	Scoring system: assessing impact on species needs (*e.g.* diet, forage habitats)	*E.g.* spring to autumn sowing Agrochemical inputs Loss of non-cropped habitat Land drainage	Regional

Table 1. Continued.

Name reference	Taxonomic group (species)	Species number	Expression of result	Aggregation function	Input variable or component	Scale
MASC (Craheix *et al.*, 2012)	Not precised	No	4 qualitative classes	Decision tree implemented in DEXi software (Bohanec *et al.*, 2008)	Crop diversity, Non-sprayed area, Treatment frequency index[a]	Cropping system
DEXiPM (Pelzer *et al.*, 2012)	Flying natural enemies, Polllinators, Soil natural enemies, Weeds, Flora of semi-natural area	Not explicit	5 qualitative classes	Decision tree implemented in DEXi software (Bohanec *et al.*, 2008)	E.g. deep tillage, Treatment frequency index[a], Habitat management for soil natural ennemies	Cropping system and neighboring habitat
Sattler *et al.* (2010)	Bird, Amphibian, Mammal, Hoverfly, Field flora	1, 1, 1, Not explicit, Not explicit	Between 0 (low) and 1	Decision tree with fuzzy subsets	E.g. cropping practices, Herbicides treatment frequency index[a], Fertilization (kg N/ha)	Field
Tichit *et al.* (2010)	Birds	2	Occurrence of bird (yes/no)	Logistic regression function	11 management and 11 habitat variables	Field
IBEA Anonymous (2011)	None	None	5 qualitative classes	Decision tree implemented in DEXis software (Bohanec *et al.*, 2008)	E.g. mineral fertilization (kg N/ha), Treatment frequency index[a], Area with tillage	Farm

[a] Sum of ratio actual pesticide rate/recommended rate. Here the index is calculated separately for fungicides, herbicides, insecticides.

DOSSIER

or in form of risk or impact scores. Whereas some models tackle in an explicit way a broad number of species, for plants (Sanderson *et al.*, 1995) or several taxa (Butler *et al.*, 2009), most of them focus on a few number of species or few taxa without explicit information on these.

5.3 Examples of different French evaluation methods of sustainability

Providing an exhaustive overview of the "explosion" of evaluation methods in the last decades is totally out of the scope of the article. The reader should refer to synthesis like this of Rosnoblet *et al.* (2006) or Singh *et al.* (2012). Table 2 gives an overview of six French evaluation methods of sustainability in agriculture. The number of indicators for each method is higher than 30 with highest number for DAESE and EVAD. Strictly speaking, DAESE and EVAD are not evaluation method ready to implement by the end-users but offers respectively a broad list of indicators for the first one, and an indicator list and a methodological framework for the second. At farm level, causal indicators based on management variables are implemented to make the method easy to use with farmers. MASC and DEXiPM have been developed to be implemented in research work to evaluate *ex ante* innovative cropping systems during the design phase and to help to select and improve most performant systems for experimentation. MASC is based on quantitative predictive effect indicators (see Fig. 2) and can also be used to evaluate actual cropping systems (*ex post* evaluation). DEXiPM is based on qualitative causal indicators which are aggregated to make qualitative predictive effect indicators. This makes the evaluation work much faster than in MASC in spite of higher number of basic indicators. For some methods like IDEA, aggregation is based on a sum of scores that is questioned by many authors (Bockstaller *et al.*, 2008). MASC and DEXiPM decision trees are implemented thanks to the DEXi software (Bohanec *et al.*, 2008). This tool allows the design of such qualitative decision trees using "if then" rules and input variables organized in classes.

6 Discussion

The clarification of preliminary choices described in Section 3 is important to guide the user in his selection of an evaluation method, to avoid him a contingent selection led by availability of a method in his organization or surroundings. Indeed, such a choice can lead to the use of a non-adapted method that does not meet his needs or his means or even more to biased evaluation of sustainability. In any case, the great number of indicators within some themes like nitrogen, as well as the variety of evaluation methods makes the preliminary step of clarification essential to avoid a randomly approach equivalent to a lottery. Consequently, some authors propose a list of criteria to select the "best" method (Bockstaller *et al.*, 2009; Feschet and Lairez, 2015; Niemeijer and de Groot, 2008), or even more an interactive decision-aid tool such the PLAGE web platform (Surleau-Chambenoit *et al.*, 2013).

However, for some themes, the number of indicators may be low. This is particularly true for predictive effect indicators based on operational models. Such indicators require an important design work integrating knowledge on the effect of farmer management in interaction with soil, climate variables to give an output which expresses an effect which can be linked to emissions, or the state and furthermore the impact like in LCA. Topics like biodiversity (see Sect. 5.2), soil compaction, nuisance due to noises and odours present gaps for predictive effect indicators. A new challenge will be to evaluate ecosystemic services in such a predictive way, from management practices to biodiversity and from biodiversity to ecosystemic services. The second step is addressed, for example, by an indicator assessing flower pollination value of floral diversity in field margin (Ricou *et al.*, 2014).

In any case, users need also to know the quality of the information delivered by an indicator, especially when it is used to evaluate effect on an issue of sustainability like an environmental impact. Due to many simplifications, a direct correlation can rarely be expected excepted for a broad range of landscape conditions. Such correlations were for example pointed out between diversity within taxon (birds, bees, etc.) and causal indicators based on management variable like the nitrogen input or percentage of semi-natural area (Billeter *et al.*, 2008). Specific tests to identify more complex relations have been also proposed (Bockstaller *et al.* (2008). In every case, this raises the question of uncertainty linked to indicator results. For an indicator based on the nitrogen balance multiplied by a coefficient (see Fig. 3), Mertens and Huwe (2002) handled uncertainty of data by implementing an approach based on fuzzy logic so that no unique value but an interval is given to the user.

Many indicators are available at field and farm level as it comes out from the overview given in this article. However, for example, water quality indicators should be used at the scale of the water catchment or for a landscape. Emissions can be assessed at lower scale of cropping and farming systems. For indicator assessing emissions, results can be upscaled by aggregation of results obtained by calculation of an average value at higher scale weighted by the size or the number of entity at lower scale. Such aggregation at higher scale like a nation is not relevant for local impact, *e.g.* water quality, erosion, whereas it is possible for global impact, such as greenhouse gases. Upscaling requires some statistical skills for data management but must also integrate new processes (Stein *et al.*, 2001) and new environmental components (*e.g.* non cropped area). In any case, even without an upscaling procedure, the possibility to work at a fine resolution (*e.g.* field) for large extent (*e.g.* a water catchment) is an important challenge for agronomists to enable to work on realistic scenarios of management evolution, allowing finer description of cropping systems than only the type of crop and average fertilizer rates (Leenhardt *et al.*, 2010).

As already mentioned, sustainability is by nature a multidimensional issue which addresses a set of criteria which can be simply organized in a list or in a more complex framework. A question rises sooner or later concerning the necessity of an aggregation to facilitate the interpretation of set of results, especially when more than 30 indicators are used (see Tab. 2).

Table 2. Overview of six French evaluation methods of sustainability in agriculture (adapted from Feschet and Bockstaller, 2014).

Name of the method	Period of construction	Type of production	Scale	Dimensions of sustainability considered	Decomposition of sustainability in	Number o main criteria	Number of indicators	Comments: objectives, type of indicators, aggregation method
IDEA (Zahm *et al.*, 2008)	2000–2008	Arable, livestock, winegrowing	Farm	Economic, social and environmental	Scale/component/indicator	10	42	Evaluation of sustainability at farm level to make farmers sensitive to this stake. Use of causal indicators for environmental dimension, sum of scores within each dimension. The global sustainability score is the worst of the three dimensions
Diagnostic de durabilité – RAD (Féret, 2004)	2010	Dairy production	Farm	Economic, social and environmental	Dimension/criteria/indicator	18	34	Evaluation of sustainability at farm level to make farmers sensitive to this stake. No aggregation of indicators.
DAESE – OTPA (Guillaumin *et al.*, 2007)	2003–2007	Arable, livestock, winegrowing	Farm	Economic, social environmental	Axis/compartment/indicator	19	120	A large range of (causal) indicators to monitor change in management regarding sustainability for a farm network at regional (national) level. No aggregation.
EVAD (Rey-Valette *et al.*, 2009)	2005–2008	Aquaculture	Farm, production sector/landscape	Economic, social, institutional[a] and environmental	Dimension/principles/criteria/indicator	13	230	Decision aid tool for action, design and selection of a wide range of indicators (from different types)
MASC 2.0 (Craheix *et al.*, 2012)	2005–2012	Arable crops	Cropping system	Economic, social and environmental	Dimension/aggregated criteria/basic criteria/indicator	8	39	Evaluate and rank cropping systems by identifying strong and weak points regarding sustainability. Use of predictive effect indicators when possible. Aggregation thanks to DEXi software.
DEXiPM (Pelzer *et al.*, 2012)	2007–2011	Arable crops	Cropping system	Economic, social and environmental	Dimension/caggregated criteria/basic criteria/indicator/practices	8	61	Rapid *ex ante* evaluation of innovative cropping systems to select the most performant ones. Causal qualitative indicators aggregated in qualitative predictive effect indicators. Aggregation with DEXi software

[a] The institutional dimension encompasses for instance governance, participation, empowerment of community and women, justice.

DOSSIER

The relevance of composite aggregation (*i.e.* of indicators addressing totally different issues) is often discussed because of the loss of information but also due to the methodological problems it raises. A major problem is "adding apple and pear" in the case of composite indicators which can appear in scoring method (Rigby *et al.*, 2001). Several method are available to avoid this problem like the normalization technique in monetary unit or physic unit, the multivariate approach, or decision trees based – or not – on fuzzy logic (Bockstaller *et al.*, 2008). The DEXi software tool (Bohanec *et al.*, 2008) makes the design of decision tree (without fuzzy logic) quite easy but remains totally qualitative. For complex trees, special attention should be therefore put on the sensitivity of the aggregated indicator to the variability of basic ones (Carpani *et al.*, 2012). Another flaw is a certain lack of transparency when the automatic weighting procedure ot the software is used. A totally different approach is the use of multi-criteria methods based on an outranking (Cinelli *et al.*, 2014) which has the default of being based on relative comparisons and not on an absolute assessment. More recently another approach addressing, in a transparent way, compensations between indicators has been proposed (van Asselt *et al.*, 2014). In every case, we advise to use both aggregated and individual indicators, the former to compare systems or to select performant ones, the latter to to identify weak and strong points.

References

Alkan Olsson J, Bockstaller C, Stapleton L, Knapen R, Therond O, Turpin N, *et al.* 2009. Indicator frameworks supporting ex-ante impact assessment of new policies for rural systems; a critical review of a goal oriented framework and its indicators. *Environ. Sci. Policy* 12: 562–572.

Anonymous. 2011. Un outil de diagnostic de l'impact des pratiques agricoles sur la biodiversite : Outil "IBEA" : Impact des pratiques agricoles sur la Biodiversité des Exploitations Agricoles. FNE, INRA, MNHN, Paris, 105 p. (in revision).

Aveline A, Rousseau ML, Guichard L, Laurent M, Bockstaller C, 2009. Evaluating an environmental indicator: case study of MERLIN, an assessment method of the Risk of Nitrate Leaching. *Agr. Syst.* 100: 22–30.

Billeter R, Liira J, Bailey D, Bugter R, Arens P, Augenstein I, *et al.* 2008. Indicators for biodiversity in agricultural landscapes: a pan-European study. *J. Appl. Ecol.* 45: 141–150.

Bockstaller C, Girardin P, Van der Werf HGM. 1997. Use of agro-ecological indicators for the evaluation of farming systems. *Eur. J. Agron.* 7: 261–270.

Bockstaller C, Guichard L, Keichinger O, Girardin P, Galan MB, Gaillard G. 2009. Comparison of methods to assess the sustainability of agricultural systems. A review. *Agron. Sustain. Dev.* 29: 223–235.

Bockstaller C, Guichard L, Makowski D, Aveline A, Girardin P, Plantureux S. 2008. Agri-environmental indicators to assess cropping and farming systems. A review. *Agron. Sustain. Dev.* 28: 139–149.

Bockstaller C, Lassere-Joulin F, Slezack-Deschaumes S, Piutti S, Villerd J, Amiaud B, *et al.* 2011. Assessing biodiversity in arable farmland by means of indicators: an overview. *OCL* 18: 137–144.

Bohanec M, Messean A, Scatasta S, Angevin F, Griffiths B, Krogh PH, *et al.* 2008. A qualitative multi-attribute model for economic and ecological assessment of genetically modified crops. *Ecol. Model.* 215: 247–261.

Bossel H. 2000. Policy assessment and simulation of actor orientation for sustainable development. *Ecol. Econ.* 35: 337–355.

Bouleau G. 2012. Ce que nous apprend l'histoire des indicateurs environnementaux. *Rev. fr. For.* 5: 645–652.

Braband D, Geier U, Kopke U, 2003. Bio-resource evaluation within agri-environmental assessment tools in different European countries. *Agric. Ecosyst. Environ.* 98: 423–434.

Brentrup F, Kusters J, Kuhlmann H, Lammel J. 2004. Environmental impact assessment of agricultural production systems using the life cycle assessment methodology – I. Theoretical concept of a LCA method tailored to crop production. *Eur. J. Agron.* 20: 247–264.

Brooks S, Bubb P. 2014. Key Knowledge for Successful Biodiversity Indicators UNEP-WCMC, Cambridge, UK.

Buczko U, Kuchenbuch RO. 2010. Environmental Indicators to Assess the Risk of Diffuse Nitrogen Losses from Agriculture. *Environ. Manage.* 45: 1201–1222.

Butler SJ, Brooks D, Feber RE, Storkey J, Vickery JA, Norris K. 2009. A cross-taxonomic index for quantifying the health of farmland biodiversity. *J. Appl. Ecol.* 46: 1154–1162.

Carpani M, Bergez JE, Monod H. 2012. Sensitivity analysis of a hierarchical qualitative model for sustainability assessment of cropping systems. *Environ. Modell. Softw.* 27-28: 15–22.

Cerf M, Meynard JM. 2006. Les outils de pilotage des cultures : diversité de leurs usages et enseignements pour leur conception. *Natures Sciences Sociétés* 14: 19–29.

Cinelli M, Coles SR, Kirwan K. 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. *Ecol. Indic.* 46: 138–148.

CORPEN. 2006. Des indicateurs AZOTE pour gérer des actions de maîtrise des pollutionsà l'échelle de la parcelle, de l'exploitation et du territoire. Ministère de l'Ecologie et du Développement Durable, http://www.ecologie.gouv.fr/IMG/pdf/maquette_azote29_09.pdf, Paris.

Craheix D, Angevin F, Bergez JE, Bockstaller C, Colomb B, Guichard L, *et al.* 2012. MASC 2.0, un outil d'évaluation multicritère pour estimer la contribution des systèmes de culture au développement durable. *Innov. Agronom.* 20: 35–48.

Delbaere B. 2003. An inventory of biodiversity indicators in Europe, 2002. European Environment Agency (EEA), Copenhagen.

Falque A, Feschet P, Garrabé M, Gillet C, Lagarde V, Macombe C, *et al.* 2013. ACV sociales: Effets socio-économiques des chaînes de valeur. Fruitrop Thema, Montepllier, France.

Féret S. 2004. Un outil mis au point dans l'Ouest de la France : le diagnostic du Réseau agriculture durable. *Travaux et Innovations* 110: 39–42.

Feschet P. 2014. Analyse du Cycle de Vie Sociale. Pour un nouveau cadre conceptuel et théorique. Thèse de doctorat de l'Université Montpellier 1 – Sciences économiques, 298p.

Feschet P, Bockstaller C. 2014. Méthodes d'évaluation multicritère des systèmes agricoles et ACV sociale, quelle complémentarité ? In: Macombe, C., Loeillet, D. (Eds.), Social LCA in progress, 4th International Seminar in Social LCA, pre-proceedings. Fruitrop, Montepllier, France, pp. 18–25.

Feschet P, Lairez J. 2015. Guide méthodologique pour l'évaluation multicritère de la durabilité en agriculture. Dir. Aubin J, Bockstaller C, Bouvarel I, Ed. Quae-Educagri, Coll. Sciences en partage, Paris (à paraître).

Girardin P, Bockstaller C, van der Werf HMG. 1999. Indicators: tools to evaluate the environmental impacts of farming systems. *J. Sustain. Agr.* 13: 5–21.

Gras R, Benoit M, Deffontaines JP, Duru M, Lafarge M, Langlet A, *et al.* 1989. Le fait technique en agronomie. Activité agricole, concepts et méthodes d'étude. Institut National de la Recherche Agronomique, L'Hamarttan, Paris, France.

Gudmundsson H. 2003. The policy use of environmental indicators – learning from evaluaiton research. *J. Transdisciplinary Environ. Studies* 2: 1–11.

Guillaumin A, Hopquin JF, Desvignes P, Vinatier JM. 2007. OTPA: Des indicateurs pour caractériser la participation des exploitations agricoles d'un territoire au développement durable. 1ère partie: recommandation pour la mise en oeuvre. Projet CASDAR OTPA.

Hansen JW. 1996. Is agricultural sustainability a useful concept? *Agr. Syst.* 50: 117–143.

Jeanneret P, Baumgartner DU, Knuchel Freiermuth R, Koch B, Gaillard G. 2014. An expert system for integrating biodiversity into agricultural life-cycle assessment. *Ecol. Indic.* 46: 224–231.

Keichinger O. 2001. Evaluation de l'impact des pratiques agricoles d'exploitations de grandes cultures sur la valeur cynégétique à l'aide d'indicateurs agro-écologiques. Thèse de doctorat, INPL-ENSAIA, Nancy.

Lacousmes P. 2005. Le développement durable : une "illusion motrice". Le développement durable. Les termes du débat. Armand Colin, Paris, pp. 95–107.

Langeveld JWA, Verhagen A, Neeteson J, van Keulen H, Conijn JG, Schils RLM, *et al.* 2007. Evaluating farm performance using agri-environmental indicators: recent experiences for nitrogen management in The Netherlands. *J. Environ. Manage.* 82: 363–376.

Ledoux L, Mertens R, Wolff P. 2005. EU sustainable development indicators: An overview. *Nat. Resour. Forum* 29: 392–403.

Leenhardt D, Angevin F, Biarnes A, Colbach N, Mignolet C. 2010. Describing and locating cropping systems on a regional scale. A review. *Agron. Sustain. Dev.* 30: 131–138.

Le Roux X, Barbault R, Baudry J, Burel F, Doursan I, Garnier E, *et al.* (Eds.). 2008. Agriculture et Biodiversité. Valoriser les synergies, Expertise scientifique collective. Synthèse du rapport, INRA, France.

Lindenmayer DB, Likens GE. 2011. Direct Measurement Versus Surrogate Indicator Species for Evaluating Environmental Change and Biodiversity Loss. *Ecosystems* 14: 47–59.

López-Ridaura S, van Keulen H, van Ittersum MK, Leffelaar PA. 2005. Multi-scale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. *Environ. Dev. Sustain.* 7: 51–69.

Maurizi B, Verrel JL. 2002. Des indicateurs pour des actions de maîtrise des pollutions d'origine agricole. *Ingénieries* 30: 3–14.

Mertens M, Huwe BU. 2002. FuN-Balance: a fuzzy balance approach for the calculation of nitrate leaching with incorporation of data imprecision. *Geoderma* 109: 269–287.

Meyer-Aurich A, Zander P, Hermann M. 2003. Consideration of biotic nature conservation targets in agricultural land use – a case study from the Biosphere Reserve Schorfheide-Chorin. *Agric. Ecosyst. Environ.* 98: 529–539.

Mitchell G, May A, Mc Donald A. 1995. PICABUE: a methodological framework for the development of indicators of sustainable development. *Int. J. Sustain. Dev. World Ecol.* 2: 104–123.

Niemeijer D, de Groot RS. 2008. A conceptual framework for selecting environmental indicator sets. *Ecol. Indic.* 8: 14–25.

Oenema O, Kros H, de Vries W. 2003. Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. *Eur. J. Agron.* 20: 3–16.

Oenema O, van Liere L, Schoumans, O. 2005. Effects of lowering nitrogen and phosphorus surpluses in agriculture on the quality of groundwater and surface water in the Netherlands. *J. Hydrol.* 304: 289–301.

Payraudeau S, van der Werf HMG. 2005. Environmental impact assessment for a farming region: a review of methods. *Agric. Ecosyst. Environ.* 107: 1–19.

Pelzer E, Fortino G, Bockstaller C, Angevin F, Lamine C, Moonen C, *et al.* 2012. Assessing innovative cropping systems with DEXiPM, a qualitative multi-criteria assessment tool derived from DEXi. *Ecol. Indic.* 18: 171–182.

Purtauf T, Thies C, Ekschmitt K, Wolter V, Dauber J. 2005. Scaling properties of multivariate landscape structure. *Ecol. Indic.* 5: 295–304.

Rey-Valette H, Clément O, Aubin J, Mathé S, Chia E, Legendre, *et al.* 2009. Guide de co-construction d'indicateurs de développement durable en aquaculture. Cirad, Ifremer, IRNA, IRD, UM1 Montepellier, France.

Ricou C, Schneller C, Amiaud B, Plantureux S, Bockstaller C. 2014. A vegetation-based indicator to assess the pollination value of field margin flora. *Ecol. Indic.* 45: 320–321.

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

Riley J. 2001. The indicator explosion: local needs and international challenges. *Agric. Ecosyst. Environ.* 87: 119–120.

Robinson J. 2004. Squaring the circle? Some thoughts on the idea of sustainable development. *Ecol. Econ.* 48: 369–384.

Rosnoblet J, Girardin P, Weinzaepflen E, Bockstaller C. 2006. Analysis of 15 years of agriculture sustainability evaluation methods. In: Fotyma, M., Kaminska, B. (Eds.), 9th ESA Congress, Warsaw, Poland, pp. 707–708.

Sadok W, Angevin F, Bergez JE, Bockstaller C, Colomb, B, Guichard L. *et al.* 2008. Ex ante assessment of the sustainability of alternative cropping systems: guidelines for identifying relevant multi-criteria decision aid methods. *Agron. Sustain. Dev.* 28: 163–174.

Sanderson RA, Rushton SP, Pickering AT, Byrne JP. 1995. A preliminary method of predicting plant species distributions using the British National Vegetation Classification. *J. Environ. Manage.* 43: 265–288.

Sattler C, Nagel UJ, Werner A, Zander P. 2010. Integrated assessment of agricultural production practices to enhance sustainable development in agricultural landscapes. *Ecol. Indic.* 10: 49–61.

Schneller C, Lellahi A, Bockstaller C. 2013a. Caractérisation des indicateurs azote. Groupe de travail "évaluation multicritère" GIS GC-HP2E.

Schneller C, Lellahi A, Bockstaller C. 2013b. Document descriptif de la base de données INDIC : une compilation d'indicateurs au niveau mondial. Groupe de travail "évaluation multicritère" GIS GC-HP2E.

DOSSIER

Singh RK, Murty HR, Gupta SK, Dikshit AK. 2012. An overview of sustainability assessment methodologies. *Ecol. Indic.* 15: 281–299.

Smith CS, McDonald GT. 1998. Assessing the sustainability of agriculture at the planning stage. *J. Environ. Manage.* 52: 15–37.

Stein A, Riley J, Halberg N. 2001. Issues of scale for environmental indicators. *Agric. Ecosyst. Environ.* 87: 215–232.

Surleau-Chambenoit C, Morin A, Galan MB, Cariolle M, Leclercq C, Guichard L, *et al.* 2013. PLAGE, un réseau d'acteurs et une plateforme WEB dédiée à l'évaluation agri-environnementale et de la durabilité des pratiques agricoles, des exploitations agricoles et des territoire. *Innovations Agronomiques* 31: 15–26.

Tichit M, Barbottin A, Makowski D. 2010. A methodological approach to identify cheap and accurate indicators for biodiversity assessment: application to grazing management and two grassland bird species. *Animal* 4: 819–826.

Van Asselt ED, van Bussel, LGJ, van der Voet H, van der Heijden G, Tromp SO, Rijgersberg H, *et al.* 2014. A protocol for evaluating the sustainability of agri-food production systems-A case study on potato production in pen-urban agriculture in The Netherlands. *Ecol. Indic.* 43: 315–321.

Van Cauwenbergh N, Biala K, Bielders C, Brouckaert V, Franchois L, Garcia Cidad V, *et al.* 2007. SAFE – A hierarchical framework for assessing the sustainability of agricultural systems. *Agric. Ecosyst. Environ.* 120: 229–242.

Van der Werf HGM, Petit J. 2002. Evaluation of environmental impact of agroculture at the farm level: a comparison and analysis of 12 indicator-based methods. *Agric. Ecosyst. Environ.* 93: 131–145.

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.