

Review

The Use of Reference Values in Indicator-Based Methods for the Environmental Assessment of Agricultural Systems

Ivonne Acosta-Alba^{1,2} and Hayo M. G. van der Werf^{1,2,*}

¹ INRA, UMR1069, Soil Agro and hydroSystem, F-35000 Rennes, France;
E-Mail: ivonne.acosta-alba@rennes.inra.fr

² Agrocampus Ouest, UMR1069, Soil Agro and hydroSystem, F-35000 Rennes, France

* Author to whom correspondence should be addressed; E-Mail: hayo.vanderwerf@rennes.inra.fr;
Tel.: +33-223485709; Fax: +33-223485430.

Received: 6 December 2010; in revised form: 27 January 2011 / Accepted: 31 January 2011 /
Published: 16 February 2011

Abstract: Many indicator-based methods for the environmental assessment of farming systems have been developed. It is not the absolute values of the indicators that reveal whether the impact of a system is acceptable, but rather the distance between these values and some reference values. We reviewed eight frameworks for the environmental assessment of agricultural systems that define reference values for their indicators. We analyzed the methods used to establish reference values and explored how to improve these methods to increase their usage and relevance. This analysis revealed a striking diversity of terminology, sources, and modes of expression of results. *Normative reference values* allow the assessment of a single system with a previously defined value; *Relative reference values* are based on indicator values for similar systems or a reference system. Normative reference values can be *Science-based* or *Policy-based*. A science-based normative reference value can be a *Target value*, which identifies desirable conditions, or an *Environmental limit*, which is the level beyond which conditions are unacceptable. The quantification of the uncertainty of reference values is a topic which is barely explored and warrants further research. Reference values present a means of introducing site specificity into methods for environmental assessment which seems, at present, largely under-exploited.

Keywords: reference value; threshold; fair earthshare; critical flow; sustainability standard; critical threshold values; baseline value; environmental minimum requirement

1. Introduction

The sustainability of agricultural production systems has been the object of much study [1]. Sustainability has been defined in many ways, with the “triple bottom line” approach, which aims to balance the three dimensions of sustainability, being the most widespread. This approach allows for trade-offs between the biophysical, social and economic spheres. We consider a hierarchical view of sustainability to be more appropriate. The hierarchical vision considers that biophysical limits to sustaining life on earth are absolute, and that “societies cannot exist without a functioning life-support system, and economies can only flourish within a functioning social system with effective institutions and governance structures” [2].

To assess the sustainability of an agricultural production system, its impact on the environment must be quantified. Over the last decade many methods for the environmental assessment of farming systems have been developed. Several reviews have shown that these methods have become increasingly complex, integrating a variety of impacts and the latest scientific knowledge [3-10].

These methods have a similar structure consisting of seven stages which are more or less explicitly defined, depending on the method: (1) Definition of the system to be assessed; (2) Identification of the overall goal of the method and definition of the dimensions of encompassed sustainability; (3) Identification of objectives (issues of concern) to be considered in each dimension; (4) Selection or conception of indicators for each objective; (5) Establishment of reference values for each indicator; (6) Calculation of indicator values; (7) Interpretation of results, identification of improvement options.

Indicators are essential features in all methods; they are a favored tool to understand complex systems, such as agricultural systems, and their impacts on the environment [11]. Girardin *et al.* [11] describe a procedure for developing an indicator. In this procedure, the determination of reference values, norms, or veto thresholds constitutes a key stage. This issue is important, because it is not the absolute values of the indicators that reveal whether the impacts of a system are acceptable, but rather the distance between these values and some reference values. Thus, reference values help to interpret the indicator value and may guide the evolution of a system towards an acceptable level defined in the objectives of the study [12]. Reference values are requested by users, because they help to interpret the method's results.

Defining a desirable state of the environment is not easy. There is a lack of data and knowledge about ecosystem functions, and about the level of impact that may negatively affect these functions. Only a minority of environmental assessment methods define reference values which distinguish acceptable from unacceptable impact levels. Thus most assessment methods reflect a “less is better” approach, which allows their user to identify among several options the one having the lowest impact. This approach may seem environmentally responsible, but it is not good enough to evaluate impacts on natural systems [13] as it does not indicate whether “less is good enough”.

The purpose of this paper is to review a variety of frameworks allowing the environmental assessment of agricultural systems, which define explicit reference values for their indicators. We analyze the methods used to establish reference values and explore how to improve these methods to increase their usage and relevance.

2. Description of the Methods

An inventory of methods and tools for assessing environmental or overall sustainability of agricultural systems, or of economic activities in general, forms the basis of this article. We selected methods which explicitly establish reference values for their indicators from a literature review of methods published as journal articles, books, reports, conference proceedings, and on-line sources. We preferentially included methods published in peer-reviewed journals, which were actually applied to case studies. To ensure sufficient diversity we excluded methods with major similarities. The eight methods selected are described below, Table 1 summarizes the methods' major characteristics.

2.1. Framework for Evaluating Sustainable Land Management (FESLM)

FESLM is a sustainability evaluation framework created by a panel of experts [14]. Its creation was sponsored by ten international institutions involved in agricultural development and research. It was designed as a "structured, logical pathway for making decisions on whether or not a carefully defined form of land management is likely to prove sustainable in a defined period of time". The principles of sustainability evaluation used come from the *Framework of Land Evaluation* [15]. Sustainable land management is defined as maintaining production/services (Productivity), reducing the level of production risk (Security), protecting potential natural resources and preventing degradation of soil and water quality (Protection), and being economically viable (Viability) and socially acceptable (Acceptability). Sustainability is assessed for a particular type of land use during a stated period of time on a specific area. FESLM assists planning by comparing alternative forms of land use. *Indicators* are environmental statistics that measure or reflect environmental status or change in condition. *Thresholds* are critical levels for these indicators; a threshold level representing the level beyond which a system undergoes significant change. The interacting processes and factors that determine threshold levels are called *Criteria*. Criteria are standards or rules (models, tests or measures) that govern judgments on environmental conditions. Criteria can be deduced by four approaches: on-site observation, examining historic records for the site, comparison with similar sites and modeling. Gomez *et al.* [16] used FESLM to propose thresholds for the evaluation of agriculture in the Philippines at the scale of a region. Mean community-level values for various biophysical and economic indicators were used as thresholds.

2.2. Ecological Footprint (EF)

EF is a resource accounting tool used to quantify environmental sustainability, a "land-based surrogate measure of the population's demands on natural capital" [17]. It measures how much biologically productive land and water area a population or an economy use to satisfy its consumption and to absorb the waste generated, using existing technology and resource management [18]. The central concepts of EF are ecological, carrying capacity and overshoot. Overshoot occurs when an ecosystem is exploited faster than it can renew itself. The main assumptions made by its authors are: (i) earth's biocapacity is limited; (ii) every category of energy and material consumption and waste discharge requires the production or absorptive capacity of a finite area of land or water; and (iii) this

area can be defined for a specific human subpopulation or economy. Renewal and absorption rates depend on the health and integrity of ecosystems. EF considers five biocapacity components (cropland, grazing land, fishing grounds, forest area and built-up land) and also a “carbon land”, which is the amount of forest land required to take up anthropogenic CO₂ emissions to maintain a stable CO₂ concentration in the atmosphere. These six components are added up into a single value: the ecological footprint. Dividing the global bioproductive land and water area by the present world population yields a *fair earthshare*, which represents the reference value of one individual’s EF. By multiplying this value by a regional or national population, the same reasoning can be applied to a region or a nation. By definition, if the EF value exceeds the fair earthshare, the person’s or nation’s way of life is not environmentally sustainable. Because EF expresses environmental impact as a single indicator, its simplicity has been recognized as a powerful communication tool even by its critics [19]. This method has been updated several times and applied largely at national levels [20-25].

2.3. Ecological Scarcity Method (ESM)

The Ecological Scarcity Method, [26] is the latest update of the Ecopoints method [27,28]. It is a method for impact assessment in Life Cycle Assessment (LCA), which is a decision support tool for the environmental analysis of processes or products. An LCA produces an environmental inventory, which identifies resource consumption and pollutant emissions for all processes associated with a product’s life cycle: from the extraction of resources, their processing, the manufacture of the product, its use and disposal. ESM permits the aggregation of life cycle inventory data in a set of indicators of environmental impact according to the “distance to target” principle. Eco-factors, expressed as eco-points per unit of pollutant emission or resource extraction are the key parameter used by the method. Eco-factors are determined, reflecting, on the one hand, the current emissions situation in Switzerland and, on the other hand, Swiss national policy targets or international targets supported by Switzerland. The method has been adapted to other countries: Belgium, Japan, Netherlands, Norway and Sweden [26]. The more the current level of emissions or consumption of resources exceeds the *critical flow*, *i.e.*, the reference value based on policy targets, the greater the eco-factor, expressed in eco-points, becomes. The critical flow should be calculated or derived from statutory emission/ambient targets and/or from political statements of intent. If these are not available, it can be based on expert opinion or modeling assumptions of an advisory group. The Eco-factor is calculated as:

$$= K \times \frac{1 \times EP}{F_n} \times \left(\frac{F}{F_k} \right)^2 \times c$$

where *K*: characterization factor of a pollutant or resource; *EP*: Ecopoint; *F_n*: normalization flow, current annual flow for Switzerland; *F*: current annual flow in the reference area; *F_k*: critical annual flow in the reference area; *c*: constant [26].

As the ratio of the current flow over the critical flow is squared in the eco-factor formula, any growth of the current flow leads to an exponential growth of the eco-factor, as does a reduction in the critical flow. Spatial and temporal differentiation can be introduced, which is of obvious interest, as both current and critical flow may vary in space and time. Spatial differentiation has been implemented for freshwater resources, for which six scarcity categories have been defined.

A case study comparing several biofuels using ESM was published recently [29,30].

2.4. Sustainability Gaps (SGAPS)

Ekins and Simon [31] developed a method to determine whether economic activities in a region are environmentally sustainable. Environmental sustainability is defined as the maintenance of important environmental functions into the future. The method proposes indicators and reference values (*Sustainability Standards*, SS) to assess current development patterns. SS should be “derived as far as possible on the basis of objective considerations deriving from environmental science concerning the maintenance of important environmental functions, rather than being influenced by considerations of cost or political feasibility”. The level of SS should be set to respect the following principles: (a) not threaten critical ecosystems and/or biogeochemical systems; (b) not have a detrimental effect on human health; and (c) not harvest renewable resources faster than their rate of regeneration; or (d) not deplete non-renewable resources faster than the rate of development of substitutes. If this level is uncertain, it is recommended to use the safe minimum standard or the precautionary principle to avoid the risk of irreversible changes in future. This method estimates a “sustainability gap” which is the difference between the current level of environmental impact and the SS. The method was applied to assess the level of air pollution in the United Kingdom and the Netherlands. When targets used in environmental policy are based purely on science, SS become *Sustainability Targets*; however, if targets depend also on political will (e.g., science recommends stopping the emission of gasses causing global warming, but policy calls only for decreasing their emission), they become *Policy Targets*. Ekins and Simon [31] compared different targets and sustainability gaps for air emissions, and estimated “years to sustainability” by determining how long it would take, on continuation of current trends, for the sustainability standard to be attained.

2.5. Sustainability Assessment of Development Scenarios (SADS)

Nijkamp and Vreeker [32] present a framework to assess the sustainability of development strategies at a regional level, with a particular view on the treatment of uncertain information. They adopt the view that “sustainability means that the development of an economy has to take place within a set of pre-specified normative constraints or pathways”. This framework is based on a systematic multicriteria flag model capable to take into account *Critical Threshold Values* (CTV). A CTV is defined as “the numerical normative value of a sustainability indicator that ensures a compliance with the carrying capacity of the regional environmental system concerned”. The authors indicate that CTV are based on scientific information and expert opinion, more detail is not given. Exceeding a CTV would impose an unacceptably high cost on the environment. In this method, reference values are not a single value but a band width, defined by CTV_{min} and CTV_{max} , to reflect uncertainty. This band width mirrors the range of CTV values expressed by experts or policy makers. CTV_{min} indicates a conservative estimate of the threshold, while CTV_{max} refers to a maximum allowable value, with CTV_{int} being halfway between CTV_{min} and CTV_{max} . Color “flags” are attributed to indicator values: green (no reason for concern) for values below CTV_{min} ; yellow (be alert) for values between CTV_{min} and CTV_{int} ; red (reverse trends) for values between CTV_{int} and CTV_{max} ; and black (stop immediately

further growth) for values above CTV_{max} . Three development scenarios for the southern peninsular region of Thailand were compared using eighteen indicators summarizing social, economic and environmental sustainability [32].

2.6. Framework for Assessing the Sustainability of Natural Resource Management Systems (MESMIS)

MESMIS is an operational structure, used widely by different institutions in Latin America, notably to assess sustainability of natural resource management systems (NRMS) or ecosystems transformed by humans [33]. Main premises of MESMIS are: (i) sustainability is defined by seven attributes of NRMS: Productivity, Stability, Reliability, Resilience, Adaptability, Equity and Self-reliance; (ii) sustainability evaluations are valid for a specific management system on a specific spatial and time scale; (iii) evaluation of sustainability is a participatory process; and (iv) sustainability is assessed through the comparison of systems either at the same time or over time. After determination of the system's critical points (*i.e.*, features which have critical impact on the survival of the system) that have to be improved, indicators are selected. Indicator results are presented as scores between 0 and 100 in an Amoeba diagram to facilitate the comparison of analyzed systems. Most case studies using MESMIS compare systems at the farm scale, but it can be applied at different spatial scales [34]. The sustainability assessment is based on indicator values for the critical points in each system. The main advantage of this method is its flexibility and adaptability. An example is given by Brunett Pérez *et al.* [35], who compared two agro-ecosystems involving dairy and corn production. After choosing indicators, reference values (called *baseline values* or *thresholds*) for each indicator were chosen based on expert opinion or consultation of literature. The systems were compared with an "optimum" condition defined by the reference values. Then according to the distance between indicator value and optimum value, scores are given considering a system as more or less sustainable.

2.7. European Analytical Framework for the Development of Local Agri-Environmental Programmes (AEMBAC)

This framework is the outcome of a three-year (2001–2004) EU project [36]. It was tested in 15 study areas in seven European countries. The overall objective of the AEMBAC project was to create a tool for the identification, development and evaluation, of locally appropriate agri-environmental measures based on the analysis of indicators and the assessment of environmental functions [37]. Two types of indicators were used: *state indicators*, describing the state of the agro-ecosystem and its ability to perform environmental functions; and *pressure indicators*, describing pressures that the local agricultural systems exert on the environment. The state indicators depend on the important environmental issues in each area, and for each of them a reference value called *Environmental Minimum Requirement* (EMR) was identified. An EMR is a single value or a set of values (a range) that should allow a satisfactory performance of the environmental function analyzed [37]. The gaps between indicator values and their corresponding EMRs are assessed. The authors insist on the need to define local EMRs rather than EMRs at European or national levels. Because there is no single way to determine the value of an EMR, Bastian *et al.* [37] propose the following sources: natural ecosystems, past situations, expert judgment, scientific literature, and

agro-ecosystems where an environmental function is performed successfully. The authors find that for many environmental issues there is not sufficient scientific information available to know whether the performance of an environmental function is sustainable or not. They further point out that, from a philosophical point of view, an EMR cannot only be based on scientific fundamentals. Since nature has no inherent goals, it is not possible to draw conclusions (normative statements) from observations (descriptive statements). The authors conclude that EMRs should be based on the scientific knowledge available, but in the end, in addition to scientific information, targets have to be defined by human society. Thus subjective valuations and political choices will have to be made to establish reference values.

2.8. Sustainability Assessment of Farming and the Environment (SAFE)

SAFE is a framework for assessing the environmental, economic and social sustainability of agricultural systems. It does not seek to find a common solution for sustainability in agriculture as a whole, but to serve as an assessment tool for the identification, development and evaluation of locally more sustainable agricultural production systems, techniques and policies [38]. It can be applied at three spatial levels: the field, the farm, and a higher spatial level: landscape, region or nation [39]. SAFE is a hierarchical framework composed of principles, criteria, indicators, and reference values. *Principles* are general conditions for achieving sustainability, and are formulated as a general objective to be reached. *Criteria* are specific objectives, more concrete than principles and relating to a state of the system, and therefore easier to assess and to link indicators to. *Indicators* are variables of any type that can be assessed in order to measure compliance with a criterion. *Reference values* describe the desired level of sustainability for each indicator [38]. By decreasing order of preference, reference values can be based on legislative norms, scientific norms, or observations in the studied farms [39]. They can be *relative* (an average or comparison with a sector or a trend) or *absolute* (a fixed value, based on a scientific or legal source). Absolute reference values can be *target values*, which identify desirable conditions, or *threshold values* which may be expressed either as minimum or maximum levels or ranges of acceptable values, that should not be exceeded, taking the precautionary principle into account. These types of reference values can be applied in a range of spatial scales such as the field, farm, or landscape/watershed/administrative unit scale.

3. Comparison of the Methods

To compare the selected methods we first looked at their general characteristics: the object they study (farming systems or economic systems in general), their target users, their objective and, in particular, the dimensions of sustainability studied, and finally the spatial scale of the systems they study. Next we considered more specific characteristics regarding the reference values: terms used to designate them, sources used to establish them, numeric and visual approaches used to express results and the introduction of spatial differentiation.

3.1. Objects Studied and Target Users

Four methods (FESLM, MESMIS, AEMBAC and SAFE) have been designed for the assessment of farming systems, one of these (MESMIS) was specifically conceived to assess peasant farming systems (Table 1). Four methods (EF, ESM, SGAPS, SADS) have a more generic vocation, as they assess economic systems in general, SADS aims to assess socio-economic systems.

Unsurprisingly, six methods (FESLM, ESM, SGAPS, SADS, AEMBAC, SAFE) indicate decision makers among their target users; the two methods that do not refer to decision makers are more of the grass-roots type, mentioning community activists (EF) and peasant organizations (MESMIS) among their target users (Table 1). Researchers are given as target users for five methods (EF, ESM, MESMIS, AEMBAC, SAFE) and policy makers for three methods (EF, SGAPS, SADS).

Table 1. General description of assessment frameworks: object studied, target users, objective and systems studied.

Method	Object	Target users	Objective	Systems studied
(a) FESLM	Farming systems	Decision makers, farmers, land users	Evaluate the environmental, economic and social sustainability of land management	Land use systems (farm, region)
(b) EF	Economic systems	Community activists, policy makers, researchers	Assess the environmental sustainability of individuals, communities, regions, nations, humanity	Person, community, nation, world
(c) ESM	Economic systems	Decision makers, researchers	Assess environmental impacts of pollutant emissions and resource use	Products, processes, sites
(d) SGAPS	Economic systems	Decision makers, policy makers	Determine whether an economic activity is environmentally sustainable	Any activity
(e) SADS	Socio-economic systems	Decision makers, policy makers	Assess the sustainability of development strategies at the regional level	A regional socio-economic system
(f) MESMIS	Peasant farming systems	Peasant organizations, researchers, development agents	Evaluate environmental, social and economic sustainability in a systemic, participatory, interdisciplinary way.	Peasant natural resource management systems
(g) AEMBAC	Farming systems	Decision makers, researchers	Identify, develop and evaluate locally appropriate agro-environmental measures	Field, farm, landscape
(h) SAFE	Farming systems	Decision makers, researchers	Identify, develop and evaluate locally more sustainable agricultural systems, policies and techniques	Field, farm, landscape, nation

a [14], b [17], c [26], d [31], e [32], f [33], g [37], h [38].

3.2. Objective and Systems Studied

Four methods (FESLM, SADS, MESMIS, SAFE) aim to assess sustainability, considering its environmental, social and economic dimension (Table 1). The other four methods focus on the assessment of environmental sustainability. The MESMIS method assesses sustainability in a participatory way. Most methods can study systems across a wide scale or spatial range, from individual to humanity (EF) or from field to nation (SAFE). SADS is the least wide-ranging in this respect, as it focuses on a regional socio-economic system.

3.3. Terms and Sources for Reference Values

The methods use a variety of terms to designate what we have chosen here to call reference values (Table 2). The term *threshold* is used by two methods (FESLM, MESMIS), other terms used are *Fair Earthshare*, *Critical flow*, *Sustainability Standard*, *Critical Threshold Value*, *Baseline value*, *Environmental Minimum Requirement*, *Reference value*. Four of the methods differentiate two types of reference values. SGAPS distinguishes *Sustainability targets*, which are based on objective considerations derived from environmental science, from *Policy targets*, which are influenced by considerations of cost or political feasibility. AEMBAC similarly distinguishes *Science-driven* and *Society-driven* Environmental Minimum Requirements. SADS distinguishes CTV_{min} , a conservative estimate of the threshold, and CTV_{max} , the maximum allowable value of the threshold. Here, the differentiation serves to quantify the uncertainty regarding the level of the reference values. SAFE contrasts *Absolute Reference values*, which are based on a scientific or legal source and allow the assessment of a single system with a previously defined value, and *Relative Reference values*, which are based on indicator values for similar systems or a reference system.

The methods draw on a variety of sources for the establishment of reference values (Table 2). Six methods use science to establish reference values, with five methods referring to scientific literature (SGAPS, SADS, MESMIS, AEMBAC, SAFE), and one to modeling (FESLM). Expert opinion (ESM, SADS, MESMIS, AEMBAC) and or either legislation or policy targets (ESM, SGAPS, AEMBAC, SAFE) are used by six methods as a basis for reference values. Two methods use community averages (FESLM and MESMIS), or historic records (FESLM) or past situations (AEMBAC). Other sources, cited once, are: ratio of land and water area over population (EF), values for natural ecosystems and well-functioning agroecosystems (AEMBAC) and farm data (SAFE). All methods, except EF, use more than one source to establish reference values. Table 2 gives illustrative examples of reference values for the six methods for which such examples were available.

Table 2. Terms to designate reference values, sources for reference values and examples of reference values.

Method	Term to designate reference values	Sources for reference values	Example of reference value
(a) FESLM	Threshold	Community averages, historic records, modeling	<i>Yield</i> : average in the municipality plus 20%
(b) EF	Fair Earthshare	Ratio of bioproductive land and water area over population	<i>Fair earth share</i> : 2.1 ha per person
(c) ESM	Critical flow	Swiss and international policy targets, expert opinion	<i>Fresh water use</i> : 10.7 km ³ /year (20% of available freshwater)
(d) SGAPS	Sustainability Standard Sustainability Targets Policy targets	Scientific literature, legislation	<i>Atmospheric CO₂ conc.</i> : 280 ppm (pre-industrial level)
(e) SADS	Critical Threshold Value CTV _{min} CTV _{max}	Scientific literature, expert opinion	Not specified
(f) MESMIS	Baseline values/thresholds	Expert opinion, scientific literature, community averages	<i>Soil organic matter content</i> : 2%
(g) AEMBAC	Environmental Minimum Requirement Science-driven Society driven	Natural ecosystems, past situations, expert opinion, scientific literature, values for well-functioning agroecosystems, legislation	<i>Valuable biotopes</i> : 10% of total area
(h) SAFE	Reference value Absolute Relative	Scientific literature, legislation, farm data	Not specified

a [14], b [17], c [26], d [31], e [32], f [33], g [37], h [38].

3.4. Expression of Results and Spatial Differentiation

Reference values help to interpret indicator values. The methods reviewed here display a variety of numeric and visual approaches to relate impacts to reference values (Table 3). Three methods (FESLM, EF, MESMIS) express results as the ratio of the indicator value over the reference value; two of these methods also use radar graphs. ESM uses the square of the impact-reference value ratio. SGAPS and AEMBAC use the difference (*Gap*) between the indicator value and the reference value. SGAPS further quantifies the gap as *Years to sustainability*, by calculating the time necessary to reach the reference value on continuation of current trends. SADS uses color flags which reflect the value of the indicator relative to the reference value band width defined by CTV_{min} and CTV_{max}. The use of a band width rather than a single value allows the expression of uncertainty associated with the definition of the reference value. The colors convey clear messages to the method's users, e.g., green (no reason for concern) for values below CTV_{min}; and black (stop immediately further growth) for values above CTV_{max}.

Table 3. Expression of results relative to reference values, spatial differentiation of reference values.

Method	Expression of results	Spatial differentiation
(a) FESLM	Ratio of indicator value over reference value, radar graph	Reference values defined at the scale of the system studied
(b) EF	Ratio of indicator value over reference value	Reference values defined at global scale
(c) ESM	Square of the ratio of indicator value over reference value	Reference values can be regionalized, this was implemented for fresh water use
(d) SGAPS	Difference between indicator value and reference value (Sustainability gap); time to reach reference value on continuation of current trends (Years to sustainability)	Reference values defined at the scale of the system studied
(e) SADS	Color “flags” (green, yellow, red, black) indicate position of indicator value relative to reference value band width	Reference values defined at the scale of the system studied
(f) MESMIS	Ratio of indicator value over reference value, radar graph	Reference values defined at the scale of the system studied
(g) AEMBAC	Difference between indicator value and reference value (Gap)	Reference values defined at the scale of the system studied
(h) SAFE	Not specified	Reference values defined at the scale of the system studied

a [14], b [17], c [26], d [31], e [32], f [33], g [37], h [38].

Reference values may allow the introduction of spatial differentiation in environmental assessment methods. Six methods (FESLM, SGAPS, SADS, MESMIS, AEMBAC, SAFE) define their reference values at the scale of the system studied, and thus take into account the site-specific character of a desirable state of the environment. ESM uses reference values based on Swiss policy targets. The method supports spatial differentiation, which has been implemented for freshwater consumption, through the implementation of six scarcity categories. EF defines its reference value at the global level, by dividing the global bio-productive land and water area by the world population.

4. Discussion

4.1. Characteristics of Methods Reviewed

Four of the methods reviewed here were specifically designed to assess farming systems; the other four assess economic or socio-economic systems in general. The methods cover a wide range of target users, ranging from community activists and peasant organizations to researchers and policy and decision makers. Half of the methods assess environmental sustainability, the other half assess sustainability through its environmental, social and economic dimensions. With respect to systems studied, the set of methods reviewed here reveal major variability: most methods are designed to study systems across a wide spatial range, some focus on a specific spatial level. Overall, the methods reviewed here display a large diversity with respect to object studied, target users, objectives and system studied, supplying a broad basis for this review.

4.2. Classification of Reference Values

The eight methods use a total of eight terms for what we have chosen here to call Reference values. This semantic diversity may reflect the fact that reference values have so far not been a major topic in analyses of methods for environmental assessment, but have rather been treated as one among many elements in the construction of such methods. This paper thus is timely, as it takes stock of existing approaches in order to propose improvements and methodological clarity. We propose to use the generic term *Reference value* as the preferred term for “the desired level for an indicator” [38], rather than one of the other more specific terms used in the reviewed methods (*Threshold, Fair Earthshare, Critical flow, Sustainability Standard, Critical Threshold Value, Baseline Value, Environmental Minimum Requirement*). Reference values quantify the sustainable goals. Even if these goals depend on the definition of sustainability given, they make the concept of sustainability operational to stakeholders [10].

SAFE [38], following von Wirén-Lehr [40], contrasts *Absolute Reference values*, which allow the assessment of a single system with a previously defined value, and *Relative Reference values*, which are based on indicator values for similar systems or a reference system. This is a fundamental distinction, and we agree with van Cauwenbergh *et al.* [38] that it should be at the basis of a classification of reference values. However, instead of the expression *Absolute reference value*, we propose the term *Normative reference value*, since this type of values may be formulated “in a relative way” e.g., when the target for a nation’s greenhouse gas emissions is a reduction by 75%. Thus the opposition *Normative versus Relative* reference values seems more appropriate than the opposition *Absolute versus Relative*.

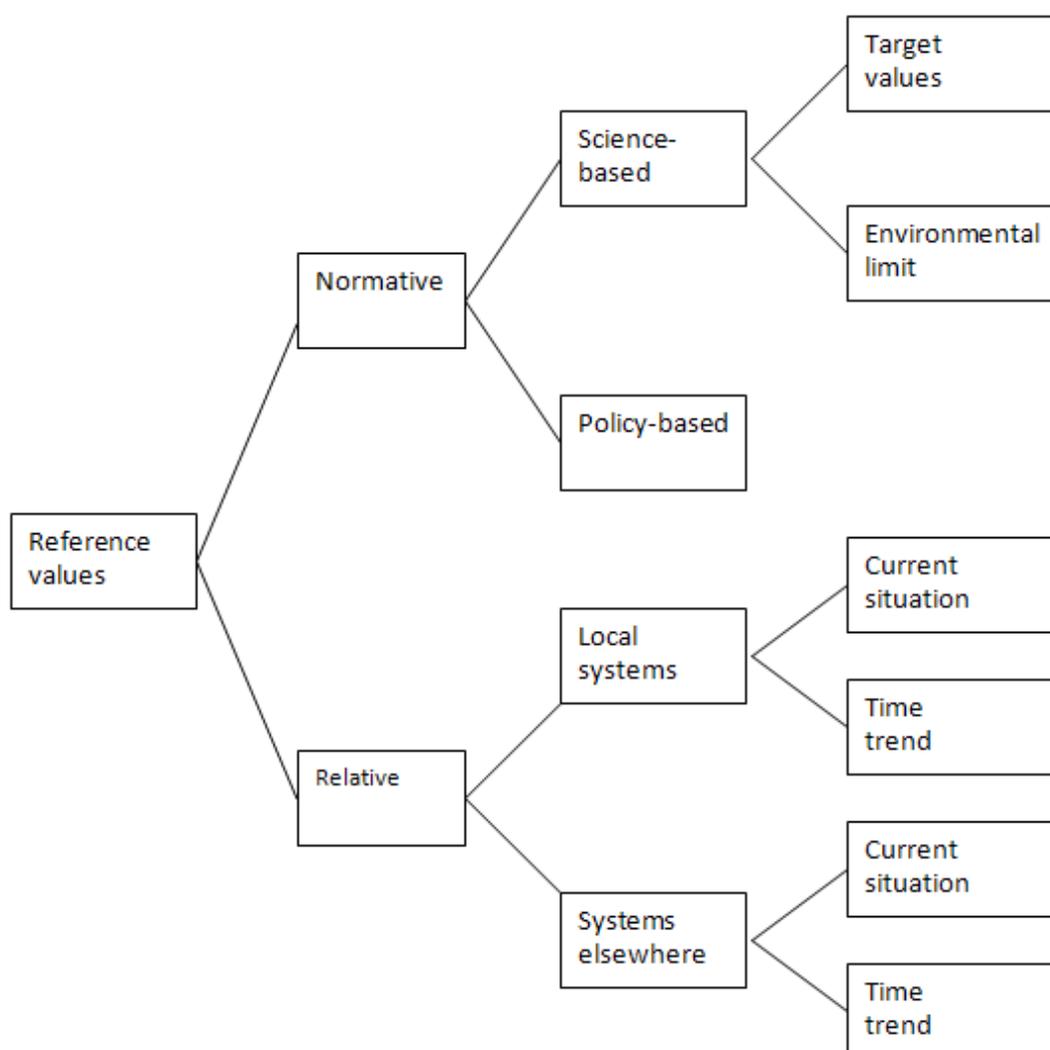
In the SAFE framework, Absolute Reference values can be based on a scientific or a legal source, but this criterion is not used in their classification of reference values. SGAPS [32], in its classification of reference values distinguishes *Sustainability targets*, based on objective considerations derived from environmental science, from *Policy targets*, which are influenced by considerations of cost or political feasibility. AEMBAC [37] distinguishes *Science-driven* and *Society-driven* Environmental Minimum Requirements. We propose the use of the terms *Science-based* and *Policy-based* to distinguish these two types of normative reference values.

The SAFE framework [38] further distinguishes two types of Absolute reference values: *Target values*, which identify desirable conditions (as proposed by Mitchell *et al.* [41]), and *Threshold values*, which may be expressed either as minimum or maximum levels, or ranges of acceptable values, that should not be exceeded. From their review of the scientific literature describing environmental limits and thresholds, Haines-Young *et al.* [42] conclude that, although the concepts have been discussed widely, the terms limits and thresholds have been applied inconsistently across different fields. They define *Environmental limit* as the level of some environmental pressure, or level of benefit derived from the natural resource system, beyond which conditions are deemed to be unacceptable in some way. The term can be applied irrespective of the type of dynamic exhibited by the system (linear response, simple non-linear response, threshold response). Haines-Young *et al.* [42] reserve the term *Threshold* to describe situations in which a distinct regime shift between alternative equilibrium regimes exists, which may or may not be reversible. The authors argue that the concept of environmental limit is more useful generally, as, while including the possibilities of system collapse

associated with the threshold concept, it focuses attention on the possibly more wide-spread, chronic or progressive loss of integrity which natural resource systems may suffer with increasing environmental pressures. Haines-Young *et al.* [42] also elaborate on the concept of *Target value*. They argue that, fundamentally, the idea of a limit involves setting a maximum level of damage to a natural resource system that we are prepared to accept. However, in management terms it might be preferable to maintain the system in “good” condition, by specifying target values that are well above the agreed limit. Based on this analysis, we propose to use the terms *Target values* and *Environmental limits* (rather than *Target values* and *Threshold values*) to distinguish these two types of science-based normative reference values.

Relative reference values can be derived from comparable *Local systems* or from *Systems elsewhere*. In both cases they can be based on the *Current situation* for these systems or on *Time trends* for these systems. The preceding propositions and terminology have been summarized in a schematic classification of reference values (Figure 1).

Figure 1. Classification of reference values.



4.3. Normative Reference Values

A variety of sources are used to establish normative reference values, with scientific literature and legislation or policy being most frequent. As discussed in the previous section, only two methods differentiate science-based reference values from policy-based reference values. We feel that any method relying on these two sources should make this distinction, as policy-based reference values usually are a compromise based on science on the one hand, and on societal considerations (cost, political feasibility) on the other. Thus, policy-based reference values might be considered as resulting from a “bottom-up” process while science-based reference values are perceived as “top-down” [10]. In consequence, policy-based reference values will generally be less strict than science-based reference values. The systematic distinction of these two types of reference values will help to reduce the uncertainty associated with reference values by reducing the heterogeneity of its sources.

It is obvious however that, depending on their implementation, the distinction between these two types of reference values can be fuzzy, as the values they yield may be close. This proximity of their values can have two causes. First the value of a science-based reference value obviously depends on the “scientific” source used, and given our imperfect understanding and the lack of consensus among scientists regarding the functioning of ecosystems this inevitably will introduce variability [43]. Secondly, policy-based reference values result from a compromise between scientific knowledge and political feasibility. Depending on the relative weight of each of these elements, the resulting compromise value will be more or less close to a reference value that would be based on science only. Policy-based reference values could be used when science-based ones are not available and should be clearly identified as such [44]. Reference values can help to improve environmental management over time [45] and guide systems towards sustainability. Reference values used should be up-dated as frequently as possible since they are defined according to present knowledge.

4.4. Expression of Results

The ratio of the indicator value over the reference value is the most common approach used (in four out of eight methods) to relate impact values to reference values. This ratio represents a simple and effective means to express results. Among these four methods, ESM represents an interesting originality, as it uses the square of the ratio value; as a result large ratios are weighted proportionately higher, relative to small ratios. The difference or *Gap* between the indicator and the reference value is used by two methods as a means of presenting results, here SGAPS proposes *Years to sustainability* as an original way to quantify results. SADS proposes a visually attractive approach to communicate results through its color flag system. This diversity of methods for the expression of results will be helpful to those interested in the implementation of reference values. All modes of expression of results encountered here seem quite straightforward and simple to interpret by users of the methods.

4.5. Expression of Uncertainty and Spatial Differentiation

Several of the papers reviewed here touch upon the question of uncertainty, and it is obvious that reference values, given the sources they are based on, will often be highly uncertain. However, among the methods reviewed here, only SADS proposes a means to capture the uncertainty of its reference value. The quantification and expression of the uncertainty of reference values clearly is a subject that warrants further work.

Reference values represent a desirable state of the environment. For many impacts, the desirable state of the environment is site-specific. Thus reference values obviously represent an interesting way to introduce spatial differentiation into methods for environmental assessment. Six of the methods reviewed here define their reference values at the scale of the system studied, and thus take into account the characteristics of the local environment, introducing spatial differentiation in the environmental assessment approach. ESM also supports spatial differentiation, by defining six scarcity categories for fresh water resources according to the region of water consumption. Reference values present a means of introducing site specificity into methods for environmental assessment which seems at present largely under-exploited.

5. Conclusions

The analysis of the use of reference values in the eight methods for environmental assessment reviewed here, has revealed a striking diversity of terminology, sources, and modes of expression of results. Based on this analysis, we formulate the following recommendations for the implementation of reference values in environmental assessment methods.

- Recommendations on terminology
 - The term *Reference value* should be used as a generic term for the desired level for an indicator.
 - *Normative reference values* allow the assessment of a single system with a previously defined value, *Relative reference values* are based on indicator values for similar systems or a reference system.
 - Normative reference values can be *Science-based* or *Policy-based*. Making this distinction explicit will contribute to reducing the uncertainty associated with reference values.
 - A science-based normative reference value can be a *Target value*, which identifies desirable conditions, or an *Environmental limit*, which is the level of some environmental pressure, or benefit, from the natural resource system, beyond which conditions are unacceptable.
- Other recommendations
 - Methods should make a clear distinction between science-based reference values and policy-based reference values, as policy-based reference values usually are a compromise based on science on the one hand, and on societal considerations (cost, political feasibility) on the other. Thus policy-based reference values will generally be less strict than science-based reference values.

- From the academic point of view science-based reference values are obviously preferable, from a practical point of view policy-based reference values, when available, will be easier to implement, as they incorporate the results of difficult choices, outside the domain of science, made by public decision makers.
- The quantification of the uncertainty of reference values is a topic which is barely explored and warrants further research.
- Reference values present a means of introducing site specificity into methods for environmental assessment which seems at present largely under-exploited.

Acknowledgements

This research was financially supported by the French National Research Agency, SPA/DD project “Animal Production Systems and Sustainable Development” (ANR-06-PADD-017) and the ADEME, the French Environment and Energy Management Agency.

References

1. Pretty, J. Agricultural sustainability: concepts, principles and evidence. *Phil. Trans. R. Soc. B* **2008**, *363*, 447-465.
2. Fischer, J.; Manning, A.D.; Steffen, W.; Rose, D.B.; Daniell, K.; Felton, A.; Garnett, S.; Gilna, B.; Heinsohn, R.; Lindenmayer, D.B.; MacDonald, B.; Mills, F.; Newell, B.; Reid, J.; Robin, L.; Sherren, K.; Wade, A. Mind the sustainability gap. *Trends Ecol. Evol.* **2007**, *22*, 621-624.
3. von Wieren-Lehr, S. Sustainability in agriculture—an evaluation of principal goal-oriented concepts to close the gap between theory and practice. *Agr. Ecosyst. Environ.* **2001**, *84*, 115-129.
4. van der Werf, H.M.G.; Petit, J. Evaluation of the environmental impact of agriculture at the farm level: A comparison and analysis of 12 indicator-based methods. *Agr. Ecosyst. Environ.* **2002**, *93*, 131-145.
5. Payraudeau, S.; van der Werf, H.M.G. Environmental impact assessment for a farming region: A review of methods. *Agr. Ecosyst. Environ.* **2005**, *107*, 1-19.
6. Galan, M.B.; Peschard, D.; Boizard, H. ISO 14001 at the farm level: Analysis of five methods for evaluating the environmental impact of agricultural practices. *J. Environ. Manage.* **2007**, *82*, 341-352.
7. Rossing, W.A.H.; Zander P.; Josien E.; Groot J.C.J.; Meyer, B.C.; Knierim, A. Integrative modelling approaches for analysis of impact of multifunctional agriculture: A review for France, Germany and The Netherlands. *Agr. Ecosyst. Environ.* **2007**, *120*, 41-57.
8. Van der Werf, H.M.G.; Tzivilakis, J.; Lewis, K.; Basset-Mens, C. Environmental impacts of farm scenarios according to five assessment methods. *Agr. Ecosyst. Environ.* **2007**, *118*, 327-338.
9. Bockstaller, C.; Guichard, L.; Makowski, D.; Aveline, A.; Girardin, P.; Plantureux, S. Agri-environmental indicators to assess cropping and farming systems. A review. *Agron. Sustain. Dev.* **2008**, *28*, 139-149.

10. Binder, C.R.; Feola, G.; Steinberger, J.K. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. *Environ. Impact Assess. Rev.* **2010**, *30*, 71-81.
11. Girardin, P.; Bockstaller, C.; van der Werf, H. Indicators: Tools to Evaluate the Environmental Impacts of Farming Systems. *J. Sustain. Agr.* **1999**, *13*, 5-21.
12. Gomontean, B.; Gajaseni, J.; Edwards-Jones, G.; Gajaseni, N. The development of appropriate ecological criteria and indicators for community forest conservation using participatory methods: A case study in northeastern Thailand. *Ecol. Indic.* **2008**, *8*, 614-624.
13. Bare, J.; Udo de Haes, H.; Pennington, D. *An International Workshop on Life Cycle Assessment Sophistication*; U.S. Environmental Protection Agency: Cincinnati, OH, USA, 2000.
14. Smyth, A.J.; Dumanski, J. *FESLM: An International Framework for Evaluating Sustainable Land Management*; Report No. 92-5-103419-2; Food and Agriculture Organization of the United Nations: Rome, Italy, 1993.
15. *A Framework for Land Evaluation*; Report No. ISBN 92-5-100111-1; Food and Agriculture Organization of the United Nations: Rome, Italy, 1976.
16. Gomez, A.A.; Swete Kelly, D.E.; Syers, K.J.; Coughlan, K.J. Measuring sustainability of agricultural systems at the farm level. In *Methods for Assessing Soil Quality*; Doran, J.W., Jones, A.J., Eds.; Soil Science Society of America: Madison, WI, USA, 1996; SSSA Special Publication 49, pp. 401-409.
17. Wackernagel, M.; Rees, W. *Our Ecological Footprint. Reducing Human Impact on the Earth*; New Society: Gabriola Island, Canada, 1996.
18. Global Footprint Network. *Global Footprint Network. Advancing the science of sustainability*, 2007. Available online: <http://www.footprintnetwork.org> (accessed on 29 July 2009).
19. van den Bergh, J.C.J.M.; Verbruggen, H. Spatial sustainability, trade and indicators: an evaluation of the 'ecological footprint'. *Ecol. Econ.* **1999**, *29*, 61-72.
20. Kitzes, J.; Peller, A.; Goldfinger, S.; Wackernagel, M. Current methods for calculating national ecological footprint accounts. *Sci. Environ. Sustain. Soc.* **2007**, *4*, 9.
21. Knaus, M.; Lohr, D.; O'Regan, B. Valuation of ecological impacts—a regional approach using the ecological footprint concept. *Environ. Impact Assess. Rev.* **2006**, *26*, 156-169.
22. Lenzen, M.; Murray, S.A. A modified ecological footprint method and its application to Australia. *Ecol. Econ.* **2001**, *37*, 229-255.
23. Moran, D.D.; Wackernagel, M.; Kitzes, J.A.; Goldfinger, S.H.; Boutaud, A. Measuring sustainable development—Nation by nation. *Ecol. Econ.* **2008**, *64*, 470-474.
24. Wackernagel, M. An evaluation of the ecological footprint. *Ecol. Econ.* **1999**, *31*, 317-318.
25. Wackernagel, M.; Monfreda, C.; Moran, D.; Wermer, P.; Goldfinger, S.; Deumiling, D.; Murray, M. *National Footprints and Biocapacity Accounts 2005: The Underlying Calculation Method*; Global Footprint Network: Oakland, CA, USA, 2005.
26. Frischknecht, R.; Steiner, R.; Jungbluth, N. *The Ecological Scarcity Method—Eco-factors 2006. A Method for Impact Assessment in LCA*; Environmental Studies no. 0906; Federal Office for the Environment: Bern, Switzerland, 2009.
27. Braunschweig, A.; Bär, P.; Rentsch, C.; Schmid, L.; Wüest, G. *Methode der ökofaktoren 1997*; BUWAL: Bern, Switzerland, 1998.

28. Braunschweig, A.; Müller-Wenk, R. *Oekobilanzen für Unternehmungen, Eine Wegleitung für die Praxis*; Paul Haupt: Bern, Switzerland, 1993.
29. Scharlemann, J.P.W.; Laurance, W.F. Environmental science—How green are biofuels? *Science* **2008**, *319*, 43-44.
30. Zah, R.; Böni, H.; Gauch, M.; Hirschler, R.; Lehmann, M.; Wäger, P. *Ökobilanz von energieprodukten: Ökologische bewertung von biotreibstoffen*; Empa Abteilung Technologie und Gesellschaft: Gallen, Switzerland, 2008.
31. Ekins, P.; Simon, S. Estimating sustainability gaps: Methods and preliminary applications for the UK and the Netherlands. *Ecol. Econ.* **2001**, *37*, 5-22.
32. Nijkamp, P.; Vreeker, R. Sustainability assessment of development scenarios: methodology and application to Thailand. *Ecol. Econ.* **2000**, *33*, 7-27.
33. Lopez-Ridaura, S.; Masera, O.; Astier, M. Evaluating the sustainability of complex socio-environmental systems. The MESMIS framework. *Ecol. Indic.* **2002**, *2*, 135-148.
34. Lopez-Ridaura, S.; van Keulen, H.; van Ittersum, M.K.; Leffelaar, P.A. Multiscale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. *Environ. Dev. Sustain.* **2005**, *7*, 51-69.
35. Brunett Pérez, L.; González Esquivel, C.; García Hernández, L.A. Evaluación de la sustentabilidad de dos agroecosistemas campesinos de producción de maíz y leche, utilizando indicadores. *Livest. Res. Rural Dev.* **2005**, *17*, article number 78. Available online: <http://www.lrrd.org/lrrd17/7/pere17078.htm> (accessed on 8 May 2010).
36. Simoncini, R. *The AEMBAC Project: Final Report*; IUCN: Brussels, Belgium, 2004. Available online: http://www.nandoperettifound.org/documents/2000_04_AEMBAC_Final.pdf (accessed on 20 August 2010).
37. Bastian, O.; Corti, C.; Lebboroni, M. Determining environmental minimum requirements for functions provided by agro-ecosystems. *Agron. Sustain. Dev.* **2007**, *27*, 279-291.
38. van Cauwenbergh, N.; Biala, K.; Bielders, C.; Brouckaert, V.; Franchois, L.; Garcia Ciudad, V.; Hermy, M.; Mathijs, E.; Muys, B.; Reijnders, J.; Sauvenier, X.; Valckx, J.; Vanclooster, M.; Van der Veken, B.; Wauters, E.; Peeters, A. SAFE—A hierarchical framework for assessing the sustainability of agricultural systems. *Agr. Ecosyst. Environ.* **2007**, *120*, 229-242.
39. Sauvenier, X.; Valckx, J.; Van Cauwenbergh, N.; Wauters, E.; Bachev, H.; Biala, K.; Bielders, C.; Brouckaert, V.; Garcia Ciudad, V.; Goyens, S.; Hermy, M.; Mathijs, E.; Muys, B.; Vanclooster, M.; Peeters, A. *Framework for Assessing Sustainability Levels in Belgian Agricultural Systems—SAFE*; Belgian Science Policy: Brussels, Belgium, 2006. Available online: http://www.belspo.be/belspo/home/publ/pub_ostc/CPagr/rappCP28_en.pdf (accessed on 7 December 2009).
40. Von Wihren-Lehr, S. Sustainability in agriculture—an evaluation of principal goal-oriented concepts to close the gap between theory and practice. *Agr. Ecosyst. Environ.* **2001**, *84*, 115-129.
41. Mitchell, G.; May, A.; McDonald, A. PICABEU: A methodological framework for the development of indicators of sustainable development. *Int. J. Sustain. Dev. World Ecol.* **1995**, *2*, 104-123.

42. Haines-Young, R.; Potschin, M.; Cheshire, D. *Defining and Identifying Environmental Limits for Sustainable Development. A Scoping Study*; Final Overview Report to Defra, Project code NR0102; Centre for Environmental Management, School of Geography, University of Nottingham: Nottingham, UK, 2006.
43. Muradian, R. Ecological thresholds: A survey. *Ecol. Econ.* **2001**, *38*, 7-24.
44. Walter, C.; Stützel, H. A new method for assessing the sustainability of land-use systems (II): Evaluating impact indicators. *Ecol. Econ.* **2009**, *68*, 1288-1300.
45. Groffman, P.; Baron, J.S.; Blett, T.; Gold, A.J.; Goodman, I.; Gunderson, L.H.; Levinson, B.M.; Palmer, M.A.; Paerl, H.W.; Peterson, G.D.; LeRoy Poff, N.; Rejeski, D.W.; Reynolds, J.F.; Turner, M.G.; Weathers, K.C.; Wiens, J. Ecological Thresholds: The Key to Successful Environmental Management or an Important Concept with No Practical Application? *Ecosystems* **2006**, *9*, 1-13.

© 2011 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).