seamless

System for Environmental and Agricultural Modelling; Linking European Science and Society

Review of approaches to establish reference levels to interpret indicators

Van der Heide, C.M., Brouwer, F., Bellon, S., Bockstaller, C., Garrod, G., et al.

Partners involved: LEI, INRA, UEVORA, UNEW, UNIABDN

Report no.: 30
August 2007
Ref: PD2.5.1
ISBN no.: 90-8585-118-1 and 978-90-8585-118-9

LEI WAGENINGEN UR

INRA Institut National de la Recherche Agronomique

Logo's main partners involved in this publication

Sixth Framework Programme

SEAMLESS integrated project aims at developing an integrated framework that allows ex-ante assessment of agricultural and environmental policies and technological innovations. The framework will have multi-scale capabilities ranging from field and farm to the EU25 and globe; it will be generic, modular and open and using state-of-the art software. The project is carried out by a consortium of 30 partners, led by Wageningen University (NL).

Email: seamless.office@wur.nl
Internet: www.seamless-ip.org

Authors of this report and contact details:

Name: Martijn van der Heide Partner acronym: LEI
Address: Agricultural Economics Research Institute
P.O. Box 29703, 2502 LS The Hague, the Netherlands
E-mail: martijn.vanderheide@wur.nl

Name: Floor Brouwer Partner acronym: LEI
Address: Agricultural Economics Research Institute
P.O. Box 29703, 2502 LS The Hague, the Netherlands
E-mail: martijn.vanderheide@wur.nl

Name: Stéphane Bellon Partner acronym: INRA
Address: INRA Ecodéveloppement, Site Agroparc, Domaine St Paul,
84914 Avignon Cedex 9, France
E-mail: bellon@avignon.inra.fr

Name: Christian Bockstaller Partner acronym: INRA
Address: UMR INPL-(ENSAIA)-INRA Agronomie et Environnement Nancy-Colmar
BP 20507 68021 Colmar Cedex, France
E-mail: bockstal@colmar.inra.fr

Name: Guy Garrod Partner acronym: UNEW
Address: Centre for Rural Economy, School of Agriculture, Food and Rural Development, The University of Newcastle-upon-Tyne, NE1 7RU, United Kingdom
E-mail: L.M.Stapleton@ncl.ac.uk

Name : Ghislain Geniaux Partner acronym: INRA
Address: INRA Ecodéveloppement, Site Agroparc, Domaine St Paul,
84914 Avignon Cedex 9, France
E-mail: geniaux@avignon.inra.fr

Name: Rosário Oliveira Partner acronym: UEvora
Address: University of Évora
Largo dos Colegiais 2, 7000 Évora
E-mail: mrgo@uevora.pt

Name: Pete Smith Partner acronym: UABDN
Address: School of Biological Sciences, University of Aberdeen, Cruickshank Building, St Machar Drive, Aberdeen, AB24 3UU, United Kingdom
E-mail: pete.smith@abdn.ac.uk

Name: Lee Stapleton Partner acronym: UNEW
Address: Centre for Rural Economy, School of Agriculture, Food and Rural Development, The University of Newcastle-upon-Tyne, NE1 7RU, United Kingdom
E-mail: L.M.Stapleton@ncl.ac.uk

Name: Emmanuelle Weinzaepflen Partner acronym: INRA
Address: UMR INPL-(ENSAIA)-INRA Agronomie et Environnement Nancy-Colmar
BP 20507 68021 Colmar Cedex, France
E-mail: weinzaep@colmar.inra.fr

Name: Chengyi Zhang Partner acronym: UABDN
Address: School of Biological Sciences, University of Aberdeen, Cruickshank Building, St Machar Drive, Aberdeen, AB24 3UU, United Kingdom
E-mail: c.zhang@abdn.ac.uk

Disclaimer 1:

“This publication has been funded under the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2). Its content does not represent the official position of the European Commission and is entirely under the responsibility of the authors.”

Disclaimer 2:

Within the SEAMLESS project many reports are published. Some of these reports are intended for public use, others are confidential and intended for use within the SEAMLESS consortium only. As a consequence references in the public reports may refer to internal project deliverables that cannot be made public outside the consortium.

When citing this SEAMLESS report, please do so as:

Van der Heide M., Brouwer, F.M., Bellon, S., Bockstaller, C., Garrod, G. et al., 2007. Review of approaches to establish reference levels to interpret indicators,, SEAMLESS Report No.30, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 44 pp, ISBN no. 90-8585-118-1 and 978-90-8585-118-9.

Table of contents

General information

Task(s) and Activity code(s): Task 2.5, Activity 2.5.1

Input from (Task and Activity codes):

Output to (Task and Activity codes): Task 2.6

Related milestones:

Executive summary

The report clarifies the terminology relevant for understanding ‘reference levels’ for the various domains of sustainable development. Based on a literature review, the report explores what interpretation is given to its meaning. This offers a basis using the terminology in the context of SEAMLESS.

The main principles guiding the definition of reference levels are identified in Section 2: ‘irreversibility’, ‘precautionary principle’, ‘risk’, ‘safe minimum standards’, ‘trade-off analysis’ and ‘uncertainty’. Terminology on reference levels is very much linked to the environmental domain and much less so to the social and economic domains. Also, absolute and relative levels of ‘reference levels’ are explored. Relative performance or benchmark has gained interest in the scientific literature. Trade-offs may exist and choices may have to be made in practical cases

The concepts of ‘reference levels’ are explained in real world cases in Section 3, addressing

- target thresholds and the precautionary principle in the context of environmental indicators like the application of nitrogen;
- threshold values and targets in the context of economic indicators like budgetary expenditures and total welfare;
- threshold values and targets in the context of social indicators like accessibility to service centres.

Some remaining methodological issues on reference levels are further discussed in Section 4, and the consequences of this report for SEAMLESS are examined in Section 5. There might be disagreement among scientists to establish an absolute reference level, and it remains to be clarified which are the (dis)advantages of using absolute versus relative threshold values.

Specific part

1 Introduction

In order to interpret indicators and assess the impact of policy and behavioural changes and innovations in agriculture and agroforestry adequate reference levels are crucial. These reference levels are thus necessary for practical assessment and comparison purposes. Comparison of indicators with pre-determined reference levels can, for example, be used to assess progress towards a defined goal. As already briefly indicated in the Description of Work (p. 58), "… reference levels are usually set as normative margins beyond which significant changes are thought likely to occur." In fact, the Description of Work lists three types of reference levels: thresholds, critical values and critical ranges, and target values.

Interestingly, the definitions of these three types of reference levels – in particular the threshold category – seem to refer mainly (and explicitly) to ecological systems and concerns. In the remainder of this deliverable, the various types of reference levels, their differences, as well as related concepts will be described and discussed in detail. To this end, the present deliverable aims to go beyond ecological considerations to address broader aspects of sustainability. For example, Jared Diamond's "*Collapse: How Societies Choose to Fail or Succeed*" (2005) tells the dramatic decline of societies. He shows in a magnificent style that these societies did not slowly fall apart; after reaching certain points, they suffered drastic reductions in population and productivity. As such, Diamond suggests that mismanagement of renewable natural resources is only part of why societies collapse.[1] There are also additional causes, which are more social and economic than ecological (such as institutional and cultural failure and relations with outsiders – i.e. individuals from outside the society). This knowledge is relevant for the SEAMLESS project, and in particular for the specific issue of reference levels for indicators.

The aim of this paper, which is defined in the Description of Work as PD2.5.1, is twofold. First, it introduces the concept of reference level as well as concepts that can guide the fixation of these reference levels, such as standards, precautionary principle, and risks and uncertainty. Its second purpose is to review existing (scientific) information available on the setting of reference levels for environmental, economic and social indicators. More specifically, by reviewing approaches to establishing critical ranges and reference levels, the paper attempts to lay a basis for the other activities within Task 2.5.

We approach our task by examining the relevant concepts from different perspectives. For a multidisciplinary project as SEAMLESS it is important to understand how the relevant scientific disciplines contribute to the work on indicators. Not surprisingly, therefore, that we here focus on reference levels that are relevant for the environmental, economic and social domains of sustainable development. From a scientific perspective, these domains often correspond with different scientific disciplines and it appears that each

[1] By collapse, Diamond means *unsustainable* trajectories that precipitously fall.

discipline has developed its own specific ideas on the concepts and characteristics of reference levels and critical ranges. As a consequence, the same words have often different meanings to different experts. Put differently, some terms are defined and treated in different ways by different disciplines, although, of course, scientific disciplines and dimensions of sustainability do not coincide exactly. Moreover, some terms are more common, or more firmly rooted, in one discipline than in another.

In order to unravel this Gordian knot of complex concept descriptions, we intend to proceed in four steps. These steps, together with the content and organisation of this paper are shown graphically in Figure 1. Each of these four steps is motivated and explained in the following sections.

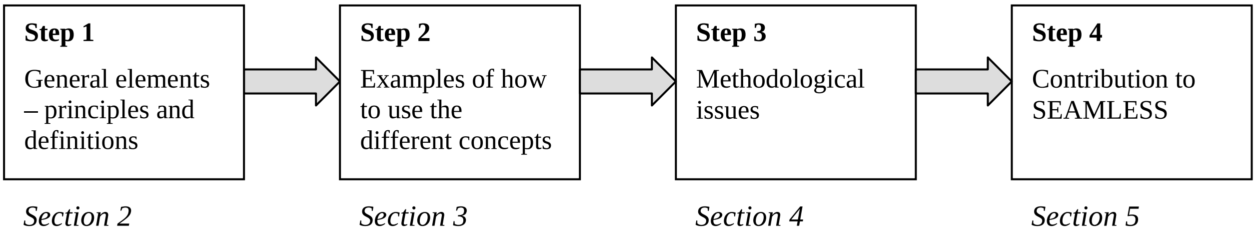


Figure 1 Various steps to achieve a SEAMLESS approach on how to develop and set reference levels

The structure of this PD is as follows. Section 2 begins with the description of principles and concepts, such as uncertainty, irreversibility, and precautionary principle, which guide the definition of the various types of reference levels. In the second part of this section, the reference levels themselves are described and defined. To illustrate the various definitions, we present in section 3 for each thematic indicator one example. That is, we present examples from the fields of agronomics and environmental sciences, economics, and social sciences. One of the examples deals with agri-environmental measures (and in particular with the application of nitrogen), which reward farmers who choose to go beyond what is legally required. As we will see, agri-environmental measures can be easily linked to the most relevant concepts that need to be understood. In addition, this section briefly explores interfaces between the environmental, economic and social dimensions. Then, in section 4, we focus on methodological issues. The main issues that are addressed in this section are reference levels and scale issues, reference levels and the so-called ‘local context’, how to establish reference values, what are those values, and who decides on what they are. Section 5 then examines the consequences of this PD for the SEAMLESS project. That is, it describes the relation with other PD’s within WP2, and a few words will be spent on the practical implementation of reference levels in the GUI. In this section, also the question of using relative or absolute reference values will be dealt with. Brief conclusions complete this section.

2 General elements

This section deals with the special terms or expressions that are considered most relevant to the activities within Task 2.5. To this end, this section describes and discusses these terms and expressions from the perspectives of practitioners in the fields of ecological, social and economic sciences. The scope of this section is thus rather broad, in that it includes the concept descriptions from various points of view. As such, we try to contextualise the core concepts that are central to Task 2.5.

As a prelude to the definition of the various types of reference levels, this section begins with the principles and concepts that steer, and are relevant for, the definition, fixation, and setting of the various types of reference levels (subsection 2.1). In alphabetical order, these principles are: irreversibility, precautionary principle, risk, safe minimum standard, trade-off analysis, and uncertainty. In subsection 2.2 the reference levels themselves are defined. That is, the subsection describes the various types of concepts that are embraced by the umbrella term of reference level. Here, the most relevant concepts are: benchmarks, critical ranges, norm, reference level, standard, target, target value, and threshold. Finally, in subsection 2.3, we say something about the difference between a single-indicator approach and a composite indicator approach. Moreover, attention will be paid to the processes that are relevant to derive the reference levels. In this subsection, we also introduce the difference between relative and absolute reference levels. A fourth issue addressed in this subsection deals with the application of the various principles, concepts and types of reference levels among the various scientific disciplines.

2.1 Principles that guide the definition of reference levels

In this subsection, some principles and concepts are presented to help the definition, fixation and establishment of the concepts that are described in subsection 2.2. We believe these principles and concepts are needed to support SEAMLESS participants in dealing with the construction of reference levels. Moreover, these principles and concepts can also help those who decide on what these levels are.

Irreversibility

Irreversibility occurs when a system cannot return to the initial state from the alternate state. Thus, irreversibility can be defined as the inability to return a system to a previous desired state given anticipated possible effects, where the state is described by key elements, (i.e. species, population or productivity). In fact, all living systems – be they ecological, economic or social – can be irreversible. Usually, irreversibility relates to environmental resources that cannot be replaced, or which could not be restored.

For environmental damage, the irreversibility (of the damage) often refers to the permanent loss of environmental assets or environmental quality, requiring preventive action rather than restoration or clean-up within a time scale for human’s observation.

Precautionary principle

Although there is no universally accepted definition of the precautionary principle, the term is generally used to describe the idea that if the consequences of an action are unknown, but are judged to have some potential for major or irreversible negative consequences, then it is better to avoid that action (see Gollier et al., 2000). As such, the precautionary principle is designed to be invoked when the level of scientific uncertainty about the consequences or likelihood of the risk is such that the best available scientific advice cannot assess the risk with sufficient confidence to inform decision making. The principle particularly relates to ecological uncertainty – for example, evolution of ecosystems, climate change, introduction of new chemicals, and loss of biodiversity – rather than to economic uncertainty – for example, business cycles and macroeconomic stability (van den Bergh, 2001; Pindyck, 2000).

The precautionary principle implies that where significant or irreversible ecological risks are involved, any lack of scientific evidence with respect to cause and effect should not be used as a reason for avoiding taking appropriate action to prevent ecological degradation. For instance, a precautionary approach to biodiversity loss would involve measures to reduce habitat fragmentation, despite uncertainty about the exact extinction rates due to the fragmentation process, or about the (cumulative) effects of species loss on the benefits that human populations derive, directly or indirectly, from them (van der Heide, 2005). Applying the precautionary principle is essentially a matter of making assumptions about consequences and likelihoods to establish credible scenarios, and then using standard procedures of risk assessment and management in order to address the hazard or threat.

Risk

A simple definition of a risk is a problem that could cause some loss or threaten the stability, viability or the vulnerable assets of a system, but which has not happened yet. From a financial point of view, a risk also refers to the chance that an investment – as a stock or commodity – will lose value, or the chance of non-payment of a debt. But in the context of the SEAMLESS project, a risk refers rather to the possibility, or probability, of suffering harm or loss. As such, a risk is related to the human expectations as it denotes a potential negative impact to an asset or some characteristic of value that may arise from some present process or from some future event.

Safe minimum standard

Making sure that a natural resource does not fall below a specified danger zone may well be the minimum goal, or safe minimum standard (SMS), of a sustainable agricultural policy. The SMS rule focuses in particular on renewable resources with a critical range such as species, scenic resources, and storage capacity of groundwater basins. It is supposed that once the resource is reduced below a critical range, it cannot be raised again due to irreversibility. With regard to renewable resources with a critical range, use decisions are particularly difficult, as the critical zone is often uncertain and dependent on biological and socioeconomic factors. Note that soil, which is a resource directly relevant to the

SEAMLESS-project, is sometimes regarded as a non-renewable resource and sometimes as a renewable resource. In fact, soil is a renewable resource, but only in the long term. That is, it is only renewable over long periods of time, measured not in days or years but in decades and even centuries.

SMS's are traditionally linked to the area of nature conservation in order to prevent as best as possible major irreversibilities (Perrings et al., 1992; Randall and Farmer, 1995; Crowards, 1998). An SMS approach to nature conservation represents a decision-making principle which suggests that there be a presumption in favour of not harming the natural environment unless the costs of that action are intolerably high (Randall and Farmer, 1995; OECD, 1999). Some argue this concept, which was introduced by Ciriacy-Wantrup in the 1950s and adopted and revitalised by Bishop (1978) in the 1970s, bridges the gap between economists and ecologists (see Spash, 1999). The SMS defines the level of preservation that ensures survival and implies a conservative approach to risk bearing (Randall, 1988). In effect, deciding to conserve today can be shown to be the risk-minimising way to proceed given the presence of uncertainty about the consequence of nature loss (Tisdell, 1991; Hanley et al., 1997). Due to scientific uncertainty about the consequences of using natural assets, an SMS approach shifts the burden of proof from those who wish to conserve to those who wish to develop (Randall and Farmer, 1995; Norton and Toman, 1997). The SMS approach is related to the precautionary principle, but it permits more scope for economic development. The barriers to economically rational actions that threaten the natural environment are under an SMS lower than when the precautionary principle is adopted (Wills, 1997, van Kooten and Bulte, 2000, van der Heide, 2005).

Trade-off analysis

Trade-off analysis is a system approach to balancing the trade-offs between time, cost and performance. The term is often used as an umbrella term for various valuation techniques, such as conjoint analysis, choice modelling and contingent valuation. What these techniques have in common is that they are decision-support (rather than decision-making) tools. They support the decision maker in 'choice situations', for example when he or she has to choose between various (opposite) policy measures. Trade-off analysis is a collection of standard statistical techniques that provide objective insight into consumer preferences using a quantifiable and repeatable approach. Trade-off analysis is, of course, not restricted to economic or material concerns. It also allows deeper analysis of trade-offs between, for example, different economic and environmental indicators. This view is supported by the increasing number of conjoint analysis studies in the field of environmental and ecological economics (see, for example, Baarsma, 2003; Sayadi *et al.*, 2005).

Uncertainty

Uncertainty is expressed as a degree to which a value (e.g. a predictive output of a model) is unknown. As such, the term uncertainty applies to predictions of future events or to the unknown. Uncertainty can result from lack of information or from disagreement about what is known or even knowable. It may have many types of sources, from quantifiable errors in the data to ambiguously defined concepts or terminology, or uncertain projections of human's perturbation to natural environment. Uncertainty can therefore be represented by quantitative measures (e.g. a range of values calculated by various models) or by qualitative statements (e.g. reflecting the judgement of a team of experts).

2.2 Different types of reference levels

Reference levels come in a wide variety of names (benchmark, standard, trend, threshold, norm); but all refer to a comparison to which an indicator can be examined or gauged. The reference value gives a point of reference to help interpret what we know about an indicator; to force discussion about what the measurement of an indicator is telling us; to help us assess whether we are moving in the desired direction and at the right pace; and to help identify what other things interact with or are affected by that indicator.

A variety of different terms are used to describe reference values. There is little consistency in the use of the terms, and they are not necessarily mutually exclusive. This subsection considers in some detail the various types of reference levels.

Benchmarks

Points of reference against which a measurement can be made and against which others may judge progress. Benchmarks can be quantitative or qualitative, input or outcome, short-term or long-term. The use of the term benchmark is fairly broad and may encompass a range of other kinds of reference values. Some view benchmark conditions as a set of intermediate conditions or points along the way to the desired future condition.

Critical range

The concept of critical range is typically related to the ecological and agronomic disciplines. As such, one can think of a critical range as the level of a nutrient below which crop yield, quality or performance is ‘unsatisfactory’. Within economics and other social sciences, the term critical range is hardly or even not in use. In Appendix A, we specify an explicit link between critical ranges and the various functions in the landscape. In other words, in this appendix, we describe three models – each of which focuses primarily on landscape and landscape changes – that represent different approaches for combining sustainability concepts with the need of thresholds definitions. In statistics, the critical range can be regarded as the values within which the calculated value of the test statistic falls when the Null Hypothesis is rejected.

UNFCCC (United Nations Framework Convention on Climate Change) has set that the greenhouse gas to be stabilized in atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. Such critical a level, should be achieved within a time-frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened and to enable economic development to proceed in a sustainable manner.

In social sciences, the concept of critical range has not been widely used, although the concept can be useful in certain contexts, as it is in the case of the marginalization process. Intuitively, we – i.e., the authors of this PD – define a critical range then as the level of marginalization in relation to, for example, age composition, education, or the number of people below which a community has no capacity to resist a process of decay and decline. Below the critical range, the community has no means or facilities to react to incentives in order to renovate or increase dynamics.

Norm

In sociology, a norm, or social norm, is a rule that is socially enforced. As such, it is closely associated with behaviour. After all, social norms are shared standards about what is appropriate or acceptable behaviour in a specific context and at a particular time. Here, however, we consider a norm as a value considered appropriate for a certain indicator. A norm affirms how things should or ought to be.

Reference level

In general, a reference level has a generic meaning, namely a level to which the indicator is referred. However, in a more specific meaning, a reference level relates to the level of ecological, social or economic 'quality', based on scientific evidence. Above the reference level, there are demonstrated effects on the ecological, social or economic system. Reference levels provide a basis for establishing goals for management and policy making. Following the evaluation and selection of the applicable scientific literature, integration of knowledge, and characterisation of exposure and risk, the reference level is determined. If not mentioned otherwise, the term reference level is used throughout the text to express the general meaning, i.e. to describe a level to which indicators are referred.

Standard

A standard can be defined as a quantitative or qualitative value or level of achievement with respect to a specific indicator which represents acceptable performance. Loosely interpreted, a standard is often used synonymously with 'performance specifications'. A standard is usually established by an authority as a rule for the measure of quantity, quality, extent or value (see 'merriam-webster online dictionary'). Standards are negotiated and they evolve usually over time.[2] Standards and norms are closely related and used interchangeably. For example, formal norms can be defined as official standards or laws that govern behaviour. Social norms are perceived standards of acceptable attitudes and behaviors prevalent among members of a community.

[2] In the longer run, also evolving indicators, the so-called 'vectorial' indicators can be included. The direction of their evolvement can also be meaningful (and not only their position as compared with a reference value).

Target
A target is defined by stakeholders or policy makers; it is a goal of improvement of an indicator result. For example, a target can be to reduce or increase the indicator value in a rate of x % in a given time. In this case, the reference is the initial raw value of the indicator. A target can be defined when it is difficult to find a norm.

Target value
Intuitively, a target value is a value specifically defined as a target for a certain indicator. It constitutes, or represents, the goal of a particular (policy) measure. The difference between a threshold and a target value is clear: while thresholds represent lines policy makers, or other stakeholders and decision makers, do not wish to cross, target values represent goals and levels that they do in fact wish to reach.

Threshold
Following PD2.6.1, we could interpret the definition of a threshold from PD2.6.1 as a minimum value for (stakeholder) acceptability which, as a concept, is applicable across environmental, social and economic dimensions of sustainable development. Here, however, a threshold is generally referred to as a point between alternate regimes in a (ecological, economic, or social) system. When the indicator reaches a value above or below the threshold, the situation of the system significantly changes. For an ecological threshold, a more specific definition is available: "the point at which there is an abrupt change in an ecosystem quality, property or phenomenon, or where small changes in an environmental driver produce large responses in the ecosystem." (Groffman *et al.*, 2006).

Unfortunately, the term threshold is not as straightforward as it seems. For example, in terms of economic indicators it may not be the case that above or below a given threshold there are significant changes, at least in the short-term. Conversely, significant changes (i.e. due to nonlinearities associated with particular indicators) above or below a particular threshold are easy to envisage for certain environmental indicators both for the short and long-term. Despite this ease of detection, analysis of thresholds is nevertheless complicated by nonlinear dynamics and by multiple factor controls that operate at diverse spatial and temporal scales. Concerning the social dimension, thresholds are difficult to define and not much elaborated on in the literature. Nevertheless, their identification could be a relevant contribution to the identification of the critical ranges that may be decisive in social terms.

In general, a threshold value comes from scientific knowledge or experts point of view (von Wirèn-Lehr, 2001). A threshold value represents an inflexion point, observed or expert-determined, in this relation.

2.3 Final remarks: relative or absolute reference levels? The issue of benchmarking

A major difficulty is encountered when attempting to produce objective relevant thresholds and, if not possible, to conceive procedures that enable the determination of absolute targets for a number of international initiatives of social, economic as well as environmental indicators. As a result, some Sustainable Development Initiatives and Frameworks (SDIF) adopt the concept of relative performance or benchmark. Perret (2002) also justifies this choice considering that “when institutions (European Commission, FAO, World Bank, OECD, PNUD) do not have the usual tools of public authorities, benchmarking – or comparative gauging – is an efficient mean to influence States’ decisions, while encouraging each country to compare itself with ‘good pupils’ and to take ‘best practices’ as a model.” Apart from this strategic dimension, benchmarking also exhibits other advantages.

The benchmark option is especially prevalent in initiatives related with social indicators (Sharpe, 1999). In the social field, it is difficult to establish absolute thresholds, and only the positive or negative feature of indicators on sustainable development is identified in the common range of variation of such indicators.

To clarify more precisely the procedure of benchmarking, we can consider an indicator I estimated with information related to the variable X, continuous, collected at the level of N spatial scales. The performance of each spatial unit ($j = 1\ldots N$) is estimated in terms of ‘distance’ between the values X_j and X_k of a spatial reference unit k. In most cases, the identification of this spatial reference unit is based on an analysis of the distribution of X in all the considered spatial units. As a result, one can consider the distance between the first decile (which is the same as the tenth percentile) and the median. This approach enables to make the indicator less sensitive to extreme values, e.g. exceptional performances due to regional specific conditions.

We obtain, with for example X_k = quantile(X, 95%):

If $X_j < X_k$ then $I=f(X_j-X_k)$, otherwise $I_j= f(X_k)$ with f monotonous.

In this case, it can be noted that an implicit threshold equals zero. Thus, whatever the value of $X_j < X_k$, X_j+epsilon is considered as favourable in terms of sustainable development.

If data are available on several periods, dynamic analysis of such ‘distances’ enables us to assess trends of convergence and divergence among regions. Its also allows, at this scale, to estimate regional trends of a region as compared to other regions. Conversely, analysis of ‘distances’ is not suited to assess the global performance of all regions in time. The difficulty relies upon the method used to assess a better convergence of various entities (regions, countries) as compared with ‘good pupils’ selected as references.

In order to aggregate detailed data into composite indicators, distances can be expressed as so-called ‘normed distances’. This concept, which can used to simplify the data normalisation process, is however only applicable when the selected indicators are those whose contribution – whether positive or negative – can be compared in terms of importance. Whether this is the case depends on the framework definition.

The mono-dimensional approach in the use of benchmarking is especially useful for dynamic analysis. Needless to say that, from a static point of view few things change concerning the direct use of X levels (with $I_j=f(X_j)$), except a mere sensitivity to extreme values as well as a shortcut for indicators normalisation. However, not only from a mono-dimensional, but also from a multi-dimensional perspective is benchmarking relevant. Indeed, performance cannot be achieved in all the dimensions of sustainable development, and some 'performance configurations' can match political trade-offs or choices, see, for example, comparative advantages in specific domains. Analysing the values of various designed indicators can facilitate the identification of independent domains, and domains where a better performance is detrimental to performances in other domains. This can open an interesting avenue to deal with relations between thresholds and aggregation, since the differentiation of independent domains can be used to select composite indicators and their aggregation method. Such an avenue, which is especially important at the regional level, can be pursued in a next step of SEAMLESS project,

3 Examples to illustrate the various concepts

Generally, the most obvious way to select thresholds or targets is to use what is already available from the scientific literature or from the policy domain. If guidelines for setting thresholds are available, it is worth considering using these guidelines or already existing thresholds, even when they still need to be discussed scientifically. At this stage of the project, no effort has been taken to select threshold values with other methods, such as the consultation of experts (Angus *et al.*, 2003) or by simulation with help of models. However, it is possible that these types of thresholds could be developed for indicators that are very context dependent especially in the social domain.

In this section, we will try to further explain the various concepts – and the problems that can arise when trying to operationalise them – with examples from the fields of agronomics and environmental sciences, economics and social sciences. These examples will be used to facilitate our understanding and to illustrate the relationship between the different concepts. Of course, other examples are imaginable, but we think that with this (broad) range of examples, we can give some flavour of the most relevant concepts and the context in which they might appear. We conclude this section with some illustrative passages on the interface between the three abovementioned scientific fields.

3.1 Example 1: nitrogen application – targets, thresholds and the precautionary principle

Based on all the considerations before and the current situation of given area (for example, an average N fertilization of 200 kg N per ha) a target for stakeholders would be to reduce the fertilisation on his or her farming land, by, for instance, 20 kg per ha in 2 years (10%), or 50 kg per year (25 %). Besides these fictitious targets, there are also standards on the application of nitrogen from organic manure defined by the European Union: the Nitrates Directive (91/676 EEC) defines for each farm or livestock unit, the amount of livestock manure applied to the land each year, including by the animals themselves, shall not exceed a specified amount per hectare. The specified amount per hectare is the amount of manure containing 170 kg N.

Modelling may provide evidence that nitrogen load in water is reduced in case the application of nitrogen fertiliser remains below 150 kg N per ha. This figure of 150 kg can be interpreted as a scientifically modelled threshold for nitrogen leaching to water. Between 150 kg and 300 kg, effects are harmful but still reversible however when the level of 150 kg is exceeded a risk is taken. If the application of nitrogen is above 300 kg N per ha, then harmful irreversible effects might occur. Thus, when the level of 300 kg N per ha is exceeded, it is expected that irreversible negative consequences will follow. However, there exists scientific uncertainty about the precise magnitude and nature of these adverse effects. The precautionary principle, then, is to assume the worst until proven otherwise and invites policymakers to take into account this worst possibility.

The critical nutrient range defines the realistic range of values of nutrients, such as nitrogen and phosphorus, required for plant nutrition. The critical range for a certain crop can be, for example, 0.25% to 0.35% nitrogen in the clippings.

Based on all these considerations and the current situation of a given area (e.g. average N fertilisation of 200 kg N per ha), a target for stakeholders would be to reduce the fertilisation of 20 kg N per ha in 2 years (10%), or 50 kg N per year (25%).

3.2 Example 2: budgetary expenditures and total welfare – threshold values and targets

In this subsection, we will discuss the issue of (in particular: the problems associated with) setting threshold values and targets for two economic indicators on the current restricted list of indicators for Test Case 1 (D 2.1.1): budgetary expenditure and total welfare based on the understanding that, semantically, in terms of economic indicators of sustainable development, a threshold should be regarded as a minimum value of acceptability.

Budgetary expenditure

Budgetary expenditures are defined as the array of monetary support provided to farmers under the first pillar of the Common Agricultural Policy. Expenditure under the second pillar of the CAP is not included. Therefore, it is perhaps wise to define the threshold for this indicator in terms of the European Commission's decision to fix expenditure on the first pillar of the Common Agricultural Policy until 2013: in 2006 this equates to € 45.5 billion per year, rising to € 50.5 billion per year in nominal terms by 2013. In real terms, however, this represents a decrease in budgetary expenditure: if we define real terms in 2004 constant prices and assume an inflation rate of 2% per year then budgetary expenditure in 2006 equates to €43.7 billion per year in 2006, falling to €42.3 billion per year in 2013 (Ahner, 2004).

A particularly pertinent alternative threshold, particularly in the context of sustainable development, might be to define acceptability in terms of meeting a specific ratio of spending on the first pillar of the CAP versus the second pillar. However, as noted above, it will not be possible to operationalise second pillar CAP expenditure within SEAMLESS. In terms of the target for this indicator, stakeholders could be consulted about whether the current fixed budget is optimal from their perspectives. However, the response to such a question is likely to vary considerably depending on the interest of the various stakeholders. This may seem to call into question the validity of using such targets but if we accept that targets should be stakeholder defined then it may be optimal to assign different targets for different stakeholder groups (e.g. farmers versus policymakers) in order to reflect the inherent subjectivity of using such targets for particular indicators.

Total welfare

It is difficult to define a threshold because of the following considerations:

- Total welfare is a function of five separate indicators – *Agricultural income*, *Money metric* (consumer surplus), *Profits of the processing industry*, *Tariff revenues* and *Budgetary expenditure* – therefore we are talking about the aggregation of five different thresholds.
- How do we aggregate these various thresholds? The *Total Welfare* function weights each component equally but attaching equal importance to producer and consumer surpluses, particularly in a developing country context, could be unwise because "agricultural development requires producer surplus gains on a sustained basis […] it would thus seem that producer surpluses are much more valuable than consumer gains." (Poonyth and Sharma, 2003, p. 17). However, this raises the question of which are the most suitable weights to use in terms of the components of the function and their associated thresholds.
- The threshold for both *Profits of the processing industry* and *Money metric* could resolve to zero but in terms of the former, this is only viable if costs are expressed in economic, rather than accounting terms whereas in terms of the latter this is only viable if calculated using a direct utility function. There is no determination in the literature of an acceptable threshold for profits described in accounting terms because positive accounting profits do not imply positive economic profits but a reasonable threshold in this case would be the minimum opportunity cost of capital which is the rate at which a firm borrows funds since one alternative to a production activity is to pay back borrowed money. This can be calculated using prevailing national or European Central Bank interest rates as appropriate. Whereas, for an indirect utility function such as the one used in CAPRI to calculate consumer surplus (after Varian, 1992) the money metric measure is the minimal expenditure needed for consumers to reach the utility level of the simulation scenario at prices of the reference scenario. Therefore, it appears that the threshold for consumer surplus is determined endogenously in CAPRI; it does not require setting exogenously.

- *Budgetary expenditure* and *Tariff Revenues* may be the easiest components of this indicator to define in terms of appropriate thresholds. *Budgetary expenditure* was discussed above in terms of the European Commission's decision to fix spending on the first pillar of the CAP until 2013. In terms of *Tariff Revenues*, the long-term existence of import tariffs is in doubt. Although no agreement was reached at the 4th World Trade Organisation Ministerial meeting in Hong Kong (December, 2005) this issue is likely to remain on the European political agenda with phasing out of tariffs possible in the medium-term (perhaps after the completion of the phasing out of export subsidies, timetabled for 2013). Therefore an appropriate threshold for this indicator might be zero. Ceteris paribus this would reduce *Total welfare* which is undesirable from an optimality perspective unless parallel gains can be made to other components of this indicator or, alternatively, we sacrifice a proportion of welfare in a European context for gains in international welfare which accrue due to the removal of these tariffs.

Perhaps the most important point from this discussion is that deriving a threshold value for a Total Welfare indicator may involve subjective judgements i.e. in relation to weights attached, or not attached, to producer and consumer surpluses. To avoid this problem an alternative method for threshold setting might be to define acceptability in terms of real-term increases in welfare i.e. if welfare in year x is greater in real-terms compared to year x-1 then the threshold for acceptability is met or by expressing Total Welfare in the policy scenario as a percentage of *Total Welfare* in the baseline scenario. In terms of target setting, stakeholders might be more interested in changes to selected components of the welfare function rather than the aggregate. However, any subjectivity introduced here – that is, greater emphasis on producer rather than consumer surplus for example – is likely to be heavily dependent upon the interest they represent.

3.3 Example 3: example related to social indicators – accessibility to service centres, and proximity between farms

This subsection deals with the main issues associated with the identification of threshold values and targets within social indicators, taking in consideration specifically the example of accessibility and proximity, which relates to issues such as Social Capital and to Quality of Life.

To make sense of any indicator the reference value is essential, i.e., what is the indicator compared with. Within the development social indicators, the identification of reference values has shown to be highly contextual, as the levels which are acceptable for one context may not be in another, and vice-versa. This is valid for all types of indicators but needs particular attention in the social dimension, as the evaluation of social aspects, by their nature, depends on the subjects opinions and values and is thus loaded with a higher level of subjectivity than other issues. In the same way, thus, identifying threshold values and targets should also be supported on a certain level of contextualization, related with the regional level considered in SEAMLESS.

If this need for contextualization of the social indicators in particular is accepted in SEAMLESS, the Regional Typologies defined by WP4 (PD4.5.2: Regional Typologies of socio-economic contexts, with routines for calculating socio-economic indicators) may serve as a basis for further calculation. Some precaution is needed though, as these typologies draw on the economic situation in the 1990s, what may in some cases not be a sufficiently strong basis for ex-ante assessments to be produced at present. It could be interesting to base certain assumptions and assessments on trends that can be explored from the data, even if the time horizon is rather short.

The examples considered here deal with accessibility. The indicator a) Accessibility to service centers relates to the distance that has to be covered by the farmer in order to get contact, with an adviser, a bank, or services related with his activity, or even other type of social contact that he can establish in a centre. The general assumption is that the closer the farmers are from the services centers, the easiest they can reach information, support, exchange, debate, and other types of services – and the easiest they can be reached by others.

In this way, the smallest the average distance from farms to service centers, the highest the intensity of relations and the probability for a more developed social capital, and, in a rural framework, the better quality of life. The indicator b) Proximity between farms, relates to the distance between farm units, i.e. the distance that a farmers and those who live in the farm have to use to get to the other closest farm unit. This is relevant as a measure of quality of life, as it shows the possible frequency of social contacts with other people also related with the farm sector. The general assumption is that the closest the farms are from each other, the easiest is the contact and thus a stronger social network can be established, reducing social isolation that often in a problem in agricultural contexts.

Indicator a) measures the average distance of all farms within a given spatial unit, from the farm to the service centre. The calculation of sensitivity to different scenarios may be based on a fixed distribution of centers and road network, and changes in the number of farms per type, but also generally in number of farms and their distribution in the region, considering that, in terms of land use, farms are located only in agricultural land. Indicator b) measures the average distance of all farms between each other, in a given spatial unit. The calculation of sensitivity to different scenarios may be based on a fixed distribution of farm centers, and changes in the number of farms and their distribution in the region, considering that, in terms of land use, farms are located only on agricultural land. Information on changes in farm types may be relevant, as it makes it possible to associate with certain dimensions or locations (soil types).

On both cases, a very low accessibility or proximity to other farms, meaning a low frequency of contacts, may have the effect that the existence of certain farms ceases to ve socially viable, as it becomes to isolated, in relation to farm sector information and services, but also in relation to social contact and networking.

In relation to thresholds and targets, they may be quite different from one region to other, as the capacity for mobility of farmers, and also their families in the case on indicator b), can be very different, as well as accepted travel time for acquiring a certain kind of service or for meeting others. In the same way, the desired contact in order to feel as part of a community, or network, is highly dependent on cultural and social rules. Just in Europe, between a farmer in Southern Finland and a farmer in Northwestern Portugal, the need and demand for farm related services and social contact may vary from a monthly to a weekly or a daily contact, for instance.

These thresholds and targets for this kind of social indicators are very little considered in literature until now. There are references from case-study research, but not from generally accepted indicators. In this way, the suggestion is to define, through a constructive combination of a theoretical and a pragmatic approach (Schafer 2004), that combines mainly concepts of Social Capital and Quality of Life with models of accessibility, some thresholds and targets to be dealt with in SEAMLESS. This has been put forward in PISIS and it would demand a closer collaboration between specialist in different workpackges, not the least WP4, WP2 and WP5. But as a starting point, it demands the acceptance of the regional contextualization of some indicators.

3.4 Example 4: the interface between the three scientific fields

Although the importance of biodiversity in agriculture and farming systems has been recognized in, a large number of studies have shown that biodiversity in agro-ecosystems has been declining and has become eroded at genetic, species and ecosystem level – in plants, animals and soils. Also of concern is the loss of biodiversity in 'natural' habitats due to land use changes (expansion of agricultural areas to frontier areas or reduction in agricultural areas).

Thrupp (2005) distinguishes two different agro-biodiversity indicators, namely (i) indicators of decline and erosion of agro-biodiversity (e.g. number of varieties contributing to a certain percentage of crop species; reduction over a certain time-span); and (ii) practices and approaches likely to conserve or enhance biodiversity in agro-ecosystems, such as Integrated Production (IP) and integrated pest management practices, and organic farming.

While referring to organic farming, Hole *et al.* (2005) describe the potential impacts of specific farming practices on biodiversity. They differentiate direct from indirect effects on target species or environmental compartments. However, the authors do not specify whether threshold effects can be evidenced, nor if the combination of practices into technical systems benefits to target species or habitats. Concerning farming practices, a possible approach would consist in using 'benchmarking', for instance considering variability among fields or cropping systems within a farm or a region, and comparing outputs with average or best practices referenced.

Indicators of biodiversity and wild life conservation have been proposed at farm level, but seldom is given any justification for their focus on certain species or communities species. At a more global level, Lawler *et al.* (2006) fund geographical and taxonomical gaps (amphibians being understudied) as well as gaps in invasive species. Callicott *et al.* (1999) suggest a dichotomy between nature conservation concepts focusing on species composition (biodiversity, resources conservation, i.e., a compositionalist view), and concepts focussing on ecosystem functions (ecosystem health, i.e. a functionalist view).

The first approach can be applied to cropping patterns and sequences, although the relation between temporal and spatial dimensions is not always clear-cut. Concerning crop diversity on agricultural land, Bockstaller and Girardin (2003) suggest a crop sequence index (ISC) at field level, including three components:

– Effect of the preceding crop (k_p), with a ranking of preceding effects
– Minimum period before repeating cultivation of a crop (k_r) (e.g. to avoid soil-borne pathogens)
– Crop diversity in the sequence (k_d)

The recommended reference value of the ISC is approximately 7, which is consistent with IP principles, namely concerning k_d This reference value can be obtained, for instance, with a good preceding crop (k_p = 4-5), a minimum period between two occurrences of one crop (k_r = 1-1,2), or matching four different crops in the sequence (k_d = 1,4).

The second approach was introduced by Becker (1997) who distinguishes three generations of environmental indicators, among which bio-indicators are the first generation.

The second generation of indicators focuses on ecosystem dynamics, on the structure and function of entire ecosystems. This includes the assessment of values such as 'ecosystem integrity', as expressed by the 'Index of Biotic Integrity' (Regier, 1993), and 'ecosystem health' presented earlier and discussed by Xu and Mage (2001). These two generations can be called 'ecological indicators', since their design is strongly anchored in ecology (for further details, see PD 2.2.1).

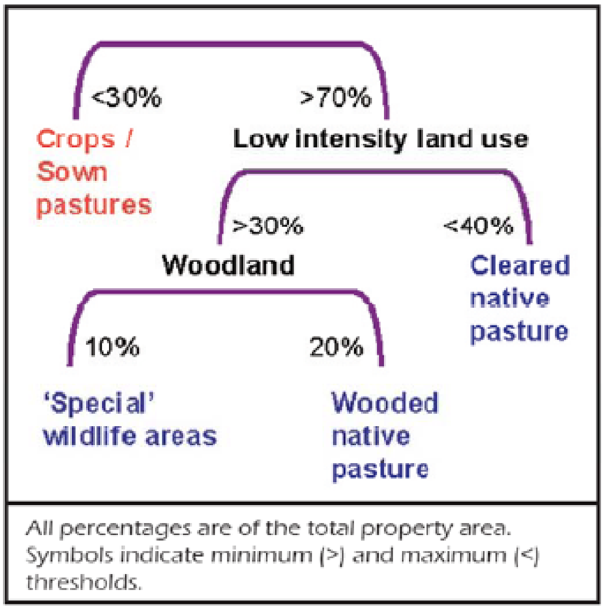
In fact, bio-indicators could preferably be used to confront indicator calculation with field observations, or to compare situations. Instead of using absolute values, the local occurrence and species and varieties can also be defined enhancing stakeholders' knowledge, as suggested by Rodrigues *et al*. (2003), whereas changes in natural vegetation and corridors of fauna were assessed separately (avoiding double counting due to overlap between corridors and natural vegetation). Likewise, Noe *et al* (2005) developed and tested indicators at farm level to facilitate the dialogue between farmers and experts on wildlife quality. 'Weed cover in cereal fields' was used as an indicator of floral and faunal biodiversity on cultivated land, and 'uncultivated biotope areas' was used as a general measure of wildlife habitat. Functional groups of herbaceous plants and low mobility butterflies were used as indicators of conservation value.

In most cases, agricultural land and natural habitats are considered as relatively independent. Thresholds are also defined at farm level, usually expressed in terms of a percentage of uncultivated land to maintain at farm level. In a project with integrated and organic farmers, Vereijken (1995) suggested a threshold of 5% of biotope area. Ecological infrastructures and 'ecological compensation areas', ranging from 8 to 10% of farm area, are also suggested by IOBC (2005). This could be of particular relevance for land use planning and Good Agricultural and Environmental Conditions (GAEC) and Agri-Environmental Measures in the CAP second pillar, where the location of such infrastructures could be prioritized to protect environmental resources. However, this entails another working scale.

At another scale, indications suggest composition and distribution of land use patterns. For example, in order to assess net changes at EU level in the middle term Weber *et al*. (2003) and SEAA (2003) use a land accounting system including land cover and land use matrixes. This approach raises questions about the nature of the transformations depicted as 'critical points' might be crossed. For example, does the gains in cover or use compensate for the losses? This issue of compensation requires an understanding of the benefits associated with a given land cover element and the status those benefits have in a new mosaic. Relations between agricultural and non-agricultural activities also entails economic thresholds. For instance, in urban planning documents, an agricultural area which covers less than 10% of the total would entail a potential dissipation of the whole agricultural sector (Geniaux, personal communication).

In Australia, CSIRO has published a set of guidelines to help to ensure a balance between conservation and development, based on thresholds relating to land cover proportions which, if exceeded, will "adversely affect ecosystems services, lead to declines in native flora and fauna and affect production in the long-term" (see the figure below).

Figure 6.3: Landscape thresholds suggested by CSIRO to balance conservation and development.



Beyond the issue of 'landscape quality', which can lead to endless discussions and subjective judgements, land use and its spatial structure can contribute to an understanding of the evolution of biodiversity. Understanding the evolution of biodiversity enables us to cope with insufficient bio-ecological data at European level, and it helps us to derive biodiversity indicators which can be used in the agricultural sector.

Three main ways can be used to simulate this evolution, which are related to the evolution in land use and agricultural practices:

- Ecological modelling, likely to determine habitat potential and subsequent biodiversity indicators (see task 3.7 in WP3).
- Landscape ecology indicators (Turner *et al.*, 2005) among which some have a scientifically established relation with functional groups at specific scales. Such indicators aim at accounting for landscape composition, fragmentation and connectivity. Thresholds exist in: fragmentation for different functional groups of biodiversity, ruderal species, density of hedges (Corine land cover or spot 5 images; confrontations of distribution of indices with empirical data and measurements).
- An assessment of the role of certain practices on biodiversity (Hole *et al.*, 2005.) could be used as a proxy of the number of conversions to such practices (preferably those with a positive role).

These options could be developed in the next steps of SEAMLESS project.

4 Methodological issues

4.1 Who defines reference levels and who decides upon them

This topic has been dealt with in detail in PD 2.6.1. In this PD, the authors write on page 8 that "…consultations should give the opportunity to each potential user and representative stakeholder of the IF to indicate, in interaction with scientists, the indicators to which they usually refer and/or they would like to refer with IF in their decision-making, the relative importance of each of them and the criteria or reasons to which they refer to use/ select them."

Moreover, the authors continue on pages 9 and 10 with stating that "It is admitted that the choice of the threshold or the qualification scale of an indicator must vary according to the area conditions and the point of view of the stakeholders. Regarding the level reached by the several components, a judgment is passed on the assessed scenario. But the judgment can be contradictory between different categories of actors (for instance (…) the qualification of the nitrogen balance may change according to the consulted stakeholders: farmers point of view about this kind of scale can be different from environmentalist). In Seamless, this judgment is devoted to help political decisions. The scientist can inform about risks associated to the different values of an indicator but has not legitimacy to arbitrate between actors issues: it is a political choice to build in light of available scientific knowledge."

All in all, consultation of stakeholders, in interaction with scientists, should provide the thresholds or qualification scales of indicators that the end-users and stakeholders of the IF usually use and/or plan to use.

4.2 How to establish reference levels

Generally speaking, this PD represents the (scientific) consensus on how in SEAMLESS, we can correctly interpret and apply the relevant principles and concepts in general and to specific cases. In order to achieve this consensus, we should address the central question: how to establish reference levels? This is an extremely important question, because these values determine the translation of indicators into policy decision criteria. For example, applying a precautionary principle means including a number of devices for addressing and reducing the effects of uncertainty so that agricultural activities can be planned and conducted in a way that is sustainable in spite of the inherent uncertainty. One of these is the definition of targets and reference levels.

In general, there are three main approaches that could be used to establish reference levels to interpret indicators: (i) analysing existing reference levels, thresholds and target values, and reviewing existing methods for estimating these values; (ii) determining reference levels, thresholds and target values by attributing values ourselves, based on existing scenarios, empirical data, expert knowledge, or model simulation; and (iii) intuition and symbolic thought. The third approach – intuition and symbolic thought – is less concrete then the other two approaches, which are based on logic and sense. Moreover, establishing reference levels solely by intuition does not happen very often; that is, intuition and symbolic thought are mostly used in combination with the other two approaches. Because this third approach is certainly not a reliable guide for the SEAMLESS-project, we focus here only on the other two approaches, which will be outlined below. Although the issue of relative values is already addressed above, the issue is also relevant for the question how to establish reference levels. Therefore, we close this subsection with a few words on relative values and benchmarking.

4.2.1 Reference levels based on existing methods and values

Although there are major differences in the types of thresholds, reference levels and target values, there is much overlap among them. As a result, values or levels can be 'borrowed' from existing methodologies or literature, and applied directly to the indicators identified in the SEAMLESS project. Thus, values, levels, targets, etc. will then be based on the information generated by previous studies and reports. An important question here is which studies and reports must be used. Is there, for instance, a difference between the scientific literature and information available from the policy domain (e.g. Commission Directives and Decisions)? In fact, useful information can be derived from Brussels. For example, for establishing a threshold for nitrates leaching, it seems to be obvious to rest on the EU-guideline for nitrate in water: 50-mg/L^{-1} nitrates. However, this threshold is not equally relevant in all European regions. Nordic countries have in general a very low level of mean nitrate concentrations (<2 mg/L^{-1} nitrates), whereas in Western Europe the mean nitrate concentrations in groundwaters are above 25 mg/L^{-1} nitrates. For sustainability reasons, it may therefore be wise to specify different values and levels for different European regions. Note that in this respect WP4 has done some work on environmental typologies and classification of vulnerable zones, which are, as a matter of fact, identified under the EU Nitrate Directive. From a visualization perspective, it is possible to decide that the areas close to the guidelines are shown in a specific colour, or to create a map which shows the relative change between baseline and policy scenarios. Or Another example relates to the pesticide leaching to the groundwater. Here again, it seems to be obvious to take the EU-guideline for drinking water: 0.1 mg/L^{-1}. Pesticide developers from other continents (Australia, United States), however, based their indicator on other guidelines, which differ according to the active ingredients (Kookana *et al.*, 2005).

All in all, due to spatial variety and complexity, special care must be taken to properly identify the various levels, values and targets. Existing methodologies and values from one source (or European region) cannot automatically be applied to other regions. That is, where data exist, there may be variations in methodology that make wildly borrowing values unreliable. This means that levels, values and targets should be determined on a 'site-specific' basis, and are – obviously – context dependent. Thus, defining reference levels on the basis of existing methods and values seem to be only part of the story. Complementary to this approach, it might be useful to establish reference levels by determining the appropriate values ourselves.

4.2.2 Reference levels based on empirical data, expert knowledge and model simulation

For a number of indicators, no theory, data or methodologies exist that would allow a range of acceptable levels, values and targets – let alone reference points or target values. This is particular the case for economic and social indicators. Moreover, the fact that levels, values and targets are context dependent implies that they can change over space and time, and adds emphasis to the need for establishing and updating them within the SEAMLESS project. This means that within the project, it is not sufficient to rely simply on existing values. Empirical trends and expert opinions should be used to adjust levels, values and targets obtained from (sparse) data.

With respect to the relation between thresholds and models outputs, several situations should be considered. In case of a deterministic model, thresholds can be investigated with a sensitivity analysis of the model; i.e. sensitivity to relevant variations of $ß_i$, in various configurations of X_l, with $ß_i$ the i model parameters, and X_l the l model variables. With a non-deterministic model, it can be interesting to simulate a set of model outputs and analyse their variability with different configurations of $ß_i$ and X_l. Investigated thresholds can refer to X_l values or directly to a model output (S_j). When relevant information on S_j thresholds is not available elsewhere, it is necessary to identify the X_l value ranges which lead to discontinuities or a great variability in S_j, or result in an 'over-increase' in the absolute values of S_j secondary derivatives.

Concerning the thresholds to infer, solutions that are based on data and indicators are also manifold. In the case of 'benchmarking', the analysis of the distribution of observed values for various regions can be enough to provide an idea of privileged reference points. With absolute thresholds, the analysis of the distribution can also be useful, provided that the unsustainable feature of observed extremes is demonstrated. Statistical analyses can also be used to improve the position of thresholds between the observed extremes.

From another viewpoint, accurate thresholds can be obtained at field or farm level, but be meaningless at regional or national level (see the abovementioned examples with nitrates or biodiversity). The discussion can then focus on the selection of statistic indicators mentioning, for instance, how many times local thresholds are exceeded. However, the limits of such an approach are obvious in the case of absolute thresholds. In order to overcome this limits, two options can be envisaged: (i) using expert-based knowledge; (ii) address differently the issue of thresholds that is not independently for each indicator, but through the analysis of the relations that exists between various indicators and phenomena. Therefore, studying correlations between indicators is a fundamental prerequisite for the identification of thresholds. Moreover, if correlations are identified, statistical modelling can facilitate the assessment of the consequences that various levels of indicators (or groups of indicators) have for other indicators. This shows the importance of a multidimensional approach of the question of thresholds, as related to aggregation issues. This will be dealt with further the proceedings tasks 2.5 and 2.6.

4.3 Reference levels and scale issues

Values assigned to a concept can be scale dependent and need not necessarily be relevant for every scale. This is for example true for water pollution caused or induced by nitrates from agricultural sources. Current EU guidelines refer to the hydro systems (groundwater, river). However, for lower scales, such as field or farm level, it can be questioned whether those nitrate values are relevant. From emission of pollutant at field or farm scales to pollution of the hydro system, different processes (hydrologic, mitigation, compensation by natural area, etc.) can occur. Consequently, applying those targets to a lower (or higher) scale can for some pollutants lead to an overestimation of the risk, whereas for other pollutants the same target level at different spatial scales can lead to an underestimation of the risk. However, EU guidelines are well known, especially the Water Framework Directive and the Nitrate Directive, so that most policymakers know that these guidelines refer to the level of catchments or regions, rather than to a particular field level. But in general, guidelines are used by many stakeholders in an unconscious way independently of the scale.

For other pollutants, using the same target level at higher scales can lead to an underestimation of the risk as for nutrient balances. When those are calculated at higher scales, compensation may occur and lead to lower values at higher scales. However, because Prototype 1 does not include information on all those processes (especially hydrological ones), at least at this stage of the project, those values can be used, being based on a worst-case situation. Such a situation then indicates that the model outcomes are provisional and subject to change – and present conservative estimates – as those values are not agreed on.

4.4 Reference levels and the local context

As already mentioned in section 4.2.1, a reference value, such as the nitrate EU guideline for drinking water, may be adapted to regional contexts. When considering the scale issue, it is possible to correct reference values at lower or higher scale with co-called contextual elements. For example, when we return to the nitrate EU guideline for drinking water, the ratio cropped area/non-cropped (natural) area in the region may be taken into consideration and would reveal the potential compensation by non-cropped area for water pollution by nitrate or pesticides.

For nutrient balances, references values should be adapted to the scale, and even to the soil type and its mineralisation potential as shown by Oenema *et al.*, (2005). In any case, the feasibility and the relevance of such an 'adaptive' approach in the frame of SEAMLESS should be assessed. The use of simulation models and expert knowledge for typical situations seems to be a reasonable and promising approach to create a typology of contexts with given reference values. Another noteworthy point is that the examples mentioned in previous sections tackle regional or local impacts (water quality, soil fertility, etc.), whereas for global impact as greenhouse impact, fossil energy depletion, this contextualisation is probably not a necessity.

A final important aspect to consider is that in SEAMLESS, the minimum spatial scale for indicators is the farm level (represented by farm types). At this level, Good Agricultural Practices (GAP) are relevant. The strategy of the European Council on the environmental integration and sustainable development in the Common Agricultural Policy in 1999 includes the objective to recognise that a reference level of GAP, which is dependent on local conditions, should be respected in all agricultural areas of the EU. This is formulated as follows: “The general principle is that where farmers provide services to the environment beyond the reference level of good agricultural practices, these should be adequately remunerated. Certain methods of agricultural production, for example organic farming integrated production and traditional low-input farming and typical local production, and typical local production, provide a combination of positive environmental, social and economic effects.” (see Council of the European Union, 1999, p.6; European Commission, 2001, p. 28).

5 Consequences for SEAMLESS

Reference levels are essential to define undesirable sustainability conditions in advance. Assigning these values to indicators is, of course, not new. Especially from an ecological point of view, there is a large literature and extensive experience with setting reference levels and targets, for example, for indicator species. For exploited species, many science advisory and management agencies use limit and precautionary reference points as core tools. As a result, the threatened or endangered species community has extensive experience with using quantitative reference levels.

Reference levels, thresholds and target values are context-dependent and cannot without careful consideration be used within the SEAMLESS project. As a result, there is limited ability to borrow existing values and apply them to the SEAMLESS indicators. It is hard to justify taking a value that is well justified as a reference level for one area, and using it as a *de facto* reference level for other areas.

Whereas there is extensive data in the ecological literature, information about reference levels, thresholds, target values, etc. for social and economic indicators is much more sparse. All in all, lack of data, in combination with values borrowed from different (regional) sources or different time periods makes it almost inevitable that (some of) the levels, values and targets should be assessed within the SEAMLESS project. Thus, not all the information needed can be compiled from existing sources and should therefore be gathered or estimated within SEAMLESS itself, for example by expert knowledge or model simulation.

In fact, for establishing reference levels, a distinction can be made between single indicators and aggregated or composite indicators. Single indicators have specific reference levels, whereas aggregated or composite indicators entail methodological and calculation issues for the definition of these levels. As a result, borrowing values from existing literature is more applicable for single indicators than for aggregated or composite indicators. Another implication is that establishing levels, values and targets, is very much related to task 2.6 (‘Development of a methodology to assess multiple indicators and compare them’) within the SEAMLESS project. Therefore, cooperation is urgently needed. Moreover, this PD entails a relation with PD 2.2.1 and, of course, with task 2.5. It also opens exciting possibilities for future work on many methodological issues.

Thus, because indicators do not necessarily progress in a linear fashion, the identification of thresholds is very important. However, in reality few actual thresholds are known.

5.1 Absolute or relative reference values?

It is expected that for several indicators the main problem will be the availability of a reference value. The subjectivity establishing reference values and the limited role scientists currently play (see PD 2.6.1) may lead some authors and researchers to refuse any absolute reference value, as they are considered in this document (Bockstaller *et al.*, 2006). An alternative to using an absolute reference value is to work with relative reference value, which can be based on:

- An initial value in case of time series
- Average conditions as benchmark
- Values of the baseline scenario

In the context of SEAMLESS, using relative reference values may be a solution, especially for indicators where no values are available. After all, establishing absolute reference values is quite a daunting task. Therefore, using relative reference values can save lots of time. Unfortunately, however, relative reference values have a main limitation, namely their relativity, which raises two important questions:

- What is the status of the reference value? In Bockstaller *et al.*, (2006) authors of the SALCA method (using a life cycle analysis) based their reference level on average environmental performance of Swiss farms. Those farms are considered as well performing because of the cross-compliance policy. In this example, appropriate expertise is needed to gauge the relative value.
- How to assess the significance of the degree of variation of a given indicator in comparison to a relative reference (e.g. value of the baseline scenario)? In SEAMLESS, this question cannot be ignored, especially if all indicators are systematically compared with the value of the baseline scenario.

Due to these questions, it may intuitively appear that absolute reference values are superior to relative reference values but this is only true where credible (i.e. from the academic literature or well-reasoned rules-of-thumb) absolute values exist. Nevertheless, an issue that is pertinent here, and which has been discussed already in this report, concerns how we define thresholds. For many environmental indicators thresholds are closely related to the concept of irreversibility and non-linear system dynamics: if the value of a given indicator falls below a certain level and drastic changes occur then the level at which this occurs would be important for threshold setting. However, in the economic domain thresholds are more often associated with a minimum value of acceptability. The notion of ‘acceptability’ is obviously rather subjective and hence the difficulty in applying absolute values.

5.2 Proposals for GUI

Reference levels in the form of thresholds and target values, should be visible when indicators are presented in the GUI for a given spatial scale. Raw values of indicators cannot be presented without their reference values, which can be given when results are presented in tables. More attractive is to add to the numerical value a qualification of value depending on the position of the indicator value relative to the reference value. This can be done through textual descriptions (e.g. positive/indifferent/negative) and/or a colour codes (e.g. green/orange/red). Topographical or thematic maps with a colour code are, of course, highly imaginative and thought provoking (see Figure 2).

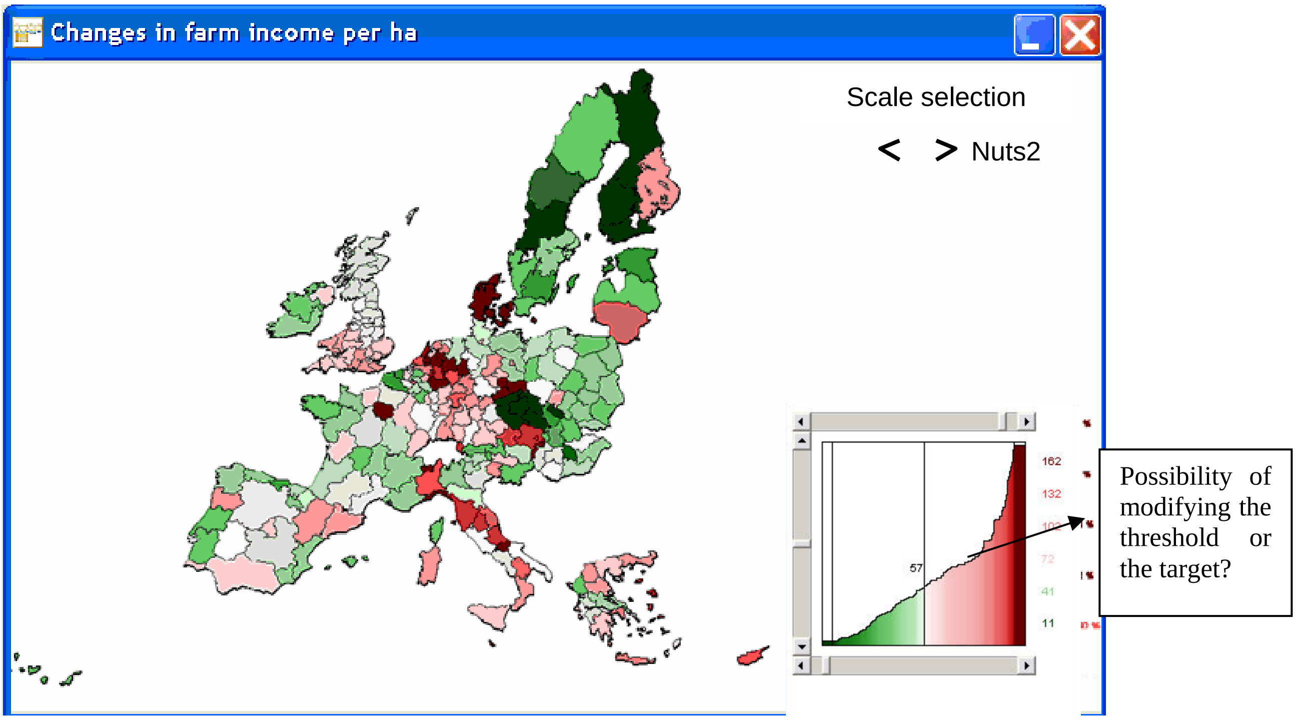


Figure 2 Example of a map-based presentation of an indicator, with a possibility to modify the reference value of the indicator

An important question related to the presentation of the indicators in the GUI, is how much flexibility will be given to the end-users. Several options are possible:

1. No flexibility
2. Total flexibility, allowing users to modify proposed values
3. Partial flexibility, allowing users to modify value within a given range
4. No flexibility in the premodelling phase but only in the postmodelling step: This is shown in Figure 2, where the users can modify in an interactive way the reference value. The result of each modification is subsequently visualised. This would allow the user to run a kind of sensitivity analysis.

These options can be arranged on a continuous spectrum with no flexibility on one side and total flexibility on the other. In addition to these three options, there is a fourth option, namely no flexibility in the pre-modelling phase and flexibility in the post-modelling phase. This is shown in Figure 2, where the users can modify in an interactive way the reference value and visualise (graphically) what the different results would look like. This would allow the user to run a kind of sensitivity analysis. The most preferable option is probably the combination of option 3 and option 4.

6 Concluding remarks

The report offers the results of a cross-disciplinary literature review on reference levels, used in the environmental, economic and social domains of sustainable development. In part, the report contributes to an extended glossary of terminology for further use in the SEAMLESS project, and also clarifies the interpretations given in disciplines on the various components related to reference levels.

The report provides a common understanding of thresholds that is adopted from the scientific literature, but it remains to be decided how such thresholds should be defined in SEAMLESS. There might be disagreement among scientists to establish an absolute reference level, and it also remains to be clarified how using absolute or relative threshold levels. Absolute levels might have limited relevance in the social domains where norms are identified and the levels of indicators are very contextual dependent. In contrast, absolute levels are far more common in the environmental domain and critical ranges are available beyond which exceeding risks are taken.

Further work on reference levels is planned, and the current report offers a basis for a more operational use of reference levels and setting critical ranges that will be reported in the coming months.

7 Glossary[3]

Absolute reference value	A reference level (see below) that is independent of other values or arbitrary standards of measurements.
Benchmarking	Benchmarking, or comparative gauging, is measuring or judging other persons or 'things' by comparing them with a benchmark; i.e. a point of reference or standard.
GUI	Graphical User Interface that allows tracking and tracing of variables and parameters in model components and throughout the model chain
Indicator	A value that can be used to evaluate or assess different types of impact. Indicators are in SEAMLESS values that have been assessed to be relevant for a specific context, a specific policy, or user group
Reference level	A general term for a level to which the indicator is referred. Reference levels come in a wide variety of names, such as benchmark, standard, threshold, norm, but all refer to a comparison to which an indicator can be examined or gauged.
Reference value	Concept that includes several types of values, namely thresholds, critical values, critical ranges, and target values. These values can be used to define the context for an indicator as well as the basis of reference to which the indicator should be compared.
Relative reference value	A reference level (see above) that has a relation to or connection with or necessary dependence on another value.
Stakeholders	Stakeholders include the potential users of the results of SEAMLESS-IF – namely, the European Commission Directorates General and the European Environmental Agency – but also regional policy makers, private companies, farmers' organisations, NGO's and the scientific community.

[3] For detailed definitions of relevant principles, concepts and the various types of reference levels, the reader is referred to the subsections 2.1 and 2.2.

References

Ahner, D. 2004. *CAP reform and EU enlargement: the future of European agricultural policy* [online]. Available from: http://europa.eu.int/comm/agriculture/events/sofia/ahner.pdf.

Angus, A.J., I.D. Hodge, S. McNally and M.A. Sutton. 2003. The setting of standards for agricultural nitrogen emissions: a case study of the Delphi technique. *Journal of Environmental Management*, 69; 4, pp. 323-337.

Baarsma, B.E. 2003. The valuation of the IJmeer nature reserve using conjoint analysis. *Environmental and Resource Economics,* 25; 3, pp. 343-356.

Becker, B., 1997. *Sustainability Assessment: A Review of Values, Concepts, and Methodological Approaches*. Consultative Group on International Agricultural Research. The World Bank, Washington, D.C., USA.

Bergh, J.C.J.M. van den. 2001. Ecological economics: themes, approaches, and differences with environmental economics. *Regional Environmental Change*, 2; 1, pp. 13-23.

Bishop, R.C. 1978. Endangered species and uncertainty: the economics of a safe minimum standard. American Journal of Agricultural Economics, 60, pp. 10-18.

Bockstaller, C., G. Gaillard, D. Baumgartner, R. Freiermuth Knuchel, M. Reinsch, R. Brauner and E. Unterseher. 2006. *Méthodes d'évaluation agri-environnementale des exploitations agricoles: Comparaison des méthodes*. INDIGO, KUL/USL, REPRO et SALCA, Colmar, ITADA, p. 112.

Bockstaller C. and P. Girardin. 2003. *Mode de calcul des indicateurs agri-environnementaux de la méthode INDIGO®* (version 1.42 du logiciel). Working Paper.

Callicott J.B., L.B. Crowder and K. Mumford. 1999. Current normative concepts in conservation. *Conservation Biology*. 13 1, pp. 22-35.

COUNCIL DIRECTIVE of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources (91/676/EEC).

Council of the European Union. 1999. *Agriculture and Environment; Council Strategy on the Environmental Integration and Sustainable Development in the Common Agricultural Policy established by the Agricultural Council.* Brussels, Agricultural Council.

Crowards, T.M. 1998. Safe minimum standards: costs and opportunities. *Ecological Economics*, 25; 3, pp. 303-314.

CSIRO. 2000. *Balancing Conservation and Production. Understanding and Using Landscape Thresholds in Property Planning.*An Initiative of the Grazed Landscapes Management Project at CSIRO Tropical Agriculture, Brisbane.

European Commission. 2001. *A Framework for Indicators for the Economic and Social Dimensions of Sustainable Agriculture and Rural Development.* Brussels, EC, Agriculture Directorate-General.

Gollier, C. B. Jullien and N. Treich. 2000. Scientific progress and irreversibility: an economic interpretation of the 'Precautionary Principle'. *Journal of Public Economics*, 75, pp. 229-253.

Groffman, P.M., J.S. Baron, T. Blett, A.J. Gold, I. Goodman, L.H. Gunderson, B.M. Levinson, M.A. Palmer, H.W. Paerl, G.D. Peterson, N. LeRoy Poff, D.W. Rejeski, J.F.

Reynolds, M.G. Turner, K.C. Weathers and J. Wiens. 2006. Ecological thresholds: the key to successful environmental management or an important concept with no practical application? *Ecosystems*, 9, pp. 1-13.

Hanley, N., J.F. Shogren and B. White. 1997. *Environmental Economics in Theory and Practice*. Houndmill-London, MacMillan Press Ltd.

Heide, C.M. van der 2005. *An Economic Analysis of Nature Policy*. Amsterdam, Vrije Universiteit, Tinbergen Instituut Research Series no. 356, Ph.D. thesis.

Hole, D., A. Perkins, J. Wilson, I. Alexander, P. Grice and A. Evans. 2005. Does organic farming benefit biodiversity? *Biological Conservation*, 122; 1, pp. 113-130.

Holling, C.S. 1973. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, pp. 1-23.

IOBC. 2005. *Ideabook on Ecological Infrastructures and Functional Biodiversity at the Farm Leve*l.

Jollands, H. (2006). How to aggregate sustainable development indicators: a proposed framework and its application. *International Journal of Agricultural Resources, Governance and Ecology*, 5; 1, pp. 18-34.

Knight, F.H. 1921. *Risk, Uncertainty and Profit*. Boston and New York, Houghton Mifflin.

Koga N, H. Tsuruta, H. Tsuji and H. Nakano. 2003. Fuel consumption-derived CO2 emissions under conventional and reduced tillage cropping systems in northern Japan. *Agriculture, Ecosystems and Environment*; 99, pp. 213–219.

Kookana, R.S., Correll, D.L. and Miller, R.B., 2005. Pesticide impact rating index- a pesticide risk indicator for water quality. Water, Air, Soil Pollution: Focus (in press).

Kooten, G.C. van and E.H. Bulte. 2000. *The Economics of Nature; Managing Biological Assets*. Oxford, Blackwell Publishers Ltd.

Lal R. 2004. Carbon emissions from farm operations. *Environment International*, 30, pp. 981–990.

Lawler, J. J., J. E. Aukema, J. Grant, B. Halpern, P. Kareiva, C. R. Nelson, K. Ohleth, J. D. Olden, M. A. Schlaepfer, B. Silliman, and P. Zaradic. 2006. Conservation Science: A 20-Year Report Card. *Frontiers in Ecology and the Environment* 4; 9, pp. 473-480.

Noe, E., N. Halberg and J. Reddersen. 2005. Indicators of biodiversity and conservational wildlife quality on Danish organic farms for use in farm management: a multidisciplinary approach to indicator management and testing. *Journal of Agricultural and Environmental Ethics*, 18; 4, pp. 383-414.

Norton, B.G. and M.A. Toman. 1997. Sustainability: ecological and economic perspectives. *Land Economics*, 73; 4, pp. 553-568.

OECD. 1999. *Handbook of Incentive Measures for Biodiversity; Design and Implementation.* Paris, OECD.

Oenema, O., L. van Liere and O. Schoumans. 2005. Effects of lowering nitrogen and phosphorus surpluses in agriculture on the quality of groundwater and surface water in the Netherlands. *Journal of Hydrology*, 304, pp. 289-301.

Perrings, C., C.F. Folke and K.-G. Mäler. 1992. The ecology and economics of biodiversity loss: the research agenda. *Ambio*, 21; 3, pp. 201-211.

Pindyck, R.S. 2000. Irreversibilities and the timing of environmental policy. *Resource and Energy Economics*, 22; 3, pp. 233-259.

Poonyth, D and R. Sharma. *The impact of the WTO negotiating modalities in the areas of domestic support, market access and export competition on developing countries: results from ATPSM* [online]. International conference - Agricultural policy reform and the WTO: Where are we heading? Capri, Italy, June 23-26, 2003. Available from: http://www.ecostat.unical.it/2003agtradeconf/Contributed%20papers/Poonyth%20and%20Sharma.pdf.

Randall, A. 1988. What mainstream economists have to say about the value of biodiversity. pp. 217-223. In: E.O. Wilson (ed.). *Biodiversity*. Washington, DC, National Academy Press.

Randall, A. and M.C. Farmer. 1995. Benefits, costs, and the safe minimum standard of conservation. pp. 26-44. In: D.W. Bromley (ed.). *Handbook of Environmental Economics.* Oxford and Cambridge, Basil Blackwell Ltd.

Regier, H. A. 1993. The notion of natural and cultural integrity. pp. 3-18. In: S. Woodley, J. Kay, and G. Francis (eds). *Ecological integrity and the management of ecosystems*. Ottawa, Canada, St. Lucie Press.

Rodrigues G.S., C. Campanhola and P.C. Kitamura. 20003. An environmental impact assessment system for agricultural R&D. *Environmental Impact Assessment Review*, 23, pp. 219-244.

Sayadi, S., M.C. Gonzalez Roa and J.C. Requena. 2005. Ranking versus scale rating in conjoint analysis: evaluating landscapes in mountainous regions in southeastern Spain. *Ecological Economics*, 55; 4, pp. 539-550.

SEEA, 2003. The Handbook of National Accounting: Integrated Environmental and Economic Accounting 2003. Report available from the Internet: http://unstats.un.org/unsd/envAccounting/seea2003.pdf.

Sharpe, A. 1999. *A Survey of Indicators of Economic and Social Well-being.* Paper prepared for Canadian Policy Research Networks, Ontario, Centre for the Study of Living Standards.

Spash, C.L. 1999. The development of environmental thinking in economics. *Environmental Values*, 8; 4, pp. 413-435.

Tisdell, C.A. 1991. *Economics of Environmental Conservation; Economics for Environmental and Ecological Management.* Amsterdam-London-New York-Tokyo, Elsevier Science Publishers B.V.

Thrupp, L.A. 2005. The importance of biodiversity in Agroecosystems. *Journal of Crop Inprovement*. New dimensions in agroecology. pp. 315-337.

Turner M.G., R.H. Gardner and R.V. O'Neill. 2005. *Landscape Ecology in Theory and Practice – Pattern and Process***.** Springer-Verlag, New York, NY, USA.

United Kingdom Parliament. 2006. *House of Commons Written Answers for 14 Feb 2006* [online]. Available from: http://www.parliament.the-stationery-office.co.uk/pa/cm200506/ cmhansrd/cm060214/text/60214w05.htm.

Vereijken, P. 1995. *Designing and Testing Prototypes: Progress Reports of the Research Network on Integrated and Ecological Arable Farming Systems for EU and Associated Countries.* DLO Research Institute for Agrobiology and Soil Fertility, Wageningen, Concerted Action AIR 3-CT920755

Weber, J.L., *et al*. 2003. *Integration of Environmental Accounts in Coastal Zones; Case Study of Tourism*. Report of the European Topic Centre on Terrestrial Environment for Eurostat and the EEA. Barcelona-Bellatera.

Wirèn-Lehr, S. von. 2001. Sustainability in agriculture – an evaluation of principal goal-oriented concepts to close the gap between theory and practice. *Agriculture, Ecosystems & Environment*, 84; 2, pp. 115-129.

Wills, I. 1997. *Economics and the Environment; A signalling and incentives approach.* St. Leonards, Allen & Unwin.

Xu, W. and Mage, J.A. 2001. A review of concepts and criteria for assessing agroecosystem health including a preliminary case study of southern Ontario. *Agriculture, Ecosystems and Environment,* 83; 3, pp. 215-233.