International Institute for Sustainable Development | Institut international du développement durable

http://www.iisd.org

Linking Farm-Level Measurement Systems to Environmental Sustainability Outcomes:

Challenges and Ways Forward

Aimee Russillo
László Pintér

October 2009

Prepared for Agriculture and Agri-Food Canada

Linking Farm-Level Measurement Systems to Environmental Sustainability Outcomes:

Challenges and Ways Forward

Aimee Russillo
László Pintér

October 2009

IISD and the authors would like to acknowledge the substantive input and comments provided by Richard Perkins and David MacLaughlin at WWF to sharpen the focus of the paper. We are also grateful to Gord Kurbis at Pulse Canada for access to data, and to select members of Pulse's network of producers and researchers for the lentils case study and for comments. Jane Barr and Angeline Gough prepared the pulse case study, available separately as a background document. Members of the steering committee of the International Forum on Assessing Sustainable Agriculture (INFASA) as well as INFASA co-chair, Jan Grenz, provided initial guidance with regard to the overall direction of the research.

During the course of developing the report the following other individuals sent additional comments, helped with coordination or provided technical assistance: Diane Conolly, Maryline Guiramand, Bob Macgregor, Eve Rickert, Charles Thrift and Dave Wilkins.

The report was prepared under IISD's framework agreement with Agriculture and Agri-Food Canada, whose support and interest in the work is thankfully acknowledged.

Published by the International Institute for Sustainable Development

IISD contributes to sustainable development by advancing policy recommendations on international trade and investment, economic policy, climate change and energy, measurement and assessment, and natural resources management, and the enabling role of communication technologies in these areas. We report on international negotiations and disseminate knowledge gained through collaborative projects, resulting in more rigorous research, capacity building in developing countries, better networks spanning the North and the South, and better global connections among researchers, practitioners, citizens and policy-makers.

IISD's vision is better living for all—sustainably; its mission is to champion innovation, enabling societies to live sustainably. IISD is registered as a charitable organization in Canada and has 501(c)(3) status in the United States. IISD receives core operating support from the Government of Canada, provided through the Canadian International Development Agency (CIDA), the International Development Research Centre (IDRC) and Environment Canada; and from the Province of Manitoba. The Institute receives project funding from numerous governments inside and outside Canada, United Nations agencies, foundations and the private sector.

International Institute for Sustainable Development
161 Portage Avenue East, 6th Floor
Winnipeg, Manitoba
Canada R3B 0Y4
Tel: +1 (204) 958 7700
Fax: +1 (204) 958 7710
Email: info@iisd.ca
Website: iisd.org

Table of Contents

Acknowledgements

IISD and the authors would like to acknowledge the substantive input and comments provided by Richard Perkins and David MacLaughlin at WWF to sharpen the focus of the paper. We are also grateful to Gord Kurbis at Pulse Canada for access to data, and to select members of Pulse's network of producers and researchers for the lentils case study and for comments. Jane Barr and Angeline Gough prepared the pulse case study, available separately as a background document. Members of the steering committee of the International Forum on Assessing Sustainable Agriculture (INFASA) as well as INFASA co-chair, Jan Grenz, provided initial guidance with regard to the overall direction of the research.

During the course of developing the report the following other individuals sent additional comments, helped with coordination or provided technical assistance: Diane Conolly, Maryline Guiramand, Bob Macgregor, Eve Rickert, Charles Thrift and Dave Wilkins.

The report was prepared under IISD's framework agreement with Agriculture and Agri-Food Canada, whose support and interest in the work is thankfully acknowledged.

Executive summary

Agriculture by its very nature has a significant impact on the natural environment. Several converging trends make it difficult for the world's farmers to keep up with the growth in food demand from rising populations and changes in consumption patterns. These include increased water scarcity, the growing conversion of cropland to non-farm uses and more extreme climate events. Based on aggregate global trends and outlooks for the future, we can conclude that human efforts are not making enough positive difference. Many unresolved issues require understanding the links between farm-level practices and outcomes and impacts at different scales and time frames, including cumulative effects. As with the financial crisis, agriculture today must not run up a huge ecological debt to be paid by the next generation.

Metrics can help in the diagnosis of problems, but also in identifying solutions and driving technological changes in a direction that supports these solutions by pricing environmental inputs and consequences that are currently unpriced. There is increasing urgency to understand what practices and policies are effective for evidence-based decision-making at all scales, from farmer to policy-maker. Without this evidence, decisions will be uncoordinated, based on poor or incomplete information. Several trends in science, policy, markets and technology create new possibilities for the development and use of agri-environmental indicators at multiple scales.

Metric initiatives respond to these data demands by developing criteria, indicators, tools and reports to "measure" sustainable agriculture. While there is now a wide awareness of the sustainability concept, there is also wide interpretation of the definitions and components of sustainability, based on different disciplines and political beliefs and values.

The different interpretations of sustainable agriculture result in multiple conceptual frameworks for organizing and orienting the development of criteria and indicators. At the heart of the debate over different indicators are not only different views of sustainable agriculture but different approaches and information needs of users, as well as differences in how the indicators will be used. This diversity is related to the different lines of accountability, considering that indicators play a key role in the establishment of accountability to different users or stakeholders. "Measure what matters" has become a mantra. But measure *what* matters to *whom*, and *how*? Those interested in using these metrics cannot distinguish one initiative from another, making it difficult to make choices around which initiative best suits their information needs.

The *process* of developing the conceptual framework and indicators is as important as the indicators themselves in learning and building consensus. Indicators start to make a real difference even before they are implemented. By participating in the process, stakeholders explore what matters to whom, building a shared sense of vision and brokering consensus.

In attempts to capture the complex socioeconomic and environmental relationships in agriculture, many metric initiatives develop lists with hundreds of indicators, which are often process-based and not linked to higher-level outcomes. A proposed solution is a key impact approach, focusing on a few select things that really matter rather than seeking comprehensive coverage, and working with a limited set of indicators and impact proxies linked to models. Users of metrics disagree over the use of a few strategic headline indicators across commodities and regions. While there is general agreement across many initiatives and stakeholders of the key environmental impacts on a very generic level, such as water availability and quality, biodiversity, climate change and waste, this approach presents some challenges: key impacts vary geographically, and sustainability principles require a holistic and integrated approach. In addition, due to numerous technical, methodological and institutional challenges, not enough science-based evidence exists to build models linking practices to these key impacts across scales.

There are enormous opportunities to link the data from the hundreds of metric initiatives in existence or in development, operating across scales. Some attempts have been made to coordinate efforts, but the growing diversity of indicator initiatives demonstrates the limited success of harmonization attempts to date and the need for sustained efforts.

This paper reviews a number of the issues, constraints and challenges to using indicators at different levels to build this evidence. Assessment is a continuum from farm-level output to higher-scale outcomes and impacts across spatial and temporal scales, with different roles and responsibilities of stakeholders at each level.

The farm level is where sustainable agriculture practices can contribute to or mitigate negative environmental outcomes and impacts. Farming is an economic activity, within legal land ownership boundaries. Farmers will have little incentive to collect data on results and impact indicators that have no specific value for them or are off-site, difficult to attribute to particular practices or farmers, or used primarily for controlling farmer activity.

One approach to studying key impacts would be to focus on ecosystem services provided to and impacted by agricultural practices. These occur at higher levels and highlight the relationships among the different scales. While these measures are only in development, this approach also enables the monetization of these ecosystem services as part of the solution to more sustainable production practices.

Metric initiatives working at farm level need to develop links to higher-scale data sets to understand the sustainability of farming *systems*. Collection and analysis of large amounts of data will allow production of science-based models linking practices to impacts across crops, regions and systems. In order to accomplish this, harmonized approaches, partnerships and data-sharing are essential.

Definition and agreement on a set of higher-level goals with broad participation across levels is an ambitious task and poses numerous technical, methodological and institutional challenges. Strong leadership and a move beyond coordination to collaboration, consensus and compromise will be important factors in moving forward.

This does not mean one size fits all, but the process would benefit from standardization of terminology, criteria and indicator development standards based on international guides such as the Bellagio Principles and the International Social and Environmental Accreditation and Labelling Alliance (ISEAL) Code of Good Practice for Setting Social and Environmental Standards. Coordination and integration of various metric initiatives under common frameworks can help understand how the different levels are related and, together with open data platforms, can build the science and knowledge base.

In order to deal with the technical, institutional and methodological challenges, capacity needs to develop at various scales. Some of the needed capacity is technical, but capacity development is also critical for managing the process of indicator development; linking with other scales; integrating indicators into decision-making; and outreach, communication and branding.

A stepwise, evolutionary approach to measurement, reporting and verification systems should focus on a few strategic indicators as a starting point, embedded in a broad conceptual framework.

We make the following key recommendations for ways forward:

1. Standardize terminology and develop common conceptual frameworks to understand how the different levels, from farms to entire agroecosystems, are related.
2. Coordinate and integrate the various metric initiatives under common frameworks and provide open data platforms to build evidence of what works under what conditions. This knowledge is the foundation for building science-based metrics.
3. Base criteria and indicator development standards on international guides such as the Bellagio Principles or the ISEAL Codes of Good Practice.
4. Develop capacity.
5. Understand that developing indicators is an evolutionary process, and a perfect, comprehensive set will be elusive—but also unnecessary. The process is a political one, requiring collaboration, consensus and compromise.

List of acronyms

ARTEMIS	Advanced Real Time Environmental Monitoring Information System
BellagioSTAMP	Bellagio SusTainability Assessment and Measurement Principles
DMI	Dairy Management Inc.
DPSIR	driving forces, pressures, state, impacts and responses
EISA	European Initiative for Sustainable Development in Agriculture
FAO	United Nations Food and Agriculture Organization
GDP	gross domestic product
GIS	geographic information systems
IAASTD	International Assessment of Agricultural Knowledge, Science and Technology for Development
IDEA	Indicateurs de Durabilité des Exploitations Agricoles
IISD	International Institute for Sustainable Development
INFASA	International Forum on Assessing Sustainable Agriculture
ISEAL	International Social and Environmental Accreditation and Labelling Alliance
LEAF	Linking Environment and Farming
MDGs	Millennium Development Goals
NGO	non-governmental organization
OECD	Organisation for Economic Co-operation and Development
RISE	Response-Inducing Sustainability Evaluations

1.0 Introduction

The purpose of this paper is to define and help design tools and models to better measure and link environmental outcomes at the farm, agroecosystem[1] or landscape,[2] and national levels. While we look at information needs and information generation at various levels, a major objective is to explore the issues and the information infrastructure requirements of sharing and understanding the links between farm-level practices and higher-scale outcomes.

This paper is intended for use by those developing and using indicator systems that measure the environmental performance of sustainable agriculture. Another key audience comprises those who see the growing diversity of metrics initiatives and pursue harmonization as a strategic direction. While we address issues and information needs of key actors across levels, the main audience for the paper will be information generators and information users. These same stakeholders may both generate and use information; however, generally each group has different roles and responsibilities.

We will clarify some of the barriers that have prevented the various metric initiatives and standards systems from being better harmonized across multiple scales and more fully taken up by those who use and report on the information collected. Different stakeholders need different systems that help link impacts associated with farm-level activities with ecosystem-scale outcomes in order to understand the trade-offs of different management practices.

Problems and solutions to agricultural impacts lie in actions at the local level. Sustainable agriculture depends on understanding these linkages.

Many studies dealing with the environmental or socioeconomic impacts of production processes focus on an either/or approach that results in isolated data—*either* from the macro or global level (using statistical data) or the micro or local level, for example, case studies involving detailed house surveys. Voluntary standards systems[3] and the majority of current agricultural metric systems focus on collecting data at the farm level. These systems use means-based indicators, which measure compliance with specific farm production practices. Linking of these farm-level practices to the environmental effects of the farmer's practices is still only undertaken to a very limited extent, using proxies such as area certified. Ecosystem reporting, such as state of the environment reports or the Millennium Ecosystem Assessment, focuses on the macro or global/subglobal level without accounting for local conditions, actions or innovations. We have limited understanding of the contribution of individual actors to global threats or the state of specific local resources such as water.

1 Agroecosystems are a subset of ecosystems and may share some functions, but they include processed inputs such as fossil fuel energy and human and animal labour, and their outputs are not restricted to the immediate site of agricultural activity (the farm), but rather includes the region that is impacted by this activity.

2 Many definitions of landscapes exist, here we define them as any area that is spatially heterogeneous in at least one factor of interest (Turner, Gardner & O'Neill, 2001).

3 In this paper we define a *standards system* as including all component functions involved in the application of a specific standard, ranging from its definition via capacity building to verification and monitoring (International Social and Environmental Accreditation and Labelling Alliance [ISEAL], 2009a).

Both micro and macro approaches have limitations: national statistical data are in most cases too broad to attribute changes to specific sectors or commodities, while farm-level micromanagement and local case studies usually do not provide insights into important dynamics and impacts such as land-use changes, overwithdrawal of water for irrigation, and global warming. It is not possible to simply sum micro into macro. More importantly, the meso level (that is, the landscape or region) is often the functional unit within which biodiversity and natural resources are conserved. Initiatives, tools and clear responsibility for metrics are often lacking at this level. The objective of this paper is to examine the technical, methodological and institutional barriers to these linkages and what information infrastructure is needed to facilitate these linkages across levels.

The paper will illustrate how these barriers manifest themselves in the context of trying to measure and communicate performance related to some specific environmental issues from the farm up or the ecosystem down. Key lessons, as well as recommendations for the different actors, will provide the pathways for achieving improved environmental sustainability outcomes from agriculture.

Our focus is not on defining the impact areas of agriculture or recommending a specific set of common indicators; however, it gives clear guidance to those who are designing indicators on the principles that they should apply to develop meaningful (robust) and yet cost-effective schemes of measurement. Many initiatives are developing key impact indicators for sustainable agriculture, and other initiatives are working on harmonization efforts, discussed in Section 3.5.

The focus of this paper is on the "information infrastructure," the context for data collection and systems of information use. The concept is commonly understood in terms of information technology—the communications networks and associated software that support interaction among people and organizations. The term was popularized by Al Gore in the 1990s (Clarke, 2009). Here we use it to emphasize the need for interconnectedness between stakeholders and data at different levels for different purposes: accountability, communication, learning and decision-making.

This paper presents a framework for developing a core set of questions to define the systems needed and to identify the players who should be working together better, as well as their interests and capacities.

We limit our scope to environmental indicators associated with farm-level production and their linkage to ecosystem-scale outcomes. It is important to consider the entire supply chain and life cycle of a product; production of inputs for agriculture, packaging, processing, distribution, preparation, consumption and disposal all contribute to a product's overall impact on the environment and on communities. Strategies to measure and minimize these impacts are important considerations. Social, economic and ethical issues are also important parts of the sustainability equation. However, to do an analysis of the issues associated with information collected, reported and analyzed at different scales, this paper focuses on the growing and production stages at the farm level, linked to environmental performance at the ecosystem level. Most of the issues discussed are relevant to other product stages and socioeconomic aspects.

Why measure?

Agriculture or farming is a high-risk business, subject not only to pests and weather but also to changes in resource availability (scarcity or deterioration), market conditions and government policies.

Managing risk is an ongoing enterprise and requires constant learning and adaptation, as most farmers can attest. This involves regular monitoring of crops, reserves, markets and environmental conditions. Globalization of food production and supply has created longer supply chains with new risks and vulnerabilities. Transparency along the entire value chain can help to understand the socioeconomic status of the various actors, often far removed from one another, and the environmental sustainability of their actions.

Measuring serves as a foundation that promotes several goals. These include better decision-making (lowering risks and costs), an early warning system for emerging issues, sustainability balanced with development, understanding impacts and allowing for corrective action, identifying limits and opportunities, continuous improvement, accountability and communication.

You can't manage what you don't measure.

Recognizing the vital role of measurement, in 1987 the World Commission on Environment and Development (Brundtland Commission) called for the development of new ways to measure and assess progress toward "sustainability." In response corporations, non-governmental organizations (NGOs), academics, communities, nations and international organizations have made significant efforts to assess performance, often without any coordination or harmonization. In 1996 an international group of measurement practitioners and researchers first developed a set of guiding principles, known as the Bellagio Principles, focused on priorities that are applicable on global, national, regional and local scales, to form a common ground (IISD, n.d.). The principles emphasize openness (accessibility and transparency), comprehensive approaches, practical focus, communication (meeting needs of stakeholders, simple and plain language), broad participation, an assessment process for learning, sufficient institutional capacity, and the need for a coherent framework and goals. A second version is expected in late 2009.

Different key actors along the supply chain have specific information requirements that drive the need for metrics. The need for quality and regularly produced information to help understand the impacts of farming practices on the environment is increasing at all levels.

Measure what matters

What to measure is a fundamental question. Measure what is important. Activities and results that are essential to achieve specified goals should be measured. However, as discussed in the following sections, determining what is important to different stakeholders and agreeing upon goals can be complicated by the need for information at different levels within complex systems, different uses for information, *who* determines *what* matters, and *how* the information will be used and by *whom*.

How information is used by different stakeholders

Producers need metrics to

- improve performance;
- manage input costs and resource availability;

- manage outputs (waste streams, emissions);
- understand the longer-term effects of farming practices on their operations, including resource use intensity, soil and water quality, and market access, and calibrate practices based on this understanding;
- understand how their operations are impacting common environmental property resources such as a shared water source or pollinating insects;
- communicate to buyers about the key environmental impacts of production;
- mainstream market access by reducing the reputation and operational risks that producers face as supply chain partners;
- access niche markets with potential price premiums;
- access capital at better terms (because of the reduced risk they represent as borrowers); and
- demonstrate responsible stewardship to policy-makers, public interest journalists and concerned citizens.

Buyers (exporters, importers, procurers and so on) need metrics for

- assurance that the goods and services were produced under socially and environmentally sustainable conditions;
- mitigation of operational and reputation risks in their supply chains;
- identification of environmental risks and opportunities;
- access to particular markets that have stringent environmental standards as a precondition of participation; and
- addressing consumer demand and activist pressure.

Uses of information

The different uses and purposes of information can create tensions among stakeholders. Evaluation research and metric initiatives use various classifications of use, with sometimes conflicting terminology. Most agree, however, on three distinct types of use: (1) instrumental use (2) conceptual use and (3) legitimizing use (see, for example, Amara, Ouimet & Landry, 2004; Beyer, 1997).

Many particpants involved in indicator development assume the indicators are for instrumental uses, such as decision-making—that is, as basis to change actions or behavior. However, they are often used for conceptual use (learning and education) or legitimizing use (to convince others of an intervention or to defend actions such as a project). Often the use has a combination of purposes, without an explicit understanding of these uses or the tensions that arise because they assign very different roles and purposes for indicators (Rosenström, 2009).

Manufacturers and retailers use metrics to

- access new business opportunities, such as new markets, distribution innovations and supply chain localization;

- mitigate corporate operational and reputation risks;
- find means to reduce supply chain costs by improving efficiency of resource use;
- improve relations with communities through better information;
- improve relationships between business operations and governments and regulators;
- offer transparent and credible information about products, services and brands to substantiate claims that the company is "walking the talk";
- strengthen management processes;
- inform corporate strategy; and
- increase market share (World Business Council for Sustainable Development, 2008b).

Certification schemes and NGOs use metrics to

- mitigate the impacts of agricultural practices on the natural environment;
- report environmental performance objectively to potential buyers of certified products (food manufacturers and retailers);
- report to growers who might wish to join a certification scheme on how successfully adopting practices required by a particular certification or management scheme will help achieve improvement in key environmental impacts;
- improve certification standards and project interventions;
- provide accountability and transparency in communications to public and private donors regarding results and the extent to which goals are being met;
- demonstrate the credibility of a particular standards system or certification label, such as organic, Fairtrade, Rainforest Alliance Certified or the LEAF label; and
- demonstrate to country governments the positive impacts of standards and that they are not trade barriers.

Consumers and citizens use metrics to

- inform their purchasing decisions by understanding the environmental impact of different agricultural production systems and
- address concerns about food safety, such as water contamination or the use of chemical inputs.

Policy-makers need metrics to

- understand the effects of policy on agricultural practices;
- understand the needs for new policies to mitigate the environmental impacts of agricultural practices;
- evaluate competing green claims and the extent to which the environmental components of sustainability initiatives meet public environmental policy objectives;
- identify emerging issues and detect broad changes;
- plan and manage public goods and services; and
- ensure that global suppliers meet national regulations and policy objectives.

The audit and evaluation community[4] uses metrics for

- understanding and analyzing key environmental drivers;
- including environmental targets and objectives in performance appraisals; and
- monitoring and reporting on environmental performance.

People engaged in wide range of activities and a multitude of indicator systems at local, national and international levels are dedicating large amounts of resources to providing analysis and developing core sets of indicators based on their objectives and the needs of key stakeholders. The lack of harmonization or coordination between systems results in different procedures, processes and indicators. At the same time, information technology has created an overwhelming amount of information to sift through to make decisions. This trend is likely to continue, resulting in higher volumes of information and fragmentation and creating competing measures. This increases the risk of making decisions with incomplete information or misleading data. A good example of this is the recent collapse of financial markets in the United States, when many financial experts, using lots of data, were caught off guard or made investment decisions based on indicators that painted a misleading picture of the U.S. housing markets. Thus the need arises to differentiate between information that matters and information that is of secondary importance or irrelevant—for a specific context. If decisions are to improve the environmental performance of agriculture, then it will be necessary to provide a framework for analysis, measurement and response to understand the links between farm-level practices and higher-level outcomes. Defining what matters regarding the environmental aspects of agriculture begins with looking at the key environmental impacts.

4 Government audit and assessment agencies such as the Canadian Commissioner of the Environment and Sustainable Development or voluntary private initiatives such as the Prince of Wales' Accounting for Sustainability Project.

2.0 Key concepts

In the past 50 years globalization, population growth, consumption shifts, market pressures and technological innovations have spurred agriculture worldwide to increase output and productivity. In the words of Agriculture and Agri-Food Canada (2005, p. iv), "This has engendered structural changes in the industry, characterized by the adoption of new technologies and a gradual shift towards larger, more intensified operations."

Intensification of agriculture—increasing yields for a given quantity of inputs such as labour, land, time or fertilizer through modern agronomy techniques such as breeding, pesticides, fertilizers and technological improvements—has fed our expanding societies. Increased yields enable less conversion of land to cultivation and may cause reduced rates of deforestation. The World Bank estimates that GDP growth in agriculture is at least twice as effective in reducing extreme poverty as GDP growth originating in other sectors. As the largest managed ecosystem in the world, agriculture holds the potential to halt, if not reverse, rapid ecosystem degradation through environmentally sound practices and policies. Because more than half of all species exist primarily in agricultural landscapes outside protected areas, biodiversity can be preserved only through initiatives conducted with and by farmers (World Bank, 2007). So can we consider such initiatives win-win?

They have some big losers. Changes in agricultural practices, particularly the intensification and concentration that has occurred over the last century as part of the green revolution to feed growing populations, have caused widespread ecological damage and negative socioeconomic effects (Atkinson et al., 2004). In modern agriculture, factors that limit production, specifically water and nutrients, are provided through synthetic chemicals and irrigation. The doubling of agricultural food production during the past 35 years is associated with huge increases in nitrogen and phosphorus fertilization and in the amount of irrigated cropland and total land under cultivation (Tilman, 1999). Cumulative effects on biodiversity, water quality, water availability and climate change (through nitrogen and carbon dioxide emissions) are apparent worldwide.

Demand for agricultural primary products is predicted to double by 2050 due to population increases and changes in prosperity and increased consumption of meat and other livestock products. The consequent rise in agricultural output is expected to increase pressure on natural resources and price volatility. In 2007, for the first time, the world's urban population exceeded its rural population. The demand for high-value primary and processed products is rapidly increasing, driven by rising incomes, faster urbanization, liberalized trade, foreign investment and advancing technology. Speculation in food stocks and competing uses such as biofuels and cereals for non-human consumption contributes to pressures.

The agricultural green revolution, which began in the 1950s, was based on research and focused on increasing productivity—growing more green things—not, as the phrase is more commonly used to mean today, on environmentally friendly, chemical-free production methods. India experienced massive famines in the early 1960s and embraced these revolutionary new scientific techniques for high-intensity, chemical-based, high-yield agriculture systems. Wheat, rice and cotton varieties replaced traditional crops of grains, beans and vegetables. Today the small region of Punjab is India's breadbasket for wheat and rice (Zwerdling, 2009).

Growing pressures

- World population to exceed 9 billion by 2050
- Urban population exceeds rural population
- World demand for food to increase 70% to 80% by 2050
- Agriculture single largest global user of fresh water, accounting for 70% of withdrawals
- Use of water for agriculture expected to double by 2050
- Loss of fertile land through processes such as erosion and salinization
- More than 60% of ecosystem services degraded or in decline
- 1.5 billion more people by 2030, with 1.5 billion undernourished
- 2.5 billion people living in water-stressed or water-scarce conditions
- Demand for cereals and tubers increased by more than 50%
- Prosperity driving increases in meat and livestock product consumption: demand doubled in developing countries
- Global food prices up 83% in 2008, and a 50% to 200% increase in select commodity prices
- 25% of population food insecure due to land degradation
- Top-10 fished species fully fished or overexploited
- Food production estimated to be short of demand by up to 25% in 2050
- Agricultural involvement in rural areas of developing countries where 73% of the world's poorest live

Sources: Millennium Ecosystem Assessment (2005); United Nations Food and Agriculture Organization (FAO, 2006); Natural Value Initiative (n.d.); Worldwatch Institute (2006; 2009); WorldBank (2007); Häni, Verschuur & Goodlass (2005); Nellemann (2009).

The long-term consequences of adopting the new practices are now, after 30 years, creating devastating social, economic and environmental impacts. Groundwater tables are sinking as much as a metre a year. Chemical fertilizer use has tripled over this period to achieve the same yields. Chemical-resistant pests continue to ruin large amounts of crops (Zwerdling, 2009).

For India to continue on this route while trying to feed a fast-growing population would have serious implications. As one local farming commission director predicts, "If farmers don't drastically revamp the system of farming, the heartland of India's agriculture could be barren in 10 to 15 years" (Zwerdling, 2009).

An impact defined as "key" is one that is important to stakeholders or for the sustainability of the system. This again introduces values and the need to define these key impacts in terms of local vulnerabilities and risks. One example of a key impact in a local context comes from Lake Balaton, Hungary. This Central European lake's physical and chemical properties made it especially vulnerable to the eutrophication caused primarily by heavy fertilization used in farming during the Communist era. In particular, its shallow depth caused high nutrient levels to accumulate from runoff, resulting in algae blooms. The wake-up call for the local populations was not only the physical effect on the lake but also the economic impact of the resulting eutrophication on declining local tourism.

Climate change is also a key impact, and it can be measured using the proxy of one or more of the principal greenhouse gases. In practice, for cropping systems this means measuring the nitrogen use efficiency of a particular crop. The locality in which impacts accumulate is not important for greenhouse gas emissions and climate change; what is released at the farm level contributes to a globally fungible pool of emissions. Deforestation matters where there is forest and where a lot of deforestation is occurring, water use is important where water is scarce. In both these cases the impacts are defined as key because stakeholders' interests, such as livelihoods or cultural practices, are adversely affected by the use of natural resources for agriculture. Where local, cumulative impacts are important, the definition of what is key should start with defining the natural resources that are scarce or critical (that is, which natural resources are close to thresholds and limits) in the locality, or with the values attributed to natural resources that are endangered by agricultural activity.

Impacts occur at multiple scales and generally over several years or even decades. For example, effects on biodiversity may lag behind changes in agricultural practices. Pesticides may bioaccumulate, and some agricultural pesticides are now found even in polar organisms and marine sinks far distant from the places where they are used. Furthermore, impacts are interrelated, with many long-term cause-and-effect relationships not well understood. For example, erosion or contamination of soil can directly affect aquatic populations and livelihoods dependent on them, but can also have indirect impacts, such as on recreational activities or the costs of water treatment and cleanup.

Some consensus has emerged among stakeholders working in metric initiatives on the overarching categories for key agricultural environmental impacts such as biodiversity, climate, soil, toxicity and water (Clay, 2004), but it lacks details of particular goals or links to specific agricultural practices. The establishment of targets has distinct methodological challenges, and the process is often not explicit in metric initiatives. Ideally all stakeholders at various levels could measure, report on, analyze and compare one set of indicators for each of these categories of impact and use the results to improve practices. Many metric initiatives began with the intention to develop a core set of 6 to 10 or 10 to 15 key impact indicators. An internal analysis conducted on metric initiatives by the Sustainable Food Laboratory showed a general consensus with respect to the overarching environmental impacts of agriculture as well as specific impacts associated with certain initiatives and commodities.

In 2006 the International Forum on Assessing Sustainability in Agriculture (INFASA) convened a forum and produced a report, From Common Principles to Common Practice (Häni, Pintér & Herren, 2007), to begin a dialogue on synergies and tools for common key impact assessment tools and frameworks.

2.1 Why measure the environmental impacts of agriculture?

"The ecological footprint of industrial agriculture is already too large to be ignored and projected increases in future global environmental changes could make the footprint even larger." (IAASTD, 2009b, p. 23)

Two international studies, the International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD) Agriculture at a Crossroads reports and the Millennium Ecosystem Assessment, have different objectives and audiences but reach some important common conclusions. Both outline several key trends and impacts that highlight that business as usual is just not a feasible option, and both emphasize the need for better understanding and monitoring of key impacts.

Both reports found that consequences were often not foreseen, as they occurred over time and some occurred outside of traditional farm or business boundaries, distant from the original activities. For example, inappropriate fertilization intended to increase yields may be washed away, triggering eutrophication and large dead zones (areas of low oxygen) in coastal areas, such as the Gulf of Mexico, and some lakes. A critical observation that should be emphasized as one of the key obstacles to sustainable agriculture is that *sustainability initiatives focus on controllable practices within the parameters of a farm, but the impacts often lie outside the farm's boundaries and the farmer's direct control.*

Based on the aggregate global trends and outlooks for the future outlined in these reports, it is clear that our efforts are not making enough positive difference. The rapidly changing and globalizing food economy is changing where and how food is produced and processed. The growing complexity of the system and distance between user and provider raises the needs of stakeholders along the supply chain for transparency regarding the environmental impacts of food production.

Over the long term, increased food production is a prerequisite for a growing and shifting world population. Managing agricultural systems in a way that avoids deterioration in the supply of agricultural products to well over 6 billion people without irreversibly degrading the integrity of natural and agroecosystems is a huge challenge. As Nelleman (2009, p. 5) writes, "The experts argue that unless more sustainable and intelligent management of production and consumption are undertaken, food prices could indeed become more volatile and unaffordable as a result of escalating environmental degradation." This recent United Nations Environment Programme study on the food crisis proposes several short-, mid- and long-term options—all of which require robust monitoring and assessment of progress.

Whether agriculture is viewed from the perspective of a small family farmer or large-scale agribusiness, "sustainability" has become the code word for balancing the need to feed rising populations while considering environmental and socioeconomic impacts. However, we lack a clear understanding of the complex relationships and trade-offs between economic development and environmental integrity.

Sustainability rests on the principle that we must meet "the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987).

While there is now a wide awareness of the sustainability concept, there is also wide interpretation of the definitions and components of sustainability that we will purposely seek to sustain over time, based on different disciplines, political beliefs and values. Because there is no widely accepted, pragmatic definition of sustainable agriculture (*what* to measure), there is no single, globally applicable set of metrics to measure the sustainability of agricultural practices. Even a precise global definition existed, agricultural practices that will mitigate impacts also depend on site-specific variables such as climate, ecology, geography, demography, affluence and regulations.

Reflecting these differences, numerous sustainable development metric tools and indicators have evolved across all levels. Since 1995 more private, international, food-related standards have emerged than in the previous five decades combined (Giovannucci & Purcell, 2008). These are discussed in more depth in Section 3.5. The proliferation of methods and metrics signals not only the differing needs for information but also the underlying confusion about what needs to be done. We have failed to communicate the paradigm shift that those involved in agriculture need to undertake in order to meet the growing food demand using "sustainable practices." A cacophony of voices represents different interests trying to define "what is needed." Without a clear direction, management practices may fall back on what has worked to double food production over the past 35 years—not sustainable management, but sustained yield management—with the consequence that cumulative impacts on common environmental property resources such as water or pollinators continue to be ignored. The United Nations Food and Agriculture Organization predicts that 80% of the needed increase in food production will come from greater input use and improved technology on existing land (FAO, 2007b).

So do we need more metrics? A more strategic approach to metrics may be the key. The proliferation of metric systems and definitions has created confusion on the part of decision-makers and members of the public who are interested in information about impacts in a simple format.

The lack of harmonization or even standardized definitions, tools or reporting is a barrier to passing information across levels for decision-making. Other risks include the possibility that decision-makers will waste resources reinventing the wheel or developing various information technology platforms, or that they will become frustrated with information overload, competing initiatives and lack of comparability; that the lowest common denominator will become the bar; or that customers and consumers will become confused as to the credibility of green claims.

A common framework and set of definitions also allows for comparisons across initiatives and for building bodies of information to aid long-term learning and improvements in decision-making, a key to sustainability.

2.2 Why key impact indicators?

"The acceleration of change reduces the time from recognizing the need to make a decision to completing all the steps to make the right decision. The number and intricacy of choices seem to be growing beyond leaders' abilities to analyze and make decisions." (Glenn, Gordon & Florescu, 2009, p. 4.)

The complexity of the issues underlying sustainable agriculture, and the costs of measuring them, are much larger than commonly thought. The stakeholders involved in developing indicators need a stepwise, evolutionary approach to measurement, reporting and verification systems. This approach should focus on a few simple but strategic indicators as a starting point within a common conceptual framework. Many metric initiatives have developed exhaustive lists of indicators in order to monitor the complexity of sustainability issues and increase buy-in from their stakeholder groups. The costs and effort involved in gathering and analyzing this information are high. These initiatives are increasingly recognizing the need for a small subset of credible indicators that are manageable and feasible, rather than extensive sets that stretch capacity. "At the end of the day, it is better to have indicators that provide approximate answers to some important questions than to have exact answers to many unimportant questions" (United Nations Development Programme, 2002, preface).

Critical, "keystone" or headline indicators that may be able to illustrate the wider health of systems and monitor key economic, social or environmental changes have been discussed for more than a decade now. This approach focuses on a few select things that really matter rather than seeking comprehensive coverage.

This pragmatic approach is described in the report for the United Nations Division for Sustainable Development's report *Sustainable Development Indicators: Proposals for the Way Forward* (Pintér, Hardi & Bartelmus, 2005). The report looked at the potential of short, core indicators, calling them "headliner indicators" and looking at the experiences of organizations using smaller core sets from larger sets produced by organizations such as the World Bank and the Organisation for Economic Co-operation and Development (OECD) for national-level sustainable development indicators, or the United Nations Human Development Report. Other organizations, such as the Consultative Group on Sustainable Development Indicators, propose focusing on only a small set from the outset; for example, the Royal Society for the Protection of Birds calls for 10 global headline indicators (RSPB, 2003). An approach using fewer strategic indicators is pragmatic, but one important conclusion of the report is that "the composition of such a set, even if it is the result of a consensus based process, will always reflect the preferences and biases of those involved in making the selection" (Pintér, Hardi, & Bartelmus, 2005, p. 7). At the global level, many powerful institutions have specific priorities and interests, potentially resulting in use of indicators for legitimizing purposes rather than for decision-making, learning and improving. One of the most well-known sets of global indicators, the Millennium Development Goals (MDGs), contains close to 50 indicators that are not universally applicable or accepted, with criticisms that they are missing areas (such as the environment) or too focused at the national level.

A WWF study of existing agricultural metric systems highlighted the fact that there is no global system of evaluation for the environmental impacts of agriculture (Perkins, Payen & Potashnick, 2008). However, currently stakeholders across levels have consensus on the need for a simpler, strategic approach and agree this approach can be applied to indicate environmental impacts, even if the picture is imperfect. They agree that they need to work with focused indicator sets and impact proxies linked to models rather than with long, all-inclusive, holistic approaches, which are impractical. Accompanying this need is the requirement, brought in part by the information revolution, to better differentiate between information that matters in any given context and information that is of secondary importance.

Several initiatives and commodity roundtables are working together to define the key impacts of particular agricultural commodities. The WWF has spent over a decade convening multistakeholder commodity groups (roundtables) to identify the key social and environmental impacts of agriculture. They have identified six categories—climate change, water pollution, water shortages, toxicity, biodiversity loss and soil degradation—and have spent several years testing the key impact approach (Perkins, Payen, & Potashnick, 2008). These categories can serve as a framework for defining *what* matters to *whom*, at least from the perspective of a global environmental conservation organization concerned about the environmental impacts of agriculture. In the short term this approach may result in partial coverage; however, it can be valuable if it acknowledges with stakeholders that it is a learning process (conceptual use).

There is also debate about the possibility of using a few strategic headline indicators across commodities and regions. Many initiatives generally agree on the key impacts at a very generic level, such as those on water, biodiversity and climate change. These are often referred to as "meta issues." Selection and definition of one or two indicators for each of these meta issues has several benefits: comparability, simplicity, ease of understanding and general agreement on the major environmental impacts of agriculture. The three main issues challenge this approach.

Key impacts vary geographically. Some counter that the key impacts of any particular commodity vary from growing area to growing area, and that even within one country different growing areas face site-specific environmental conditions. While this is true, some impacts are still very widespread. Additional or adapted indicators may be needed for some producers in some areas, within the overall framework.

Sustainability principles require a holistic and integrated approach. The tenet that farming needs to be managed in a holistic and integrated fashion is one of the core Bellagio Principles and underpins almost all international frameworks of sustainable development. Just as with the growing awareness that environmental impacts cannot be ignored within sustainable development strategies, the socioeconomic and environmental issues are interrelated, with trade-offs and multiplier effects. The risk when focusing on a few environmental indicators is that seeking to reduce specific key environmental impacts could distort the management of the farm system and even produce adverse outcomes elsewhere in the system—outcomes that are not monitored. In addition a number of external factors, such as climate change and pest pressures, are independent from farm management and can influence environmental outcomes.

Key impacts are not necessarily linked to practices. Most studies of the environmental effects of agriculture relate to inputs and outputs. These are commonly the basis for most agricultural commodity metric initiatives (including certification standards), which focus on prescriptive indicators of farming practices—that is, indicators that measure compliance with prescribed practices. Few indicator initiatives or studies look at the higher-level key environmental outcomes or impacts of farm practices.

Other challenges to key impact indicators

A smaller metric set risks leaving out an important element. Indicators need to be simple to understand and implement; however, they represent complex, dynamic relationships. A simplified approach risks ignoring complexities or important issues. Particularly if incentives such as market access are involved, there is also the risk of farmers chasing numbers to get positive results with one key impact. The key impact approach may not capture the interdependence and dynamic processes involved in the agricultural production system.

In addition to methodological issues, a focus on a few key impact indicators also raises technical issues. Many metric initiatives, including certification standards, have focused on farm-level management activities using input and output indicators. However, the biggest environmental problems are linked to common property resources at the landscape and catchment levels. Management practices and actions take place at the farm level. These are associated with longer-term, larger-scale outcomes and impacts. The dilemma is how to measure and manage activities at the farm level at one time scale and connect to larger-scale, longer-term impacts (meta issues).

Researchers consulted in the 2008 WWF study thought it was not possible to make the small set globally applicable, and that scientists had a lot of work to do to make the indicators more efficient. The report concluded that the external conditions that are specific to a region make it difficult to have a standardized set of indicators. Furthermore, different impacts occur according to the commodity that is produced, also complicating the development of an indicator set that is common for different products. The researchers recommended development of a small common set of indicators complemented by two or three indicators specific to each sector or product, with regional indicators in the case of specific threats or risks (Payen, 2008).

Glocalization

The term *glocalization* has been used to show the human capacity to bridge scales (local and global) and to help overcome meso-scale, bounded, "little-box" thinking. The conceptual framework for metrics should be widely applicable but reflect local realities. The system and approach can use a holistic framework. Implementation, however, focuses on a limited set of issues and indicators. This dichotomy follows the principles of thinking globally (the holistic framework) and acting locally (a few key issues are critical; looking at what matters within the context). The system will evolve over time through learning and adaptation. This approach allows farmers, buyers and policy-makers to build up evidence resulting from monitoring, creating the incentive to pursue wider measurement and assessment practices. Coverage can be increased, but with increased complexity and cost of transmission. Moreover, this does not assume the smaller subset of indicators can be applied universally.

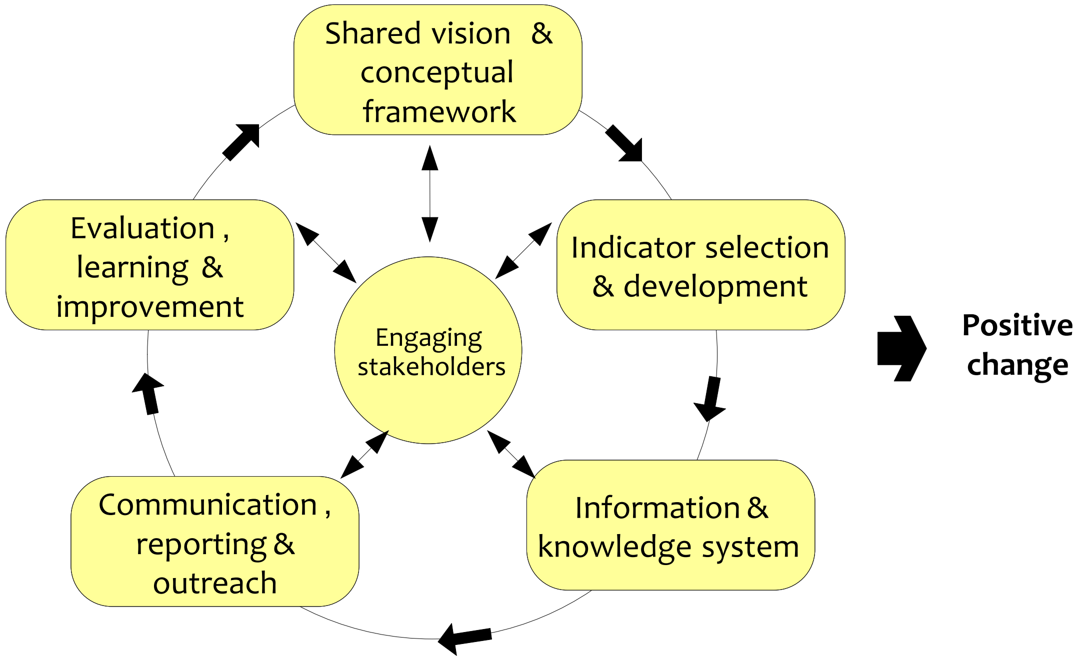
2.3 What makes a good indicator?

Users of indicators often forget that indicators just indicate something such as change; they do not explain why it has occurred. Precisely what they indicate depends on the context and can be interpreted by different users for different purposes. Understanding the role of indicators as a tool assists in defining a good indicator.

The lack of agreement on what to measure has led to systems with hundreds of indicators—many focused only on practices, not outcomes. The uncoordinated process of developing indicators also leads to confusion. Most metric initiatives begin with a general framework of sustainability along social, environmental and economic themes. Participants generate exhaustive lists of indicators based on existing criteria and indicator systems. Then, in a series of workshops and conference calls, they pare down these lists to a smaller set, which is then divided up or assigned to expert groups to define. The interdependence of the issues and the focus on priority key impacts and risk areas gets diluted. An underlying premise of the work of INFASA was that current measurement tools and processes are often not well-suited to the task. One study recently reviewed current initiatives that began with the ambition of selecting 10 to 15 key impact indicators, and revealed that each of the initiatives had actually settled on 50 to 100 indicators, the majority of which were input/output based (ISEAL, 2009a).

The organizations or initiatives developing the indicators tend to focus on the indicator itself rather than the overall system and its necessary pieces. The majority of time and resources are spent on the indicator selection and development stage, most likely because it is tangible and easier to understand.

Figure 1 Indicator system. Adapted from IISD and the United Way of Winnipeg (2007).



The other elements cannot be an afterthought but should be considered from the outset, noting several important points. The process starts with identifying the elements of a vision or values, in order to establish the framework. An information-and-knowledge system must be in place when data collection starts. Too often it is an afterthought, without dedicated resources and capacity. Engagement in the centre is key at all stages. Communication and outreach is an important investment for information use and uptake. Evaluation, learning and improvement are also key, and are discussed more in Sections 2.3.3 and 2.4.

The following section does not go through these steps, but looks at the challenges and obstacles specific to indicator selection and development, as this is where most metric initiatives have focused their resources.

2.3.1 *Indicator selection*

Indicators are often selected based simply on data availability and ease of understanding. Inputs (such as amount of fertilizer applied) and outputs (such as the nitrogen content of soil) are tangible and easier to measure than outcomes. Few studies look at actual long-term impacts because of the lack of data that could help understand the cause and effect connecting agricultural practices with environmental impacts.

Several factors complicate this:

- The impacts themselves vary based on a wide range of factors, including the type of agricultural system, other activities, land preparation, farming operation practices, agro-processing, climate, and socioeconomic issues such as population densities and levels of economic development.
- The impacts are cumulative, so that the first farm in a forest leaves plenty of room for wild nature, while the last one leaves none.
- A time delay occurs between an action and its ecological consequences. The delay may be twenty years or more, but it can also be quite short.

If the analysis originates from an ecosystem perspective, it needs to take into account not only agricultural activities but also all other pressures that contribute to the same impacts (for example, the contribution of agriculture and human sewage to phosphate-loading of aquatic systems).

Indicate what?

Numerous tools and criteria are used in developing, defining and selecting good indicators. It is fundamental to understand explicitly *what* is being measured and for what purpose. What goals and objectives, in terms of the key environmental impacts of agriculture, need to be addressed? The indicator must be constantly reviewed for relevance in terms of the objectives, processes, institutional context, capacities and stakeholders.

Too often metric initiatives begin with excellent intentions and a broad concept of sustainability goals, and then create a long list of indicators based on existing sets of indicators. Large multistakeholder advisory committees are set up to create a sense of inclusiveness and stakeholder involvement. However, these participants seldom have an explicit framework or common understanding of the goals and objectives, even among the participants in the initiative (Russillo, 2008).

At the heart of the debate over different indicators are not only different disciplinary approaches and views of sustainable development but also the information needs of users, as well as differences in how the indicators will be used. This is related to the different lines of accountability, considering that indicators play a key role in the establishment of accountability to different users and stakeholders.

Different disciplinary approaches, views of sustainable development, information needs of users and differences in how indicators will be used all lead to the proliferation of various metric initiatives and lack of harmonization of indicator sets. Measure *what* (key impacts) matters to *whom* (stakeholders), and *how* (indicators)?

Those interested in using the information from data gathered on indicators cannot distinguish one system from another, making it difficult to make choices around which initiative best suits their needs. As a response, other initiatives have emerged in recent years to promote harmonization or transparency of the various metric initiatives and systems discussed in Section 3.5.

In the following section we outline some of the basic concepts for developing good indicators, but it is important to understand what indicators are and are not important before selecting and qualifying them. Indicators are "partial reflections on reality, based on uncertain and imperfect models" (Meadows, 1998); tools, not ends; context-dependent; pieces of quantitative or qualitative data that need interpretation; based on models that are evolving; and selected based on values, experiences and world views, and they may often seem like trying to aim at a moving target.

Equally important to robust indicators are guiding principles such as the Bellagio Principles, which emphasize openness (accessibility and transparency), key indicators and standardized measurement methods, communication (that meets needs of stakeholders and is in simple, plain language), broad participation, an assessment process for learning, sufficient institutional capacity and a coherent framework and goals.

All indicators are not created equal

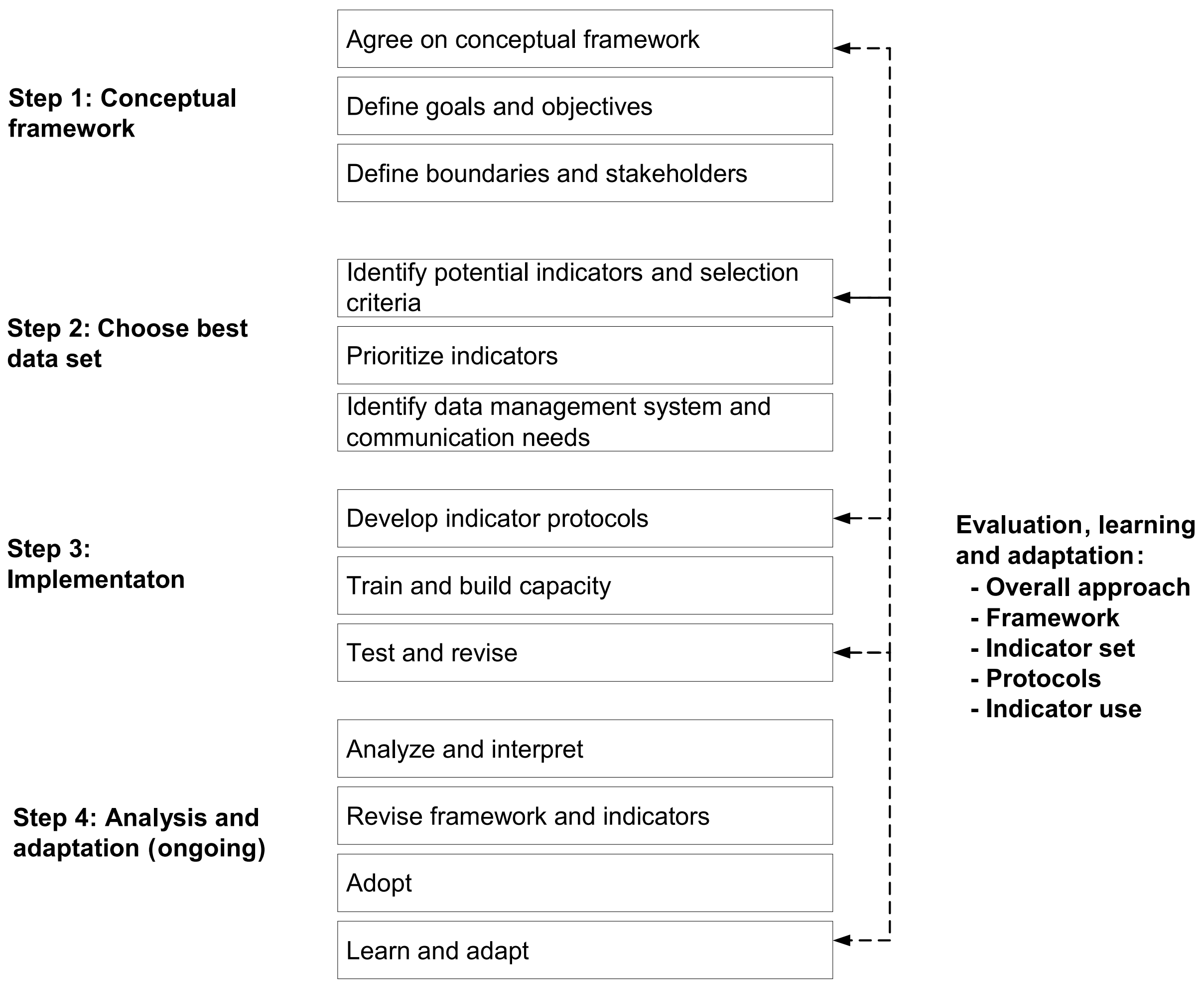
Developers of indicators must understand that there are different spatial and temporal scales of indicators. Ideally a small set of core impact indicators would be applied across scales and regions, becoming the key impact indicators. Long-term, broad changes in the environment and social conditions are difficult to isolate, measure and attribute to practices at the production unit level. It may take five years or more to realize impacts, and many other factors outside the scope of the farming operation can affect farm- and higher-level outcomes. Ecosystem approaches advocate spending decades to monitor change. Can we wait that long when we are not even sure of the effectiveness of the indicators themselves? A more practical approach is a combination of input-output, outcome and proxy indicators that are clearly linked to outcome models and impact objectives, as discussed in the following section.

Indicators can be quantitative or qualitative. Some argue that quantitative indicators are more comparable, rigorous and credible. However, qualitative aspects are equally valuable when monitoring dynamic aspects of sustainability such as vulnerability or risks. Qualitative indicators can be made quantitative indicators through scales or scoring. Indicators can also be normative or descriptive. Different types serve different purposes for decision-makers; for example, as early-warning, stress or state indicators. They can also be portrayed differently; for example, as raw data, ratios, percentages, scales or indexes. Decision-makers can set benchmarks or minimum standards of performance with respect to a particular metric. These indicator characteristics—their typology—often create debate during indicator selection over which indicator is "better."

2.3.2 Indicator development

Developing indicators is a phased process that requires adjustments and improvements over time. The process of indicator development is as important as the indicators themselves, often serving as a catalyst for dialogue, focusing attention on key issues, obtaining information on the state of action conditions, setting goals and identifying strategies. The process should be participatory at all stages. Participatory processes increase buy-in, ensure relevance and serve as tools for communicating with and training stakeholders.

Figure 2 Indicator life cycle.



Numerous books and manuals cover indicator development, but the following table outlines the main steps, highlighting the challenges and issues for metric initiatives at each stage. During each stage there will be compromises between what is ideal and what is practical, what is scientifically sound and what is technically and economically feasible. Many of the challenges can be overcome by taking a more holistic approach to indicator selection and development (Figure 2) rather than approaching it as an isolated exercise.

Table 1 Steps and challenges in the indicator life cycle.

Step	Challenges
Define goals and priorities: the conceptual framework	• Trying to develop one-size-fits-all indicators within limitations; for example, limiting indicators to the farm level while trying to address landscape- and national-level issues such as watersheds and food security. • Obtaining participation from key stakeholders with different values; perspectives and information needs can create competing interests. • Resolving conflicts arising from single- or limited-interest groups pushing their champion indicators. • Aligning these competing interests with the key elements and risks (issues such as threats, impacts, thresholds and dependencies). Defining goals and priorities in broad, generic terms to meet all stakeholder expectations.
Choose the best available data set	• Focusing and developing a core set of indicators based on the defined goals and priorities collected during the first phase. It is important to continually ask, What do decision-makers, who could be farmers, buyers, marketers or policy-makers, really need? Are they aware of what they need? Would they have the capacity use the information if they had it? • Establishing a rigorous and transparent selection process. • Refining poorly defined protocols (what, how, when, who, where and why). • Defining data management systems, including communication and reporting needs (considering the user and the appropriate format).
Implement the metric and collect data	• Developing and defining protocols for data collection, reporting and analyzing (who, what, where, how, when and why). • Developing and designing a data-management system during the conceptual phase. • Allocating resources and training for data collection and methodologies. • Adapting and refining indicators needed in early stages based on learning. • Aligning the values of data generators and data users.
Analyze and adapt	• Addressing the information and knowledge system, communication, reporting and evaluation from the outset to make sure data is usable for analysis and of value for decision-makers. • Building in review, analysis, reporting and dissemination to avoid being data rich but information poor. • Clearly and concisely analyzing the data, including the indexes, maps and visuals. • Regularly reviewing the indicators to ensure that they are relevant and reflect changing conditions and learning.

One of the continual tensions in indicator selection is balancing between the ideal and the practical, defining measurable and credible indicators that are also simple enough to be useful. Overly complicated or technical indicators may be difficult for users to understand and interpret and have also not been very effective in communications. Simple indicators that are measurable will keep those involved in collecting and using data committed and motivated.

"Practical" can mean selecting fewer, strategic indicators with a stepwise, evolutionary approach. In the next section we discuss how to recognize the potential trade-offs between holistic and strategic and the danger of unintended negative consequences from using a few simple indicators.

Main objectives during this phase of indicator development are to find robust indicators and verifiable measures to support the selected indicators. These must be simple to understand, cheap to implement and effective in measuring progress toward goals. Stakeholders must understand that indicator selection is not just about balancing interests but also about actual trade-offs, such as scientific rigor versus feasibility.

Sustainable agriculture systems, like all systems involving direct interaction with ecosystems, are complex and dynamic, that is, a moving target. Provision of the information required to make informed decisions may require technical, complex indicators. Relying on simple indicators may not satisfy the informational needs of some users. Thus, simplicity of the indicators requires that they be explicit in their meaning. New tools can assist in collecting, analyzing and communicating complex indicators in simpler terms. For example, several of the online carbon calculators use sophisticated modelling to take a few basic pieces of information and express them using a simple index: the "carbon footprint."

The costs of measurement and reporting are often a decisive factor in indicator selection. Measurement and assessment has ongoing costs and is often underestimated, creating system failures. The use of secondary data when available and appropriate is preferable owing to the higher costs associated with primary data collection. Many monitoring and evaluation manuals include the advice that existing data sources can be adapted as indicators. One screening criterion for the Keystone Alliance for Sustainable Agriculture's Field to Market indicators is the existence of publicly available data. Use of existing data is efficient and may sometimes be the only option. Those working on developing metrics should review the adequacy of existing information, however, since "current approaches are in danger of falling into the 'availability trap'—since the results are usually needed immediately, high priority is given to using existing or easily obtainable information" (Brang et al., 2002, p. 114). Existing data is often gathered for other purposes and may not provide the information needed.

Balancing ideal, comprehensive lists of indicators developed with stakeholders against time, budget and feasibility constraints has often led to lists representing the "lowest common denominator": initiatives that are bogged down in endless discussions of feasibility, leave out issues, or fall back on input/output-level indicators as proxies.

Ideally the priority concern is the indicator's efficiency in measuring progress toward goals. The challenge during the initial selection process is to understand up front how effective the indicators will be in measuring progress. Often this criterion can't compete with data availability and cost in the indicator selection process.

Quality criteria for screening

Producing a limited, comprehensive list of key indicators to measure sustainability has been an elusive goal for most metric systems. A list of ideal indicator characteristics, developed with an understanding of the trade-offs outlined above, is used to screen the longer lists to create a key or core set. The selection criteria commonly include requirements that the indicators be specific, measurable, achievable, realistic, timely, science based and objective. The screening criteria will reflect the values and priorities of individuals or the initiative, thus they should be transparent and documented. They do not determine the feasibility on one indicator versus another, but provide the framework for asking questions on the key issues for the initiative and facilitate distinguishing the various indicator properties (Kurtz, 2001).

There will be trade-offs between feasibility, cost-effectiveness, practicality and credibility. Recognizing and accepting that few if any of the indicators selected will meet all of the requirements is critical.

The list of selection criteria is an internal and external tool. Externally, it communicates to stakeholders not directly involved in the selection process the context of the indicators, including priorities and constraints. Internally, the criteria are useful for routinely monitoring the relevance of the indicators. However, in practice, beyond the creation of the initial list during the development phase, these criteria are seldom used. Particularly in international membership organizations, the tendency is to seek consensus and settle on generally agreed-upon areas or define minimum "feasible" indicators in terms of current data availability, collection practices or data management systems—without circling back to the actual selection criteria.

The processes for developing selection criteria are tied to stakeholder norms and values, with the same potential pitfalls occurring in indicator selection itself—namely, the tendency to jump directly to a long list and use "feasibility" as the main criterion. Current selection criteria do not go far enough to imagine what kind of information we really need to assess performance in a paradigm of sustainability. Nor do they acknowledge that the issues are complex and costs to collect and manage information are very high. The processes take a long time to include stakeholders' perspectives; this also increases the funding needs for coordination, workshops and meetings, as well as the risk of volunteer participants dropping out.

Positive outcomes from the indicator selection and development process

A stepwise, evolutionary approach focused on a few strategic indicators has several potential advantages: streamlined tools and methodologies as well as focused learning to gain knowledge to improve the measurement system.

Indicators start to make a real difference even before they are implemented through the development process and assessment of lessons learned. By participating in the process, stakeholders explore *what* matters to *whom*, building a shared sense of vision and brokering consensus.

The development of a conceptual framework and indicators should involve participatory processes and clear statements of the criteria used in calculating each indicator. These factors are important to promote transparency and dialogue. If the assumptions and methods used in developing an indicator are transparent and understood, even if organizations or individuals disagree with the method or the result, an open and flexible process can become the basis for dialogue and adjustment.

2.3.3 *Indicator application and management*

"Results-based management at the United Nations has been an administrative chore of little value to accountability and decision-making." (United Nations Office of Internal Oversight Services, 2008)

The summary sentence of this United Nations report sums up many of the challenges presented by poorly designed indicator selection and overall mismanagement of indicators. The previous section focused on the development of indicators, the internal learning process and adaptation of the indicators. Even the best-designed indicators only measure (indicate) change; they do not explain why or how change occurs, provide scientific evidence or make decisions. Indicators must be applied and managed with the necessary information infrastructure for ongoing measurement and assessment of change in order to understand the cause and effect of specific management practices and activities at different temporal and spatial levels. As with indicator development, the use of the indicators—the application and management of information—must be adaptive.

To date, many indicator systems have been closed, proprietary and not transparent in their results. Some of this is due to data confidentiality issues. Particularly at the farm or business level, information about performance and inputs or outputs can be used in commercial negotiations by the buyer. This raises the question of who owns data about farm performance and who should store that data. The inability to aggregate data within and across systems has hindered the usability of the information across scales and levels. This is mainly due to the methodological and technical challenges addressed above, including lack of comparability (the system can't aggregate the data), manual data management systems that create silos of information in paper format, even within single organizations, and the lack of a standardized language and definitions for aggregating across initiatives or at different levels.

A stepwise approach to indicators focuses on using an overall conceptual framework with common definitions, but measuring *fewer* things *better*. Adaptive learning in the use of indicators allows the addition of more indicators and more sophisticated indicators over time to improve the understanding of causes and effects.

2.4 Adaptive management and learning

Regular review of indicators is required to identify whether the information obtained is useful and relevant for problem-solving and to determine if the issues have changed. Regular review of the information generated by the indicators is equally important to build knowledge about models, theories of change and links between farm-level practices, and to improve decision-making. Generally this is a key function of the monitoring and evaluation system. Ongoing review, learning and adaptation should be built and budgeted into the measurement system. The indicators should also be embedded into a planning and management process. These are important concepts of adaptive management. Environmental sustainability is a continuous and dynamic process that cannot be monitored using rigid indicator sets.

The central ideas of adaptive management have parallels in business (continuous improvement, total quality management and organizational learning), systems theory (feedback control), social sciences (institutional learning and design) and political science. In a complex world where change is occurring more and more rapidly and the future is largely unknown, a focus on learning by identifying and monitoring the appropriate indicators, providing feedback on the system, is essentially "learning to manage by managing to learn" (Bormann, 1993).

Adaptive management is in essence an adaptive approach to the planning cycle. Environmental and development goals are often at odds with one another, and both involve complex systems. Responses to interventions can be unpredictable, especially over long time frames. Using adaptive management techniques for managing this uncertainty is critical. Within an adaptive management framework, indicators are primarily used for instrumental (using results to change behaviour) and conceptual (learning and improving) purposes. As discussed previously, the different uses and purposes of information can create tensions. Results should not be seen as performance evaluations but as an opportunity to learn, including from unexpected results. This is very difficult when the results are part of the process whereby buyers determine whom to buy from based on environmental performance (legitimizing use). Commercial pressure is seen as a powerful tool to provide the incentives for farmers to improve performance. Assessment may risk focusing on pursuing specific results rather than on continual improvement by adapting what is learned.

3.0 Key emerging trends

The following builds upon the Bellagio SusTainability Assessment and Measurement Principles (BellagioSTAMP), released at the 3rd OECD World Forum on Measuring the Progress of Societies in October 2009 (IISD and OECD, 2009). BellagioSTAMP is based on the recognition that, due to a number of key emerging worldwide trends in the domains of science, policy, technology and civil society, new opportunities and challenges are arising for the development and use of sustainability indicator and assessment systems, including agri-environmental indicators. The opportunities come from better scientific understanding of the interactions and dynamics of coupled socioecological systems (such as agroecosystems), improved technical capacity (such as remote sensing and monitoring or computing power) and increased attention to measurement by the public, policy-makers and industry. The challenges are associated with critical capacity gaps, the growing diversity of metrics initiatives and the increasingly critical state of sustainability trends in various parts of the world.

3.1 Trends in science

The emergence of "new" concepts of science and the increasing demand for science-based evidence in recent decades has implications for metrics and metric initiatives, including:

- Increasing demand for policy-relevant science (see Section 3.2).
- Increasing demand for new scientific approaches that embrace uncertainty and cross disciplinary approaches, reflecting real-world circumstances and challenges.
- Increasing emphasis on place-based, integrative sustainability science.
- Increasing need to focus agricultural knowledge, science and technology on the multifunctionality of agriculture and understanding economic and non-economic benefits and costs (IAASTD, 2009a).

The urgency of environmental issues connected to global warming highlights the fact that it is no longer feasible to wait around until science "proves" or "disproves" something, particularly when direct measurements are not possible and the consequences may take years to manifest. The concept of post-normal science, developed by Silvio Funtowicz and Jerome Ravetz, is gaining some proponents. Basically, post-normal science attempts to characterize a methodology of inquiry that is appropriate for cases where "facts are uncertain, values in dispute, stakes high and decisions urgent." It is primarily seen in the context of the debate over global warming and other similar, long-term issues where we possess less information than we would like.

According to its advocates, post-normal science is simply an extension of situations routinely faced by experts such as surgeons or senior engineers on unusual projects, where the decisions being made are of great importance but where not all the factors are known. Although their work is based on science, they must always cope with uncertainties, and their mistakes can be costly or lethal. Given the great importance of climate systems and the fact that little is known about them, conventional methods of inquiry, based on determining all relevant information before proceeding, are too slow and uncertain to enable dealing with an issue that is too complex to be fully understood and too important to wait on.

Sustainability science is also an evolving field. Its name reflects a desire to provide the generalities and broad-based approach of "sustainability" with a stronger analytic and scientific underpinning. To do so, it

> can be usefully thought of as "neither "basic" nor "applied" research but as a field defined by the problems it addresses rather than by the disciplines it employs; it serves the need for advancing both knowledge and action by creating a dynamic bridge between the two. (Clark, 2007, p. 1737)

Sustainability science provides a critical framework for sustainability (Komiyama & Takeuchi, 2006), while *sustainability measurement* provides the evidence-based, quantitative data needed to guide sustainability governance. These new scientific approaches need to be underpinned by strong monitoring and adaptive management systems to address uncertainty, change over time and space, and manage the unexpected.

3.2 Trends in policy

"Most of the world's poor people earn their living from agriculture, so if we knew the economics of agriculture, we would know much of the economics of being poor."—Theodore W. Schultz, Nobel Prize Laureate in Economics

Historically, development has focused on economic indicators, and progress has been measured by GDP. The recent global crises—financial, food, housing, energy and so on—are highlighting the vulnerability of large populations and the need to shift toward policies that incorporate human well-being and environmental issues. Governments provide the enabling policy and regulatory framework for this shift, particularly concerning food safety, agricultural production and trade, while seeking to meet food-security objectives. Numerous trends in the policy arena pose new opportunities and challenges. These trends include

- a rising need for good, transparent information to support democracies;
- movement toward more evidence-based policies;
- increased use of national indicators for accountability;
- audit processes that serve as drivers of metric development (for example, Canada's Commissioner for the Environment and Sustainable Development and the U.S. General Accounting Office);
- use of information technology for aggregation across political boundaries, allowing comparisons and transparency;

- emergence of the OECD's Global Project on Measuring the Progress of Societies as a high-level coordinating mechanism;
- increases in green or sustainable public procurement policies, including mandated percentages;
- increased demands by consumers and consumer groups for protection from "greenwashing" through government regulations such as the Green Claims Code in the United Kingdom or the Roundtable for Green Products in the United States; and
- increased economic opportunities through transformation to low-carbon, sustainable economies (U.K. Department for Environment, Food and Rural Affairs, 2007).

3.3 Trends in markets

Numerous trends in markets, both at the corporate and consumer end, pose new opportunities and challenges requiring better, verifiable information. These trends include

- growing global consumer concern over the environmental, economic and social sustainability, public health implications and safety of agricultural practices and products;
- lengthening food chains with dominant megaplayers;
- metric initiatives filling the gaps in mandatory regulation of social and environmental practices in other countries;
- emerging voluntary private certification standards promising market access and better prices as an award for "sustainable" environmental and social stewardship practices;
- mainstreaming of certification systems creating a shift from small, niche, premium markets to cost savings, improved quality and market access;
- increased public-private partnerships creating higher demands for transparency and accountability (using indicators) to address concerns of "greenwashing";
- information technology accelerating emerging or unexpected market trends, creating higher levels of uncertainty and risk; and
- a rapid rise in corporate responsibility reporting, with over 1,500 companies using the Global Reporting Initiative format.

Farmers need to have the capacity to make new farming and technology choices to meet demands for a safe and healthy diet while navigating new regulations and standards, changing global consumption patterns, improved market access and potential value-added opportunities.

In addition, there businesses are increasingly aware of their dependence on ecosystems and their use of and impact on natural resources. The Millennium Ecosystem Assessment highlighted this, stating: "Businesses cannot function if ecosystems and the services they deliver—like water, biodiversity, fibre, food and climate—are degraded or out of balance" (Millennium Ecosystem Assessment, 2005).

> Customers want and will expect retailers to have greater transparency in the future. This is going to accelerate. We might as well get used to it.
>
> Despite all the work that's been done, we see only bits of information, but not the full picture across the supply chain. We don't know the patterns, hidden costs and impacts of the products we make and sell. Nor do we have a single source of data or a common standard for evaluating the sustainability of products.
>
> If we want to help the customer of the future live better, we need that data. We need that big-picture view. So today, we're announcing that we will lead the creation of a Sustainability Index. The index will bring about a more transparent supply chain, drive product innovation and, ultimately, provide consumers the information they need to assess the sustainability of products.
>
> —Mike Duke, President and CEO of Walmart
>
> Sustainability Milestone Meeting, July 16, 2009
> (quoted in Block, 2009)

3.4 Trends in technology

Advances in information and communication technologies offer statistical data providers, research centres and metric initiatives new tools and methodologies for defining, collecting, managing, analyzing and reporting on indicators. The Internet has changed the way people all along the supply chain look for and use data. The Millennium Project's *2009 State of the Future* (Glenn et al., 2009) reported that nearly 25% of humanity is connected to the Internet. More people are using the Internet in China than compose the total population of the United States. Mobile phones equipped with sensors and new software such as GeoChat allow geospatial ground truthing and data to be gathered, uploaded, downloaded or confirmed. Technology is making it easier to manage large amounts of data from different sources, formats and regions. The availability of cheaper monitoring tools combined with pervasive wireless technology and growing access to the Internet enable a type of civic science, where data collected through traditional methods and institutions of science can be organically combined ("mashed up") with both quantitative and qualitative information gathered by citizens for use in public policy and even individual decision-making (Backstrand, 2003). Although some concerns have arisen over data-quality due to lack of training and technical capabilities of "amateur" data collectors, projects benefit from the sense of ownership developed by these "citizen scientists."

Some of the opportunities presented by information technology advances include

- better use of available technology through new approaches to demand-led information needs (for example, extension and market information);
- increasing access to and declining costs of information and communications technologies to bridge information gaps;
- increasing use of information and communications technologies for transformation of information to knowledge by analyzing and reporting using complex models, with simple interfaces to access data;
- increasing use of networks and platforms for sharing tools, applications, codes and ideas such as ICT4Progress.org, wikis and GoogleEarth;
- ability to display statistical data in a meaningful way and to reach out to new audiences with a clear and easy-to-understand presentation of information, "transforming statistical information into societal knowledge" (Pintér, Hall, Scrivens, Hardi & Giovannini, 2009, p. 9);
- acceleration of improvements in instrumentation, communication among scientists and synergies across scientific disciplines; and
- increasing data-collection coverage due to the involvement of the general public (for example, through annual bird counts and Waterkeepers)—but there are trade-offs involving data quality.

Some of the new tools and methodologies include

- remote sensing and geospatial monitoring, as well as increasing possibilities to use real-time observations;
- complex, global modelling projects such as EnviroGRIDS;
- advanced technologies such as GeoNetwork and RANET, which allow visualizations such as maps and charts to present data in user-friendly forms;
- statistical Data and Metadata Exchange standards, established 2001 to maximize exchange of data and interpretability (GeoNetwork standardizes metadata and data collection);
- remote-access, Internet and wireless technologies that allow extensive data collection and management through low-cost or low-tech tools such as cellphone text messages or DigitalFormz.com;
- more efficient information infrastructure and improved tools for data organization and analysis, such as Miradi and Muddy Boots software;
- easier access to georeferenced data sets such as those available through the FAO, the FAO's ARTEMIS (Advanced Real Time Environmental Monitoring

Information System) and the Global Agro-Ecological Zones system offered by the International Institute for Applied Systems Analysis;

- systems of systems (structured systems composed of various other systems), such as the Global Earth Observation System of Systems; and
- online index tools based on complex ecosystem modelling, such as the carbon calculator COMET-VR.

The main challenge with technology advances is that metric initiatives, particularly non-profits, have limited resources and capacity. Furthermore, their uncoordinated efforts can lead to incompatibility of software and hardware platforms.

3.5 Who is doing what now?

Understanding who is working at what levels on what issues would be an ideal starting place for building consensus. However, an online compendium created by IISD in 2002 categorized close to 850 initiatives (www.iisd.org/measure/compendium). Since the compendium was developed seven years ago, the authors and others involved in metric initiatives estimate that the number of initiatives has doubled. This multitude of systems has created demand for new metric initiatives targeted at rating, benchmarking or reporting on the metric initiatives themselves. This section draws on Wunderlich and Russillo (2007).

Due to the sheer volume of initiatives and numerous classifications, the following section does not attempt to map or inventory the initiatives, but looks at the mechanisms, synergies and differences using prominent examples. This builds upon previous work done by INFASA, including a review of 25 key metric initiatives and a workshop on synergies and cooperation (Häni, Pintér & Herren, 2007). The initiatives reviewed are working at different levels, many with different target audiences and objectives. These include codes of conduct, projects, reporting initiatives, and verification and standards systems. All are driven by the basic need for information on the impacts of sustainable agriculture, though for different purposes and users, as outlined in Chapter 2.

Even with a small sample, the metric initiatives reviewed have a wide range of drivers. In some organizations development is the main priority, while in others the main driver is protecting biodiversity or providing farmers with tools for improved management. The use and purpose of the information is also highly varied and includes combinations of conceptual (learning), instrumental (decision-making) and legitimizing (demonstrating or proving impacts) uses.

The context in which these metrics are developed is important to understand, as it affects the characteristics and use of the metrics. That context includes the purpose of the indicators, framework, geographical focus, level (for example national), target users and selection criteria. For some groups, such as the Keystone Field to Market Alliance, critical selection criteria may be ease of use (certification systems) or data availability. Others, such as the OECD, may require comparisons across countries (OECD, 2001a). Most of the systems note that the information will be of value to all types of stakeholders; however, understanding the purpose and primary user of the information is important to defining the indicator and its calculation.

The primary users and uses of the information differ and include all levels, from policy-makers, for tracking progress against sustainability targets such as Canada's National Agri-Environmental Health Analysis and Reporting Program, to the producers themselves, for improving practices (such as through the Stewardship Index for Healthy Crops). Buyers are trying to figure out how to relate all these factors and decide which metrics might be most appropriate for them and their supply chains. Some buyers have developed their own metrics, which are not just used to identify areas of improvement for farming practices but also have become a requirement for market participation. This blurring of objectives for the use of information creates new tensions between information for learning and information as a yardstick.

New initiatives emerge from perceived weaknesses or gaps in existing systems. The increase is particularly noticeable in the agricultural sector with its potentially large environmental or natural resource footprint, processes that pose significant health and safety risks, labour-intensive production in developing countries and unequal trading relations for some commodities, such as cotton or coffee.

Because these initiatives often emerge on a project basis, the data on environmental practices and links to results are piecemeal, with buyers, consumers and policy-makers unable to make comparisons between the environmental performance of different production systems. The data may be useful in a narrow context, but cannot be aggregated to provide a landscape-level or ecosystem perspective. Instead of adding to the evidence, the result is too often competing metrics.

The following examples are not exhaustive or meant to qualify specific systems, but are intended to illustrate how different context leads to different content, collection and reporting systems—the information infrastructure. All of which creates barriers to using data across initiatives for comparison, benchmarking and aggregation. The search for one size that fits all continues; although based on different purposes and users, this approach has not only been elusive but may even be detrimental, by providing poorly matched data for decision-making and relying on producers who are reluctant to participate in the system.

Many discussions of metric initiatives begin with a typology of the initiative. However, this has become less distinct as initiatives embrace wider multistakeholder approaches and adapt to new issue areas over time. For simplicity this paper looks at two broad categories: public and private.

Public metric initiatives

In the early 1990s the United Nations' call for countries to measure their progress toward sustainability under various international commitments resulted in the development of various sustainable development indicator sets covering socioeconomic as well as environmental issues. The search for the ideal set of sustainable development indicators has continued for almost two decades, and the debate over identifying sustainable development indicators includes hundreds of projects (McCool and Stankey, 2004), which has led to a depiction of "an indicator industry" (King, Gunton, Fairbairn, Couts & Webb, 2000; Rydin, Holman & Wolff, 2003). Agriculture is linked to many global environmental issues and is a key element of global trade, so policy-makers need quality information to monitor progress, identify issues and track trends. Many international agencies have developed or are working on environmental indicators for agriculture as a priority issue.

Some common characteristics of public metric initiatives include higher scales (regional, global or national) and basis in some variation of pressure-state-response frameworks using socioeconomic and environmental themes. Frameworks, criteria and indicators are often linked to internationally recognized and accepted principles or agreements. These initiatives generally assert instrumental use (decision-making) for the indicators. Due to the scale and longer time frame of the information, the more common use is conceptual—for tracking progress and identifying trends and changes. Most initiatives have members collect the data using some common framework and platform. Many of them include web-based, open interfaces to the data. They create periodic reports, which are distributed to members, stakeholders and the general public.

One of the main challenges for the global and regional initiatives is adapting these procedures to local scales to make them useful at other levels for other stakeholders. Several private, farm-level initiatives reference or align with international sustainable development indicators such as the MDGs and the Convention on Biological Diversity, without actually linking to them.

Private metric initiatives

The terms *private standards* and *voluntary standards* are frequently used interchangeably. Private standards are highly varied with respect to who develops them, who adopts them, the parameters of the agri-food systems they address and so on. Reflecting this diversity, the functions they perform and the potential impacts they have are not clear (Henson and Humphrey, 2008). Two main functions of private standards include risk management, such as food safety, protection of brand capital and consumer trust, and product differentiation, to communicate with consumers. The main adopters of private standards are dominant supply chain actors, predominantly large food retailers and food service companies. Recent trends in private standards include increasing policy-level support and partnerships.

The multiple use and users of indicators have been discussed in several previous sections. The metric initiatives reviewed also differ in whether they measure actual input/output, the models they use, their inherent risks and whether they use point systems (scales). The majority of the indicator sets are based on farm management practices (associated with process-based standards), using hundreds of prescriptive criteria (in one case close to 400), versus recording the consequences of the farming system (outcomes or impacts), as in performance standards. Process standards tend to be prescriptive instead of results-based, making the linkages among management systems and outcomes difficult to assess. To address this issue and the information gap on environmental impacts, new initiatives keep emerging.

Some initiatives include third-party certification and the use of product labels for business-to-business and business-to-consumer communication. However, the sheer volume of indicators and information gathered, and their incompatibility, makes it difficult to pass on information to stakeholders and enable consumers to make individual purchasing decisions in a transparent, cost- effective manner. The majority of the initiatives reviewed collect the data themselves, often through appointed auditors or technical assistance providers. The overwhelming majority have closed, proprietary data management platforms, with some still functioning informally with paper-based systems.

Conclusions

The organizations and initiatives reviewed were at different stages in their development, with some having been around more than a decade and with extensive data sets, and others in a pilot stage with no testing to date. Some function as a project, others as a program or stand-alone legal entity. Both the historic and current organizational structure influence the indicators and information infrastructure with for-profit, non-profit, academic, non-governmental, industry and government agencies all working at different levels with varying organizational and coordinating structures.

These differences can cause some confusion between organizations or initiatives and external stakeholders, as the criteria and indicator systems, definitions and understanding of sustainable agriculture are applied differently based on the different drivers. Unfortunately, the end result is overlap of tools developed, different indicator definitions and an inability to leverage the different data sets and tools or compare impacts. For example, the data might become less useful to policy, as multiple sources of information now all vie for attention but cannot be pieced together in a meaningful way. In the same way, information could also become less useful to local producers, who may not know what tools to choose. However, overlap is not necessarily bad per se, as it can create an adaptive system that addresses different needs of end users. What is required, as will be discussed later, is the ability to harmonize terminology, identify the focus and interrelatedness of the data being collected, leverage the various tools and compare impacts. To achieve this, effective coordination at the macro level is necessary.

Critical success factors identified in the review include

- a multistakeholder approach with mixed geographical representation;
- a formal centralized coordination function to move the initiative forward;
- a clear focus;
- a strategic or business plan that includes growth and transition;
- transparency, consensus and stakeholder input into the decision-making process, requiring more time and resources, but improving buy-in and ownership of the criteria and indicator system in the long run;
- a process for developing indicators, since consensus built, values shared and foundations laid determine the long-term success of an initiative;
- a conceptual model as a common foundation, even if not explicitly communicated;
- coordination with other metric initiatives, though this adds an additional level of coordination that is not budgeted into any of the initiatives; and
- consideration during initial indicator development of how data will be used.

The new metrics: Benchmarking for transparency

The continuing development of uncoordinated indicator systems has created a new industry of benchmarking, scorecarding and reporting on metric initiatives and eco-labels. The multiplication of initiatives makes it increasingly difficult for stakeholders inside and outside of the initiatives, including policy-makers, civil society organizations and the finance, food manufacturing and retail sectors, to assess their utility and performance. The high growth in the number of "green" or sustainable products has attracted new businesses and led to concerns of "greenwashing" and questioning of the credibility of the metric initiatives. To address this, new metric initiatives are emerging to monitor and report on the metric initiatives. Ironically these benchmarking initiatives are following much the same uncoordinated pathway of criteria and indicator development, resulting in different frameworks, different indicators and competing scorecards. Selection criteria, as well as information on who developed them and how, are not always transparent. A number of platforms are also working to create a common vision, but without a specific indicator set such as the INFASA initiative and the Sustainable Food Laboratory. Table 2 is not exhaustive but illustrates the different levels of measurement and assessment of these meta-metric initiatives.

Table 2 Meta-metric initiatives

Initiative	Lead organization(s)	Target users	Stage	Tools	Criteria and indicators
Platforms					
ISEAL Impact Code	ISEAL	Standards initiatives	2010 launch	Impact code	15 or fewer; impact
Sustainable Agriculture Management System	Sustainable Agriculture Initiative	Industry members	In discussion	Bench-marking study	Under discussion; Unilever model
Sustainability Consortium	Walmart; University of Arizona; University of Arkansas	Industry	In development	Global database	Full life-cycle assessment
Green Claims Code	U.K. Department for Environment, Food and Rural Affairs	Industry; consumers	1998	Paper and electronic guides	Principles and ISO 14025
Green Products Roundtable	Keystone	Policy-makers	In development	n/a	n/a
Tools					
Ecolabelling.org	Big Room	Industry; consumers	285 labels	Web-based	About 15; general;* implementation
T4SD	United Nations International Trade Centre	Different interfaces for producers and procurers	2009 pilot; 2010 launch	Web-based portal with different user interfaces	About 600; general; governance; implementation; content; impact
State of Sustainability Reporting (SSI)	IISD; International Institute for Environment and Development; AidEnvironment; United Nations Conference on Trade and Development	Supply chain actors	2010	T4SD	About 150; general; governance; implementation; content; impact
OneWorldTrust	OneWorldTrust, European Centre for Not-for-Profit Law	Corporate social responsibility organization, donors	Launched 2009: 309 initiatives in several languages	Online, searchable database	About 30 categories; general; implementation; governance; monitoring and evaluation

*"General" refers to organizational information, scope, reach, coverage and industry.

4.0 What to measure at different levels

Understanding the links between farm-level practices and outcomes and impacts at different scales and time frames, including accumulated effects, is key to agriculture that is sustainable. There are enormous opportunities to link the data from the hundreds of metric initiatives in existence or in development, operating at different scales. For this paper we examined over 50 different metric initiatives (see Section 3.5) to better understand not the indicators themselves, but how information is collected, analyzed, shared and reported—the information infrastructure.

There have been attempts to coordinate across scales, such as country-level coordination with international frameworks (Agriculture and Agri-Food Canada national metrics and the OECD), as well as working to connect to the farm level through partnerships and tools. Some farm-level initiatives are working to connect at other levels; for example, the French Indicateurs de Durabilité des Exploitations Agricoles (IDEA) or Farm Sustainability Indicators (www.idea.portea.fr) tool for farm-level sustainability assessments is currently working to connect indicators relevant to farm-level practices to data from the higher-level Farm Accounting Data Network and the national census. In order to move from the conceptual to concrete examples, IISD conducted two rapid assessments in a developing country (cocoa in Ghana) and a developed country (lentils in Canada).

The case studies were intended to illustrate the report by assessing the environmental information infrastructure related to crops grown in a distinct geographical and socioeconomic setting. Each study was to use the lens of indicators of environmental outcomes related to that crop to examine how, by whom and to what end environmental sustainability is measured during the crop's production and at each geographical level. Rapid assessments were designed to explore the usefulness of the indicators and identify gaps in the environmental information structure. The process itself was a good learning experience, though the results were very limited and not as useful as planned. Interesting internal lessons were learned about the communication challenge of understanding the information needed to understand and assess the information infrastructure: what is collected, analyzed and reported, and how information is used. Lessons learned included the following:

- Information exchange at various levels can occur both formally and informally. The channels of information exchange need to be appropriate to both the information generators and users. However, the lack of an information-and-knowledge system hinders the ability to aggregate and use information outside an informal circle of influence.
- Use of experts to assist in interpreting and applying data is key for farm-level information uptake and application of knowledge.
- Indicators most directly relevant for farmers, such as yields or use intensities for chemicals and fertilizers, can be aggregated, but at best these are indirect measurements of environmental impacts.
- Proxy indicators used for donor and external accountability, such as land area under best management practices or certified, are not particularly useful for farmers or for regional or national governments for understanding outcomes or improving practices.

- Indicators most directly related to environmental outcomes are collected at the regional or national level. While this data is important in measuring the end results, it does not necessarily link the farm practices and the myriad efforts occurring at all levels to enable more sustainable agricultural practices to take shape.
- High-level aggregated indicators and proxy indicators can "hide" impacts and usually do not give a very accurate picture of the farming or environmental conditions at a specific location.
- The system must provide value for generators of information and transparency regarding how information will be used.

These efforts to link data face many of the same technical, methodological and institutional challenges and obstacles:

Technical

- Data quality and data gaps
- Lack of baselines and thresholds
- Dispersed nature of the data resources or sources over different scales and time frames—difficult to compare or aggregate, or require modification
- Need to simplify complexity through information technology tools and communication
- Different agroecological system sensitivities
- Linking farm-level activity to outcomes and impacts (attribution)
- Non-substitutability of information
- Inappropriate format of information, causing access constraints

Methodological

- Lack of explicit conceptual framework
- Knowledge gaps: understanding of ecosystem processes involved is incomplete and still evolving
- Need to understand functions, transformations and interrelationships
- Parameters not defined, trying to have one size fit all
- Farming as an economic activity with stakeholders interested in outputs and outcomes that are unpriced
- Complex, holistic systems that lose value as communications, management and evaluation tools—the right level of information is important

- Need to have clear objectives for data
- Understanding of the value of data at different levels
- Lack of knowledge about causalities and linkages
- Methodologies in different stages
- Lack of testing at broader scales and different systems

Institutional

- Management of resources at the local level, while impacts may occur at a larger scale
- Social, environmental and economic nature of efforts, with political involvement
- Cost of data collection, analysis and dissemination, requiring partnerships
- Building of information networks and capacity
- Consensus-building for common frameworks, goals, definitions, metrics and reporting
- Incentives for environmental and social performance
- Compromises between cost, accuracy, intended use and data collection methodologies
- Data ownership, confidentiality and privacy issues
- Players in the agriculture sector who usually don't have jurisdiction over all processes taking place in the agroecosystem
- Variation in accessibility of data across organizations

4.1 Continuum of assessment

Assessment is a continuum from output to outcomes to ultimately long-term impacts. Defining the goals within an overall conceptual framework provides a common understanding of what is being measured, in terms of terminology, level, unit and type of assessment at various levels. A conceptual framework is simply a representation of a shared vision and serves as a high-level direct reference to the basic concepts of sustainability, relationships between the different dimensions and goals. It's the big picture. Too often, people think they are speaking a common language and working under the same assumptions, but they have not explicitly agreed on these, leading to potential conflicts later. Explicitly defining the framework is particularly important for harmonization and standardization efforts. In addition, the conceptual framework helps to focus and clarify what to measure, becoming the backbone structure for the initiative in developing indicators and for communication purposes.

A number of frameworks of sustainability differ mainly in the way the different issues of sustainability are presented (conceptualized), the relationships between them, the way in which issues to be measured are grouped and the rationale for selection of indicators. Some general frameworks include pressure-state-response and its variations, issue- or theme-based frameworks, asset accounting, and results- or outcomes-based frameworks. The various frameworks have pros and cons, and some are more appropriate for different goals and levels of measurement. The framework chosen will affect the indicators used.

Conceptual frameworks for indicators help to focus and clarify what to measure, what to expect from measurement and what kinds of indicators to use. Diversity of core values, indicator processes and sustainable development theories have resulted in the development and application of different frameworks.

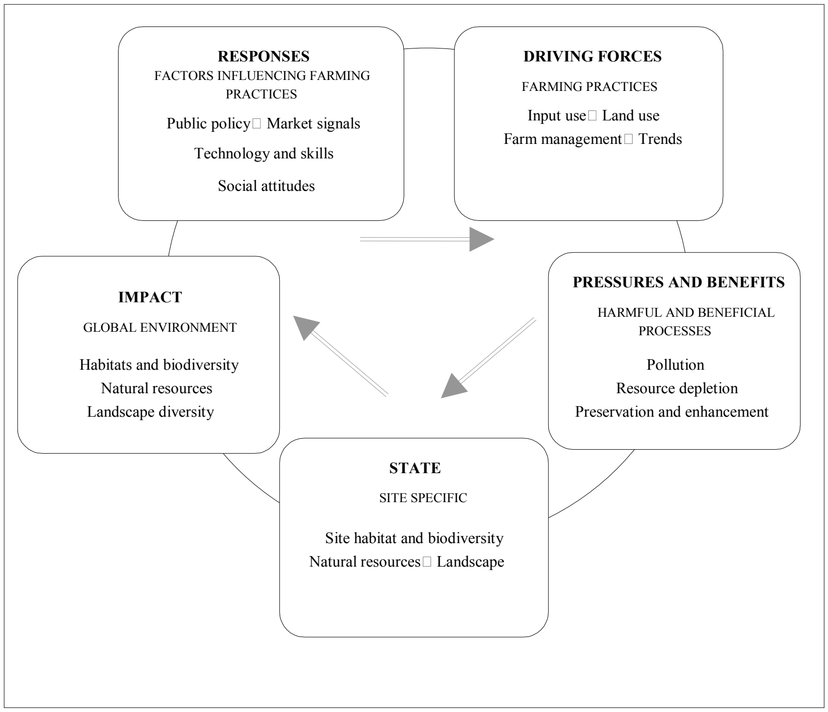
The main differences among them are the ways in which they conceptualize the key dimensions of sustainable development, the inter-linkages among these dimensions, the way they group the issues to be measured, and the concepts by which they justify the selection and aggregation of indicators.

(United Nations, 2007, p. 39)

Some examples of conceptual frameworks

DPSIR (driving forces, pressures, state, impacts and responses) frameworks and their variations give a structure for stakeholders to facilitate the identification of the key impacts. The framework can be used to identify the state (such as soil and water quality) and pressure (use of resources, wastes) variables in relation to impacts (changes in state variables), which reflect the physical part of the framework. Driving forces and responses are more linked to human aspects, the first factor being related to the decisions made on farming practices, and the second to the reactions of producers, consumers and other supply chain actors.

Figure 3 Agricultural DPSIR model from the European Environment Agency.

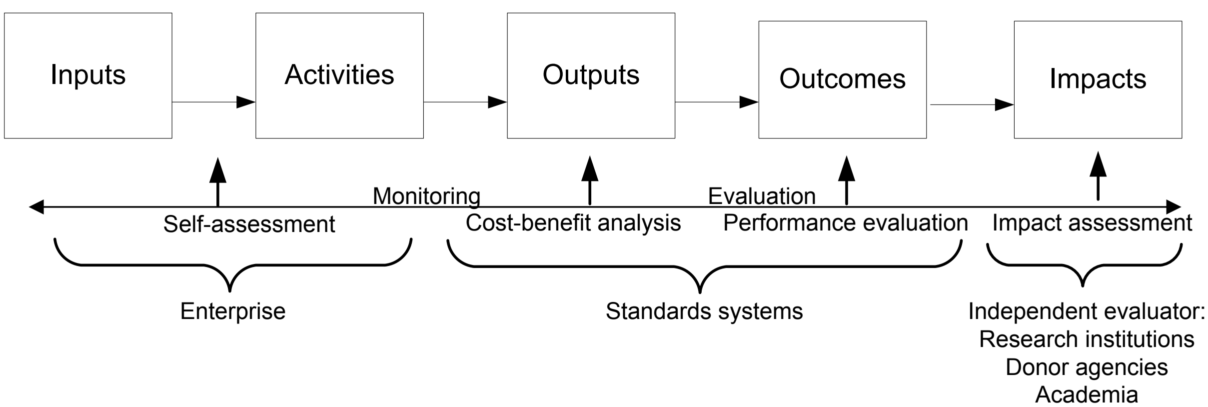


The value of a framework is not necessarily in the categorization of individual indicators, but rather in the provision of a common framework within which indicators can be presented, debated and analyzed.

The key question is how conceptual frameworks can help make cross-scale linkages between the landscape level and the farm or community level. Practically, when identifying indicators at a certain level (particularly lower, such as the farm), cross-scale linkages can be take into account by considering higher-level indicators as a menu, and where possible, measures should be synchronized. This does not mean using a "single-list indicator" approach, which, as discussed previously, has some major issues, but using a conceptual framework to adapt to the context. The framework identifies the goals and is used to identify the core indicators to assess progress toward these goals. It is also necessary to identify where along the continuum of outputs, outcomes and impacts the performance is being measured.

Models or other tools should be used to clearly define links to longer-term outcomes and impacts. For example, through diagrams that map out a series of causal statements that link factors in an "if…then..." fashion, the "logic model" provides a generic conceptual framework for defining how a particular activity or strategy contributes to reaching a long-term goal.

Figure 4 Roles across the logic model. From Russillo, 2008.



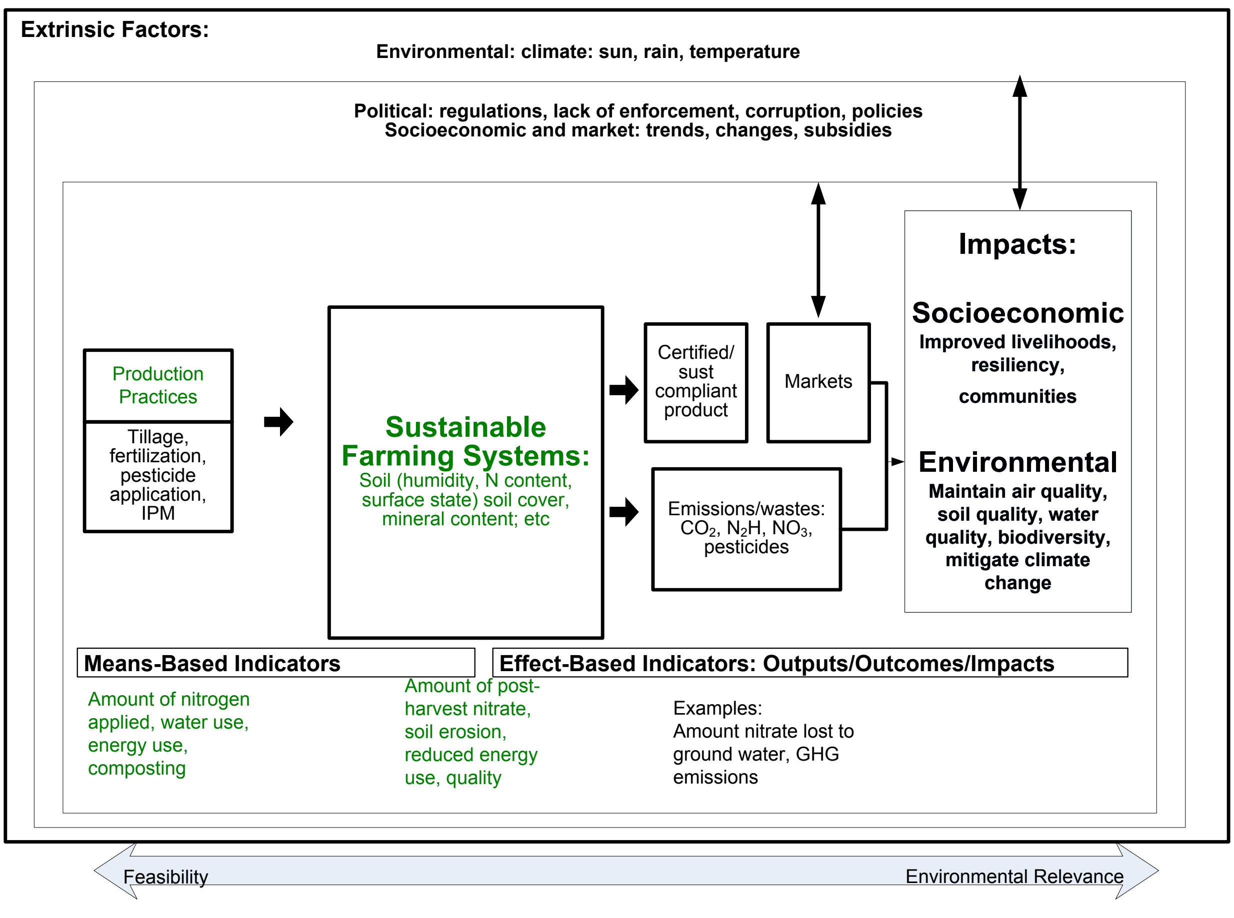
Determining what to measure, how to measure it and to whom to report the results to can be more of an art than a science.

4.2 Indicators at different levels of the system

As discussed previously throughout the paper, data and information are needed at different levels for different reasons and for different uses (instrumental, conceptual or legitimizing), which influences the nature of the indicator. The following illustrates the cause-and-effect chain from production practices to impacts. Indicators at different levels of the system related to environmental impacts include those measuring farmer production practices (requiring means-based indicators such as water use or nitrogen use) and emissions and wastes released into the environment (requiring effect-based indicators).

The interests of the different stakeholders may reach beyond specific boxes, with some groups working at one level but interested in the macro-level category of impacts, even if they don't really know how to make the link. And vice versa, some macro-scale agencies, such as governments, may be interested in some enterprise-level practices, such as fertilizer use, and try to influence it through programs like environmental farm plans.

Figure 5 Indicators at different levels of agriculture-environment systems. Adapted from Van der Werf & Payraudeau (2004).

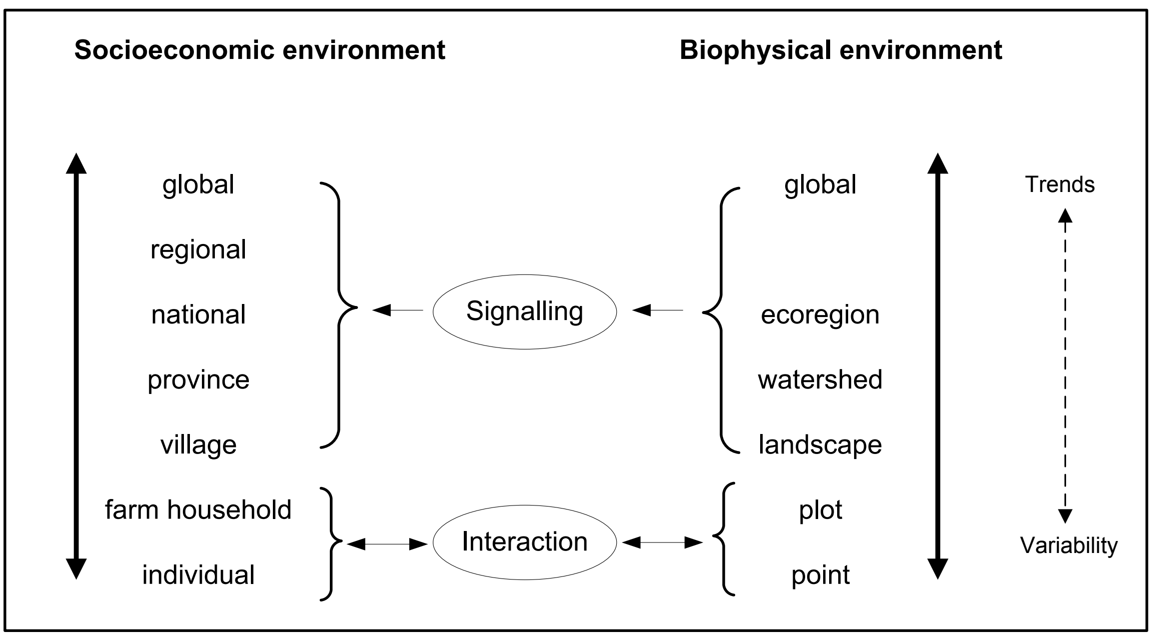


Both issues are related to attribution or linkages. External factors can also influence outcomes, such as rainfall or temperatures, so even if farm practices across several seasons are constant, their contribution to outcomes may shift due to other changes in the system that may have little or nothing to do with agriculture. The source of some major impacts may even be external, such as climate change or transboundary air pollution. An additional aspect of reality is that, depending on the impact pathway, long time lags can occur between specific practices and their results. Most decision-makers want information *now*.

Indicators should be defined at a meaningful spatial scale, but defining appropriate system boundaries is not straightforward. Ecological, political and geographical boundaries are often mismatched, and there is also a nesting or hierarchy of not only geographic units but also actors who are located within and care primarily for issues within their unit, where they have jurisdiction and where their primary interests lie. Farmers are interested in data related to practices or management of resources that they have control over, but the same is true for regulators or agribusiness. In some cases interests across scales and organizational levels overlap, and in such cases defining common metrics is more straightforward. But often they don't, and then setting up common metrics and assessment procedures is problematic.

Decisions and activities at the farm level interact with and affect the biophysical environment (Roetter, Keulen, Kuiper, Verhagen & Laar, 2007) (see Figure 6). The farm-level activities and management practices also affect higher scales such as watersheds (through impacts such as fertilizer runoff) or global scales (for example, tillage affecting carbon emissions that contribute to global warming), immediately or over time. Governmental and non-governmental (such as consumer) groups notice the higher-scale environmental effects—outcomes and impacts—and adopt practices or policies, such as regulation, incentives or disincentives, or eco-certification, that affect lower scales.

Figure 6 Interactions between humans and the environment at various scale levels. From Verhagen, Wösten and de Jager (2008, p. 71).



The ecosystem scale is widely regarded as the most appropriate scale at which to address many questions of ecosystem health, but the most relevant data are not necessarily collected at this scale. And the farm level is where practices regarding sustainable agriculture can contribute to or mitigate impacts on ecosystem goods and services. However, locally important and often specific details may not all be relevant when calculating ecosystem-scale averages.

Understanding the differences between these interactions is key to linking them across scales. There is no one-size-fits-all set of indicators. Measures at more than one scale can create a comprehensive understanding of the system. The following sections examine the different scale levels, looking at specific examples for trends, challenges and implications for indicators to have value across scales. Defining the roles and responsibilities of stakeholders at each level is also key. Where does each level fit with others? How can this be determined? How is information collected, and how is it made public? How is it used? Clarifying the objective of the data, what it will be used for and by whom, is important in defining the indicator and its scale and determining where it should be measured.

4.2.1 Farm-level indicators

Information users and generators. At the farm level, producers both generate and use information. Some of the information they generate is passed on to other stakeholders, including supply chain actors, metric initiatives, local government agencies and certification organizations. Producers use information generated by metrics in order to:

- Manage inputs, throughputs and outputs.
- Manage implementation of sustainable practices in the most cost-effective manner.
- Gain a better understanding of the financial and other risks associated with production and entry into sustainability-conscious markets.
- Understand the social, environmental and economic costs and benefits associated with adopting one program or another.
- Access new revenue streams (for example, through payments for ecosystem services).

Traders, manufacturers and retailers. Each of these stakeholders has a direct interest in maintaining the stability, quality and good management of their supply base. They also share an interest in responding to consumer demand, including the demand for evidence related to environmental impact. They need information generated about farm-level activities to:

- Manage risks. Operational risks include scarcity, costs, natural hazards exacerbated by environmental degradation, and higher insurance. Other risks include those associated with regulations, access to capital (if the investment community adopts more rigorous policies) and reputation (Grigg, n.d.).
- Identify and use opportunities, including market differentiation, increased efficiencies and supply chain resilience.

Metric initiatives, certification labels and benchmarking schemes. Operating at the farm level, these have evolved over the last 20 years in order to provide the tools and information for producers and buyers to accomplish these objectives. Common characteristics of initiatives operating at the farm level include:

- Private, voluntary initiatives developed by academia, industry and NGOs.
- Multistakeholder approaches that include producers.
- Proprietary data-collection systems and data sets.
- Minimum analysis of metadata, making the initiatives data rich but information poor.
- Increased use of web-based tools and spreadsheets for data management and reporting.
- Lack of testing for practicality on a large number of farms or regions in order to build models.
- Focus on compliance or progress with recommended or required farming practices, not on the results of those practices.

Process indicators are straightforward, relatively easily measured and controllable by the farmer. However, these indicators are not suitable for recognizing errors or guiding change, especially when the evaluation concerns farming practices that have been defined ahead of time (Van der Werf & Petit, 2002), such as with certification systems. Recognizing these limitations, initiatives are reacting in several ways:

- Developing additional "impact" indicators to monitor for communication and accountability (such as ISEAL, Linking Environment and Farming (LEAF), and Unilever's sustainability initiative).
- Focusing on results-based indicators rather than prescriptive process indicators (for example, the Stewardship Index).
- Benchmarking against other farmers (such as Response-Inducing Sustainability Evaluations [RISE] and IDEA), national averages (such as the Keystone Field to Market Alliance) or regional results (such as Canada's National Agri-Environmental Health Analysis and Reporting Program).

These strategies attempt to address the immediate pressure felt by the individual metric initiatives to "prove" impacts or show results; however, they still do not address many of the bottlenecks, identified in earlier sections, of linking practices to results or impacts at different scales.

Even the level of measurement on the farm itself must be considered and defined for each indicator depending on the information needs, feasibility and relevance. The metric initiatives look at different levels of "farm": whole farm, farm enterprise (cost/operation centre, market garden enterprise and different crops) and crop. The distinction is critical because some systems (such as commodity roundtables, ISEAL certification members and Keystone) focus on the crop level, while others (such as the European Initiative for Sustainable Development in Agriculture [EISA] and LEAF) focus on agroecological systems or integrated farming. Policy-makers are interested in the whole farm, while procurers are more often interested in the crop level. Measurement may be technically possible and disaggregated at the crop level or at the whole-farm level. Resources are generally managed at the whole-farm level, taking into account the relationship between the different crops and production systems. It is not necessarily possible to consider all the indicators on the same level. This becomes an issue when the information is aggregated or compared with other systems.

Differences

The main difference among the initiatives that are focused on farm-level indicators is whether their objective is evaluation of the environmental impact or of adherence to good agricultural practices. This affects what data is collected and how it is collected, reported and used.

The majority of the certification systems reviewed evaluate adherence to good agricultural practices using third-party auditors and pass/fail criteria. They collect actual farm-level data such as water use and aggregated scores, hiding areas of underperformance. Audit reports are confidential and can be used as guides to identify areas for improvement; however, to avoid conflict of interest, the standards system cannot offer interpretation or technical assistance.

Some initiatives, particularly those not rewarding with a seal or label, use tools such as self-assessment tools, some of which are web based. Adherence to good agricultural practices is still the main evaluation criterion; however, the initiatives are focused on providing tools for better farm management (conceptual use) rather than awarding a seal (legitimizing use). Initiatives such as (LEAF) and SureHarvest use evaluations for comparisons from year to year and for benchmarking against other farmers, rather than to pass or fail a farmer. Many of the metric initiatives are involved in interpretation of the data, though farmers differ in how they use quantitative information in their management. Some use technical data or graphical representations (such as the spider diagrams used by RISE and IDEA); others rely on practical skills and verbal exchange (such as in the lentils case study). Environmental farm planning in Canada, supported by Agri-Food Canada and several of the provincial governments, represents a centrally coordinated initiative where metrics are understood to serve the purposes of farm management as informed by environmental farm plans. All of these factors affect the perceived value of the information to the farmer and whether or not information is used as a motivating factor for positive change.

Data collection and use

The objective of the evaluation affects the data collection. Indicators based on farmer practices cost less for data collection but do not allow an actual evaluation of environmental impact. Certification programs and company codes generally take advantage of regularly scheduled audits and visits to collect farm-level data. Others, such as Dairy Management Inc. (DMI) and Unilever, use the collection or sales point for data. The resources and origin of the initiative generally determine if the data is collected in a standardized format and in an electronic or paper format, and whether it is aggregated, analyzed and reported on. Proxy indicators, such as acreage and number of certified operations, are still the primary data reported externally. Many, particularly the non-profits, still use non-standardized paper or spreadsheet formats and manage little of the information electronically, with some exceptions (for example UTZ CERTIFIED and LEAF). Generally to date this information has been used for external communications such as donor accountability and marketing. Industry initiatives generally have built information infrastructure in some form into their programs, including hand-held wireless data collection at the farm level connected to web-based platforms. This information has been of particular value for use in managing supplier relationships and reducing risks, and for external communications. Data collected by national or state and provincial agencies or extension services follow statistical protocols and are usually entered into dedicated databases.

Lessons learned about farm-level indicators

Key findings related to use of farm-level indicators include the following:

- The producer does not want to be held accountable for outcomes he or she cannot control, though outcomes primarily matter from the point of view of the regulator and consumer.
- If information is used by stakeholders, such as buyers, who have a direct influence on the farmers with results that affect decision-making, there must be a practical way to consider external influence (Payen, 2008).
- Those collecting the data need to compensate the farmer through information that has value or other incentives.
- Initiatives need to ensure farmers have the capacity to collect and interpret data and to make the best use of indicators in management and farm-level decision-making.
- Initiatives need simple databases and information collection protocols.
- Information by itself is insufficient, but technical assistance providers and interpretation are needed to enable knowledge transfer.
- Producers need to be included in the indicator development process, and the initiative needs to clarify who defines "best practices."
- Advisors and farmers need reference or threshold values, which often don't exist or require a significant technical or scientific effort to establish.
- The input/output relationship of farmer practices linked to higher-level outcomes must be understood, keeping in mind that this may require technical expertise in modelling and thus assistance by extension services.
- Improved practices are knowledge intensive and require research and extension systems that can generate and transfer knowledge and decision-making skills to farmers.
- Using a final numerical score allows compensation across issues and scales, hiding potential negative impacts; that is, favourable practices could offset harmful practices.

- Scoring systems require the construction of value scales and a means of interpretation.
- Informal systems of information sharing and knowledge transfer should also be fostered where applicable.

Roles and responsibilities

Farmers. Farm-level indicators should focus on controllable factors such as inputs and outputs. Priority should be given to quantitative indicators that show changes on the farm over time and space at the scale relevant for a given farm, and that are relatively easy to calculate, audit and understand. The farmer can collect this information using standardized, accessible tools. This data is beneficial to the farmer in improved management and sustainability through the interpretation of information and links to management practices through advisors or technical assistance. Use of management indicators should complement, not substitute for, outcome metrics. In cases where the information infrastructure allows, farmers may be able to input data directly into higher-level (such as ecosystem-level) databases, which they could also use to access either their own or others' data, as permitted, or data at a higher scale.

Metric initiatives. Should provide farmers with the tools and interpretation of results for more sustainable practices. Transparency in how farmers can improve and become more sustainable and economically viable under the metric initiative's system is critical. In addition, the initiative is responsible for transparency surrounding the costs, benefits and results of the different approaches in specific contexts. It needs to collect a lot of data from a lot of farmers and aggregate it in several different ways to set references or thresholds and build in statistical analysis. Initiatives need to develop links to higher-scale data sets to understand the sustainability of farming *systems*. In order to accomplish this, harmonized approaches, partnerships and sharing of data within and across initiatives is essential. Besides simply accessing what data is available, higher level organizations such as national statistical agencies may also need to proactively tailor their data-collection activities so that they better respond to information needs at the level of the farm.

Key environmental impact indicator approach at farm level

It is hard to have consensus, even at a generic level, on the choice of indicators because of local and regional differences or significance and different values placed on landscapes or biodiversity. Farmers will be reluctant to collect data on results and impact indicators that have no specific value for them because of the additional costs of data collection or because they are used primarily for control purposes.

However, this approach can be the key link to integrating farm-level practices to landscape-level impacts—not as a yardstick for individual producers, but integrated into the conceptual framework as goals for farm-level practices. A slightly different interpretation of the key environmental impact approach involves using metrics for ecosystem services.

Ecosystem services are the resources and processes supplied by the natural ecosystems that sustain human well-being. These are not ecosystem products or assets, but functions that have long been considered free and infinite. The Millennium Ecosystem Assessment (2005) and IAASTD (2009a) classify different categories of services: provisioning, such as the production of food and water; regulating, such as the control of climate and disease; supporting, such as nutrient cycles and crop pollination; and cultural, such as recreational benefits. Many of these functions are in the public domain and currently are "free" in economic terms. The recent decline in some of these services has created competition and new markets.

Typically agriculture has been managed to maximize the provisioning services to produce food and fibre. Now the prospect of markets for ecosystem services represents a new role and potential revenue stream for farmers, with the precondition that farmers, or whichever other agent is arranging the transactions, can *prove* that specific services are delivered for the payments received.

Agriculture depends on ecosystem services and also generates them. Agricultural ecosystem services can be grouped into three categories:

1. Services that *directly support* agricultural *production* (such as maintaining fertile soils, nutrient cycling and pollination).
2. Services that *contribute directly* to the *quality of life* of humans (such as cultural and aesthetic values of the landscape).
3. Services that *contribute* to *global life-supporting* functions (such as carbon sequestering, maintenance of biogeochemical cycles, supply of fresh water and provision of wildlife habitats).

Focusing on ecosystem services provided to and impacted by agricultural practices will highlight the relationship of the different scales. While measures are only in development, this approach also offers the prospect of *monetization* of some of the services as part of the solution to more sustainable production practices. One of the best examples of this is the evolving carbon market, in which a global public good has been made "private" and linked to market-based mechanisms. Experience shows that accounting for carbon is an essential precondition for the functioning of carbon markets, otherwise transactions cannot be verified. This is working reasonably well in the case of point-source greenhouse gas emissions, but accounting is more complex for carbon associated with non-point sources in land use, land use change and forestry.

Farm-level indicators link to other levels

Information regarding farm-level practices can be used to link to landscape or ecosystem levels and have value beyond the producer. Farm-management indicators are "raw" data that can be directly linked to activities. As such:

> They can provide an early indication of likely changes in the direction of environmental impacts sometimes well before they can be measured by other indicators, such as those pertaining to soil and water quality. They can also serve as a proxy for "state" indicators where the latter are difficult or costly to monitor. Measuring farming practices is often more practical and cheaper than measuring actual changes in the environment. Monitoring the trends in management practice indicators alongside appropriate "state" indicators, such as water quality, can also help policy-makers to evaluate directly the success of policies aimed at environmental improvement. (OECD, 2001b, p. 85)

4.2.2 Landscape- or ecosystem-level indicators

Governmental and non-governmental agencies are both interested in information at the landscape level, particularly as they relate to ecosystems goods and services. With increasing populations and consumption pattern shifts, demands on landscapes are increasing and even conflicting. Decision-makers need indicators that measure conditions, pressures, resiliency and societal responses.

Generally, landscape-level initiatives are partnerships among various stakeholders with different interests (and values). National and international government agencies often are involved in or lead these initiatives because of the role of public goods within many landscapes. Many agencies and countries are working to develop landscape-level frameworks and indicators, with different approaches in terms of objectives, scope and methodologies. Numerous state-of-the-environment or ecosystem reports have been prepared to appraise the condition and trends of ecosystems, including the global Millennium Ecosystem Assessment, the Global Environment Outlook or the many subglobal reports and assessments that have been produced based broadly on the methodology of integrated environmental assessment (such as the United Nations Environment Programme's Integrated Environmental Assessment Community Platform, www.unep.org/ieacp/iea). While these reports are important for taking the pulse and highlighting trends, many critical information gaps and issues remain, hindering the usefulness of this information at other farm-level scales.

Most voluntary standards in agriculture do not directly measure environmental issues and outcomes at the landscape level, though some do include socioeconomic issues such as community benefits. Some of the newer initiatives, including the commodity crop roundtables, do take some landscape-level environmental issues, such as watersheds, into consideration. Ecoagriculture Partners is one of the few non-government-led initiatives focusing on metrics at the landscape level. Their Landscape Measures Initiative, being developed with Cornell University, works to specifically link farm-level practices to landscape-level results. The conceptual framework, tools and indicators exist but, as the methodology has not been extensively tested, no data exists yet regarding the links of practices to results.

Models, surveys, census data, remote sensing, governmental and custom data sets are the main sources of information. In many cases data collection using advanced technology to build complex landscape typologies based on existing data sets is now possible, though limited by resource and capacity constraints in many developing countries. Even in developed countries, different departments and regions use different systems and indicators. Data collected at the landscape level is generally used by regional or national governments for periodic analysis of past trends to prepare state-of-ecosystems reports for policy-makers, researchers and NGOs in order to monitor changes and identify key issue areas.

Critical issues with landscape-level indicators

Technical. Issues include application of remote sensing methods for the area-wide monitoring of land-use change; standardized data-processing techniques, which are vital for the spatial and temporal comparability of results (data often comes from various agency monitoring programs for specific purposes); and the selection of a manageable set of indicators that reflects the structural properties of landscapes.

Institutional. Institutions muse choose appropriate spatial units that allow for an integration of landscape indicators (which tend to relate to cross-border phenomena) and socioeconomic indicators (which are usually available for administrative entities or areas) (Lausch & Herzog, 2002). They must also handle ownership and management of the data, which often comes from different sources and scales, as well as how information is made public and disseminated, including the cost of data and access limitations.

Roles and responsibilities

Farmers. While farmers have economic motivation to improve management practices or address on-site problems, they have weak incentives to mitigate off-site effects. Avoiding such externalities requires regulatory mechanisms, negotiated solutions or transferring payments between those causing the damage and those affected by it, possibly involving large numbers of people separated in space, time and interests (World Bank, 2007). Farmers should participate in the development of landscape-level indicators that are linked with on-farm management decisions and ecosystem service payments.

Metric initiatives. Monitoring programs of non-governmental, academic and government institutions should "nest" landscape-level indicators within a conceptual framework of sustainability, with links both up and down the scale. This increases the need for harmonized or standardized frameworks, definitions and data-management tools. For example, the Agriculture and Agri-Food Canada national report uses 12 ecoregions across Canada, drawing upon the soil landscape database, which includes not only information on soil properties but also information on agricultural production and management (Agriculture and Agri-Food Canada, 2005). Metric initiatives must provide analysis and clear links to agricultural systems and practices for use by advisors and technical assistance providers to enable farmers to adopt more sustainable practices. Farmers will need greater ecological literacy to better understand interactions in complex ecosystems.

Key environmental-impact indicator approach at the landscape level

The key environmental-impact approach is appropriate at the landscape scale when tied to specific farming systems and practices. The key elements of landscapes include structure, function and value, all of which can be tied to ecosystem services.

Landscape-level indicators have value for farmers when results are linked to practices for improvement (conceptual use) and payments for ecosystem services (instrumental use). Landscape-level information can also help farmers better identify broader regional or ecosystem trends affecting their operations in the present or potentially in the future, which they may need to adapt to. Adaptation may become a critical element of success and sustainability, particularly in cases where the external risks are significant and their only possible response is through adaptation.

Landscape-level indicators provide policy-makers, researchers and NGOs with important information regarding patterns of environmental risks and conditions in specific regions. For policy-makers, they provide information for land-use planning and conservation at the national level, helping to resolve trade-off conflicts between "natural" habitat protection and agricultural development (such as in wetlands), supplying information to help generate forecasts of future agricultural land-use patterns (OECD, 2001b). They can point out areas for further research and action and show trends over time.

4.2.3 National or regional level

The United Nations' Agenda 21 calls for countries, international organizations and NGOs to develop and use sustainability indicators as part of an integrated approach to accounting. Many countries collect data based on international agreements such as Kyoto or the MDGs, or coordinate national efforts with international agencies such as the OECD's agri-environmental indicators. Policy-makers need clear and objective information on how different sustainable practices impact producers and their communities as well as food security, and to monitor trends and changes in the condition of ecosystems. Policy-makers need to understand

- the relationship between policy changes and farmer behaviour; for example, choice of inputs;
- the relationship between input use, production and environmental quality; and
- the relationship between changes in environmental quality and changes in social well-being (Sinner, n.d.).

The increasing number of cases where reports and indicators are becoming mandatory (that is, legislated) in the context of accountability regimes in line with national and international agreements and conventions is also increasing. The European Union is using Aarhus Convention[5] principles in legislation for its Water Framework Directive, the Kyoto Protocol and legislation on greenhouse gases and national water and air quality.

Data collection and use

Data collection is generally done through partnerships of agencies at local, regional or state, and national levels, using existing monitoring programs, data sets and information infrastructure for modelling and analysis. In some instances, primary data is collected through surveys; however, the scale and time required to gather this data requires substantial resources. New opportunities are arising for farmers to become more direct suppliers of data through Web 2.0 technologies, with which they can record data about their operations in public statistical databases that they can also directly access.

5 Signed in 1998 and ratified by 41 countries, the Aarhus Convention grants the public rights regarding access to information, public participation and access to justice in governmental decision-making processes on matters concerning the local, national and transboundary environment.

Farmers have also served as a source of statistical information in the past through farm surveys, but new web technology could radically transform how this works. Besides collecting data themselves, public agencies often act as synthesizers and data analyzers, though once out in the public domain, all other players may also carry out their own, contextually more focused data analysis (that is, at the firm or farm level).

Due to the broad scale of data and longer time frames of changes, reporting (including analysis) is done periodically, typically every five to ten years. Indicators are predominantly lagging, indicating what already happened, as opposed to leading, signalling what should happen within the model. This poses the risk that resources or services will move "beyond the tipping point" of recovery before even being identified or understood.

Challenges with national-level indicators

Technical. Landscape-level indicators require reference or threshold values to measure progress; have long time lags for the environmental impacts they measure; require a massive volume of information; are based on data sources developed for other purposes, resulting in missing or incompatible data; have a top-down perspective that is not measurable at the local level, where actions and practices can be changed; have few links with on-farm management decisions; work at broad temporal and spatial scales that do not correspond to the local level or serve as guides to on-farm best management practices; aggregate data, which hides risks, vulnerabilities and sensitivities of specific localities; and present difficulties and costs associated with data quality and quality assurance due to the possibility that data was collected by many parties using different protocols.

Methodological. The long timelines to gather fragmented information over broad scales make it difficult to assess accumulating impacts or stresses on ecosystems and services. Furthermore, the indicators are lagging rather than leading.

Institutional. Responsibility for the collection, analysis, and dissemination of the data needed for key policy and decision-making is fragmented, resulting in a profusion of insufficiently coordinated federal, state, local and non-governmental efforts (H. John Heinz III Center for Science, Economics and the Environment [Heinz Center], 2008). "Private" data on public goods creates tensions over who pays for and owns data. For instance, in some cases industry may also collect regional-level data where factors directly affect production. This may be particularly true in developing countries where public agencies are weak. For example, the flower industry has some of the best data on groundwater in Kenya. Different agencies often manage the resource base with competing objectives—Costa Rica was one of the first countries to put energy, mines, environment and water under one ministry. Institutions need support, including funding and capacity building. Finally, policy-makers generally plan along short time frames.

Roles and responsibilities

Farmers. Farmers can use information for benchmarking (such as with Keystone Field to Market), and for understanding environmental issues and conditions in a broad sense (such as Agriculture and Agri-Food Canada).

Metric initiatives. These have similar roles and responsibilities as at the landscape level, providing the overarching framework and standardization of approaches, tools and definitions. Metric initiatives at this level must provide analysis and clear links between policies and agricultural systems and practices, including market-based mechanisms such as payments for ecosystem services.

Key environmental impact indicator approach at the national level

The use of key environmental impacts is useful at the national scale when tied to ecosystem services, nested ecoregions, landscapes, and specific farming systems and practices. It is also essential for tracking current and emerging environmental risks and assessing the adequacy and impact of policy responses. In some cases, for example in Canada, reporting on environmental performance at the sector, department or ministry level is a legislative requirement that is periodically reviewed by national audit agencies.

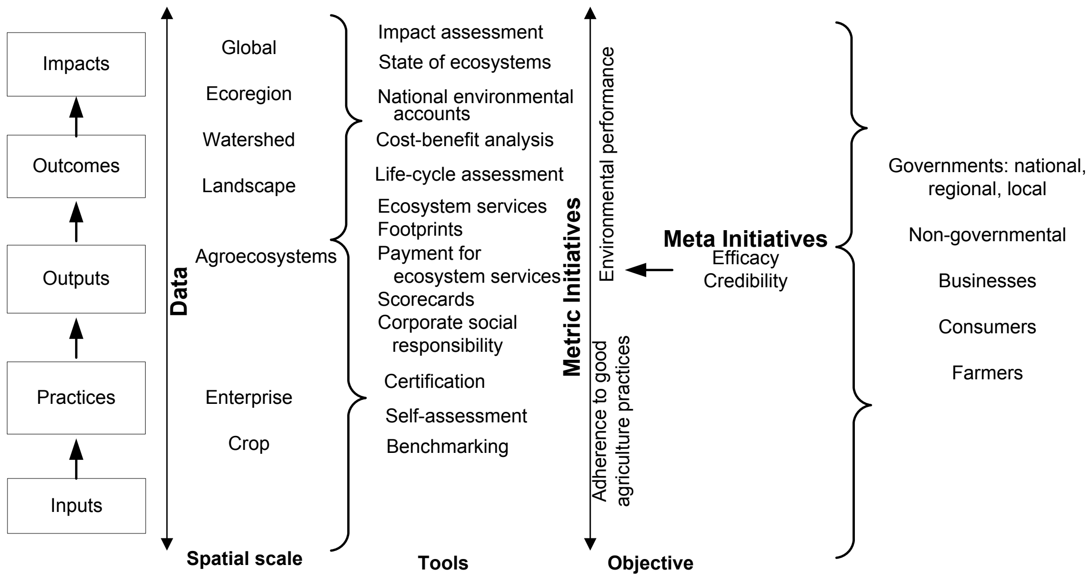
Linking economic and ecological models is a huge challenge—including issues discussed in previous sections regarding aggregation, choice of environmental indicator and interpretation. Strong leadership is needed. Are national-level indicator initiatives like global initiatives such as the OECD's agri-environmental indicators the answer? The Heinz *Road Map to the Future* report recommends that the U.S. federal government adopt a national environmental indicator initiative, guided by the federal government, states, the private sector, environmental organizations, universities and others. This effort would link national indicators with information used by local, state, corporate and other decision-makers and drive an agenda for improving data collection and reporting (Heinz Center, 2008).

4.2.4 Linking the levels

Some of the areas where linkages could be established include

- a common framework or common elements of a framework;
- use of comparable indicators and metadata;
- collaborative capacity building and data platforms that allow data submission by a distributed network of actors, including, potentially, farmers;
- clarity on impact pathways across scales and actors;
- attribution of impact to various drivers or pressures at the different levels;
- complementary capacities;
- platforms and openness for cross-scale dialogue;
- comparable legislative requirements;
- a clear incentive mechanism and cost-sharing across the many levels; and
- multi-scale oversight, audit pressure and mechanisms.

Figure 7 Metric initiative tools and objectives at different spatial scales.



The responsibilities at each level were outlined in the previous sections. While considering linkages, it is important to keep in mind the roles and boundaries of each level:

Farms and companies act

- within legal boundaries;
- voluntarily—with their own interests, to improve sales, adapt to given external conditions and so on; and
- as private economic players.

Goals are often set at a higher scale. They are influenced by

- regional and local governments (through legislation and financial incentives) and
- consumers and consumer organizations (Ruijter, 2009).

For the farm level, indicators and certification schemes related to adherence to good agricultural practices have value for farmers through lower inputs and improved yields and market access. These process-based indicators should be integrated with outcome and impact measurements that can be used by other stakeholders, such as buyers, NGOs and government agencies. Incentives for farmers to manage or monitor higher-level and off-site outcomes and impacts must be developed, such as through ecosystem service payments.

For the wider community, landscape- and ecosystem-level indicators based on the environmental *effects* of farmer practices, both on- and off-site, are preferable to indicators based on farmer practices themselves, as the link with the objective is direct and the choice of means is left to the farmer (Van der Werf & Petit, 2002). For example, the adoption rate of environmental farm plans in Canada is used as a farm-level indicator, but in order to interpret their impacts, there is also the need to look at spatial and temporal correlations between adoption rates and various practices under the plans and specific agroecosystem trends using models and georeferencing.

National-level indicators provide the overarching framework for local, regional and national organizations, including transboundary ecosystems and global issues. Some problems, such as climate change, require cooperation at the global level.

Finally, not only are different indicators needed at different levels, but the same indicator may need to have a different level of precision based on how the information is to be used. Establishing indicators to be used as proxies for measuring carbon dioxide emissions is one example of how different measurements are useful at different scales, as Table 3 illustrates.

Table 3 Measuring carbon at different scales. Adapted from Jonathan Hillier, University of Aberdeen, Sustainable Food Lab Carbon Webinar.

Complexity	Models	Requirements, data/money	Aggregation level, uncertainty	Notes
Tier 1	IPCC Tier 1	N application rates, previous land use etc.	National, annual resolution	Suitable for rough overviews; limited data available
Tier 2	"Not" The Unilever Calculator	N application rates, fertilizer type, soil properties etc.	Regional, annual resolution	Suitable for regions, e.g. counties, cooperatives, large landowners
Tier 3	DAYCENT	N application rates, soil properties, climate data etc.	Site scale, weekly resolution	Suitable for individual farmers, known crops, yields, good availability of climate data
	DNDC	N application rates, soil properties, climate data etc.	Site scale, weekly resolution	Suitable for individual farmers, known crops, yields, good availability of climate data
	Measurement	Money!	Site scale, daily resolution	

Developing a common framework

Definition and agreement on a set of higher-level goals with broad participation and input is an ambitious task and poses numerous technical, methodological and institutional challenges. Strong leadership can help organizations move beyond coordination to collaboration, consensus and compromise.

The government did it….

The Millennium Development Goals (MDGs) are one of the highest-profile core sets of common goals and indicators. They consist of eight goals, to be achieved by 2015, that respond to the world's main development challenges. The eight MDGs are broken down into 21 quantifiable targets that are measured by 60 indicators. The process took years, but in 2000, the MDGs were adopted by 189 nations and signed by 147 heads of state and governments. The MDGs are not universally applicable and accepted, they focus on the least developed countries, and they are weak in the environmental dimension (though a new target was added in September 2008). However, even for organizations with different evaluation models and goals, the MDGs have become a standard reference set for organizations including the World Bank, FAO, United Nations Development Programme, and many other international and national organizations.

None of the objectives of the MDGs are new, but they were repackaged with time-bound targets and indicators. Critics have argued that the scientific basis for developing and realizing the MDGs is weak. The MDGs might not be all-inclusive, address every aspect of sustainable development, fit into existing frameworks or be consistently applied to date—but almost 190 governments signed on, and every donor agency is aligning their monitoring and reporting to the MDGs.

The tourism industry did it…

The Partnership for Global Sustainable Tourism Criteria is a coalition of 27 organizations working together to foster increased understanding of sustainable tourism practices and the adoption of universal sustainable tourism principles. The initiative began over two years ago, reached out to close to 100,000 tourism stakeholders and analyzed more than 4,500 criteria from more than 60 existing certification systems and other voluntary sets of criteria.

The initial goal was to create a core set of critical issues and no more than 15 to 20 key indicators. This grew over time to a theme-based conceptual framework with 37 criteria. The criteria are not universally applicable, are sometimes vague and are not performance based. However, the process of developing the criteria brought together stakeholders from very different backgrounds, needs and priorities who reached agreement on a common language of sustainable tourism.

The UN Food and Agriculture Organization and ISEAL have begun a process to develop a sustainability framework as a move toward a shared understanding of key sustainability goals and core indicators.

In summary, developing a common framework is a political process requiring compromise and consensus. There is no perfect set of core goals, criteria or indicators for sustainability. The process of defining a common language is valuable, though it takes time and is ongoing and evolving. Issues can be added in a stepwise approach.

4.2.5 Building the base: bottom up

In order to understand the relation between specific practices and outcomes and impacts at higher levels, there is the need to build bodies of evidence of farming practices across systems, regions and crops. This cannot be done in isolation by one institution or metric initiative. Large sample sizes are needed to provide quantitative data that are statistically reliable, but this also increases costs. Data-sharing and partnerships between metric initiatives will be essential at all levels. Coordinated data collection, analysis and reporting require common frameworks, criteria and indicator development standards based on international guides such as BellagioSTAMP or, at the sector level, the ISEAL Code of Good Practice for Standard-Setting.

A general capacity question arises related to all of the technical, institutional and methodological challenges. In order to deal with these, capacity needs to be developed at various scales. Some of the needed capacity is technical, but the need is also related to managing the indicator process, linking with other scales, integrating indicators into decision-making, and outreach, communication and branding.

Sharing or harmonizing information infrastructures may be the key to addressing the challenges presented in the paper. A recent study by the Foundation Strategy Group looked at twenty efforts to develop shared approaches to performance, outcome, or impact measurement across multiple organizations and identified shared platforms as one of the major breakthroughs.

As Kramer, Parkhurst and Vaidyanathan (2009) say of shared measurement platforms:

> These systems allow organizations to choose from a set of measures within their fields, using web-based tools to inexpensively collect, analyze, and report on their performance or outcomes. Benefits include lower costs and greater efficiency in annual data collection, expert guidance for less sophisticated organizations, and improved credibility and consistency in reporting… And, as the field gains experience in developing these systems, the effort and investment to launch new systems will likely decrease. (p. 1)

The GeoPortal Case

Started in 2006 by several leading conservation organizations and with a minimal budget, this project provided a catalogue and common platform for GIS data across dozens of institutions. In July 2009 it was decommissioned, despite having registered 800 users and collecting over 4,000 metadata records over its three-year existence. The following is reprinted with minor edits from a July 1, 2009, email from Frank Biasi, director, Conservation Projects, National Geographic Maps.

Here are the top 10 lessons learned from this initiative. These lessons could be applied to similar efforts:

1. Outreach and promotion through various means are essential.
2. Data or metadata publishing should be required and supported by managers and funders.
3. In-kind support is great, but funds for maintenance, upgrades, curation and marketing are essential.
4. Centralized metadata creation is effective and efficient if funds are available.
5. Portals should allow filtering by organizations, including branded subportals.
6. Portals should support organizations' internal and external publishing needs.
7. Without dedicated stewards, web channels should be populated automatically, not manually.
8. Usability and simplicity in finding and posting content is essential.
9. Map-viewing software should be simple and usable for non-technical staff.
10. The *concept* of sharing data is much more advanced than the *practice*.

5.0 Recommendations

This paper has reviewed a number of the issues, constraints and challenges involved in using indicators consistently and in a harmonized way at different levels to understand sustainable agricultural practices and impacts. The review revealed that strengthening the coherence and consistency of indicator systems across different scales that, in turn, can help track the contribution of farm-level practices to agroecosystem-level outcomes is necessary but difficult and has several potential advantages. We do not wish to give the impression that nothing can or should be done. The need is clear, but many financial, political and scientific issues are unresolved and will require strong leadership and collaboration.

The complexity of the underlying issues around sustainable agriculture, and the challenges and costs to measure them, are much larger than commonly thought. In fact, the challenge is not simply what to measure and how, but to determine the implications of more relevant and credible indicators for accountable decision-making. A precondition for this is much stronger and more systematic integration and harmonization to share and build indicators. We need a stepwise, evolutionary approach to measurement, reporting and verification systems, and it should focus on a few strategic indicators as a starting point, embedded in a broad conceptual framework.

Our recommended path forward is composed of five main elements:

1. Standardized terminology and common conceptual frameworks.
2. Coordination and integration.
3. Criteria and indicator development standards.
4. Capacity development.
5. An evolutionary process.

1. Standardized terminology and conceptual frameworks

- We must standardize terminology and develop common conceptual frameworks to understand how the different levels, from farms to entire agroecosystems, are related. We need clarity on the expected outcomes, which includes a vision of the overall system and how the information system will work—not just the environmental outcomes. This process should build upon existing frameworks to affirm, modify and adopt what is working.
- We need to more clearly understand the levels where better alignment is needed and how we can achieve alignment of the metrics themselves. Because of the many differences between farms, crops, supply chains, markets, geographies and agroecosystems, finding commonality must start at the level of values, principles, criteria, categories of metrics and their relationships as outlined in conceptual frameworks.

- We need alignment at the level of processes: responsibilities, cost sharing, and who benefits and how from the process.
- We need to find commonalities from a practical standpoint while recognizing the challenges of not just these differences but the different governance structures of the metric initiatives themselves.
- Stakeholders with different interests (farmers, retailers, governments, NGOs and so on) need to come to agreement on at least common categories of concerns and linkages that can serve as a basis for developing indicators.

2. Coordination and integration

- We need coordination and integration of various metric initiatives under common frameworks and open data platforms in order to build evidence of what works and under what conditions. This knowledge is the foundation for building science-based metrics. We need to build non-traditional networks and partnerships across disciplines.
- We need to learn from metric initiatives in other sectors or operating at different levels.
- We need to coordinate with national statistical agencies and agencies in charge of geospatial data infrastructure related to the environment and national environmental accounts.
- We need to establish open platforms with data-sharing facilitation, not just on the basis of research projects and initiatives, but as operational systems with the required technical and institutional infrastructure.
- We need to establish common protocols, data collection, analysis and reporting mechanisms throughout the supply chain.
- We need to collect data on large numbers of farms in order to build models and establish targets and thresholds for science-based accepted norms based on local factors.
- We need to define and establish baseline data and information.

3. Criteria and indicators

- Criteria and indicator development should be done using best practice standards and international guides such as the BellagioSTAMP or the ISEAL Code of Good Practice. We need to emphasize participation and transparency at all levels.
- We need to identify common denominators and metadata standards.

- We need to link farm-level management indicators to higher-level and off-site outcomes and impacts, including defining indicator sets linked to ecosystem goods and services.
- We need to disaggregate indicators as far as possible, in order to work at different spatial and temporal levels.
- We need to review the value of indicators during development at different levels for different purposes and users.
- We need to adapt indicators to local contexts while continuing to comply with key principles of the framework.
- We need to identify indicator targets, with timelines, linked to the goals and objectives both at the higher (ecosystem) and lower (farm) levels. This would help make a link to expected outcomes and clarify—to some extent—how macro-scale outcomes should be disaggregated.
- We need to define links of metrics to incentives, payments and accountability mechanisms.

4. Capacity development and resources

- We need to cultivate existing capacity rather than building from the ground up or reinventing the wheel.
- We need to include resource commitments beyond indicator selection and development, for information and knowledge systems (next point), communication and outreach, evaluation, learning and improvement.
- We need to share resources and platforms to build expertise and learning platforms (Kramer, Parkhurst & Vaidyanathan, 2009).
- We need emphasis on building upon and developing capacity at all levels with regard to processes (such as framework and indicator development), information systems (for data collection and management) and products (such as indicators, analysis and dissemination). This is particularly critical for farmers.
- We need to identify key tools and facilitate sharing them across initiatives using common capacity development materials.
- We need to use information technology to manage complexity and coordination demands.
- We need to use GIS for presenting information in a geocoded, spatial system, including collecting and presenting non-GIS information in the same kind of system.

5. Evolutionary process

- We need to understand that the process of developing an indicator system requires collaboration, cooperation and even compromises. It is an evolutionary process, as we learn from the system and the different stakeholders.
- We need to adopt a stepwise approach with a few (two to three) indicators related to high-priority issues, using one commodity crop as a common starting point to learn, adapt and improve.
- We need to accept that the process is not holistic or comprehensive.
- We need to keep in mind the cost and effort involved in gathering and analyzing required information. It is better to develop a manageable and trackable set of indicators than have too many, which will stretch capacity.
- We need to link to evaluating, proving and auditing the impact of standards and certification systems.
- We need to think in terms of intellectual scaffolding: beginning with the familiar and building upon this as a form of problem-based learning, adding new information as we learn.
- We need to build metrics into organizational, operational planning cycles.
- We need to include calibration and adaptive management within the process.
- As we learn, we need to move to other commodities and agroecosystems.

References

Agriculture and Agri-Food Canada. (2005). *Environmental sustainability of Canadian agriculture: Agri-environmental indicator report series, report #2*. Ottawa: Agriculture and Agri-Food Canada.

Amara, N., Ouimet, M., & Landry, R. (2004). New evidence on instrumental, conceptual, and symbolic utilization of university research in government agencies. *Science Communication, 26*(1), 75–106.

Atkinson, G., Baldock, D., Bowyer, C., Newcombe, J., Ece Ozdemiroglu, D. P., & Provins, A. (2004). *Framework for environmental accounts for agriculture: Final report*. London: Economics for the Environment Consultancy.

Backstrand, K. (2003). Civic science for sustainability: Reframing the role of experts, policy-makers and citizens in environmental governance. *Global Environmental Politics, 3*(4), 24–41.

Beyer, J.M. (1997). Research utilization: Bridging a cultural gap between communities. *Journal of Management Inquiry, 6*(1), 17–22.

Block, B. (2009, July 20). Wal-Mart Scrutinizes Supply-Chain Sustainability. WorldWatch Institute, *Eye on Earth.*

Bormann, B. T., Cunningham, P. G., Brookes, M. H., Manning, V. W., & Collopy, M. W. (1993). Adaptive ecosystem management in the Pacific Northwest. U.S. Forest Service, General Technical Report PNW-GTR-341.

Brang, P., Courbaud, B. C., Fischer, A., Kissling-Näf, I., Pettenella, D., Schönenberger, W.,... Grimm, V. (2002). Developing indicators for the sustainable management of mountain forests using a modelling approach. *Forest Policy and Economics, 4*(2), 113–123.

Clark, W.C. (2007). Sustainability science: A room of its own. *Proceedings of the National Academy of Sciences, 104*(6), 1737–1738.

Clarke, R. (2009). *Roger Clarke's information infrastructure home-page.* Retrieved from http://www.rogerclarke.com/II

Clay, J. (2004). *World agriculture and the environment: A commodity-by-commodity guide to impacts and practices*. Washington D.C.: Island Press.

Giovannucci, D., & Purcell, T. (2008). *Standards and agricultural trade in Asia. ADBI discussion paper 107*. Tokyo: Asian Development Bank Institute.

Glenn, J. C., Gordon, T. J., & Florescu, E. (2009). *2009 State of the Future*. Washington DC: Millennium Project.

Grigg, A. (n.d.). *Dependency and impact on ecosystem services—unmanaged risks, unrealised opportunities*. Cambridge: The Natural Value Initiative.

H. John Heinz III Center for Science, Economics and the Environment. (2008). *Environmental information: Road map to the future*. Washington, D.C.: Heinz Center.

Häni, F. J., Pintér, L., & Herren, H. R. (2007). *Sustainable agriculture: From common principles to common practice. Proceedings and outputs of the first Symposium of the InternationalForum on Assessing Sustainability in Agriculture (INFASA)*. Bern: IISD.

Henson, S., & Humphrey, J. (2008). *Understanding the complexities of private standards in global agri-food chains.* Sussex, U.K.: Institute of Development Studies.

International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD). (2009a). *International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD): Global report.* Washington, D.C.: Island Press.

International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD). (2009b). *International Assessment of Agricultural Knowledge, Science and Technology Development. Global summary for decision-makers.* Washington, D.C.: Island Press.

International Institute for Sustainable Development (IISD) and the United Way of Winnipeg. (2007). *Business plan for the Winnipeg Community Indicator System.* Winnipeg: IISD.

International Institute for Sustainable Development (IISD) and Organisation for Economic Co-operation and Development (OECD). (2009) *BellagioSTAMP—Sustainability assessment and measurement principles*. Winnipeg: IISD. Manuscript in preparation.

International Institute for Sustainable Development (IISD). (n.d.). *Complete Bellagio Principles.* Winnipeg: IISD.

International Social and Environmental Accreditation and Labelling Alliance (ISEAL). (2009a). ISEAL Alliance Strategic Plan 2009–2013. London, U.K.: ISEAL.

International Social and Environmental Accreditation and Labelling Alliance (ISEAL). (2009b). *ISEAL. R083 State of the art in measuring the impacts of social and environmental standards: Methodology.* London, U.K.: ISEAL.

King, C., Gunton, J., Fairbairn, D., Coutts, J., & Webb, I. (2000). The sustainability indicator industry: Where to from here? A focus group study to explore the potential of farmer participation in the development of indicators. *Australian Journal of Experimental Agriculture, 40*(4), 631–642.

Komiyama, H., & Takeuchi, K. (2006). Sustainability science: Building a new discipline. *Sustainability Science, 1*(1), 1–6.

Kramer, M., Parkhurst, M., & Vaidyanathan, L. (2009, July). *Breakthroughs in shared measurement and social impact.* Retrieved from http://www.fsg-impact.org/ideas/item/breakthroughs_in_measurement.html

Kurtz, J. L. (2001). Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. *Ecological Indicators, 1*(1), 49–60.

Lausch, A., & Herzog, F. (2002). Applicability of landscape metrics for the monitoring of landscape change: Issues of scale, resolution and interpretability. *Ecological Indicators, 2*(1–2), 3–15.

McCool, S.F., & Stankey, G.H. (2004). Indicators of sustainability: Challenges and opportunities at the interface of science and policy. *Environmental Management, 33*(3), 294–305.

Meadows, D. (1998). *Indicators and information systems for sustainable development.* Hartland, Vermont: Sustainability Institute.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human well-being: Opportunities and challenges for business and industry*. Washington, D.C.: World Resources Institute.

Natural Value Initiative. (n.d.). *Agriculture & biodiversity*. Retrieved from http://www.naturalvalueinitiative.org/content/002/201.php

Nellemann, C. M. (2009). *The environmental food crisis: The environment's role in averting future food crises. A UNEP rapid response assessment.* Arendal, Norway: UNEP/GRID-Arendal.

Organisation for Economic Co-operation and Development (OECD). (2001b). *Environmental indicators for agriculture: Concepts and framework, Volume 3*. Paris: OECD.

Organisation for Economic Co-operation and Development (OECD). (2001a). Agri-environmental indicators. In *Environmental indicators for agriculture: Concepts and framework, Volume 3*. Paris, France: OECD. Retrieved from http://www.oecd.org/dataoecd/28/1/1890235.htm

Payen, L. (2008). *Developing and testing indicators of key environmental impacts of agriculture*. 66. European Engineer Degree CAH Dronten thesis.

Perkins, Payen, & Potashnick. (2008). *2008 Measure what matters.* Internal discussion paper. Gland, Switzerland: WWF.

Pintér, L., Hall, J., Scrivens, K., Hardi, P., & Giovannini, E. (2009). *Bellagio Principles Revisited: Working Draft paper*. Winnipeg: IISD.

Pintér, L., Hardi, P., & Bartelmus, P. (2005). *Sustainable development indicators: Proposals for the way forward.* Winnipeg: IISD.

Roetter, R. P., Keulen, H. V., Kuiper, M., Verhagen, J., & Laar, H. H. (2007). *Science for agriculture and rural development in low-income countries*. Netherlands: Springer Netherlands.

Rosenström, U. (2009). *Sustainable development indicators: Much wanted less used?* Monographs of the Boreal Environment Research. Helsinki: Edita Prima Ltd.

Royal Society for the Protection of Birds (RSPB). (2003, December 11). *Measuring real progress: Headline indicators for a sustainable world.* Retrieved from http://www.rspb.org.uk/ourwork/policy/economicdevelopment/indicators.asp

Ruijter, F. D. (2009). *Sustainability and biomass production: Why—what—who—where?* Holland: Wageningen UR.

Russillo, A. (2008). *ISEAL R082 state of the art in measuring the impacts of social and environmental standards: Issues.* London: ISEAL.

Rydin, Y., Holman, N. & Wolff E. 2003. Local sustainability indicators. *Local Environment, 8*(6), 581–589.

Sinner, J. (n.d.). *The development of agri-environmental indicators at the OECD*. Wellington, New Zealand: MAF Information Services. Retrieved from http://www.maf.govt.nz/mafnet/rural-nz/sustainable-resource-use/land-management/sustainable-agriculture-conference/susconf3.htm

Tilman, D. (1999). Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proceedings National Academy Science, 96*(11), 5995–6000.

Turner, M., Gardner, R., & O'Neill, R. V. (2001). *Landscape ecology in theory and practice: Pattern and process.* New York: Springer Verlag.

U.K. Department for Environment, Food and Rural Affairs. (2007). *Commission on Environmental Markets and Economic Performance*. London, U.K.: DEFRA.

United Nations Development Programme. (2002). *RBM in UNDP: Selecting indicators*. New York: UNDP.

United Nations Food and Agriculture Organization (FAO). (2006). *World agriculture: towards 2030/2050*. Rome: FAO.

United Nations Food and Agriculture Organization (FAO). (2007b). *The State of Food and Agriculture: Paying Farmers for Ecosystem Services*. Rome: FAO.

United Nations Office of Internal Oversight Services. (2008). *Review of results-based management at the United Nations*. Report of the Office of Internal Oversight. Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=A/63/268

United Nations. (2007). *Indicators of sustainable development: Guidelines and methodologies* (3rd Ed.). New York: United Nations.

Van der Werf, H. M. G., & Payraudeau, S. (2004). Environmental impact assessment for a farming region: A review of methods. *Agriculture, Ecosystems & Environment, 107*(1), 1–19.

Van der Werf, H. M. G., & Petit, J. (2002). Evaluation of the environmental impact of agriculture at the farm level: A comparison and analysis of 12 indicator-based methods. *Agriculture, Ecosystems & Environment, 93*(1–3), 131–145.

Verhagen, J., Wösten, H., & de Jager, A. (2004). Agriculture and environment. In R.P. Roetter, H. Van Keulen, M. Kuiper, J. Verhagen, & H.H. Van Laar (Eds.), *Science for agriculture and rural development in low-income countries* (57–75). New York: Springer.

World Bank. (2007). *World Bank development report 2008: Agriculture for development.* Washington, D.C.: World Bank.

World Business Council for Sustainable Development. (2008a). *Measuring impact: A framework methodology version 1.0*. Washington, D.C.: World Business Council.

World Business Council for Sustainable Development. (2008b). *Measuring impact: Beyond the bottom line*. Switzerland: Atar Roto Presse SA.

World Commission on Environment and Development. (1987). *Our common future.* Oxford: Oxford University Press. Retrieved from http://worldinbalance.net/agreements/1987-brundtland.php

Worldwatch Institute. (2006). *State of the World 2006*. Washington, D.C.: Worldwatch Institute.

Worldwatch Institute. (2009). *Vital Signs 2009*. Washington, D.C.: Worldwatch Institute.

Wunderlich, C., & Russillo, A. (2007). *Survey of prominent criteria and indicator initiatives: Lessons learned and needs assessment*. Winnipeg: IISD.

Zwerdling, D. (2009, April 13). India's farming "revolution" heading for collapse. *National Public Radio*. Retrieved from http://www.npr.org/templates/story/story.php?storyId=102893816

Further reading

GTZ. (2007). *Outcomes and impacts of standards and standards initiatives: Overview and analysis of available outcome and impact studies*. Bonn, Germany: GTZ.

Halberg, N., Verschuur, G., & Goodlass, G. (2005). Farm level environmental indicators: Are they useful? An overview of green accounting systems for European farms. *Agriculture Ecosystems and the Environment , 105*(1–2), 195–212.

United Nations Food and Agriculture Organization (FAO). (2007a). *Private standards in the United States and European Union markets for fruit and vegetables*. Rome: FAO.