FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Guide for the Sustainable Intensification Assessment Framework

24 October 2017

By:

Mark Musumba (UF[1]), Philip Grabowski (MSU[2]), Cheryl Palm (UF[1]) and Sieglinde Snapp (MSU[2])

Steering Committee:

Gundula Fischer (IITA[3]), Bruno Gerard (CIMMYT[4]), Jerry Glover (USAID[5]), Fred Kizito (CIAT[6]), Generose Nziguheba (IITA[3]), Vara Prasad (KSU[7]), Peter Thorne (ILRI[8]), and Bernard Vanlauwe (IITA[3]).

Acknowledgments:

We are grateful to the Feed the Future Innovation Lab for Sustainable Intensification, Africa Research in Sustainable Intensification for the Next Generation (Africa RISING), and Consortium for Improving Agriculture-based Livelihoods in Central Africa (CIALCA) researchers for their support and contributions during the field visits and trainings.

This study is made possible by the support of the American People provided to the Feed the Future Innovation Lab for Sustainable Intensification (SIIL) through the United States Agency for International Development (USAID). The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Program activities are funded by the United States Agency for International Development (USAID) under Cooperative Agreement No. AID-OAA-L-14-00006.

Affiliations:

[1]University of Florida; [2] Michigan State University; [3]International Institute for Tropical Agriculture; [4]International Maize and Wheat Improvement Center; [5]United States Agency for International Development; [6]International Center for Tropical Agriculture; [7]Kansas State University; [8]International Livestock Research Institute.

USAID FROM THE AMERICAN PEOPLE

FEED THE FUTURE INNOVATION LAB
Collaborative Research on Sustainable Intensification

KANSAS STATE UNIVERSITY

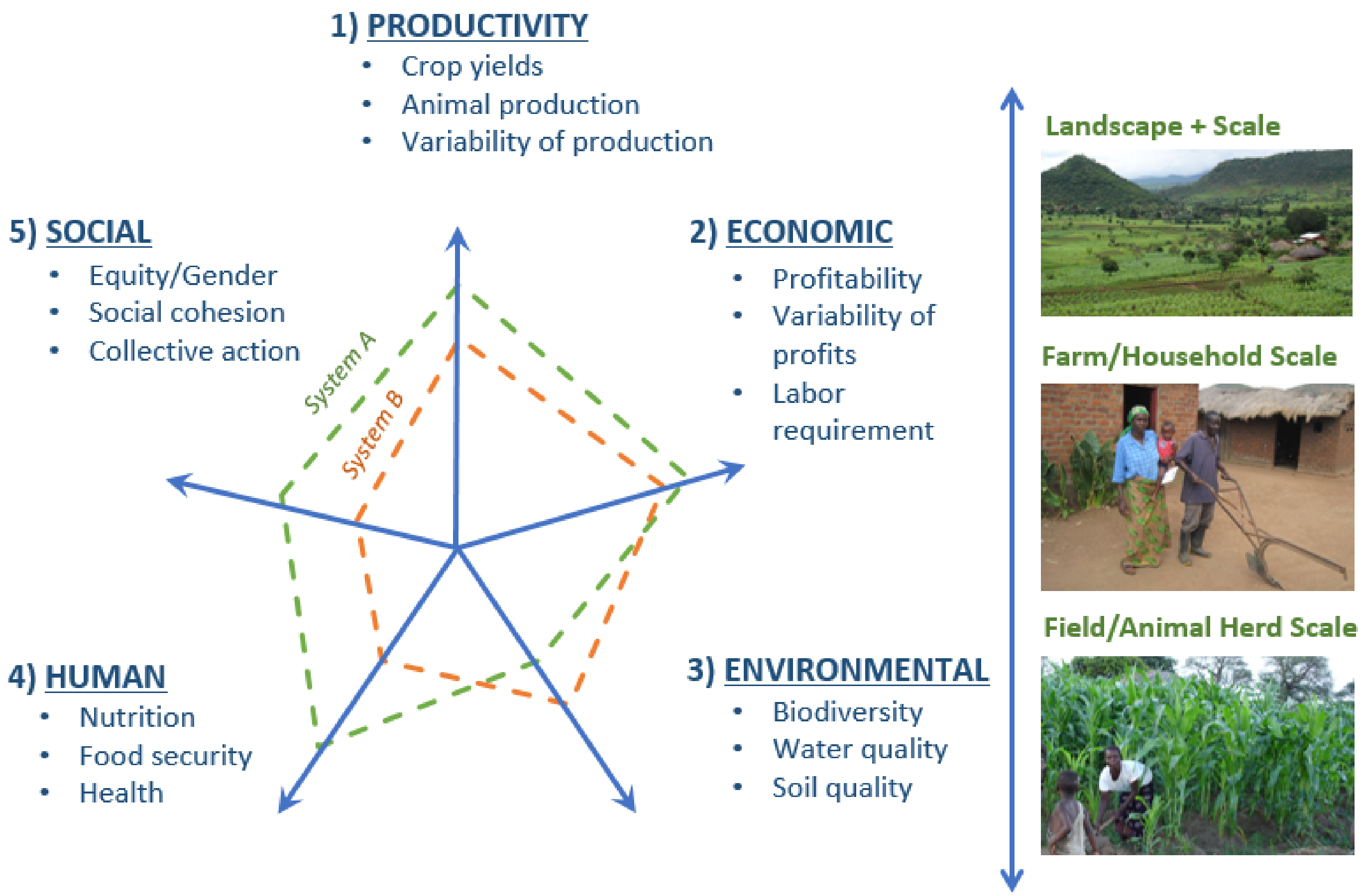
Table of Contents

Guide for the Sustainable Intensification Assessment Framework

1. Introduction

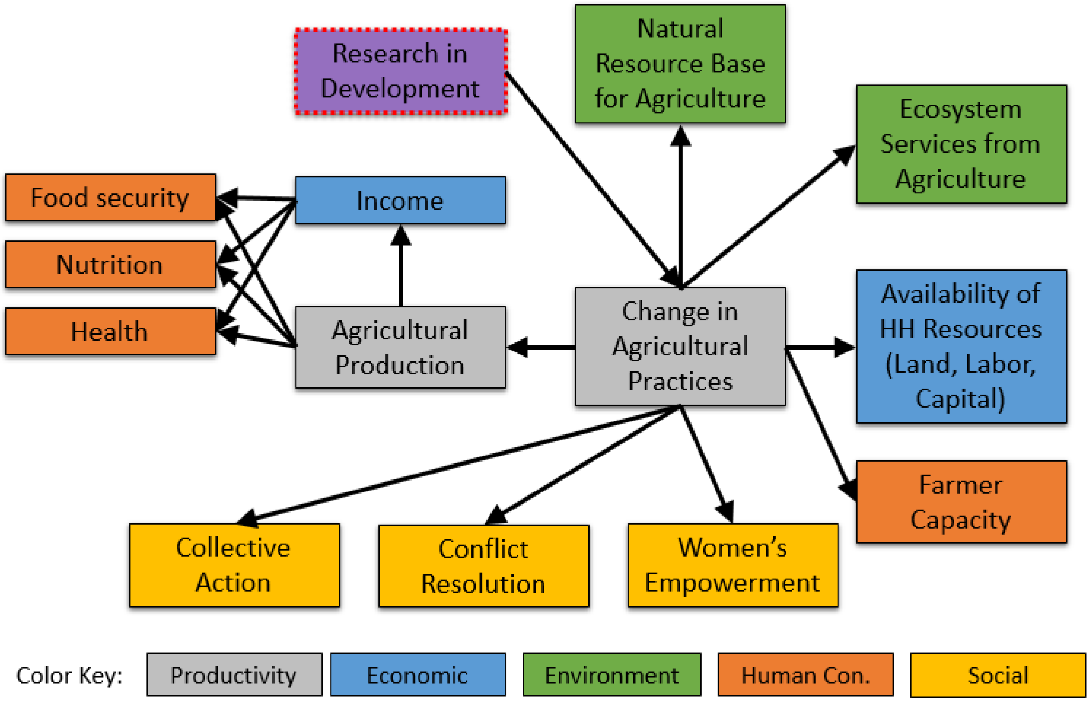
Sustainable Intensification (SI) offers a means to balance the environmental, economic, and social objectives of agriculture. Agricultural intensification may be defined as increasing output per unit input per unit time. A narrow definition of sustainable intensification is “production of more food on the same piece of land while reducing the negative environmental impacts and at the same time increasing the contributions to natural capital and flow of environmental services” (Zurek et al., 2015). The definition of SI has evolved to include non-environmental dimensions such as social issues, economics, and the human condition (Loos et al., 2014). The inclusion of social aspects helps ensure a balanced approach to the intensification process. In this guide, we present a framework of objective-oriented SI indicators organized into five domains critical for sustainability: productivity, economic, environment, human condition, and social domains. The objective-oriented indicator assessment is similar to the goal-oriented framework proposed by Olsson et al. (2009) in which objectives of the innovation are identified and then indicators are linked to the objectives to assess performance in a balanced approach across domains. The metrics for each indicator are categorized across spatial scales: field, farm, household, and landscape, so that the assessment can be used for innovations at any scale and so that cross-scale linkages can be considered (Figure 1).

Figure 1: Interlinkages across the five domains of sustainable intensification and across spatial scales with examples of indicators for each domain.



The framework was developed to provide indicators for assessing the relative sustainability of an innovation across the five domains. Our target audience is researchers involved in developing and analyzing innovations for SI, particularly in the research for development context. The framework was developed primarily for use in smallholder farming contexts where changes in agricultural production can have positive or negative effects on development goals such as alleviating poverty, avoiding land degradation, increasing food security and nutrition security, and supporting women's empowerment (Figure 2). To develop innovations that can support these diverse goals, research to assess SI innovations needs to be interdisciplinary, drawing upon the theories and methods of the biophysical and social sciences. Ideally, this framework will support assessment of sustainable agriculture intensification innovations through interdisciplinary, iterative co-learning approaches. The framework is not designed for project evaluation, though it can contribute towards that goal.

Figure 2: Conceptual framework for potential effects (positive or negative) from agricultural research to development goals across the five domains.



This document aims to strengthen researchers' ability to holistically assess the performance of an innovation in terms of the direct and indirect consequences within and across domains. In the following section, we introduce key concepts related to the assessment of sustainable intensification, and we describe the process used for developing the framework and how it relates to other sustainability assessments. We provide a wide range of indicators that we judge are useful metrics in a SI assessment process. Objectives of users of the SI framework are expected to vary; this framework helps put the objectives into context so that indicators will be chosen for a specific purpose and objective. The criteria for inclusion of the indicators is a follows: 1) the indicators are deemed scientifically sound, with broad acceptance by scientists working on sustainable agricultural intensification; 2) the indicators are clearly defined and easy to understand; and 3) the indicators are sensitive to changes in innovation or management practices; and 4) the indicators are measurable by researchers. This guide and framework of indicators was developed through an inclusive process that included a survey of scientists working on sustainable intensification, a literature review on critical indicators for sustainable intensification (Smith et al., 2017), and discussions with scientists during site visits and workshops on the important indicators used in their research work. The indicators described are part of a living document, and we anticipate that this will continue to be expanded and refined over time by users of the SI framework.

In section 3, we provide step-by-step instructions to enable scientists and partners engaged in research for development to assess innovations across multiple objectives. This includes selection of relevant indicators across the multiple objectives and domains and a practical analysis of tradeoffs, synergies, and relative sustainability. We support the use of a transparent approach to selecting indicators across the five domains with explicit reasoning for choosing indicators and excluding others. Indicators not listed in this framework may also be used, for example, through a participatory process with stakeholders to identify locally applicable metrics for assessing SI (see, for example, Eele, 1994). We provide guidance to researchers in selecting metrics that are feasible given resource constraints, which are often highly limited. We also outline an exercise using a causal loop diagram for considering the tradeoffs and synergies from an innovation within and across domains and scales; this exercise will highlight possible additional indicators that are important to assess.

In section 4, we provide tables for each domain with lists of indicators and metrics organized by spatial scale. A brief description of each indicator is included. Further information about the metrics and measurement methods can be found in the SI indicator manual (http://www.k-state.edu/siil/resources/assessmentframework/index.html).

2. Assessment of Sustainable Intensification

2.1. What is sustainable intensification?

Sustainable Intensification focuses on improving the efficient use of resources for agriculture, with the goal of producing more food on the same amount of land but with reduced negative environmental or social impacts. The term "sustainable intensification" originated in the 1990s in the context of how to achieve improved yields over the long-term in fragile environments of Africa (Pretty, 1997; Reardon et al., 1995). Intensification has the potential to reduce pressure for conversion of natural lands to agriculture (Cook et al., 2015).

The need for this intensified production of food, fuel, and fiber to be “sustainable” comes from the realization that intensification may not provide long-term stable production, especially if it degrades soil or water resources. Recent SI work has put a major emphasis on management strategies that can reverse land degradation and reduce yield losses despite climatic changes (Dahlin and Rusinamhodzi, 2014). Much of this SI research focuses on environmental aspects of sustainability using biological and ecological principles to improve the ecosystem services of a given farming system and to reduce the environmental problems associated with it (Petersen and Snapp, 2015).

Production practices that are environmentally sound and economically profitable may have complex social dimensions that affect sustainability. SI is often presented as a solution to food insecurity and malnutrition. However, achieving those goals requires fair distribution of the net benefits from increased productivity. For this reason, SI interventions need to explicitly consider issues of equity, poverty alleviation, and gender empowerment (Loos et al., 2014). A shift towards intensified production can indirectly result in problems regarding poverty, food security, nutrition, health, and/or social issues. When these problems persist, they can even decrease the stability of the increased productivity.

Sustainable intensification is not a particular set of agricultural practices. There can be many pathways to sustainable agricultural intensification that will vary by location and scale based on the agro-ecological zone, farming system, cultural preferences of farmers, institutions and policies, as well as other factors (Pender et al., 1999). Each of these pathways will have a unique set of changes in management practices or technologies that will lead to varying environmental and socioeconomic tradeoffs and/or synergies across and within domains. Thus, SI should be used as a conceptual framework for guiding how to achieve balanced outcomes from changes in agriculture (Garnett and Godfray, 2012). Unfortunately, the term “sustainable intensification” is often used to describe any type of agricultural intensification that may have potential environmental benefit (Godfray, 2015). In contrast, the SI indicator framework presented here aims to provide practical means to consider multiple dimensions of sustainability.

A variety of indicator frameworks have been developed to assess progress towards sustainability in agriculture. Many of these frameworks provide information on a single aspect of a system, such as soil health, nutrition, or poverty alleviation (Eele, 1994; Bockstaller and Girardin, 2003; Niemeijer et al., 2008; Gustafson, et al., 2016). Other system-based frameworks evaluate multiple attributes of the system such as resilience, stability, adaptability, self-reliance, equity, and reliability (Lopez-Ridaura et al., 2005; Conway, 1994). Using a system based indicator framework requires a thorough understanding of the agricultural system in which the innovation is implemented (Van Cauwenbergh et al., 2007); this approach may have limited application without systematic guidance. Here we present a goal-oriented framework, where the researcher lists the primary goal of the innovation or project and identifies several operational goals under each domain that are then used to select indicators and evaluate the innovation.

2.2. The Sustainable Intensification Assessment Framework

2.2.1. Purpose of the SI assessment framework

Sustainability assessment has progressed towards the use of indicator frameworks that provide a basis for selection of a core list of indicators from a comprehensive list of indicators. Numerous indicators have been used and recommended for assessing sustainable agricultural intensification (Lopez-Ridaura et al., 2002; Speelman et al., 2007; ISPC, 2014; Smith et al., 2017; Mahon et al., 2017), but few have explicitly explored the needs scientists face in using sustainability indicators in research for development(Smith et al., 2017). The sustainable intensification indicator framework described in this document aims to provide a synthesized list of indicators and metrics and means to explore all the domains of sustainability. The indicators and guidelines presented should not be seen as the only way to assess SI. Instead, the goal is to provide a common framework that can guide research on SI and facilitate cross-program learning and assessment on the factors that lead to successfully working towards sustainable intensification (AAAS, 2015).

The framework is primarily intended to guide agricultural scientists working in research for development but is flexible and can be used by scientists interested in sustainable intensification more broadly. Scientists may use this framework for a pre-adoption assessment of the potential sustainability of their innovation. This pre-adoption assessment provides important information for use in the adoption phase (roll out or scale up phase) of the innovation. The framework of indicators and metrics provided below includes both ‘gold standard’ approaches, as well as, simplified methods and metrics as options that may be more feasible to use considering the spatial, temporal, and cost limitations. From these tables of indicators researchers and stakeholders can select those most relevant to their programs.

This indicator framework can be used to analyze the relative sustainability of innovations for intensification by collecting data for the most relevant indicators for an innovation and comparing them with the status quo. The status quo is often some form of practice common in the same location. It is important to have a fair comparison so that potential benefits of the innovation are not overstated. In some cases, multiple comparisons may be needed. For example, in section 3.5.1 we summarize a study (Snapp et al., submitted) where the relative sustainability of intercropped and fertilized maize and legumes is compared to both unfertilized sole maize (the most common farmer practice) and fertilized sole maize (another farmer practice that aids in distinguishing the effect of the legumes from the effect of the fertilizer). Where long-term data is available, the SI indicators framework can also be used to quantify trajectories of sustainable intensification by comparing indicators from all domains across time.

Data on SI indicators can be presented through visualization techniques, such as radar charts to compare performance of innovations. Instead of combining indicators into an index (where important details become obscured), we recommend presenting the results for each indicator separately. This allows communities, scientists, implementation partners, and policy makers to objectively evaluate the research results based on the importance they assign to each indicator. Different stakeholder groups may have different priorities regarding sustainability related goals (e.g., biodiversity conservation, agricultural production, food security, and gender equity). There is a growing move towards developing composite indicators for each sustainability pillar or domain and for all domains (Gómez-Limón and Sanchez-Fernandez, 2010; Haileslassie et al., 2016). Although such composite indices can be estimated using this framework, we believe that estimating and presenting individual indicator to stakeholders provides a transparency and parsimony to identifying change and performance.

A critical component of this assessment is to identify potential tradeoffs and synergies from an SI intervention. In the exercise provided in Section 3.2, researchers can consider how the various indicators listed under each domain might be affected positively or negatively by an intervention that they are investigating or planning to research. This exercise provides a structured means of considering the broader farming and livelihood systems and selecting the indicators that reflect these potential tradeoffs and synergies. This type of qualitative assessment should be informed by the scientific literature as well as by discussions with farmers, fellow researchers, NGOs or other stakeholders about the potential direct and indirect effects of a SI innovation. By using this exercise, researchers can anticipate potential synergies and tradeoffs and minimize unintended negative consequences by mitigating them through the research design and implementation.

The SI indicators framework can also be used to guide monitoring and evaluation (M&E) efforts in development projects. All of the key concepts and methods for measuring or estimating the indicators are presented in this framework and the accompanying manual of SI indicators. Several considerations are needed to effectively scale up or aggregate plot and household level indicators to assess the project-level effect (such as at the village, watershed, or sub-district level [Marinus et al., forthcoming]). Nevertheless, the same process for selecting the most relevant indicators and reflecting on synergies and tradeoffs can be applied to M&E for development projects.

2.2.2. Five domains of Sustainable Intensification

The five domains of sustainable intensification, which emerged during discussion by stakeholders in a meeting in Accra, Ghana, in 2013, are productivity, economic, environment, human condition and social domains (Glover, 2016). This framework of five domains distinguishes important aspects of sustainable intensification compared to the three domains used by many sustainability assessments: economic, environmental, and social domains (Lopez-Ridaura et al., 2002; Van Cauwenbergh et al., 2007). The five-domain framework ensures that important aspects such as equity (gender, age, class), nutrition and community factors such as social cohesion and collective action are not overlooked in the indicators selection process.. We are aware that there is overlap among the indicators in the different domains and these overlaps indeed provide additional insights. For our purpose, the domains are described and organized as follows:

Productivity: The productivity domain is critical in capturing productivity both in cropping and livestock systems. Following the SI literature, this domain focuses on land as a critical input. Increasing productivity is the essential characteristic of intensification, with the goal of increasing output per unit of input for a given time period (season or year). In livestock systems, stocking rates or offtake may be used as a measure of intensification, while in cropping systems intensification focuses on yields (Mahon et al., 2017). This domain also captures postharvest losses and cropping intensity (the number of crops per year from the same piece of land). It also contains indicators that may be used to assess the production potential of the land as well as, potential variability due to biophysical aspects. Other inputs associated with intensification (such as labor, water quality, fertilizer, and capital) are captured in the economic domain.

Economic: This domain focuses on issues directly related to the profitability of agricultural activities and returns to factors of production (land, labor, and capital). In addition to profitability, this domain includes indicators related to the productivity of inputs, apart from land, and includes water, nutrients, labor, and capital. Furthermore, indicators likely to affect the probability of investment in enhancing productivity (market participation) are included. Farmers' decisions to choose which crop to grow and how to allocate resources to different activities are affected by marketability of a given commodity and livelihood strategies chosen to improve wellbeing. This domain captures farmers' market orientation, diversification of income sources, and extent and movement towards high value crop production.

Environment: This domain focuses on the natural resource base supporting agriculture (e.g., soil, water, air), the environmental services directly affected by agricultural practices (e.g., habitat, soil water holding capacity, biodiversity) and the level of pollution coming from agriculture (pesticides, eutrophication, greenhouse gases). Improved efficiency metrics are described under the economic domain but are also critical for tightening nutrient and energy cycles, a key principle for sustainable agriculture.

Human condition: This domain contains indicators related to the individual or household, including nutrition status, food security, and capacity to learn and adapt. While some of these concepts are dependent on social interactions (such as within the household or community), they are distinct from those in the social domain that directly focus on interpersonal relationships.

Social: This domain focuses on social interactions of the farming communities or society, including equitable relationships across gender, equitable relationships across social groups, the level of collective action, and the ability to resolve conflicts related to agriculture and natural resource management.

2.2.3. Scales of analysis

Measuring indicators to assess sustainable intensification typically requires observing parameters at a given scale, which determines the unit of analysis, sampling design, and protocols to be used. The tables of indicators presented in section 4 categorize the indicators into four spatial scales: plot level, farm level, household level, and the "landscape or administrative unit." The landscape or administrative unit scale can be defined as community, watershed, district, province, or even the nation as a whole. Observing only one scale can be useful for a specific indicator and domain but assessing the broader implications and interactions with other domains usually requires investigating multiple scales

2.2.4 Definitions

We use the following definitions to distinguish between indicators, metrics, and measurement methods:

Indicator – A "quantitative or qualitative factor or variable that provides a simple and reliable basis for assessing achievement, change or performance" (ISPC, 2014).

Metric – "This represent[s] the values on which indicators are built." These are computed by aggregating and combining raw data, for example, yield (harvest per hectare) or height for age. (ISPC, 2014).

Measurement method – A set of activities to generate raw data (observations such as weight, height, plot size, etc.) that can be used to compute metrics. This can include modeling and the output generated from modeling.

It is important to note that a metric can be an indicator if it is used to assess performance and decision making. "Thus all indicators are metrics, but not all metrics are indicators" (ISPC, 2014).

2.3. Approach used to refine sustainability indicators

To develop a flexible framework, we explored the literature and interacted with scientists to obtain a list of important indicators and then analyzed them for precision and ease to measure. We also carried out field visits to interact with scientists and stakeholders (farmers and other partners) to gather insight in the process of stakeholder engagement, data collection, indicator generation, and perceptions by participants in the process.

The suite of indicators in the framework was generated from our visits and interactions with scientists that include the following:

- Africa Research in Sustainable Intensification for the Next Generation (Africa RISING) meeting with steering committee members – September 2015
- Africa RISING Project in Mali and Millennium Villages Project in Mali – September 2015
- Africa RISING project sites in Ethiopia – November 2015
- Interaction with scientists at the annual meetings for the Sustainable Intensification Innovation Lab (SIIL) – January 2016 and January 2017
- Consortium of Improving Agriculture-based Livelihoods in Central Africa (CIALCA) project in Rwanda
- An online survey of SI researchers – June 2016
- Africa RISING East and Southern Africa (ESA) Phase 1 Legacy meeting – Tanzania, July 2016
- Africa RISING ESA Phase 2 planning meeting – Malawi, October 2016
- Africa RISING West Africa planning meeting – Ghana, February 2017
- Center of Excellence for Sustainable Agricultural Intensification and Nutrition (CE SAIN) – Cambodia, April 2017
- Sustainable Intensification Innovation Lab (SIIL) sub-awardees from Senegal and Burkina Faso – Senegal, May 2017

During this process, we also collected information on the data, methods, and protocols that scientists use to estimate indicators for their various projects. In some projects we found data gaps, meaning that the projects are not collecting data from a sufficient number of indicators to evaluate sustainable intensification. We are proposing data collection methods to fill this gap. Where gaps were identified and new indicators proposed, we presented them to the teams and discussed their relevance and measurability. A similar approach has been used by earlier studies (Zurek et al., 2015; Taylor et al., 1993; Van der Werf and Zimmer 1998) to refine indicators in situations where no other possibility of validation exists (i.e., a new indicator is proposed but with no data to estimate it).

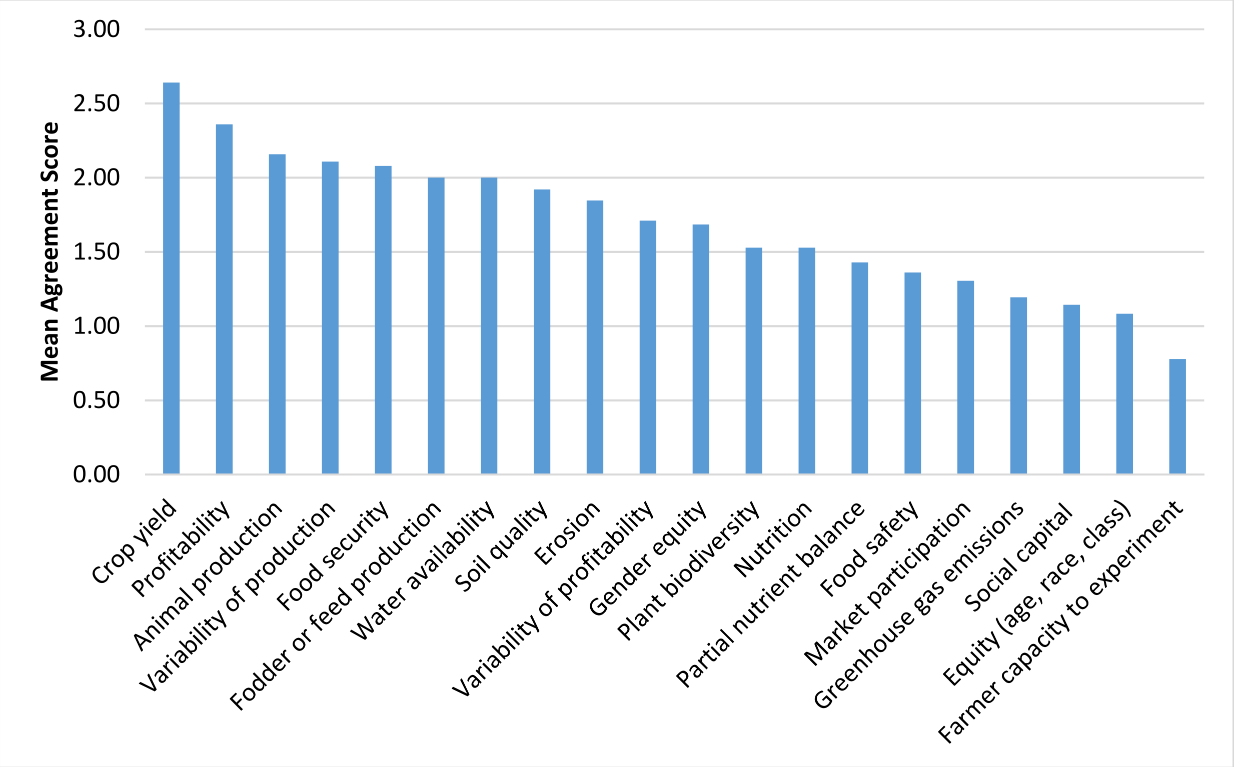
An online survey in 2016 focused on scientists involved in agricultural research for development related to sustainable development. The 44 scientists who participated in the survey identified themselves as follows: 60%, biophysical agricultural scientists; 20%, social scientists; and 20%, interdisciplinary or 'other' scientists. The scientists were asked to indicate the most frequently used indicators by domain. Table 1 provides the indicators most frequently used by the scientists by domain. The human condition and social domain indicators were not measured as frequently by the scientists. The results support what a number of studies have indicated, that there are gaps in indicator selection and use across the domains. One takeaway is that scientists working in sustainable development should make special efforts to consider indicators across all domains in order to overcome potential disciplinary biases. Scientists were provided with a list of indicators and asked to determine on a Likert scale their level of agreement on the criticality of indicators to assessment of sustainable intensification. The results are presented in Figure 3. On average indicators in the social and human condition domain did not receive high of levels of criticality

Table 1. Commonly measured indicators used by SI researchers who participated in an on-line survey [1]

Productivity	**Economic**	**Environment**	**Human Condition**	**Social**
Yield (75%)	Profitability (59%)	Soil Carbon (34%)	Production of nutritious foods (25%)	Gendered rating of innovation (43%)
Yield variability (50%)	Labor requirements (52%)	Crop water availability (30%)	Capacity to experiment (23%)	Gender equity impact (27%)
Crop residue production (45%)	Input use efficiency (48%)	Nutrient Partial Balance (27%)	Dietary diversity (18%)	Conflicts over resources (11%)
Cropping Intensity (35%)	Market Participation (34%)	Soil acidity (27%)	Nutrition awareness (16%)	Equity (youth, ethnic, etc.) (9%)
Yes, better. I changed. _EKS	Yes, better. I changed. _EKS	Yes, better. I changed. _EKS	Yes, better. I changed. _EKS	Yes, better. I changed. _EKS
Animal Production (16%)	Variability of profitability (27%)	Erosion (18%)	Food availability – production (14%)	

Notes: [1] The number in parentheses indicates the percentage of the 44 respondents who measure that indicator

Figure 3: Indicators of sustainable intensification, ranked by average level of agreement (3 = strongly agree and -3 = strongly disagree).



3. How to use the SI assessment framework

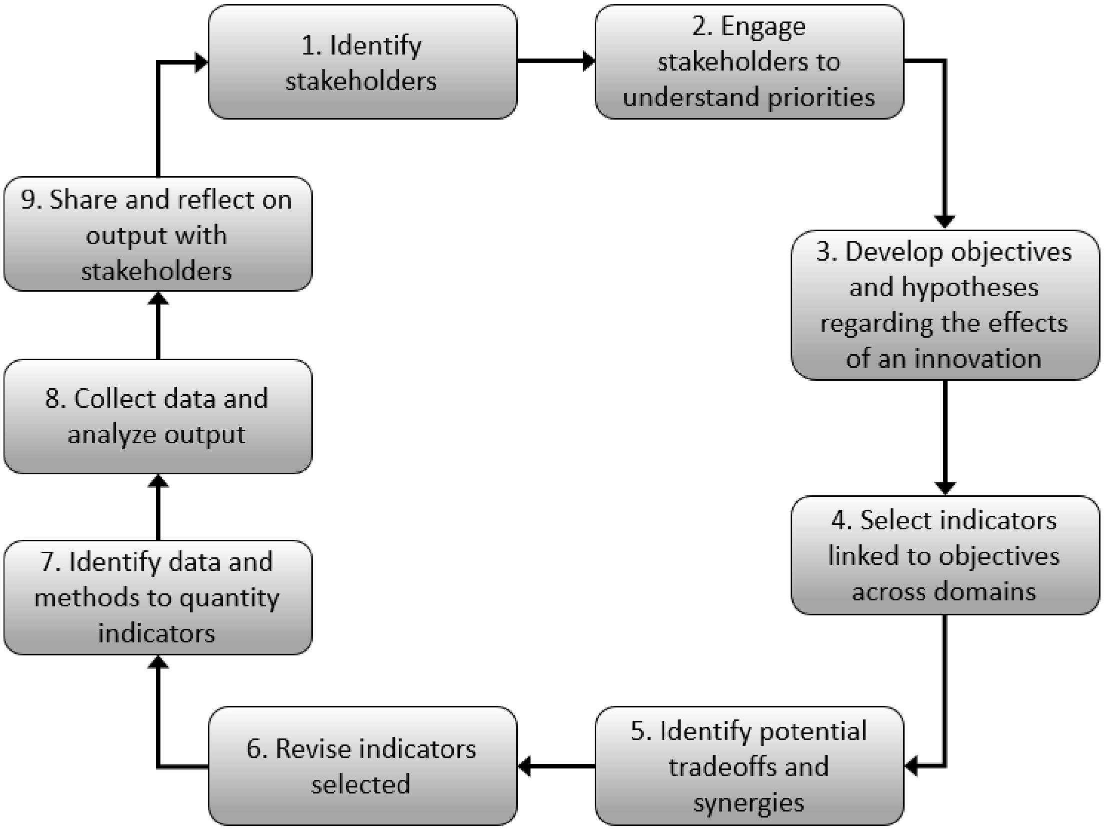
The framework provides documentation for two main processes in indicators assessment: 1) indicator selection process that is objective oriented and 2) identification of tradeoffs and synergies across the five domains.

A strength of this assessment framework is that it provides scientists with tools to examine the process of selecting a balanced set of indicators across domains and an exercise to assess *a priori* the tradeoffs and synergies that may be caused by an innovation.

3.1. Indicator selection

Selection of a core set of indicators is an essential process that determines what will and will not be measured as part of the sustainability assessment. We recommend that selection involve engagement with stakeholders and scientists working in different disciplines. This process will bring divergent views and perspectives but will ensure an improved understanding of different aspects of sustainability and lead to a robust set of indicators. The process of indicator selection should be transparent, well defined, and robust to ensure that it is credible (Latruffe et al., 2016; Dale and Beyeler, 2001). It is critical to select indicators that are balanced to consider all the domains of sustainability and ensure that the relevant stakeholders are involved. In this section, we provide instructions for processes of stakeholder engagement and indicator selection. A structured set of steps for selecting indicators is provided in Figure 4 and described below.

Figure 4. Illustration of process of selecting indicators and presenting the output to stakeholders. (Adapted from Dale et al., 2015).



<u>Step 1. Identify stakeholders:</u> Identify the stakeholders that should be involved in the indicator selection process. The relevant stakeholders are likely to change depending on the stage of development of the innovation. At the very early stages, a small group of researchers may be the key stakeholders as research questions and interventions are discussed. At the other extreme, when an innovation has already been widely promoted then a broad set of stakeholders should participate, such as the end-users (e.g., farmers), those who have promoted it (e.g., extension workers), and other affected groups (e.g., private sector representatives, policy makers, chiefs, farmer cooperatives, etc.). The following steps focus on an innovation at an intermediary stage where stakeholders may include an interdisciplinary group of researchers, extension workers and development professionals working together on research for development.

<u>Step 2. Engage stakeholders to understand priorities:</u> Organize a meeting to engage the stakeholders in the process. Steps 3 through 8 can be accomplished at this meeting if time allows.

-- Introduce the sustainable assessment framework – focusing on the “principles” or the five domains of sustainable intensification that underlie its theoretical framework.

-- Discuss what the domains entail and what indicators are considered in each domain. Ensure that individuals have a copy of the list of indicators per domain from section 4 of this SI assessment guide.

-- Ask stakeholders which components of SI are of primary concern to them and to those around them. Of all the goals embedded within sustainable intensification, which ones are of highest priority for them?

Step 3. Develop objectives and hypotheses regarding the effects of an innovation:

-- At the meeting, discuss the goals, main objectives, and sub-objectives of the innovation or project to the stakeholders.

Develop hypotheses about how the innovation would relate to the objectives and sub-objectives across the five domains.

Step 4. Select indicators linked to objectives across domains: Use Table 2 (the indicator selection worksheet) to guide indicator selection. First, fill out the first column of the table with the main objective of the innovation and the sub-objectives to be achieved per domain.

Next, identify the indicators and metrics that will appropriately measure the objectives and sub-objectives listed.

Finally, choose the data collection methods that will be appropriate to obtain the data needed to estimate these indicators. Although this is revisited in more detail in step 7, consideration of data collection methods early in the process helps ensure that the assessment will be feasible given resource constraints.

Step 5. Identify potential tradeoffs and synergies: Use the tradeoff exercise (section 3.2) in small groups and then share the results. Discuss how the tradeoffs and synergies relate to the goals and objectives prioritized by the stakeholders.

Step 6. Revise indicators selected: Discuss any indicators that may need to be added based on the tradeoffs and synergies identified in Step 5. Table 4 in section 3.2 can be used to help guide that process for a thorough assessment of the given innovation. If possible, members should be provided a chance to discuss their proposed indicators and work towards consensus on which indicators should be included.

Refine the indicator list per domain. Please ensure that you discuss which indicators were not selected and why, especially if there are very few selected for a given domain.

Step 7. Identify data and methods to quantify indicators: Discuss how the innovation will be implemented, who will collect the data, if expertise is available to collect and estimate the indicators, and if the budget is sufficient to collect and estimate the suite of indicators. If there are gaps in expertise and funding, then provisions should be made to fill these gaps or address them within the current team and budget. The process is documented to ensure that there is justification about why an indicator was not included.

Please note that during this step, it also is useful to determine how indicator baselines and targets will be set. (This is not fully addressed here as it is likely to be context specific and thus is beyond the scope of this SI assessment framework guide.)

Step 8. Collect data and analyze output: Specific guidance on data collection for each of the indicators is provided in the SI indicator manual.

Step 9. Share and reflect on output with stakeholders: Once the data has been collected, it is important to share the results of the analysis with stakeholders. If possible, the data should be evaluated for trends over time. Also, the identification of unforeseen tradeoffs is an important consideration when assessing how the innovation performed across stakeholders' goals and objectives. Stakeholders should be encouraged to discuss what implications they see for how the development of the innovation should be adapted.

The cycle can then repeat for further refinement of assessing the innovation. New stakeholders may need to be invited based on what was learned during the previous cycle. New priorities may arise from stakeholders and new hypotheses may be generated based on what was learned from the data collected. The innovation itself may have been adapted, thus altering its effects across the five domains and the indicators needed to assess its performance.

Table 2: Indicator selection worksheet

Define the main project objective:

There are multiple objectives that need to be achieved in the process of sustainable intensification. List the primary and then other objectives (sub-objectives) that you would like to achieve by domain. If possible, list the indicators that might be used to assess each sub-objective and the method of measurement.

Domain	Sub-objectives	Indicators for assessing sub-objectives	Measurement Method	Scale of assessment (field, household, etc.)
Productivity				
Economic				
Environmental				
Human Condition				
Social				

3.2. Assessing tradeoffs and synergies

Smallholder farming systems have complex interactions. For instance, there are purchased and nonpurchased inputs; production can be sold or consumed; landscapes are heterogeneous with poorer and richer households; some fields close to home get more manure and kitchen waste and fields far from home get practically no inputs.

A key principle of systems theory is that changes to one part of a system often have complex and linear or nonlinear effects in other parts of the system. These relationships come in the form of either tradeoffs or synergies (Table 3). As we consider how to design research that provides practical solutions for complex farming systems, we need to carefully consider how innovations may have direct and indirect effects outside the primary focus of our research. Tradeoffs are especially important to SI because there are competing uses of the land, limited resources available, and conflicting objectives across stakeholders that affect sustainability outcomes.

Table 3: Categories of tradeoffs and synergies, with examples

Category	Decision	Example tradeoff	Potential synergy
Within a domain	Land allocation	Legumes vs. maize	Intercropping increases harvest of both
Across domains	Crop residues	Fodder vs. soil fertility	Integrated system with effective manure use
	Level of input use	Production vs. pollution	Fertilizer stimulates improved soil carbon cycling
Across spatial scales	Land use – intensification or extensification	Farm-level profitability can lead to landscape level habitat loss via agricultural expansion	Investing in diversified agriculture expands habitat (land sharing)
Across time	Time preference in soil management	Immediate gain and long-term loss vs. short-term loss and long-term gain	Multipurpose legumes for food, fodder, fuel, income, and/or soil fertility
Across types of farmers	Community grazing norms during dry season	Crop growers control residues vs. herders with free access	Manure from herders enriches soils of farmers

This exercise aims to help explore tradeoffs and synergies related to an agricultural innovation.

Figure 5 provides a list of many of the indicators organized in groups by each domain. This figure can be used as a worksheet to reflect on the synergies and tradeoffs related to an agricultural innovation.

One goal is to generate research questions or identify data gaps that may help implement a systems perspective to the research.

1. Pick one innovation to focus on and describe it in the box at top left.
2. Choose one context where this innovation is being targeted, and focus it for this exercise. Summarize it in the box at top left.
3. CIRCLE the indicator(s) representing the primary direct effects of the innovation in that context. In many cases, there will only be one indicator, but in some cases an innovation will have direct effects on two or more indicators. You may have to add an indicator to the list or revise the wording of the list – for example splitting out maize productivity from legume productivity.
   Note: These should be indicators that are the immediate effect of the innovation – not just the primary goal. For example, an Integrated Pest Management (IPM) project that aims to improve yields would circle pest levels as the immediate effect rather than yields. The indirect effect from reduced pest levels to increased yields would be considered in the next step.

4. Identify the secondary indirect effects that are likely to result from a change in the circled indicators. To represent these indirect effects, draw arrows from the circled indicators. Additionally indicate if an arrow represents a positive synergistic relationship (plus sign) or a tradeoff (negative sign). You can differentiate the importance of these effects by using one plus sign, two plus signs, or three plus signs to show the strength of the synergy and the same for minus signs on tradeoffs. You can also write the word DELAY on the arrow if there is a significant time lapse between cause and effect.
5. Continue the process by considering the next set of effects and continuing to draw arrows. Continuing the IPM example – increased yields could then result in both improved food security and increased profitability. Although almost everything can be related to everything else somehow, we suggest limiting yourself to 10, or at most 15, arrows that seem to be the most important.
6. Be sure to consider possible feedback loops in these relationships. For example, increased yields result in increased residues, which increase soil organic matter, which affect subsequent yields.
7. It is useful to write down assumptions on the diagram, using a pen of a different color. In the example from the previous step it would be worth highlighting the assumption that residues only result in soil organic matter if they are not burned or grazed.

An example of two completed worksheets is presented in the Appendix.

Note: Identifying positive and negative feedback loops is a critical component of systems analysis. Often these feedback loops can be opportunities to intervene and change system performance. With this in mind, it may be valuable to add key factors or drivers to the diagram that influence the performance of the innovation, in addition to the effects of the innovation. These factors or drivers are often the factors researchers need to control for to determine how much the innovation contributes to the indicator of interest. For example, rainfall, soil carbon levels, and germplasm are key factors affecting yield. Some of those factors/drivers may also be indicators relevant to SI (such as soil carbon levels).

Once the tradeoff and synergy exercise has been completed, it is important to consider the implications for indicator selection. Table 4 provides a worksheet for identifying indicators that should be added to the assessment based on the tradeoff and synergy exercise.

Figure 5: Tradeoff and synergy exercise worksheet

Project Name:

Research focus, objective, and scale:

Productivity

Crop productivity

Crop residue productivity

Animal productivity

Variability in production

Input use efficiency

Yield gap

Cropping Intensity

Environment

Vegetative cover

Plant biodiversity

Fuel security

Pest level

Insect biodiversity

Water availability

Water quality

Soil erosion

Soil carbon

Soil chemical quality

Soil physical quality

Greenhouse gas emissions

Pesticide use

Social

Gender equity

Age equity

Equity of marginalized groups

Social cohesion

Collective action

Economic

Profitability

Variability of profitability

Income diversification

Input use intensity

Returns to land, labor and capital

Labor requirement

Poverty rates

Market participation

Market orientation

Human condition

Nutrition

Food safety

Food security

Capacity to experiment

Human health

Draw arrows for connections --------------->

Use +, ++, or +++ to show synergies

Use -, -- , or --- to show tradeoffs

Table 4: Worksheet for adding indicators based on the tradeoff and synergy exercise

There are often tradeoffs and synergies within and across domains that may have important repercussions for sustainable intensification.

Tradeoff or synergy description	Additional indicators for assessing tradeoff or synergy	Measurement method	Scale of assessment

3.3. Operationalizing the assessment

3.3.1. Choosing data collection methods

In choosing data collection methods there are gold standard and alternative measures. Gold standards are methods of measurement and estimating metrics that are considered the most reliable and scientifically sound based on the state of the art. In many cases these methods are widely used in research; for example, in agronomy crop cuts are the gold standards to estimate crop yield. However, there are also several alternative measures that might be more practical for a given situation. Continuing the example of crop yield, an alternative to crop cuts commonly used by social scientists is a recall survey, which is more feasible for large sample sizes.

One key issue for selecting a data collection method is cost. The actual cost for design, collection, and analysis will likely depend on how much of the necessary expertise exists within the research team and how much of it will need to be contracted. Scientists should be careful to select an appropriate sample size (based on the indicator's intended use) so that the data will provide a reliable estimate to assess change of performance of the innovation. If the sample size is too large, there will be unnecessary cost of data collection; if it is too small then the results may not be statistically valid.

For example, assessing the nutritional status of a household or population, anthropometric measurements (the gold standard) could be used, but this may require obtaining a list of households in the area to ensure that an appropriate number of households with children under the age of five is sampled (IFAD, 2017); it also requires a very large sample number. An alternative measure that could provide some indicator of nutrition is a survey on food groups consumed by members of the households. If the study's focus is not mainly on children, using recall survey to estimate dietary diversity as a proxy for nutrition may be a cost-effective alternative approach. Both measurement methods have advantages and limitations and an in-depth explanation for each metric is provided in the SI indicator methods manual.

Some of these alternative measures provide proxy indicators that are related to what is directly measured by the gold standard method. Proxy indicators are those used when data or information is not observed directly (Riley, 2001). Proxy indicators are used in situations where no direct measurements exist or the cost of direct measurement is too high. An example is ownership of assets (such as car, tin roof, or television) as a proxy indicator for income, when there are no reliable data on household income, such as from a detailed survey. A similar example is the use of nitrogen fertilizer as a proxy indicator for nitrous oxide emissions from crop production. For example, some estimates have determined that 1% of nitrogen used on fields is emitted as nitrous oxide emissions (IPCC, 2007). (However, recent field work is showing this may be an overestimate for many areas in the tropics.) Work continues to determine and document available proxy indicators from indirect measurements, modelled output, and remote sensing. Ideally, information that is relatively easy to observe (like household assets or fertilizer use) will be highly correlated with information that is difficult to observe (income and nitrous oxide emissions). Scientists are particularly interested in linking agricultural practices (like soil conservation measures), which are easy to observe, with difficult to measure environmental indicators (like erosion or soil carbon). The feasibility of developing such proxy indicators depends on the exact relationship between the proxy and direct measure and how context-specific that relationship is. In cases where assessing sustainable intensification is cost-prohibitive, developing effective proxies will be vital for parsimonious, comprehensive evaluation of technologies.

3.3.2. Developing an action plan

After selecting indicators and choosing data collection methods, it is important to plan for the detailed implementation of the SI assessment. Table 5 is a worksheet that can be used by a research team to develop those plans. The data collection action plan is organized by method (rows) because often one method (e.g., a survey, an experiment) may be used to measure multiple indicators. For each method, the research team should decide how and when the data will be collected, who will collect the data, and what resources will be needed. In addition, it is imperative that responsible parties for each task be identified.

One goal of this planning process is to ensure adequate resources (e.g., expertise, funds) are available for the SI assessment. Where resources are lacking, adjustments may be made to reduce the cost by exploring low cost methods that may include alternative data collection methods or sharing expertise across projects or departments.

Table 5: Worksheet for planning to implement the assessment

Planning data collection for SI indicators

Data collection method	Indicators	How? (sample size, questions, design)	When?	Who?	Resources needed

3.4. Presenting indicator output

Presenting multiple sustainability indicators to an audience is complex and requires that relative changes among indicators be captured. There is debate on the optimal number of indicators that can be presented on a given visual aid to prevent information overload (Miettinen, 2014) but also to demonstrate relationships between the indicators that portray differences relative to a reference value or target, as well as, to one another. In addition, if temporal aspects are part of the assessment, an illustration is required to show change over time of given indicators and how they move towards or away from sustainability targets. Similarly, if the impact of context is being considered, an illustration of indicator performance for different contexts needs to be presented such as the presentation of results for various locations across a continuum (e.g., high, medium, and low agricultural potential). A number of studies have used various methods to present indicator output results to examine tradeoffs and/or relative changes using biplots, bar charts, and matrices. One of the most widely used approaches are radar charts, with variants including spidergrams, star plots and petal diagrams (Frelat et al., 2016; Zurek et al., 2015; Snapp et al., 2010) that can be used to visualize simultaneously a large number of studies or locations, if detail is minimized (Droppelmann et al., 2016). Indicator output presentation may also require setting and presenting thresholds below and above which a target indicator is red- flagged for either policy or technological intervention. Zurek et al. (2015) propose use of a traffic light system to indicate whether a given indicator is below, above, or near a critical threshold.

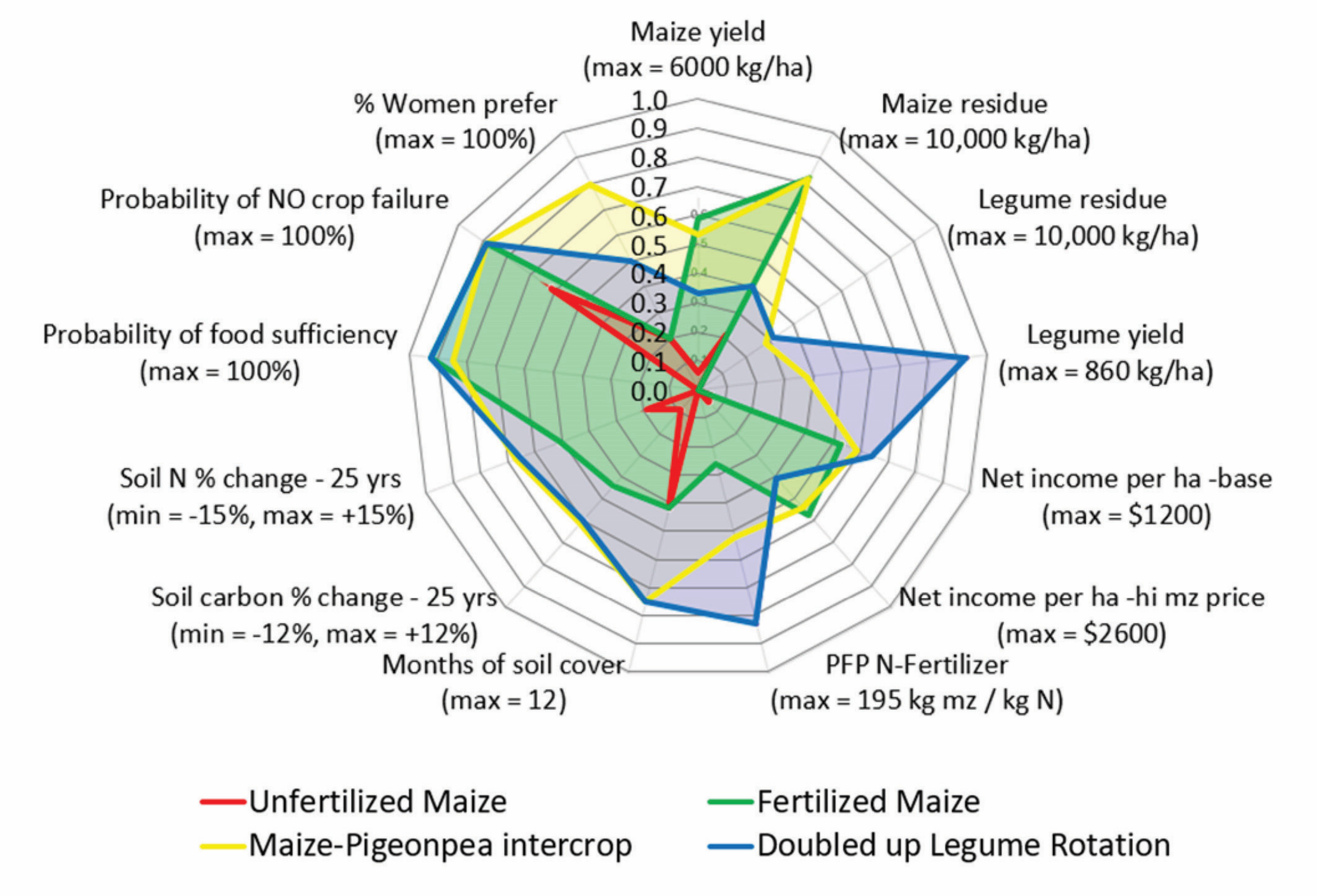
In addition, farm typologies can be used to compare performance across given levels of intensification. Hammond et al. (2017) present farm level performance in Tanzania that categorizes farms by size (hectare) and intensification level (nitrous oxide emissions). In situations with limited data, models have been used to examine tradeoff and synergies using bio-economic models, like FarmDESIGN (Groot et al., 2012). In this study, twoway scatterplots are used to present performance of the household relative to scenarios in which multiple objectives are considered (Groot et. al., 2012; Kanter et al., 2016). Marinus et al. (forthcoming) point out the importance of understanding the variability as well as the mean values for indicators. They also point out the challenge of losing information when aggregating from lower spatial scales. Next we present a few examples for visualizing the output from the SI assessment framework.

3.5 Case study: Applying the SI indicator framework in Malawi

The SI indicator framework was used for guiding research in a USAID-supported Africa RISING project in central Malawi (Snapp et al., submitted). In brief, a participatory action research approach was undertaken to evaluate sustainable intensification trajectories for key household types (Chikowo et al., submitted). Based on literature review and stakeholder engagement, research hypotheses were developed, related to the following overall questions: Can enhanced legume crop diversification increase farming system productivity, profitability, and the environmental sustainability of maize-based cropping? Does this vary with household characteristics and within household? Through an iterative action research process, including a mother-baby trial design (Snapp, 2002) and systematic implementation of panel surveys, research questions were developed and refined. Further information of the current status and research directions can be found at this website: http://globalchangescience.org/eastafricanode.

Goals of the project included increasing legume cover and promoting diversity of legume functional type on maize-based farms. We considered SI indicators in each of the five domains to assess sustainability and consider synergies and tradeoffs. For productivity indicators, we relied on cropping system trial data and modeled Net Primary Productivity (NPP) and yield data associated with maize and grain legumes. Modelling also allowed us to expand the range of weather conditions under consideration and thus the inference zone. For economic indicators, we considered profitability based on two scenarios and the previously mentioned productivity values. The environmental domain indicators relied largely on experiment data and modeled values, based primarily on the extent and diversity of legume cover based on technology attributes (e.g., pigeonpea interventions had greater duration of carbon (C) and nitrogen (N) fixation, due to the nature of pigeonpea growth characteristics (see Snapp et al., 2010). Indicators in the human condition and social domains were primarily based on modeled food supply relative to the demand of typical households, and farmer participatory rating of technologies and experimentation underway with farmers. SI indicators are presented for Golomoti (Figure 6), a lakeshore location and relatively marginal site.

Figure 6: Comparison of maize and maize/legume technologies in Golomoti, Malawi, across all five domains utilizing data from Africa RISING trials, surveys, and crop models.



Source: Snapp et al. (submitted)
Note: PFP N-fertilizer is the partial factor productivity of nitrogen fertilizer

The indicator assessment across SI domains illustrates that when maize-pigeonpea and doubled-up legumes, two diversified technologies, are compared to continuous maize, they have many features associated with greater sustainability. Note the diversity of crops produced, including high-nutrient value pigeonpea and other grain legumes (human condition), and the extended ground cover and nitrogen fixation (environment) associated with these technologies. These are associated with higher soil N and C accrual than either sole maize system (environment). This is suggestive of greater natural resource base amelioration associated with legume diversification, as hypothesized. However, a challenge emerged in that the annualized maize production was only high for fertilized maize. Lack of fertilizer and high legume integration systems were both associated with modest annualized maize yields, which makes it difficult to achieve food security in this environment. The doubled-up legume system with judicious fertilizer had high maize yields in that phase of the rotation but required sufficient land size. So, for farmers with sufficient land size this may be a feasible option and a profitable one that also includes other aspects of sustainability (Figure 6). One assumption highlighted was that legume residues can be utilized in a sustainable manner, but this concept requires more detailed follow-up studies.

4. Indicators by domain and scale

In the next section, we present tables of selected indicators for each domain across scales that are important for consideration by researchers working in research for development. Researchers are not limited to the use of only the listed indicators, but those in the tables provide an initial group of those mostly likely to be important for assessing sustainability of innovations. This list is based on feedback from researchers working on sustainable intensification. The tables also include the methods that can be used for data collection to generate the indicator values (last column to the right of each domain table).

A brief definition of each indicator is provided after the tables of indicators by domain. The SI indicator manual has further details, including a summary of the methods for data collection for each metric, the units of analysis, and the limitations for interpreting each method. This section has the indicators listed by domain as follows: productivity, economics, environmental, human condition, and social domain.

4.1. Productivity Domain (Note: The superscript letters (a,b,c) after each metric refer to the measurement methods in the right-hand column)

PRODUCTIVITY DOMAIN					
Indicator	**Field/plot level metrics**	**Farm level metrics**	**Household (hh) level metrics**	**Community/ Landscape + metrics**	**Measurement method**
Crop productivity	Yield (kg/ha/season) [a,b,c](including tree product/area under crown) Rating of yield [d]	Yield (kg/ha/season) [a,b,c]		Remotely sensed measures of crop productivity (kg/ha/yr) [e]	[a] Yield measurements [b] Recall survey [c] Crop models [d] Farmer evaluation [e] Remote sensing
Crop biomass productivity	Forage production (kg/ha/season) [a,b,c] Residue production (kg/ha/season) [a,b,c] Rating of residue production [d]	Residue production (kg/ha/season) [a,b,c]		Remotely sensed measures of crop biomass (kg /ha/yr) [e]	Same as for yield
Animal productivity	Animal products (amt./animal/yr.) [a,b] Animal by-products (amt./animal/yr.) [a,b] Rating of animal productivity [c]	Animal product per unit land (amt./ha/yr) [a,b] Animal byproduct per unit land (amt. /ha /yr) [a,b] Herd composition [a]	Animal product per hh (product/hh/yr) [a,b] Animal byproduct per hh (product /hh/yr) [a,b]	Net commercial offtake (product /ha/yr) [a]	[a] Recall survey [b] Production measurements [c] Farmer evaluation
Variability of production	Coefficient of variability [a] Probability of low productivity [a]	Coefficient of variability [a] Probability of low productivity [a]	Rating of variability [b] Rating of production risk [b]	Variability of NPP [c]	[a] Productivity over time [b] Farmer evaluation [c] Remote sensing
Input use efficiency	Product per input [a,b]				[a] Survey and productivity measures [b] Models
Yield gap	Yield gap per crop (kg/ha/season)	Yield gap per crop (kg/ha/season)		Yield gap (kg/ha/season)	Same as Yield
Cropping intensity	# of cropping seasons per year [a]			#of cropping seasons per year	[a] Recall survey
Post-harvest losses			% harvest lost [a,b]		[a] Direct measurements [b] Survey

Description of Productivity Indicators and metrics

Crop productivity

Crop productivity is a measure of the total sum of annual plant production per unit area per unit time, which is also known as net primary productivity (NPP). Crop productivity can be partitioned by tissue type (e.g., grain, leaves, stems) based on how the plant is used. The unused portions of crops are often referred to as crop residues, which is the next indicator.

Yield

Yield is the most commonly used measure of crop productivity and equals the production for a given land area, generally measured at the field scale.

Farmer rating of Yield

In cases where farmers have tried a new technology but on-farm yield measurements and detailed surveys are not available, it is possible to obtain farmers' qualitative evaluations of yields, for example by asking them to rank or rate yields for varieties or management practices.

Remotely sensed crop productivity

Remote sensing can be used to measure vegetation indices and productivity across landscapes through spectral signals and multi-input algorithms. Inferences can be made about yield but require extensive ground-truthing of the information.

Crop biomass productivity

For sustainable intensification, the productivity of the land needs to be assessed in terms of all that is produced (not just grain yield). This is especially important where vegetative biomass is used for fodder or returned to the soil.

Residue production

Crop residue productivity and vegetative tissues are measures of the 'non-grain' biomass from plant production on a known area.

Remotely sensed measures of crop biomass

Remote sensing can detect optically measured vegetation indices, which can then be used to estimate crop biomass production through Net Primary Productivity (NPP). A portion of this NPP at crop maturity can be estimated as crop grain yield, and the balance is the crop's non-grain biomass production.

Animal productivity

Animal productivity is the total sum of products and services from animals. In the context of sustainable intensification, the efficiency of that production is important, for example the amount of land required to produce the animals and also the area needed to produce feed for the animals. Though there are dozens of species of animals raised for various purposes, the great majority of products are derived from the primary livestock species – cattle, pig, sheep, goats, and poultry.

Variability of production

The risk of low yields or low animal productivity may be even more important in some contexts than the average productivity. Quantifying the variability of productivity over time and space is an important measure of this risk.

Input use efficiency

The concept of efficiency focuses on avoiding or reducing wastage of a resource. Input efficiency is supposed to increase the performance of the system and minimize losses to the environment. It is important to note that input use efficiency should not be used by itself as an indicator but should be used along with yields. for the reason for this is that input use efficiency is often highest at low yields. Farmers are striving to increase yields—not just input use efficiency.

Yield gap

Yield gap is measured as the difference between yield potential or water-limited yield and actual yield. Actual yield may be obtained through surveys or crop cuts; potential yields may be obtained from models and existing secondary sources, such as maps and experimental data.

Cropping intensity

Cropping intensity is defined as the number of crops a farmer grows in a given agricultural year on the same field (Raut et al., 2011) and is another means for intensification of production from the same plot of land. Cropping intensity is likely to be important to monitor where the intervention affects the likelihood of irrigation or planting during short season rains.

Post-harvest losses

Post-harvest losses such as from insects, mold and rodents can be substantial, thus reducing any benefits from increased productivity.

4.2. Economic Domain (Note: The superscript letters (a,b,c) after each metric refer to the measurement methods in the right-hand column)

ECONOMIC DOMAIN

Indicator	Field/plot level metrics	Farm level metrics	Household level metrics	Community/ Landscape + metrics	Measurement method
Profitability	Net income [a] ($/crop/ha/season) Gross margin [a]	Net income [a,c] (total net income for all farm activities) Gross margin [a]	Net income [a,c] (total net income for all farm activities)	Contribution to regional or national GDP [b]	[a] Survey [b] Regional and national statistics [c] Participatory evaluation
Variability of profitability	Coefficient of variability of net income [a] Probability of low profitability [a,b]	Coefficient of variability of net income [a] Probability of low profitability [a,b]	Coefficient of variability of net income [a] Probability of low profitability [a,b]		[a] Survey [b] Farmer evaluation
Income diversification	N/A	Diversification index [a]	Diversification index [a] Number of income sources[a]		[a] Survey
Returns to land, labor, and inputs	Returns [a] (monetary value of output/input used)	Returns [a] (monetary value of output/input used)	Returns [a] (monetary value of output/input)		[a] Survey and productivity measurements
Input use intensity	Input per ha [a]	Input per ha [a]	Input per ha [a]		[a] Survey
Labor requirement	Labor requirement (hours/ha) [a,b] Farmer rating of labor [c]	Labor requirement (hours/ha) [a,b] Farmer rating of labor [c]	Labor requirement (hours) [a,b] Farmer rating of labor [c]		[a] Recall survey [b] Direct observation [c] Farmer evaluation
Poverty	N/A	N/A	Asset index [a] Per capita hh consumption expenditure [a] Wealth categorization [b]	Poverty headcount ratio [a] Asset wealth categorization [b]	[a] Survey [b] Participatory exercise
Market participation	N/A	N/A	% production sold [a]	Total sales [a]	[a] Survey
Market orientation	N/A	N/A	% land in cash crops [a] Market orientation index [a]		[a] Survey

Description of Economic Indicators and Metrics

Profitability

Profitability measures the gains from the agricultural enterprise using data on revenues and expenses.

Net income

One metric for profitability is net income that is derived from incomes and costs of production including fixed and overhead costs.

Gross margin

Gross margin is an alternative profitability that calculates profits from incomes and variable costs of production (ignoring the fixed and overhead costs as sunk costs).

Variability of profitability

Variability of profitability is an important metric because it provides a measure of variation from the mean that can be attributed to either the production or consumption side and reflected changes in markets, climate and other factors.

Income diversification

Income diversification[1] is a measure of household's income generation from multiple sources. This livelihood strategy is observed among households and may insure food security in times of production or weather risk.

Returns to land, labor, and capital

Returns to land, labor, and capital is related to the profitability. These metrics enable one to assess the opportunity cost of the resource, which may be even more influential than overall profitability when the resource is the most limiting factor. Comparing the rate of return to a resource with a new technology is important for this analysis.

Input use intensity

Input use intensity measures the amount of a given input used per unit of area (e.g., kg nitrogen input/ha; liter of irrigation water/ha). Input intensity provides a measure of assessing two important issues: 1) whether a given input is used and 2) amount used per unit area.

Labor requirement

Labor requirement focuses on the amount of human labor needed for agricultural production. This measure includes a count of number of work hours, cost of labor, and seasonal supply and demand for labor.

Poverty

Poverty is estimated as a welfare measure to ascertain the minimum level of income that is adequate to sustain a livelihood.

Per capita household consumption expenditure

Per capita household consumption expenditure is the total household consumption expenditure on food and nonfood items that is adjusted by the number of household members. This provides a measure of how much a member of the household consumes or spends per day.

Asset index

---

[1] Income diversification is not synonymous with livelihood diversification

An asset index is a poverty measure used to categorize household's wellbeing using data from their ownership of a list of assets.

Asset wealth categorization

Wealth categorization is a commonly used participatory exercise where key informants develop categories of wealth or well-being and then rank or group the households in their community.

Market participation

Market participation examines whether a farmer sells his or her agricultural commodities or buys inputs from the market.

Market orientation

Market orientation is defined as production with an intention to sell to the market. Market orientation examines the agricultural production by the household that is destined for the market (Gebremedhin and Jaleta, 2010). This indicator is used to distinguish between market participation of the household (indicator described previously) versus degree of commercialization. Commercialization indices have focused on the ratio of the total value of agricultural production sold to that produced (Carletto et al., 2017). Note that the market orientation index presented in this document also considers the share of land allocated to the crop by the household.

4.3. Environment Domain **Part 1: Biodiversity and water** (The superscript letters ([a,b,c]) after each metric refer to the methods in the right-hand column)

ENVIRONMENT DOMAIN (Part 1: Biodiversity and water)					
Indicator	**Field/plot level metrics**	**Farm level metrics**	**Household level metrics**	**Community/ Landscape + metrics**	**Measurement method**
Vegetative cover	% Vegetative cover by type (tree, shrub, grass, invasive) [a,b] % Burned land [a,b] % Bare land [a,b]	% Vegetative cover by type [a,b] % Burned land [a,b]	N/A	% Vegetative cover by type [c] % Burned land [c] % Bare land [c]	[a] Quadrats, transects, or visual estimate of cover [b] Participatory exercise [c] Remote sensing
Plant biodiversity	Alpha Diversity Index [a,b] # Species or varieties [a,b]	Beta Diversity Index [a,b] # Species/varieties [a,b]	N/A	Gamma Diversity Index [a,b] % Natural habitat [c]	[a] Vegetation sample [b] Transects [c] Satellite images
Pest levels	Pest abundance and severity by type [a,b]				[a] Seasonal transects [b] Traps
Insect biodiversity	# Pollinators [a,b,c] Diversity index [a,b,c] # Beneficial insects [a,b,c]			# Pollinators [a,b,c] Diversity index [a,b,c] # Beneficial insects [a,b,]	[a] Traps [b] Direct observation [c] Seasonal transects
Fuel availability	Fuel biomass (e.g., wood, residues) produced on plot [a,b,c]	Fuel biomass (e.g., wood, residues) produced on farm [a,b,c]	% hh fuel by type (wood, charcoal) [a,b] # months energy security [a,b] Fuel collection time [a] % of hh fuel from farm [a]	% of fuel from off-farm [a,b] Spatial arrangement of fuel sources [b] % of hhs with energy security [a]	[a] Survey [b] Participatory exercise [c] Biomass measurement
Water availability	Irrigation use by crop [b] Soil moisture [a,b,c] % of plants wilting [b,d] Infiltration rate [a,d]	Irrigation use [b] % of fields wilting [b, d]	Water sufficiency [b] Water security index [b] Water security rating [d]	% of irrigated land [b, e] % flow not diverted [f] % hh with sufficient water [b]	[a] Field and lab tests [b] Survey [c] Crop models [d] Participatory exercise [e] Remote sensing [f] Stream sampling
Water quality			Rating of water quality [b]	% Clean water sources [a,b] % Pop. w/clean water [b] Salinity [a] Phosphate /nitrate /pathogenic microbe concentration (mg/L) [a]	[a] Water sampling [b] Household survey

4.3. Environment domain part 2: Soil and pollution (Note: The superscript letters ([a,b,c]) after each metric refer to the methods in the right-hand column)

ENVIRONMENT DOMAIN (Part 2: Soil and pollution)					
Indicator	**Field/plot level metrics**	**Farm level metrics**	**Household**	**Community/Landscape + metrics**	**Measurement method**
Erosion	Soil loss (tons/ha/yr) [a,b,c] Rating of erosion [a,d]		N/A	Sediment load (mg/L) [e] Erosion (tons/ha/yr) [b]	[a] Direct measurement [b] Models [c] Survey [d] Participatory exercise [e] Stream sampling
Soil biology	Total carbon (% or Mg/ha) [a,c] Labile or 'active' carbon (POXC) [a] and/or CO_2 mineralization [c] Partial carbon budget [b,c] Earthworms [d]	Relative measures of plot-level metrics across farm fields Total Carbon[e]	N/A	N/A	[a] Soil test [b] Survey [c] Modeling [d] Farmer and plant assessment [e] Remote sensing
Soil chemical quality	Soil pH (acidity) [a] Electrical conductivity [a] Soil nutrient levels [a] Nutrient partial balance [b] Biological nitrogen fixation [a]	Nutrient partial balance [b] Biological nitrogen fixation [a]	N/A	Nutrient partial balance [a,b]	[a] Soil tests [b] Survey and lookup tables
Soil physical quality	Aggregate stability [a] Bulk density [a] Water holding capacity [a] Infiltration rate [a]				[a] Soil tests
GHG emissions	CO_2 equivalent emitted per ha [ab]	CO_2 equivalent emitted per ha [ab]	N/A	CO_2 equivalent emitted per ha [ab]	[a] Lookup tables by activity or input [b]Models
Pesticide use	Active ingredient applied per ha [a]	Active ingredient applied per ha [a]	N/A	Pesticides concentration in water [b]	[a] Agricultural survey [b] Water tests

Description of Environmental Indicators and metrics

Vegetative cover

Vegetative cover is the percentage of ground area that is covered by vegetation (canopy cover) that may be in natural landscapes or agricultural areas. It can be from canopy variety of sources, including field crops, cover crops, and trees.

Plant biodiversity

Biodiversity is examined in terms of the number of plant species and is focused on the richness (number of species in a sample) and evenness (the distribution of species with in the sample). It can be measured at field, farm, and landscape scales.

Pest levels

This indicator focuses on examining the pest levels that may have an effect on agricultural productivity.

Insect biodiversity

This indicator focuses on number of species of pollinators and other beneficial insects plus the richness of these species (evenness and abundance).

Insect biodiversity

Insect biodiversity can be used as an indicator of health of the ecosystem and for assessing how changes in the population may affect agricultural production or the sustainability of productivity. Insects are crucial in the functioning of terrestrial ecosystems and can have both beneficial and negative consequences for agriculture. In agricultural production, insects play an important role in plant reproduction and also may help maintain a balance of pest population (Kevan et al., 1999).

Fuel availability

Fuel is very important for livelihood in developing country agricultural landscapes. In many cases in rural areas, biomass energy is the dominant form of fuel and includes wood, charcoal, crop residues, manure.

Water availability

The water availability indicator assesses if there is sufficient water for both plant growth and human consumption. At the field level, the indicator examines the presence of adequate moisture in the growing season and also excess water that would be detrimental to plant growth.

Water quality

Water quality indicators are meant to provide information on the fit for use of water for crop, livestock, and human consumption.

Erosion

Soil erosion examines movement of soil from a field; it can be moved by water or wind and is often one of the main processes leading to soil degradation. It is difficult to quantify and so is often ranked or evaluated from a qualitative assessment.

Soil biology

Soil biology has direct impacts on plant growth and indirect effects as it mediates soil chemical and physical properties and processes. The biological quality of the soil may be observed through macro-fauna like earthworms and through microscopic soil fauna and its diversity. Soil carbon is a critical indicator of soil quality that is a direct result of soil biological activity and is important for soil moisture and nutrient retention and livelihood of soil microbes (Doran and Jones, 1996; Reeves, 1997).

Soil chemical quality

Soil chemical quality refers to both the acidity of the soil and the nutrient contents in the soil. Nutrient partial balance is a useful and parsimonious metric that examines nutrient output and inputs for a given area (output minus input to the soil).

Soil physical properties

Soil provides the physical medium in which plants grow and roots penetrate. The physical structure of the soil allows the infiltration and storage of water, the movement of air into and out of the soil, all of which are critical to maintaining a physical environment in which the plant grows. Physical factors that are important to maintaining the soil structure and the processes related to structure include aggregate stability and a light, non-compacted and friable soil.

Greenhouse gas (GHG) emissions

Agriculture is a major greenhouse gas (GHG) emitter and therefore has implications for climate change (IPCC, 2007; Vermeulen et al., 2012). The main sources of GHG from agriculture include fertilizer use that leads to nitrous oxide emissions from the soil and manures, methane emissions from ruminants and rice production, and land use change (IPCC, 2014; Gustafson et al., 2016). GHG emissions from most smallholder agriculture in the tropics is relatively small compared to large scale agriculture in the tropics and temperate zone

Pesticide Use

Pesticide use focuses on the risk and environmental impact of pesticides on water quality and death of species.

4.4. Human Condition Domain (Note: The superscript letters (^{a,b,c}) after each metric refer to the methods in the right-hand column)

HUMAN CONDITION DOMAIN					
Indicator	**Field/plot level**	**Farm level**	**Household level metrics**	**Community/ Landscape + metrics**	**Measurement method**
Nutrition	Protein production (g/ha) ^{a,b} Micronutrient production (g/ha) ^{a,b}	Total protein production (g/ha) ^{a,b} Total micronutrient production (g/ha) ^{a,b} Availability of diverse food crops ^{a}	Access to nutritious foods ^{a} Dietary diversity ^{a} Food consumption score ^{a} Nutritional status (underweight, stunting, wasting) ^{c} Uptake of essential nutrients ^{d}	Market/landscape supply of diverse food ^{a,e} Dietary diversity ^{a} Rate of underweight, stunting and wasting ^{c} Average birthweight ^{c}	^{a} Survey ^{b} Look up tables ^{c} Anthropometric measurements ^{d} Blood tests ^{e} Participatory mapping
Food security	Food production (Calories/ha/year) ^{a,b}	Food production (Calories/ha/year) ^{a,b}	Food availability ^{a} Food accessibility ^{a} Food utilization ^{a} Food security composite index ^{a} Months of food insecurity ^{a} Rating of food security ^{c}	Total food production ^{a} % population food secure	^{a} Survey ^{b} Look up tables ^{c} Participatory assessment
Food safety			Biological contaminants Mycotoxins (toxicity units per gram) ^{a} Chemical contaminants Pesticide contamination ^{a,b} Heavy metal contamination ^{a} Physical contaminants Quantity of rocks per ton of grain ^{c}	Incidence of food borne diseases (E.coli, Salmonella, Campylobacter)	^{a} Laboratory testing ^{b} Health center data ^{c} Sorting and weighing
Human health				Incidence of zoonotic diseases ^{a} Incidence of vector borne diseases ^{a}	^{a} Health center data
Capacity to experiment			# of new practices being tested ^{a,b}	% of farmers experimenting ^{a,b}	^{a} Individual survey ^{b} Focus group

Description of Human Condition Indicators

Nutrition

Nutrition plays an important role in sustainable agriculture as both an output and input. Good nutrition may improve the productivity of farmer, and production of nutritious food may improve nutritional status through consumption of own production or increased incomes, enabling household to buy nutritious foods from the market (IFPRI, 2014). Dietary indicators focus mainly on women and young children who are the groups most vulnerable to malnutrition.

Micronutrient production

This indicator is important in areas or populations where there is a nutrient deficiency and where the innovation being assessed is likely to affect the availability of that nutrient (Burchi et al., 2011).

Nutrition awareness

Nutrition awareness is used to indicate the percentage of population that has received information on how to improve production, preparation, and consumption of nutritious foods

Food security

Measuring food security has been a challenge but the concept has been defined as a state in which “all people at all times have the both the physical and economic access to sufficient food to meet dietary needs for a productive and healthy life” (USAID, 1992). Food security has evolved from food availability to examining nutritional capabilities of the food that is produced or accessible to a household (Burchi and De Muro, 2016). The Food and Agricultural Organization has defined the three main pillars of food security as food availability, food access, and food utilization.

Food availability

Food availability is defined as the availability of sufficient quantities of food of appropriate quality supplied through domestic production or importation. The food availability indicator measures the amount of food produced by the household and the amount that is sold and purchased per capita to come up with an estimate of calories and nutrients available per capita (Remans et al., 2013).

Food accessibility

Food access may be defined as the ability to acquire sufficient quality and quantity of food to meet the nutritional requirements of individuals within the household for a productive life (Swindale and Bilisky, 2006). Food access indicator tends to focus on the economic aspect and examines the ability of a household or a person to purchase food.

Food utilization

Food utilization refers to an individual’s capacity to make use of food for a productive life (Swindale and Bilisky, 2006). Food utilization focuses on the diversity of the food consumed in the households and assesses the food groups available, calories consumed from staples versus nonstaples, and an evaluation of protein and micronutrient composition of food consumed.

Months of food insecurity

The months of food insecurity is a metric used to assess the frequency of household food insecurity and is the months in which these incidents occur.

Food safety

Food safety is a key issue that ensures a fit for consumption and quality of the food. The food safety metrics can be grouped by the type of contaminant: biological, chemical, or physical. Mycotoxins are often cited as potential biological contaminant that affect food safety and may lead to chronic illness if excessive mycotoxins exist in a given product (Milicevic et al., 2010). Pesticide and heavy metal contamination are chemical contaminants that also require additional attention in food safety and quality. Physical contaminants include rocks or other inedible objects mixed with the harvest.

Human health

Human health may be at risk from agricultural activities due to interaction with animals or through vector borne diseases that affect both animals and humans. Some of the interventions may directly or indirectly alter these vectors, such as an increase in malaria due to irrigation.

Capacity to experiment

Capacity to experiment is the ability of the household to test innovations or management practices that are new to them.

4.5. Social Domain: (Note: The superscript letters (a,b,c) after each metric refer to the methods in the right-hand column)

SOCIAL DOMAIN			
Indicator	**Field, farm, and household level metrics**	**Community/Landscape + metrics**	**Measurement method**
Gender equity	Resources: Land access by gender [a-d] Livestock ownership by gender [a-d] Capacity: Access to information [a-d] Agency: Time allocation by gender [a-d] Management control by gender [a-d] Market participation by gender [a-d] Achievements: Income by gender [a-d] Nutrition/Food security by gender [a-d] Health status by gender [a-d] Cross-cutting: Rating of technologies by gender [b]	Women Empowerment in Agriculture Index [a, d] (measures absolute and relative empowerment across five domains: production, resources, income, leadership, and time)	[a] Individual survey [b] Participatory evaluation [c] Focus group discussions [d] Household survey
Equity (generally)	Access to resources (land and livestock ownership) [a-d] Capacity (access to information) [a-d] Agency (leadership roles) [a-d] Achievements (income, nutrition, food security, health, well-being) [a-d] Rating of technologies by group [a-d]	Variability and distributions resources, agency, and achievements [a-d]	[a] Key informant interviews [b] Participatory evaluation [c] Focus group discussions [d] Household survey
Social cohesion	Participation in community activities [a,b,c] Level and reliability of social support [a,b,c] Family cohesion [a,b,c]	Social groups [c] Participation in social groups [a,b,c] Incidence of social support [a,b,c]	[a] Household survey [b] Focus group discussions [c] Key informant interviews
Collective Action	Participation in a collective action group [a]	Collective action groups [a,b] Capacity of groups [a,b] Incidence of conflicts related to collective action [a,b] Effectiveness of conflict resolution measures [a,b]	[a] Household survey [b] Key informant interviews

Description of social indicators and metrics:

Gender equity

Drawing from the gender empowerment literature, we developed a conceptual framework for gender equity in agriculture that is detailed in Appendix 2. Following Hemminger et al. (2014), we use the empowerment framework from Kabeer (1999) to categorize gender equity metrics as follows:

- Resources – Metrics that measure differential access to resources for agriculture.
- Agency – Metrics that measure differential levels of control over resources.
- Achievements – Metrics that measure gendered differences in realizing various benefits from agriculture.

Land access by gender and livestock ownership by gender

Land and livestock are critical resources for production and differences in ownership across groups can reveal systemic inequities in how these resources have been allocated. Other key resources may be of interest in specific locations, such as irrigation water, credit, or machinery.

Time allocation by gender

This metric can be used to assess gender equity through the quantitative measurement of differences in time spent on various tasks. Division of labor by gender is not inherently negative when it allows for specialization. However, time allocation differentials can reveal gender inequities by comparing amounts of leisure time for each gender or comparing time spent on the least desirable or most taxing tasks. Also, this information can be combined with other metrics in the agency and resource categories to assess who benefits from how the time is spent.

Management control by gender

This metric aims to capture differences in decision-making power between men and women. To be operationalized, it is necessary to choose the most important decision in the given context. For cropping systems, one could measure the land area where women report that they are the primary decision maker about crop management (solely as well as jointly) compared to the land area where men report being the primary decision maker (solely as well as jointly).

Market participation by gender

Within a household, this metric can be a comparison of who markets which products. At the landscape + scale the incidence of men and women participating in the market can be compared.

Income by gender

Income is both a resource for and an achievement of women's empowerment. When considering it as a resource the focus is on access to finances and can be measured by asking who participates in the decisions to buy items such as agricultural inputs and daily goods. When considering income as an achievement, it can be measured based on net income from crops or animals controlled by each gender. If detailed time allocation has been collected, then returns to labor can be calculated and compared across genders.

Nutrition, food security, and health by gender

These metrics simply use disaggregated data from the human condition domain to compare achievements across gender.

Ratings of technologies by gender

Technologies that are used at the farm and field scale may be evaluated differently by men and women. The data collection happens at the household scale so the gendered rating is listed at the household level.

Women Empowerment in Agriculture Index

This index is calculated by following a specific data collection methodology where male and female responses are compared. This survey process may be too demanding for many programs, but it does provide a great deal of information about the various facets of empowerment at the community or regional scale. The Women Empowerment in Agriculture Index (WEAI) has five domains for the empowerment sub-index: production, resources, income, leadership, and time (Alkire et al., 2013).

Equity (generally)

This indicator draws on the conceptual framework for gender equity indicators described previously in this section. Often there may be a reason to focus on equity concerns across specific groups such as by livelihood strategy (crop growers, livestock herders, fishermen) or by ethnicity. In other cases, the focus may be on comparing how the technology performs across wealth and age groups. Usually these comparisons can be done across households which reduces much of the complexity from intrahousehold decision making that is essential for gender equity analysis.

Access to resources

These metrics are concerned with fair allocation of physical resources. They measure differential access to resources for agriculture.

Capacity

These metrics are concerned with fair allocation of information and training resources. They measure differential access to information about markets or agricultural practices.

Agency

These metrics are concerned with fair procedures. They measure differential levels of control over resources.

Achievements

These metrics are concerned with fair exchange. They measure differences in how various benefits from agriculture are realized.

Social cohesion

Direct indicators of social cohesion are “membership rates of organizations and civic participation” and “levels of trust” (in other people).Proxies for this metric are income distribution and ethnic heterogeneity (Easterly et al., 2006). Social cohesion is seen as society level issue while social capital is micro-level focused.

Collective action

Collective action is common in many areas for managing natural resources (e.g., irrigation, water, fisheries) and also can be found in areas such as agriculture for marketing, processing, procuring inputs.

References

Alkire, S., Meinzen9-Dick, R., Peterman, A., Quisumbing, A., Seymour, G. and Vaz, A., 2013. The women's empowerment in agriculture index. World Development, 52, pp.71-91.

American Association for the Advancement of Science (AAAS). 2015. Workshop on Sustainable Intensification Indicators. Notes from AAAS Annual Meeting Workshop San Jose, California. February 12, 2015

Bockstaller, C. and Girardin, P. 2003. How to validate environmental indicators. *Agricultural Systems*, 76(2), pp.639-653.

Burchi, F., Fanzo, J. and Frison, E., 2011. The role of food and nutrition system approaches in tackling hidden hunger. *International journal of environmental research and public health*, 8(2), pp.358-373.

Burchi, F. and P. De Muro. 2016. From food availability to nutritional capabilities: Advancing food security analysis. *Food Policy*, 60, 10-19.

Carletto, C., Corral, P. and Guelfi, A. 2017. Agricultural commercialization and nutrition revisited: Empirical evidence from three African countries. *Food Policy*, 67, pp.106-118.

Chikowo, R., Snapp, S. Grabowski, P., Odhong, J., Hoeschle-Zeledon, I. Bekunda, M. Submitted. Farm typologies use in targeting sustainable intensification.

Conway, G.R., 1994. Sustainability in agricultural development: trade-offs with productivity, stability and equitability. *J. Farming Systems Res. Ext*., 4(2), pp. 1-14.

Cook, S., Silici, L., Adolph, B., Walker, S., 2015. Sustainable intensification revisited, IIED Issue paper. IIED

Dahlin, S., Rusinamhodzi, L. 2014. Review of interventions and technologies for sustainable intensification of smallholder crop production in sub-humid sub-Saharan Africa (Report No. 2014:5). Uppsala.

Dale, V.H., Efroymson, R.A., Kline, K.L. and Davitt, M.S., 2015. A framework for selecting indicators of bioenergy sustainability. Biofuels, Bioproducts and Biorefining, 9(4), pp.435-446.

Dale, V.H. and Beyeler, S.C., 2001. Challenges in the development and use of ecological indicators. *Ecological Indicators*, 1(1), pp.3-10.

Doran, J.W. and Jones, A.J., 1996. *Methods for assessing soil quality.* Soil Sci Soc of Am (SSSA), Madison, WI, 410 p.

Droppelmann, K.J., S.S. Snapp and S. Waddington. 2016. Sustainable intensification technologies for smallholder maize-based farming systems in Sub-Saharan Africa *Food Security,* Ms. in press.

Easterly, W., Ritzen, J., and Woolcock, M. 2006. Social cohesion, institutions, and growth. *Economics & Politics*, *18*(2), 103-120.

Eele, G., 1994. Indicators for food security and nutrition monitoring: A review of experience from Southern Africa. *Food Policy*, *19*(3), pp.314-328.

Frelat, R., Lopez-Ridaura, S., Giller, K.E., Herrero, M., Douxchamps, S., Djurfeldt, A.A., Erenstein, O., Henderson, B., Kassie, M., Paul, B.K. and Rigolot, C., 2016. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. Proceedings of the National Academy of Sciences, 113(2), pp.458-463.

Garnett, T., Godfray, C., 2012. Sustainable intensification in agriculture; navigating a course through competing food system priorities. Food Climate Research Network and the Oxford Martin Programme on the Future of Food, University of Oxford, UK 51.

Gebremedhin, B. and M. Jaleta. 2010, September. Commercialization of smallholders: Is market participation enough. In Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa.

Glover, J. 2016. An Indicator Framework for Sustainable Intensification Assessment. Presentation at "Meeting on Indicators for the Sustainable Intensification of Agriculture", Michigan State University, December 14, 2016.

Gómez-Limón, J.A. and Sanchez-Fernandez, G., 2010. Empirical evaluation of agricultural sustainability using composite indicators. Ecological economics, 69(5), pp.1062-1075.

Godfray, H.C.J., 2015. The debate over sustainable intensification. Food Sec. 7, 199–208. doi:10.1007/s12571-015-0424-2

Grabowski, P., Musumba, M. Palm, C. and Snapp, S. Forthcoming. Sustainable agricultural intensification and measuring the immeasurable: Do we have a choice? In: Bell, S. and Morse, S. (eds.) *Sustainability Indicators*. Routledge.

Groot, J.C., Oomen, G.J. and Rossing, W.A., 2012. Multi-objective optimization and design of farming systems. Agricultural Systems, 110, pp.63-77.

Gustafson, D., Gutman, A., Leet, W., Drewnowski, A., Fanzo, J. and Ingram, J., 2016. Seven Food System Metrics of Sustainable Nutrition Security. Sustainability, 8(3), p.196.

Haileslassie, A., Craufurd, P., Thiagarajah, R., Kumar, S., Whitbread, A., Rathor, A., Blummel, M., Ericsson, P. and Kakumanu, K.R., 2016. Empirical evaluation of sustainability of divergent farms in the dryland farming systems of India. Ecological Indicators, 60, pp.710-723.

Hammond, J., Fraval, S., van Etten, J., Suchini, J.G., Mercado, L., Pagella, T., Frelat, R., Lannerstad, M., Douxchamps, S., Teufel, N. and Valbuena, D., 2017. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. Agricultural Systems, 151, pp.225-233.

Hemminger, K., B. Bock, J. Groot, M. Michalscheck, and C. Timler. 2014. Towards integrated assessment of gender relations in farming systems analysis. MSc thesis, Wageningen University (also presented as a poster with the co-authors at Tropentag 2014).

IFAD, 2017. Tools for conducting An Impact Survey. https://www.ifad.org/documents/10180/b476b436-0034-4537-95d5-e65b9bf8b7af

International Food Policy Research Institute (IFPRI). 2014. Global Nutrition Report 2014: Actions and Accountability to Accelerate the World's Progress on Nutrition. Washington, DC.

Intergovernmental Panel on Climate Change (IPCC). 2007. Fourth Assessment Report. Working I Report: The Physical Science Basis. Cambridge University Press, Cambridge.

Intergovernmental Panel on Climate Change (IPCC) 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability: Summary for Policymakers; IPCC: Geneva, Switzerland, 2014.

ISPC. 2014. Data, metrics and monitoring in CGIAR – a strategic study. Rome, Italy, CGIAR Independent Science and Partnership Council (ISPC). 88 pp.

Kabeer, N., United Nations and United Nations Research Institute for Social Development, 1999. The conditions and consequences of choice: reflections on the measurement of women's empowerment (Vol. 108, pp. 1-58). Geneva: UNRISD.

Kanter, D.R., Musumba, M., Wood, S.L., Palm, C., Antle, J., Balvanera, P., Dale, V.H., Havlik, P., Kline, K.L., Scholes, R.J. and Thornton, P., 2016. Evaluating agricultural tradeoffs in the age of sustainable development. *Agricultural Systems.*

Kevan, P.G., 1999. Pollinators as bioindicators of the state of the environment: species, activity and diversity. Agriculture, Ecosystems & Environment, 74(1), pp.373-393.

Latruffe, L., Diazabakana, A., Bockstaller, C., Desjeux, Y. and Finn, J., 2016. Measurement of sustainability in agriculture: a review of indicators. *Studies in Agricultural Economics*, 118(3), pp.123-130.

Loos, J., Abson, D.J., Chappell, M.J., Hanspach, J., Mikulcak, F., Tichit, M., Fischer, J., 2014. Putting meaning back into "sustainable intensification." Frontiers in Ecology and the Environment 12, 356-361. doi:10.1890/130157

López-Ridaura, S., Masera, O. and Astier, M., 2002. Evaluating the sustainability of complex socio-environmental systems. The MESMIS framework. *Ecological indicators*, 2(1), pp.135-148.

López-Ridaura, S., van Keulen, H., van Ittersum, M.K., Leffelaar, P.A., 2005. Multiscale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. *Environment, Development and Sustainability* 7 (1), 51-69

Mahon, N., Crute, I., Simmons, E. and Islam, M.M., 2017. Sustainable intensification–"oxymoron" or "third-way"? A systematic review. Ecological Indicators, 74, pp.73-97.

Marinus, W., Ronner, E., van de Ven, G.W.J, Kanampiu, F., Adjei-Nsiah, S and Giller, K.E. (forthcoming) The devil is in the detail! Sustainability assessment of African smallholder farming. In: Bell, S. and Morse, S. (eds.) Sustainability Indicators. Routledge.

Miettinen, K. (2014) Survey of methods to visualize alternatives in multiple criteria decision making problems. OR Spectrum, 36(1): 3-37.

Milićević, D.R., Škrinjar, M. and Baltić, T., 2010. Real and perceived risks for mycotoxin contamination in foods and feeds: challenges for food safety control. Toxins, 2(4), pp.572-592.

Mohammed, B., Gabel, M. and Karlsson, L.M., 2013. Nutritive values of the drought tolerant food and fodder crop enset. African Journal of Agricultural Research, 8(20), pp.2326-2333.

Negash, M., Starr, M. and Kanninen, M., 2013. Allometric equations for biomass estimation of Enset (Ensete ventricosum) grown in indigenous agroforestry systems in the Rift Valley escarpment of southern-eastern Ethiopia. Agroforestry systems, 87(3), pp.571-581.

Niemeijer, D. and de Groot, R.S., 2008. A conceptual framework for selecting environmental indicator sets. *Ecological Indicators*, 8(1), pp.14-25.

Olsson, J.A., Bockstaller, C., Stapleton, L.M., Ewert, F., Knapen, R., Therond, O., Geniaux, G., Bellon, S., Correira, T.P., Turpin, N. and Bezlepkina, I., 2009. A goal oriented indicator framework to support integrated assessment of new policies for agri-environmental systems. *Environmental Science & Policy*, 12(5), pp.562-572.

Pender, J., Scherr, S. and Durón, G., 1999. Pathways of Development in the Hillsides of Honduras: Causes and Implications for Agricultural Production, Poverty, and Sustainable Resource Use. EPTD discussion paper 45. IFPRI, Washington D.C. http://www.ifpri.org/publication/pathways-development-hillsides-honduras

Petersen, B., Snapp, S., 2015. What is sustainable intensification? Views from experts. Land Use Policy 46, 1-10. doi:10.1016/j.landusepol.2015.02.002

Pretty, J.N., 1997. The sustainable intensification of agriculture. Natural resources forum 21, 247-256.

Raut, N., Sitaula, B.K., Aune, J.B. and Bajracharya, R.M., 2011. Evolution and future direction of intensified agriculture in the central mid-hills of Nepal. International Journal of Agricultural Sustainability, 9(4), pp.537-550.

Reardon, T., Crawford, E.W., Kelly, V.A., Diagana, B.N., 1995. Promoting Farm Investment for Sustainable Intensification of African Agriculture. Food Security International Development Papers.

Reeves, D. W. 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil and Tillage Research 43* (1-2) pp.131-167.

Remans, R., DeRosa, K., Nkhoma, P. and Palm, C., 2013. 16 Assessing and improving the nutritional diversity of cropping systems. Agro-Ecological Intensification of Agricultural Systems in the African Highlands, p.210.

Riley, J., 2001. Multidisciplinary indicators of impact and change: key issues for identification and summary. Agriculture, Ecosystems & Environment, 87(2), pp.245-259.

Smith, A., S. Snapp, R. Chikowo, P. Thorne, M. Bekunda and J. Glover. 2017. Measuring sustainable intensification in smallholder agroecosystems: A review. *Global Food Security* 12:127-138.

Snapp, S., Grabowski, P., Chikowo, R., Smith, A., Anders, E., Sirrine, D., Bekunda, M. Submitted. Maize yield and profitability tradeoffs with social, human and environmental performance: Is sustainable intensification feasible?

Snapp, S.S., Blackie, M.J., Gilbert, R.A., Bezner-Kerr, R. and Kanyama-Phiri, G.Y., 2010. Biodiversity can support a greener revolution in Africa. Proceedings of the National Academy of Sciences, 107(48), pp.20840-20845.

Snapp, S., 2002. Quantifying farmer evaluation of technologies: The mother and baby trial design. In: Bellon, M.R. and Reeves, J., (Eds.) *Quantitative analysis of data from participatory methods in plant breeding*. CIMMYT.

Speelman, E. N., López-Ridaura, S., Colomer, N. A., Astier, M., & Masera, O. R., 2007. Ten years of sustainability evaluation using the MESMIS framework: Lessons learned from its application in 28 Latin American case studies. The International Journal of Sustainable Development & World Ecology, 14(4), 345-361.

Swindale, A. and Bilinsky, P., 2006. Household Dietary Diversity Score (HDDS) for measurement of household food access: Indicator guide, Washington,

Taylor, D.C., Mohamed, Z.A., Shamsudin, M.N., Mohayidin, M.G. and Chiew, E.F., 1993. Creating a farmer sustainability index: a Malaysian case study. American Journal of Alternative Agriculture, 8(04), pp.175-184.

Tefera, A. 2016. Ethiopia Grain and Feed Annual Report. Global Agricultural Information Service Report ET1608. USDA Foreign Agricultural Service. Retrieved from: https://www.fas.usda.gov/data/ethiopia-grain-and-feed-annual-0

U.S. Agency for International Development (USAID), 1992. Definition of Food Security, Policy Determination, PNAAV468, Washington D.C

Van Cauwenbergh, N., Biala, K., Bielders, C., Brouckaert, V., Franchois, L., Cidad, V.G., Hermy, M., Mathijs, E., Muys, B., Reijnders, J. and Sauvenier, X., 2007. SAFE—A hierarchical framework for assessing the sustainability of agricultural systems. *Agriculture, ecosystems & environment*, 120(2), pp.229-242.

van der Werf, H.M. and Zimmer, C., 1998. An indicator of pesticide environmental impact based on a fuzzy expert system. Chemosphere, 36(10), pp.2225-2249.

Vermeulen, A.T., Hensen,A. and Bleeker, A., 2012. InGOS-a FP7 Infrastructure project for non-CO2 greenhouse gas Observations. In: *EGU General Assembly Conference Abstracts* (Vol. 14, p. 4263).

Yemataw, Z. 2014. The Africa RISING enset research initiative in Ethiopia: Enhancing the productivity of farming systems. Africa RISING Brief 10. Nairobi, Kenya: International Livestock Research Institute.

Zurek, M., Keenlyside, P., and Brandt, K. 2015. Intensifying agricultural production sustainably: A framework for analysis and decision support. Amsterdam, The Netherlands: International Food Policy Research Institute (IFPRI); Climate Focus. http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/130125

Appendix – Examples of tradeoff exercises

Example of Enset in Ethiopia (adapted from Grabowski et al., forthcoming)

In this appendix, we present an example tradeoff analysis with a real-world complex SAI innovation in Ethiopia. The example comes from our interactions in November 2015 with researchers from AfricaRISING working to improve food security through the crop enset in Upper Gana and Jawe Kebeles, in the region of Southern Nations, Nationalities, and Peoples, Ethiopia.

Enset (*Ensete ventricosum*), also known as the false banana, is a native food crop of Ethiopia. Like a banana plant, it is tall with large fleshy leaves. However, rather, the part of enset consumed by humans is the short, fleshy underground root (corm). The corm is cooked like a potato and the stem is squeezed and the output fermented and then baked to form thin bread and other food products (Mohammed et al., 2013). Enset is drought tolerant and improves food security in drought prone areas of Ethiopia. Other benefits derived from enset include leaves as animal fodder during droughts, fiber from the stem to make ropes and strings, leaves as mulch to improve soil fertility and moisture (canopy cover), and stems to make glue (Negash et al., 2013). Enset has cultural medicinal uses and ownership of a field of enset is prestigious in some communities since it is an indication of high social economic status. Cooking and preparing enset is labor intensive, and the majority of the work is done by women.

The AfricaRISING project works with communities to improve the productivity of enset through disease control, crop management, and improved genetics. The primary goal is to improve food security directly through enset consumption and indirectly through improvement of animal feed availability) and indirectly through soil fertility enhancement, for example. The project also works to improve marketing of enset, which has a high market value and is used for produce sacks, glue, glucose (syrup), and food products.

The potential for improved enset production to contribute to sustainable intensification is compelling for a variety of reasons. First, the crop matures over several years and its harvest is less susceptible to annual rainfall fluctuations, such as the 2015-16 El Niño droughts that reduced cereal production by 20% in Ethiopia (Tefera, 2016). The need for resilient staple food production is an important strategy for disaster risk reduction and food security. The potential for commercializing multiple products from enset (which is not yet fully exploited) also could contribute to resilient income. These risk reductions strategies may be just as important considerations as effects on annual net income.

Second, enset and livestock are synergistic and land productivity can be very high. Enset produces feed that can be harvested throughout the year that can help overcome dry season feeding constraints, which allows herd size to increase and manure production to increase. The animal manure in turn is used to fertilize enset. This synergistic relationship may have complex consequences (positive or negative) on the landscape. Positively, enset production can reduce pressure in pasture land during the dry season; however, it could also enable higher stocking rates, which may actually increase pressure on pasture lands.

Local project staff explained that enset production has been in decline over the past decade for several reasons. Enset production is a long-term investment as the primary harvest occurs 7 to 10 years after planting. With increased market linkages, farmers are looking for higher and more immediate returns. One response to this has been increased effort to develop commercial use of enset products, e.g., starch for glue and fiber for ropes.

Another reason for decline is a severe bacterial wilt that has been killing enset plants. Farmers have been asking for a quick chemical solution to the disease, but bacterial plant infections often require prevention through careful management (Yewataw, 2014). The bacteria may be spread through cutting tools and through the manure of animals that are fed the infected material, which is a common use for diseased plants. Areka Agricultural Research Center is carrying out research to address this constraint to enset production.

Using the information obtained from discussion with the site team, we developed a tradeoff diagram to explore the current enset production system across the five domains (Figure A1). Enset production provides food during drought, regular feed for animals, and some income. The improved soil fertility is a delayed effect mentioned by farmers who stated that they do not need to fertilize a field for several years after enset is harvested, probably due to the manure application during the enset production years as well as the accumulation of decomposed enset roots and residues.

We also assume there is some erosion prevention (compared to annual crops) due to the year-round leaf cover and living root system. There are mixed effects on the social domain: high labor requirements for women but also high social prestige for households with larger plots of enset.

Next, we considered changes that might occur if the research efforts to improve market linkages and to control the bacterial wilt were successful (referred to as 'scenario'; Figure A2). This exercise provides an avenue for stakeholders to discuss the context, objectives, and indicators that may be needed to assess performance of the innovation holistically (across all domains). In addition, linkages are identified across indicators to assess tradeoffs and synergies. It emerged during this exercise that gains in production were expected to positively correlate with market orientation and food security.

Additionally, a key learning was the need to avoid unintended consequences for gender such as high labor requirements for post-production processing, perhaps by shifting towards an emphasis on mechanization. The project should consider possible interventions to reduce the demands on female labor both for the cropping and processing.

This highlights how this exercise can help identify tradeoffs and select a complete set of indicators for SAI assessment. Furthermore, it opens discussions on possible additional interventions that are needed to reduce tradeoffs and enhance synergies. It is important to have these discussions prior to implementation in case there are not obvious interventions that would mitigate tradeoffs.

Figure A1: Baseline diagram of tradeoffs and synergies for enset (false banana) in Ethiopia

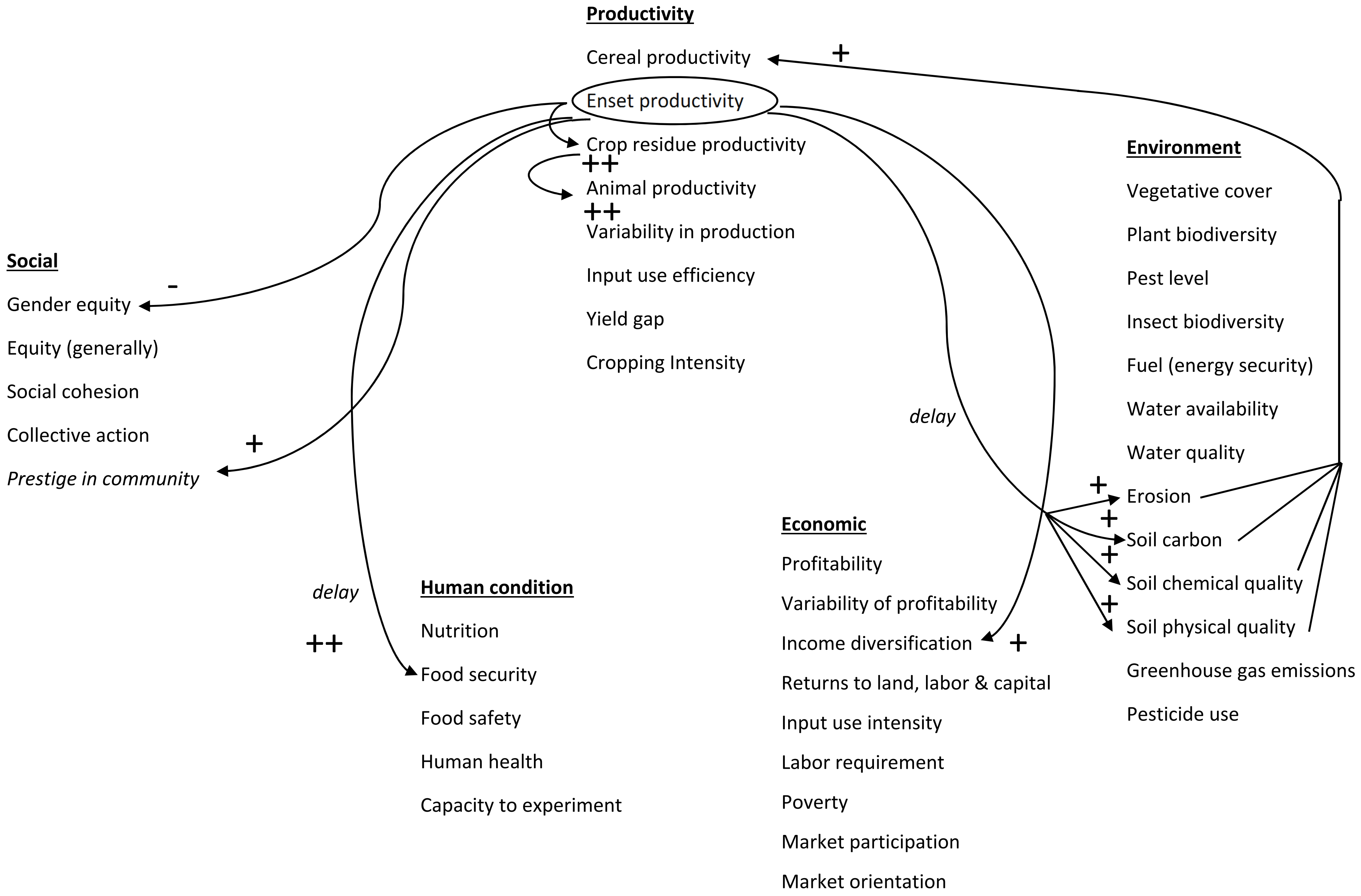


Figure A2: Diagram of tradeoffs and synergies for enset for the scenario of successful productivity and marketing research (from Grabowski et al. forthcoming).