

50th GEF Council Meeting
June 07 – 09, 2016
Washington, D.C.

**A REVIEW OF INDICATORS USED TO ASSESS THE SUSTAINABILITY OF
COMMODITY AGRICULTURAL PRODUCTION**

A review of indicators used to assess the sustainability of commodity agricultural production

Review paper for the Scientific and Technical Advisory Panel of the Global Environment Facility

Authors: Laura Vang Rasmussen¹, Arun Agrawal¹, Johan Oldekop¹

October 2015

Acknowledgments

This review paper was commissioned by the Scientific and Technical Advisory Panel (STAP) of the Global Environment Facility (GEF) to underpin their work on development and selection of indicators for the Integrated approach Pilot (IAP) on 'Taking Deforestation out of Commodity Supply Chains'. It was funded by the GEF and it contributes to the DFID project 'Assessing forest-based livelihoods dependence and understanding the impact of DFID's forest investments on livelihoods'. We thank STAP (Thomas Hammond, Guadalupe Duron, Annette Cowie, and Michael Stocking) for providing very constructive review comments on an earlier report draft. Special thanks to chair of the STAP Rosina Bierbaum for valuable guidance through the literature review and the development of this report. Thanks to the agencies and organizations interviewed for sharing their knowledge and experience on indicator development and selection.

1. INTRODUCTION

1.1. Sustainability indicators and the agricultural commodity sector

Many sustainability indicator sets have been assembled since the Rio Earth Summit in 1992. For example, the Compendium of Sustainable Development Indicator Initiatives, a worldwide directory of who is doing what in the field of sustainability indicators, mentions more than 600 current efforts to identify sustainability indicators, with hundreds of new indicators still being proposed (Parris & Kates, 2003). More recently, 100 Global Monitoring indicators accompanied by additional national indicators have been proposed to track the full range of the 17 Sustainable Development Goals (SDGs) and their 169 targets (Sustainable Development Solutions Network, 2015). However, a reliable set of indicators to assess progress toward the SDGs remains to be agreed upon, and urgent attention is therefore needed to the process of indicator selection and all the challenges it entails. 1

Yet, a number of difficulties remain in the selection of appropriate indicators. For example, the majority of sustainability indicators developed are generally claimed to be from separate analyses of economic, social and natural processes, albeit in the context of sustainability, is the integration of indicators across different arenas and scales that is most critical (Hak, 2009). Furthermore, defining an appropriate set of

¹ International Forestry Resources and Institutions (IFRI), School of Natural Resources & Environment, University of Michigan

indicators is a difficult task as too few indicators may result in important developments being missed and trade-offs will thus not be properly taken into account (Bossel, 2001). Conversely, too many indicators can make the data collection and data processing costly, redundancies might appear, and the message expressed by the set of indicators becomes challenging to interpret (Van Cauwenbergh et al., 2007). The essence of the dilemma related to too many indicators resides in the fact that the indicators are seldom integrated and that they often entail new data collection efforts. The need for an explicit focus on the process of indicator selection is therefore not just relevant in the ongoing selection of indicators to track progress towards the SDGs - it is pertinent for most applications of sustainability indicators.

The agricultural commodity sector represents a particular interesting case. Even though agricultural systems have been at the heart of much of the sustainability debate for decades as these systems occupy large areas of land – far more than any other industry – and the environmental impacts may accordingly be more profound, there nevertheless appears to be limited agreement among various stakeholders as to exactly which indicators should be included in measures hereof (Sachs et al. 2010). Even with projections that the agricultural commodity production will have extensive and increasingly adverse effects on a variety of interlinking aspects of the global environment including, and arguably most significantly, deforestation (Angelsen, 2010; Newton et al., 2013; Hosunuma et al. 2012; Persson et al. 2014), the selection of a limited set of indicators is not a straightforward exercise. Regardless of whether the production increase takes place through expansion or intensification, it will have subsequent environmental impacts relevant to e.g. biodiversity, land and water quality, and carbon sequestration; these environmental impacts are often concomitant with socioeconomic impacts on human livelihoods and well-being. But the increase in commodity agricultural production may at the same time create development opportunities for local and global communities. This situation makes it inherently difficult to assess the overall sustainability of the rapidly expanding commodity agricultural production.

Accordingly, a wide spectrum of views exists about which sustainability dimensions to track. Examples include the ‘Initiative on Global Biodiversity Impact Indicators for Commodity Production’ launched in late 2014 by the Secretariat of the Convention on Biological Diversity. Since many agencies and organizations working on sustainable production of agricultural commodities typically develop internal monitoring systems that are adapted to specific sustainability standards, comparisons between the assessments carried out by different practitioners becomes problematic. Although sustainability standards have emerged as potentially promising avenues for encourage compliance with various sustainability guidelines (Milder et al., 2015; Rueda & Lambin, 2013), it may be that this flexibility in the sustainability guidelines has allowed – and possibly reinforced - the application of various and very diverse standards and indicators. For example, the sustainability standards use different indicators to measure agricultural performance and compliance depending on the desired outcome (e.g. encourage local employment, protect old forest growth, conserve natural habitats, sequester carbon, and maintain watershed services), and a review of 12 agricultural sustainability standards revealed that not all standards prohibited clearance of certain land cover types such as primary forest (UNEP-WCMC, 2011). Even though many of the standards unite values related to poverty, environment, and health outcomes

(Barham & Weber, 2012), the very definition of sustainable agriculture depends to a large extent on the standard despite the existence of more commonly acknowledged definitions such as agro-ecology, alternative agriculture, ecological food production, low-input sustainable agriculture and organic agriculture (Bell & Morse, 1999), and more recently deforestation-free agriculture (see e.g. Macedo et al. 2012).

1.2. Introducing the GEF's IAP on 'Taking deforestation out of the supply chain'

This report was commissioned by the Scientific and Technical Advisory Panel (STAP) of the Global Environment Facility (GEF). The STAP is committed to supporting the GEF's Integrated Approach Pilot (IAP) on 'Taking Deforestation out of Commodity Supply Chains' and this report contributes to the development of metrics and indicators applicable across target countries and key commodities. The overall program objective of the IAP is: 'to reduce the global impacts of agricultural commodities on climate change and biodiversity by meeting the growing supply and demand of palm oil, soybean and beef through means that do not lead to deforestation'. The GEF has long worked at the interface between forest management, biodiversity conservation, and the wellbeing of forest-based communities. The commodity agriculture program builds on this experience, seeking to achieve far greater conservation impact and co-benefits through direct engagement with the many actors involved in affecting demand and supply in agricultural commodities.

1.3. Study objectives and anticipated contribution to the GEF's IAP

The main objectives of this study are to:

- synthesize how indicators have been applied in the peer-reviewed literature to assess the sustainability of commodity agricultural production
- identify key priorities in indicator selections among practitioners through interviews with organizations and agencies that feed into the GEF's IAP on 'Taking Deforestation out of the commodity Supply Chains'
- propose indicators relevant for the GEF's IAP on 'Taking Deforestation out of the commodity Supply Chains'

By providing a system-wide overview of previous indicator applications among scholars and comparing them with the key priorities highlighted by a subset of relevant agencies and organizations, we wish to shed some light on the way forward for indicator development and selection. We explore which of the indicators identified as most common in the peer-reviewed literature that are actually embraced by practitioners and the reasons why. By doing so, we identify opportunities and pitfalls related to the common indicators and we explore how indicators can be applied (separately or integrated) in a cost effective manner that makes trade-offs transparent between, for example, the different pillars of sustainability, the short and the long term, and different spatial scales. Based on the literature review and the assessment of useful indicators, we finally outline a subset of indicators relevant to track outcomes related to the GEF's IAP.

Since our effort explicitly serves to integrate the findings from the peer-reviewed literature and the interviews with practitioners, we try to bridge the scholar-practitioner gap that seems very present in indicator selection criteria and indicator selection. As we address some of the main challenges to select sustainability indicators and improve future monitoring of the rapidly growing commodity agricultural sector, the findings are not only relevant for the GEF's IAP but also the ongoing indicator selection process for the SDGs (UN, 2014) which are anticipated to be an opportunity to align approaches and common objectives, contribute to integrated strategies, and support common reporting between the Rio Conventions (O'Connell et al., 2015).

As there is no ideal indicator set that fully encompasses all the desired perspectives of sustainability, we make no claim that the indicators proposed in this paper are definitive indicators of sustainable production of agricultural commodities. Rather, we acknowledge that the key question both among and across practitioners and scholars has shifted from 'how do we develop, yet again, a set of indicators?' to 'which existing indicators can we use to better align approaches?' (Lebacqz et al. 2013). We caution that indicators are by definition communication tools, and the challenge is how to simplify without misrepresenting reality.

For the purposes of this study, agricultural commodities are synonymous with a class of goods derived from terrestrial ecosystems. The GEF's IAP focusses specifically on beef, palm oil and soybean as these commodities in addition to cocoa and rubber are associated with high rates of deforestation (Newton et al. 2013) concomitant with their rapid growth in the countries in which they are principally produced (beef and soybean are dominant commodities in South America, while palm oil is grown largely in Southeast Asia, FAO Stat (www.faostat.fao.org). Since the goal of our review is to provide a system wide overview of indicators previously applied across key commodities and countries, we expand the literature review to also include cocoa, coffee, corn, cotton, rice, rubber, sugar cane and wheat.

2. PERSISTENT CHALLENGES AND RECOMMENDATIONS IN INDICATOR SELECTION

Since this report aims to shed some light on the way forward for indicator development and selection, the following sections will provide an overview of the challenges that remain. These pertain primarily to the thematic sustainability areas that receive more or less attention in the peer-reviewed literature and the failure of scholars to base indicator selections on criteria deemed of key importance by practitioners – i.e. cost-effective measures based on available data and comparable across countries and countries. Below, we provide a more detailed discussion of these challenges and by doing so, we propose some next steps on the pathway to develop a common subset of indicators of value to the GEF's IAP as well as broader sustainability assessments of commodity agriculture.

2.1. A tendency to remain focused on environmental aspects in the peer-reviewed literature

As with any sustainability assessment, one should not be surprised that there are differences in the set of indicators prioritized by various scholars. Reasons for the absence of agreement as to what key facets of sustainability to be addressed and which matching indicators to apply likely have to do with the inherent ambiguity in the sustainability concept. Further, the reviewed studies are undertaken by a multiplicity of scholars whose views about the sustainability of commodity agricultural production are

not coordinated. But despite the lack of consensus among scholars as to exactly which indicators to use, the reviewed articles have some striking commonalities which in fact represent challenges in taking the findings beyond the phase of research.

Firstly, the reviewed studies typically attend to less than five indicators and focus remains primarily on the environmental facets of commodity production. Although the broader literature surely contains studies applying a broad range of indicators to track environmental, economic and cultural/social facets of the production in conjunction, we can at least defend the conclusion that environmental focused studies based on a few indicators constitute the bulk of the published literature describing indicators to assess the sustainability of commodity agricultural production. Although these studies are valuable in providing detailed information about particular environmental indicators, they are also limited in the extent to which the common indicators can be the basis of general monitoring of commodity agricultural production (FAO, 2014). These limits are primarily compounded by the tendency in the studies to remain focused on only the environmental dimensions of the production. Indeed, environmental aspects should be a core focus, but conservation and land management actions are highly determined by political, societal and economic interest which makes the downplaying of such dimensions problematic (Velasco et al. 2015). It is thus evident that an indicator set limited to, for example, soil health would be inadequate to assess the most pressing concerns associated with expanding commodity production. There is, however, nothing inherently problematic about the choice of particular dimensions to examine whether the commodity production is sustainable. But it is problematic that the reviewed studies do not acknowledge the simplification they introduce by not considering important aspects such as deforestation or gender issues.

We emphasize here that an indicator set limited to a separate sustainability pillar fails to comply with one of the key requirements of sustainability indicators: that the sets need to be holistic, covering all pillars of sustainability as it is the integration across different policy arenas that is most critical (Mcglade 2007, Reed et al. 2006). Thus, it must be acknowledged that all pillars are of enduring relevance for addressing overall sustainability, but certainly the subsequent choice of thematic areas under each pillar depends on the stakeholders and the aim of the sustainability assessment. We caution that the integration across sustainability pillars and thematic areas must be carefully conducted. Indeed, there will always be trade-offs between the three dimensions. But based on our identification of which key facets of commodity production practitioners consider most important to track, it is evident that different indicators may well point in different directions.

2.2. Attention needed to cost-effective measures based on ongoing data collection efforts

Another key challenge in deriving common sustainability indicators from the peer-reviewed literature is that scholars devote limited attention to explicitly describing the indicator selection process. The inattention to specify selection criteria indicates that the criteria generally are subjectively based according to the study aim. Despite repeated calls for defining appropriate indicator selection criteria within the field of agricultural production (see e.g. Pannell and Glenn, 2000), the selection process still appears insufficiently systematic and transparent (see also Niemeijer and de Groot, 2008). Further, the selection criteria typically apply to individual indicators rather than the inter-relation of indicators. It

may even be this inattention among scholars to explicitly defining selection criteria that has exacerbated the lack of agreement as to which indicators to use.

Explicitly defined criteria for indicator selection are by contrast a basic element of the monitoring frameworks set out by most of the interviewed agencies and organizations. Ultimately, the financial cost related to measurements is a corner stone for both practitioners and scholars and the selection of appropriate indicators must accordingly account for the likely paucity of data that will make it challenging to strike a proper balance between cost-effectiveness, scientific rigor, and relevance (Milder et al., 2012). This is also pertinent for initiatives tracking progress towards the SDGs and we argue that common indicators will need to rely on existing data collection programs and available datasets that provide a longitudinal dimension to track long-term trends over time. We caution against selection of indicators for which no adequate datasets exist. Yet, it should be noted some scholars might not perceive the very meaning of cost-effective indicators as a matter of making use of existing data (see also Zhen & Routray 2003; UNCSD 2001). For example, they would focus on appropriate sampling designs that are within the capability of involved stakeholders. Indeed, some studies would not necessarily benefit from being more attentive to indicators that rely on existing data collection programs as a trade-off may in fact occur between cost-effective indicators and indicators that may be more ideal with regards to scientific rigor – e.g. extensive measures of biodiversity and ecosystem health (Moldan and Dahl, 2007). But for studies that do intend to inform decision-making processes, we suggest data availability as a crucial guiding principle for indicator selection. This will be an important first step in the attempt to bridge the very apparent scholar-practitioner divide.

What then is the way to base general indicator development processes on existing data collection programs? Since it is seldom the case that data are being collected with the requested time and spatial scale, we argue that any indicator development has to be based on a scoping exercise of available data and existing data collection programs with attention devoted to the frequency and spatial scale for which those data are being collected. In particular, one should address how different data sets can be comparable and merged across time. By doing so, the development of indicators would account for potential scale mismatches and potentially filling time gaps.

2.3. Attention needed to indicators that are comparable across countries and commodities

A third key challenge to note in the attempt to derive common ‘good’ indicators from the peer-reviewed literature relates to the scale addressed. While the indicators applied in the reviewed studies tend to be locally defined and monitored, a multitude of agencies are calling for indicators comparable across countries and commodities.

Defining the appropriate spatial scale is, however, not a straightforward exercise. Faced with diversity in both geographies and commodities, no search for generic indicators can escape a high level of generalization that does not capture the complex realities of local sites (see also Doward 2013). On the one hand, Parris and Kates (2003) argue that the spatial scope of the indicator must correspond to the way in which the intended audience can affect change. On the other hand, sustainability indicators should be diverse enough to meet the requirements of different stakeholders (Freebairn and King 2003,

Bossel 2001). As for commodity agriculture, a multi-stakeholder environment ranges from local producers to national governments and global consumers which all have different perceptions of the aspects that make commodity production sustainable. A recent attempt to address aspirations for sustainability suggests concentrating on a biophysically sustainable and productive regime in which household/farm production exceeds the needs of the household (O'Connell et al., 2015). While the environmental dimension potentially could be addressed across various spatial scales, a narrow emphasis on production and household needs appears to downplay economic and cultural/social facets of importance at larger spatial scales such as the institutional context in which the commodity production takes place. Surely, the full assemblage of key and supplementary indicators should strike a proper balance between indicators well suited to small scale geographies such as the farm level and indicators applied at a large enough scale to include not only local but also long distance impacts on people and ecosystems (Deprez and Miller 2014). Thus, it appears pertinent for scholars in future monitoring efforts to consider how site-specific indicators such as those related to soil health can be incorporated in indicator sets applicable at various spatial scales.

2.4. Setting targets?

Finally, it should be noted that a characteristic which might influence the very success of indicators pertains to whether the indicator has targets or threshold (see e.g. Miller and Wahlén 2015). Yet, it is not highlighted as an explicit selection criterion by the interviewed agencies and only few of the reviewed articles attend explicitly to defining sustainability levels. The reviewed studies are typically conducted within a single time period rather than examining progress against a set of baseline data or a given target (see e.g. Ferraro et al. 2003). The studies clearly do not adopt a time horizon long enough to capture relevant human and ecosystem change (one of the Bellagio Principles for gauging progress towards sustainable development, see Hardi and Zdan, 1997), and many indicators promoted as sustainability indicators thereby serve to measure agricultural development rather than the sustainability of agricultural production (Dahl 2007). However, the complexity of the concept of sustainable commodity agricultural production as well as the fact that it reaches out into the future, tends to give a certain direction for policy making rather than serving as a benchmark that could be precisely defined when selecting indicators and setting targets.

In general, contradicting arguments exist as to whether or not targets should be set. On the one hand, it was suggested decades ago that any indicator purporting to measure sustainability should be reported with reference to any known threshold value of that indicator (Liverman et al. 1988). On the other hand, Pannell and Glenn (2000) argue that the idea of desired levels and thresholds is flawed with regards to agriculture sustainability indicators. This can, for example, be illustrated with the indicator 'protected area as percentage of total land area'. The higher the percentage of protected land, the higher the likelihood is of sustainability of these areas. However, the higher the percentage the more areas could have proven to need protection, which means that the former human activities may have had unacceptably high impacts and accordingly caused low sustainability in those areas. The latter line of argumentation stands also in stark contrast to the more recent calls for inclusion of targets or thresholds – as evident in the monitoring framework suggested for the Sustainable Development Goals (Sustainable Development Solutions Network, 2015) which places great emphasis on building consensus around

national thresholds and indicators to measure progress towards specific targets. As the majority of the proposed core indicators (Table 3) are represented in the Sustainable Development Goals monitoring framework, one can expect that national thresholds will be identified for those indicators.

3. CONCLUSION

This paper firstly examines how indicators have been applied in the peer-reviewed literature to assess the sustainability of commodity agricultural production. Secondly, we outline key production facets that a subset of relevant agencies deem important to track. And thirdly, we propose which of those facets and their matching indicators are the minimum list of practical measures for the GEF's IAP on 'Taking Deforestation out of Commodity Supply Chains'.

Although the literature review found some agreement among scholars, the set of indicators were large and diverse, creating obstacles to develop systematic and comparable monitoring efforts. The findings show how the concept of sustainability does not lend itself to clear definitions of what to measure, and this has led to different interpretations of which outcomes are deemed important to track. Further, the limited agreement about which sustainability dimensions to include thrive in the scientific literature as a result of inattention to explicitly defining criteria for indicator selection. Different scholars apply different indicators, and there is little basis on which to suggest a subset of meaningful indicators covering the three sustainability pillars. The position we argue in this paper is that the sheer number of indicators applied in the peer-reviewed literature makes it pertinent to select a subset of key indicators relevant for core facets of agriculture commodity production.

In contrast to the evidence from the reviewed articles, interviews with a subset of relevant agencies and organizations demonstrate how practitioners place emphasis on explicitly defined selection criteria. The financial costs related to measurements are considered a corner stone, and we argue that this is primarily a matter of relying on existing data collection programs and already available datasets that provide a longitudinal dimension to track long-term trends over time. When evaluating the indicators applied most frequently in the peer-reviewed literature against this key selection criterion, we find that the majority of those indicators fail to be considered 'good' indicators by practitioners. Only the indicators 'Nitrogen balance' and 'GHG emissions' have found common ground. What is troubling is that if the indicators applied most frequently by scholars continue to be used merely by scholars, it likely will not be helpful in addressing how best to track deforestation-free commodity production and may further exacerbate the scholar-practitioner divide in this domain. To address this problem, we advocate careful attention among scholars to select indicators that are easy to operationalize by governmental agencies, non-governmental organizations, donors and other stakeholders tracking progress, and we suggest that the selection and development of indicators should explicitly integrate data availability as a crucial guiding principle.

The fact that there is no consensus on which indicators to use makes it also pertinent that the scientific and policy communities first jointly articulate an agenda for sustainable commodity agriculture production and identify which core facets of sustainability need to be tracked. The findings presented in this paper serve to inform this process as we provide a list of 12 core facets of sustainable commodity

agriculture deemed important to track by the interviewed agencies and organizations (See table below). We envision that this list can be used as a reference to ensure that the major sustainability attributes of commodity agricultural systems are considered and relevant indicators measured. We make no claim that this list of core facets and their matching indicators is definitive of sustainable production of agricultural commodities. But the suggested indicators do comply with what must be considered a key principle in indicator selection: that the indicators are cost-effective to measure.

Finally, we propose that five of the 12 core production facets(see bolded text) can be the minimum list of practical measures for the GEF's IAP (See table below). The suggested candidate indicators are all considered cost-effective and comparable across countries and commodities, and their selection is based on their ability to track progress towards the overall program objective: to reduce the global impacts of agricultural commodities on climate change and biodiversity by meeting the growing supply and demand of palm oil, soybean and beef through means that do not lead to deforestation.

Key production facets to track		Suggested core indicators deemed cost-effective, comparable across countries and commodities, and a measure with a clear direction	Examples of supplementary program or project-level indicators
<i>Environmental</i>	Is the production deforestation-free?	a) Annual conversion of forests to agricultural land (crops and pastures) (ha/year)*(4)	% of agricultural land enrolled in programs aiming to produce deforestation-free commodities
	Does the production potentially result in biodiversity loss?	a) Red list Index*(1) b) % of terrestrial area designated for conservation*(1)	% of agricultural land under biodiversity-friendly practices
	Does the production cause pollution?	a) Nitrogen balance for agricultural areas (kg/ha)*(1)	% of agricultural land under nutrient management practices
	Does the production cause GHG emissions?	a) GHG emissions from production*(3)	
	Does the production result in pressure on water resources?	a) Water use efficiency b) Crop water productivity (tons of harvested product per unit of irrigation water)*(2)	
<i>Economic</i>	Is the production efficient?	a) Yield gap*(1)	
	Are the markets stable?	a) % of production that is produced and sold to various standards and certification schemes	% of production that is produced and sold to a specific standard or certification scheme
	Does the production potentially contribute to poverty reduction?	a) Multidimensional poverty index*	# of farmers in targeted sites with increased net income
<i>Social, cultural and institutional</i>	Does the production result in improved access to e.g. health care and education?	a) Multidimensional poverty index*(6)	% of production in compliance with animal welfare standards
	Does the production respect basic rights?	a) % of farmers who perceive their rights are recognized and protected*(5)	# of farmers receiving at least a minimum pay
	Is there a functioning extension system in place?	a) # of extension workers per 1000 farmers*(1)	
	Is there a functioning institutional system in place?	a) Documented positive changes in regulation/ enforcement that promotes sustainable production of key commodities	