Perspective

The Challenge of Measuring Agricultural Sustainability in All Its Dimensions

Pietro Gennari *, Dorian Kalamvrezos Navarro

FAO, Viale delle Terme di Caracalla, Roma, 00154, Italy

* Correspondence: Pietro Gennari, Email: pietro.gennari@fao.org.

ABSTRACT

SDG indicator 2.4.1, the "Percentage of agricultural area under productive and sustainable agriculture" aims to measure the degree of sustainability of each farm with reference to 11 distinct sustainability attributes, and hence provide an overall national assessment through a dashboard approach. Today, this indicator has an internationally agreed methodology, and dozens of countries around the world are receiving technical assistance from FAO in its measurement and implementation. However, what we may now take as a given is the result of a long and arduous process of methodological development that involved a series of difficult decisions on numerous methodological aspects. This paper reviews and provides supporting documentation for these key methodological decisions, particularly with regard to the definition of agricultural sustainability, the choice of the scale of the sustainability assessment and the data collection instrument; the sub-indicators within each dimension; the criteria to assess the sustainability level of the farm with respect to each sub-indicator; and the modality of synthesizing the information. These decisions were further encumbered by the need to faithfully capture progress towards a multidimensional SDG target determined by a political process, with negligible input from statistical experts.

KEYWORDS: composite indices; indicators dashboard; SDG indicators; Delphi Technique; farm surveys

Open Access

Received: 17 June 2019
Accepted: 16 September 2019
Published: 18 September 2019

INTRODUCTION

The 2030 Agenda for Sustainable Development has no shortage of references to the integrated and indivisible nature of the SDGs and the three dimensions of sustainable development. However, measuring sustainability in all its dimensions, including for agricultural activities, is an ambitious and complex task. The official SDG indicator 2.4.1 "Percentage of agricultural area under productive and sustainable agriculture" is the result of a long and hotly debated development process in which various alternative indicators were considered by the Interagency and Expert Group on SDG indicators (IAEG-SDG), the relevant UN forum established under the aegis of the UN Statistical Commission to lead the selection and implementation of the global indicators to monitor the 169 targets of the 2030 Agenda.

Fast-forward three and a half years, and SDG indicator 2.4.1 has an internationally established methodology [1] approved by the IAEG-SDG only in November 2018 [2] and now being implemented by dozens of countries around the world thanks to FAO technical assistance. In this article, the origins, evolution and final structure of SDG indicator 2.4.1 will be reviewed, shedding light on the critical choices underlying the current methodology. By describing the challenges encountered with each of these choices and explaining the rationale for the final approach selected, the purpose of this article is to provide the user community with an essential instrument to understand the strengths and limitations of this indicator and to interpret its results at the national, regional and global levels.

The article will initially delve into the separate processes that defined the SDG targets and SDG indicators, which explain some of the limitations of SDG indicator 2.4.1, and hence review the milestones of its methodological development, particularly with regard to: the definition of agricultural sustainability; the choice of the scale of the sustainability assessment and the most appropriate data collection instrument; the choice of sub-indicators within each dimension; the sustainability criteria to assess the sustainability level of the farm with respect to each sub-indicator; and the modality for synthesizing the multidimensional information in one value.

DEFINITION OF THE TARGET AND LIMITATIONS OF THE INDICATOR ON AGRICULTURAL SUSTAINABILITY

A key difficulty for defining SDG indicator 2.4.1 is linked to the formulation of SDG target 2.4. Though not the lengthiest of SDG targets (this distinction goes to SDG target 3.b at 81 words), it is certainly one of the most complex for attempting to combine such a multiplicity of dimensions, drivers and factors:

> *"2.4 By 2030, ensure sustainable food production systems and implement resilient agricultural practices that increase productivity and production, that help maintain ecosystems, that strengthen capacity for adaptation to climate change, extreme weather, drought, flooding and other disasters and that progressively improve land and soil quality"*

The length and complexity of the target are not the only striking features of SDG target 2.4. Out of the five principles underpinning a S.M.A.R.T target (Specific, Measurable, Attainable, Realistic, Timely), target 2.4 could realistically only presume to satisfy the "time" element, with a clearly defined implementation horizon by the year 2030. On the other end of the spectrum, "measurability" was evidently not the top priority of the drafters of this target—or of most of the other SDG targets. The reasons for this lie in the process that was established to define the SDG framework.

In the authors' view, the ideal approach would have been that a political process would define the SDG Goals, a mixed political/statistical process would determine the targets, and a statistical process would define the SDG indicators. Instead, a political process determined both the Goals and targets, and only then a statistical process was established to define indicators to monitor those targets. This resulted in two main weakness of the SDG target formulation, which often coexisted in the same target, as in the case of target 2.4: the lack of a defined quantitative level of achievement, which makes it impossible to assess when the target has been reached; and the juxtaposition of many diverse policy and measurement dimensions that require multiple metrics and a complex aggregation method in order to determine whether the target as a whole has been achieved.

The political process was centered on the so-called "Open Working Group" (OWG), an intergovernmental forum with selected membership set up in January 2013 by the UN General Assembly. The OWG was mandated with the thorny task of proposing a set of SDGs by September 2014, which should be "limited in number, aspirational and easy to communicate, addressing all three dimensions of sustainable development". The process was fully in the hands of policy-makers, diplomats and permanent representations, who usually have a limited understanding of whether a target is measurable, in principle, and whether relevant data are available for the target to be properly monitored, in practice. Having negligible input by statistical experts (only one representative from the UN Statistical Division was included in the OWG), considerations of S.M.A.R.T. and other desirable characteristics yielded to the pressures of political balancing, which were often resolved by including more and more policy dimensions in the same target. As a result, only a minority (about 30 percent) of the SDG targets have an explicit quantitative level of achievement, which hampers the ability of the international community to assess whether the world is on track to reach the bulk of the targets set by the 2030 Agenda. According to the OECD, the "heterogeneous nature of SDG targets means that setting desirable levels of achievement by 2030 on each indicator requires a variety of approaches" [3], which creates the risk that different Institutions may come up with different assessments of where the world stands in achieving the 2030 Agenda.

The complexity of SDG target 2.4, as defined through the OWG process, and the need, at the same time, to severely limit the number of official indicators to monitor it (the pressure to limit the overall number of global indicators in order to minimize the response burden on countries has led to the selection of "only" 232 unique indicators in total to monitor the 169 SDG targets, which means an average 1.37 indicators per target), meant that the methodological development of SDG indicator 2.4.1 was an uphill battle from the beginning. A number of countries suggested a host of alternative indicators, including Total Factor Productivity, Area under organic agriculture over total cultivated area, Nutrient balance indicator, Percentage of agricultural households using irrigation systems, and Percentage of agricultural households using eco-friendly fertilizers. These alternative indicator proposals, which had been tabled at different times by various countries, were dropped one after the other during dedicated proceedings at the IAEG-SDG. "Total Factor Productivity", in addition to interpretability issues, clearly cannot capture the sustainability aspect of agricultural production, which is arguably the key focus of target 2.4 and SDG 2, more broadly. A similar type of limitation affects the "Percentage of agricultural households using irrigation systems". By contrast, "Area under organic agriculture" misses the productivity aspect and is essentially limited to the environmental dimension of sustainable development, not to mention that there are currently no international standards or global certification bodies for organic agriculture. In a similar vein, the "Nutrient balance indicator" is severely restricted to environmental sustainability, with the added complication of requiring rather high cost for its data collection, not affordable in many countries (*i.e.*, physical sampling and laboratory analysis). Regarding the proposal to adopt the "Percentage of agricultural households using irrigation systems and/or eco-friendly fertilizers", the final decision was also to drop these indicators given that water availability and the sound use of fertilizers are two of the environmental sub-indicators under the broader indicator 2.4.1.

Despite its multidimensionality, indicator 2.4.1 is not comprehensive enough to properly measure every component of the SDG target 2.4, as we will see in more detail when examining its components (sub-indicators). Clearly the measurement of the capacity of the agricultural sector to adapt to "climate change, extreme weather, drought, flooding and other disasters" is not covered by indicator 2.4.1. On the positive side, the fact that the scope of the target is limited to "food production systems", delimits the applicability of the indicator only to the production stage of agricultural activities. Omitting the term "production" would have entailed the need to assess the sustainability of the entire food system, which encompasses all the stages of growing, harvesting, packing, processing, transforming, marketing, consuming and disposing of food.

BUILDING A NEW INDICATOR FOR AGRICULTURAL SUSTAINABILITY

How to Define Agricultural Sustainability

The concept of agricultural sustainability itself is not new. As early as 1988, the FAO Council, the organization's governing body, defined sustainable agriculture as "The management and conservation of the natural resource base, and the orientation of technological and institutional change in such a manner as to ensure the attainment and continued satisfaction of human needs for present and future generations. Such development conserves land, water, plant and animal genetic resources, is environmentally non-degrading, technically appropriate, economically viable and socially acceptable”. This definition was used as the basis for discussions on sustainable agriculture at the UNCED Conference in June 1992 and in further crafting Agenda 21 as it related to agriculture.

But is it possible to measure agricultural sustainability? According to some scholars [4,5], measuring sustainability is impossible because it is a dynamic concept and site-specific. This would suggest that it is futile to search for an indicator that can be applied at the global level to enable measurement and monitoring of agricultural sustainability in different countries. As sustainability is context-specific, a single measurement tool would not provide a robust basis for comparing countries in terms of agricultural sustainability. Hence, a common recommendation is to introduce a set of context-specific indicators that enabled countries to evaluate the sustainability of their own agricultural sector and subsequently compare their status with other countries.

Comparability, however, is a key concern for FAO as a UN agency producing international statistics. Leaving it up to each country to define a set of “context specific” indicators for sustainable agriculture would have been a methodologically questionable decision, as their comparison across countries would not be based on any common benchmark. Such an option would also have been difficult to accept by the IAEG-SDG member countries, as it would essentially mean that determining a global indicator for agricultural sustainability is not possible, when the whole point of the SDG indicator framework is to define a “set of global indicators…to be complemented by indicators at the regional and national levels” [6].

For this reason, since the first standard metadata document submitted to the IAEG-SDG on 14 December 2015, the FAO has steered away from suggesting that each individual country would define for itself the activities and practices conducive to agricultural sustainability. Instead, the proposal calls for “a set of common metrics” to be established “in order to ensure relevance across the whole range of possible socio-economic and bio-physical conditions”. This core set of metrics would ensure internationally comparable results, although countries would be free to supplement the official global indicators with additional “instruments that best capture the priorities most relevant to them”.

The Selection of Sub-Indicators

Regarding the selection of metrics (hereinafter “sub-indicators”), a key premise for FAO is that they should capture all three dimensions of sustainable development, *i.e.*, the social, environmental and economic dimensions. Another key premise is that SDG indicator 2.4.1 would not aim to be exhaustive, but rather practical and feasible at the same time. Seeking a perfect consensus on the full scope of sustainable agriculture would have been a rather hopeless undertaking, and even if it were possible, it would probably have rendered the indicator unmeasurable. This is because measurability is not the same thing as technical feasibility.

As the Report of the UN Task Team on Lessons Learned from MDG Monitoring underlines [7], for an indicator to be measurable, it should be possible to produce it "in a cost-effective and practical manner by countries", with a "regular data collection mechanism developed with reasonable costs". (The fact that many SDG indicator still have no agreed international methodology (the so called "Tier III indicators") bears witness to the difficulties of translating statistical concepts and methods into a viable and measurable indicator. (The global SDG Indicators have been classified in 3 Tiers: Tier I, when an established methodology exists and data are already widely available for the majority of countries; Tier II, when a methodology has been established, but data are available for less than half of the countries; Tier III, when an internationally agreed methodology has not yet been developed and data are not collected yet.)

The selection of the sub-indicators for SDG indicator 2.4.1 therefore seeks to strike a balance between the concerns, on the one hand, to adequately capture all three dimensions of sustainable development, and on the other hand, to present countries with a nimble list of metrics that would ensure measurability and feasibility at the same time. In order to achieve a broad consensus on the key aspects of sustainable agriculture to be monitored, a global consultation was organized in which experts' opinions were gathered and consensus reached following a methodologically sound mechanism.

The first proposed set of sub-indicators emerged during an expert meeting organized by FAO on 3–5 April 2017, which gathered agriculture statisticians from countries across all regions, representatives of the civil society and the private sector, as well as thematic experts from research institutions and academia [8]. In order to support and guide the choice of sub-indicators, the Delphi Technique [9,10] was used to allow participants to anonymously rate/rank the initial list of over 100 sub-indicators proposed in the literature, using the following criteria: policy relevance, universality, comparability, cost effectiveness, and independence. To this end, participants were provided with a form and asked to distribute 100 points among the proposed sub-indicators under each of the three dimensions of sustainable development. In order to reduce potential bias, three versions of the form were created with the list of proposed themes randomized. The mean rating, the corresponding 95% confidence interval and boxplots were calculated for each dimension overall and stratified by group membership to better visualize rating distributions. These results were then reported back to the group (controlled feedback) to reduce noise and allow participants to reconsider their assessment based on others' rankings. This produced a first "long list" of candidate sub-indicators, which was further trimmed into a "short list" by conducting the ranking exercise in an iterative way until convergence was reached.

It should be noted that the independence criterion (or "limited overlap between sub-indicators") only concerns the sub-indicators to be selected for SDG indicator 2.4.1, not the possible overlap between sub-indicators under 2.4.1 and other SDG indicators. On the contrary, the latter possibility, *i.e.*, relying on other SDG indicators with an established methodology as constituents of SDG indicator 2.4.1, can be regarded as an advantage, as it has the potential for reducing the reporting burden on countries, given that the same data could be used for monitoring more than one SDG indicator. There are two such sub-indicators under 2.4.1: Secure Tenure Rights to Land, which is equivalent to SDG indicator 5.a.1; and the Prevalence of Moderate and Severe Food Insecurity, which is equivalent to SDG indicator 2.1.2 (The Prevalence of Moderate and Severe Food Insecurity was included in the later stages of the methodological development process of 2.4.1 as a sub-indicator for the social dimension, replacing "Agricultural Household Income", which was deemed too highly correlated with the "Net farm income" sub-indicator under the economic dimension). Moreover, to the extent that no overall SDG index has been contemplated by the UN Statistical Commission, using the same indicator as both a standalone SDG indicator and a sub-indicator under 2.4.1 carries no risk of biasing any progress assessment.

Based on the results of the Delphi Technique exercise, a selection of nine-sub indicators (Farm output value per farm agricultural area; Net farm income; Access to financial services; Soil health; Water use; Water quality; Agricultural heterogeneity; Wage rate in agriculture; Agricultural household income; Secure rights to land tenure) was initially submitted to the IAEG-SDG in November 2017 as part of a full methodological proposal on SDG indicator 2.4.1. While the IAEG-SDG commended the FAO proposal, it requested that a sufficient number of pilot studies in different countries were conducted to demonstrate its feasibility. It also requested FAO to review once again the list of environmental indicators, by adding pesticide and fertilizer pollution while deleting nitrogen pollution, which moved too slowly to be of much use for monitoring progress within the timeframe of the 2030 Agenda.

Taking this feedback from IAEG-SDG member countries into account, FAO launched a series of pilot studies in Bangladesh, Ecuador, the Kyrgyz Republic, Rwanda and Belgium to collect more robust evidence on the measurability and relevance of the selected sub-indicators. Additional cognitive tests on the formulation of the proposed questions to collect relevant data in national farm surveys were conducted in Bangladesh, Mexico and Kenya.

On the basis of the results of the pilot tests, which confirmed its universal feasibility and relevance, FAO consolidated the methodological proposal for SDG indicator 2.4.1 that comprises a total of 11 sub-indicators (see Table 1). These sub-indicators address the minimum objectives that, according to well established literature [5,11], a food production system would need to satisfy to be defined as sustainable. For instance, the management of fertilizers and pesticides responds to the objective of minimizing the use of off-farm, external and non-renewable inputs. The adoption of biodiversity-supportive practices ensures the incorporation of natural processes, such as crop rotation, nutrient cycling, and pest/predator relationships into the agricultural production process. A positive net farm income demonstrates that the agricultural activity is profitable and can ensure maximum self-reliance among farmers and rural people. The presence of key risk mitigation mechanisms can also reduce vulnerability among farmers. Long-term sustainability of production levels is reflected in a steady or increasing farm output value per hectare, on one side, and the conservation of soil, water, and biological resources, on the other. Similarly, secure tenure rights over land reflects equitable access to the key productive resource for most farmers, while progress towards social justice is captured in the application of a minimum wage rate and a decreasing size and depth of food insecurity.

Table 1. The 11 sub-indicators of SDG indicator 2.4.1.

No.	Sub-indicator	Sustainability dimension
1	Farm output value per hectare	Economic
2	Net farm income	Economic
3	Risk mitigation mechanisms	Economic
4	Prevalence of soil degradation	Environmental
5	Variation in water availability	Environmental
6	Management of fertilizers	Environmental
7	Management of pesticides	Environmental
8	Use of biodiversity-supportive practices	Environmental
9	Wage rate in agriculture	Social
10	Food insecurity experience scale (FIES)	Social
11	Secure tenure rights to land	Social

The Scale of the Sustainability Assessment

Regarding the scale of the sustainability assessment, a key finding from the literature is that a smaller scale of assessment generally offers greater chances of identifying appropriate criteria for measuring progress. As Pretty [5,11] argues: "At the farm level, it is possible for actors to weigh up, trade off and agree on these criteria for measuring trends in sustainability. But as we move to high levels of the hierarchy, to regional, national and international levels, it becomes increasingly difficult to do this in any meaningful way." By contrast, drilling down to the individual field would probably be an overreach and not practical: even though the use of a single field, for example, may be uneconomic or unsustainable, the farm may remain economically viable. Conversely, fields on a farm may do well agronomically, but poorly in economic terms as a result of low commodity prices or high production costs.

The choice of the assessment scale is also to a large extent interlinked with the choice of the Data Collection Tool (which will be described in the next section). Thus, the farm is chosen as the main scale for the sustainability assessment not only because it makes conceptual sense according to the relevant literature, but also because it fits ideally with the data collection tool selected for the indicator, *i.e.*, the farm survey. Having chosen a different scale of assessment for the indicator as a whole, or choosing different scales for the different sub-indicator, would have entailed the use of multiple data collection tools and eventually the need to combine data from diverse assessment levels, something that was considered, but eventually rejected as explained below.

The Data Collection Tool

The FAO proposal identifies farm surveys as the preferred measurement instrument for collecting data on SDG indicator 2.4.1, even if it provides countries with the option (but not the obligation) of using alternative data sources to farm surveys, including geospatial/remote sensing data or other monitoring systems to capture environmental data, under strict criteria. Requiring the use of different data sources for different sub-indicators would have provided more flexibility to national statistical agencies, but this was rejected by developing country members of the IAEG-SDG for the lack of good alternative data sources in their statistical system and for the methodological difficulties posed by data integration.

It is clear that although some types of farm/agricultural surveys are run in many countries, none would automatically be equipped to collect the necessary information for SDG indicator 2.4.1, but would rather need to be adapted. According to a recent FAO survey, 60.2% of countries run farm/agricultural surveys with some regularity (at least once every five years), and 73.4% of countries have requested substantial technical assistance from FAO specifically to upgrade such existing data collection tools and to make them fit for SDG monitoring purposes. To this end, FAO, in collaboration with IFAD and the World Bank, has already been working towards the establishment of a harmonized programme of Agricultural and Rural Integrated Surveys (AGRIS) that can form the basis for collecting data on indicator 2.4.1. Through this program, methodological guidelines on how to conduct cost-effective and modular farm surveys that can monitor a number of agricultural related SDG indicators, including 2.4.1, are being provided to countries, together with technical support for their implementation.

The main limitation of using one single data collection tool to cover all 11 sub-indicators is that not all sub-indicators can be operationalized as impact/outcome measures (the ideal scenario); rather, some sub-indicators would need to be gauged through an assessment of the relevant farming practices [7]. This is especially the case for measuring the environmental impacts of agriculture, which is typically done through monitoring systems like soil and water sampling. By contrast, a farmer cannot possibly provide quantitative analyses of soil content or water quality. Using a farm survey instrument, instead of environmental monitoring systems, therefore implies moving from measuring outcomes/impacts to assessing trends in farming practices that any farmer in any corner of the world could easily recall.

The Sustainability Criteria for Each Sub-Indicator

Another challenging dilemma was whether and how to set criteria to determine whether each sub-indicator was being fulfilled. On one end of the spectrum, the dynamic nature of sustainable development, as highlighted by some scholars, would suggest that only measuring trends was possible, and that comparisons across countries would need to be limited to assessing whether development is progressing or regressing. This, however, is not a satisfactory solution because countries with a very poor starting point in terms of agricultural sustainability could be expected to make greater progress, and thus achieve a higher indicator score, whereas countries with already high standards would find it harder to make progress, and thus would end up achieving a lower score.

FAO's final methodological proposal therefore determines three fulfillment levels for each sub-indicator: a "sustainable" (green) and "unsustainable" (red) level, as well as an intermediate "acceptable" (yellow) level introduced to reflect an intermediate situation where a trend may be stable, or some minimum steps have been taken, though still without achieving all criteria for deeming a farm "sustainable". For instance, if a farm has access to credit, insurance, and has demonstrated on-farm diversification, it can be marked "sustainable" from the perspective of resilience and risk mitigation (sub-indicator 3). By contrast, if it has access to only one of these three mechanisms it is deemed "acceptable", while access to none of these mechanisms would place it in the "unsustainable" category. In this way, it is easier for countries to demonstrate progress towards farm sustainability, as opposed to a situation where a binary distinction, such as acceptable/unacceptable, were applied.

It should be noted that not all criteria are underpinned by a universal scientific consensus. While it would be difficult to dispute that consistent negative profitability or a declining trend in water availability indicate non-sustainability, the performance of some sub-indicators is more difficult to assess, taking also into account the selected data collection tool for indicator 2.4.1. However, as explained above, setting concrete criteria for each sub-indicator is deemed necessary for building a global SDG indicator, ensuring direct comparability across countries. In the case of sub-indicators such as biodiversity, therefore, certain criteria are supported by an extensive, though not universal, body of literature [12–17], and could potentially be open to critique by national or international experts with different views. This is an acknowledged shortcoming of the indicator methodology about which FAO has been transparent with countries, and for which it is open to further revision as soon as a new scientific consensus is reached.

A SUSTAINABILITY INDEX?

Another critical methodological factor for the development of SDG indicator 2.4.1 concerns the modality of synthesizing the multidimensional information in one single value. The literature review suggests that the construction of "agricultural sustainability measurement indices" ([5], p. 32–26) is a common approach for dealing with the complexity of the issue and trying to capture information from multiple indicators. Moreover, the development of composite indices has recently been promoted as the best solution for measuring progress towards the entire SDG monitoring framework by some institutions, such as the Sustainability Development Solutions Network [18]. However, National Statistical Offices have been more reluctant to adopt this approach in the context of the IAEG-SDG process. The methodological weaknesses that composite indices suffer from in terms of choice of indicators, correlation among them and subjectivity of the weighting attributed to each constituent indicator are well known. More importantly, composite indices do not allow one to quantitatively assess the exact degree of progress - one of the essential characteristics that SDG indicators must have. While composite indices are quantitative values, they nonetheless remain qualitative (ordinal level of measurement), as differences over time in the index value cannot be interpreted meaningfully. Last but not least, they fail on the critical criterion of relevance [19,20], *i.e.*, of being "relevant to policy making and sensitive to policy interventions at the appropriate level" [7]. In particular, a composite index cannot reveal which specific component has contributed to the overall result and therefore which specific policy measure may or may not have contributed to determine the positive outcome.

The approach selected for SDG indicator 2.4.1, instead, is to measure all relevant dimensions of sustainability at once for the same statistical unit (*i.e.*, the farm) and then display results through an indicator dashboard at the national level. Ideally, to obtain the proportion of agriculture area that is sustainable, the assessment of sustainability would be made across all sub-indicators for each individual farm that is part of the sample. The farm would then be assigned a sustainability level that is the most constraining across all sub-indicators, and the results would then be aggregated across farms at the national level. In the hypothetical example of the state of Molvania in 2018, illustrated in Figure 1, profitability is the most constraining factor as it affects 40% of the agricultural area of the country, which hence determines the overall proportion of a country under sustainable agriculture. This is the most accurate way of measuring agricultural sustainability in a scenario where one single data collection instrument (the farm survey) can be used to collect information on all sub-indicators for a given area.

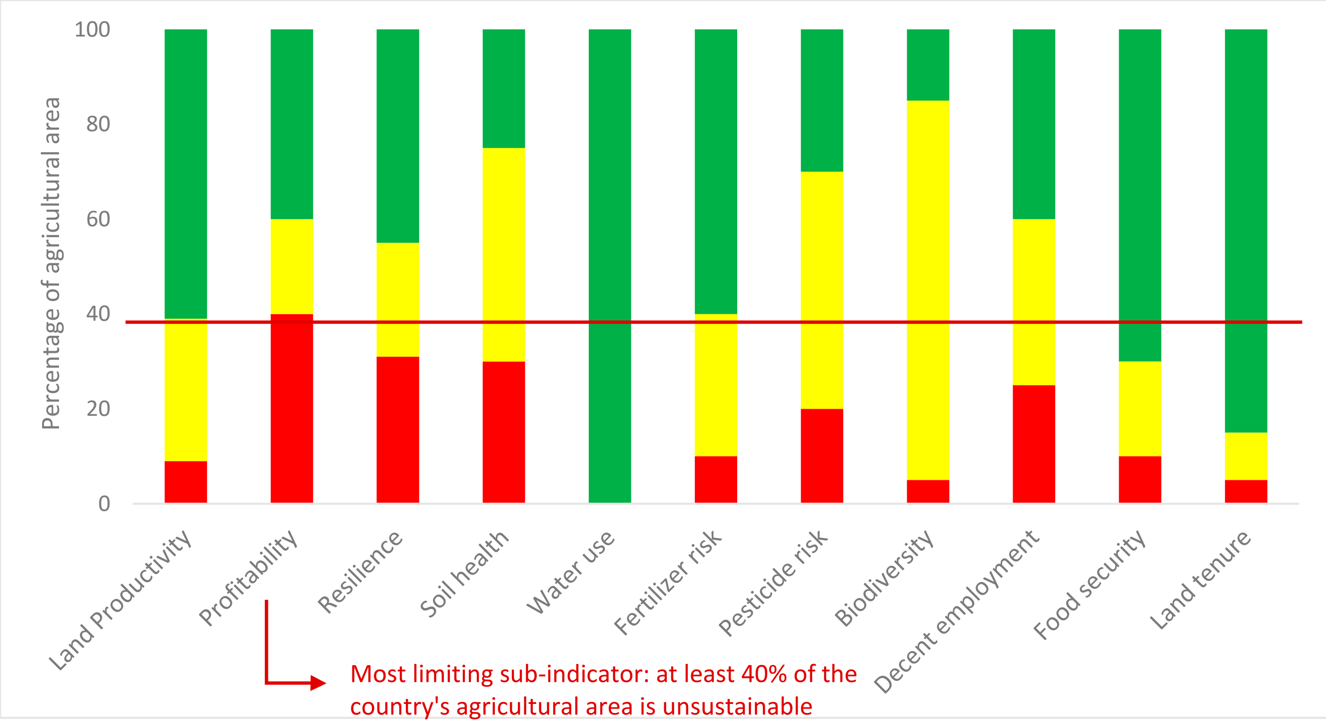

Figure 1. Percentage of agricultural area under productive and sustainable agriculture in Molvania in 2018: the dashboard approach.

Nonetheless, as explained above, the methodology does allow for the use of alternative data sources under certain conditions, and, in any case, it is not strictly necessary for all sub-indicators to be collected simultaneously by one single farm/agricultural survey. In this scenario, where different sources may be used to collect information on the different sub-indicators, indicator 2.4.1 is instead derived from a dashboard at country level, which is pegged to the results of the sub-indicator that has the lowest sustainability performances. This requires calculating, amongst all sub-indicators, the one that has achieved the highest level of unsustainability at the country level (see example above). It should therefore be noted that, while the national-level dashboard offers a tool for easily reporting indicator 2.4.1, implementable across a variety of data collection methods, it will systematically over-estimate the proportion of agricultural area under productive and sustainable agriculture, compared to a farm-level dashboard approach. The reason is that different holdings will likely be categorized as unsustainable across different sub-indicators, but this information is lost by aggregating individually at the national level.

CONCLUSIONS

The above considerations formed the core structure of FAO's final methodological proposal submitted to the eighth IAEG-SDG meeting, 5–8 November in Stockholm. A round of country case studies had already demonstrated the feasibility of the proposed approach, *i.e.*, collecting all the necessary information for measuring each of the eleven sub-indicators in a farm survey and determining their level of fulfillment at farm level. The stakes were high, not only from a narrow reputational point of view for FAO as a proposing agency, but because measuring sustainable agriculture was indispensable for the whole construct of SDG 2, which not only called for ending hunger and improving nutrition, but also for "promoting sustainable agriculture". Indeed, incorporating sustainable agriculture into a new SDG 2 was seen by many as a major leap forward for the SDG framework compared to the previous Millennium Development Goals, and testimony to the new Goal's truly "integrated" nature. Although a few IAEG-SDG member countries raised some concerns, mainly aimed at certain fulfillment criteria for one of the sub-indicators, the general consensus among member countries favored the FAO proposal, and the methodology for SDG indicator 2.4.1 was thus approved for global reporting.

While in no way presuming to offer a complete measure of all aspects of agricultural sustainability, SDG indicator 2.4.1 provides what academics, statisticians, and policy practitioners could only have imagined a few years ago: a core set of internationally agreed metrics, selected through a methodologically sound process, with thresholds defined using the prevalent scientific consensus, which can provide a thorough—albeit not exhaustive—assessment of the sustainability of each farm on any continent. This was a challenging task requiring a fine balancing act between the complexity of the SDG target, the three dimensions of sustainable development, the multiplicity of sub-indicators involved, and the operational concerns regarding their measurability. These concerns and the choice of the data collection instrument determined most of the specific decisions regarding the scale of the sustainability assessment; the sub-indicators within each dimension; the sustainability criteria to assess each sub-indicator; and the modality of summarizing the results. The knowledge and in-depth appreciation of the rationale behind the methodological approach selected can provide the user community with the necessary tools to better understand the strengths and limitations of this indicator and to adequately interpret its results.

AUTHOR CONTRIBUTIONS

PG and DKN designed the study, analyzed the literature and available experimental data and wrote the paper.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

FUNDING

While the preparation of the paper did not benefit from any financial support, the methodological development of SDG indicator 2.4.1 was funded by FAO corporate budgetary contributions.

ACKNOWLEDGMENTS

The authors are grateful to the FAO technical team supporting the methodological development of SDG indicator 2.4.1, including—but not limited to—Jean Marc Faures, Amy Heyman, Asfandiyar Khan, Francesco Tubiello, Flavio Bollinger, Irene Hoffman, Riccardo Biancalani and Yonca Gurbuzer.

REFERENCES

1. FAO. SDG Indicator 2.4.1. Proportion of Agricultural Area under Productive and Sustainable Agriculture. Methodological Note. Rome (Italy): FAO; 2018.
2. United Nations General Assembly. Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. Resolution adopted by the General Assembly on 6 July 2017. New York (US): United Nations General Assembly; 2017. A/RES/71/313.
3. OECD. Measuring Distance to the SDG Targets: An assessment of where OECD countries stand. Paris (France): OECD; 2017.
4. Ikerd J. Two related but distinctly different concepts: Organic farming and sustainable agriculture. Small Farm Today. 1993;10(1):30-1.
5. Hayati D. A Literature review on frameworks and methods for measuring and monitoring sustainable agriculture. Rome (Italy): Global Strategy to improve agricultural and rural statistics (GSARS); 2017. Technical Report n.22.
6. United Nations General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development. Resolution adopted by the General Assembly on 25 September 2015. New York (US): United Nations General Assembly; 2015. A/RES/70/1.
7. United Nations. Lessons Learned from MDG Monitoring From A Statistical Perspective: Report of the Task Team on Lessons Learned from MDG Monitoring of the IAEG-MDG. New York (US): United Nations; 2013.
8. FAO. SDG Indicator 2.4.1 Percentage of Agricultural Area under Productive and Sustainable Agriculture. Expert Meeting; 2017 Apr 3–5; Rome, Italy. Available from: http://www.fao.org/3/a-br908e.pdf. Accessed on 2019 Sep 16.
9. Linstone HA, Turoff M. The Delphi Method: Techniques and Applications. Glenview (US): Addison-Wesley Educational Publishers Inc; 1975.
10. Habibi A, Sarafrazi A, Izadyar S. Delphi Technique Theoretical Framework in Qualitative Research. Int J Eng Sci. 2014;3(4):8-13.
11. Pretty JN. Regenerating Agriculture: Policies and Practice for Sustainability and Self-Reliance. Washington DC (US): Joseph Henry Press; 1995.
12. Weibull AC, Östman Ö, Granqvist Å. Species richness in agroecosystems: the effect of landscape, habitat and farm management. Biodivers Conserv. 2003;12:1335. doi: 10.1023/A:1023617117780
13. Holzschuh A, Steffan-Dewenter I, Kleijn D, Tscharntke T. Diversity of flower-visiting bees in cereal fields: Effects of farming system, landscape composition and regional context. J Appl Ecol. 2007;44(1):41-9.

14. Abson DJ, Fraser EDG, Benton TG. Landscape diversity and the resilience of agricultural returns: a portfolio analysis of land-use patterns and economic returns from lowland agriculture. Agric Food Secur. 2013;2:2.
15. Belfrage K, Björklund J, Salomonsson L. Effects of Farm Size and On-Farm Landscape Heterogeneity on Biodiversity—Case Study of Twelve Farms in a Swedish Landscape. Agroecol Sustain Food Syst. 2014;39:2.
16. Fahrig L, Girard J, Duro D, Pasher J, Smith A, Javorek S, et al. Farmlands with smaller crop fields have higher within-field biodiversity. Agric Ecosyst Env. 2015;200:219-34.
17. Redlich S, Martin EA, Wende B, Steffan-Dewenter I. Landscape heterogeneity rather than crop diversity mediates bird diversity in agricultural landscapes. PLoS One. 2018;13(8):e0200438. doi: 10.1371/journal. pone.0200438
18. Sustainable Development Solutions Network. Indicators and a Monitoring Framework for Sustainable Development Goals: Launching a data revolution for the SDGs. New York (US): SDSN; 2015.
19. Booysen F. An Overview and Evaluation of Composite Indices of Development. Soc Indic Res. 2002;59:115-51.
20. Moreira SB, Crespo N. Composite Indicators of Development—some recent contributions. In: Jeremic V, Radojicic Z, Dobrota M, editors. Emerging Trends in the Development and Application of Composite Indicators. Hershey (US): IGI Global; 2016.

How to cite this article:
Gennari P, Navarro DK. The Challenge of Measuring Agricultural Sustainability in All Its Dimensions. J Sustain Res. 2019;1:e190013. https://doi.org/10.20900/jsr20190013