When less is more: innovations for tracking progress toward global targets

Todd S Rosenstock1,2, Christine Lamanna1, Sabrina Chesterman1,3, James Hammond1, Suneetha Kadiyala3,4, Eike Luedeling1,5, Keith Shepherd1, Brian DeRenzi6 and Mark T van Wijk2,7

Accountability and adaptive management of recent global agreements such as the Sustainable Development Goals and Paris Climate Agreement, will in part rely on the ability to track progress toward the social and environmental targets they set. Current metrics and monitoring systems, however, are not yet up to the task. We argue that there is an imperative to consider principles of coherence (what to measure), standardization (how to measure) and decision-relevance (why to measure) when designing monitoring schemes if they are to be practical and useful. New approaches that have the potential to match the necessary scale of monitoring, with sufficient accuracy and at reasonable cost, are emerging; although, they represent a significant departure from the historical norm in some cases. Iterative review and adaptation of analytical approaches and available technology will certainly be needed to continuously design ways to best track our progress.

More is more

Social, economic and environmental change has proceeded over the past century at an unprecedented pace, resulting in improved human well-being at the expense of the environment [1,2,3]. In 2015, multiple major international agreements were signed, setting out explicit objectives to end hunger, combat climate change and build resilient societies, amongst many other goals [4,5]. The ability to monitor progress toward the set targets is vital for accountability and adaptive management of these international agreements, although some have argued that targets themselves are usually arbitrary and do not improve performance; rather focus should be on monitoring progress toward achieving the overall goal [6]. Regardless, the design and operation of the required monitoring systems presents theoretical, technical and logistical challenges. For example, the scientific community struggles to find consensus on the best and most practical ways to measure progress for complex outcomes such as food security and resilience for or multiple outcomes simultaneously [8,9,10], despite efforts to describe best practice [11].

Faced with the need to measure changes rapidly across multiple diverse outcomes, monitoring programs are generally following the principle that *more is better* [5,12,13]. That is, monitoring programs will be better able to track progress toward targets if they collect more data, more frequently, across more indicators. Despite advances in the ability to acquire and handle large amounts of data, budgets as well as human and institutional capacity are typically insufficient to deal with the scale and complexity of even current monitoring efforts [14–16] suggesting that the ‘more is more’ approach is overly burdensome. Even if it were possible to measure everything, everywhere, all the time, we may still not have an actionable and useful understanding because the complex nature of the human and natural systems being measured [17].

Given limited resources, monitoring approaches will necessarily reach a practical limit in terms of data that can be collected. Therefore, we argue for increased coherence (coordination in what is measured), standardization (coordination in how measurements are made) and decision-relevance (why measurements are made) in current and emerging monitoring systems. Then we summarize...
advances in metrics and monitoring systems that could help to generate robust and actionable information for both tracking and adaptively managing global progress. Specific terms used in the article have been defined in Box 1.

### More is less

*Coherence* relates to what is measured: the selection of metrics and indicators, which are the core of monitoring systems. Currently, individual organizations, national programs and international agreements select metrics and indicators for monitoring according to their own priorities, resulting in a proliferation of metrics and indicators. More than 2000 metrics are being used in global assessments of agricultural sustainability [18] and 186 indicators are used across 12 major international environmental agreements [19]. Rarely are monitoring efforts aligned despite overlapped in objectives. Only the Sendai’s Agreement and the United Nations Convention to Combat Desertification explicitly link to other initiatives’ indicators and monitoring efforts [20,21]. A lack of coherence among monitoring systems with similar objectives limits our ability to compare progress or outcomes across time or place, which in turn limits the utility of data collected.

Even with coherence in what is measured, there can be a lack of *standardization* in how to measure agreed upon metrics. Oftentimes, a metric can be measured in multiple ways, altering the value and perhaps meaning derived [8*,22,23]. When defining metrics and indicators for monitoring programs, tension develops between rapid and cheap but sometimes overly simplistic frameworks and systems that use the best indicators that science can offer, but which are often too data-demanding and costly for practical monitoring purposes. Objectives and constraints of cost, scale and accuracy collide and easy solutions that meet all three goals are typically unavailable [24]. Differences among monitoring objectives, the interests of data collectors and the perspectives of end users of data can lead to a large divergence in the types of information collected across teams or projects, even where there are similar aims [25].

Considering a scenario where the objectives of coherence and standardization are met, the *decision-relevance* of the data is not guaranteed. That is, we may produce consistent and reproducible information, but without a strong reason why. Evidence from agricultural and environmental monitoring suggests that passive, mandated monitoring schemes, in which data are gathered as a stipulated requirement of government legislation, or a political directive, have had poor success [26]. Such monitoring systems are rarely designed to inform specific decisions, or answer scientific questions that are relevant to specific development contexts [26]. Selection of targets, metrics and monitoring systems is more often driven by opportunism and visibility [27] and the need to report progress towards goals, as opposed to their prospects for supporting practical decisions. Shepherd et al. [28] reviewed 103 agricultural and environmental monitoring systems globally and found few provided a clear mechanism for how the amassed data were going to move from information to action, had a clear conceptual framework or theory of change and were designed with the statistical rigor necessary to ensure internal and external validity of results.

Nutrition and health surveillance approaches used to diagnose systemic risks offer an informative counterpart. First, appraisals are conducted to determine what decisions the collected data is meant to inform, and what knowledge gaps currently prevent effective decision making, ensuring that data collected is meaningful and useful [31]. Then, sampling frames and indicators are explicitly selected based on practical considerations of implementation such as cost-effectiveness, sensitivity and usefulness to inform programming [29–31]. Significant effort is made to build broad coalitions of scientists, governments and other partners to come together and agree on metrics and their implementation such as with the recent development of the Minimum Dietary Diversity for Women metric [32], though the process can sometimes take years to complete.

### Less is more

These challenges illustrate that the global community is far from meeting the monitoring needs of current agreements [33]. Yet somehow, monitoring frameworks need to
develop coherent, standardized and decision-relevant information at scale. Recent innovations in indicator design and selection (what and why to measure) and technology for data collection (how to measure) provide reasons for optimism. Novel approaches designed by decision and data scientists and new applications from the private sector have the potential to revolutionize our ability to acquire, analyze and interpret data for decision-making in a cost-efficient manner and with lower capacity requirements.

In stark contrast to ‘more is more’ philosophy, ‘lean data’ has emerged from the private sector to fill a need for cost- and time-efficient performance metrics that can be used to make decisions. Lean data approaches develop and apply simplified metrics, such as rapid score-card methodologies, to quantify complex multi-criteria indicators and generate high-resolution data on factors important for tracking business development, operations and impact [34,35]. This methodology has generated and applied indicators for poverty (e.g. the Progress out of Poverty Indicator), food insecurity (e.g. Hunger and Food Insecurity Access Scale, HFIAS) and even indicators quantifying environmental sustainability (e.g. the Natural Resource Integrity Assessment). This has enabled the creation of new household survey tools that use these indicators to monitor smallholder farm households across multiple outcomes simultaneously [36,37], generating uniquely multidisciplinary datasets for low cost (Box 2).

A critical decision, however, is which relatively small sets of indicators to track. This is where advances in decision science can contribute. Probabilistic modelling techniques, such as Monte Carlo simulations, Bayesian Belief Networks and value-of-information analysis, construct causal models of the system being monitored through representative multi-stakeholder processes [38,39]. These participatory modelling techniques integrate qualitative and quantitative data on the system, as well as uncertainties, to identify what is important to measure from a decision making perspective, thereby creating the case for decision-relevant indicator sets. These methods overcome serious limitations of using subjective weighted scores on arbitrary scales with no standardization, which tend to increase rather than reduce error in risk assessment and decision-making. Probabilistic techniques are already in use in Earth observation, malaria control and predicting human disease [40,41] though application in development has been limited thus far [42]. Lessons from public health surveillance are also relevant, where much of the monitoring has shifted from outcome monitoring, which is slow and expensive, to monitoring of risk factors, which is much cheaper and quicker and gives earlier warning on likely trends [43]. For example, monitoring of heart disease is now heavily focused on behavioural risk factor surveillance, such as telephone interviews on diet and exercise habits.

After thoughtful selection of a relatively small and standard set of indicators to monitor, the challenge becomes how to monitor them at the desired spatial and temporal resolution. The rapid increase in mobile technologies and—equally important—mobile penetration now allow unprecedented data collection in almost all parts of the world. Household survey data collected with mobile technology reduces costs, increases data quality and reduces time between collection, analysis and decisions by comparison to conventional paper surveys [44]. Using text messaging, interactive voice response and call centres for live voice calls can further facilitate collection of data at the temporal frequency and spatial scale needed for decision-making [45]. Because of this, mobile data collection is rapidly becoming the norm [46], evidenced by the recent increase of open source tools and proprietary add-ons available in the market for almost every sector.

Not all indicators need be collected via survey, however. Remote sensing offers additional opportunities to reduce the monitoring burden for tracking social and environmental change. Though remote sensing has been commonplace for some indicators such as deforestation and land degradation [47,48], new applications allow tracking

---

**Box 2 Impact of method selection on cost, accuracy and scale of data collection.**

Data on welfare indicators of rural households can be collected in many different ways. Household level surveys, one widely used approach, try to characterize both agricultural activities and management strategies while at the same moment trying to gain insight into welfare indicators like food security and nutrition. Surveys range from rapid appraisals, usually less than one hour, to extensive questionnaires sometimes taking multiple sessions to complete. Each approach serves a different purpose. But it is important to realize the length and complexity of the survey have consequences for costs and data quality. Costs of administering a targeted one hour survey will be on the order of 20 USD per household, whereas a longer 4 hours survey costs roughly 50 USD per household, while the extensive World Bank Living Standard Measurement Survey—Integrated Survey on Agriculture is likely to be 100-200 USD per household [37,45**,** 69]. In terms of quality, data quality decreases with increasing survey duration. A golden rule for survey length is not to surpass the one and a half hours. Beyond this length, the attention span of the interviewee declines with consequential effects on the quality of the answers [80].

There are relatively easy ways to assess the overall data quality, either by looking at individual answers, for example at reported yield levels, field sizes and farm gate prices, or at more integral indicators like food self-sufficiency or food availability, which gives a quick insight into the potential of the farm household to generate enough food energy to feed the family [25]. Two problems with household data are commonly encountered (Figure 1): (1) An unrealistically large gap between potential supply and household energy need indicating that the survey is missing essential information that contributes to overall income or food consumption and (2) A substantial over-estimation of consumption of crop and livestock products indicating systematic problems with yield and consumption estimates.
Food self-sufficiency and availability analyzed for 200 households in northern Ghana surveyed in 2012 [Data published in Ref. 70]. The households are ordered on the x-axis according to their food availability score, with the dashed line equal to food self-sufficiency. Households on the left show a severe food gap, while on the right people interviewed are consuming 3–4 times what they need in terms of calories; both results are highly improbable and point to concerns with data quality.

Figure 1

Food self-sufficiency and Other activities

Food availability ratio

0 2 4 6 8 10

Current Opinion in Environmental Sustainability

of economic development (based on artificial light seen from space), ecosystem services (from plant functional traits, soil properties, carbon stocks, biomass, freshwater quality), infrastructure status (such as pavement area and damage), and even human health [49–55,56–57]. A case has already been made to use remote sensing to monitor the Aichi Biodiversity Targets set by the United Nations Convention on Biological Diversity [58]. With the number and diversity of applications increasing at the same time as the cost of high-resolution satellite imagery products is decreasing, remote sensing is emerging as a clear opportunity to monitor not only traditional biophysical values but social change as well.

One of the most disruptive innovations in data collection goes beyond remote data collection (e.g. mobile and remote sensing), and instead empowers ordinary citizens to be the agents rather than just the objects of monitoring efforts. Two such approaches, crowdsourcing and citizen science are increasingly being used to monitor everything from water quality, to biodiversity, conflict, road traffic and other indicators [59–62]. Participatory approaches are most effective when combined with mobile technologies to allow real time monitoring at unprecedented scales at little to no cost. For example, crowdsourcing was used to validate a global map of cropland and farm size based on individuals’ assessments of Google Earth images to evaluate accuracy of classification [63]. In addition to reaching scale, these platforms allow monitoring of rapidly changing processes such as conflict, disease transmission and disasters (Figure 2). Though the mode of data collection is new, the collaborative and mutually reinforcing social innovations of crowdsourcing and citizen science, build on earlier monitoring successes. Many REDD+ project have promoted and used community engagement to protect tropical forests [64,65] which propelled citizen science to the forefront of monitoring sustainability [45+]. Crowdsourcing may well emerge as the common mechanism for monitoring generally and specifically for tracking the SDGs [61+,66].

Less is less

Advanced analytical approaches and remote technology offer great potential to revolutionize data collection and analysis for a wide variety of indicators allowing near real-time input into decisions. Whilst promising, the approaches have limitations.

Many innovations require a shift in our way of working. At the most basic level, the suggestions above necessitate a shift from paper-based to electronic surveys, which would come with concomitant challenges in data security, storage, sharing and archiving. Furthermore, even using simple and pervasive technologies and mobile devices, implementation may be limited by the capacity of enumerators, survey administrators or institutions. Some parallels may be drawn from remote sensing where there has been significant activity using these technology for monitoring for some time, yet capacity to use these tools in developing countries is still wanting [14]. The transition, however, may not only be practical but also philosophical. For example, Bayesian statistics and inference represent a departure from the dominant paradigm of statistics, one not typically taught or promoted until recently, and so has both a steep learning, but also acceptance, curve [67+].

Before use, risks arising about data quality need to be mitigated. For example, when using mobile data methods, the data are often limited to persons that have access to mobiles. Though penetration is increasing significantly across much of the world, there are still large numbers of people, especially in rural Africa that either do not have access to mobile phones or network coverage is too poor to contact them reliably. This could result in the collection of biased data [68]. When using crowdsourcing and citizen science, despite some positive results [69], assessing the quality of the provided estimates and reporting bias is difficult. With remote sensing, approaches need to give sufficient attention to calibration and validation using ground measurements with statistically valid sampling schemes for the target areas to provide robust estimates. And if monitoring risk factors versus outcomes, clear
relationships need to be established via research between risk factors and outcome variables, which are often unavailable. In all cases, the scientific community needs to come together to highlight challenges, set conditions of best practice and to provide guidance to governments, private sector and other programs.

Conclusions
The success of global initiatives will in part depend on developing coherent, standardized and decision-relevant monitoring systems. There is clearly a long way to go. The upside is that the political will, funding and innovations are becoming available. The scientific community and private sector must continue to innovate, bring together lessons learned from pilots and build partnership models that foster sustainability of these initiatives. The development community needs to embrace the new approaches being developed that have the potential to match the necessary scale of monitoring, with sufficient accuracy and at reasonable cost, although they represent a significant departure from the historical norm. As the world changes, however, so will the best approaches to meet monitoring demands. Iterative review and adaptation of analytical approaches and available technology will almost certainly be needed to continuously design ways to best track our progress.

Acknowledgements
We thank colleagues including E. Girverz, C. Corner-Dolloff, E. Wollenberg, M. Richards, M. Rufino, J. Tanega, and E. Kumpel, amongst others, who helped develop these ideas through conversations over the past few years. The Surveillance of Climate-smart Agriculture for Nutrition (SCAN) project funded by UK Aid through the Innovative Metrics and Methods for Agriculture and Nutrition Action (IMMANA) supported the writing of this manuscript. We acknowledge the CGIAR Fund Council, Australia (ACIAR), Irish Aid, European Union, International Fund for Agricultural Development (IFAD), Netherlands, New Zealand, Switzerland, UK, USAID and Thailand for funding to the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), specifically the Partnerships for Scaling Climate-Smart Agriculture Product (P4S), and the Research Program on Water, Land and Ecosystems (WLE) which provided funding for some authors (TSR, CL-CCAFS and EL, KS-WLE).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


Authors update the seminal ‘planetary boundaries’ framework and provide a status report of environmental condition relative to the boundaries.


The paper details some challenges, such as providing a cost estimate, for monitoring the SDGs.


This paper proposes a decision-focused approach to monitoring the SDGs using probabilistic models of alternative actions and value-of-information analysis to define high-value metrics that have the potential to improve decisions and performance.


The authors demonstrate the importance of understanding indicators, and their measurement, for deriving meaning as data show that food insecurity indicators generate significantly different values with similar context.


The author reports on key insights from a seminar workshop on monitoring sustainable development. It is among the first to highlight the importance of coherence, standardization and decision-relevance in monitoring.


17. Perretti CT, Munch SB: Regime shift indicators fail under noise levels commonly observed in ecological systems. Ecol Appl 2012, 22:1772-1779.


43. Use of Bayesian Networks for the development of a generic probabilistic impact model for agricultural development interventions.


This paper explains how principles common in nutrition and health surveillance can be translated into the land sector, which historically has taken the outcome monitoring versus risk monitoring approach.


The authors provide a detailed and state-of-the-art analysis of the many obstacles to using mobile technologies for research and data collection ranging from technical to cultural issues.


The authors demonstrate the potential of using data exhaust from mobile phones to fill gaps in socioeconomic data not available in developing countries.


The paper examines the mode of data collection for the Aichi targets through a remote sensing lens and demonstrates the effectiveness of remote sensing to monitor 11 of the 20 targets, which has potentially large implications for tracking progress.


In a short opinion, authors present the case that an alternative approach to monitoring SDGs is needed.


The book describes the historical path of Bayes’ Theorem and Rule in vivid detail. This book should be required reading for anyone interested in Bayesian approaches.


A working paper that discusses many of the challenges and opportunities for using mobile phones for surveys. The authors promote the technology for this purpose but not without a clear eye toward the limitations and implications for research and monitoring.


A discussion of a number of mobile-based efforts by the World Food Program for monitoring provides a good assessment of the use of a number of methods.


This paper describes one of the seminal uses of Geo-Wiki, a landuse crowdsourcing tool, that has become increasing common in scientific investigations.


http://dx.doi.org/10.1016/j.biecon.2016.08.011.