

MONITORING RESULTS OF AGRICULTURAL AND FOOD SECURITY PROJECTS: THE INDICATOR CHALLENGE

A results framework, which specifies objectives, indicators and targets, is the basis to monitor progress of development projects. Selecting appropriate indicators and measuring results, however, can be extremely challenging. If indicators are not well selected or data quality is low, their use to support decision-making and to report on results may lead to wrong conclusions. This article provides an overview on criteria to select indicators and explores the main challenges when measuring results in agricultural and food security projects.

By Sarah Holzapfel

Various issues and criteria need to be taken into account when selecting indicators and organising the data collection process. A few of the most important ones are listed in the Box.

- **Direct:** The indicator clearly represents the intended result.
- **Objective:** The indicator is unambiguous about what is being measured.
- **Use of country systems:** The indicator is part of a country-led results framework, and data is provided by country-level M&E systems.
- **Owned:** Partners and stakeholders agree that the indicator is useful.
- **Attributable:** The indicator can be plausibly associated with the intervention.
- **Practical:** Data can be collected easily, on a timely basis and at a reasonable cost.
- **Reliable:** Data is consistent and comparable over time.

There are often trade-offs between the criteria, which have to be carefully weighted. In the following, the most common trade-offs as well as the challenges when defining indicators and measuring results are described.

DEFINING INDICATORS

Indicators often include a variety of concepts, which have to be defined and clarified. One example is an indicator formulated as “the hectares of land cultivated by smallholders (20 per cent women) under sustainable agricultural practices have increased by 10,000 ha”.

First, the project has to define which practices are understood as sustainable and who is considered a smallholder. Second, it has to be clarified what “20 per cent women” (of 10,000 ha) means. Does the 20 per cent refer to land



Indicators have to be clearly designed. Does “20 per cent women” (of 10,000 ha) refer to land cultivated by female-headed households or to land over which women have primary decision-making rights?

cultivated by female-headed households or to land over which women have primary decision-making rights? Ambiguities often arise with respect to indicators, which measure agricultural productivity, income or food security. If, for example, an increase in productivity by 30 per cent is set as a goal, the project has to specify whether the increase is expected among all households in the intervention area, as a mean among all direct (and indirect) beneficiaries or only among a share of the beneficiaries. These different approaches imply a high difference in the level of ambition, especially if the intervention area is large.

USE OF COUNTRY SYSTEMS

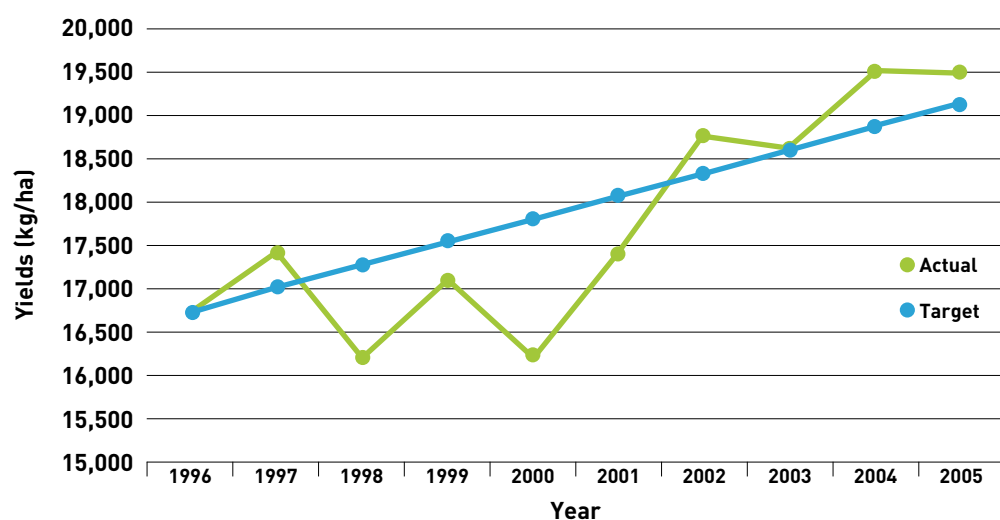
As a best practice and in line with the aid effectiveness agenda, country-led results frameworks and Monitoring & Evaluation (M&E) systems should be used as a common tool among all concerned actors to assess project performance. In practice, donors often introduce their own, parallel monitoring systems and only partly rely on country-level systems. While the objectives of development interventions are usually derived from country-results frameworks, the percentage of indicators that draw on data from country-level M&E systems is much lower. Indicators are often determined by donor priorities and selected with limited involvement of partner countries. However, country systems can best be strengthened by using them. Studies show that parallel monitoring systems create inefficiencies such as parallel and uncoordinated data collection efforts. Harmonising results measurement systems among partner countries and donors could increase data availability and quality.

DATA AVAILABILITY AND DATA QUALITY

In general, measurement errors for key outcome indicators in agriculture, such as yields, gross margins and smallholder incomes, tend to be particularly high. Among the diverse problems when measuring smallholder crop areas, production and crop value are ill-defined or even non-existent plot boundaries, intercropping, irregular planting density, non-standardised measurement units (e.g. bunch or pieces), a high share of subsistence production and significant post-harvest losses.

Data availability and quality when relying on data from partner countries' own systems are often low, which is one of the main reasons why donors set up parallel M&E systems. In

Detecting a trend in maize yields



Source: Global Donor Platform for Rural Development (2008, p. 28)

many developing countries, continued underinvestment in statistical systems has led to low technical and institutional capacity for producing data from administrative systems or household surveys. As a result, there are no regular, reliable estimates of many key indicators that are commonly used to monitor progress in agricultural and food security development projects. Often, data is only available at the country level but not for the specific area where the project is active. A related problem is that although national statistical agencies, donors, universities and research institutes carry out a high number of surveys, data is often not published or shared. One positive example of how to address this problem is the introduction of the Open Data Policy and Development Data Library by the United States Agency for International Development (USAID).

To be comparable over time, data has to be collected using the same data collection methods and sampling frames. Often, however, a variety of data sets and methods are used within the course of a project. For example, as baseline data, national survey data are used. Later on, a survey is carried out among project beneficiaries using a different sampling frame and questionnaire.

BASELINE DATA

When collecting baseline data, there is often a trade-off between collecting data for the purpose of planning a project and for monitoring and evaluating it. To plan a project and to set targets, data should be collected as early as possible. At a very early stage of project planning, however, the intervention area might

not yet be narrowed down, and it might not be clear who will be targeted by the project. This creates challenges when results of a project are monitored through follow-up surveys. For example, it might be that baseline data was collected among households who in the vast majority have not participated in the intervention.

One weakness often observed in results frameworks is that projects do not collect baseline data for indicators such as “hectares of land under sustainable agricultural practices” or “number of households who have adopted an innovation”. Instead, it is assumed that the value is zero before the start of the project. This is unlikely to be true and can lead to an overestimation of results.

MEASURING LONGER-TERM RESULTS AND THE ATTRIBUTION PROBLEM

Development co-operation aims to contribute to long-term goals, such as Sustainable Development Goal (SDG) 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture. When monitoring development co-operation projects, the aim is therefore to go beyond measuring outputs, i.e. products, capitals goods and services which result from a development intervention, and their short-term effects (e.g. access to and use of services, adoption of innovations). There are, however, two problems when measuring medium- and long-term results.

First, medium- and long-term results are influenced by a variety of external factors, and it is challenging to assess whether observed chang-



A lack of data availability frequently complicates project evaluations. Often, data can only be obtained at country level, and not for the specific area where the project is active.

Photo: Arne Hoel/World Bank

es are a result of the intervention or are due to other influencing factors. For example, agricultural yields or smallholder income are highly variable over time and subject to a variety of external factors such as climate variability and changes in world market prices. Rainfall, or the lack of it, has a particularly strong influence on production, and consequently, yields can vary considerably from one year to the next, especially in non-irrigated conditions, which are prevalent in sub-Saharan Africa. Attribution is practically impossible if country-level data is used to measure progress of an intervention, which is only active at a sub-national level. The challenge of attribution implies that

existing data on the SDG indicators, which is often only available at country level, is in most cases not suitable to measure the results of an intervention.

Second, agricultural development programmes usually take several years until changes at outcome level can be observed because the process of promoting and adopting innovations is time-consuming. Usually, several harvest cycles are needed until changes can be observed. In addition, with the high variability in production, it can take several years until trends in yield levels can be detected (see also Figure on page 27).

EVALUATION DESIGN

The use of a before-after-comparison among the target group or a comparison of outcomes among programme participants and non-participants is common. However, such methods are not suitable for indicators at medium-term outcome or impact level and may lead to wrong conclusions. A before-after-comparison shows changes among the target group and does not provide information on the causal effect of an intervention. For example, a smallholder's income might have also increased in the absence of a programme – because of his or her own efforts, favourable climatic conditions or high market prices.

It is also problematic to compare outcomes among the treatment group with outcomes among a control group if participation in a programme is based on preferences, decisions and unobserved characteristics (self-selection). Because of self-selection, the control group is unlikely to be statistically identical (on average) to the treatment group. As a result, differences in outcome variables cannot be attributed to the project. It is probable that the two groups would have performed differently even in the absence of the programme.

Causality can be established if rigorous impact evaluation techniques are used, such as randomised assignment, difference-in-differences, and matching. While these are used to evaluate projects, they are only rarely applied for monitoring development projects because they are complex and time-consuming and can be very costly.

If medium-term outcomes and impacts cannot be attributed to a project, are indicators at that level suitable to monitor progress of development co-operation projects? Often, it is advisable to monitor short-term outcomes instead, for example, if studies have shown that the adoption of a specific practice promoted by a project leads to desired changes at outcome and impact level. Moreover, it can generally be expected that if individuals have adopted a practice or technology for a longer time, it has positively affected their lives. Otherwise, disadoption would be the logical consequence.

Sarah Holzapfel, PhD, is a research fellow at the German Development Institute in Bonn.

In her research she focuses on how to increase the results orientation of agricultural, rural development and food security interventions in line with the aid effectiveness agenda.

Contact: Sarah.Holzapfel@die-gdi.de

RECOMMENDATIONS

- **Indicator definition sheets** should be developed for each indicator in results frameworks explaining all relevant information allowing for a proper analysis and interpretation of data (e.g. definition of concepts, data collection method, evaluation design).
- Donors should strengthen their support to **country-level M&E systems**, use them as much as possible, and **harmonise** their results measurement systems.
- Data collected by donors through surveys or as administrative data should be **openly available, accessible, usable and unrestricted**.
- If **baseline data** is collected, surveys should be carried out only after the intervention area is known and sufficient information on who will be targeted is available. Before, secondary data, expert opinions and qualitative methods can be used to plan an intervention and formulate targets.
- **Indicators measuring medium-term outcomes and impacts** should only be used for the purpose of monitoring if they can be plausibly attributed to the intervention (e.g. in case rigorous methods can be applied), and if changes are measurable within the project timeframe. If this is not the case, short-term outcome indicators are more suitable.