

Africa RISING – How will we know we have been successful?

Overview

Africa RISING (AR) is a research for development program designed to pilot potential interventions for the sustainable intensification of mixed crop-tree-livestock systems and provide data and information that will lead to the better design of development projects. It comprises three linked projects (West Africa, East & Southern Africa and Ethiopian Highlands) with a separate Monitoring, Evaluation and Impact Assessment Component. The primary hypothesis of the Africa RISING Program is that sustainable intensification of mixed crop-tree-livestock systems leads to increased whole farm productivity, which in turn leads to development outcomes (improved welfare) such as improved livelihoods (income, assets, capacity etc.) and better food and nutrition security for those who depend on these systems. It is further hypothesised that a combination of relevant interventions is more likely to increase whole farm productivity than single interventions.

The hypothesis will be tested by implementing baskets of interventions in selected communities.¹ Within a community, interventions will be ‘offered’ to volunteers. The type of interventions -and delivery methods- is expected to vary across time, space, and local context, even across farms depending on the choice of the individual household. They will also vary according to the farm/household typology that will classify farm households ‘sufficiently similar’ in relation to the expected effects of AR. Farming systems analysis and modelling will be used to help identify and target appropriate interventions across different farm types and to perform ex-ante impact analysis. Crop modelling analysis can also be applied.

Methods

To test the hypothesis that Africa RISING interventions lead to improved whole farm productivity and development outcomes, one would need to answer the counterfactual question of “how would farm productivity and development outcomes have fared for farmers who are offered (and accepted) the intervention(s) in the absence of the intervention(s)?”. Since it is impossible to know the answer to this question, one needs to establish a credible group of farmers who would have had characteristics (farm productivity, welfare, etc.) similar to those who were exposed to the intervention(s) but who were not treated by the intervention(s). The specific approach to be pursued for testing the above hypothesis will be guided by the scale, nature, and timing of (planned) interventions by individual research teams, especially since site- and context-specificity and own-adaptation by beneficiaries are integral parts of the Program.

While Randomised Control Trials (RCTs) are becoming the standard way by which the impacts of a new technology can be assessed, such approach is not applicable in the context of Africa RISING, except in some specific cases limited to the delivery mechanisms, rather than the type of intervention *per se*. It is argued that - a) the intervention households² are not selected at random

¹ The definition of ‘community’ varies between countries depending on the local administrative and geographical arrangements.

² *Intervention households* are households in the target communities to which Africa RISING interventions are directly applied; *Non-intervention households* are households in the target communities in which interventions are not applied,

but volunteers (therefore, self-selected) or selected purposively by the researchers: b) the interventions are not unique, but multiple technologies are at play, which additionally vary from community to community and even from household to household; c) the attribution of impact to specific actors or actions is not possible given the multiplicity of actors and partnerships as well as on-going interventions.

IFPRI would like to emphasize that it is possible to design RCTs –and to address potentially non-random refusal of “treatment”- even when some randomly selected subjects refuse to participate; instead of focusing in the average treatment effect (ATE), the RCT will address the intention-to-treat (ITT) impact. In addition, randomization can possibly be done at different levels and stages in the delivery process (e.g., a phased-in RCT design), and the fact that we cannot randomize the intervention households does not completely rule out other possible types of randomizations.³

To correctly assess the extent to which changes in outcomes of interest, if any, can be attributed to Africa RISING research activities, IFPRI strongly believes on the need for designing and implementing an evaluation strategy that ensures measurement of these economic phenomena with accuracy or robustness. A properly and scientifically designed impact evaluation is also necessary for well-informed decisions about scaling up. Unlike project monitoring, which examines and tracks whether targets have been achieved, impact assessment examines how outcomes of Africa RISING beneficiaries have changed as a direct (and, if modelled explicitly, indirect) effect of the program. It seeks to provide cause-and-effect evidence and quantifies changes in development outcomes that are *directly or indirectly attributable* to Africa RISING, and not to other confounding factors.

For Africa RISING projects at an early stage of implementation or for planned interventions, IFPRI believes that Africa RISING team should be open to seizing possibilities for embedding an RCT evaluation design (including a pipeline evaluation⁴), for example, to assess the impact of a specific technology or delivery mechanism.⁵ In the absence of a random selection of target communities and households, various non-experimental designs can be explored to construct a plausible counterfactual group.⁶ For example, if selection determinants are known (or believed to be observable), then various regression-based approaches (e.g., matching) can be employed to construct an acceptable comparison group and mitigate selection bias. If selection determinants are (believed to be) unobserved but are thought to be time invariant, panel data approaches (including simple difference-in-differences) can be employed. When none of the above is possible, the

but which may benefit from spill-over effects of facilitated transfer; finally, *Non-target households* are households outside the target communities but in the same development domain.

³ Of course, this is not to say that *only* RCTs allow for identification of causal effects.

⁴ The idea of this type of evaluation is to use, as the comparison group, people who have applied for a program but not yet received it. The key assumption is that the timing of treatment is random given application. In practice, one must anticipate a potential bias arising from selective treatment amongst the applicants or behavioural responses by applicants awaiting treatment. This effect could represent a great concern in AR, depending on how interventions are conducted.

⁵ For example, in partnership with another USAID-funded project (the Cereal Systems Initiative for South Asia (CSISA) initiative) Lybbert *et al.* (2012) designed and implemented a field experiment to better understand heterogeneity of farmers’ demand for Laser Land Leveling (LLL) services and another field experiment that combines some of CSISA’s new technologies (abiotic stress resistant rice varieties) with weather index insurance policies. Travis Lybbert, Nicholas Magnan, Anil Bhargava, Kajal Gulati, David Spielman. 2012. Farmers’ heterogeneous valuation of Laser Land Leveling in Eastern Uttar Pradesh: an Experimental Auction to Inform Segmentation & Subsidy Strategies, *American Journal of Agricultural Economics*.

⁶ The essential requirement for this evaluation would involve the research teams to carefully document the criteria used to identify households eligible for receiving the intervention(s).

problem of selection bias cannot be addressed and any “impact evaluation” effort will have to rely heavily on the program theory. Qualitative and participatory approaches would therefore build an argument towards plausible association (but **not** causality). These different approaches are of course not mutually exclusive. It should also be noted that the internal validity of the causal evidence will depend on the quality of the match between target and comparison groups, while the external validity of the results will depend on the representativeness of the sample from which the evidence is drawn.⁷

Irrespective of the specific evaluation design, however, target households and communities need to be selected⁸ so as to be statistically representative of households and communities within the IFPRI-delineated ‘development domains.’⁹ Representativeness is necessary (but not sufficient) to ensure external validity of results and assist in informed decision making on scaling.¹⁰ An intervention would need to be offered to ‘enough’ number of farmers to precisely estimates its effect.¹¹ In the absence of a credible and well-thought evaluation approach as well as target households and communities that are not representative of the population they are drawn from, estimates of the effect of interventions on whole farm productivity and development outcomes will be inaccurate and imprecise and, therefore, cannot be extrapolated.

In addition to the direct interventions on the target farms, research teams will put in place mechanism to facilitate extension of the interventions to other farms in the community, which will not be subject to the same degree of study. While measuring potential indirect effects of the program (i.e., spillover effects¹²) and understanding of transmission mechanisms certainly provide insights on how AR operates, careful thought should be given to data requirements for correctly measuring spillovers within the context of the Program.¹³ At the moment, researchers have identified (or are in the process of identifying) “intervention” households in target communities.

Baseline and follow-up data should be collected from “non-intervention” households, which are those in current target communities to which interventions are not applied. If the purpose of data collection from current non-intervention households is to measure spillover effect from research activities, then one also needs to think through how current non-intervention fare during scaling-out of research activities. Specifically, whether the distinction between intervention and non-intervention households within current targets will prevail over the time-horizon of the Program

⁷ Internal validity refers to the accuracy of the evidence, while external validity refers to the generalizability of the evidence to the population from which the sample is drawn or to another “similar” sample or population.

⁸ Selection criteria need to be documented and shared with the M&E team to help inform selection of ‘comparable’ control communities and households.

⁹ The ‘development domain’ refers to the original IFPRI designation from the site selection process, which takes into account locally relevant market, biophysical, and demographic indicators, and is therefore defined in a country-specific context.

¹⁰ External validity refers to the generalizability of results about impact of the intervention(s) on farm productive and development outcomes to other settings.

¹¹ If multiple interventions are being offered in a focus country and in a given period of time but no single intervention is offered to ‘enough’ number of farmers, evaluation efforts will have to focus on assessing the ‘overall effect of Africa RISING’ in the focus country, rather than the effect of the single intervention.

¹² In this document, spillovers refer to a situation where farmers not eligible to receive AR intervention, or who are eligible to receive the intervention but have not received it, benefit from the intervention indirectly through a variety of ways – such as externalities (e.g., when channeled by successful AR farmers), general equilibrium effects (e.g., depressed maize price through increased maize production due to AR interventions), social and economic interactions (e.g., neighbors and relatives interacting with and learning from a successful AR farmer), and behavioral changes.

¹³ Manuela Angelucci and Vincenzo Di Maro. 2010. Program Evaluation and Spillover Effects. *IDB Impact-Evaluation Guidelines Technical Notes*, no. IDB-TN-136 (available [here](#)).

depends on the nature and timing of planned research activities. In this regard, there appears to be at least two options.

Option 1. A research team expects to scale up interventions *within currently identified target communities (current targets) in the future (at time T+1) BUT not sure yet which of the “non-intervention” households will be targeted.* When this is the case,

- i. Baseline **AND** follow-up surveys of (a sub-set of) non-intervention households in current targets will need to be conducted ***before*** the start of scaling activities at time T+1, if there is no guarantee that the “non-intervention” households from whom baseline data was collected would remain (directly) unaffected by the program during scaling.
- ii. If the research team expects the AR activities to have a spillover effect on households outside (but in proximity of) the currently targeted communities, then non-intervention households can be sampled from non-target communities that are ***adjacent*** to the current targets.

Option 2. A research team expects to scale out research activities to *new target (and non-control) communities at time T+1.* When this is the case,

- i. Baseline and follow-up survey of (a sub-set of) non-intervention households in currently identified target communities will be part of the overall baseline and follow-up surveys in currently identified target and control communities.

Measuring success (or failure)

To assess program success, or otherwise, the following information/evidence needs to be generated at the specified scales.

1) What are the processes (technical, social, institutional) by which Africa RISING is improving farm productivity?

Relevant scale: the household/farm scale.

Who is responsible for providing evidence: led by the regional/country project research teams, using a variety of agronomic, systems, participatory action research and other approaches.

2) What are the implications of these productivity-enhancing processes for environmental, social, economic sustainability?

Relevant scale: the household/farm scale, with some natural resource management (NRM) implications at the landscape level.

Who is responsible: led by the regional/country research teams, using a variety of agronomic, participatory systems analysis, modelling, farming systems, and other approaches.

3) What are the impacts of these productivity-enhancing processes on development or research outcomes, at a variety of scales?

There is a strong need to define what development or research outcomes to assess, in order to properly define the type of data to collect at baseline and follow-up.¹⁴ For example, in order to provide evidence on the impact of Africa RISING on child and women nutrition, data are needed on child and women anthropometry. Once the outcomes are clearly defined, then an evaluation strategy needs to be designed and implemented to ensure measurement of program effect on farm productive and welfare with accuracy and robustness.

Sub-Task 1. Household/farm livelihoods scale:

- What effect has Africa RISING had on, for example, farm practices and welfare of the intervention households?¹⁵
- This take will be led by IFPRI, with input from the regional/country research teams on the type and intensity of interventions, selection process of target communities (from within IFPRI-identified development domains) and households, and any other relevant information.
- Regional/country research teams will employ agronomic techniques and farming systems modelling approaches (facilitated by the Wageningen team) to characterize households in target and control communities.
- Baseline and follow-up data will be collected from intervention households (to be identified by the regional/country project research teams) and non-target households in control communities (to be selected by IFPRI, with input from research teams about criteria they used to select intervention households).

Sub-Task 2. Community scale:

In addition to household surveys, community surveys will need to be conducted to provide quantitative evidence on the effect of the Program at community scale. Collection and analysis of qualitative data (in an iterative manner) may also be necessary to generate knowledge about common patterns and themes at community scale. As discussed above, research projects expect to put in place mechanisms to facilitate extension of the interventions to other farms in the community, although the other farmers will not be subject to the same degree of intervention. Nonetheless, correctly accounting for community-level effects of the Program (including direct and indirect effects) may be a challenge without a clear path for the “scaling up” process and the extension.

While the distinction between intervention and non-intervention households within community makes sense in year T, some or all of the households currently identified as non-intervention may be affected by the program at future time T+1, 2, etc. Hence, expectations about an accurate and robust evidence of the effect of the program at the community scale should be reconciled with what can/cannot realistically be delivered within the context of realities on the ground.

¹⁴ As such, the length and depth of the survey tool proposed by the M&E team needs to be evaluated against what the team is expected to deliver, among other things.

¹⁵ Here an actual effect is sought, hence simulation models cannot help much, unless we rephrase in “What effect is estimated (or simulated) Africa RISING to have on [...]”

- This task will be led by the regional/country research teams and IFPRI.
- For Africa RISING projects with clear (and persisting) distinction in the status of households within target and control communities, IFPRI will employ various non-experimental techniques to provide evidence on program effect at community-level.
- Researcher teams will employ a variety of agronomic, farming systems, and other approaches to assess the effect of AR on input suppliers, traders, natural resource management, etc.
- Baseline and follow-up data will be collected, right after the harvesting time, from:
 - Intervention households (to be identified by the regional/country project research teams),
 - Non-intervention households in target communities or communities that are adjacent to target communities (to be identified jointly by regional/country project research teams and IFPRI), and
 - Non-target households in control communities (to be selected by IFPRI)
- IFPRI will conduct structured community surveys to capture community-level characteristics, if needed.

Sub-Task 3. Development domain scale:

To the extent Sub-Tasks 2 and 3 are based on representative households selected, in turn, from representative target communities, evidence from Sub-tasks 2 and 3 will provide insight on the scope for scaling up (similar) research activities to other (similar) households in different communities within the development domains under study. It is worth noting that generating a credible evidence to help inform scaling decisions goes beyond simulation and requires cause-and-effect evidence drawn on statistically representative households.

Application of the survey tool

The proposal is to administer a household survey to:

- a) **All the intervention households** in the target communities (or a sub-sample depending on their number -but this is currently deemed unlikely-).
- b) **A randomly selected group of non-intervention households** in the target communities. For this sampling to be feasible, however, there is a need for:
 - i. Clarity on planned scaling activities (whether the scaling will involve more households within current target communities and, if so, which ones, versus households in new target communities)
 - ii. Clarity on planned “mechanism to facilitate extension of the interventions” and their expected coverage (only within current target communities or including adjacent communities)

Without (i), there is no guarantee that a household identified as “non-intervention” at time T (at baseline) may still remain a “non-intervention” household at T+1 and later on and, if so, the proposed survey of “a randomly selected group of non-intervention households in the target communities” would be flawed. Without (ii), selecting the ‘right’ (statistically similar) kind of non-intervention households (for correctly measuring indirect effects of the program) will be difficult.

- c) **A random selection of households in non-intervention communities** within the same development domain.

At the minimum, the Africa RISING evaluation household survey will be conducted at the start and at the end of the project, with a possible mid-point.

The evidence to be generated by the household surveys and the research teams will help assess sustainable intensification trajectories for different household typologies as they occur, and to inform the development of scaling up and out strategies. In order to generate this evidence, data need to be collected on composition of households, crops grown at the plot level, livestock systems, farm and crop management practices, use of inputs, and key livelihood strategies employed, all dimensions that the survey instrument currently captures. These are crucial data to evaluate sustainable intensification trajectories, and evolution of changes in farm practices within the development domains of interest. The sampling strategy and survey instrument will follow the overall approach described here, and will be further fine-tuned for each country through input from the research teams.

Monitoring Outcomes

In order to assist in the periodic (bi-annual) reporting of the progress with regard to the 9 aggregate USAID FtF indicators and additional project-specific indicators, the HarvestChoice team, working with Spatial Development International (SDI), is currently updating its M&E platform created in 2012. The Africa RISING M&E Platform is a web-based mapping application intended to visualize where development work is taking place, and to match it with a wide suite of biophysical and socio-economic spatial layers. The platform was intended to serve four distinct audiences:

- **Research teams and collaborators**, who can use the website to share data, stories (incl. multi-media content), lessons-learned, and as a one-stop-shop for up-to-date M&E information (**RT**).
- **Harvest Choice's M&E Officers**, who will provide curated maps, datasets and summary reports outlining the approach and methods used, as well as outcomes over time (**HC**).
- **M&E Officers at USAID Offices**, who will have access to both project-level and aggregate information to be further imported into the Agency's own reporting systems (**USAID**).
- The **general public**, who will have access to sections of the website upon approval (**PBU**).

As part of this round of platform update, a number of new user requirements will be embedded in support of AR M&E function. Among other features, users¹⁶ will be able to upload data onto the platform, view summary information (key characteristics of sites and communities) and reports for each action site (including baseline statistics), securely enter and document periodic performance indicators, view monitoring indicators for each action site, and edit project- and site-specific information. In addition, the M&E platform will be linked to AR wiki page to provide users with

¹⁶ Some of the new user requirements (e.g., data editing rights) will be restricted to specific group of users (e.g., M&E officers or CG colleagues).

additional background information about the program and the projects in each mega-site. Annex A summarizes proposed user requirements in support of Africa Rising M&E function.¹⁷

Conclusions

Summing up, the available options and tools for M&E activities are:

- a. **Farming systems characterization** and the analysis of production economics of integrated crop-livestock systems, according to, e.g., the Erenstein & Thorpe (2010) or Tiftonell *et al.* (2010). The key here is to characterize systems that have thus far eluded good economic characterization and good economic analysis. Use of both village and household data is feasible.
- b. **Household livelihood strategies** and poverty characterizations that accompany the AR baseline survey. With good data, there is some descriptive analysis that can be done to inform the project of the precise nature of their interventions, e.g., whether the purposively selected beneficiaries are systematically different than the general or average population, how different households deal with risk, etc.
- c. **Ex-ante impact assessment** and **crop-modeling simulations** that the HarvestChoice team can provide, by using primary household and agro-ecological information to calibrate ex ante impact assessment modeling.
- d. **Specific case studies** or experiments that can be conducted in some specific sites for a unique (or limited set) of intervention(s). We can focus on one field experiment that combines a predefined set of AR new technologies, following the CSISA model in which, e.g., they conducted a specific laser land leveling field experiment (Lybbert *et al.*, 2012).
- e. **Project monitoring** activities on, e.g., recording, tracking, and reporting FTF indicators.

Responsibilities associated with the options above can be assigned to:

- a. the Wageningen team, who will provide analysis and modelling.
- b. the IFPRI team, in collaboration with the research teams, through the Africa RISING evaluation household surveys.
- c. the IFPRI team, in collaboration with our in-house crop modelers.
- d. the IFPRI team, in collaboration with the research teams, and the local M&E coordinators. Precise design and experiment still need to be set, but they will likely revolve around integrating an innovative agronomic/agricultural technology with economic analysis, possibly in an area where only AR is or will be intervening.
- e. the IFPRI team, in collaboration with the research teams, and the local M&E coordinators through the Africa RISING M&E Platform, designed and facilitated by Spatial Development International (SDI).

¹⁷ HarvestChoice team welcomes feedback from CG colleagues regarding user requirements the M&E platform should encompass.

Annex A. Africa RISING M&E Platform User Requirement

Ref #	Requirement/action	Category	Priority	RT	HC	USAID	PUB	Notes	Deadline
R1	Explore AR's missions, rational, milestones, partner profiles and contact information.	general	high	x	x	x	x	HC: provide links to existing AR wiki, and partner contact information for all action sites.	05/17/2013
R2	Visualize (map), print and download lists of AR's action sites	explore	high	x	x	x	x	HC/SDI: current maps need updating. Also make sure we publish existing socio-eco indicators (by district)	05/17/2013
R3	Visualize, print and download AR's stratification layers on a map (incl. legends) for all target countries	explore	high	x	x	x	x	HC/SDI: current maps need updating.	06/14/2013
R4	Visualize, print and download development domains (i.e. strata) on a map (incl. legends) for all target countries	explore	high	x	x	x		Joe: provide stratification layers <i>by country</i> (broken out by district and/or EPA ¹⁸ , and/or section)	06/14/2013
R5.1	View summary information (key characteristics of sites and communities)	reporting	high	x	x	x	(x)	Carlo/Mel: provide a site report	06/28/2013

¹⁸ Extension Planning Areas (applies to Malawi)

Ref #	Requirement/action	Category	Priority	RT	HC	USAID	PUB	Notes	Deadline
	and reports for each action site (incl. baseline statistics)							template.	
R5.2	View monitoring indicators for each action site	reporting	high	x	x	x	(x)	<u>Strong deadline</u> (but no editing features needed). HC: provide data in tabular format for ingestion.	07/26/2013
R6	View summary information and reports for each country, mega site and for the entire AR program	reporting	high	x	x	x	(x)	HC: provide summary template by mega-site and for the entire AR program. Analytics will be performed off-line.	08/30/2013
R7	Filter (query) action sites by key characteristics and by spatial location	explore	low	x	x	x	x	HC: provide list of possible filtering parameters (in addition to topics/sub-topics – likely candidates are technologies, commodities and performance measures)	
R8	Securely edit information on project sites (incl. multi-media material)	content management	med	x	x			Mostly link back to information on mega-sites available on AR wiki	08/28/2013
R9	Securely enter and document periodic performance indicators	reporting	high	x	x			<u>Strong deadline</u> . We want to at least demo that feature to M&E team.	08/28/2013

Ref #	Requirement/action	Category	Priority	RT	HC	USAID	PUB	Notes	Deadline
R10	Show/hide project attributes (descriptions, indicators) and keep track of changes	content management	low		x				
R11	Create, update user profiles and invite new users to join in	user management	low		x			By end August only HC will need editing rights.	
R12	Control view/edit access to sections of the websites	user management	low		x				
R13	Website analytics (GA)	content management	high		x			Maria: Add a new site to HarvestChoice GA account	05/17/2013