About FARA

FARA is the Forum for Agricultural Research in Africa, the apex organization bringing together and forming coalitions of major stakeholders in agricultural research and development in Africa.

FARA is the technical arm of the African Union Commission (AUC) on rural economy and agricultural development and the lead agency of the AU's New Partnership for Africa's Development (NEPAD) to implement the fourth pillar of Comprehensive African Agricultural Development Programme (CAADP), involving agricultural research, technology dissemination and uptake.

FARA's **vision**: reduced poverty in Africa as a result of sustainable broad-based agricultural growth and improved livelihoods, particularly of smallholder and pastoral enterprises.

FARA's **mission**: creation of broad-based improvements in agricultural productivity, competitiveness and markets by supporting Africa's sub-regional organizations in strengthening capacity for agricultural innovation.

FARA's Value Proposition: to provide a strategic platform to foster continental and global networking that reinforces the capacities of Africa's national agricultural research systems and sub-regional organizations.

FARA will make this contribution by achieving its *Specific Objective* of **sustainable improvements to broad-based agricultural productivity, competitiveness and markets**.

Key to this is the delivery of five *Results*, which respond to the priorities expressed by FARA's clients. These are:

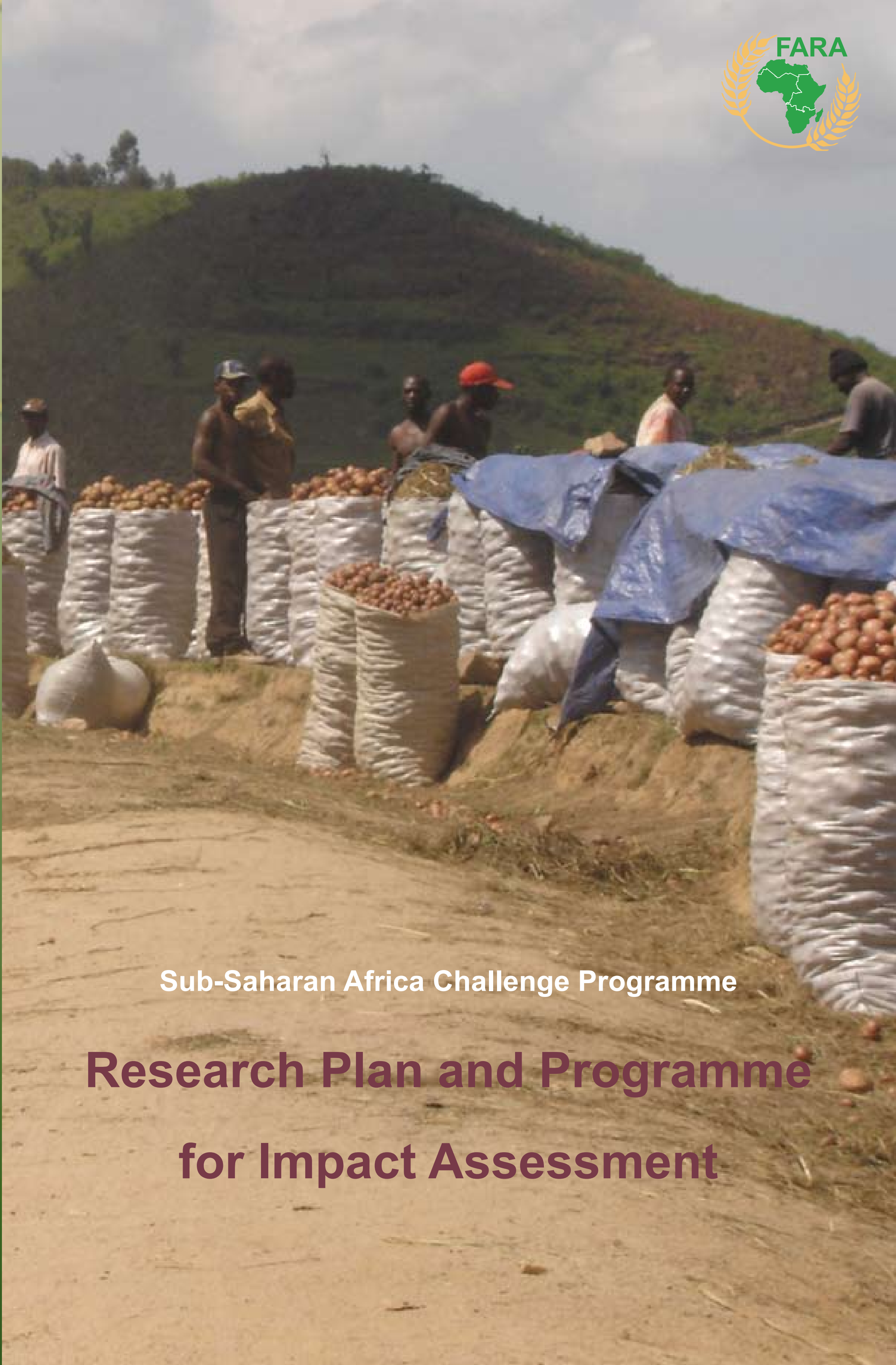
1. Establishment of appropriate institutional and organizational arrangements for regional agricultural research and development.
2. Broad-based stakeholders provided access to the knowledge and technology necessary for innovation.
3. Development of strategic decision-making options for policy, institutions and markets.
4. Development of human and institutional capacity for innovation.
5. Support provided for platforms for agricultural innovation.

FARA will deliver these results through the provision of networking support to the SROs, i.e.

1. **Advocacy and resource mobilization**
2. **Access to knowledge and technologies**
3. **Regional policies and markets**
4. **Capacity strengthening**
5. **Partnerships and strategic alliances**

FARA's major donors are The African Development Bank, The Canadian International Development Agency, European Commission, the Governments of the Netherlands, United Kingdom, Italy, Ireland, Germany and France, the Consultative Group on International Agricultural Research, the Rockefeller Foundation, Bill and Melinda Gates Foundation, the World Bank, and the United States of America Agency for International Development.

Forum for Agricultural Research in Africa (FARA)
PMB CT 173 Cantonments, 2 Gowa Close
Roman Ridge, Accra, Ghana
Telephone: +233 21 772823 / 779421 / 774561
Fax: +233 21 773676 • Web site: www.fara-africa.org
ISBN 978-9988-8470-1-2 (print) ISBN 978-9988-8470-0-5 (pdf)



Sub-Saharan Africa Challenge Programme

Research Plan and Programme for Impact Assessment

PHOTO GALLERY OF KANO-KATSINA-MARADI (KKM) ACTIVITIES IN 2008

PLAR Module on "Layout of nutrient omission trial, phosphorus source trial".

Parcitipatory observations of nutrient omission trial on rice cropping by Dantankari farmers.

Pilot farmer presenting phosphorus source trial outputs for soybean during the a famers' field day in Ikara LG.

Participatory evaluations of nutrient omission trial on rice cropping by Gyazama farmers.

Field visit during a farmers' exchange visit in Ikara LG.

Field visit during a farmers' exchange visit in Dandume LG.

PHOTO GALLERY OF KANO-KATSINA-MARADI (KKM) ACTIVITIES IN 2008

Group working on "Performance indicators of rice cropping and participatory observations" during the PLAR-ISFM Training.

Group working during the IP Actors Workshop (November 20–21, 2008).

Farmers' presentation of nutrient omission trial during the IP Actors Workshop.

Group working during the Participatory Value Chain Analysis (PVCA)Training in Zaria (November 25–28, 2008).

Training of facilitators on "Participatory learning and action research for integrated soil fertility management (PLAR-ISFM Training)" in Zaria, August 5–8, 2008.

Sub-Saharan Africa Challenge Programme

Research Plan and Programme for Impact Assessment

Forum for Agricultural Research in Africa (FARA)
PMB CT 173 Cantonments
2 Gowa Close, Roman Ridge
Accra, Ghana
E-mail: aadekunle@fara-africa.org

2009

Citation: FARA (Forum for Agricultural Research in Africa). 2009. *Sub-Saharan Africa Challenge Programme: Research Plan and Programme for Impact Assessment*. Accra, Ghana. 36 pp.

Forum for Agricultural Research in Africa (FARA)
PMB CT 173 Cantonments
2 Gowa Close, Accra, Ghana
Tel.: +233 21 772823 / 779421
Fax: +233 21 773676
E-mail: info@fara-africa.org
Web site: http://www.fara-africa.org

ISBN 978-9988-8470-1-2 (print)
ISBN 978-9988-8470-0-5 (pdf)

Design: www.bluepencil.in
Print: www.pragati.com

Contents

List of Figures iv

List of Tables iv

Foreword 1

Acknowledgements 3

1. Introduction 4

2. SSA CP's research questions and expected outputs and outcomes 6
- 2.1. Research questions 6
- 2.2.Expected outputs 6
- 2.3.Expected outcomes 7

3. Integrated Agricultural Research for Development (IAR4D) 8
- 3.1. Structure of IAR4D 8
- 3.2. Process principles of IAR4D 9

4. IAR4D impact pathway 10

5. Research hypotheses 12

6. Evaluation design 13
- 6.1. Characterisation of treatment and counterfactual sites 13
- 6.2. Census of district/local government area/commune sites and characteristics 14
- 6.3. Sampling method 14
- 6.4. Sample sizes at IP, taskforce, PLS and SSA CP scales 16
  - *6.4.1.Sample sizes for taskforces* 16
  - *6.4.2. Sample sizes for PLS and programme* 16

7. Testing the research hypotheses: the outcomes of focus 18
- 7.1. Testing the first hypothesis 18
- 7.2. Testing the second hypothesis 19
- 7.3. Testing the third hypothesis 19

8. Data Collection 20
- 8.1. Baseline surveys for IP and community level characteristics 20
- 8.2. Baseline survey for household and village community characteristics 21
- 8.3. Evaluation surveys 21

9. Data Analysis: estimation Issues 22
9.1. Quantitative analysis 22
9.2. Qualitative assessment approaches 25
10. SSA CP project portfolio and logframe 26
References 29
Acronyms and abbreviations 30

List of Figures

Figure 1. Organization of ARD actors in a linear configuration 8
Figure 2. Organization of ARD actors in a network configuration 9
Figure 3. IAR4D impact pathway 11
Figure 4. Illustration of stratification of a PLS by four development domains 15
Figure 5. Random selection of sites and households 16
Figure 6. Levels at which SSA CP data will be collected and pooled 20
Figure 7. Cause-and-effect attribution problem of impact 22

List of Tables

Table 1. Research questions and hypothesis 12
Table 2. Sample sizes at IP, taskforce, PLS and programme scales 17
Table 3. SSA CP portfolio of projects 27
Table 4. SSA CP logframe 28

Foreword

The Sub-Saharan Africa Challenge Programme (SSA CP) is a bold African-led research initiative seeking to increase the developmental benefits from agricultural research and development (ARD). It aims to achieve this objective by proposing, testing and evaluating a more effective alternative to the conventional ARD approaches. The research thrust of the SSA CP was shaped by feedback from numerous consultations carried out by the Secretariat of FARA about the main reasons behind the underperformance of ARD in Africa. These consultations observed that agricultural research in Africa has produced numerous excellent research outputs that have not generated the expected developmental benefits across the continent.

The consultations attributed the limited impact of ARD to the failure by the intended users of research outputs to put them into use beyond the local domains within which the outputs are generated and tested, leading to "islands of success". The low uptake of research outputs has in turn been attributed to some features of the way ARD is currently conducted, whereby researchers, the end users of research outputs and other providers of services that support agricultural production, value addition and marketing, largely work in isolation from one another. The alternative ARD approach proposed by the SSA CP is designed to address the institutional and other process-related impediments associated with the conventional approach. The proposed approach is known as Integrated Agricultural Research for Development (IAR4D).

The idea that the way ARD is organized accounts for a large part of its underperformance in Africa has been acknowledged for some time. However, efforts aimed at addressing this shortcoming have lacked: (i) the scope necessary to bring about the required wide-scale institutional changes on the continent, and (ii) concrete and generalizable evidence about the wide-scale feasibility and cost–benefit effectiveness of the proposed alternative approaches. FARA has leveraged it mandate as the apex organization for ARD in Africa to develop a programme designed to overcome these twin challenges. The FARA Secretariat is currently coordinating the SSA CP and proposes to leverage its networking support functions[7] to support the wide-scale uptake of the Programme's recommendations.

The SSA CP has undergone several design changes since its inception in January 2005. It was initially conceived as a large-scale action-research and capacity-building initiative aimed at testing and scaling out IAR4D. During its 18-month inception phase, it established governance and management structures and drew up its first research plan. After the inception phase, it was evaluated by the Science Council (SC) of the Consultative Group on International Agricultural Research (CGIAR)—the funders of the SSA CP. The SC recommended that the SSA CP should focus on rigorously establishing the proof of the IAR4D concept; that is, testing the viability of the approach

7. (i) Advocacy and Resource Mobilization, (ii) Access to Knowledge and Technologies, (iii) Regional Policies and Markets, (iv) Capacity Strengthening, and (v) Partnerships and Strategic Alliances.

and evaluating its cost–benefit effectiveness relative to the conventional method. This recommendation required FARA to redesign the Programme's research plan. After two iterations, a new research design was accepted by the SC—and is reported here.

The new research design of the SSA CP depicts the Programme as a large-scale experiment for evaluating the effects (benefits, costs and risks) of IAR4D and comparing them with the effects of conventional ARD approaches. The evaluation also attempts to identify the effects of the various components of IAR4D. The research is also expected to contribute to the emergent body of knowledge on methodologies for combining quantitative evaluation of processes with their better-established qualitative counterparts.

The SSA CP is structured into four projects and nine subprojects spread across Sub-Saharan Africa. In addition to generating evidence about whether IAR4D works and is more cost effective than the conventional ARD approach, the SSA CP will provide tested guidelines for implementing IAR4D and a critical mass of ARD stakeholders familiar with IAR4D. FARA proposes to use the outputs of the SSA CP to advocate for the evolution and reform of African ARD institutions. The evidence about the effectiveness of IAR4D will be used to inform the advocacy efforts and to craft policies for change. The Secretariat of FARA proposes to widely disseminate the guidelines on application of IAR4D and to support the integration of IAR4D in agricultural training and capacity-strengthening programmes.

On behalf of the numerous partners involved in the conception and implementation of the SSA CP, FARA wishes to acknowledge: (i) the organs of the CGIAR (that is, the Executive Council, Secretariat, and Science Council) for supporting the creation of the SSA CP and guiding its evolution into a creditable research programme, and (ii) the governments of the United Kingdom, Italy, the Netherlands, Norway and Denmark, as well as the Commission of the European Union, for funding the SSA CP.

FARA wishes to thank all the individuals and organizations involved in the SSA CP—especially the sub-regional agricultural research organizations (SROs), project lead institutions and the various organizations responsible for implementing the subprojects—for believing in SSA CP's vision and working tirelessly to make possible the achievements registered by the SSA CP to date. FARA counts on the continued commitment of the CGIAR, funders and implementing stakeholders to realize the SSA CP's cause.

Dr Monty P. Jones

Executive Director, FARA

Acknowledgements

FARA gratefully acknowledges the team that assembled the first draft of the SSA CP's research design described herein. This team comprised the following:

- Dr Adewale Adekunle, Coordinator of the SSA CP and Director for Partnerships and Strategic Alliances at the FARA Secretariat;
- Dr Aliou Diagne, the Impact Assessment Economist at the Africa Rice Center
- Dr Tahirou Abdoulaye, an Agricultural Economist at the International Institute of Tropical Agriculture and Coordinator of the SSA CP in the Kano-Katsina-Maradi Pilot Learning Site
- Dr Joseph Rusike, an Agricultural Economist at the International Institute of Tropical Agriculture and Coordinator of the SSA CP in the Zimabwe-Mozambique-Malawi Pilot Learning Site
- Dr Aggrey Agumya, Programme Officer in the SSA CP's Coordination Unit at the FARA Secretariat
- Dr Jean-Claude Legoupil, a Technical Advisor to the SSA CP's Coordination Unit at the FARA Secretariat.

The first draft was subsequently revised, drawing on comments provided by the SSA CP's task-force teams led by Dr Paul Mapfumo (CIMMYT); Dr Mikkel Grum (Bioversity International); Dr Robert Delve (TSBF-CIAT); Dr Robin Buruchara (CIAT); Dr Andy Farrow (CIAT); Dr Rose Njeru (ISAR); Dr Moses Tenywa (Makerere University); Dr Nouri Maman (INRAN); Dr Abdoulaye Mando (IFDC); and Dr Alpha Kamara (IITA). Additional comments were provided by Dr Jemimah Njuki (CIAT) and Dr John Pender (IFPRI).

FARA also acknowledges feedback from the CGIAR's Science and Executive Councils on earlier designs of the SSA CP's research, which was instrumental in shaping the current research design.

FARA wishes to thank the Commission of the European Union and the governments of the United Kingdom, Italy, the Netherlands, Norway and Denmark for providing funding for the SSA CP. The Commission of the European Union and the governments of the United Kingdom and Italy have pledged to continue funding the research described herein.

1. Introduction

The Sub-Saharan Africa Challenge Programme (SSA CP) was initiated in 2004 following extensive consultations with numerous agricultural stakeholders (researchers, extension and development agents, policy makers, farmers and the private sector) to diagnose the main reasons behind the underperformance of agricultural research in Africa. These consultations, which were convened by the Secretariat of the Forum for Agricultural Research in Africa (FARA), established that, besides inadequate funding, the main impediment to African agricultural research's contribution to development impact lies in the way the research is organized and conducted, in other words, the impediment is institutional. Thus, the SSA CP was formulated in response to the need to dramatically increase the development impact of agricultural research on livelihoods in Africa. It aims to achieve this goal by developing, testing and promoting an approach for conducting agricultural research for development (AR4D) in Africa, which overcomes the shortcomings of traditional approaches. The approach proposed by the SSA CP is known as *Integrated Agricultural Research for Development (IAR4D)*. A characterization of IAR4D is presented in Section 3.

The SSA CP is coordinated by FARA. It was launched in January 2005, starting with an 18-month inception phase during which it was tasked with establishing a governance and management structure, developing a coherent research plan and identifying the multi-institutional teams that would be charged with implementing the plan. The programme's inception phase work was implemented in three Pilot Learning Sites (PLSs), namely: (i) Kano-Katsina-Maradi (KKM)—which straddles Niger and Nigeria; (ii) "Lake Kivu", which covers the area where Democratic Republic of Congo, Rwanda and Uganda meet; and (iii) Zimbabwe-Mozambique-Malawi (ZMM)—a transect that runs from north-east Zimbabwe through central Mozambique into southern Malawi.

At the end of the SSA CP's inception phase, in June 2006, the CGIAR Science Council (SC) commissioned an evaluation of the programme to establish its readiness to proceed to implementation of its research plan. Specifically, this evaluation sought to assess: (i) the progress registered by the programme in designing a compelling research plan from which international public goods (IPGs) would emerge, and the feasibility of implementing this plan in the PLSs; (ii) the knowledge gained from the institutional learning and its contribution as an IPG, particularly for Sub-Saharan Africa; and (iii) how the proposed IAR4D approach adds value to the identification of appropriate agricultural interventions.

This evaluation concluded that the progress made by the programme towards achieving its inception phase objectives provided sufficient grounds for it to be granted the approval to continue for a further phase (the research phase) of 3 years. It recommended that the SSA CP's research should focus on a proof of the IAR4D concept. This would involve rigorously establishing: (i) whether IAR4D works and delivers IPGs; (ii) whether IAR4D is superior to traditional agricultural research and development (ARD) approaches in delivering benefits to end users, and (iii) whether IAR4D is replicable outside its test environment.

Accordingly, the SC directed the SSA CP to develop a coherent research plan, which would include a detailed narration of how the evaluation to identify the effects of IAR4D (and its separate components) will be carried out. The SC also endorsed research on the *interfaces* of processes driving (a) productivity gains, (b) efficient use of resources, (c) the care of the environment, (d) agricultural policies and (e) markets, as the problem and opportunity space within which IAR4D will be implemented and evaluated. The SC further recommended that the SSA CP's research plan should be described in its 2008–2010 medium-term plan (MTP).

FARA welcomed the SC's recommendations and acted on them. It coordinated the development of a new research plan which was integrated in the SSA CP's 2008–2010 MTP. FARA submitted this MTP to the SC in May 2007. The SC's commentary on this MTP was critical of the lack of specificity and targets by which progress towards establishing the proof of concept could be monitored. It characterized the impact pathway as too generic. The SC concluded by observing that the SSA CP's research design did not sufficiently address the key requirement of identifying the effects of the different components of the IAR4D approach in a scientifically and statistically-based manner.

FARA has acted on the concerns raised in the SC's 2007 commentary by radically revising the SSA CP's research plan to firmly align the SSA CP's research with the SC's recommendations aimed at rigorously generating evidence for a proof of the IAR4D concept. This document summarises the SSA CP's revised research plan. It describes IAR4D—in particular, its structure, process principles and impact pathway. It then outlines the hypotheses that the SSA CP sets out to test in establishing a proof of the IAR4D concept and describes how they will be tested, that is, the framework for sampling, data collection, and methods for data analysis. It concludes by presenting logframes indicating the outputs expected from this research, as well as the corresponding annual output targets, intended users, outcomes and impact.

2. SSA CP's research questions and expected outputs and outcomes

2.1. Research questions

During its 3-year research phase, the SSA CP will focus on answering the following three research questions[7]:

1. Does the IAR4D concept work and can it generate IPGs and regional public goods (RPGs) to end users?
2. Does the IAR4D framework deliver more benefits to end users than conventional approaches (assuming conventional research, development and extension approaches have access to the same resources)?
3. How sustainable and usable is the IAR4D approach outside its test environment, that is, concerning its scaling out for broader impact?

The SSA CP's research will centre around evaluation of IAR4D, specifically the processes it entails and their benefits compared to those delivered by traditional approaches. Accordingly, the formulation of the programme's research design sought to make a clear characterization of IAR4D, and particularly its components. In addition, the design lays out the methods for monitoring and evaluating IAR4D processes and evaluating their development impact relative to traditional ARD approaches.

The SSA CP's research will be carried out by nine taskforces across the three PLSs—three taskforces per PLS. Each taskforce will essentially implement a case study that uses IAR4D to resolve constraints on the development, dissemination and uptake of research results in order to generate substantially greater impact in a specific context. The nine taskforces are, however, organized as elements of a coherent programme that will pool together their data across different sites and make inferences to address the three research questions.

2.2. Expected outputs

The principal outputs and IPGs of the SSA CP are:

(i) Innovation platforms[8] for use across a wide range of conditions in Africa;

(ii) Technological, institutional and policy options targeted on the interface of processes driving productivity gains, efficient use of resources, the care of the environment, and policies and markets;

7 A fourth research question regarding the effects of the individual components of IAR4D was mentioned in the SC commentary, but will not be addressed in this research. Indeed, all the IAR4D components will be applied at the same time at each site and by each taskforce. If all taskforces implement all the components then we cannot identify the effects of different IAR4D components. To identify the separate effects of the five IAR4D components, a 5 × 5 factorial design would be needed. This would require 25 IPs per taskforce per PLS or 225 IPs for the whole programme (given that each one of the 3 PLSs has 3 taskforces). The number of sample villages and farmers this would require is not feasible given the timeframe and resources available to the SSA CP. Therefore, the research will focus on a "black box" assessment of whether IAR4D works. Nevertheless, because IAR4D is flexible and adaptable, it is likely that we will have variation in its implementation by the different taskforces, and thus the treatment will be variable during implementation. The expected variation in implementation will be monitored and used *ex post* to decompose the black box and identify effects of separate components, to assess components that work well and components that work better than others. In this way, the impact of components of IAR4D will eventually be compared with that of similar elements of the conventional ARD approaches.

8 The definition of innovation platform is presented in section 3.2. The innovation platform IPG is a set of validated principles and procedures for setting up and operationalizing innovation platforms.

(iii) An evaluation of whether IAR4D works and the benefits it delivers compared to those delivered by traditional approaches;

(iv) A database of process and impact indicator variables for 36 innovation platforms and their associated research communities and households;

(v) Methods and tools for designing, implementing and analysing social experiments in Sub-Saharan Africa.

The procedures for implementing IAR4D include those concerned with: (a) setting up an Innovation Platform (IP); (b) making the IPs work, and (c) institutionalizing the use of IPs within organizations and institutions.

2.3. Expected outcomes

The outcomes expected from the SSA CP's implementation and evaluation of IAR4D include:

(i) increased diffusion and adoption of IAR4D by research and development organizations

(ii) increased awareness of the impact pathway for IAR4D

(iii) increased investment towards supporting IAR4D processes

(iv) increased human and institutional capacity for innovation among ARD actors

(v) increased adoption of sustainable productivity and profitability enhancing innovations.

The impact of the SSA CP is synonymous with the impact of IAR4D, which includes improvements in food security, income, livelihood assets, the natural resource base and resilience to shocks.

3. Integrated Agricultural Research for Development (IAR4D)

The conception of IAR4D was driven by dissatisfaction with traditional approaches for organizing ARD in Sub-Saharan Africa. These approaches are widely blamed for contributing significantly to the unsatisfactory performance of ARD in improving the livelihoods of its end users—the smallholder farmers. The poor performance of traditional ARD approaches is manifest in low adoption rates of technologies, poor linkages among agricultural value-chain actors and the pervasive unprofitability of farm enterprises in SSA. It has been hypothesized that these indicators of unsatisfactory ARD performance are traceable to the organization of research and development as a linear process (Figure 1). This configuration of ARD actors limits interaction with researchers and timely intervention in research process and direction.

IAR4D aims to transform this configuration by embedding research within an innovation system comprising all actors in agricultural value chains. Within such a system—a network configuration—innovation does not follow a linear path that begins with research, moves through the processes of development, transfer, diffusion, adoption, production, and ends with successful introduction and use of new products and processes; rather, it tends to involve continuous feedback between different stages (Dantas, 2005), thus drawing on the knowledge of all relevant actors at each stage. The network configuration (Figure 2) facilitates timely interaction and learning, and aims at generating innovations (rather than research products per se). Here, innovation refers to the activities and processes associated with the generation, product distribution, adaptation, and use of new technical and institutional/organizational knowledge. It therefore adds value to products of research, thus catalysing the achievement of development impact.

IAR4D is characterized by a structure and several process principles.

3.1. Structure of IAR4D

The IAR4D structure is an Innovation Platform (IP)—an informal coalition, collaboration, partnership and alliance of public and private scientists, extension workers, representatives of farmers, farmers' associations, private firms, non-governmental organizations, and government policy makers who communicate, cooperate and interact (often across sectoral and ministerial lines) motivated by the common belief that increasing agricultural productivity can help improve the welfare of all members of society (Eicher, 2006). The core competencies brought to bear by the IP are greater than the sum of the IP's constituents acting independently.

Figure 1. Organization of ARD actors in a linear configuration.

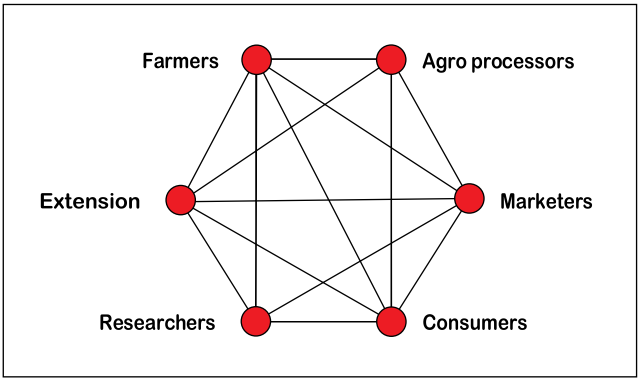


Figure 2. Organization of ARD actors in a network configuration.

3.2. Process principles of IAR4D

The key process principles that characterize IAR4D comprise the following.

1. *Existence of an IP*, which serves as the platform for diagnosing problems, exploring opportunities and investigating solutions. The actors in the IP:

(a) are organized in partnerships/teams to bring about mutually desirable change;

(b) are competent and have incentives to jointly innovate;

(c) are constituted to include sources of the key competences and knowledge required to address the problems, opportunities and/or entry points that prompt its establishment.

2. *Non-linear (network) collective and collaborative interaction* among IP actors (rather than linear researcher–extension–farmer transfer of technology model (see for example Figure 1).

This enhances:

(a) direct and continuous interaction, communication and knowledge-sharing among the IP actors;

(b) quick and continuous feedback from end users (farmers) at all stages of the research for development;

(c) timely integration of new knowledge into the innovation process using experiential learning, monitoring and evaluation, and the continuous feedback.

3. *Research that addresses key constraints and opportunities agreed by IP actors in the context of entire value chains* (from input supply through production to consumption) and sustainable livelihood systems.

4. A research process that is *multidisciplinary* and *participatory*.

5. *Institutional and human capacity building* for IAR4D actors to effectively participate:

(a) the capacity building needs are identified by IP actors;

(b) training (formal and non-formal) is provided by the appropriate partners.

4. IAR4D impact pathway

The point of departure of IAR4D from traditional ARD approaches lies in how innovations are generated. While traditional ARD approaches *exogenously* bring innovations into the system, IAR4D instead establishes an institutional innovation—the Innovation Platform—which in turn, *endogenously* generates the innovations (technological, market, institutional and policy). summarizes the research-to-impact pathway used to hypothesize the causal relationships between research inputs, and the research outputs—the IP institutional innovation and its results (knowledge increase, behavioural change and innovations at the interfaces of processes driving productivity, environment, policies and markets); knowledge and behavioural outcomes at the household, community and market levels, and impact outcomes. This is the hypothesized generic impact pathway for IAR4D. Impact pathways for individual SSA CP taskforces exhibit minor variations on Figure 3 depending on specificities of the problem or opportunity they address.

The main outcomes at the IP level are increased awareness, increased knowledge drawn from several IP sources and behavioural changes at the individual and system levels. These combine to generate innovations at the interfaces of productivity, care for the environment, policies and markets with a potential to demonstrably increase the delivery of benefits to end users. This will in turn lead to outcomes at farm household, village community, and market levels. The main outcomes at the household and community levels are:

- increased awareness and knowledge;
- behavioural outcomes (such as adoption of relevant innovations, more effective supply of inputs to satisfy demand, increased and better-expressed demand for inputs, and increased volume of input sales);
- market outcomes (increased and more effective supply of outputs, increased demand by consumers);
- efficiency outcomes (increased yields, technical efficiency and profit).

These outcomes lead to impacts in the form of welfare and equity outcomes (such as increased incomes, poverty reduction, improved health and nutrition, and equity) and environmental outcomes (for example, imputed soil fertility and erosion). It is hypothesized that evidence provided by the SSA CP's research comparing the benefits of IAR4D against conventional ARD approaches will determine whether communities and other organizations more directly involved in development will seek to adopt and scale up IAR4D.

The outcomes and range of IAR4D's impact are influenced by several conditioning factors (see Figure 3). These factors complicate the attribution of changes in impact indicators to IAR4D alone. Factors exogenous at the household level but endogenous at the community level

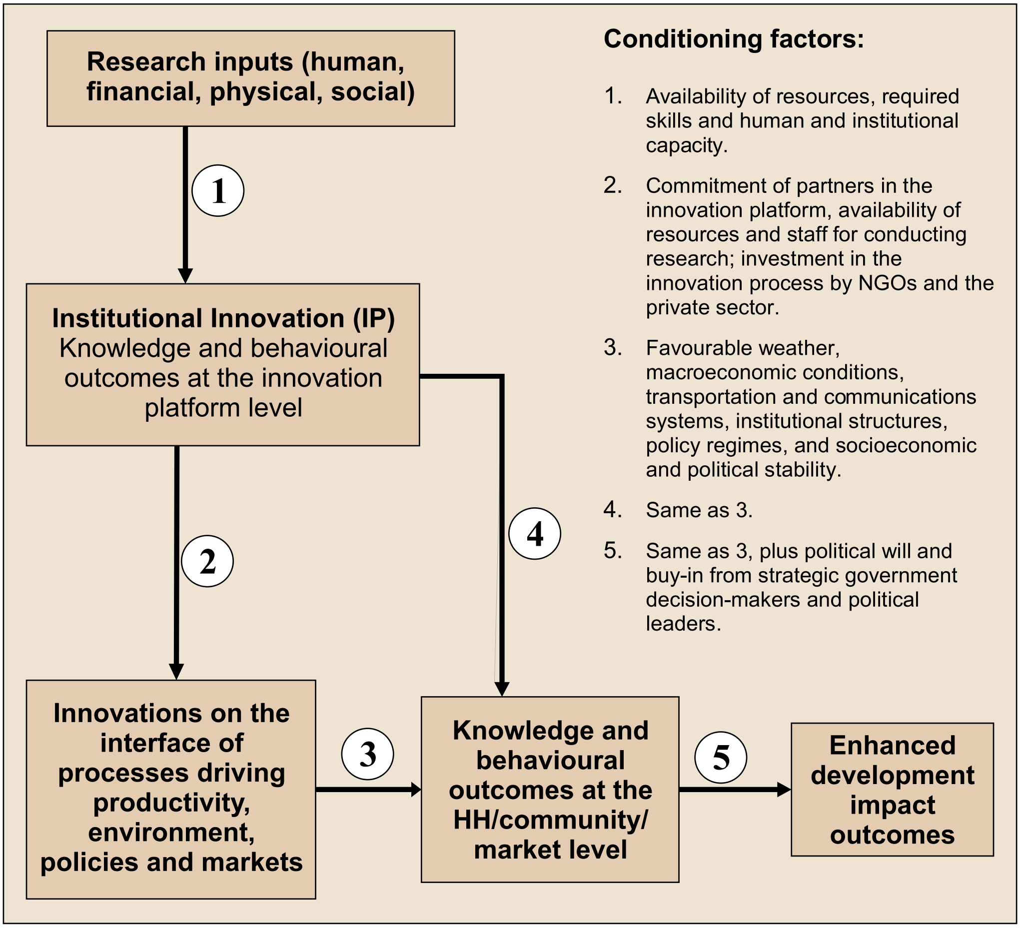


Figure 3. IAR4D impact pathway.

include infrastructure (public and privately supplied), institutions (governance and market structures), policies (macroeconomic, sectoral, pricing, social), technologies, and information.

Factors exogenous at the community level include agroclimatic conditions and external market conditions (world prices and access to foreign markets).

5. Research hypotheses

The SSA CP will test three hypotheses that flow from the three research questions (see Section 2).

Table 1. Research questions and hypotheses

Research question	Hypotheses
Does the IAR4D concept work and can it generate international and regional public goods to end users?	**H1:** If an innovation platform (IP) is created and is functional with the 5 components characterizing IAR4D, then it will lead to increased interactions among partners in the IP compared to where there is no IP; increased interactions among farm households in communities, and better developmental outcomes where IAR4D is in operation compared to communities where IAR4D is not in operation
Does the IAR4D framework deliver more benefits to end users than conventional approaches (assuming conventional research, development and extension approaches have access to the same resources)?	**H2:** IAR4D delivers more benefits to end users compared to conventional approaches (if the conventional ARD approaches have access to the same resources)
How sustainable and usable is the IAR4D approach outside its test environment, that is, concerning its scaling out for broader impact?	**H3:** If the design and estimation shows that IAR4D works in the different PLS contexts then it can be extrapolated outside the test environments

6. Evaluation design

In order to test the three hypotheses in a statistically robust fashion and empirically determine whether IAR4D works and whether it delivers more benefits than conventional approaches, a multiple-treatments experimental design will be used. This design compares household- and community-level outcomes under: (i) IAR4D, (ii) the conventional approach, and (iii) no intervention. In other words, the SSA CP experiment will comprise three treatments carried out in three blocks (the PLS) and nine repetitions (three per block—the taskforces).

Following White and Chalak (2006) we take the set of *counterfactuals* to be the set of all possible states of the world with outcomes taking different values under different possible states of the world. We also define an *intervention* as the move from one possible state to another. So there are as many counterfactuals as there are possible states of the world[7]. However, under the SSA CP we are limiting ourselves to comparing outcomes under IAR4D and under only two other possible states, namely: the conventional approach and under non-intervention. So, our set of counterfactuals is limited to the set $\{\omega_0, \omega_1, \omega_2\}$ where ω_0 is the non-intervention state consisting of having neither IAR4D nor the conventional approach in operation, ω_1 the state consisting of having the conventional approach in operation, and ω_2 is the state consisting of having IAR4D in operation[8].

7 But among all the possible states of the world only one gets realized (the factual) in any given situation, all the other are counterfactuals.

8 Only one of the three possible states gets realized in any given site. The realized state will then be the factual and the two unrealized ones the counterfactuals.

The effectiveness and impact of IAR4D will be assessed throughout the impact pathway all the way to the farmer level. The hypothesis about whether IAR4D works will be tested by comparing the values of relevant knowledge, behavioural, efficiency, welfare, equity and environmental outcomes under ω_2 and under ω_0. Similarly, the hypothesis about whether IAR4D delivers more benefits than the conventional approach will be tested by comparing the values of relevant knowledge, behavioural, efficiency, welfare, equity and environmental outcomes under ω_2 and under ω_1. The "with" and "without" IAR4D comparison will be made by comparing the values of the same outcomes as above under ω_2 and under the composite possible state "ω_0 or ω_1".

Characterisation of treatment and counterfactual sites

Innovation Platforms will be evaluated at the district, local government area or commune level, because it is conceptualized that the innovation process is best organized through geographically decentralized sites. The geographical area of influence of the IP is conceptualized to be mostly within district/local government areas/commune jurisdictional boundaries because of clustering of activities and interactions among government administration units, public research and extension organizations, farmers, farmers' organizations, NGOs, agricultural input suppliers and output marketing firms, credit and finance organizations, and service providers.

- The treatment communities will consist of organizations and farm households in areas where IAR4D will be practised.
- The non-treatment communities will consist of similar organizations and households in other sites.

The PLS will be zoned into development domains—areas with comparable development potential. The development domains used by the SSA CP are based on two factors that usually have the largest influence on agriculture-driven development, namely agroclimatic potential and access to markets (see Figure 3). The development domains combined with population data will be used to target areas most likely to provide the highest returns on the SSA CP's investment. They will also provide a basis for stratifying the PLS in order to capture its variation and to delineate similar domains from which comparable sites will be selected.

Research sites (districts, communes, local government areas) will be allocated to IAR4D and non-IAR4D treatments through stratified random sampling. The strata within which the randomization will be carried out are four development domains delineating the combination of market access potential and agroclimatic potential. Each IAR4D treatment site (district, commune, local government area) will have a corresponding counterfactual site randomly selected from the same stratum as the IAR4D site (see Figure 4 for an example). Taskforces will spread IAR4D treatment sites across various strata in order to investigate the performance of the approach across a wide range of conditions. Each taskforce will establish four Innovation Platforms in four separate districts/communes/local government areas. Thus, each taskforce will work in eight sites.

Within IAR4D and non-IAR4D sites, focal villages will also be selected randomly. The focal villages will be screened prior to implementation of IAR4D to establish whether or not they have had conventional ARD or IAR4D-type of projects in the past 2–5 years. Villages will be classified into two types: (a) "*clean*" villages that have neither had IAR4D nor conventional projects in the last 2–5 years[9]; and (b) conventional ARD villages that have had projects identifying, promoting and disseminating technologies in the past 2–5 years.

IAR4D will be introduced in "clean" villages within the IAR4D sites. The SSA CP's hypotheses will be tested by determining whether outcomes differ among households in the IAR4D, "clean" and conventional ARD villages.

6.2. Census of district/local government area/commune sites and characteristics

A list of districts within the PLS and their characteristics, including their representation of the four development domains, will be used to identify clusters for targeting under the research and classification of villages into "clean" and "non-clean" status. Some of this information will be collected from national statistical offices and used for stratification and multistage sampling of villages. Field visits will be made to sampled villages to collect information on their "clean" or "non-clean" status using key focus group discussions and informant interviews.

An alternative method entails using geographical information systems (GIS) and quota sampling to select district/local government area/commune sites for assignment to treatment. GIS tools will be used to randomly pick five points, followed by field visits to check their "clean" or "non-clean" status of the quadrant in which they fall. If the quadrant is not "clean" then it will discarded and another one sampled until the quota is met. The advantage of this method is that it does not require a prior census.

6.3. Sampling method

Multistage stratified random sampling will be carried out within the selected districts (IAR4D

9 Whereas "clean" villages are defined as those that have not had any intervention or initiative (conventional or IAR4D-like) over the last 2–5 years, the categorization of any village as "clean" will depend on the local context.

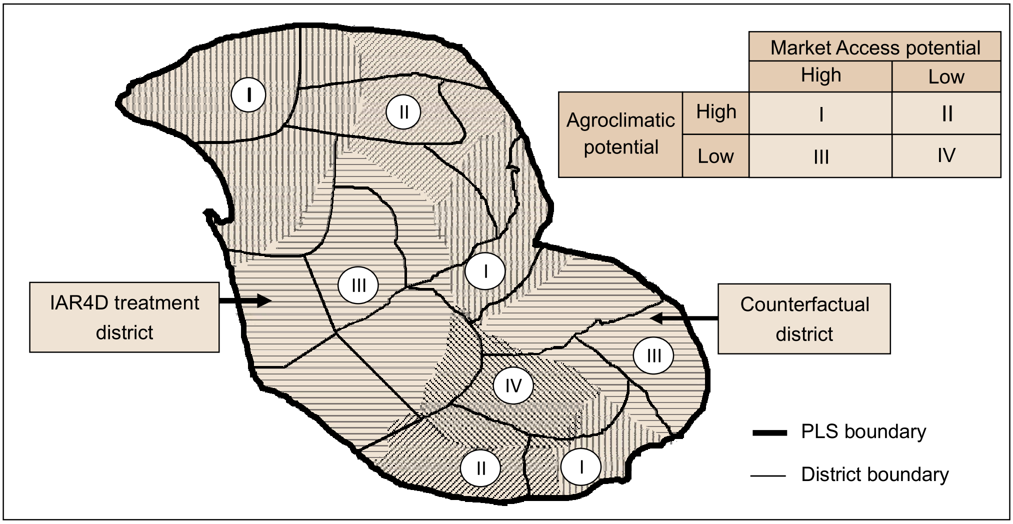


Figure 4. Illustration of stratification of a PLS by four development domains.

and counterfactual) to select the villages where the treatments will be applied, that is villages where IAR4D will be introduced, village communities where conventional approaches are in operation, and villages where no interventions have been carried out over the last 2–5 years (see Figure 5).

The Miguel and Kremer (2004) methodology of randomizing treatments across schools (districts and village communities) and not individual farm households will be used, because it captures spillover and externality benefits that would be underestimated if treatment was only randomized at the individual level. All districts, local government areas and communes within the PLS will first be listed and grouped according to their representation of the four development domains. Depending on the context and its specific requirements, each taskforce will define the strata within which it will randomly select the four districts that will serve as its IAR4D treatment sites; that is, where IAR4D will be introduced. Within the IAR4D sites, a census of the village communities will be conducted to develop a village sampling frame and stratify the villages into "clean" and "non-clean" villages. At least five focal villages per IAR4D site will be randomly selected from "clean" villages. These villages will become the theatres for action research aimed at developing innovations at the interface between productivity, care of the environment, policies and markets. Within the focal IAR4D village communities, at least 10 households per village will be randomly selected for monitoring and evaluation.

Four counterfactual districts/local government areas/communes that are similar to the IAR4D sites—that is, share the same development domain—will be assigned to conventional and non-IAR4D-non-conventional ("clean" village) treatments. As for IAR4D sites, a village census will be carried out and villages stratified into "clean" and "non-clean". For each counterfactual site matching an IAR4D site, five focal villages will be randomly selected from "clean" villages only and assigned to the non-IAR4D-non-conventional treatments. Similarly, five focal villages will be randomly sampled from "non-clean" villages and assigned to conventional approach treatment. At least 10 households per focal village will be randomly selected for monitoring and evaluation.

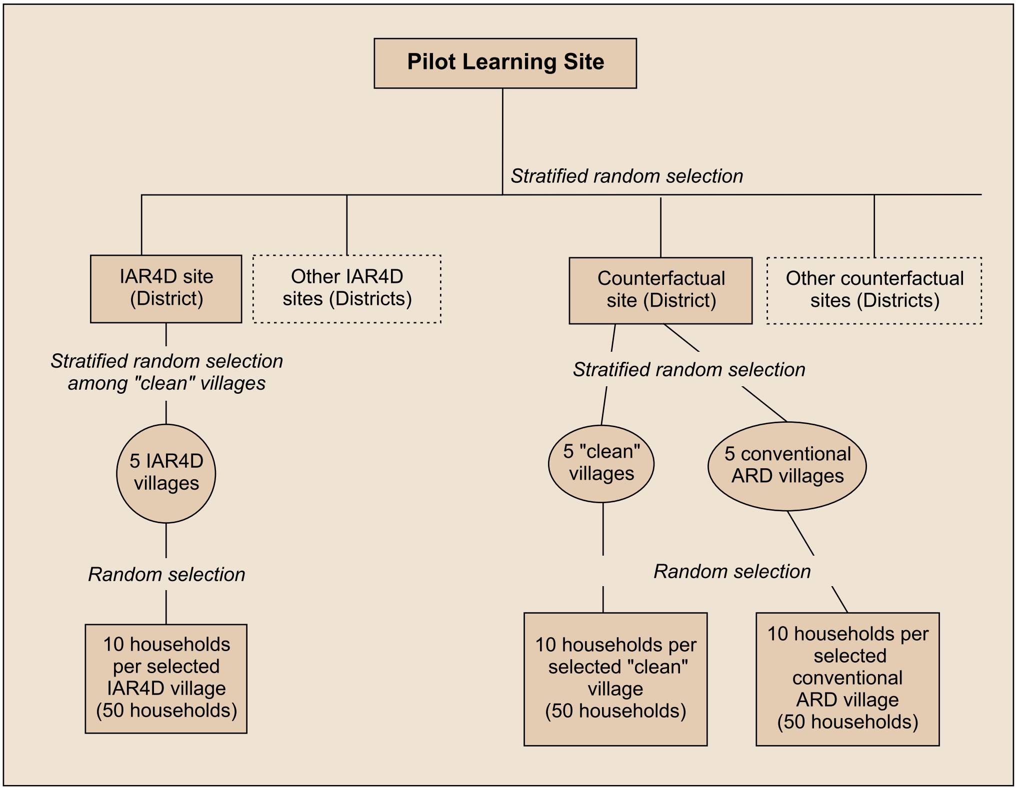


Figure 5. Random selection of sites and households

6.4. Sample sizes at IP, taskforce, PLS and SSA CP scales

6.4.1. Sample sizes for taskforces

For each of the three treatments (IAR4D, conventional ARD and no intervention at all), taskforces will initiate action research in five focal villages. Thus, each taskforce (4 sites) will work in 60 villages. Within each village the taskforces will monitor 10 households. Over the 60 IAR4D villages, each taskforce will monitor 600 households.

6.4.2. Sample sizes for PLS and programme

At PLS level, the taskforce sample sizes are multiplied by a factor of 3 reflecting the three taskforces in each PLS. At programme level, the taskforce sample sizes are multiplied by a factor of 9 reflecting the nine taskforces constituting the programme, that is, 36 Innovation Platform sites and 36 comparator (counterfactual) sites; 540 village, and 5400 households consisting of 1800 IAR4D treatment, 1800 conventional and 1800 non-IAR4D-non-conventional counterfactual villages.

The sample sizes at the four scales, that is, IP, taskforce, PLS and programme) are summarized in Table 2.

Observations and results across taskforces will be pooled using meta-modelling to evaluate

Table 2. Sample sizes at IP, taskforce, PLS and programme scales.

	No. of IPs	No. of IAR4D villages	No. of conventional ARD villages	No. of non-IAR4D-non-conventional villages	Total no. of households
IP	1	5	5	5	150
Taskforce	4	20	20	20	600
PLS	12	60	60	60	1800
SSA CP	36	180	180	180	5400

the site-to-site variation of IAR4D treatment effects. Observations and results will be pooled across multiple sites and PLSs to evaluate the programme and predict the impact of IAR4D at new sites.

7. Testing the research hypotheses: the outcomes of focus

7.1. Testing the first hypothesis

The first hypothesis states that *"If an innovation platform is created and is functional with the five components characterizing IAR4D, then it will lead to increased interactions among partners in the IP compared to where there is no IP; increased interactions among farm households in communities, and better developmental outcomes where IAR4D is in operation compared to communities where IAR4D is not in operation"*. This will be tested at the first level of analysis by analysing if implementation of IAR4D results in:

- non-linear collaborative interaction among IP actors;
- research plans to address key constraints and opportunities agreed by the IP in the context of entire value chains;
- multidisciplinary and participatory research;
- skills to effectively participate.

And if this, in turn, leads to first-level outcomes, namely:

- increased sharing of knowledge among IP members and communities;
- increased number and composition of disciplines involved in the research;
- level of congruence between research conducted, and constraints and opportunities identified by the IP;
- congruence between training conducted, the needs identified and behavioural outcomes;
- number of technological and institutional innovations developed, identified and promoted.

The taskforces will be in control of the conditioning factors, since they will provide IPs with resources and process guidelines for implementing participatory learning and action research (PLAR). PLAR involves joint planning, research on the interfaces, monitoring and evaluating the research process and its outputs, analysis and learning, seminars, debriefing and collective assessment sessions, and exchange visits. A participatory monitoring and evaluation (PM&E) system will be established and used for communication and experiential learning, resulting in more salient, credible and legitimate research findings, and institutional and human capacity outcomes[7].

It should be recalled that the main output of the SSA CP's research is the innovation platform (including "routinized" processes for its successful operationalization). The IPs will facilitate the generation of innovations including technological, market and policy, training, networking and information-sharing options resulting from properly targeted research at the interface of the processes driving productivity gains, efficient use of resources, care of the environment, policies and markets. This will be tested by establishing whether the five

7 Meinke *et al.* (2006) argue that the translation of research information into real-life action requires three essential components: (a) salience (the perceived relevance of the information); (b) credibility (the perceived technical quality of the information); and (c) legitimacy (the perceived objectivity of the process by which information is shared).

characteristics of IAR4D lead to innovation outcomes, namely:

- Awareness
- Knowledge
- Behavioural change
- Innovations on interfaces of processes driving productivity, efficient use of resources, care of environment, policies and markets.

7.2. Testing the second hypothesis

The second hypothesis states that *"IAR4D delivers more benefits to end users compared to conventional approaches (if the conventional ARD approaches have access to the same resources)"*. This will be tested by comparing outcomes of the three treatments (IAR4D, conventional, and non-IAR4D-non-conventional). The comparison, which will use established and emerging methods for comparing treatment outcomes, will control for the costs of delivering services under IAR4D and conventional approaches. Because there may not be enough cost–benefit data per approach, the hypothesis testing will also use cost–benefit per farmer reached or served in the community or cost–benefit per farmer who benefits from IAR4D, say, through adoption of IAR4D-derived innovations.

7.3. Testing the third hypothesis

The third hypothesis states that *"If the design and estimation shows that IAR4D works in the different PLS contexts then it can be extrapolated outside the test environments"*. This will be tested by implementing the frameworks developed by Heckman (2005) and Dehejia (2003)[8]. Linear parameterization models will be used to extrapolate and generalize IAR4D to other larger contexts. Structural models of how treatment effects differ by households, village community and area will be developed using the results on how impacts vary along these dimensions. These models will be applied to forecast the effects of IAR4D in new contexts. The data to be collected will include household, farming community and area characteristics that can be used with the conditional independence assumption to identify structural models for extrapolation.

To assure variability in contexts and implementation of IAR4D, each taskforce will experiment with four innovation platforms. Country or cross-country analyses can be carried out using the synthetic control matching methods developed by Abadie *et al.* (2007). Synthetic control methods for case studies provide inference regardless of the number of comparison units, the number of available time periods, and whether the data are individual (micro) or aggregate (macro).

Another test of this hypothesis will be the analysis of adoption and diffusion of the IAR4D by organizations more directly involved in development, including national research and extension programmes, farmers' organizations, NGOs and agribusiness firms. The SSA CP will conduct surveys to identify organizations taking up IAR4D interventions, and to establish the organizational determinants of and constraints to the adoption of IAR4D.

8 Heckman (2005) argues that using structural models permits interpretability in terms of theory and recognition of missing variables (unobservables), and provides the basis for forecasting the impacts of interventions implemented in one environment, to other environments, including their impact in terms of welfare. Dehejia (2003) has developed a hierarchical model for data that has a group structure, which can be used to evaluate and predict the treatment impact of a programme that is implemented at multiple sites. Moffitt (2004), however, argues that purely statistical models alone are unlikely to be satisfactory because of the numerous causal effects involved in the scale-up problem. Moreover, purely statistical models do not adequately separate the different confounding factors.

8. Data collection

Taskforces will assemble data to establish baselines, monitor IAR4D processes, monitor the generation and use of innovations, and to evaluate their impacts. These data will be collected at several levels, namely plot, household, village, innovation platform and district. Taskforce data will be pooled at PLS level to obtain a PLS perspective, and subsequently at programme level to obtain a Sub-Saharan Africa perspective (see Figure 6).

8.1. Baseline surveys for IP and community level characteristics

Baseline surveys, field observations and focus group discussions will be conducted to benchmark pre-treatment characteristics of IPs, site characteristics and baseline levels of outcomes predicted under the IAR4D approach: number, variety and time to develop innovations; knowledge and behavioural outcomes (adoption, input supply, input demand, volume of sales), market outcomes (output supply and consumption demand), and productiv-

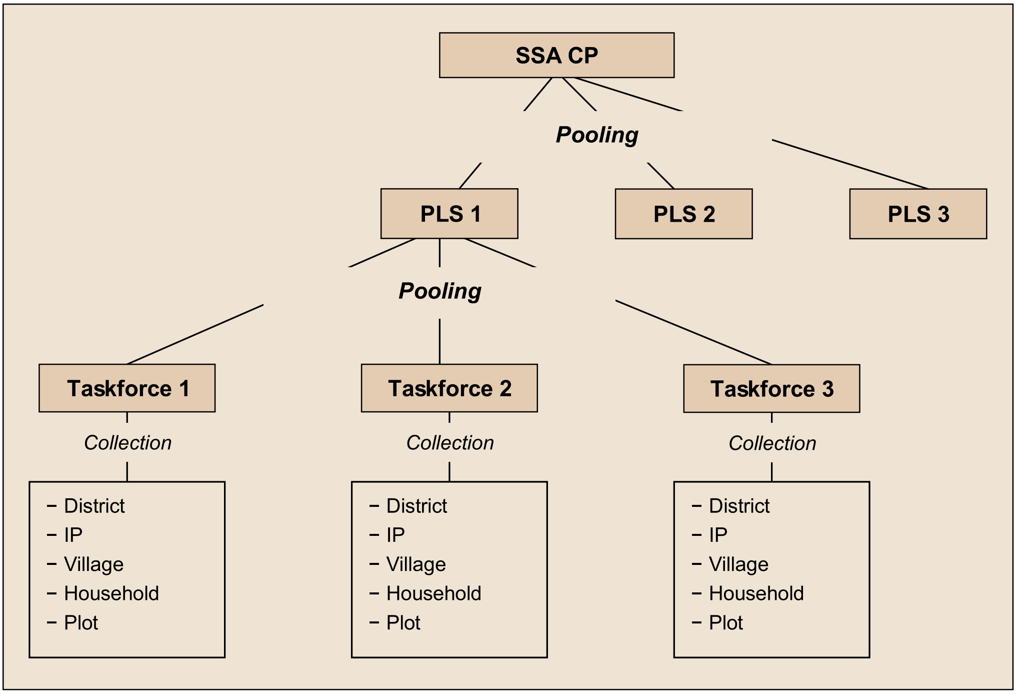


Figure 6. Levels at which SSA CP data will be collected and pooled

ity outcomes (yields, technical and allocative efficiency, and profit); and impacts (incomes, livelihood assets and equity). Several indicators will be used to measure outcomes. These will differ with context. Questionnaires will be designed for comparison within an IP over time and across IPs.

To generate counterfactuals, surveys and field observations will be conducted in the comparison sites and villages assigned to conventional and non-IAR4D-non-conventional treatments. Key players in the innovation systems—such as public and private agricultural researchers, extension, farmer leaders, traders, dealers, lenders and key informants—will be interviewed to benchmark characteristics of innovation systems and baseline levels of outcomes as for the IP sites.

8.2. Baseline survey for household and village community characteristics

Baseline surveys, observations and focus group discussions will be conducted to collect data on household- and village-community-level characteristics, and behavioural, efficiency, environmental and welfare outcomes. Surveys will track feedback, information diffusion, awareness and knowledge changes, adoption, and market effects of innovations and spillovers using the Miguel and Kremer (2004) approach and other methodologies.

8.3. Evaluation surveys

Follow-up evaluation surveys and qualitative assessment studies will be conducted in the third year (2010) to assess the implementation process; document all the intermediate steps of the research-to-impact pathway and conditioning factors; assess participants' subjective reactions to IAR4D; identify subgroups experiencing greater or lesser impact than the sample as a whole; and measure changes in outcomes at the levels of the IP, household, community and market. Follow-up surveys will use indicators used in the baseline surveys to measure outcomes.

9. Data analysis: estimation Issues

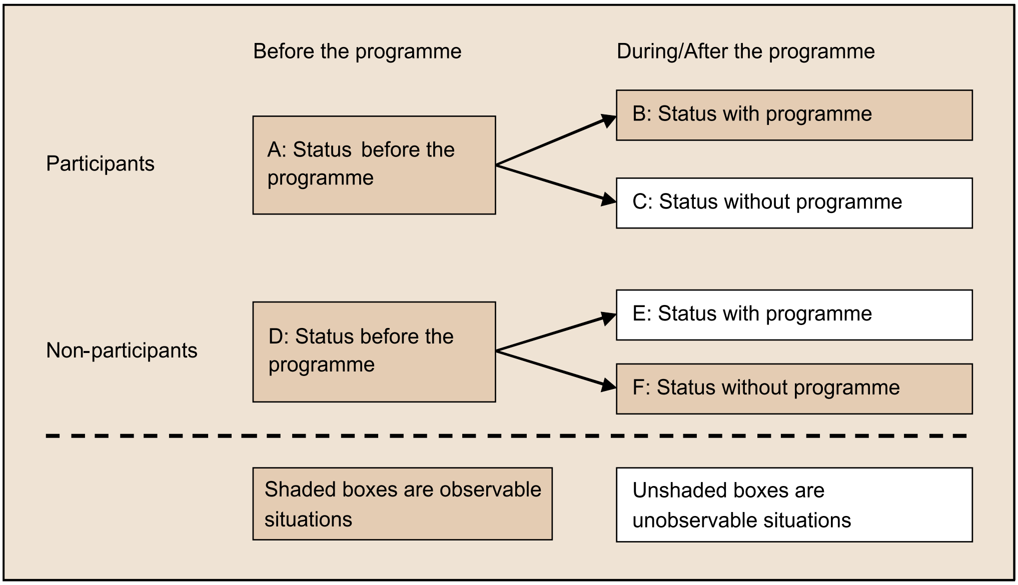
9.1. Quantitative analysis

The fundamental evaluation problem in estimating the effects of the IAR4D approach is the attribution problem and constructing counterfactuals. The counterfactuals, that is, what would have happened to participants and non-participants without the programme, are never observed. In Figure 7, **A** and **D** and **B** and **E** can be observed, but not **C** and **F**. How can **B–C** be estimated if there are no observations? An assumption often made is that **E=F**, that is, there is no self-selection among programme participants, scale effects or spillovers. But the programme might affect prices in general, and there may be social and economic interaction effects on participants and non-participants. Potential spillovers and scale effects determine selection of with and without programme analytical approaches.

Estimation methods include the following.

- Longitudinal comparisons of participants' outcomes **(B–A)**, that is *before and after* treatment.
- Cross-sectional comparisons of participants' outcomes versus non-participants **(B–E)**.
- Social experiments **(B–E,** with **A=D** and **C=F)**, i.e., random assignment ensures treatment households participating in the programme and non-treatment households are statistically equivalent. The key assump-

Figure 7. Cause-and-effect attribution problem of impact.



tion is **E=F**, that is there are no effects of the programme on non-participants.

- Difference-in-differences estimator **((B–E)–(A–D))**, which accounts for fixed differences between outcomes of participants and non-participants.
- Matching **(B–E|A(X)=D(X))**, i.e., compare outcomes of participants and non-participants who are similar in observed characteristics.
- Econometric methods **(B–E|X)**, which account for impacts of observable and unobservable confounding factors **(X)** on outcomes.
- Combinations of the above (e.g., difference-in-difference with econometrics: **(B–E|X1)–(A–D|X0)**).

Longitudinal comparisons of participants' outcomes are in general most commonly used. Differences between outcomes of project participants after and before the programme are used to measure the effects of the programme. Pre-programme data are used to infer the missing counterfactual outcomes for participants. An advantage of this before–after estimator relative to other estimators is that it can be implemented even when data are available only on participants (Todd, 2008). However, longitudinal comparisons are likely to produce biased estimates of treatment effects, because they do not separate such effects from possible confounding factors **(A≠C)** such as bad or good weather, prices of crops, and long-term trends.

Cross-sectional comparisons of participating and non-participating households can improve attribution to treatments. They involve using data on a comparison group of non-participants to infer the counterfactual outcomes for participants. This estimator has the added advantage of requiring minimal data. However, its estimates will be biased if participants and non-participants are very different (presence of selection bias, e.g., poor and rich households), because the estimated cross-sectional differences will not be due to the programme alone **(A≠D, E≠C)**.

A social experiment overcomes selection bias by randomizing placement so that all observational units within some defined set have the same probability *ex ante* of receiving the treatment. If the treatment assignment is random and there is full compliance, that is, focal villages in IAR4D sites cannot reject treatment, and focal villages in control approach and non-IAR4D-non-conventional site cannot implement IAR4D, then the assessment of the effect of IAR4D on outcomes such as yield at the household level can be done by taking the difference in means. Random assignment ensures that observed and unobserved characteristics of treatment and non-treatment households have the same distribution, that is, **A=D and C=F**. Statistically, **B–C=E–F=F–C**. Thus, it resolves the problem of purposive placement based on unobserved factors. The assumption of no spillover and scale effects is very important for social experiments: large spillovers invalidate the purpose of counterfactuals.

The randomized design proposed for this study seeks to minimize the spillover problem by assigning the IAR4D treatment at the district/local government area/commune level instead of individual treatment (using control village communities at sufficient distance from treatment village communities). However, randomized social experiments can alter the way the programme works in practice, institutional and political factors may delay randomized assignment, and randomization only yields mean outcomes for the counterfactuals rather than distribution of outcomes and gainers and losers.

The difference-in-differences estimator nets out pre-project differences between participants and non-participants, such as initial differences in wealth, from the final difference. This measures the pure effect of the programme **((B–E)–(A–D))**. The approach can be generalized to multiple periods. The double-difference method has the advantage that it removes selection bias if the effects of selection bias are additive and time invariant and outcomes are not affected by expectations of participation. Therefore, the method can be used even if purposive sampling is used to select households. But double differencing has several problems, including selection bias that may not be time invariant, such as differential growth rates due to different initial endowments, and sensitivity to data quality, since measurement errors are more serious in comparing changes in variables than comparing levels.

Propensity score matching methods involve identifying a sample of comparator non-participants that are as similar as possible to participants in their predicted likelihood of participation, and then comparing mean outcomes. For example, baseline data can be used to select participants and non-participants who are similar in observed characteristics and to compare differences across space or over time. The strengths of this approach are reduced dependence on parametric assumptions and reduced bias from comparing non-comparable observations. The weaknesses are that only selection on observables is addressed and selection bias resulting from unobservables may still remain; reliance on the parametric model to predict participation; heavy reliance on extent and quality of data to predict participation; and difficulties finding comparable non-participants which, in turn, results in sample truncation. Other weaknesses with the approach are that it requires a larger baseline survey, since non-comparable participating and non-participating households are dropped and it affects the population for which impacts are assessed.

Econometric methods account for predictive effects of other factors, that is, the impact of observable and unobservable confounding factors **(X)** on outcomes across individuals or over time. Consequently, they are often used in studies of impacts. Econometric modelling has the advantage that it can control for selection biases by accounting for observable differences between programme participants and non-participants ("selection on observables" and "selection on unobservables"). But econometric approaches suffer from several problems:

- parametric approaches depend on valid parametric assumptions;
- non-parametric approaches rely on large sample size and good data;
- identification of suitable instrumental variables (IVs) is often difficult;
- estimation based on IVs only evaluates impacts due to variation in IVs;
- biases may result from comparing non-comparable observations.

Several developments have taken place in the econometric modelling treatment effects literature that make it a powerful approach for analysing evaluation problems addressed in this study (Heckman, 2005), namely:

1. Development of an explicit framework for outcomes, measurements and choice of outcomes where the role of unobservables in creating selection problems and justifying estimators is modelled.
2. Extensions to analyse subjective evaluations of outcomes and using choice data to infer outcomes.
3. Extension to model *ex-ante* and *ex-post* realizations and evaluations of treatments, regret and anticipation by agents.
4. Development of models for identifying entire distributions of treatment effects (*ex ante* and *ex post*) rather than the mean parameters traditionally estimated by statisticians. These distributions can be used to determine the proportion of people who benefit from the treatment.
5. Identification of distributional criteria allowing for analysis of alternative social welfare criteria for outcome distributions comparing different treatment states.

6. Modelling of simultaneous causality that relaxes recursive frameworks and relaxation of strong "ignorability" assumptions, which allows analyses of social interactions, general equilibrium effects and scale-up effects.

A number of quantitative statistical approaches will be combined to improve the robustness of the analysis (Ravallion, 2008). This is important since no single evaluation method is ideal in all circumstances.

- Multiple methods will be used to increase confidence in conclusions, since each method has different strengths and weaknesses;
- Econometrics or propensity score matching will be used with double-difference estimator to limit or account for effects of pre-project differences between participants and non-participants;
- Econometrics will be used with instrumental variable methods to address potential biases caused by selective participation;
- Propensity score matching will be combined with econometrics to limit sample analysed econometrically to comparable units.

9.2. Qualitative assessment approaches

Qualitative approaches will also be used, including impact pathway analysis, outcome mapping, participatory evaluation, and developmental evaluation. These approaches involve engaging partners and stakeholders to lay out their theory of change and hypotheses about how they expect impact to be achieved. Monitoring and evaluation will then be linked to the realization of the expected impacts. Monitoring will focus on key essential factors to enable outcomes take place, for example, for a market network monitoring will include how household indicators change, and also network indicators if these are part of impact pathways.

The strengths of qualitative assessments include better understanding of processes by which impacts come about and of stakeholders' perceptions. For example, farmers and traders can provide insights about which mechanisms are most important in generating impact. Furthermore, understanding programme details and processes is a precursor for understanding selection issues and identifying instrumental variables in econometric modelling. Because IAR4D is flexible and adaptive, qualitative approaches are important since they allow for adaptation over time. Qualitative approaches are especially useful for organizational learning and change, and understanding the determinants and constraints of IA4D adoption and diffusion.

Qualitative methods will be used to reinforce quantitative methods rather than substitute them. While the quantitative methods will address the question of what the impact is, their qualitative counterparts will address the question of why and how impact is or is not being achieved. Qualitative methods enable better understanding of programme theory and context, which provides knowledge on what is working well and what is not, thereby making assessment more relevant to decision makers. They can be used to establish how to apply lessons learnt elsewhere and therefore resolve external validity.

Understanding the diffusion of information will be essential in assessing spillover effects, which may be revealed by quantitative baselines. However, combining this information with qualitative analysis will produce more solid conclusions. The main weakness with qualitative approaches is that by themselves, they are unable to attribute impacts to interventions. Another problem is that they result in sampling and interviewer biases. For reasons enumerated above, a combination of quantitative and qualitative evaluation approaches is proposed for monitoring, evaluation and impact assessment of IAR4D.

10. SSA CP project portfolio and logframe

The SSA CP project portfolio comprises four projects: one Meta-Analysis project and three PLS projects (see Table 3). The Meta-Analysis project draws on the data and outputs generated by the three PLS projects to test the SSA CP's hypotheses and to derive general principles for the effective implementation of IAR4D. The three PLS projects contribute to the overall objective of the SSA CP by: (a) establishing, facilitating and monitoring the operation of innovation platforms (12 platforms per project); (b) developing and introducing technological, market, policy and other institutional innovations; and (c) evaluating whether IAR4D works in their specific contexts and whether it delivers more benefits than the traditional research–extension approach. The four projects complement one another.

The overall logframe of the SSA CP showing the pathway that links the outputs and their targets for each year; the expected user of the outputs and the intended benefit by end users is summarized in Table 4.

Table 3. SSA CP portfolio of projects.

<table>
<tr><td colspan="3">Three PLS projects:
Objective: Derive process guidelines for implementing IAR4D and generate data for evaluating its feasibility and impacts by: (i) creating functioning innovation platforms; (ii) generating innovations in farmers' fields, and (iii) evaluating IAR4D impacts in specific contexts and for specific objectives.</td></tr>
<tr><td>LK PLS (Eastern & Central Africa)</td><td>KKM PLS (West Africa)</td><td>ZMM PLS (Southern Africa)</td></tr>
<tr><td>Specific objective: More food production and agricultural productivity through diversification and improved market access while improving the use of natural resources.
The LK PLS comprises the following subprojects:
1. More food products and better nutrition at reduced cost and minimal degradation of the natural resource base (subproject lead institution: ISAR)
2. Beneficial conservation and sustainable use of natural resources (subproject lead institution: Makerere University)
3. Wealth creation through agro-enterprise diversification and improved market access (subproject lead institution: CIAT)</td><td>Specific objective: Improve the productivity of farming systems and ensure an efficient use of resources through technical, administrative, marketing and management improvements.
The KKM PLS comprises the following subprojects:
1. Improving livelihoods of rural population in the Sahel through intensification, access to markets, and sustainable management of natural resources (subproject lead institution: INRAN)
2. Sustainable agricultural intensification and integrated natural resources management to improve rural livelihoods in the Sudan Savanna (subproject lead institution: IITA)
3. Developing a multi-stakeholder approach to linking technical options, policy, and market access for improved land productivity in the Northern Guinea Savanna (sub-project lead institution: IFDC)</td><td>Specific objective: Improve the performance of the agricultural value chains through intensification and other technical innovations in high & low potential farming systems.
The ZMM PLS comprises the following subprojects:
1. Improving human nutrition and income through integrated agricultural research on production and marketing of vegetables in Malawi and Mozambique (subproject lead institution: Bioversity International)
2. Integrating sustainable soil fertility management innovations in staple cereal systems and other value chains to enhance livelihoods and environmental systems in Southern Africa (subproject lead institution: SOFECSA/CIMMYT)
3. Efficient water and nutrient use in cereal grains systems in market-based conservation agriculture systems (subproject lead institution: TSBF-CIAT)</td></tr>
<tr><td colspan="3">Meta-Analysis (of the IAR4D Concept) Project
Objective: Derive generalizable principles for implementing IAR4D and evaluate its feasibility and impact across Sub-Saharan Africa.
Implementation of the meta-analysis project will be led by the cross-site research support team, which will work in collaboration with the PLS project teams.</td></tr>
</table>

Table 4. SSA CP logframe

Outputs	Output targets	Intended Users	Outcomes	Impact
1. Empirical evidence of whether IAR4D works, the extra benefits it delivers compared to those delivered by traditional approaches given the same resources and whether it is replicable beyond the SSA CP's test sites	**2008:** Starting conditions (institutional, market, technological, biophysical and assets) characterized; IAR4D evaluation framework developed **2010:** Feasibility of IAR4D evaluated; Cost–benefit effectiveness of IAR4D compared with traditional approaches established; Replicability of IAR4D established.	SROs, NARS, IARCs, national, regional and international policy makers, agribusiness actors, NGOs, farmer organizations, donors	Increased adoption and reliance on IAR4D (increased involvement of non-traditional actors in ARD) Increased investment towards supporting IAR4D processes Increased human and institutional capacity for innovation among ARD actors	Improved returns from agricultural research and development contributing to improved food security, increased household incomes, reduced poverty, and sustainable natural resources management
2. Guidelines/principles for implementing IAR4D	**2008:** Framework for deriving principles and guidelines from social experiment **2009:** Guidelines for establishing innovation platforms tested **2010:** Principles for implementing IAR4D identified.	Researchers (NARS, SROs, IARCs), national policy makers, agribusiness actors, NGOs, farmer organizations		
3. A database of process and impact indicator variables for 36 innovation platforms and their associated research communities and households	**2008:** Database of baseline conditions established **2010:** Time series database of process and indicators variables for 540 villages and 5400 households established and made accessible to the public.	Researchers (NARS, SROs, IARCs), national, regional and international policy makers, agribusiness actors, NGOs, farmer organizations, donors		
4. Methods and tools for designing, implementing and analysing social experiments in Sub-Saharan Africa	**2008:** Good practices for site selection and sampling in social experimentation identified **2009:** Good practices and tools for tracking learning and institutional change **2010:** Tool and methods for impact evaluation in social experiments.	Researchers (NARS, SROs, IARCs)		
5. Potential technological, market, policy and institutional innovations identified, developed and mechanisms for putting them into use analysed	**2009:** At least nine technological and nine institutional innovation identified and tested **2010:** The mechanisms through which at least nine technological and nine institutional innovations can be successfully "put into use" identified and documented.	Farmers and other actors along the value chain Researchers in national and international institutions Advisory service providers		

References

Abadie, A., A. Diamond and J. Hainmueller (2007). Synthetic control methods for comparative case studies: estimating the effect of California's Tobacco Control Program. http://ksghome.harvard.edu/~aabadie/ccs.pdf.

Dantas, E. (2005). The "system of innovation" approach, and its relevance to developing countries. Science and Development Network. http://www.scidev.net/dossiers/index.cfm?fuseaction=policybrief&dossier=13&policy=61.

Dehejia, R. (2003). Was there a Riverside Miracle? An hierarchical framework for evaluating programs with grouped data. *Journal of Business and Economic Statistics 21*(1): pp. 1–11.

Eicher, C.K. (2006). The evolution of agricultural education and training: global insights of relevance for Africa. *Staff Paper* 2006-26. East Lansing, Michigan: Department of Agricultural Economics, Michigan State University.

Heckman, J.J. (2005). The Science Model of Causality. http://www.nuff.ox.ac.uk/users/nielsen/res/Heckman/Heckman_sci-causality.pdf.

Meinke, H., R. Nelson, P. Kokic, R. Stone, R. Selvaraju and W. Baethgen (2006). Actionable climate knowledge: from analysis to synthesis. *Climate Research 33*: pp. 101–110.

Miguel, E. and M. Kremer (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1): pp. 159–218.

Moffitt, R.A. (2004). Forecasting the Effects of Scaling Up Social Programs: An Economics Perspective. Baltimore, MD: Department of Economics, Johns Hopkins University. http://www.econ.jhu.edu/People/Moffitt/progEv.html.

Ravallion, M. (2008). Evaluating anti-poverty programs. In *Handbook of Development Economics* Vol. 4 (ed. by T.P. Schultz and J. Strauss). Amsterdam: North-Holland.

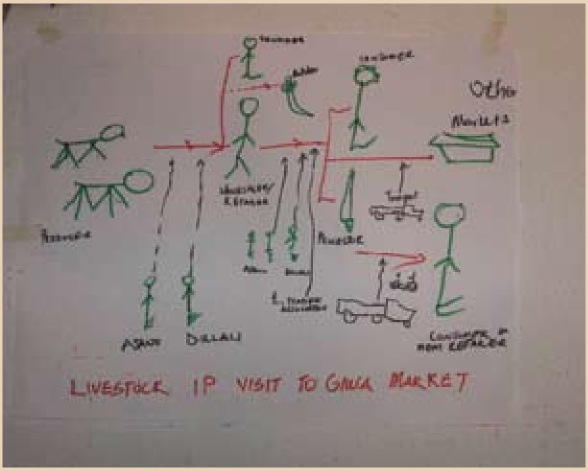
Todd, P.E. (2008). Evaluating social programs with endogenous program placement and selection of the treated. In *Handbook of Development Economics* Vol. 4 (ed. by T.P. Schultz and J. Strauss). Amsterdam: North-Holland.

White, H. and K. Chalak (2006). Parametric and Nonparametric Estimation of Covariate-Conditioned Average Causal Effects. *UCSD Department of Economics Discussion Paper*. San Diego: University of California.

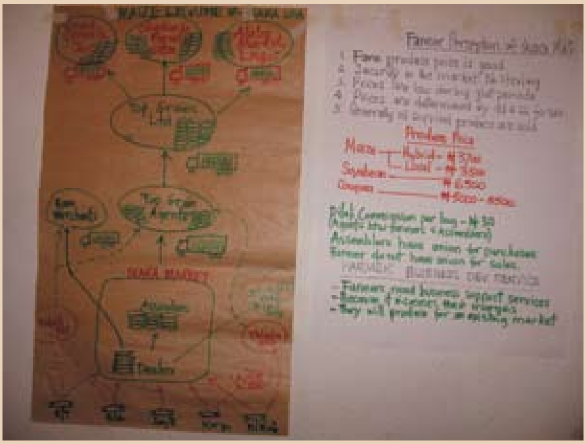
Acronyms and abbreviations

ARD	agricultural research and development
CGIAR	Consultative Group on International Agricultural Research
CIAT	International Center for Tropical Agriculture
CIMMYT	International Maize and Wheat Improvement Center
FARA	Forum for Agricultural Research in Africa
GIS	geographical information system
IARC	international agricultural research centre
IAR4D	integrated agricultural research for development
IFDC	International Fertilizer Development Center
IFPRI	International Food Policy Research Institute
IITA	International Institute of Tropical Agriculture
INRAN	Institut National de Recherches Agronomiques du Niger
ISAR	Institut des Sciences Agronomiques du Rwanda
IP	innovation platform
IPG	international public good
KKM	Kano-Katsina-Maradi (PLS)
LK	Lake Kivu (PLS)
MTP	medium-term plan
NARS	national agricultural research system(s)
NGO	non-governmental organization
No.	number
PLAR	participatory learning and action research
PLS	Pilot Learning Site
PM&E	participatory monitoring and evaluation
RPG	regional public good
SC	Science Council of the CGIAR
SOFECSA	Soil Fertility Consortium for Southern Africa
SRO	sub-regional research organization
SSA CP	Sub-Saharan Africa Challenge Programme
TSBF-CIAT	Tropical Soil Biology and Fertility Institute of CIAT
ZMM	Zimbabwe-Mozambique-Malawi (PLS)

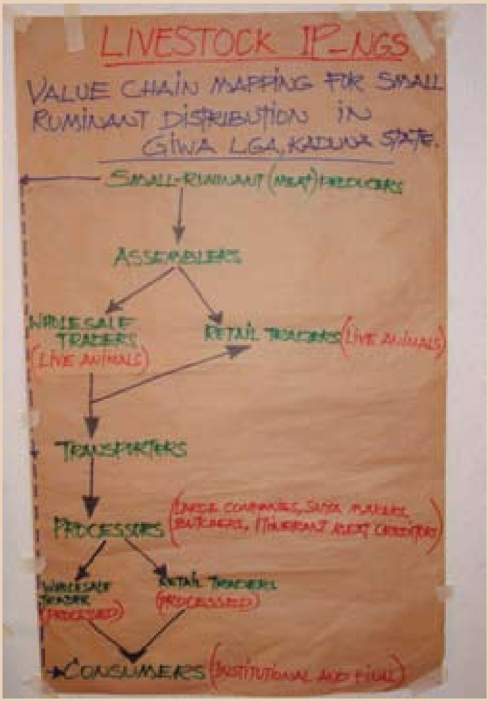
Presentation of cost structure of Tomato Value Chain by Vegetable IP group representative during the PVCA training.

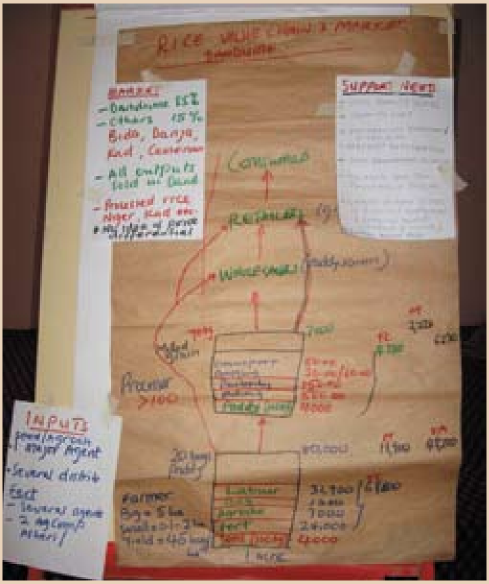


Above and top right: Livestock Market Value Chain in Giwa developed during the PVCA training.



Maize Market Value Chain and farmers' perception of Ikara market (PVCA training).





Rice Market Value Chain in Dandume developed during the PVCA training.