

# **Implementation of the CGIAR Performance Measurement System in 2005: Moving forward**

**(October 2006)**

## **Executive Summary**

2005 was the second year of the CGIAR Performance Measurement System after the pilot year of 2004.

The CGIAR Secretariat, the Science Council and CGIAR Internal Audit Unit jointly assessed this year's exercise with input from Centers. Several new indicators have been piloted this year before including them fully in the PM System. The key findings and suggestions for further refining selected indicators are as follows:

### **1. Seven Performance Elements**

#### **RESULTS**

##### **Outputs -**

- The output indicator based on Center self-assessment will be able to reflect research performance and annual achievement well when it follows a planning process designed for performance reporting. The need for external (SC) assessment of the output reports is likely to be vastly reduced and the steps will include self-reporting and external verification. The quality of this self-assessment indicator depends on the Center researchers' ability to define ambitious but realistic, goal oriented and therefore significant output targets. The achievement of which can be monitored (identified and measured) by the Center itself and subsequently reported as achieved/not achieved and verified within the PM system.
- The SC will continue to comment on relevance, quality and ambition of the outputs and output targets in connection with its Medium-Term Plan (MTP) assessment. It is not the SC's role to accept or reject individual output targets, but rather the SC guides the Centers to aim at a meaningful number of clearly defined targets that suit the processes of planning, assessment of the plans by the SC for their relevance, and self monitoring in the PM system.
- The evidence for fully achieved outputs should continue to be subject to review on a small sample basis, and the outputs to be verified should continue to be selected randomly and in a way that is not advertised to the Centers in advance of submission of the PM System results. This will ensure that appropriate attention is given by Centers to the submissions for all outputs.

##### **Outcome**

- The quality score applied in 2006 (on 2005 data) will need to be reassessed. A more refined rating scale could help differentiate between (i) initial and early

outcomes from different types of research and (ii) major outcomes that will lead to change towards impact.

### **Impact**

- Criteria/Indicators of performance (indicator 3a) should be refined for greater precision. SPIA should consider how the quality and relevance of studies and activities reported under indicator 3a can be more readily and clearly derived from the information and data provided by the Centers.
- Information/data for indicators of performance may be requested in the form of “check-off” lists (with supporting information/data/description) rather than relying on the submitted narration, in order to (i) make the evaluation process less subjective, and (ii) ease the report writing burden for the IA focal points.
- In rating the submissions, a standard method of adjustment for size of Center should be developed, but for this to be meaningful, resources allocated to epIA activities must be provided by Centers.
- The weights applied across the four criteria used in the assessment of indicator 3a should be re-considered.
- SPIA suggests that the scores received by Centers in 2006 for 3B (covering 2003-2005 period) will carry-over for the next two years until 2009 when the next assessment is undertaken (covering 2006-2008 period).

### POTENTIAL TO PERFORM

#### **Quality and Relevance of Current Research**

- Merits of the alternative approaches to calculating the publication indicators should be examined (e.g. rolling averages)
- Additional evidence material for verification purposes should be submitted at the time of reporting. The provision of a template listing the needed evidence data would be helpful
- Better definition of boundaries of what constitutes a “center publication”, and refinement of the definition of “full-time equivalent” of scientist time.

#### **Institutional Health**

- **Governance** – The questionnaire should be further refined by (i) further clarifying definitions and formulations as indicated in the IAU report (ii) capturing to additional Board responsibilities and practices (e.g. Board meeting practice and attendance). Centers should be requested to submit additional evidence material along with the Board statement..
- **Culture of Learning and Change** - The questionnaire should be further refined by further clarifying definitions and formulations as recommended in the IAU report. Centers should be encouraged to develop internal systems for tracking staff development data.
- **Diversity** – For verification purposes Centers should be requested to submit evidence material at the time of submission (a template could be provided). A cut-

off date at which the indicator should be calculated (e.g. end of the year) should be clearly indicated in the guidelines.

- **Exploring a new indicator** - Given the relevance and importance of adequately maintaining and rehabilitating global public goods (e.g. genebanks) a logical consequence is to measure a Centers ability of maintaining global public goods (e.g. genebanks, databases) generated by Centers in the past. Possible new indicators should be explored.

### **Financial Health**

- The two piloted indicators should be maintained;
- Benchmarks used for indicator 6a/b (solvency and stability) will be reviewed again by the finance community and recommendations should be made;
- A fifth indicator is proposed: audit opinion on the financial statement.

### **STAKEHOLDER PERCEPTIONS**

- Measures of Stakeholder Perceptions about Center performance provides a relevant “customer satisfaction ” dimension in the PM System, a critical dimension for assessing the performance of any institution
- Given that a comprehensive Stakeholder Perception survey constitutes a major undertaking, this kind of survey should only be undertaken every 2-3 years.
- Survey planning should be coordinated with other surveying activities conducted by the SC and Centers

## **I. Introduction**

2005 was the second year of the CGIAR Performance Measurement System after the pilot year of 2004. Lessons from the pilot year have been used to refine the indicators and definitions used for the 2005 exercise.

The PM system is a dynamic M&E tools, where adjustments over the initial couple of year after its establishment can be expected aiming at continuously improving the validity and reliability of the performance measures. The following paper describes some of the key observations made with the indicator set used for 2005, and draws lessons and conclusions for refining the indicators to be used in the coming year.

## **II. The Seven Performance Elements: Observations and lessons learnt**

The PM system comprises a diverse set of performance indicators grouped under seven performance elements which fall under three dimensions: (i) Results (ii) Potential to perform, and (iii) Stakeholder Perceptions.

The following describes the key observations made about the reliability and validity of each performance indicator, and draws some conclusions for planning next year's exercise.

### **A. RESULTS**

#### **Element 1: Outputs**

The PM indicator for outputs is the “% MTP output targets achieved in the reporting year”. It is based on the Centers' self-assessment of full achievement of each output target. This year was still a year in transition, because the base document for reporting the output targets was the MTP 2005-2007, which was not designed for the PM purpose. The MTP 2006-2008 was the first MTP aligned in its design with the PM System.

#### **Observations**

In the instructions for the PM 2005 submissions, the disconnect between the definitions of the MTP 2005-2007 and the PM was acknowledged. Consequently, Centers were instructed not to report every milestone/verifiable indicator listed in the MTP 2005-07, but to describe in the PM report achievements, deliverable to intended users, that represent as far as possible research achievements planned for 2005. Centers were asked to report on all MTP projects and outputs listed in MTP 2005-2007. They were also given the opportunity to explain major departures from the plans and subsequent changes to the expected targets and their achievement.

The SC determined that the suitability and quality of the indicators provided by the Centers to reflect output performance depended on two independent parameters:

- (1) Match between the PM output target submission and the MTP of 2005-2007;

- (2) Acceptability of output targets determined by whether the reported output targets meet the definition of an output target, i.e. whether they were deliverables to the intended users.

**(1) Matching of PM output target submissions with the MTP 2005-2007**

Because the 2005-2007 MTPs used a different set of definitions than the PM system is now using, Centers had leeway to decide what to identify as an output target following the definition. The SC decided to comment on the matching of the PM Output submission with the MTP 2005-2007 only at the Project and Output level to help assess the extent to which the output data reflect Center's major research programs. In a few cases where the PM report did not show either the Projects or the Outputs of the MTP 2005-7, the SC examined whether the reported output targets covered the research results planned for 2005 within the original milestones and verifiable indicators. With the exception of a few Centers that had changed their project structure, the submissions reflected the MTPs of 2005-2007 well.

**(2) Acceptability of output targets**

The SC examined all the output targets listed by the Centers in their PM submission against the definition. The SC compiled its assessment criteria in detail into a set of generic criteria for output targets given in Annex 1. The targets were identified as either acceptable or unacceptable.<sup>1</sup> Because it is a transitional year, the SC broke the acceptable category out into "okay" targets, those that met the definition of an output target, and "unclear" targets, those that implied a deliverable research result, even if the nature of the deliverable was somewhat unclear. The purpose of the "unclear" category was to illustrate the importance of clarity in both planning and reporting output targets and to provide individual feed-back to each Center on the quality of their output targets. The SC views this year as a transitional year, as mutual learning will help to improve the planning of identifiable, measurable, clear and sufficiently sizeable research output targets to be reported in the annual PM exercise.

- ***Quality of output targets***

Several Centers reported their 2005 milestones as the output targets. While in many cases the milestones indicated an intended research achievement, this approach tended to yield a large number of targets, some of which were minor. Centers also reported activities that had been listed in the MTP 2005-2007. Consequently, even if many output targets were considered acceptable because they implied a deliverable research result, many were described in vague and generic terms, creating ambiguity about what achievement had been intended (e.g. case studies relating to innovation systems executed; new uses of crops for food and non-food products are developed). It is likely that this is a problem of finding the adequate formulation rather than a problem of the objectives of research and achievement of results per se.

---

<sup>1</sup> Note: The SC followed the same main criteria as for its MTP 2006-2008 assessment for identifying unacceptable output targets. The following types were excluded: 1) outcomes to be reported under different PM indicators; 2) impacts to be reported under different PM indicators; 3) publications defined by count, not described by content, to be reported mostly under a different PM indicator; 4) on-going activities; 5) process milestones that do not represent a deliverable research achievement

Of all the output targets listed in the PM submissions 75% were considered acceptable; of those 72% were considered “okay” and 28% “unclear”. Furthermore, a common problem even of the “okay” targets was that they lacked clarity about the nature of the deliverable achievement and often stated a completed activity (e.g. “carried out research on empowering communities through food-based programs in Ethiopia” instead of specifying the policy strategies intended to be achieved for delivery). This reflects the problem of retrospective reporting which is likely to be overcome when Center reports are based on a forward looking annual plan in the MTP following the agreed definitions.

Nearly all of the 25% unacceptable output targets were process related milestones or described initiation of or on-going activities. Commonly, unacceptable output targets described internal communication, such as project meetings. With some Centers there was large variability between Projects in the quality of the output targets. Examples of acceptable and unacceptable output targets in different categories are given in Annex 1 and 2.

In its commentary the SC calculated what the indicator would be for acceptable output targets only. In nearly all cases the re-calculated indicator was only marginally different, higher or lower, from the indicator based on the Center submissions. This shows that while accuracy in defining clear output targets was and is essential for improving the quality of the indicator, it did not influence the indicator scores for any practical purposes. Centers obviously did not include on-going activities and process milestones to improve their scores, as achievement of these kinds of targets was not reported more often than achievement of the appropriate output targets.

- ***Classification of output targets***

In the PM exercise the output targets are classified as materials, practices, policy strategies, capacity or other kinds of knowledge. Frequently, the output targets were misclassified, which prevents any meaningful summary analysis of what the Centers produce. It also made it more difficult to understand what the intended deliverable output targets were and consequently will complicate the verification of those achievements. In particular, output targets of capacity or practices were often erroneously classified as materials (e.g. publications released on participatory research and participatory monitoring and evaluation theme; multimedia training product prepared on farmer groups conducting adaptive research for agroenterprise development; establishing new market information systems). The PM instructions need to be revised to provide clearer description of output target categories.

- ***Magnitude of output targets***

The SC did not comment on the number or significance of the output targets. However, the MTP and the PM system are aimed at defining and measuring progress in delivering output targets that are of sufficient magnitude to be used by external users and contribute to an impact on the overall CGIAR goals. The SC observed that in the PM submission Centers reported very different numbers of output targets

irrespective of Center size or the type of research they do, ranging from 30 output targets by CIFOR to 148 by IFPRI and from 31 by CIMMYT to 303 by IITA. While acknowledging that different kinds of research have different time lines and different resource requirements, the SC encourages Centers to set output targets that represent achievement of intended solutions to a defined problem or challenge. In the MTP guidelines the SC has suggested as a very rough guidance that for an annual investment of US\$ 500 000 the Center should be expected to achieve one output target annually. This does not mean that each output target requires an investment of 500K, or that output targets should represent a year's work.

### **(3) Verification**

Two outputs were selected randomly by the Internal Audit Unit (IAU) of the CGIAR from the population of outputs considered acceptable by the Science Council and which were classified by the Center as fully achieved. The verification that followed was not to confirm the composite result submitted for the indicator but to test check the availability of evidence and identify any challenge that Centers and reviewers may encounter in any future verification exercise for this indicator. IAU did not conduct a technical evaluation of the evidence, but made an effort to select outputs from a spread of output categories.

### **Conclusions**

- The output indicator based on Center self-assessment will be able to reflect research performance and annual achievement well when it follows a planning process designed for performance reporting. The need for external (SC) assessment of the output reports is likely to be vastly reduced and the steps will include self-reporting and external verification. The quality of this self-assessment indicator depends on the Center researchers' ability and motivation to define ambitious but realistic, goal oriented and therefore significant output targets the achievement of which can be monitored (identified and measured) by the Center itself and subsequently reported as achieved/not achieved and verified in the PM system.
- The SC will continue to comment on relevance, quality and ambition of the outputs and output targets in connection with its MTP assessment. It is not the SC's role to accept or reject individual output targets but rather the SC guides the Centers to aim at a meaningful number of clearly defined targets that suit the processes of planning, assessment of the plans by the SC for their relevance, and self monitoring in the PM system.
- The evidence for fully achieved outputs should continue to be subject to review on a small sample basis, and that the outputs to be verified should continue to be selected randomly and in a way that is not advertised to the Centers in advance of submission of the PM System results. This will ensure appropriate attention is given by Centers to the submissions for all outputs.

## **Element 2: Outcome**

The PM indicator for research outcome is a SC's rating of Centers' submission of five statements of their most important outcomes in the reporting year based on outputs that were achieved in the three preceding years. Outcomes are the external use, adoption, or influence of a Center output(s) by partners, stakeholders and clients.

This indicator was tested in 2005 but Centers' submissions were not rated. Very large variability was then observed in the submissions, many of which described outputs and activities rather than outcomes. The SC identified a need for Centers to more carefully distinguish between outputs and outcomes. Activities are the work that generates outputs, outputs are the products of Center activities, and outcomes are the adoption, use of or influence of those products. These outcomes in turn lead to the impacts related to the overarching goals of the CGIAR; that is, reducing poverty, increasing food security and enhancing the environment.

This year the instructions were improved and Centers were asked to provide information about the outcome, in addition to the outcome statement:

- What output(s) resulted in the outcome? In which MTP was the output first identified?
- Who used/adopted or was influenced by the output? Identify the kind of people or entities affected and their location by country.
- How was the output used or adopted? What was the nature of the influence of the output?
- What is the evidence for this outcome? Specifically, what kind of study was conducted? Who conducted it?

Centers were also asked to provide evidence of the outcome either in form of reference of a formal document or source of other evidence.

### ***Process of rating by the SC***

This year the assessment criteria were focused on the quality and appropriateness of the Outcome submissions for 2005. The SC followed three assessment criteria for deriving an overall rating for the Center submission:

- Criterion 1: Of the five cases, how many are truly outcomes? Each case is scored 1 if it is an outcome and 0 if not. Total possible score for this item = 5.
- Criterion 2: Of the cases that scored "1" in the first item, how many have been clearly and specifically articulated? Each case that scored "1" in the first item is scored "1" on this item if it meets the criterion and "0" if it does not. Total possible score for this item = 5.
- Criterion 3: Of the cases that scored "1" in the 2nd item, is there a verifiable source for the outcomes described? Total possible score for this item = 5.

### **Observations**

The outcome reports were considerably better this year than in the pilot year. Nearly 70% of the submissions received were accepted, i.e. the statements represented true outcomes, and were clearly and specifically articulated, and there was a verifiable source documenting the outcome.

Out of the total of 75 statements, 10 did not describe an outcome of a research output. Several described expected outcomes, lessons from an activity or outputs. A few of these described a workshop or a conference as the output and developments deriving from these as the outcome. However, an event *per se* is an activity that may lead to outputs that are deliverable to the intended users among whom the outcomes occur. Similarly networks and partnerships were not considered outcomes but collaborative arrangements that are expected to produce research outputs that in turn lead to outcomes. Also raising awareness alone was not considered an outcome.

It is important to distinguish research outputs and outcomes planned in the logframes and reported in the PM system from process and process related milestones that lead to internal “outcomes” on research management or similar internal processes. These activities and processes are essential for managing projects and for collaboration processes but not to be included in the output monitoring of the MTP and PM that focuses on external users and changes influencing them along the impact pathway. The identification of acceptable outcomes depends on the planning of clear outputs and output targets in the MTP. The SC believes that careful planning at the MTP stage will help Centers to monitor the expected external outcomes from their research and subsequently report documented outcomes in the PM system. This year’s outcome submissions showed the importance of specificity in the reports. The submissions that were specific could more clearly be identified as cases of use, adoption, or influence. In addition, more specific reports of use, adoption, or influence tend to be more easily verified.

Further discussion is needed to clarify the nature of outputs produced by partnership and capacity-building activities so that meaningful output targets that respond to identified needs are planned. Then, outcomes based on those outputs can be more clearly assessed. In this year’s PM data, some outcome statements appropriately identified how capacity building outputs had led to NARS adoption and use of new approaches in their own research planning.

The SC observed large differences in the magnitude of the outcomes and in the stage of adoption and use. In the more border line cases, Centers were reporting raised interest among potential users of their research outputs, which could be considered a precondition for having the Center’s research products used and adopted by or influencing the intended users, but not outcome as such. They were also reporting outcomes from pilot testing, when it would be more appropriate to estimate and document the outcomes once the final output has been delivered. An example of the difference in the significance of the outcome is found in breeding: use of genetic materials by NARS in breeding is an outcome from Centers making their improved materials, the output, available to the NARS. However, it is a much earlier step in the impact pathway than release of varieties following the above mentioned breeding and testing or the subsequent adoption of those varieties by the farmers.

### **Conclusions**

- The quality score applied in 2006 (on 2005 data) will need to be reassessed. A more refined rating scale could help differentiate between (i) **initial and early outcomes**

from different types of research and the (ii) **major outcomes** that will lead to change towards impact. Such a scale can now be developed that there is a common understanding of the expectations for outcome studies among the Centers. Examples of clear outcomes at different stages of use are given in Annex 3.

### **Element 3: Impact**

The PM indicator for impact focuses on Centers' capacity to measure the impacts of their previous research through *ex post* impact assessments (epIAs) and to promote an impact culture. EpIA is here defined as empirical analysis that endeavors to assess the contribution of the research to CGIAR goals of poverty alleviation, food security and environmental protection. Within the PM exercise, epIA performance has been appraised by SPIA members through analysis of the following two indicators:

- overall IA performance evaluated annually using four broad criteria (3A indicator)
- evaluation of two Center epIA case studies for rigor using seven explicit criteria (3B indicator), assessed every three years.

For indicator 3A all SPIA members, and for indicator 3B six external impact assessment experts were involved in compiling the impact indicators. The assessment results and scores were submitted to the Science Council task force for confirmation.

#### ***Impact indicator 3A: Overall Impact Assessment performance***

Each Center was requested to submit a 1000-word narrative documenting their epIA activities, outputs and outcomes during 2005 with a particular emphasis on four key elements (criteria): (a) epIA studies; (b) innovation in and advancement of epIA; (c) communication /dissemination and capacity enhancement; and, (d) impact culture. Elaboration on each of these elements, specific indicators of performance and relative assessment weights were provided to Centers in annex 3 of the "Description of the Performance Indicators for CGIAR Centers (2005 data)".

SPIA members independently evaluated and scored each submission using a 1-10 scale, where the higher the rating the better the result of the assessment.. Assessments were made for each of the four criteria and three-to-five sub-criteria, on the basis of data provided by the Center against the indicators of performance. Individual evaluations and scores were shared between SPIA members in a closed session meeting of SPIA in Benin. After a thorough discussion of the scoring processes used, SPIA members re-evaluated and submitted their revised scores before a final consolidation. The revised scores were averaged, using the agreed weighting factors and formed the basis of a collective SPIA score for each of the four criteria and for the overall 3A impact score for the Center.

#### **Observations**

Although much improved, compared to last year's submission, there remains considerable variability in the type and quality of indicator 3A-related information

reported. Some Centers followed the guidelines carefully and provided relevant information on the different indicators of the four criteria; others seem to follow neither the format nor the instructions provided. SPIA has rated the submissions on the basis of how comprehensive the information was in the context of the guidelines. Instructions and guidelines for each criterion should be carefully followed in future.

There was a mix of IA studies reported for Criterion 1 (epIA studies), including completed, current and planned IA activities, but also from system-wide program initiatives. Only Center-led, completed studies have been counted towards evidence of effort--only these should be reported. In some cases, Centers reported the IA studies by author and year. This is inadequate. A title and brief description of the study (one sentence minimum) must also be provided, sufficient to assess meaningfully the scope, depth and type of impact analysis. Completed studies should be ex-post, not ex-ante, and not M&E (or process) type of analysis or donor project evaluations.

Criterion 2 (innovation in epIA methodology) proved difficult to evaluate in many cases due to: (i) the limited information provided by Centers, and (ii) the “indicators of performance” for Criterion 2 lacked sufficient specificity. SPIA will consider changes and improving these. Another problem relates to double counting: often the same IA studies were listed under both Criteria 1 and 2, just as methods workshops were listed in some cases under both Criteria 2 and 3 (Communication). SPIA will be considering whether to discount information when submitted under more than one criterion. Issues related to quantity vs. quality of evidence provided under each criterion, e.g., communication dissemination, will also need to be thought through more carefully in next year’s guidelines. There was considerable variability in the information provided for Criterion 4 (impact culture). Some were mere statements/claims/narratives with no evidence to back them up. Such statements will not be considered as evidence of adequate efforts. Solid evidence/data/examples are required to back-up statements or claims.

## **Conclusions**

The following should be considered for implementing 3A impact indicator next year:

- Re-word Criteria/Indicators of performance for greater precision. For e.g., under epIA studies, only completed reports, publications, etc., are acceptable. For other criteria, simple statements are insufficient; only specific titles of reports, training workshops, media/communication venues (with brief descriptions) will be counted.
- Information/data for indicators of performance may be requested in the form of “check-off” lists (with supporting information/data/description) rather than relying on the submitted narration, in order to (i) make the evaluation process less subjective, and (ii) ease the report writing burden for the IA focal points. Perhaps an annex with all references of the papers, publication briefs, conferences, etc. should be available (though not necessary to submit).

- Consideration should be given by SPIA for evaluating the quality and relevance - not just quantity - of studies reported under epIA, conferences, etc. and giving differential rates to some higher value (peer reviewed) publications, communication events, etc. Also, a distinction should be made between synthesis type studies and new benefit-cost analyses (latter carrying more weight).
- In rating the submissions, a standard method of adjustment for size of Center should be developed, but for this to be meaningful, resources allocated to epIA activities must be provided (only about 1/3 of the Centers provided that information).
- The weights applied across the four criteria should be re-considered, e.g., weight on methodological advancement may be reduced -- difficult to really measure, and a more consistent approach to scoring by SPIA members will be adopted.
- The word limit for 3a narrative submissions should be expanded to 2000 words (from 1000 words) to allow sufficient (but brief) descriptions of epIA studies completed and other epIA related activities.

### ***Impact indicator 3B: Evaluation of two Center epIA case studies for rigor***

Each Center was asked to submit their two best epIA studies completed during the period 2003-2005 that document the difference that the Center has made in achieving CGIAR goals. This should be a once-every-three-year assessment (in recognition that conducting high quality epIA is time and resource exacting). The criteria SPIA used for evaluating these epIAs were provided in the annex of the general PM guidelines (i.e. Checklist for SPIA Rating of a Center's EpIA Study Rigour).

SPIA relied on the input from external reviews carried out by six independent IA experts in making its assessment of the two Center IA case studies. In addition, for scoring (0-4 scale) each of seven quality-related criteria (reflecting transparency and appropriate methods considerations), the external reviewers also provided brief commentaries in support of their scores against each individual criterion. SPIA was able to use these in adjudicating, where necessary, between widely disparate evaluations, in forming its own judgment. A simple scaling-up adjustment (x 2.5) was applied in converting the scores from a 0-4 scale to a 0-10 scale (in this case 0 = not satisfactory). SPIA provided feedback to each of the Centers in the form of brief summaries of the external reviews of the two case studies.

### **Observations**

As with the indicator 3A submissions, there was considerable variation in the type and quality of epIA case studies submitted. As expected, the studies varied in the nature of research output considered (agro-biological, policy recommendation, NRMR, extension, etc), but more importantly in the rigour of analysis for assessing research impact. Some studies focused only on research products and their initial outcomes. Further analysis of economic and other benefits using empirical data on level of technology adoption, changes in cost and productivity would have improved the quality of these assessments.

Some studies confounded the effects of technology, efforts to promote it and partners associated with these efforts. Readers are sometimes left to guess the specific role the Center has played in realizing the impact (Question 1). On methodology, a few studies applied appropriate, rigorous analytical methods to assess the impact of technology (or extension approach or policy recommendation). These are studies published or sent for publication in professional journals. Others are rather weak in their analytical content and hence the impact assessments are not as credible. Presentation of the counterfactual scenario is a critical aspect of research impact analysis, but only some studies provided this (explicitly). Clarity of analytical approach, data used and attribution of benefits to research would have provided a good basis for counterfactuals. Several of the submissions were not considered by SPIA to be epIAs at all. For example, reviews of past IAs, consultant reports, studies in progress, etc. do not constitute epIAs. Notwithstanding these shortcomings, there were a number of good and even some outstanding examples of high quality studies (CIMMYT, IRRI, IFPRI, CIFOR, ICRISAT, CIP, WorldFish) which could well serve as models of good practice in conducting epIA studies.

Overall, SPIA was very pleased with the consistently high quality and comprehensive nature of the 60 external peer reviews (2 reviews per study, 2 studies per Center, 15 Centers). This provided a good basis to “objectively” evaluate the impact indicator evaluating quality and rigour of epIA studies. There is still a danger of “pairing” two reviewers who are consistently “far apart” in their scores – but this was not evident among the pool of this year’s reviewers. SPIA provided feedback to each of the Centers on their case studies in the form of an overall summary evaluation extracted from peer reviewers’ comments. More detailed assessments for each criterion/question could be made available to individual Centers, should that be helpful.

More specifically, SPIA will have to clarify what exactly is meant in Question 1 by the “necessary condition?” From this year’s experience it is clear that the peer reviewers had difficulties in answering this question, and SPIA had to change the weights to accommodate inconsistent responses to this question. Eventually, this indicator was not used as a “necessary condition,” especially when the two reviewers responded differently to this question. The information to assess this ‘necessary condition’ should be explicitly sought from the Center for each study they submit, i.e., there should not be any ambiguities on whether a study meets this necessary condition or not. If it does not—the study should not be reviewed/evaluated.

### **Conclusions**

The following are specific points for consideration when implementing impact indicator 3B in next round (2009):

- Need to re-iterate to Centers the importance of providing an adequate response to Question 1 - specifying clearly the role of the Center in contributing research towards the intervention responsible for change/impact and accompanying cost data. If not included in the publication sent by the Center, that information must

accompany the submission, as an annex. Without that, the submission will not be rated.

- Information on Question 7 (externally reviewed, authorship) must be provided by the Center along with the submission of each case study. It will not be included as a criterion for external peers to adjudicate.
- Submissions should be individual case studies of epIAs, and not books or lengthy manuscripts, or even compilations of case studies (and not exceeding 50 pages). The date of actual publication must be within the defined period, i.e., the last three years. The question of whether re-submissions (three years from now) or revisions of these same epIAs but in a different format (e.g., peer reviewed next time) would be permitted needs further discussion.
- SPIA will suggest to the SC that the scores received by Centers in 2006 for 3B impact indicator (covering 2003-2005 period) be carried over for the next two years until 2009 when the next assessment is undertaken (covering 2006-2008 period).
- A suggestion was made at the last SPIA meeting that SPIA members should select from the PM 3B submissions the top-ranked epIA case studies for producing Impact Briefs. These top 5-6 epIAs would be 'showcased' on the CGIAR Impact website. It is proposed that SPIA cover the costs of preparing these five Briefs as an incentive for Centers.
- Experiences with the PM exercise (3A and 3B impact indicators) will be the topic of a designated session at the upcoming IA Focal Point meeting in Nairobi in October, 2006.

## **B. POTENTIAL TO PERFORM**

### **Element 4: Quality and Relevance of Current Research**

This element is composed of four indicators:

- 1) Number of peer-reviewed publications per scientist in 2005
- 2) Number of peer-reviewed publications per scientist in 2005 that are published in journals listed in Thomson Scientific/ISI
- 3) Percentage of scientific papers per scientist that are published with developing country partners in refereed journals, conference and workshop proceedings in 2005
- 4) Science Council endorsed EPMP rating of Center research quality

### **Observations**

#### ***Publication indicators***

The reported results of the three publication indicators were reviewed for all Centers by IAU. IAU found that despite a tightening of the criteria in the 2005 PM System instructions, the calculation of these indicators continued to be subject to some varying

interpretations. Further efforts will also be required by the Centers to quality control the results submitted. There were 10 Centers where the actual result differed from what was initially reported by the Center for one or more of the indicators as some publications clearly did not meet the criteria or computations were incorrect. In most cases the differences were minor, and in three cases the audited result was higher than posted by the Center. The adjusted figures have been agreed with the Centers concerned and corrected in the overall results report.

While there is broad consensus on the relevance of publications indicators in the CGIAR PM System, the appropriateness of calculating the indicator on a per scientist basis was raised during the course of the verification. On the basis that different Centers have different staffing profiles and modalities for achieving publication outputs a number of suggestions came from Centers:

- to either calculate the indicators **using total expenditure per year as the denominator**, or calculate **publications per million dollars of expenditure** per year
- to use a **rolling average** measure for the publications.

During the verification process still inconsistencies in the calculation of “Full-time Equivalent” of scientists time was observed, which requires a stricter definition.

Finally, the requested evidence material supporting the publication measures was not sufficient to allow an adequate verification of the reported indicators without going back to Centers and requesting more material.

### ***EPMR rating by the SC***

As of now, there is no EPMR rating of Center research quality available, but the Science Council has introduced rating guidelines to be used by the EPMR panels reviewing Centers starting in 2006. These guidelines include four main criteria which are also used by the panel in evaluating the quality of science for the EPMR report. When there is a rating for at least one third of the Centers, the SC will conclude whether this indicator allows any credible differentiation between Centers to be useful as an indicator of potential to perform.

### **Conclusions**

- The merits of the alternative approaches to calculating the publication indicators should be examined
- Additional evidence material for verification purposes should be submitted at the time of reporting. The provision of a template would be helpful. The template should include (i) Citation of the publication; (ii) Year of publication; (iii) Name of Center scientist who authored or co-authored; (iv) A consecutive number, so as to cross check against the total number used in the computation. Also the actual computation should be indicated.
- Better definition of boundaries of what constitutes a “center publication”, and refinement of the definition of “full-time equivalent” of scientist time.

- Regarding indicator 4C, guidance is needed on the target range of co-publishing with developing country partners so as to encourage optimal development of different kinds of partnerships

## **Element 5: Institutional Health**

This element is composed of three sub-elements: Governance, Culture of Learning and Change, Diversity.

### **a. Governance**

This component is composed of

- a 26 item check list aimed at verifying whether adequate mechanisms for good governance of Centers are in place and are used. It is composed of questions related to (i) effective organization and management of Board Business (ii) full Board engagement with Center's strategic business, and (iii) demonstration of transparency and accountability.
- a Board statement describing the three most important actions taken by the Board in 2005 for improving the performance of the Board and the difference these actions have made in the performance of the Board. A peer-review panel composed of three experienced Center Board members/former EPMR panel member assessed the submissions in terms of relevance/merit and the anticipated impact of the actions taken.

## **Observations**

**Checklist** - The formulation of the checklist questions had been further streamlined after the pilot phase as some were ambiguous or lacked clear definitions. Formulations of a few questions and corresponding definitions remain to be further refined, and the verification exercise has helped to determine those (details can be found in the report "2005 Data verification by the CGIAR Internal Auditing Unit).

In addition, the checklist could be expanded to also include " Board meeting practice and attendance".

**Board statements**- This second component was piloted for the first time this year, and was meant to complement the checklist (which is mostly capturing whether certain processes are in place) with a more qualitative assessment. The panel used 3 criteria for the assessment

- Is it an important action for improving Board performance?
- Is the action clearly and specifically articulated?
- Is the impact of the action (the difference this action has made in Board performance) clearly and specifically articulated?

The panel observed the following:

- The assessment had to be based on the information available in the board statements and the governance checklist only. In some cases the information provided was very limited making an assessment very difficult. As a result the panel did not think that it would be adequate to make a judgment about the significance of the “anticipated impact” of the described action, as this would require more in-depth insights into Center and Board processes.
- During the assessment the panel also realized that it would have been useful to have additional evidence material supporting the statements available, and that this material should be taken into account in the assessment.
- It should be clarified whether Center Board statements should be assessed in terms of the Board’s action towards fulfilling Board’s **BASIC responsibilities**, i.e. these could be actions that are generally considered as part of a Boards responsibility anyway, or whether the assessment should focus on “**extra**” **actions** that the board is undertaking for improving its performance.
- Both components, the checklist and the board statement, should be better linked, i.e. the checklist could be used for informing the assessment exercise of the board statement.

### **Conclusion**

- The questionnaire should be further refined by (i) further clarifying definitions and formulations as indicated in the IAU report (ii) capturing to additional Board responsibilities and practices (e.g. Board meeting practice and attendance).
- Centers should be requested to submit additional evidence material along with the Board statement.
- A fourth assessment criterion for the Board statements should be included: “Is there a verifiable source for the action described?”
- The request for Board statements should be refined taking the experience of the panel into consideration.

### **b. Culture of Learning and Change**

The 11 item checklist on culture of learning and change covered a range of questions, including the conduct of a staff satisfaction survey, leadership development, staff training (budget and days), internal program planning and review, and number of CCERs completed, and the preservation of institutional memory.

### **Observations**

Separating staff development activities into (i) attendance at international conferences and professional society meetings etc, and (ii) staff training seemed to have yielded more differentiated results. During verification it became apparent that some Centers provide estimates, others provide detailed computations supporting their reports. Therefore Centers should develop systems to capture this data throughout the year so that a special exercise for the PM System will not be required.

Similar to the governance checklist the verification of this checklist detected additional limitations of certain formulations and definitions, which need further attention (details can be found in the report “2005 Data verification by the CGIAR Internal Auditing Unit).

After two cycles of measuring Center performance, it can be expected that Centers (i) have been changing their internal culture of performance measurement and (ii) have put in place internal databases with performance information to be also used as verification trails. These efforts as such would also reflect a Centers culture of learning and change and could be included in the checklist as well.

### **Conclusion**

- The questionnaire should be further refined by further clarifying definitions and formulations as recommended in the IAU report
- Centers should be encouraged to develop internal systems for tracking staff development data

### **c. Diversity**

This sub-element comprised 4 indicators capturing diversity in terms of gender, nationality and seniority.

### **Observations**

All four indicators were fully verified by IAU. The verification process could be simplified if Centers are requested to submit the following information at the time of reporting (a table template could be provided):

- documented G&D goals, a reference to the Board meeting at which the goals were approved, and the text of the relevant Board resolution or equivalent
- list of all the management positions, the rank (e.g. director, senior manager, manager or assistant manager) of the person occupying the position, information on the persons’ gender, and the formula used in computing the percentage.
- define the evidence as a complete list of all the Center’s IRS showing nationality, and the formulae used in computing the percentages.

A **cut-off date** at which the figures should be determined (e.g. end of the year) has to be clearly indicated in the guidelines.

### **d. Exploring a new indicator**

Given the relevance and importance of adequately maintaining and rehabilitating global public goods (e.g. genebanks) a logical consequence is to measure a Centers ability of **maintaining global public goods** (e.g. genebanks, databases) generated by Centers in the past.

Possible new indicators should be explored and could, for example, relate to Center investment in rehabilitation and upgrade of global public goods (i.e. genebanks, databases), or the achievement of milestones identified in systemwide rehabilitation plans of global public goods.

## **Element 6: Financial Health**

Following the pilot year of the performance measurement system, one of the lessons learned from the financial health measurement was that while the initial two indicators used were well established in the system, their adequacy for reflecting the financial health of the Centers should be further examined. As a result of this, the CGIAR finance community has adopted two additional indicators of financial health. The four indicators used are:

1. **Short term solvency (in days)** - current assets plus long term investment minus current liabilities divided by per day operating expenses excluding depreciation. 90-120 days is the recommended acceptable range
2. **Long-term financial stability (in days)** - unrestricted net assets less net fixed assets divided by per day operating expenses. 75-90 days is the recommended acceptable range.
3. **PILOT - Cash Management on restricted operations** – this indicator measures the extent to which a Center pre-finances its restricted portfolio (that affects its overall cash flow). It is computed as restricted donors' accounts receivable divided by restricted donors' accounts payable expressed as a ratio.
4. **PILOT- Indirect Cost Ratio** – this ratio seeks to measure how much research program activity a center is able to support on its institutional cost base. It is computed as indirect costs divided by direct costs expressed as a percentage. The computation of this indicator is audited by an independent public accountant.

## **Observations**

The 2006 experience (using 2005 data), has demonstrated that the additional indicators provide a more comprehensive picture of the financial condition of the Centers. However, the need to continuously evaluate the usefulness of the indicators, including the benchmarks, was also clear from the 2006 experience. The finance community has therefore, agreed to review the benchmarks used for the solvency and stability indicator and refinements could follow accordingly.

## **Conclusions**

- The two piloted indicators should be maintained;
- Benchmarks used for indicator 1 and 2 (solvency and stability) will be reviewed again by the finance community and recommendations should be made;
- A fifth indicator is proposed: **audit opinion on the financial statement**. An audit opinion is an authoritative opinion made by an auditor independent and external to the organization, on the fairness of the content and presentation of the financial statements being presented. Third parties rely on the audit opinion as a

measure of the present financial condition of an organization as a basis for decision-making.

### **C. STAKEHOLDER PERCEPTIONS**

In June/July 2006, GlobeScan Incorporated, a survey research firm, conducted a stakeholder perceptions survey on behalf of the CGIAR. The survey sample built on the collaboration survey conducted by the SC last year.

#### **Observations**

In total 1175 CGIAR Partners and 94 CGIAR Member representatives were contacted. In total 348 completed the survey. The survey completions from CGIAR members represent approximately 75 percent of the CGIAR's 2005 Member funding.

The survey captured perceptions about the performance of the CGIAR as a whole and about all 15 individual Centers. The results provide important insights about the drivers of perceptions about the CGIAR and the Centers. It provides both the CGIAR membership and each of the Centers an important tool for finding out and understanding how their performance is perceived among partners and how to build on identified strength, weaknesses and opportunities.

#### **Conclusion**

- Measures of Stakeholder Perceptions about Center performance provides a relevant “customer satisfaction ” dimension in the PM System, a critical dimension for assessing the performance of any institution
- Given that a comprehensive Stakeholder Perception survey constitutes a major undertaking, this kind of survey should only be undertaken every 2-3 years.
- Survey planning should be coordinated with other surveying activities conducted by the SC and Centers

## **IV. Performance Data Collection**

### **a. On-line Performance Data Collection System**

An on-line Performance Data Collection System for Center data and an assessment form by the Science Council and external panels is in place. For ensuring confidentiality and security, access has been limited to a small group of people, including Directors General, PM focal points, SC, SC Secretariat and CGIAR Secretariat. The site is administered by staff of the CGIAR Secretariat and an IT Specialist advising the CGIAR Secretariat.

The CGIAR Secretariat has already entered into collaboration with the IT group at ICARDA, which will help with redesigning and upgrading the data collection system to meet the mid-term needs of the users of the PM System.

### **Observations**

The PM system requires that Centers report the achievement/not achievement of all outputs and output targets and the category they fall under. This makes the entry form quite comprehensive, but also tedious as data entry cascades from (i) project name to (2) output to (3) output target. Typically Centers would “cut and past” the data from their internal data sheets.

The setting of the data entry for the output element was such that it automatically sorted the output targets under a given project along two criteria “achieved/not achieved” and the “category”, which disturbed the sequence of the output reporting. The setting was not adjusted during the data entry phase in order not to disturb the ongoing data input.

### **Conclusions**

- Data on outputs and output targets should not be sorted in any way, but saved in the order of entry.
- The upload of data from templates should be tested and introduced in the next cycle.
- In addition, it should be explored how the New Financial Information System that includes data on financial planning and reporting as well as the CGIAR Project Portfolio can be expanded to also include MTP output/output target information, which could then be pre-loaded into the PM data collection system at the time of data entry. While both data bases – the NFIS and the PM data base – are enhanced it is important to make sure the both systems are compatible.

## Annex 1 - Guide for Assessing Outputs: Generic Acceptable and Unacceptable Targets

Type of Output	Acceptable Targets	Unacceptable Targets
General	<ul style="list-style-type: none"> <li>▪ New methods and tools</li> <li>▪ New practices</li> <li>▪ New protocols</li> <li>▪ New technologies</li> <li>▪ New databases</li> <li>▪ New software</li> <li>▪ New web tools, modules and interactive applications that are more than dissemination</li> <li>▪ Enhanced Genetic materials</li> <li>▪ New knowledge (information about biological processes, ex ante analysis results etc.)</li> <li>▪ Compilations of information</li> <li>▪ Prototypes when deliverable</li> </ul>	<ul style="list-style-type: none"> <li>▪ On-going activities or processes</li> <li>▪ Initiation of work</li> <li>▪ Completed activities without a deliverable result</li> <li>▪ Outcomes (defined as use of output)</li> <li>▪ Impact of Center work (research related to impact may yield results that are deliverable outputs)</li> <li>▪ Publications (contents should be described as any of the acceptable targets)</li> <li>▪ Initial or preliminary results</li> <li>▪ Research planning and proposals</li> <li>▪ Funding applications</li> <li>▪ Routine services or analyses</li> <li>▪ Application of tools and methods</li> </ul>
Partnership	<ul style="list-style-type: none"> <li>▪ Formal agreements to coordinate activities such as consortia signed</li> <li>▪ Research plan developed with partners involving firm resource commitments?</li> </ul>	<ul style="list-style-type: none"> <li>▪ Partnership planning activities</li> <li>▪ Partnership management and coordination</li> <li>▪ Interaction among partners</li> <li>▪ Support to and establishment of infrastructure</li> </ul>
Genebank	<ul style="list-style-type: none"> <li>▪ New collections made from a defined ecoregion</li> <li>▪ Blocks of accessions received from other genebanks</li> <li>▪ Characterizations or other knowledge accumulation completed</li> <li>▪ Nurseries for a give trait /ecoregion distributed</li> <li>▪ New management procedures that add value to the collection and its use</li> </ul>	<ul style="list-style-type: none"> <li>▪ Routine services and maintenance such as regeneration, cleaning, viability testing, sub-culture</li> </ul>
Capacity	<ul style="list-style-type: none"> <li>▪ Training events delivered</li> <li>▪ Degrees completed</li> <li>▪ Training materials/modules made available</li> <li>▪ Virtual training completed</li> </ul>	<ul style="list-style-type: none"> <li>▪ Meetings, conferences, and general workshops</li> <li>▪ Promotion and demonstration of practices; field days</li> <li>▪ General support to partners</li> <li>▪ Scientist visits</li> </ul>

**Annex 2 - Examples of clear, unclear and unacceptable output targets in each category**

Clear, measurable and sizable output targets	Unclear output targets	Unacceptable output targets
<b>Materials</b>		
Gene systems for conditioning asexual reproduction in cereals identified and Nucellar/ovule-specific and inducible promoters available.	Field results on starch quality from a transgenic clone with waxy starch developed with anti-sense technology ( <i>appears to be an interim result on a clone</i> )	Transgenic cassava for modified starch properties authorized for field tests and planted ( <i>process milestone; the output will be the cassava made available to intended users</i> )
At least twenty bean lines resistant to stem maggot, BCMV and angular leaf spot available to partners in five countries in Africa (Uganda, Tanzania, Rwanda, DR Congo, and Kenya).	Double haploids: interspecific crosses (386), accelerated breeding populations (815), somaclones (3758 in Venezuela; 4440 in Colombia). ( <i>the specific nature of the output to users is not clear</i> )	Production of quality seed potato increased in Afghanistan, Bhutan and North Korea ( <i>outcome</i> )
<b>Practices</b>		
Crop establishment guidelines for direct seeding and transplanting developed for hybrid rices.	Environmentally and Socially Sensitive Area analysis methodology applied in prioritized watersheds ( <i>not clear the output is a new methodology</i> )	Community-based participatory plant breeding activities in Nicaragua and Honduras ( <i>on-going activities</i> )
Options for improved management of trypanocide resistance in West Africa developed	Together with the regional team, an INRM research approach, processes and associated methods that ensure quality, impact and improving synergies of strategic and action research promoted ( <i>the nature of the achieved deliverable output is unclear; promotion is activity</i> )	Use of organic amendments and paring to control plantain and banana nematodes adopted by farmers in Nigeria ( <i>outcome</i> )
<b>Policy strategies</b>		
Economic evaluation of hydrological services aiming to overcome problems of water scarcity by looking at incentives for sustainable practices and resource preservation.	Joint strategy to influence EU Novel Food Regulation developed with UNCTAD and GTZ ( <i>the content of the deliverable policy strategy is not given; the statement could refer to an internal strategy of being influential</i> )	Consortia and alliances of stakeholders to champion conducive policies for the promotion of agroforestry technology nurtured and maintained in Zambia ( <i>activity</i> )
Analysis completed of growth and distributional and poverty impacts of different types of public investment in Western China.	Contributions to the 2006 draft of the WHO guidelines on safe wastewater use in agriculture ( <i>nature and content of the contribution or the policy aspects of the guidelines not specified</i> )	Knowledge on food security and diet quality shared at workshops ( <i>dissemination; the policy related deliverable achievement is not given</i> )
<b>Capacity</b>		
INRAB and University staff, Benin and INERA staff, Burkina Faso	A strategy to address priority global needs of poor small scale crop-	Training in the use of KS tools and techniques provided to over 40

trained in laboratory and field experimentation to test and implement biocontrol strategies	livestock producers revised and updated ( <i>a very broad output target and the delivered capacity not defined</i> )	CGIAR Center staff in 2005 ( <i>internal milestone</i> )
Electronic Training resource (CD) on AnGR developed and used for training in Africa and Asia and made available to NARI's and Universities	Publications, workshops, documentaries, field days implemented by each project ( <i>training workshops can be an output targets, but statement is very about the deliverable targets</i> )	Support participation of national partners in international fora linking environment and development ( <i>activity with undefined targets</i> )
<b><i>Other kinds of knowledge</i></b>		
Potential impacts of biofortified beans and cassava on human health and productivity estimated.	Genotyping and phenotyping databases advanced ( <i>the specifics of the databases and the advancement not given</i> )	International workshops conducted to disseminate results of empirical impact studies ( <i>process related activity</i> )
New global modeling modules developed linking water scarcity and food production and security.	Rainfall simulation experiments and KINEROS modelling of overland flow ( <i>not clear what the deliverable output is</i> )	Dual-purpose food-feed trials progress report ( <i>report on activity</i> )

**Annex 3 - Demonstrated outcomes from Center outputs – examples from PM exercise**

<b>Output type</b>	<b>Output</b>	<b>influence/use in the beginning of impact pathway</b>	<b>outcome further along the impact pathway</b>
material	Introgression lines carrying genes from wild species of rice for resistance to biotic and abiotic stresses (IRRI)	NARES in 6 countries used 40 IRRI-bred introgression lines in breeding and evaluation programs and in marker-assisted breeding	
material	QPM made available in Asia (CIMMYT)		Adoption of QPM in Asia; 200,000 hectares under cultivation in India in 2005
material	Tropical maize inbred lines with resistance to gray leaf spot and adapted to the mid-altitudes (IITA)	The national agricultural research systems in Burkina-Faso, Cameroon, Congo, Ethiopia, Côte d'Ivoire, Kenya, Mali, Nigeria, South Africa, Tanzania, Zambia, and Zimbabwe and private seed companies requested a total of 47 sets of these trials for testing and incorporation into their research programs in 2005	
material	Maize varieties adapted to east African incorporating a natural form of herbicide resistance maize that allows herbicide use for Striga combating; seed coating technology suitable for poor farmers (CIMMYT)	NGOs and extensionists conducted over 7,000 demonstrations of Imazapyr resistant (IR) maize in Kenya in 2005, and Kenyan farmers declared IR maize, combined with seed-coating with this herbicide to provide an effective protection against Striga, as the most effective Striga control measure available.	
material	Lowland NERICA varieties with resistance to major lowland stresses, particularly Rice Yellow Mottle Virus (WARDA)	Over 60 of the new varieties have been selected by farmers in West African countries through Participatory Varietal Selection (PVS). Four of these lowland NERICAs were officially released in Burkina Faso and two in Mali in 2005.	
practices	DIVA-GIS free geographic information system for biodiversity analysis available through a public open-source site (CIP)	Program is used by scientists throughout the developed and developing countries for international and national research and training	
practices	New prototype		Raised bed planting system has

	conservation agriculture implements and practices (CIMMYT)		been adopted on nearly 50,000 ha in Shandong Province. Major savings in irrigation water use is occurring in farmer fields. Many farmers are obtaining higher wheat yields as a result of reductions in crop lodging and the incidence of powdery mildew and sharp eyespot. The bed planting system reduces production costs by allowing farmers to direct seed maize. Several small factories have begun to build appropriate bed planters which are being purchased by individual farmers and service providers.
practices	near infrared methods for soil characterization	Wider use and knowledge of a new infrared spectroscopy method among national partners in Africa and India and international development agencies. The research technology slashes the cost of conventional laboratory analyses and in doing so enables the study of chemical, biological, and physical characteristics of soils and plant materials at landscape scales, opening up a whole new opportunity for research	
practices	Conservation cropping practices including zero-tillage, stubble mulching, crop diversification (ICARDA)		Adoption on 100,000+ha in wheat-growing areas in Kazakhstan and 50-60% of the wheat-phase in irrigated cotton-wheat systems in Uzbekistan, Tajikistan and Turkmenistan. Crop diversification is also being taken up with areas of summer crops increasing to 70,000ha for safflower, 400,000ha for sunflower and 43,000ha for soybean in southern Kazakhstan and Kyrgyzstan.
practices	Integrated management of Andean weevil (CIP)		Technology is used by farmers in more than 30 communities located near or around the experimental site; and hundreds of farmers, agronomists, technicians, school and university teachers of government, private sector, and

			NGOs in the Cusco and Apurimac Departments of Peru
practices	Cowpea-cereals strip cropping system (IITA)		New system has been tested by over 3000 primary farmers and already spread to over 250,000 secondary farmers in Nigeria during 2002-2005.
practices	Research results showing that application of “pour-on”, even to sub-sets of cattle herds, allowed farmers to manage the risk of tsetse-transmitted trypanosomosis. (ILRI)		Wide adoption of this technology as an important component of tsetse control strategy by several Ethiopian state agencies, Agricultural bureaus, NGOs and local farmers’ service cooperatives.
practices	Model for Integrated Aquaculture and Agriculture (WorldFish)		Adoption of Integrated Aquaculture and Agriculture (IAA) in Malawi and out-scaling to Mozambique (188 farmers adopting)
policy strategies	Technological, socio-economic, institutional and policy options to promote in situ/on-farm conservation (ICARDA)		Four countries use wild relatives of almonds, pistachio, pear, prune and other neglected forest species as <i>Crateagus</i> , <i>Rhus</i> , etc in at least 30% of the planned afforested areas. Government forestry nurseries in Syria, Palestine and Jordan multiply annually since 2004 approximately 3.5 million, 1 million and 180,000 seedlings, respectively, of these species. Introduction of biodiversity conservation into educational curricula in the four countries.
policy strategies	Results from research on rattan management and trade showing the high biodiversity value of rattan gardens, and that rattan garden systems can be economically competitive with other cash crops, especially when non-financial benefits are taken into account. (CIFOR)		Policymakers in Indonesia removed a ban on rattan export that threatened livelihood options for poor rattan farmers.
policy strategies	Analysis on the impact of global cotton markets on	Influenced the WTO to rule in 2005 in favor of Brazil in the Brazil-U.S.	

	poverty in Benin. (IFPRI)	cotton dispute.	
policy strategies	More effective watershed development and a range of livelihood management strategies for sustainable use of the natural resource base in Asia (ICRISAT)	The National Commission on Farmers of the Ministry of Agriculture, Government of India has recommended the adoption of the watershed consortium approach for the enhancement of rainfed areas in 250 drought prone districts in India. Under new legislation many of the components for enhancing agricultural productivity and micro-enterprises, for enhancing the incomes of women farmers and the landless developed by the ICRISAT Consortium, have been recommended for widespread adoption nationally in disadvantaged areas.	
policy strategies	Research results from analyses of the informal irrigation sector, especially in urban and peri-urban areas (IWMI)	New draft National Irrigation Policy in Ghana; the first national policy in Africa that supports urban agriculture and recognizes “informal irrigation” as equally important as the Governmental supported “formal irrigation sector.”	
policy strategies	Research results on financial due diligence in forest-related investments and the links between money laundering and forest crimes (CIFOR)	Indonesian government policymakers and multilateral agencies have takes a new approach to curtailing illegal logging and other forest-related crimes.	
capacity	Strengthened institutional and collaborative capacity of NARS and sub-regional networks in Africa (CIAT)		Over 2.4 million rural households (~ 13 million people) in Eastern and Southern Africa were reached with seed of new bean varieties in the last 18 months. Farmers in 12 countries – mostly women – now grow and evaluate 108 new varieties and 9 other bean-based technologies mostly for pest and disease management
capacity	Negotiation support concept; community mapping tools; rapid hydrological appraisal methods (W.	Adoption of Negotiation Support System concepts, methods, and tools to articulate different knowledge and perceptions to negotiate landscape scale natural resource regulation by	

	Agroforestry)	funding agencies and national research and development practitioners in southeast Asia.	
--	---------------	---	--