

## ISC workshop with FP leaders – 28-29 June 2017

FOR DISCUSSION

(PAPER NO. 3)

DRAFT VERSION 20 JUNE 2017

### **Subject: Performance assessment within the CGIAR Program on Forests, Trees and Agroforestry (FTA)**

#### **Issue**

Participants are invited to discuss the issue of performance assessment within FTA, in order to make a recommendation to the ISC. This paper proposes options for such an assessment, which includes assessment of past programmatic performance and assessment of FP leaders' performance in their role in FTA. It proposes to link performance assessment to the forward-looking prioritization process and W1+2 allocation mechanism within FTA. This assessment mechanism would be put in place for the construction of the FTA 2018 POWB.

- **The ISC is invited to discuss the mechanism proposed in this note, both the overall approach and the set of performance indicators (see Annex 1 for a long list)**
- **The ISC is invited to discuss the desired features of the mechanism, and to advise on next steps**

#### **Request from the ISC**

At its meeting in November 2016, the ISC decided that *"After 2017 FTA will continue implementing a performance based allocation for W1+2 funds, in line with the rules agreed in 2015. It will update this rule by **improving the performance assessment process** along the lines suggested by Bioversity and CIAT and by introducing a way to deal with programmatic priorities, once this process is established"* (ISC/M13/D7).

How can FTA assess the performance of each FP in a transparent and equitable manner? The process used in 2016 was a transitional process and has been criticised by two key partners as not sufficiently transparent. FTA needs such a process, which should reflect a common understanding of expectations within the CRP. If possible, it should rely on a set of FTA indicators that are aligned with FTA's research for development priorities and that will also be used for the POWB and in FTA's Annual monitoring reports to the SMO.

#### **Summary of latest ISC discussions (ISC#4)**

The issue of performance assessment of the Flagships and Centers had been discussed over time by the ISC, and came into the fore as FTA needed to find ways to deal with recurrent budget reductions (see

minutes of ISC#2 meeting. As a result, ISC (and later CIFOR BoT) approved a performance based allocation rule for W1/W2 funds, that was put into practice for the exercise 2016, based on 2015 data (see Annex 2).

Concerns were raised that the performance based rule may not automatically result in the desired effects (i.e. improved performance by a weak performer). At the ISC meeting#4 in Paris, in November 2016, Ravi Prabhu, representing CG centers (CIAT, Bioversity, and ICRAF) voiced the concerns of Bioversity (and those of CIAT with regard to a decision of the gender team). CIAT asked for three conditions to be met (i): absolute clarity on performance-based management decisions; (ii) having a warning period; (iii) quantitative matrix of external evaluation. Bioversity stated that there was not sufficient clarity and felt unfairly treated. ISC recognized that there is room for improvement in the performance assessment process and in making the process more robust. Peter Holmgren added that a lot of energy was dedicated by ISC for a small amount of money in this discussion. Given that we are facing even more reductions, ISC should be even stronger in reducing funds for low performance.

AMI suggested that at its next meeting ISC needs to discuss how to better take into account performance at various levels in FTA by having the Director of FTA providing his inputs into the performance of FP leaders. Additionally, FP leaders also need to contribute to the assessment of performance of senior scientists in their flagship.

## **Proposal**

### **1) Accounting for past programmatic performance**

We propose hereby a new mechanism, to take into account the shortcomings of the previous mechanism, in introducing the following elements:

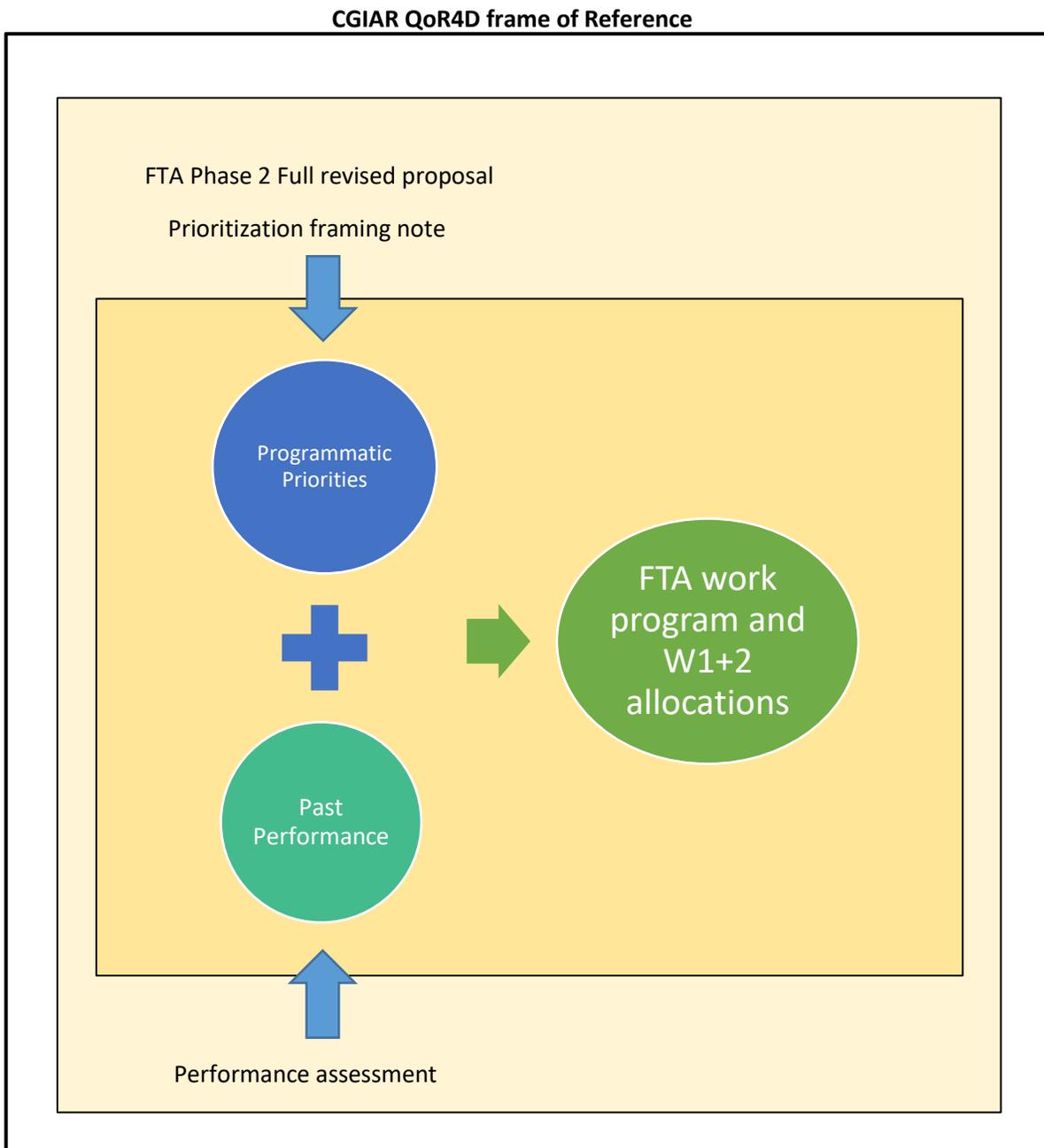
- Performance should be assessed **at FP level**.
- Suppression of the ex-post financial correction mechanism, **be substituted by a “forward-looking” adjustments on the POWB (see Figure 1)**. Past performance would impact on future budget allocations, not on current ones.
- The categories against which performance is assessed ex-post would be the same than the ones framing the prioritization (seven categories, see paper 1), with the addition of an 8th category on Managerial performance:
  1. Relevance/Delivery
  2. Scientific credibility
  3. Legitimacy
  4. Effectiveness
  5. Contribution to IPGs
  6. Strategic value
  7. Program building
  8. Managerial performance

Each category could contain one or several indicators.

A long list of proposed draft indicators, a total of 15, is found in Annex 1. They would apply to FPs, to clusters of the support platform (MELIA, data, capdev, gender), and, if deemed necessary, CoAs within FPs.

- The list in Annex 1 is presented deliberately as a long list, and a basis for discussion.
- Some of these indicators could be simply monitored (or even self-assessed by FPs), and not entering the performance rating process. The ISC is invited to reflect on which ones are important to link to performance rating.
- The ISC is also invited to discuss whether FP performance is measured at FP level, or if the assessment should go down at CoA level (depending of the number of indicators this might increase the workload beyond reason).

Figure 1. Priority and performance-based allocations in FTA



## Computing the overall performance score

Each indicator will be put in correspondence to a 0-10 performance scale (see Table 1), which will help for appropriate averaging across disparate indicators.

**Table 1: Correspondence between indicators/scoring and performance mark used for averaging**

Percentage indicator	Scoring indicator	Associated performance mark (used for averaging)
<50%	Scoring scale 1	0
50-70%	Scoring scale 2	3
70-80%	Scoring scale 3	5
80-90%	Scoring scale 4	7
90-100%	Scoring scale 5	10

The proposed performance mark incentivizes performance, and dis-incentivizes underperformance.

An overall 0-100 performance score for each FP would be computed, with weights that could be applied to the different categories. Table 2 proposes a list of weights in case all categories in Annex 1 are retained for assessment.

**Table 2 Scoring categories and weights**

Categories of indicator	Weight of category	Max score
Relevance/Delivery	3	30
Scientific credibility	2	20
Legitimacy	1	10
Effectiveness	1	10
Contribution to IPGs	1	10
Strategic value	1	10
Program building	1	10
Managerial performance	1	10
Total	10	100

## Effect of performance score

The performance mark at FP, CoA and center level would be used

1) as an input to the prioritization and planning process (see paper 1).

- **For discussion: How the specific performance scores should influence the priority-setting?**  
**One possibility could be that the performance score is used to adjust the priority score of any W1+2 funding proposal (see paper 1) by up to +/- 20%.** Each percentage point over (resp. under) FTA performance average could attract a bonus (resp. a penalty), calibrated so that a performance equal to the average plus (resp. minus) the variance of performance results across FPs attracts 10 points priority bonus (resp. malus). With this, over-performing at the level of the FTA performance variance would attract a +10% bonus of priority score. There could be a limit to this operation set at +/- 20%.

Proposals in the priority-setting process would then be ranked, after taking into account the above corrections for performance.

2) to adjust allocation of management and coordination W1+2.

➤ **The following proposal is for discussion.** A performance-based adjustment parameter to the allocation of W1+2 management and coordination funds would be introduced.

The 200k management and coordination costs a priori granted to each FP would be adjusted in the following way:

- 60% unadjusted (120k)
- 20% (or 40k) adjusted according to the relative performance of the FPs : each point over (resp. under) FTA average attracts a bonus (resp. a penalty), calibrated so that a 50 percent (20k) bonus (resp. malus) over the 40k is obtained for a performance equal to the average plus (resp. minus) the variance of performance results across FPs. Bonus cannot exceed 100%
- 20% (or 40k) adjusted according to the relative performance of the CoAs within the FPs: each point over (resp. under) FTA average attracts a bonus (resp. a penalty), calibrated so that a 50 percent (20k) bonus (resp. malus) over the 40k is obtained for a performance equal to the average plus (resp. minus) the variance of performance results across CoAs (all FTA CoAs considered). Bonus cannot exceed 100%

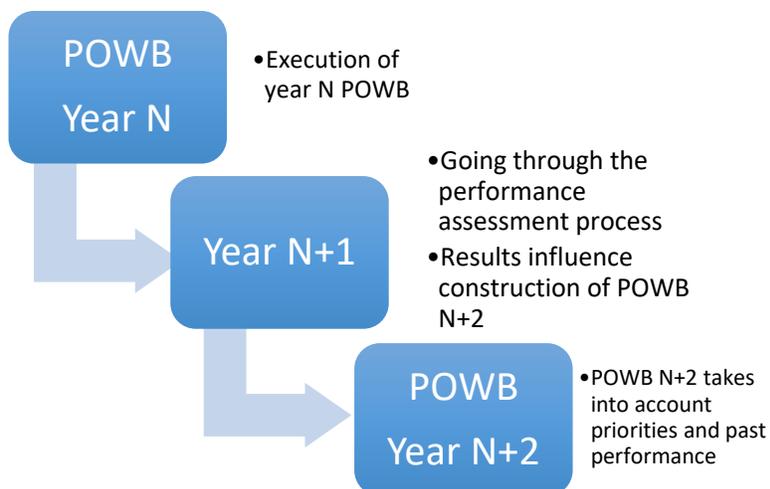
Such bonuses are to be used by FPs and partners for upscaling FTA impact and partnership activities.

#### **Timing considerations: year of measure, year of impact**

There is fundamentally a time lag between getting performance data (on a specific year), reporting, running the evaluation process, and using the data to influence priority setting.

It is proposed that data of performance of year N be assembled during the first three months of year N+1, and that the overall performance be assessed (see Annex 1 for details on responsibilities in that process) at the latest at the end of the first semester of the year N+1. This will inform the POWB N+2. (see Figure 2).

**Figure 2. Performance evaluation as inserted into the planning/reporting calendars**



**Annex 1. Draft LONG LIST of proposed draft indicators of FP performance in FTA**

Category	Index	Area	Indicator	Unit	Calculated by	Source of verification
<b>Delivery (and relevance)</b>	1	Delivery of W1/2 activities listed in the POWB and PPA	Percentage of green W1+2 outputs in traffic light report	%	Director with MSU support	<i>Traffic light report, currently non-existing, will be put in place for W1+2 funded activities</i>
	2	Delivery of bilateral activities	Percentage of green bilateral outputs in traffic light report	%	Director with MSU support	Traffic light report, financial reports
	3	Vitality Blue sky research activities carried by the FP/CoA/Center	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max report by FP/CoA leader
<b>Scientific credibility</b>	4	Adoption of sound protocols in FP research	Projects and W1+2-funded activities adopting a sound research protocol	%	Director based on a proposal by MELIA	1 page max report by FP/CoA leader
	5	Importance and quality of scientific FP production (NP = number of peer reviewed publications)	Productivity : Max (NP/W1+2; NP/total mapped)		Director with MSU support	Publication list provided by FP/CoA leader
	6	Gender and youth integration	% of FP research integrating OR scoring scale (to be defined)	% OR 1-5	Director based on SP/Gender proposal	GEIRS (score of zero if not filled)
<b>Legitimacy</b>	7	Taking into account stakeholders' perspectives and values in the bilaterals and W1+2 projects	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max narrative provided by FP/CoA leaders with associated evidence

<b>Effective-ness</b>	8	Progress towards outcomes, as part of the FP ToC	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	As many outcomes stories as CoAs provided with associated evidence, provided by FP leaders
	9	Ability to question the FP ToC, paradigm shift (changing knowledge, practices, technologies etc.)	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max narrative provided by FP leaders with associated evidence
<b>Contribution to IPGs</b>	10	Production of IPGs	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max narrative report by FP leaders with indication of CoA and Center related data
<b>Strategic value for W1+2 funds</b>	11	The work has demonstrated added value at program-level and funds have been used to strategically to help realize the FTA ToC	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max narrative report by FP leaders with indication of CoA and Center related data in the FP
<b>Program building</b>	12	The proposed work has contributed to the growth of FTA, through strengthening partnerships, generating additional development opportunities and leveraging new resources	Scoring scale (to be defined)	1-5	Director based on a proposal by MELIA	1 page max narrative report by FP leaders. Nr of new contracts signed resulting in new partnership
<b>Managerial performance</b>	13	Compliance with CGIAR and FTA branding guidelines	% of research outputs properly branded (includes bilaterals)	%	Director based on a proposal by Comms	Research outputs (material) sent to comms.
	14	Responding to MSU requests, respecting deadlines	Scale 1-5 (to be defined)	1-5	Director with MSU support	Lists of monitored requests/deadlines to be defined and transparently indicated
	15	Compiling information in projects database	% of completed fields in FTA Projects database / MARLO	%	Director with MSU support	Database statistics

## **Annex 2: Current rule (decision of ISC#2 in February 2016) for performance-based allocations of Window 1-2 funds to Flagship Projects/Centers.**

### **Background**

The issue of allocating Window 1-2 (W1-2) funds to Flagship Projects (FP)/ Centers has been discussed albeit informally in several previous FTA Steering Committee meetings without any real decision taken. As a result the allocation of W1-2 funds across Centers remains a largely mechanical exercise outside of the cross-cutting themes where it is done by the teams in close coordination with the FTA Director.

In the wake of the series of W1-2 budget shortfalls (Oct 2014: 11%; March 2015: 19%; October 2015: 38%) and given the introduction of performance-based indicators by the Consortium Office in the development of the 2016 Financing Plan and the prominence of the concept (yet ill-defined) of performance-based management for the CRPs, it becomes necessary for FTA to adopt some transparent and objective performance assessment to assist the ISC in making proposal for the allocation of W1-2 funds to CIFOR Board. To this end, the ISC requested FTA to design a performance based allocation system that can start being implemented in 2016.

Introducing a mechanism that clearly ‘rewards’ financially the more successful components of a CRP is a good way to increase internal efficiency in research, as long as the definition of ‘success’ is shared across the CRP.

This note describes principles, performance measures and an allocation rule for implementing such a performance based mechanism, starting with the allocation of the 2016 W1/2 funds.

### **Principles**

- FTA allocates W1-2 funds in a manner which reflects the relative performance of FP/Centres.
- The performance of a FP/Centre is assessed through a series of agreed, objective and transparent indicators. These indicators are measured over a three year rolling period, to average out possible out of the ordinary circumstances influencing performance.
- To ensure transparency in the collection of information, the FTA Director coordinates data collection on the indicators of performance, the Management Team<sup>1</sup> is closely associated with the exercise and teams within FTA have access to all information, at all times
- The assessment of the relative performance of FP/Centres is examined by the ISC, based upon a proposal from the FTA Director to the ISC
- The ISC proposes a W1-2 allocation plan to CIFOR Board at its 2<sup>nd</sup> meeting of the year “n-1”. CIFOR Board will make the final decision to approve, request modification, etc. for the year “n”

---

<sup>1</sup> Flagship leaders and main partners

## Indicators of performance and their measures

In its draft financing plan for 2016, the CO assessed the performance of the CRP based on the IEA evaluation reports, the CRP annual reports, the commentaries on the extension proposals and the 2014 Elsevier bibliometric survey. Several CRPs and Centres commented on the opacity and on the lack of relevance to the bases for assessment as well as the incomplete dataset used in the Elsevier surveys.

FTA assesses performance in the following areas:

- Science outputs and their efficiency
- Progress and cost in delivering outputs from the operational plans
- Progress towards outcomes Value for money” (assessed by the capacity to leverage W3 and bilateral funding to achieve outputs and outcomes, the \$/publication, \$/output; it is unlikely we can at this stage have \$/outcome)

Assessing performance in these areas should give a fair view of the overall FP performance and can be consolidated at the CRP level to answer any request for performance indicators by the CO or FO.

Based on the performance indicators proposed in the following table, the average value of each indicator will be determined for the Centers or the FPs. Some indicators may be difficult to assess by Center (e.g. achieved output or scored outcomes), at least initially. In order to keep the process simple, clear and with low transaction costs, the following indicators were selected.

Table of indicators of performance, their relative weight and source of verification				
Area	Indicator	Weight	Assessment period	Source of verification
Science efficiency	W1/2 \$ by non-refereed publication	1	2012-2014	Publication list provided by Centers and FP
	W1/2 \$ by refereed publication	1		
Delivery	\$/green or yellow outputs as in operational plan	2	2013-2014	Traffic light reports and financial reports
Outcomes / Impacts	Progress towards outcomes, as ranked by the ISC independent members, from 1 to 5	2	2012-2014	Narrative provided by FP with associated evidence
Leveraged funds	Amount of W3-bilateral leveraged by W1-2	2	2012-2014	Consolidated financial reports

Based on the comparison of the value of quantitative indicators (publications, outputs, leverage) to the Centres or FP average, a score will be given to the Center or FP for this indicator following the table below:

Indicator value	Score
51% or more below average	-1
Average + or -50%	0
51% or more above average	+1

For outcomes, the maximum is 5 and the minimum 1, so outcome stories below 2.5 are considered unsatisfactory and receive -1 and outcome stories above 4 are considered excellent and receive +1.

The score is then weighted according the weight of the indicators.

Given the information currently available, Centers can be scored against two areas: “Science quality” and “Leveraged funds”. In the future, it will be possible to measure the “Centers’ outputs” via the traffic light reports but this is not easily done now. FP can be scored against the whole set of areas.

#### **Definition of relative levels of performance**

Performance is a relative concept and is defined for resource allocation purposes by FTA as follows. Centers weighted score range from -4 to +4. A score of less than -2 is categorized as underperformance and one greater than +2 as high performance. Average performance is a score between -2 and +2.

The FP weighted score ranges from -8 to +8 and scores below -4 are categorized as underperformance and those greater than +4 are called high performance. Average performance for a FP is a score between -4 and +4.

#### **Performance based allocation rule for W1-2 funds**

In order to ensure that the basic operations of all partners are not disrupted, a flat sum is allocated off the top of the W1-2 funds received by FTA in a given year<sup>2</sup>, whatever the level of performance. For 2016 for instance, this amount is set at \$100K for all partners, and is complemented by another \$100K per FP coordinated by a partner institution. These amount are deemed sufficient to ensure that FP coordination costs are covered so the integrity of FTA as a whole is preserved.

The remaining W1/2 amount is allocated based upon the above definition of high, average and low relative performance (itself based upon the above indicators and their measures) and upon

---

<sup>2</sup> In fact, this sum is net of the costs of FTA management, cross-cutting activities and ISC costs.

the principle that W1-2 allocations should reflect the relative performance of FP/Centres. The performance based allocation rule used by FTA is:

- Low FP/Centre performers in any one year receive 30%<sup>3</sup> less than they would have received otherwise (i.e., following the allocation rules currently existing in FTA)
- The sum thus made available is distributed as follows:
  - If there are high performers: the money goes to the high performers using a pro-rated formula based on their “size” in the program as defined by their bilateral mapping with a minimum of \$100K (to remain a meaningful amount) and up to a maximum of \$1M (this amount might have to be revised based on actual implementation of the rule). The remainder if any is divided among the average performers pro-rated with the size of partner in the program (based on W3/bilateral);
  - If there is no high performer, the money goes to the average performers or to specific cross-cutting actions considered as important (using a pro-rated formula based on size of the partner in the operational plan) by ISC.

### **Revision of this rule**

This rule will need to be revised in the course of the preparation of the FTA II full proposal, in particular with regards to the determination of the “baseline” allocation at the start of FTA II and the expected performances of the Centres during the implementation of the program.

It might also be revised upon request from the Lead Center BOT or the Chairperson of the ISC after recommendation by the FTA Director.

---

<sup>3</sup> The 30% value is a compromise in order to send a significant message but not completely hamper the delivery of the FP/Centre.