GLOBAL ENVIRONMENT FACILITY
EVALUATION OFFICE

MID-TERM REVIEW OF THE GEF SYSTEM FOR TRANSPARENT ALLOCATION OF RESOURCES

Technical Document # 2

Design of STAR

October, 2013

CONTENTS

1. INTRODUCTION

The System for Transparent Allocation of Resources (STAR) is a framework adopted by the GEF for allocation of its GEF-5 replenishment resources to eligible countries to support activities to generate global environmental benefits in the biodiversity, climate change and land degradation focal areas. The implementation of STAR began in July 2010.

The GEF Council requested the GEF Evaluation Office to conduct a mid-term evaluation of STAR (STAR MTE) to provide feedback on its design and implementation. The evaluation aims to assess STAR design, its implementation, the extent it has met its objectives, and the areas for further improvement. It also aims to determine whether the changes adopted in STAR vis-à-vis the previous Resource Allocation Framework (RAF) have been effective. The approach paper of the evaluation provides more detail on the key questions and the methodological approach of the evaluation.

This paper is based on sub-study on STAR design carried out as part of the mid-term evaluation. The sub-study addresses issues such as the scientific and technical validity of the GEF Benefits Index, the GDP based Index, and, the GEF Performance Index; ceilings and floors; the flexibility features; set asides including the sustainable forest management incentive scheme. The key conclusions of the assessment on STAR design are as follows:

a) STAR has adequately addressed most of the weaknesses in RAF design. However, there are areas where its design needs to be fine-tuned for GEF-6 period.

b) STAR allocations continue to be driven by GBI, which is appropriate given the mandate of the GEF.

c) STAR indices are scientifically and technically valid, although there is scope for further improvement.

d) The market exchange rate based GDP indicator was effective in directing additional resources to least developed countries (LDCs). Nonetheless, use of a purchasing power parity (PPP) based indicator would have been more appropriate for capturing socio-economic conditions in recipient countries.

e) The present STAR approach to calculate terminal evaluation review outcome rating based component of PPI is sensitive to number of observations.

f) Removal of the 50 percent rule increased the level of utilization under STAR at mid-point for several countries. The countries that had full flexibility to STAR resources across STAR focal areas greatly benefited from it. However, the countries with marginal flexibility, especially those with aggregate allocations in the US $ 7 m to 20 m range did not benefit as much although their need for flexibility was comparable to the countries that had full flexibility.

g) Incentive schemes may be effective in directing resources from country allocations to specific GEF priorities.

2. BACKGROUND

The policy recommendations of the third replenishment identified the need to establish "a system for allocating scarce GEF resources within and among focal areas with a view towards maximizing the impact of these resources on global environmental improvements and promoting sound environmental policies and practices worldwide."[1] In September 2005, the GEF Council agreed to implement "a resource allocation framework based on an index of country's potential to generate global environmental benefits in the biodiversity and climate change focal areas and an index of performance" for the GEF 4 replenishment period.[2]

The mid-term review of the RAF (RAF MTR), conducted by the GEF Evaluation Office in 2009, noted several concerns related to design and implementation of RAF. It found that: the RAF provided limited incentives for improved performance; the ceiling on the level of resource utilization by the mid-term of GEF-5 resulted in lower levels of resource utilization; unclear guidelines limited the access of the group allocation countries to GEF resources; rules for RAF's implementation were complex and did not encourage flexibility and dynamism; and, although RAF increased country ownership in countries with individual allocations it had negligible or negative effect on ownership in the countries with group allocations.

The mid-term review of RAF recommended: reallocation of unused funds during the last year of the GEF-4; the implementation of the resource allocation framework during remaining period of GEF-4 with full public disclosure, transparency, participation, and clear responsibilities; simplification of implementation rules; and, improvement in the design and indexes to be used for the period covered by the next replenishment.[3] Other than the recommendation on simplification of implementation rules, the Council adopted all of the recommendations. The Council decided not to adopt the recommendation on simplification because of the risk that any change at that late stage in GEF-4 would not have been practical.

The preliminary proposals for the revised resource allocation framework, now rechristened as STAR – the System for Transparent Allocation of Resources – were presented at the Council's meeting in June 2009. In its November 2009 meeting the Council reviewed the revised proposals and decided to extend the STAR to the land degradation focal area and adopted new design features that provided greater flexibility in utilization of allocated resources.[4] In its June 2010 meeting the GEF Council reviewed the document on operational procedures for STAR (GEF/C.38/9/Rev.1) and approved the procedures described in the document.

[1] Summary of Negotiations on the Third Replenishment of the GEF Trust Fund (GEF/C.20/4), Annex C, page 50, para 16.
[2] Joint Summary of Chairs – Special Meetings of the Council, August 31 – September 1, 2005 (GEF/C.26/Joint Summary).
[3] Mid Term Review of the Resource Allocation Framework, GEF EO. July 2009.
[4] Joint Summary of Chairs – GEF Council Meeting, November 10-12, 2009.

3. CHARACTERISTICS OF STAR DESIGN

The objective of GEF's resource allocation framework is to function as "…a system for allocating resources to countries in a transparent and consistent manner based on global environmental priorities and country capacities, policies and practices relevant to successful implementation of GEF projects" (GEF/C.27/Inf.8/Rev.1). Under STAR, the procedure to determine a country's allocation for a focal area involves the following steps:

a) Calculate the **country's score** for a given focal area using a composite formula that combines a focal area specific GEF Benefits Index (GBI), a GEF Performance Index (GPI), and a GDP-based Index.[5]

$$\text{Gross Score} = \mathbf{GBI}^{0.8} \ \mathbf{X} \left(\frac{\mathbf{GDP}}{\mathbf{capita}}\right)^{-0.04} \ \mathbf{X} \ (\mathbf{0.65CEPIA} + \mathbf{0.15CPIA_D} + \mathbf{0.2Portfolio})$$

b) Calculate the **country's share** for each focal area by dividing the country's score for the focal area by the sum of the country scores for all countries eligible to receive STAR allocation for that focal area.
c) Compute the **preliminary allocation** for the country for a given focal area by multiplying the country share with the total amount of GEF resources available for that focal area after deducting the set asides.
d) Determine the **adjusted allocation** for the country after application of ceilings and floors.

Compared to RAF where a benefits index and a performance index had been used for calculation of a country score, under STAR, in addition to these indices, a GDP-based index with a preference for countries with lower per capita income is also part of the composite index. The benefits indices and the performance index under STAR are also different from those used under RAF in terms of the weights and indicators used for composing these indices. While the STAR's approach to calculating a country's share and preliminary allocation is identical to that used by RAF, the floors and ceilings have changed, while there was also a slight shift in the relative share of the climate change and biodiversity focal areas.

The mid-term review on RAF found that utilization of GEF resources among group allocation countries was lower than among countries with individual allocations. It also found that while RAF had increased country ownership in individual allocation countries, it had a negligible or detrimental effect in countries with a group allocation. As a response to these findings, group allocations were eliminated in the STAR's design – under STAR all eligible countries have an individual country allocation.

[5] The document, "System for Transparent Allocation of Resources (STAR)" (PL/RA/01; 2012) provides details on calculation of these indices. The document is available online at: http://www.thegef.org/gef/sites/thegef.org/files/documents/document/PL.RA_.01.System%20for%20Transparent%20Allocation%20of%20Resources.doc%20.pdf

A major criticism of RAF was that it provided limited flexibility in the design of the allocation system. The STAR's design introduced greater flexibility by removing the rule that restricted utilization of a country's focal area allocation to 50 percent at the end of the second year, and by allowing use of allocations across focal areas.

The total commitments made by the donor countries for the GEF-5 replenishment was $ 4.34 billion. This is considerably higher than the $ 3.14 billion replenishment for the GEF-4 period. Availability of higher levels of resources for the GEF-5 period led to an increase in the aggregate allocations for focal areas and to increased average country allocations under STAR.

4. KEY QUESTIONS

The key questions addressed by the sub-study on STAR design are:

1. To what extent does STAR design of STAR facilitate allocation and utilization of scarce GEF resources to enhance global environmental benefits?

2. To what extent is STAR design flexible and addresses weaknesses noted in the RAF mid-term review?

The sub-study assesses the scientific and technical validity of the global benefits indices (GBIs) of STAR. It also sheds light on the rationale for the chosen indicators and trade-offs that have been made in designing the indices. The sub-study assesses the manner and the extent to which the performance index actually influences resource flows and creates incentives for improved performance. It also addresses the extent to which the GDP-based index is technically sound and a good proxy for socio-economic conditions. The sub-study also aims to address the interaction of various indices, and determines the drivers of the allocations. It also assesses effectiveness of the design features such as set asides, floors and ceilings, that determine a country's share.

The Mid Term Review of RAF identified lack of flexibility as one of the key weaknesses of RAF. In response to the critique, two major design features and procedures have been included in STAR. These include dropping the 50 percent rule, which limited the level of a country's utilization for a focal area to 50 percent of its focal area allocation, and provision for a country's allocation for a focal area for activities in other focal areas. The evaluation assesses the extent to which these additions have been effective.

5. METHODOLOGY

A variety of methodological approaches were used to gather information for the sub study. A mix of quantitative and qualitative tools and methods was used. These include desk review of relevant documents, review of the three GEF Benefit Indices by expert panels, and statistical analysis and simulations.

The document review was undertaken to establish STAR objectives, its design, rules and procedures for implementation. The document review also covered Council documents on STAR and GEF project cycle; Assembly documents; RAF Mid-Term Review report; and, various versions of the STAR proposals.

Three expert panels – one for each focal area covered under STAR – were constituted to assess the technical and scientific merits of the focal area GBI. Each panel had two experts, which were identified in consultations with GEF's Scientific and Technical Advisory Panel (STAP) (see annex 1). The reports prepared by these panels are presented in annex 2 of this paper. These reports were also reviewed by other independent experts or evaluators. For biodiversity these peer reviews were undertaken by four independent experts identified in consultations with DIVERSITAS, for land degradation these were undertaken by two experts including an expert identified by International Geosphere-Biosphere Programme (IGBP), and for climate change it was done in-house within the GEF EO. The findings and conclusions on quality of GBI thus reflect the opinion of both the expert panels and independent reviewers. The review of GPI and GDPI was conducted by the core team of the evaluation.

Simulations were carried out to assess effects of STAR indices, floors and caps, and flexibility features. Comparisons were made assuming a non-STAR scenario and by manipulating the formulae.

6. FINDINGS

6.1 GEF Benefits Index (GBI)

The three focal areas covered by STAR have separate indices to calculate the benefits potential of a country for the given focal area. Overall, indices were assessed to be scientifically and technically valid. The indicators for biodiversity and climate change are directly linked with global environmental benefits pursued by the GEF. Although in absence of better alternatives proxy indicators have been used for the land degradation focal area, their validity has been confirmed in research linking the proxy indicators to land degradation issues of global relevance observed in countries. There are several areas where there is scope for improvement. Nonetheless, the suggested improvements are incremental in nature and do not require a complete redrawing.

As was the case under RAF, country allocations under STAR are determined primarily by a given country's potential for generating global environmental benefits. Although the GBI component has an exponential weight of 0.8 compared to 1.0 for performance, due to larger variations in the observed values on the indicators that constitute GBI it ends up playing a much larger role in determining allocations across countries. Given the overall mandate of the GEF, this is focus is appropriate. STAR being driven by the GBI is in line with the trends in other multilateral organizations to align their performance based allocation (PBA) system more closely with their mandate. The IFAD and the Caribbean Development Bank have recently updated their PBA systems to include indicators that are more effective in capturing their allocation priorities and mandate.

Biodiversity GBI

The Biodiversity GBI has two components: terrestrial biodiversity which has a weight of 75 percent; and, marine biodiversity which accounts for 25 percent of the weight. The terrestrial score of a country is determined through several steps. Firstly, all components of distinct terrestrial eco-regions within a country (CECs) are identified. Thereafter, each of the CEC in a country is scored using four characteristics – represented species (55 percent weight), threatened species (20 percent weight), eco-region representation (15 percent weight), and threatened eco-regions (10 percent weight). The composite score for each terrestrial CEC is calculated using a weighted average of the four characteristics scores. The scores for all the CECs are added to give the total GBI score for a country.

The marine score of a country is based solely on represented fish species. Each species has a uniform credit of 1. This credit is distributed across countries proportional to the estimated habitat within the respective country. The marine score for a country is calculated by adding up the credits from all of the fish species located in the territorial waters of the country.

The Biodiversity global environmental benefit index is assessed to be conceptually simple and based on scientific evidence. The index gives a lot of weight to number of represented species. However, GEF investments in the focal area are primarily directed to ecosystem scale interventions indicating a minor disconnect between the GEF priorities and weights in GBI index.

The coverage of GEF-eligible countries in terms of data richness is uneven across recipient countries. This creates a situation where countries that may have rich biodiversity but poor documentation of it receive lower allocation. For example, Angola which is widely regarded to be among the countries with rich biodiversity is assessed to have received a lower allocation due to poor documentation of its biodiversity.

The present split of 75 percent weight to terrestrial biodiversity and 25 percent to marine biodiversity is assessed to be appropriate. While it is true that marine areas account for 70 percent of the global surface, much of the marine biodiversity related national projects are

focused on shore or near shore activities. Further, GEF provides support to areas beyond national jurisdiction through set-asides for regional and global projects.

The scientific and technical validity of the biodiversity GEB index could be improved and strengthened by giving greater attention to ecosystem functions and freshwater species. Although measures of ecosystem services and the quantification of the value of biodiversity and ecosystem services are difficult, it needs to be explored further. Finer-scale measures, than those that have been used in STAR, are also available for at least some dimensions of species distribution (see annex 2.a). Wherever possible incorporation of the finer scale data will help in strengthening the biodiversity GEB index. Inclusion of only fish species data for the marine component of the biodiversity index is another area for improvement. Incorporation of data on other aspects of marine biodiversity will strengthen the index, although it will require considerable effort to ensure equitable and transparent treatment of all GEF-eligible coastal countries.

Climate Change GBI

The STAR GBI for climate change focal area is composed of two components. The first component, which accounts for 95 percent of the GBI weight, is based on countries' emissions of greenhouse gases in tons of CO2 equivalents in the year 2007 multiplied by an adjustment factor, which rewards countries that show a decrease in the amount of emissions of CO2 relative to GDP or "Carbon Intensity." The adjustment factor is expressed as a country's Carbon Intensity in 1990 divided by the country's Carbon Intensity in 2007. The second component, which accounts for 5 percent of the GBI weight, uses forest cover as a proxy for LULUCF related climate change mitigation benefits potential. It incentivizes increase in forest cover between 1990 and 2000.

Since 95 percent of GBI is accounted for by the emissions related factor, despite the adjustment factor, the index leads to high allocations to countries with high GHG emissions. However, it is also true that potential of climate change mitigation is also higher in such countries. Therefore, concentrating resource in these countries for activities that reduce GHG emissions is likely to lead to generation of greater amount of global environmental benefits (i.e. carbon emissions reduction). Moreover, the scale of GEF support to these countries is relatively small and moderated through an adjustment factor that encourages reduction in carbon intensity for a given level of production. Consequently, there is very little likelihood that greater GEF support to countries that have high carbon emissions will create negative incentives that lead to increased carbon emissions.

The indicators used for determination of the GEB potential are linked with the overall objective of the GEF-5 strategies for climate change mitigation. However, linkage with each of the climate change mitigation strategies pursued in GEF-5 is not as clear. For example, while GEF strategies may focus on sectors such as transportation or renewable energy for climate change mitigation, the index does not incorporate direct indicators from these areas.

Strengthening linkages with the climate change mitigation focal area strategies is likely to be difficult as increasing linkages also increases the risk of making the GEB index too complicated. Nonetheless, the STAR GEB index may be further improved by strengthening the adjustment factor to provide greater allocation to countries that have a good record of reducing their GHG emissions in recent years.

Land Degradation GBI

Of the three focal areas the GBI of land degradation focal area is probably the simplest. The three proxy indicators – land area affected by land degradation (20 percent weight), proportion of dry land area in a country (60 percent weight), and vulnerable population (20 percent weight) – that have been used to determine the global environmental benefits potential for land degradation are valid. Due to data availability related concerns, proxy indicators were used. Therefore, the validity may be verified in statistical terms based on results that these indicators provide.

A weakness in the index in its present form is a weight of 60 percent given to the proportion of dry land area in countries. The rationale provided in the STAR paper that consolidates the Council decisions (PL/RA/01) is that "dry-lands are an important indicator because they are predisposed to desertification and are a major factor influencing livelihoods of nearly a third of the world's population." Although the use of this proxy indicator is aligned with UNCCD's core interests and directly reflects each country's opportunity regarding dry-lands, the 60 percent weightage accorded to it is probably too high. Given the high weightage, countries with higher proportion of dry lands tend to obtain superior allocation weighting, compared to countries with a significant land degradation record but lower proportion of dry land. Indeed, it has been argued that investments in semi-arid zones especially bring lowest returns because of the limited options for sustainable land management and because the degradation processes are naturally far greater than in, say, humid areas. Comparing similar sized African countries, one comprising almost entirely dryland adjacent to another which has a high percentage of humid degraded forest, yet has a low percentage of dry land, the former attracts almost double the allocation in spite of the likelihood that the latter country can deliver more GEBs.

More than other factors the choice of the weight accorded to the indicator on proportion of dry land seem to have been driven by precedent. The allocation share derived using the land degradation GBI formulae closely follows the respective shares of countries in different regions during the GEF-3 period, especially for Africa. During GEF-4, when land degradation focal area was not covered under STAR, the share of African countries was nearly two thirds of the total utilization for the focal area. The land degradation GBI for GEF-5 does has the effect of allocating resources more evenly across regions.

6.2 GDP-based Index (GDPI)

During the STAR ad hoc committee meeting in March 2009 in Paris and GEF replenishment meeting in June 2009 in Washington DC, several participants requested inclusion of a socio-economic indicator for resource allocation. Given that there are large variations among the recipient countries in terms of GDP per capita, and the intent that this indicator should not drive the allocations, based on simulations -0.04 was chosen as the exponent for this indicator. For this exponent value, plugging the values of GDP per capita countries for the year 2008, there is a premium for countries that had a GDP per capita of less than US$ 3000 per annum. The premium is considerably higher for countries whose per capita GDP is much below US$ 3000. However, the premium decreases as GDP per capita (current prices) approaches US$ 3000. For countries with GDP per capita higher than US$ 3000 this leads to lower than business as usual allocations. Simulations show that inclusion of this indicator has led to some changes in the allocations. On average allocations to the countries classified as LDCs or HIPIC increased by 5 percent due to inclusion of GDPI. These shifts were primarily on account of 'other countries' and SIDS (Table 1).

Table 1: Effect of GDPI on STAR Country Allocations

Country Category	Share in actual STAR Allocations	Simulated share without GDPI	Net change in share due to GDPI	Percentage change in share due to GDPI
Fragile	9.3%	8.9%	0.4%	4.0%
LDC	18.1%	17.2%	0.9%	5.1%
SIDS	9.9%	9.9%	-0.1%	-0.6%
LLDC	12.5%	12.2%	0.3%	2.3%
HIPC	16.4%	15.6%	0.8%	5.0%
Other countries*	63.5%	64.4%	-0.9%	-1.3%

*Other countries include countries that have not been classified as fragile, LDC, SIDS, LLDC, HIPC countries.

Compared to market exchange rate based GDP per capita, GDP per capita based on purchasing power parity (PPP) is better at capturing socio-economic conditions as they are less volatile than the market exchange rate and are based on a comparison of production of similar goods and services across countries. In general exchange rate based GDP understates the standard of living in developing countries and, based on country specific circumstances, there are wide variations across countries in terms of the extent to which their standard of living is under stated. This limits the effectiveness of the market exchange rate based GDP per capita indicator in capturing socio-economic conditions in the countries. PPP measures are often used as a basis for comparing incidence of poverty across countries.

6.3 GEF Performance Index (GPI)

The performance index used during GEF-4 was revised taking into account the recommendations by the RAF MTR. The aggregate weight for GPI component that is based on the two indicators from the World Bank's Country Policy and Institutional Assessment (CPIA) was decreased from 90 percent to 80 percent. The weight of GEF Portfolio Performance Index

(PPI) increased from 10 percent to 20 percent. The exponent for the index remained the same at 1.0.

Inclusion of CPIA indices in GEF Performance Index is in line with the trend across the multi-lateral institutions to harmonize their PBA systems through use of IDA's CPIA indicators. Multilateral organizations such as the African Development Bank, Asian Development Bank, Caribbean Development Bank, International Fund for Agricultural Development and Inter-American Development Bank use CPIA indicators or indicators harmonized with CPIA indicators. This has been done to reduce the burden upon recipient countries, in-line with the Paris Declaration on Aid Effectiveness, and to reduce costs.

In STAR GPI two sub-components of the CPIA index have been used: The Country Environmental Policy and Institutional Assessment Index (CEPIA) that has a weight of 65 percent in the GPI, and The Broad Framework Indicator (BFI) that has a weight of 15 percent. Given that GEF activities relate more to environmental concerns greater weightage to CEPIA is appropriate. There is no scientific reason for the weightage for CEPIA at 65 percent and not 50 percent or 70 percent – however, given that this has been arrived at after deliberations provides it wider acceptance. Nonetheless, it may be difficult to establish an empirical link between the CEPIA and BFI indicators and the policy and institutional change that these indicators are aimed at rewarding and incentivizing. Simulations show that inclusion of CPIA indicators affects allocations to countries categories such as LDCs, Fragile, and HIPC, where scores on CPIA indicators included in the index tend to be lower.

The Project Performance Index (PPI) of STAR GPI has an aggregate weightage of 20 percent. Out of this 12 percent is accounted for by the index on GEF EO terminal evaluation review (TER) based Outcome ratings and 8 percent by the index on PIR ratings for implementation progress for projects under implementation. In comparison, in the formula for RAF a 10 percent weightage had been provided for the PPI: 5 percent each for the GEF PIR based rating and IEG ICR review ratings for completed projects in recipient countries.

Use of PIR ratings on implementation progress for projects under implementation poses a challenge. The indicator measures implementation progress. Therefore, it is to a large extent a reflection of the performance of an implementing agency than of recipient countries. While agency performance and project implementation progress is linked with and affected by country ownership and capacities, the link is not as direct as might be required for it to incentivize country performance. More importantly, it may create disincentives for candid reporting through PIRs.

The RAF Mid Term Review suggested that inclusion of GEF EO's TER based Outcome rating for completed projects in the PPI instead of IEG ICR review ratings should be considered for STAR. The RAF MTR had suggested that sufficient number of terminal evaluations were available for most of the recipient countries. While GEF EO rating indeed replaced the ICR ratings, it is not clear whether it strengthened the PPI index. Due to major gaps in data coverage the utility of GEF EO ratings in STAR for GEF-5 is assessed to have been limited.

The APR 2008 TER data (prepared in FY 2009) was used to determine the country specific values for the TER rating based component of PPI. In the given dataset there were 205 listed projects. However, after regional and global projects are excluded from the list, 147 projects in 72 countries remained. Furthermore, due to the graduation of countries that became member of the European Union or had no GEF activity in the preceding five years, several countries became ineligible for GEF grants for the GEF-5 period. When this was taken into account, the number of completed national projects with ratings dropped to 134 and the number countries that were eligible for STAR allocation covered through these projects reduced to 65. Of these 65 countries only 12 had at least four completed national projects.

Table 2: Availability of GEF EO TER Outcome Ratings for Completed Projects

Country category based on number of terminal evaluation review with outcome ratings (of countries eligible for GEF grants through STAR in GEF-5)	Based on TER 2008 dataset (for GEF-5)	Based on TER 2012 dataset (for GEF-6)
Countries without any TER with outcome rating	80	50
Countries with only one TER with outcome rating	35	32
Countries with two TERs with outcome rating	12	19
Countries with three TERs with outcome rating	5	11
Countries with Four TERs with outcome ratings	6	10
Countries with Five or more TERs with outcome ratings	6	22
Total number of eligible countries	144	144[6]

To some extent this weakness will be mitigated for the GEF-6 period because a greater number of terminal evaluation review based outcome ratings are now available. For example, the TER 2012 dataset includes 486 completed projects that have received TER outcome ratings. When global and regional projects, and projects in countries that are no longer eligible for GEF grants or have graduated, are excluded, the number reduces to 314. When the updated data would be taken into account, there would still be no observation for 50 countries and for 32 there would be only one observation (table 2). This underscores the point that despite improvements in the dataset for the GEF-6 period, it would still form a weak basis to provide information of performance of completed projects in the recipient countries and reliance on global portfolio average may have to continue. This said the countries that account for a bulk of GEF funding would tend to be increasingly well covered.

Effect of the PPI on country allocations is marginal. Simulations show that if the allocations were provided after dropping the entire PPI component of the STAR, the change in allocations for various country groups based on the size of STAR allocations (i.e. up to US $7 million; US $ 7 million to 20 million; US $ 20 million to 100 million; and more than US $ 100 million) range from – 1.1 percent to 1.3 percent of the allocation for that respective category. Although the change in allocations of individual countries may range from – 6 percent to 12 percent.

[6] The number of eligible countries for GEF-6 might be different than that for GEF-5. The TER 2012 data has been used to give an indication of the TER outcome rating data coverage for likely eligible countries for the GEF-6 period.

Simulations show that because of its lower weight within GPI and lower variance across countries the effect of PPI is lower than that of CPIA based indicators. They also show that inclusion of PPI in the GPI has a counter intuitive effect of increasing the allocations to the country categories with lower PPI ratings. Although CPIA indicator based score and PPI score for countries are positively correlated (0.23[7]), normalized variation among country scores on CPIA indicators is considerably higher than that on PPI score. When PPI is removed from the GPI, the CPIA indicators take the entire value of the GPI and their weight increases from 80 percent (65 percent for CEPIA and 15 percent for BFI indicator) to 100 percent (81.25 percent CEPIA and 18.75 percent for BFI). This attenuates the effect of the CPIA. On the other hand when PPI is included, it has the effect of moderating the differences across country categories.

6.4 Floors and Ceilings

Compared to RAF (GEF-4), the floors and ceilings for country allocations were changed for STAR (table 1). In all, 71 countries had climate change allocations that were equal to the floor (lower bound) of US $ 2.0, 33 for biodiversity, and seven for land degradation. As discussed earlier in this paper, due to a high weight of 60 percent assigned to proportion of dry land in a country, there is a convergence in the allocations of the countries. As a result, the floor for land degradation focal area is applicable only in a few instances.

Table 3: Floors and Ceilings under RAF and STAR

Allocations	RAF		STAR		
	Biodiversity	Climate Change	Biodiversity	Climate Change	Land Degradation
Floors	US $ 1.0 m	US $ 1.0 m	US $ 1.5 m	US $ 2.0 m	US $ 0.5 m
Ceilings	10 % of total	15 % of total	10 % of total	11% of total	10 % of total

Although several stakeholders argued against the higher allocations for large economies that are fairly resourceful, they also felt that allocations under STAR were more equitable than those under RAF. To some extent this may be attributed to the higher level of replenishment for the GEF-5 period. However, there was indeed some increase in the floors even when the increased level of replenishment is taken into account. In contrast, in absolute terms the ceiling remained more or less the same but in percentage terms it declined for the climate change focal area.

While there is some merit to the argument against resources for the larger economies, it's also true that level of overall development in these countries is not yet at a stage where activities supported by the GEF may be considered as part of their baseline. The incremental

[7] The correlation coefficient is 0.39 if the analysis is restricted to countries that have both CPIA and PPI scores based on actual ratings, and the countries for which average values (PIR/TER) or the bottom value (CPIA) have been plugged are ignored.

global environmental benefits potential provides a rationale for continued support. However, expectations in terms of level of contributions from the country may be anchored to their ability to mobilize additional resources for GEF supported activities.

The levels at which floors were set for STAR allowed a recipient to undertake at least one full size project (through pooling of focal area allocations). This approach seems to be appropriate. The rationale for setting up the ceilings should continue to be based on deliberations across the partnership and the comfort levels of the GEF Council.

6.5 Flexibility

Based on the recommendation of RAF MTR greater flexibility was introduced in the STAR design. This included removal of the constraint that only up to 50 percent of the focal area resources might be used up to the mid-point of the replenishment period; scope for usage of country allocations for activities across focal areas based on aggregate allocation size. Both these features have worked well – abolishment of the 50 percent rule more so than the provision for flexibility in use of resources across focal area.

If the rule limiting utilization of a country's focal area allocation to only 50 percent was applicable under STAR, countries that utilized more than 50 percent of their allocated resources for a focal area by the end of second year of GEF-5 would not have been able to do so. Consequently, GEF's global utilization rate for the focal areas covered under STAR at the half-period mark (i.e. June 30th 2012) would have fallen from the 48 percent (actual utilization) to 35 percent (simulated utilization using the 50 percent utilization ceiling constraint). Abolishment of the 50 percent rule allowed 67 countries to use more than 50 percent of their allocation for the biodiversity, 37 countries for climate change, and 62 countries for land degradation focal area.

Of the recipient countries, those with allocation up to 7 million dollars had full flexibility in using their STAR allocation across focal areas covered by STAR; countries with allocations from US $7 million to 20 million had flexibility of using up to US $ 0.2 million; those with allocations from US $ 20 million to 100 million could use up to US $ 1.0 million; and, those with allocations over 100 million could use up to US $ 2 million. The Secretariat was expected to manage the global utilization in such a manner that at the global level at least 90 percent of the allocations for a focal area were used for activities within that focal area. The provision for flexibility was an unqualified success for countries that had full flexibility. It had limited success in countries that had marginal flexibility.

Utility of the flexibility for countries with full flexibility (for focal areas under STAR) is borne out by empirical data. Of 63 countries that had full flexibility to use resources across focal areas, 38 countries (60 percent) had used 21 percent of their aggregate focal area allocations across focal areas by the end of the third year of GEF-5 (Table 4). For countries with marginal flexibility, the utilization across focal areas was at a much lower level.

Table 4: Utilization of country focal area allocation for activities in other focal areas

Category	Total number of Countries (Allocation)	Utilized cross-focal resources (utilization)	Recipient focal areas: Number of Countries (utilized through funds from other focal area)		
			Biodiversity	Climate Change	Land Degradation
Countries with Full Flexibility	63 ($334.42 m)	38 ($ 70.84 m)	19 ($ 30.67 m)	11 (20.17 m)	17 ($ 22.00 m)
Countries with Marginal Flexibility	81 ($ 2045.57 m)	15 ($ 2.26m)	5 ($ 1.26 m)	3 ($ 0.33m)	9 ($ 0.67m)
Flexibility: $ 0.2 m Allocation $ 7-20 m	*53 ($ 589.53 m)*	*10 ($ 1.08 m)*	*3 ($ 0.29 m)*	*3 ($ 0.33 m)*	*6 ($ 0.46 m)*
Flexibility: $ 1.0 m Allocation: $ 20-100 m	*24 ($ 861.29 m)*	*5 ($ 1.18 m)*	*2 ($ 0.97 m)*	*0 ($ 0.0 m)*	*3 ($ 0.21 m)*
Flexibility: $ 2.0 m Allocation: > $ 100 m	*4 ($ 589.99 m)*	*0 ($ 0.0 m)*	*0 ($ 0.0 m)*	*0 ($ 0.0 m)*	*0 ($ 0.0 m)*

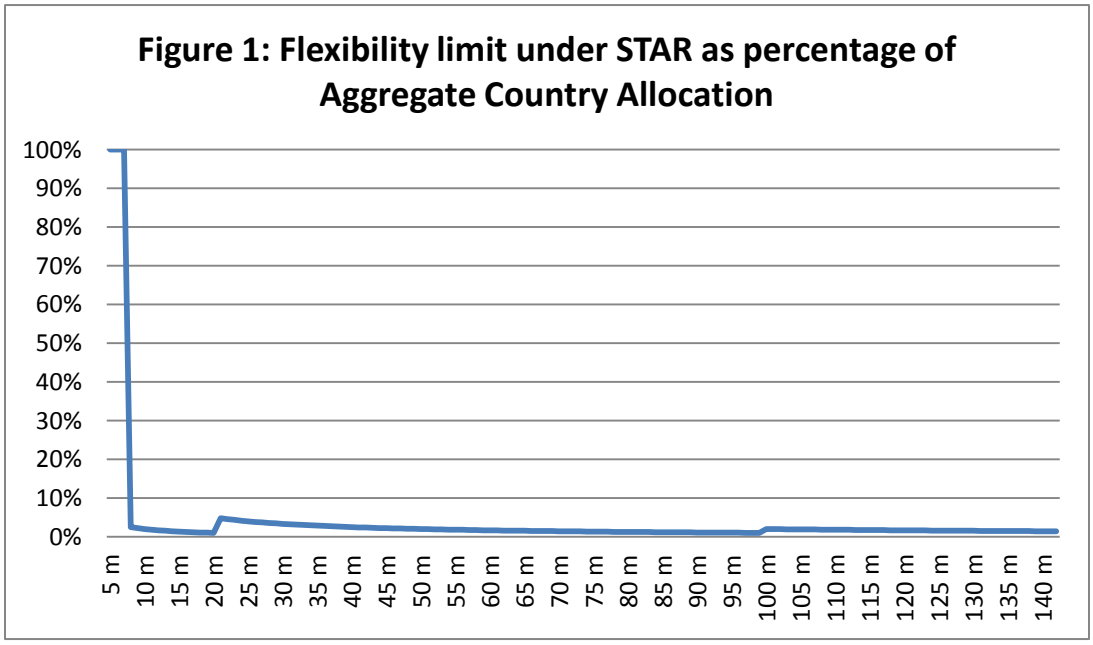
The marginal flexibility related rule for the countries with $ 7 m to $ 20 m in aggregate allocations for GEF-5 has resulted in a soft 'natural experiment', which provides an opportunity to learn more about country response to lack/lower levels of flexibility. Of the 53 countries that had aggregate STAR allocations in the range of $ 7 m to 20 m, 10 countries (19 percent) used the option to use allocations across focal areas and used about 0.2 percent of their STAR resources across focal areas. Thus, countries in this category made much less use of the flexibility related provision than the countries that had allocations in the US $ 4.0 m to US $ 7 m range. Similarly, of the other countries that had marginal flexibility very few made use of the flexibility feature by the end of the third year of GEF-5. While lower levels of utilization of this provision is understandable for countries with higher aggregate allocations, for countries that had aggregate allocations in the range of US $ 7 to 20 million this was primarily because the allowed flexibility of US $ 0.2 million was too low.

Low level of flexibility was one of the factors that were reported to have led the countries with aggregate allocation ranging from US $ 7 to 20 million to use their STAR allocations for multi-focal area projects – for the countries that belong to this category, of the total STAR resources used by them in national projects, multi focal national projects accounted for 57 percent of the share, which is similar to that of the countries that were fully flexible but considerably higher than that for other flexibility categories (table 5). Stake holder interviews revealed that in countries that had aggregate allocation of $ 7 to 20 m the marginal flexibility was not really effective and led them to prefer multi-focal projects so that resources from different focal areas could be pooled to design viable projects.

Table 5: Use of STAR resources for multi-focal projects

Country category by flexibility	Multifocal Area
Fully flexible (4-7M allocation)	140.4 (55.4%)
0.2 million flexibility (7-20M allocation)	204.7 (57.5%)
1 million flexibility (20-100M allocation)	164.6 (31.8%)
2 million flexibility (over 100M allocation)	111.2 (23.6%)
Total	621.0 (38.9%)

A country with low aggregate allocations may need flexibility to use its allocations across focal areas because its allocation for a given focal area may be too low to allow development of a viable project in that focal area. For countries with larger allocation, after it has programmed most of its allocation for a given focal area, they may be left with residual amounts that are not sufficiently large to allow it to program another viable project in that given focal area. Therefore, there is a need to pool resources across focal areas. In both these cases, the level of marginal flexibility should have, therefore, been based on the some notion of the funds required for a full size project (say half the amount of a median full size project). The actual approach adopted for determining flexibility based on aggregate country allocation was quite the opposite. When compared to other country categories based on flexibility, the approach had the effect of penalizing countries that had allocations that were in the US $ 7 m to US $ 20 m range. This led to a situation where countries, especially those in US $ 7 million to 20 million range had residual amounts in the focal areas left that they found difficult to use for other activities as the flexibility was limited to US $ 0.2 million.



6.6 Set Asides

Set asides are an important instrument for the GEF to provide resources for activities that required coordinated transboundary actions at regional and global scale. The RAF MTR indicated that the set side for focal areas covered under RAF was low and that this limited GEF's flexibility in directing resources towards activities that need coordinated transboundary action. Set asides were increased significantly under STAR – i.e. from 5 percent under RAF to 20 percent under STAR. This increase was in line with the trend seen across multi-lateral organizations – the African Development Bank and the Asian Development Bank increased the size of their set asides for regional projects due to increased demand. However, the mandate of these organizations is quite different from that of the GEF. Given the GEF's mandate for global environmental benefits it has an even stronger reason for set asides.

As was the case with RAF, STAR also adopted a uniform approach to set asides – an equal share of resources for each focal area was set aside. Of the total allocation of US $ 2,975 million for the three focal areas under STAR, US $ 595 million (20 percent) was set aside of which Sustainable Forest Management accounted for US $ 250 million (8.4 percent) and other activities for $ 345 million (11.6 percent). However, the share of the SFM set aside, and the set aside for other activities was different for the three focal areas.

Sustainable Forest Management Set Aside

In 2007, the GEF launched a pilot financial incentive scheme promoting country investments in multi-focal area projects with a focus on forests in Amazonia, the Congo and Papua New Guinea/Borneo. During GEF-5 the financial incentive scheme was expanded to cover all the forests of global importance. The $ 250 million set aside for SFM is being operated as an incentive mechanism for recipient countries willing to undertake SFM projects using their STAR allocations for biodiversity, climate change and land degradation focal areas. To access a dollar from the SFM set aside a beneficiary country is required to allocate three dollars from its STAR allocations to a project that addresses SFM related concerns. Individual countries are allowed to invest a maximum of US $ 30 million from their combined allocations for GEF-5, which means that the maximum a country may access through the SFM incentive scheme is US $ 10 million.

At the end of the third year of GEF-5 total utilization of the SFM set-aside was US $ 125.6 million (50.2 percent) through 66 projects with activities spread over 79 countries. Of the US $ 662.7 million in GEF funds invested in SFM projects, funds from the GEF Trust Fund accounted for 94 percent whereas the remainder is accounted for by other trust funds such as LDCF, SCCF and NPIF that are managed by the GEF.

Countries from Africa and Latin America and Caribbean have been able to utilize a relatively higher percentage of SFM set aside funding than their share in STAR allocations and the STAR resources utilized by them so far. A key achievement has been the utilization of the SFM set

aside funding by countries in Europe and Central Asia region, which had not been able to access these incentives during the GEF-4 period. Countries that have total STAR allocation of less than US $ 10 million are accessing relatively more SFM set aside resources. Similarly, LDCs and land locked countries have accessed a relatively higher percentage of SFM resources.

Since GEF-5 period is still under implementation, the utilization figures for the period are not final. However, the GEF resources provided for SFM have already exceeded the amounts provided during the GEF-4 period even when the larger replenishment for the GEF-5 period is taken into account, by the end of the GEF-5 period the funding for SFM projects is likely to be significantly greater than that during GEF-4 period. Despite these achievements, the overall utilization of SFM resources is likely to be considerably lower than the total set-aside envelope of US $ 250 million.

While it's too early to determine the extent to which the SFM incentive scheme has been effective in generating global environmental benefits, the experience so far does show how an incentive scheme may work in GEF. Considerable effort may be required upfront to bring countries and agencies up to speed as they may require a lot of information before they become familiar with the approach. During the first year of GEF-5 the recipient countries and to some extent key staff of the implementing agencies had little knowledge and understanding of how this incentive scheme is likely to operate. This led to poor utilization during the first year and much of the utilization took place during the second year. It is expected that by the end of GEF-5 the total utilization of the SFM set aside might increase to about 60 to 65 percent.

A low ceiling for individual countries at $ 10 million has prevented countries with large STAR allocations from accessing more resources. Application of a ceiling in utilization of funds from the SFM envelope is appropriate as there is a risk that without a ceiling it might lead to a net flow of resources to countries that have higher allocations. However, it also seems that the ceiling has been set on a rather conservative side and there is a case for a slight increase in it. In countries with smaller aggregate allocation, utilization of resources for SFM faced a different barrier. By the time recipient countries and agencies fully understood how resources from SFM may be utilized most countries with smaller allocations had already programmed their STAR allocations. Consequently, they now have little STAR resources left to access funding from the SFM set aside.

Other activities

Compared to 5 percent (US $ 100 m) of the focal area resources being set aside for other activities under RAF, 11.6 percent (US $ 345 m) was set aside for other activities under STAR. Compared to a utilization rate of 71 percent (US $ 71.3 m) up to the end of the third year of

GEF-4 under RAF, the utilization rate was 47 percent (US $ 163.2 m) under STAR[8]. In absolute terms the utilization of STAR set aside has increased. However, in percentage terms the utilization levels are much lower than during GEF-4. Thus, resources available from set asides are no more a constraint in terms of programming of regional and global projects from the set asides.

[8] The set aside utilization under STAR for other activities was US $ 163.2 m (47%) for all three focal areas and US $ 147.4 m (52%) for climate change and biodiversity focal areas – that had been covered under RAF - together.

7. ANNEXES

ANNEX 1

List of Panel Members and Independent Reviewers

Biodiversity GPI

Panel Members

1) Brian J. Huntley
2) Jeffrey A. McNeely

Independent Reviewers

1) Patricia Balvanera
2) Mark J. Costello
3) Melodie Alyce McGeoch
4) Brian William van Wilgen

Climate Change GPI

Panel Members

1) Ralph E H Sims
2) Christine Woerlen

Independent Reviewers

1) Neeraj Kumar Negi
2) Steve Thorne

Land Degradation GPI

Panel Members

1) Michael Stocking
2) Douglas Taylor

Independent Reviewers

1) Anna Tengberg
2) Christine Fürst

ANNEX 2

ANNEX 2.A

Review of Biodiversity GPI
Brian J. Huntley and Jeffrey A. McNeely

1. What is the scientific and technical validity of the global environmental benefits indices for biodiversity? To what extent do the results provided by the specified indices reflect the importance of a country in terms of global environmental benefits potential for the biodiversity focal area?

The RAF Mid-Term Review included a Delphi survey among 36 biodiversity specialists, plus other evaluations, and concluded that there was general positive consensus on the formula and information base used to deliver the GEB index for biodiversity. The present reviewers agree with this general assessment, but times have changed, information has strengthened and GEF priorities have been revised since the RAF MTE, and several modifications might be considered in future improvements to the index.

The structure of the approach is based on national level information on four 'values': the proportional representation of ecoregions; threats to these ecoregions; represented species; and threatened species. The data for these component values were drawn from the extensive databases of the IUCN, WCMC, CI, BLI, WWF, FishBase, FAO, and others (most of which in turn depend on national databases or sources of data).

The approach meets the GEF STAR objective of being 'conceptually simple, scientifically based', but is less responsive to the GEF and CBD mandates relating to sustainable use and genetic resources; nor is it 'comprehensive in its coverage of GEF-eligible countries' (in the sense that all such countries be covered in roughly comparable detail), as discussed in further detail below in response to the questions addressed in this review.

The biodiversity index is primarily based on species-level data that are given very heavy weight (by a factor of 0.55+ 0.20). However, GEF investments in the biodiversity Focal Area are primarily (and properly, in our view) directed to ecosystem scale interventions, especially in protected areas and to a lesser (but growing) degree to mainstreaming biodiversity conservation across production landscapes (see, for example, Miller et al., 2012 for justification for such a shift). Very little GEF funding goes to species-level activities, even though the motivations for GEF funding (PIFs and Project Documents) often emphasize the importance of flagship, threatened or endemic species. Plant species, on which the life of all animals is dependent, are weighted equally with each of the five vertebrate groups (mammals, birds, reptiles, amphibians, fish). This weighting seems arbitrary, based perhaps on access to data rather than on any particular predictive power or theoretical basis. Further, in terms of conserving the evolutionary richness of faunas, assessments at taxonomic levels higher than species (i.e. Genus or Family levels) might be more appropriate than species number (Lotz et al. 2013). The same might apply especially to plants, where some rich floras

are dominated by single speciose genera (some 57 plant genera contain at least 500 species – Frodin, 2004).

The scientific and technical validity of the indices could also be improved by giving greater attention to freshwater species, including invertebrates. Global patterns of freshwater species diversity, threat, and endemism are now receiving greater attention (Collen et al. 2013,), and as this work matures it could make a useful contribution to strengthening the biodiversity indices.

The CBD concept of biodiversity covers the continuum of genes/species/ecosystems/landscapes and the processes by which they interact. Emergent properties (ecosystem services) of biodiversity are not explicitly identified in the original CBD definition of biodiversity, but following the 2005 Millennium Ecosystem Assessment (MA, 2005) have been embraced in the CBD Strategic Plan and Aichi Targets. Ecosystem services form a fundamental component of GEF investments in mainstreaming interventions – and both within protected areas and in their surrounding lands. Ecosystem functions (providing goods and services, at local habitat to large landscape/seascape scale) deserve higher weighting in the further development of the STAR.

The relationship between biodiversity (usually measured as number of species because this is the only component of biodiversity that is readily counted) and ecosystem services (measured by processes such as carbon sequestration, water yield, nutrient cycling, conservation of pollinators, and many others) remains somewhat uncertain. Ninan, 2009 and Kumar, 2010, provide detailed examples of the values of ecosystem services, while Mace et al. 2012 and Ingram et al. 2012 argue that such measurement is not all that straight-forward. The assumption that conserving species begets conserving ecosystem services is complex; for example, a species-rich tropical forest may store less carbon than a monospecific, intensively-management plantation, though the biodiversity value of the latter is very low and it may harbor few pollinators.

So measures of ecosystem services and the quantification of the value of biodiversity and ecosystem services remain difficult. But this critically important topic deserves special consideration in STAR, given the statement in the GEF 6 BD strategy that "*the strategy addresses the most critical underlying driver of biodiversity loss; the failure to account for the full economic value of ecosystems and biodiversity, through systemic biodiversity mainstreaming approaches that have high potential for far-reaching and sustained impact*".

WWF ecoregions provide a rather crude surrogate for ecosystems and ecosystem services, but remains the most comprehensive system available for such global assessments. WWF's ecoregions cover often-vast areas stretching across numerous countries, and with considerable variation within them. The concept of 'threat' to such complex units seems questionable to those working on the ground. Even so, STAR uses the proportion of ecoregions and threatened ecoregions, represented in each country as a value (weighted by a factor of respectively 0.15+0.10). It refers to the 'areas that remain un-cleared for agriculture and urban development' as the basal unit. How such 'un-cleared areas' are

calculated remains obscure, although the source of information is provided (the International Food Policy Research Institute, which obtains most of its data from national sources). For most developing countries, such data do not exist at a meaningful scale or accuracy. For many ecoregions, especially those containing substantial arid systems, advanced remote sensing tools may be required to determine where overgrazed, degraded or transformed landscapes are located, and even then, range condition is very sensitive to recent rainfall or other seasonal factors, degree of threat is difficult to measure, and even 'overgrazed' remains a controversial concept (Mysterud, 2006).

Given these limitations, it may be worthwhile to consider whether the 'ecoregion' is the most appropriate scale for STAR, when other, finer-scale, measures are also available for at least some dimensions of species distribution (Jenkins et al., 2013). But until a better option becomes available, it may be best to apply greater weighting to ecoregion representation and threat (as the best available surrogate for ecosystem services, landscape species, system resilience, connectivity, etc). Refining the ecoregion concept will require more sensitive measurement of ecoregion distribution, condition and characteristics at species and habitat diversity levels. Such refinement would enable ecoregions to continue to serve as useful global-level units for comparative assessments, while recognizing that they are still crudely delimited in many countries.

It might be argued that country-based biogeographic units would be more useful, but for the purposes of STAR, comparison would be difficult unless each country used the same system, an unlikely outcome judging from the diversity expressed in National Biodiversity Strategies and Action Plans.

A detailed review of the BD index is frustrated by the difficulty of access to the input information and implementation methodology, which is described only in general terms in available documentation. No indication is given in available documents of the 'comprehensiveness' of the datasets used. The RAF MTE mentions "88 countries without sufficient data to compute meaningful allocations", but the 2008 GEF Benefits Index for Biodiversity still includes all GEF-eligible countries. It would be helpful, and transparent, to make the input data for each country available on the GEF website, following confirmation or at least referral of the datasets to each recipient country for verification.

Current data provided for this review are sometimes difficult to interpret. For example, data for Namibia, an arid, species-poor country with limited ecoregional diversity, is scored 71 points for plant representation, 44 for ecoregional representation, and 64 for threatened species. Its neighbor, Angola, with perhaps the richest ecoregional diversity in Africa (including parts of 15 WWF ecoregions), ranging from Namib desert to Guineo-Congolian and Afromontane forests, has scorings of 75, 55, and 72, only marginally higher than Namibia. While the biodiversity of Namibia is very well documented, that of Angola is poorly known. Thus information biases can markedly influence the result. As just one example, in GEF 5, Namibia was allocated $6,5 million; Angola $7,3 million, despite the latter having vastly greater biodiversity assets. (Note that the original data used in the RAF/STAR was not accessible to the reviewers, so mis-interpretation of the application of the formula is

possible; but we still believe that the STAR criteria should lead to Angola receiving a far higher GEF allocation than Namibia. Further research might reveal better examples of this inconsistency).

Despite the shortcomings referred to above, in general the ultimate funding levels allocated to individual recipients did not seem widely out of proportion to their overall, and generally recognized, biodiversity values *and potential effectiveness of conservation projects*. (The performance measures used in the linked GPI [GEF Performance Index] may well account for the discrepancy between Namibia and Angola noted above. The original RAF allocated 30 countries about 76% of the global biodiversity benefits, while the remaining 118 countries account for 24% of the global biodiversity benefits).

2. **Are weights accorded to terrestrial and marine biodiversity balanced?**

With 70% of the globe covered by marine ecosystems, one cannot argue that a 25%/75% split in GEF allocations between marine and terrestrial systems is 'balanced'. However, given that the STAR is targeted at effectiveness, its impact will be limited to on-shore and near-shore activities, which probably comprise much less than 25% of the global surface. Further, GEF funding streams for International Waters, which include the Large Marine Ecosystems, falls outside STAR, and is additional funding. Investments in Areas Beyond National Jurisdiction are included in set-aside funding.

The balance between marine and terrestrial biodiversity is further complicated by the significant gap in knowledge about marine biodiversity, even when considering only the coastal zone (Costello et al. 2010). STAR's GBIBD is therefore unlikely to be treating marine biodiversity (except perhaps fish species) as comprehensively as it treats terrestrial biodiversity. Aichi Target 11 poses more challenges, calling for '*10% of coastal and marine areas, especially areas of particular importance for biodiversity and ecosystem services, are conserved though effectively and equitably managed, ecologically representative, and well connected systems of protected areas and other effective area-based conservation measures, and integrated into the wider landscapes and seascapes*.' Achieving Target 11 will call for significant investments in even building an understanding of the major variables included in the targets, much less achieving them. Designing cost-effective projects with achievable targets is no simple matter in the light of considerable ignorance about such fundamental concepts as marine ecosystem services, effective management of marine systems, connectivity of sites containing a high percentage of highly mobile species, area-based conservation measures, and how to integrate marine biodiversity investments into wider landscapes and seascapes (noting that many of the most serious threats to coastal marine systems are land based).

Therefore, many of the approaches to marine biodiversity conservation need to be tested, drawing on principles of adaptive management where projects are treated as experiments to provide important insights into achieving the intent of the Aichi Target 11. Given these constraints, a 25% allocation is deemed reasonable, not least because over 60% of the global

human population lives within 100 km of the sea, and the intensity of use of coastal zones and the near-shore is exerting major and negative impacts on the biodiversity of these systems. However, the use of only fish data for the marine component of the BD index needs addressing. Available information on the area of coral reefs, mangroves, sea-grass beds; the breeding grounds of sea turtles and marine mammals, and the new information collected by the recently-completed Census of Marine Life (2000-2010), could strengthen the index, though this will require considerable effort to ensure equitable and transparent treatment of all GEF-eligible coastal countries.

3. **To what extent are the weights accorded to different global environmental indicators and their constituents appropriate?**

The weights within terrestrial systems are – Represented species = 0.55; Threatened species = 0.20; Represented Ecoregions = 0.15; Threatened Ecoregions = 0.10.

These weightings are said to respond to the greater accuracy of species-based information, which is constantly improving; the July 2013 update of the IUCN Red List added 4,807 additional species, bringing the total assessed to 70,294, of which 20,934 are considered threatened with extinction. This supports the argument on data availability, but given the targeting by GEF towards ecosystem-level interventions, one might expect a stronger weighting for Represented Ecoregions and for Threatened Ecoregions, especially those that hold high levels of endemic species under threat (hotspots), while giving due attention to the limitations discussed in our answer to question 1, above. The identification of a wider range of values should be more readily achieved with the rapid increase of synthesized data that are being gathered by IUCN, WCMC, BirdLife, FAO, CGIAR, etc., and drawing from National Reports prior to the CBD COP 11. Technological improvements will enhance the inter-operability of these datasets, and the increasing sophistication of remote sensing of habitats may provide new and useful sources of data to ensure that the weights are appropriate apportioned.

The general trends in STAR allocations are consistent with the pattern of expenditure by donors prior to the advent of STAR. In a comprehensive review of donor investment in biodiversity conservation projects over the 1980 – 2008 period, Miller et al. (2013), the indicators chosen included threatened mammal, bird and amphibians; total species richness per country, endemic species, governance, population size corrected for area, plus GDP-PPP. Given that the database for biodiversity was from IUCN, and that 28% of the $18.55 billion invested over the review period came from GEF, the similarity of results in terms of country targets might be expected. Miller et al. (2013) indicated the need for finer grained data sets.

4. **To what extent are the chosen indicators for STAR aligned with the GEF-5 priorities for the biodiversity focal area?**

The GEF 5 strategic objectives for the biodiversity focal area were -

1. Improve Sustainability of Protected Area Systems (with project support to improve sustainable financing to PAs, expand marine and terrestrial ecosystem representation, expand threatened species representation, and improve management effectiveness of PAs).
2. Mainstream biodiversity conservation and sustainable use into production landscapes and seascapes and sectors (with project support to strengthen policy and regulatory frameworks, implement invasive species management frameworks, and produce biodiversity-friendly goods and services).
3. Build capacity for the implementation of the Cartagena Protocol on Biosafety.
4. Build capacity on Access to Genetic Resources and Benefit Sharing (now the Nagoya Protocol).
5. Integrate CBD obligations into national planning processes through enabling activities.

Objectives 3-5 were supported by Focal Area set-aside funds, so fell outside of the STAR and are not considered here.

With regard to objective 1, the chosen indicators respond well to the two project support components that cover PA expansion and threatened species inclusion. With regard to objective 2 the indicators probably had little significance or relevance to the issue of mainstreaming biodiversity conservation and sustainable use into production systems. Given the high priority that the CBD and GEF have given to mainstreaming in the GEF investments made since 2002 (US$1.6 billion in 327 projects, with $5 billion in co-financing between 2002 and 2012), further consideration should be given to the choice of indicators in STAR to more closely align with the programmes of GEF 6 (see below).

5. **What new developments have taken place in terms of scientific understanding of environmental problems that need to be addressed in the GEB indices for upcoming replenishment period (GEF-6)?**

The GEF 6 Biodiversity Focal Area Strategy has not yet been formally approved, but is likely to have sharper biogeographic and operational targets than GEF1 through 5. Many of the recent publications cited in this report will contribute to a new focus in the GEF biodiversity strategy that responds to expectations of donors for more cost-effective programmes; greater emphasis on sustainable financing mechanisms; urgent responses to wildlife crises such as those of African elephants, rhinos, and great apes, and Asian elephants, tigers, and great apes; greater attention to organized illegal harvesting of threatened species (for example, using CITES data as part of a revised STAR index); climate change impacts on coral reef systems; impacts of invasive alien species on island ecosystems and CBD guidance on mainstreaming approaches; influences of climate change and other factors on aquatic ecosystems, perhaps giving such systems greater attention in a revised STAR index; and the relationships between areas dedicated to biodiversity conservation and those dedicated to production of biomass to meet human needs for food, biofuels, and other biological products.

All of these issues underline the need for careful targeting of investments not only to the most diverse ecosystems with high levels of threat and endemism (the result from using the

existing STAR indices) but also to systems with special characteristics, such as supporting significant numbers of wild relatives of crops; species of high economic potential; governance systems that achieve biodiversity conservation across production sectors; ecosystems that provide especially significant ecosystem services, especially related to water; and systems that provide useful insights into new products, a process known as biomimicry or biomimetics. Investments also need to be allocated to management approaches or processes that do not necessarily lend themselves to indices for allocations (though some indices can be used to measure progress after a project is being implemented).

As indicated above, the current indicators used for STAR indices are both too coarse and not sufficiently sensitive to specific GEF programme objectives. While useful in providing broad-based guidance to resource allocation at a global scale, and therefore important for the goals of STAR, they can be improved to make them more widely applicable to GEF 6.

This conclusion can be strengthened by considering the GEF 6 Biodiversity Focal Area strategic objectives and programmes, and the role of STAR in directing investments. GEF 6 BDFA includes:

Objective 1: To improve sustainability of PA systems (with a focus on 'globally significant ecosystems').

- Programme 1. Sustainable financing;
- Programme 2. Expand the coverage of the global PA estate in terrestrial and marine ecosystems and of species;

Objective 2: Reduce threats to globally significant biodiversity.

- Programme 3. Improved management effectiveness in and around PAs.
- Programme 4. Reduce poaching to African elephant and rhinos.
- Programme 5. Avoid imminent extinction of island ecosystems, with a focus on Invasive Alien Species management on selected islands.
- Programme 6. Implement the Cartagena Protocol

Objective 3. Sustainable use of biodiversity.

- Programme 7. Ridge-to-Reef, focused on Coral Reef Ecosystems.
- Programme 8. Secure agriculture's future, through sustainable use of plant and animal genetic resources.
- Programme 9. Implement the Nagoya Protocol.

Objective 4. Mainstreaming of biodiversity in production landscapes and seascapes.

- Programme 10. Valuation of biodiversity and ecosystem services into development and finance planning.

The existing STAR indicators help target areas considered under objective 1 which explicitly targets 'globally significant ecosystems'. But as indicated above, the weighting for ecoregional representation and for the level of support given to PAs in ecoregions with high

levels of endemic and threatened species need to incorporate new approaches that will make the STAR index more useful, drawing on recent advances in both science and practice.

Programmes 4 and 5 also need much finer-scale indicators than those in the current version of STAR, which is not suitable for targeting resource allocation within these activities. Programme 6 falls outside STAR Programme funding. For Programme 7, the current indicators are too coarse to identify priorities, but the inclusion of data on the area of coral reefs; area of mangroves; and area of sea-grass habitats might help address this weakness. Incorporating more information on the ecosystem services provided by these systems would also help. Programme 8 shows a similar limitation, but could be better informed through sources of information on centres of diversity of wild relatives of crops (for example, Hajjar and Hodgkin, 2007 and continuing work by FAO). Programme 9 is resourced from set-aside funds, but in any case the current indicators do not have any predictive capacity on genetic resources of potential economic value. Programme 10 has no direct or obvious link to the current indicators.

It is therefore of concern that the current course-grained indicators used in the STAR do not guide the specific needs of GEF 6 investment strategies. The actual relationship between total investment and each programme needs to be considered. If GEFSEC 'ring-fences' funding allocations for specific programmes, then this acts as an additional filter for targeting funds, first by programme and then to countries that meet the STAR criteria and priority ranking. It is not obvious, or transparent, how this a priori allocation by GEFSEC would be determined.

6. **To what extent are the GEB indices – within the context of performance and socio-economic indices included in the composite country score indices – effective in directing the GEF resources to countries where there is greater realizable potential to generate global environmental benefits?**

Albert Einstein is attributed with the comment that " Not everything that counts can be counted, and not everything that can be counted counts". STAR helps to target countries according to their perceived biodiversity conservation need. It is not necessarily effective in ensuring effectiveness of investments.

While objective scientific criteria and indices can guide targeting of resources to the countries with the greatest biodiversity resources, existing measures or predictors of performance are less convincing. The socio-economic indices used might be the best available. Project performance measures, based on the Mid-Term and Terminal Reviews of GEF projects, would be a good source of input, but our experience has shown that well executed projects do not necessarily serve to deliver the highest return of biodiversity outcomes or impact. A well-executed project in a country of limited biodiversity importance does not necessarily result in greater impact for GEBs than a struggling project in a dysfunctional state rich in threatened species and habitats. Changing policies may be a necessary precursor to improved performance in the field, but the need for policy changes are not necessarily revealed by the STAR indices. Poor performance in past GEF projects can

provide points for discussions between agencies and national Focal Points on the kinds of investments that will be most productive in achieving the GEF objectives.

By design, GEF has not proven to be risk-averse, and for as long as many biodiversity-rich countries continue to suffer from poor governance systems, risks will be inherent in GEF investments. Indeed, GEF investment in the high transaction costs (including 'hands-on' management and capacity building interventions) may well continue to be essential in many countries (though by no means all; projects in "mature" GEF-eligible countries such as Brazil, China, Costa Rica, Malaysia, Mexico, Russia, and South Africa should have lower transaction costs). The current 10% 'agency fee' for GEF executing agencies should be made more flexible, possibly based on the STAR socio-economic indices. (It is noted that global indices such as those on the ease of doing business, transparency, corruption, etc. are not included in the GPI. These factors might be contained within the other indices, but key predictors of project performance in many developing countries are governance related).

With regard to cost-effectiveness, the conclusion reached by Miller et al. (2013) is pertinent -

"Because the impacts of biodiversity aid are likely shaped by many country- specific factors, inferring the pathways through which aid affects conservation outcomes will require finer-grained data and more sophisticated analyses. Systematic evaluation of biodiversity aid effectiveness remains a research frontier, one that assumes particular importance as the international community seeks to take the steps necessary to reach shared biodiversity conservation targets".

7. **To what extent is reliable data available on chosen indicators? Have different types of datasets and data sources emerged that provide more reliable data than used for STAR?**

The system of updating the GEF STAR database is not clear from documentation provided for this review. It might be assumed that the key partners in the development of the RAF are in close communication with the GEF, and will bring the wealth of their experience and resources to bear on the revision of the database. More up to date general assessments of the state of biodiversity at global scales, based on national level detail, are available (Butchard et al., 2010, 2012; Hoffmann et al. 2012; McCarthy et al., 2012; McGeoch et al., 2012; and the various CGIAR databases, among others mentioned earlier in this report), and it would be appropriate for GEF to convene a workgroup of such specialist institutions to revise the STAR. However, this should be preceded by the distribution for comment of the country-level raw data used in the RAF to all Focal Points and national and international experts and institutions for comment.

8. **What are the areas where STAR GEB indices for the biodiversity focal area may be improved upon and how?**

This report has indicated numerous ways that the STAR GEB indices could be improved. The first step in improving the database and the acceptance of the STAR system would be to send the raw data sets to each GEF/CBD Focal Point and national institutions to consult on the values given (though in at least some cases, these raw data sets will have originated by the national institutions). This may have been done during RAF development, but the iterative process of consultation with a network of experts, both national and international, can only improve quality and ownership of the system. The STAR GEB indices should be dynamic, and regularly improved as additional data become available and new technologies improve the objectivity and transparency of the data.

Accurate datasets will also strengthen the GEF project cycle from first funding applications (i.e. at PIF level) where too often weak or inaccurate cases are made on the biodiversity assets of applicant countries. Improvement of the ecoregion concepts and delineation, coupled with the use of other tools where these are relevant (see the discussion under Question 5, above), will lead to their better definition, better characterization, and better use in GBI indices as the National Focal Points become more familiar in their use.

Conclusions

The present resource allocation framework for biodiversity has been an effective way of channeling funds to countries on a basis of their individual biodiversity attributes, governance, socio-economic status and performance measures. Improvements to the structure of the model, through more weight being given to ecosystem versus species values; the use of a more comprehensive, nationally verified database; and further learning from country performance in GEF projects, will strengthen the system's science base and ownership.

References

Butchart, Stuart H. M. et al. 2010.Global Biodiversity: Indicators of Recent Declines. *Science* 328, 1164 DOI: 10.1126/science.1187512

Butchart SHM, Scharlemann JPW, Evans MI, Quader S, Arico S, et al. 2012. Protecting Important Sites for Biodiversity Contributes to Meeting Global Conservation Targets. *PLoS ONE* 7(3): e32529. doi:10.1371/journal.pone.0032529

Collin, Ben, et al. 2013. Global patterns of freshwater species diversity, threat, and endemism. *Global Ecology and Biogeography.* DOI: 10.1111/geb.12096

Costello, Mark J. et al. 2010. A census of marine biodiversity knowledge, resources and future challenges. *PLoS ONE* 5(8): e12110. Doi: 10.1371/journal.pone.0012110.

Frodin, David. 2004. History and concepts of big plant genera. *Taxon* 53(3): 753-776.

GEF 2011. GEF 5 Focal Area Strategies.Global Environment Facility, Washington DC. 110 pp.

Hajjar, Reem, and Tony Hodgkin. 2007. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. *Euphytica* 156:1-13

Hoffmann, Michael et al. 2010.The Impact of Conservation on the Status of the World's Vertebrates *Science* 330, 1503 DOI: 10.1126/science.1194442

Ingram, J.C., K.H. Redford and J.E.M. Watson. 2012. Applying ecosystem services approaches for biodiversity conservation: benefits and challenges. S.A.P.I.E.N.S 5.1 URL : http://sapiens.revues.org/1459

Jenkins, Clinton N. et al. 2013. Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Science*. Dooi/10.1073/pnas.1302251110

Kumar, Pushpam (ed.). 2010. *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations*. Earthscan, London.

Lotz, C.N., Caddic, J.A., Forner, M., & Cherry, M. 2013. Beyond just species: Is Africa the most taxonomically diverse bird continent? *South African Journal of Science* 109: 98-101.

MA (Millennium Ecosystem Assessment). 2005. *The Millennium Ecosystem Assessment*. Island Press, Washington D.C.

Mace, G.M., K. Norris and A.H. Fitter. 2012. Biodiversity and ecosystem services: a multilayered relationship. *Trends in Ecology and Evolution* 27: 19-26.

McCarthy, Donal P. et al. *2012*. Financial Costs of Meeting Global Biodiversity Conservation Targets: Current Spending and Unmet Needs. *Science* 338, 946
DOI: 10.1126/science.1229803

McGeoch, Melodie A. , Stuart H. M. Butchart, Dian Spear, Elrike Marais, Elizabeth J. Kleynhans, Andy Symes, Janice Chanson and Michael Hoffmann. Global indicators of biological invasion: species numbers, biodiversity impact and policy responses. *Diversity and Distributions* 16, 95–108

Miller, James R., et al. 2012. Nature reserves as catalysts for landscape change. *Frontiers in Ecology and Environment* 10(3): 144-152. DOI: 10.1890/100227

Miller, Daniel C., Arun Agarwal and Timmons Roberts. 2013. Biodiversity, governance and the allocation of international aid for conservation. *Conservation Letters* 6, 12 -20.

Mysterud, A. 2006. The concept of overgrazing and its role in management of large herbivores. *Wildlife Biology* 12: 129-141

Ninan, K.N. (ed.). 2009. *Conserving and Valuing Ecosystem Services and Biodiversity: Economic, Institutional and Social Challenges*. Earthscan, London.

GEF Documents consulted –

2005 – Oct 17 – GEF 27 /Info 8/Rev 1 GEF RAF – Annex 1 – GBI-BD 33 pp

2008 - 6 October MID-TERM REVIEW OF THE GEF RESOURCE ALLOCATION FRAMEWORK Design of the RAF
Technical Paper #2

2008 3 November MTE - Delphi Study of the RAF Benefits and Performance Indices Technical Paper #5 Prepared by World Perspectives Inc. 205 pp

2009 – May – Midterm Review of RAF – 150 pp

2009 – Oct 9 – GEF/C36/6 – STAR Options and Scenarios. 34 pp

2010 – July 1 – GEF/C 38/9/Rev 1 – Operational Procedures for STAR 6 pp

2010 – October 19 - GEF/C39/Inf. 10 – Set asides outside STAR

2013 – Feb 12 – Mid term Evaluation of the System for Transparent Allocation of Resources. Approach paper

**2013 – March 14 – POLICY PL/RA/01 - STAR *This paper consolidates Council decisions on the System for Transparent Allocation of Resources and constitutes the Policy that will govern its application. 38 pp*

ANNEX 2.B

Review of Climate Change GPI

Ralph E H Sims and Christine Woerlen

1. What is the scientific and technical validity of the global environmental benefits indices? To what extent are the results provided by the specified indices likely to reflect the importance of a country in terms of its GEB potential for climate change mitigation and adaptation? To what extent do the present indices reward the polluters and create perverse incentives?

1.a Calculation of country scores

A composite formula was used in GEF-5 that combined a Global Benefits Index (GBI) specific to the focal area, a GEF Performance Index (GPI), and a gross domestic product-based Index (GDPI). The GDPI was added to the GBI and GPI used in the RAF in order to provide additional support to LDCs by giving a preference for countries with lower per capita incomes and that possibly have greater barriers to generating GEBs.

$$\text{Country score} = GBI^{0.8} * GPI^{1.0} * GDPI^{-0.04}$$

The scientific and technical validity of the specific indicators used in these indices, are now considered, followed by a discussion of the opportunities to update them.

GBI

$$GBI = \text{Baseline}_{2007}\ \text{GHG emissions}\ (tCO_{2eq}) * \frac{\text{Carbon intensity factor}\ (tCO_{2eq} / \$\ GDP_{1990})}{\text{Carbon intensity factor}\ (tCO_{2eq} / \$\ GDP_{2007})}$$

Since tracking GHG emissions from land use change was too uncertain, as a proxy, a second GBI indicator was used in the GEF-5 GBI (weighted at 5%). This fairly basic quantitative adjustment factor was related to a country's forest cover in 2005 and introduced to reward countries that have a decreasing rate of loss of forest area over time[9].

$$GBI = 0.95 * \left[tCO_{2eq\,2007} * \frac{tCO_{2eq} / \$\ GDP_{1990}}{tCO_{2eq} / \$\ GDP_{2007}} \right] + 0.05 * \left[\text{Forest cover}_{2005}\ (ha) * \frac{\text{change in cover}_{1990\text{-}2000}}{\text{change in cover}_{2000\text{-}2005}} \right]^{10}$$

The CO_2 intensity factor accounts for 95% of the weight for all countries when accumulated together and the forest cover related component account for 5 %. The square brackets indicate that the values are normalized to ensure the assigned values dictate the de-facto weights and not the absolute values.

GPI

This performance index is based on the annual World Bank International Development Association's (IDA) Country Policy and Institutional Assessment (CPIA) and the GEF's Portfolio Performance Index (PPI). For the GEF-5, CPIA data from 2008 were used to rate countries. As a proxy for a country's commitment to environmental policies and institutional frameworks, governance and financial management, and its actual performance from GEF projects, the

[9] Where forest cover had increased (as shown by a positive change in cover in both periods) the adjustment factor was taken to be 1.

[10] Not clear exactly when these periods cover: if for a decade then it should be from 1 January 1990 till 31 December 1999. So periods as stated should then become "1990-1999" and "2000-2004".

GPI is a counter-balance to the GBI. When the index was revised in GEF-5, its appropriateness, manner and extent was reassessed. In the GEF-5 it was used to determine the performance potential of a country and hence contribute to its climate change resource allocation.

The PPI used in GEF-5 was calculated from progress ratings taken from GEF project implementation reports (PIRs), which gave 40% of the PPI. The other 60% weighting came from the terminal evaluation reports (TERs) and hence was mainly based on outcomes since they come at the completion of a project. It therefore provided a better indication of whether the objectives were achieved or not.

$$PPI = (0.4 * PIR) + (0.6 * TER)$$

The Country Environmental Policy and Institutional Assessment index (CEPIA, based on indicators within the Policies and Institutions within Environment Sustainability) and the Broad Framework Indicator (BFI, as constructed from the Public Sector Management and Institutions cluster indicators) were also used to produce the GPI.

$$GPI = (0.65 * CEPIA) + (0.15 BFI) + (0.20 * PPI)$$

GDPI

The STAR-documentation refers back to the OPS4 study to point out that LDCs have historically had smaller GEF investments, and that there are multiple barriers to working in these countries. In the RAF, LDCs – due to their typically low energy-related emissions – received relatively small allocations and were mostly in "groups". To address the concerns raised, an evaluation was made of how technically sound the indices are that relate to key socio-economic factors, and whether they could be a good proxy for socio-economic conditions. As a result, a premium factor was introduced to account for country capacity as derived from the per capita nominal value of GDP in addition to the "floors" imposed to provide a minimum allocation to a country (see step **4)** below).

The indicator used in the GDPI is the nominal value of GDP per capita annual income adjusted to a chosen negative exponential. Thus the lower the GDP per capita of a country and the greater the exponential used, then the higher the GDPI score and resulting in a higher allocation.

$$GDPI = (GDP \text{ per capita})^{-0.04}$$

Calculation of individual country shares

As for the RAF, the country score for a focal area in STAR in the GEF-5 was divided by the sum of the country scores for all countries eligible to receive STAR allocations for that focal area. This approach aimed to give added predictability to likely allocation of the resources available in order to facilitate the advanced planning of GEF deliverables. This targetted countries with relatively small allocations that could then predict more easily their minimal allocation.

$$\text{Country share (\%)} = \frac{\text{Country score}}{\text{Sum of country scores for all eligible countries (100\%)}}$$

Computing the preliminary allocations

The country share was multiplied by the total amount of GEF resources available for that focal area (after deduction of the 20% set-asides[11] from the total available finance for that

[11] The aim of the set-asides is to provide additional GEF resources outside of the STAR (up to USD 0.5M), to each eligible country that requires greater assistance for enabling activities that are obligations under the UNFCCC. Set-asides are also a means for countries without access to individual

focal area). The mid-term review of the RAF identified several constraints resulting from group allocations which, consequently, were eliminated from the STAR design. There is no apparent valid reason to re-introduce group allocations into GEF-6.

Preliminary allocation ($) = Country share (%) * (GEF resources available – set-asides) ($)

Determine the adjusted allocations

The floors (minimums) and ceilings (caps) for climate change financing are applied at this stage in order to generate GEBs most cost effectively. Under the RAF, a floor was a minimum allocation of USD 1.0M and a ceiling, a maximum of 15% of the total allocation. These were revised in the STAR to USD 2.0M and 11% of the total. Any country with an allocation calculated to be lower than the minimum floor receives the minimum allocation amount that in GEF-5 was set at USD 2.0M for climate change (compared with USD 1.5M for biodiversity and USD 0.5M for land degradation); conversely, should any preliminary application be higher than 11% of the total, the country allocation is capped at 11%. These levels were selected, together with the flexibility option (see below), so that every eligible country can receive an allocation sufficient to contribute to maximising the GEBs. The total minimum USD 2.0M level as set for the climate change focal area was around 0.19% of the total GEF-5 resource allocation. If based on the total replenishment received, the GEF-6 strategy eventually results in a higher total allocation for the climate change focal area, then consideration should be given to maintaining this ~0.2% share as the minimum allocation for small countries.

Adjusted allocation ($) = Preliminary allocation adjusted for ceiling and floor where appropriate.

Since there is no scientific basis for the setting of these maximum and minimum allocations, appropriate cap and floor levels can be chosen based on operational issues and political decisions resulting from country negotiations.

2. What new developments have taken place in terms of scientific understanding of environmental problems that need to be addressed in the GEB indices for the upcoming replenishment period (GEF-6)?

The November 2012 STAP information document "Climate Change – a scientific assessment for the GEF" [12] was produced knowing that the IPCC 5th Assessment Report would not be published till late 2013, early 2014.[13] It provides a summary of the present knowledge of climate change and also outlines the many mitigation opportunities available through energy efficiency, renewable energy, forestation, and other low-carbon technology options. In essence, the science, now supported by many observations, has grown stronger and confirms that climate change is already occurring and is almost certainly anthropogenic. This has not changed the fundamental understanding and so does not drastically affect the basic objectives of the GEF in terms of GEBs or the STAR process.

STAR resources to receive some funding. In GEF-5, $272 M of the $1360M CCM allocation was used for set-asides – see http://www.thegef.org/gef/sites/thegef.org/files/documents/C.39.Inf_.10_Programming%20Approach%20for%20Utilization%20of%20Set-asides,%20October%2019,%202010_0.pdf

[12] http://stapgef.org/sites/default/files/Climate%20Change-A%20Scientific%20Assessment%20for%20the%20GEF_2.pdf

[13] The IPCC Climate Science report will be available in September 2013; the Adaptation report in February 2014, and the Mitigation report in April, 2014.

Mitigation to avoid a mean temperature rise greater than the 2°C maximum is the objective of the UNFCCC, as agreed by consensus of all countries at the 15th Conference of Parties held in Copenhagen in 2009. It remains technically feasible with many mitigation actions being cost effective. However, the political will and national commitments remain relatively weak so that annual global GHG emissions continue to rise and atmospheric concentrations of carbon dioxide and other GHGs are nearing critical levels.

Several OECD countries have managed to reduce their carbon intensity by decoupling GHG emissions from economic growth, but this is more challenging for non-OECD countries with key sustainable development goals and population growth to also consider. The business-as-usual dependence on fossil fuels for their future economic growth needs to be overcome by the rapid deployment of low-carbon "leapfrog technologies". This will avoid lock-in resulting from any current and future investments in new, high-carbon emitting heat and power plants and transport infrastructure.

Regarding land use, land use change and forest sinks (LULUCF), there has been little advance in recent years through IPCC/UNFCCC developments. The negotiations for the second commitment period of the Kyoto protocol (2013-2020) acknowledged the deficiencies in the potentially unbalanced approach that enabled only selected areas of land to be included. The need to move to all-inclusive accounting for national carbon accounting systems has therefore been set as a future goal. The loophole of voluntary accounting for carbon stock change in "forests that remain as forests" has been closed, and forest management has become a mandatory category for Annex I parties. Carbon stored during the life of useful wood products (such as building materials) has been recognised, but not the carbon in wood product wastes that end up in landfills. Force majeure provisions for unplanned losses (such as through fires and storms) have enabled countries to be exempted from accounting for an area affected by such events, thus excluding those emissions from the GHG inventory until the carbon stock has been restored.
Forest data being reported by non-Annex-1 parties remain limited, but the latest IPCC guidelines for agriculture, forests and other land use (AFOLU) contain reasonable generic data that can be used to better estimate carbon stocks per hectare for different forest types. However, soil carbon datasets remain uncertain, although a framework has been proposed to fill the gaps[14].

3. To what extent are the chosen indicators for the STAR aligned with the GEF-5 priorities for the climate change focal area?

GEF-5 priority areas for climate change mitigation[15] were to promote:

1) the demonstration, deployment and transfer of innovative low-carbon technologies;
2) market transformation for energy efficiency in the industry and building sectors;

[14] Smith P et al., 2012. Towards an integrated global framework to assess the impacts of land use and management change on soil carbon: current capability and future vision. Global Change Biology 18, 2089–2101, doi: 10.1111/j.1365-2486.2012.02689.
[15] http://www.thegef.org/gef/sites/thegef.org/files/documents/document/GEF-5_CC_strategy.pdf

3) investment in renewable energy technologies;
4) energy efficient and low-carbon transport and urban systems; and
5) conservation and enhancement of carbon stocks through sustainable management of land use, land use change and forests.

These strategies can be grouped into fossil fuel related projects and LULUCF-related projects, in line with the two components of the GBI.

Fossil fuel related projects come under the priority area strategies 1 – 4 above. They can all lead to GHG savings[16] and are therefore linked to the GBI indicator. At first glance, larger GHG emissions might imply larger mitigation opportunities so could absorb higher amounts of GEF funding. On the other hand, the ease of mitigating emissions depends on the source of these emissions; some sectors being more ripe with technical alternatives than others (e.g. low-carbon electricity generation is easier to achieve than low-carbon transport). A new strategic aspect of the GEF CCM strategy is the explicit placement of innovative, new, low-carbon technology demonstrations as the first option. While this strategy has long-reaching roots going back as far as Operational Program 7, it has consistently lacked the necessary concentration of resources to be effectively implemented. The strategy document itself points out that "in large countries and emerging economies with strong technical capacity and market potential, emphasis will be placed on market demonstration and commercialization of innovative, emerging technologies." This requires a certain concentration of resources in these countries, which potentially can be facilitated by the STAR.

Thus, the strategies as a group and the GBI indicator are to some degree aligned – the perceived need to compile all indicators into one composite index is fudging the picture to some degree - and the STAR might facilitate the implementation of some of the strategies. But the strategies are also allocated funding envelopes and achieve strategy-specific outcomes. If STAR allocations are used by the countries in their specific country programs, these will eventually determine the distribution of GEF funds across the different strategies. This might prevent the achievement of strategy-specific outcomes by avoiding having two set of contradicting outcomes which might be problematic. In addition, large GHG emissions might lead to large allocations but do not necessarily attest to the readiness of a country for implementing significant mitigation measures (see below).

LULUCF -related projects in GEF-4 and, with some modifications, in GEF-5 have been partially implemented through country envelopes and with the possibility to receive a top-up

[16] Albeit at different marginal abatement costs.

incentive from the cross-focal Sustainable Forest Management/REDD+ set-aside program. When a country undertakes a LULUCF-related project (the forest type being consistent within the IPCC definitions) and combines significant shares of their STAR allocations from at least two of the three focal areas (climate change, biodiversity, land degradation) but totaling no more than $30M, then GEF provides additional funding through the SFM/REDD+ program at a ratio of STAR=3 : SFM =1[17]. As 5% of the GEF CCM strategy funds were supposed to go to forestry, the 5% weight seems appropriate. On the other hand, and as admitted by the designers of the STAR, the forest-only indicator of forest cover change would not properly represent any non-forest-related land use carbon stock changes.

4. To what extent is reliable data available on chosen indicators? Have different types of datasets and data sources emerged that provide more reliable data than used for the STAR?

A review of the databases used in the RAF was undertaken by the STAP in 2009[18]. STAP recommended to continue with the climate analysis indicator tool (CAIT database) of the World Resource Institute but also to link it with the JRC Emissions Database for Global Atmospheric Research (EDGAR). This recommendation is reviewed below given the changes in the databases. Whatever model is used, the challenge will be that of measurement and data collation. Developing new indices may encourage countries to develop standardized internal systems for measurement but this is outside the TOR for this report.

GBI

Since the start of the GEF-5 period, a new version of CAIT has been produced as also for EDGAR (2.0 Beta). The updated EDGAR currently provides time series data from 1990-2011 for annual CO_2 and CO_2/capita emissions arising from fossil fuel combustion (excluding international transport) plus cement manufacture. For non-CO_2 GHGs, available data now includes total annual emissions per capita from 1990-2010. So unless better datasets become evident, EDGAR and CAIT should continue to be used and the GBI updated using 2010, or even 2011, data.

As part of this STAR review, a comparison has been made for 8 selected countries between the most recent EDGAR CO_2 data; CAIT CO_2 data to 2010; CAIT all-GHG gases excluding land use change and forests (LUCF); CAIT all-GHGs including LUCF; and EDGAR all-GHG data. (Germany, New Zealand and United States are ineligible for GEF funding but were included to give a broader overview). Several discrepancies between the databases became evident during this comparison (including inconsistencies with some energy CO_2 data taken from a

[17] http://www.thegef.org/gef/sites/thegef.org/files/documents/C.38.Inf_.4.Rev_.2%20-%20Investment%20Guidelines%20for%20Forest%20Management_0.pdf

[18] Michaelowa A, 2009. An evaluation of the current global benefits index in climate change (GBICC) used in the resource allocation framework (RAF) of the Global Environment Facility (GEF). In STAP document (undated) "Assessing Global Climate Benefits for Allocation of GEF Resources: conclusions and recommendations of the STAP"

New Zealand government web site used as a quick check[19]). However, in most cases there appears to be a reasonable correlation between EDGAR and CAIT CO_2 data, the exceptions being New Zealand and Vanuatu (one of the small island developing states - SIDS).
The all-GHG data, particularly where LUCF data were included, were more variable and appeared to confuse the picture for some countries. More recently, the producers of the FAO Global Forest Resources Assessment 2010[20] claimed it to have higher accuracy than previous assessments. An updated version based on more remote-sensing activities, and hence possibly with even higher accuracy, is expected in 2015. These developments seem to reconfirm the decision not to include LULUCF data in the GBI for the GEF-5 until other more accurate databases appear (possibilities include http://www.planetaryskin.org). The current approach, based on changes to forest cover area, although not ideal, should probably be retained until the claimed improved accuracy of forest data reducing the wide uncertainties can be determined.

GPI
In the GEF-5, CPIA data from 2008 were used to rate countries. Since then a 2011 IDA Review has been undertaken and the criteria reviewed[21] and in 2012, the IDA Resource Allocation Index for 80 countries (up from the 75 used in the GEF=5) became available.[22] For GEF-6, any data gaps for countries lacking CPIA data will once again need filling from a direct request by GEF to the World Bank.

It was not possible to easily compare in detail the revised version of the CPIA and the 16 criteria used with those selected indicators ("one plus five"), as used in the GEF-5. However, from a brief assessment, there appears to be no reason why the CEPIA, (based on indicators within the Policies and Institutions within Environment Sustainability) and the BFI (as constructed from the Public Sector Management and Institutions cluster indicators), should not continue to be used in the same manner. It is assumed that small island developing states (SIDS) will once again need to be given an average score for the CEPIA and BFI (instead of the minimum) to avoid being unfairly penalised due to limited previous engagement with the World Bank.

GDPI
One concern of this methodology was that some SIDS have small economies and low populations but still have a relatively high GDP/capita. In these cases, the GDPI does not properly address the vulnerability of SIDS to natural disasters. However, this has been taken into account by the STAR process with the use of the minimum allocation floors.

5. To what extent are the weights accorded to different global environmental indicators, and their constituents, appropriate?

Overall, assessing the country scores and allocations has evolved to become a complex process due to attempting to accommodate a wide range of national circumstances. Compared to the RAF, several of the weightings assigned to the various components of the indices were varied in the STAR to give a fairer outcome. These now need to be reviewed to

[19] http://www.med.govt.nz/sectors-industries/energy/energy-modelling/data/greenhouse-gas-emissions
[20] http://www.fao.org/news/story/en/item/40893/icode/
[21] http://www.worldbank.org/ida/papers/CPIAcriteria2011final.pdf
[22] http://www.worldbank.org/ida/IRAI-2012.html. For GEF-6

ascertain whether they are the most appropriate. The perceived need to compile all indicators into one composite index makes the formula less transparent than, for example, a step-by-step allocation process would be. By using this approach, countries could possibly more clearly understand what aspect of their statistical data as used in the indices is responsible for them achieving a higher or lower allocation. On the other hand, this approach would not necessarily make the indices, nor the allocation process, less complex.

Overall country score formula

The relative weights of the GBI and GPI were originally set in the GEF-4 in an attempt to balance the recent performance of a country against its potential to further reduce GHG emissions. They remained the same in the GEF-5 but the GDPI was added to provide additional support to LDCs. As discussed below, a country's allocation is very sensitive to the choice of the exponential of the GDPI, which indicates that a lower weight might be given to that indicator.

A detailed review of the country scores given under the GEF-5 should be conducted at the end of the GEF-5 period to compare them with actual project uptake with the objectives of:

- assessing whether or not the weightings used were appropriate,
- evaluating whether the barriers being experienced by LDCs and SIDS received due recognition, and
- checking whether countries with limited GEF project experience, (and hence unable to receive PIR and TER ratings), were fairly treated by receiving the average PIR and TER country scores in the PPI which seems to be a somewhat tenuous method of assessment.

  A firm recommendation on how to account for the allocations for such countries in a fair manner cannot be given without knowing the detailed outcomes obtained from GEF-5 experiences.

Lacking the resources for empirical analysis, this review is limited to theory-based considerations. In the following section, the GEB indices for climate change are discussed in more detail.

Global Benefit Index GBI

The GBI contains two types of mathematical weights (as outlined in section A) above):

1. a ratio of past to present carbon intensity of the economy; and
2. weighted land use changes through a forest cover indicator as well as an emissions trend.

1) The inclusion of the ratio of carbon intensities of the economy in 1997 and 2007 was meant to avoid "rewarding the polluters". If carbon intensities have increased between 1997 and 2007, then a country's allocation is corrected downward. The adjustment factor used has some high and low extremes which seem to be mostly related to conflicts and general governance disruptions. The five highest adjustment factors belong to post-Soviet countries, thereby enhancing their allocations due to their carbon intensities shrinking after the collapse of the Soviet bloc. The lowest adjustment factors affect countries with (relatively) high economic growth rates and where this might depend on using energy sources with high carbon content. The value of this indicator is therefore questionable in terms of it being a proxy for "polluters" and its contribution to a rational and useful resource allocation.

It is questionable whether "paying the polluters" could, in fact, be a valid funding strategy. Some of the countries that get penalized are experiencing extensive economic growth which leads to higher carbon intensity, for example because they have to use local high carbon fuels in old power plants. For these cases it could be argued that they could make particularly good use of the GEF resource as they are facing large investment booms, and therefore could – for example – leverage large co-financing ratios for clean energy purposes. Whether or not this is a systematic trend cannot be determined at this point. More research would be required to understand whether the carbon intensity adjustment is a good proxy for "polluters" or not. They might in fact offer larger and more cost effective mitigation potential as well as more opportunity to combine climate mitigation with economic or social development. The strongly dynamic development in these countries leads to high rates of learning and significant financial flows, both of which are good pre-conditions for sustainable mitigation action.

2) As mentioned above, the relatively low weight given to forest cover seems to be in line with the GEF-5 strategic priorities at first glance and if the changes in forest cover would be acceptable as a proxy for the overall LULUCF-related changes in carbon stocks. The designers of the GBI added this forest component with reservations so gave it a low weight until the concept could be further investigated. Sustainable forest management was also supported through a cross-focal area set-aside (cf. C.39.Inf._10) outside of the STAR. This means that the absolute total of grants that can be attracted through the 5% shares from the GEF focal areas is a multiple of the allocated envelope. Theoretically, therefore, these projects were able to attract resources over proportion to sustainable forests as they had this additional leverage. More concerning is that the mathematical behaviour of the formula appears flawed:

- It is non-robust so that if countries have had a negative trend in one period and a positive trend in another, the formula technically would lead to negative weight of the forest component. For these cases, special rules had to be designed which demonstrates the somewhat low power from applying the formula.
- The change in forest cover calculations have been normalized as proposed, but exactly how this has been achieved is unclear from the related documents.
- Countries are penalized if they lose their natural resource faster. For example, if forest cover was reduced by 5% in the decade 1990-1999 and then another 5% in the following 5 years from 2000-2004, the change adjustment is 1. If the rate of loss increases (for example, if the loss in absolute hectares stays constant), then the funding decreases (*ceteris paribus*). This extends the logic of using the carbon intensity adjustment to avoid paying the polluter. It is not specified whether this is the intended rationale of the correction. A key point can be made that these countries need more help to stop the loss of their forest cover. On the other hand, it is probably hard to stop trends of forest loss through an environmental program like the GEF as the root causes might lie deeper than the reach of these projects.

GPI

The STAR policy document specifies that "the GPI is a proxy for performance, considering actual performance from GEF projects, commitment to put in place environmental policy and institutional frameworks, and governance and financial management"[23]. The Council has accepted this rationale, so it has become the role that the GPI is expected to play. As mentioned above, the GPI uses the Country Environmental Policy and Institutional Assessment (CEPIA) indicator and the Broad Framework Indicator (BFI) from the World Bank's Country Policy and Institutional Assessment (CPIA), and a revised calculation of the GEF Portfolio Performance Index (PPI).

a. CPIA-derived indicators

Studies on the robustness of the CPIA exist (e.g. GtZ, 2008[24]; IEG, 2009[25]). GtZ (2008) determined that some of the claims of the World Bank regarding the CPIA, (in particular the claim that in a historic simulation the CPIA was able to predict growth patterns), could not be reproduced with publicly available data. On the other hand, the effort put into the acquisition of the ratings as well as the important use of these ratings for calculating IDA scores imply that they are produced with the necessary reliability. The criticisms that the studies provide the CPIA-components that the STAR uses[26] do not necessarily question the use of these sub-indicators for the STAR. While the CEPIA covers all areas of environmental governance rather than just those related to the global environmental benefits, it is well chosen as a measure of national environmental policy and performance which is its function in the STAR. Characterized on this basis, the CPIA-derived indicators are well suited to represent the aspects described above.

b. Calculation of the PPI

The method used to calculate country scores for GEF-5 included the PPI[27] that reflects the past performance of a country on GEF projects. While the PPI in the RAF was based on PIR ratings only, the PPI in the STAR also includes TER ratings that contribute 60% of the overall ratings. The rationale given for the higher weight for the TER rating is well placed. We agree that the TERs are more indicative for overall country performance and are less dependent on idiosyncracies than the PIRs, and should therefore be given higher weight.

[23] http://www.thegef.org/gef/sites/thegef.org/files/documents/document/PL.RA_.01.System_for_Transparent_Allocation_of_Resources.031413.pdf Page 11.

[24] GtZ (2008): Adaptation and Refinement of the World Bank's "Country Policy and Institutional Assessment" (CPIA). Discussion paper.
http://www.gppi.net/fileadmin/gppi/CPIA_Study_Online_Dec2008.pdf

[25] IEG (2009): The World Bank's Country Policy and Institutional Assessment -an Evaluation.
http://www.scribd.com/doc/48069713/The-World-Bank-s-Country-Policy-and-Institutional-Assessment-An-IEG-Evaluation

[26] GtZ (2008) critiques regarding the BFI is that it is too heavily reliant on economic governance – for the GEF however, this focus can be considered appropriate. IEG (2010) notes that the CEPIA is highly compounded as 85 questions go into this rating.

[27] See Annex 3 of GEF paper "POLICY:PL/RA/01, 14 March 2013

A concern arises for countries with a short GEF track record. This results in weak data bases – a challenge recognized before and accommodated with a formula.[28] Even if the data base for the PPI has improved since the RAF, it can still be very weak, for example by being based on one rating from only one PIR. In the basis for calculating the PPI for the STAR in GEF-5, 50 countries had no TER rating and 4 countries had no PIR rating. Countries without any ratings get the default values, but countries with one measurement only get that measurement. Judging from the number of countries that have no ratings, an even higher number is likely to have only one rating each. This therefore has the potential of skewing the index markedly. In the worst case, one project that happened to hire an underperforming project manager received a "Highly Unsatisfactory" rating in the first year. Such a rating can lead to a reduction of the country allocation for the following 4 years and penalize a whole country (and the global environment) for something that was actually a small incident in the implementation of a single project, maybe even the mistake of an implementing or executing agency. Even worse, small database errors in the PIR database might have similar impacts. Hence the PPI indicator is not considered sufficiently robust. A simple remedy for this would be to set a minimum number of necessary ratings (e.g. 4), and if only a smaller numbers of ratings are available, the average score should account for the rest. For example, if two actual PIR ratings are available, the weighted average should include twice the default value.

c. Calculation of the GPI and the relative weights of the components

Compared with the 10% weighting used in the GEF-4, the PPI was given a 20% weighting within the GEF-5 GPI to better link this index with the historic performance of GEF projects. However, this has had limited effect since, in reality, project performance does not vary much for most countries and few observations lead to extrapolated data based on the global portfolio average.

The weightings used for CEPIA (representing environmental policy) and BFI (representing general governance) were consequently reduced to 65% and 15% respectively in the GEF-5 GPI formula. The impact of this difference to the RAF (that had 70% and 20% respectively) was small. However, while the CEPIA expresses the quality of environmental policy, the BFI incorporates all non-environment-related aspects. From a theoretical point of view therefore, this change overall reduced the influence of general governance aspects in favour of environmental policy aspects. This was not the original intention of the proposed change which related to enhancing the GEF portfolio impact only. Therefore, in order to fully comply with the intent of the change of weights, the BFI weight should have been kept at 20%, the enhancement of the PPI maintained at 20%, and the weight of the CEPIA reduced to 60%.

GDPI

The GDP indicator was introduced to give higher relative allocations to LDCs. As the LDC definition is tagged to GDP per capita, this is an appropriate measure. However, the exponential formula used might be considered sub-optimal for at least two reasons:

[28] Described in PL-RA, Annex 3, p. 24.

- The exponential selected for the GEF-5 was -0.04. As a result, resources were shifted from richer to poorer countries around a base level GDP income of approximately USD 3000/capita year. This means that LDC countries with incomes significantly below this level (normally defined as USD 1000/capita year) are rewarded to the detriment of richer countries. But even countries that are not LDCs anymore receive more than they would without this correction.

- The allocation is rather sensitive to the chosen exponential, as already pointed out by the GEF Secretariat. The weight of this index used in GEF-5 was set lower than that commonly used by multi-lateral development banks to reflect the different mandate of the GEF. Changing the exponential by say $\pm$ 25% (to -0.03 or -0.05) would cause a major shift in allocations of several percentage points.

Both phenomena might be considered fair and justifiable but their distributional effects, and the intentions behind them, are not explicitly discussed in the rationale of the formula which intended to support to LDCs only. This implies that some consequences might be unintentional and so it could be worth spending some time on further deliberations. There is no evidence whether adjusting the exponential would be beneficial for LDCs or not In essence, investments are made in low carbon development pathways rather than in mitigation. So for the LDC adjustment, alternative formulations might be considered, e.g. changes in the carbon intensity of GDP over time could be a sufficient indicator or linking the floor to the population of a country. Once again, analysis of GEF-5 experiences might serve to help identify whether to adjust the chosen exponential or change the overall approach to better support LDCs.

To summarize, overall the indicators, as well as their weightings, seem sound. The most important point for further discussion is whether the ratio of historic to recent carbon intensity is an appropriate proxy for "not rewarding the polluters" and whether that is a strategically important dimension in the first place. A look at the data confirms that at least the extremes in this distribution were benefiting or suffering from external shocks that affected their carbon intensity. Among all other indicators, only two minor aspects have given reason for critical comments:

- GPI: the collective weight of the environmental indicators (PPI and CEPIA) should probably stay at 80% and governance (BFI) at 20% rather than losing weight – governance is an important aspect of being able to run successful projects.

- GDPI: the present formula is extremely sensitive and does not reward only LDCs but every country with a GDP per capita under 3000 USD.

6. To what extent are the GEB indices – within the context of performance and socio-economic indices as included in the composite country score indices – effective in directing the GEF resources to countries where there is greater realizable potential to generate GEBs?

The overall objective of the GEF is to generate global environmental benefits (GEBs) and this overarching mission should therefore be embedded in any policy that GEF formulates or uses. The objective of the STAR, however, is to allocate scarce GEF resources to countries in order to maximize transparency and consistency with respect to governance and environmental policy aspects. Lining up these objectives creates major challenges in two pertinent dimensions that need to be addressed:

a) Do the CCM country allocations mirror the actual ranking of countries in their potential to generate GEBs?

b) Are country-based allocations the best way to generate GEBs in the CCM focal area, or might it be better to revert back to a first-come, first-served basis using competition on a project by project basis?

Unfortunately, assessing the utility of country allocations is not possible as long as the "potential ability of a country to generate GEBs" cannot be quantified. How to measure this potential is the challenge. For example, maybe a country has more influence on global benefits generation if it leads a negotiation group in the UNFCCC negotiations, or hosts a COP? Having influence and gaining larger GEF allocations may not encourage a country to move towards a more constructive form of collaborative behaviour in order to gain GEBs. Therefore, it is difficult to measure how much "potential" a country has for generating GEBs. GHG emissions and changes in GHG emission trends can be seen as a proxy. This review can only serve to trigger some initial discussions on these aspects.

It was not necessarily a main objective of the STAR to direct resources to countries that have greater realizable potential to generate GEBs, even though this is the over-arching objective of the GEF. Rather, the Council documents state that the STAR (as well as the RAF) should constitute*"...a system for allocating resources to countries in a transparent and consistent manner based on global environmental priorities and country capacity, policies and practices relevant to successful implementation of GEF projects."* A brief summary assessment by the panel on how the STAR attained this objective (Table 1) indicates that the STAR is tailored to adequately address this objective.

Table 1: Summary assessment of the STAR GEB's contribution to STAR objectives

A system for allocating resources to countries in a:	**Represented by....[component of GEB or aspect thereof]**	**Assessment and Rationale**
..transparent manner...	Public accessibility of the methodology.	Yes, the STAR is transparent. Almost all data are publicly accessible and the calculation can be controlled by other experts and NGOs.

...consistent manner...	Consistency across countries regardless of their economic status and stage of development.	Yes, the STAR is applied consistently. The formulae are chosen such that only in a few cases do special rules need to be invented. Counter examples do exist however; for example the forest adjustment factor is not applicable if the trend changes direction, and the maximum allocation cap restriction on larger countries.
...based on global environmental priorities...	GHG emissions, carbon intensity improvements, forest cover, ratio of trends in forest cover.	These indicators are related to global environmental priorities. They are not very focused on the CCM programming modalities, strategies and options, but that is not a specific requirement of the objective.
...and country capacity, policies and practices relevant to successful implementation of GEF projects	PPI, CPIA, BFI	Yes, these indicators have been found to represent the country capacity, policies and practices relevant to successful implementation of GEF projects. Critiques regarding the weights have been formulated above.

The evaluation question F) formulated above is slightly different, and asks whether the STAR also maximizes impact on the Global Environmental Benefits. A number of factors in the setup of the GEF need to be considered in the STAR and limit the possibilities for this optimization. For example, from an economist's perspective, funding should address the least cost mitigation opportunities but the costs of these are often so low (often negative) that no grant funding would be required to address them. Grant funding would not be spent on these options even though more GHG benefits could theoretically be achieved in these options than in those options that have real (positive) incremental costs. This type of optimization according to the marginal abatement costs is not implementable in the GEF as it ignores the specific value of the "GEF resource" that is grant funding.

In addition, there is a common understanding that countries that are willing to implement GEF-eligible mitigation measures cannot and should not, *a priori*, be excluded from GEF funding. These might not necessarily be the countries that offer the most cost effective or largest mitigation opportunities. The current allocation system could therefore have limited their access to funds[29].

The theoretical optimum distribution from an economist's perspective would be the Pareto-optimum. In this situation the marginal benefit of additional funds in one country would be exactly the same as the marginal benefits of additional funds in another country. The allocations would be of different sizes and depend on the mitigation opportunities in the countries. As we have no way to assess (given the dearth of empirical measurement) whether we are at this point or not, we cannot judge if the allocation is optimal. Whether or not these mitigation opportunities that could not be satisfied from GEF resources due to the limited access, would have resulted in higher GEBs than those that were funded instead in countries with higher allocations, cannot be determined on a theoretical basis in the context of this analysis. –This needs to be assessed on the basis of empirical research.

[29] Even as fair as the present measures have attempted to be, they can skew the original allocation towards the poorer countries.

Also, in a situation of limited resources and under equity considerations as presented, the Pareto-optimum might not be achievable. Overall, the GEF funds are very limited compared to the size of the challenge, the various assessments on the funding needs and even the pledges of the Annex-I countries at the COP-15 in Copenhagen. The GEF has therefore adopted in its funding paradigm the need to be catalytic and leverage the limited resources with co-financing and multiplication effects.

Mitigation opportunities do exist in each country, but there are a number of factors that increase their number, size and accessibility (and thus potentially also lower their unit costs). In general, countries with higher GDP, more people, more cars, and more oil resources typically have more weighty mitigation opportunities than others. Overall GHG emissions are then a good proxy for the number and scale of these mitigation opportunities. A correction using the carbon intensity adjustment factor, however, contradicts this logic. It rewards countries that have already implemented mitigation opportunities, and reduces the resources available to those countries that still offer a large number of opportunities. As an example, countries with dynamic growth (like Vietnam) often exhibit less than perfect scores on this count, but offer more sizeable mitigation opportunities due to significant investment activities. Reducing the attention that the GEF receives in these countries through reducing the GEF envelope might further weaken the priority that country policy makers give to these mitigation opportunities. This might lead to lost mitigation opportunities and result in lock-in to highly emitting, carbon-based, capital stocks.
It becomes more complicated since the GEF funds are so limited that large impacts on GEBs can only be achieved through catalytic and leveraging action. This means that the GEF's definition of "global benefits" (or the benefits that are measured to determine the Pareto-optimum) needs to be carefully considered. In the light of the need to be "catalytic", is it really possible to measure whether the global benefit of a role-model, zero emission, small island developing state (SIDS) is higher than the replacement of a coal-fired power plant with wind farms in say China? Potentially, the absolute GHG savings of building wind farms in China would be much higher, not to mention that they might also create more local benefits in terms of jobs and development opportunities. But demonstrating that a zero-emission SIDS is possible might have stronger impacts on the general perception of the potential for renewable energy and energy efficiency, and encourage many more players to do more for the global climate in more powerful ways than displacing a single electricity generating plant could ever do. The GEF mandate of being catalytic and leveraging, which is caused by the limited amount of funding available, is exactly the reason why a more refined definition of Global Benefit Impacts of the GEF is required than purely direct GHG emission reductions.

SFM/REDD+ With respect to the forest aspects, the statements of the GEF Secretariat already raise a number of caveats that serve to confirm that it is necessary to keep comparing the indicators with international developments both on the policy level as well as on the data level. The potential of a country to conserve forest cover certainly depends on the existing forest cover. Other carbon flows, including from natural or semi-natural sources (like cattle, or burning coal seams) are not included in the allocation formula. On the other hand, the linkage between the country allocation and the cross-focal SFM/REDD+ fund demonstrates an interesting leveraging opportunity. The specific impacts of this construct are not discussed further here but it is assumed that they will appear in another part of this evaluation.

Other types of potentials of a country to impact on the global greenhouse gas emissions situation. Some countries cannot only influence their own GHG emissions, but through South-South-Exchange, trade relations, or other types of influence, also impact on other countries' emissions. For example, China and Korea produce a significant share of the consumer appliances used all over the world, and therefore the energy efficiency of Chinese and Korean products might determine the energy consumption in many other countries. Other countries might depend on the export of wood resources, fossil fuels or beef, so that consumer behavior in downstream countries might impact GHG emissions in exporting countries through the carbon embedded in imported goods. However, including trade balance-related indicators is difficult and again, the question arises "what is a potential".

Overall, the indices are not really dynamic by necessity (constraints include data limitations, and the declared objective of facilitating medium-term planning in the countries with a known GEF envelope). Parameters do not reflect "mitigation readiness" or "climate finance readiness". Current research is conducted on these issues in a number of institutions and networks. Its results should be taken into account for including these aspects in the STAR in the future. Indicators reflecting country climate finance readiness should be included in the governance aspects of the STAR. Low values on these indicators could potentially also trigger additional technical assistance funds for enhancing climate finance readiness.

Yet another dimension of the question "where is the largest potential to deliver global environmental benefits" is related to global needs[30] – i.e. common goods that are needed to moderate the climate crisis. There is for example a distinct need to develop and disseminate new technologies to manage energy-related and other carbon stocks. Developing, testing and disseminating these technologies would be a large effort for any single player or country. Some technologies and approaches can only be developed in GEF recipient countries as their development needs to be adjusted to a distinct climate or culture (e.g. building technologies, or rural energy supply technologies). While the GEF overall climate mitigation funds would be able to support such developments, the GEF allocations and the STAR as currently tailored do not support cross-country cooperation, and limit the possibility to pool resources for joint deployment. While formerly the GEF Council was able to guide such large efforts, the influence is now limited. Countries would have to include such concerted efforts in their medium term plans and fund them from their country envelopes.[31]

7. What are the areas where STAR GEB indices for climate change focal area may be improved upon and how?

a) Short-lived radiative forcers and soil carbon

STAR allocations that would support mitigating short-lived radiative forcers such as methane, black carbon, etc. need careful review since it is generally agreed that this approach, whilst it might buy some time before reaching the 450ppm stabilization level, should not become a substitute for mitigation of longer lived CO_2 emissions. The current STAP work program includes an assessment of short-lived radiative forcers and their mitigation for possibly

[30] Regarding the option to base allocations on local/national formulated needs, see "NAMAS" below
[31] This was possible with some of the programs from the set-aside. Empirical analysis is needed on how that worked, and whether or not it was in fact more difficult than without the country envelopes.

including in the CCM Strategy. If it is agreed that they could be included in future GEF projects, then some form of indices would need to be included in the STAR. Mitigation through increasing soil carbon (including biochar) is too scientifically uncertain for GEF to support at this stage.

b) Research on GBI and allocation logic

A question often raised is whether the STAR should consider national GHG emissions or base the allocation process on a CO_2 emissions per capita basis. There is a major discrepancy between the USA, Qatar and Australia (~17-19 t CO_2 / capita year) and LDCs (~ 1-2 t CO_2/capita year) with the global average currently around 5 t CO_2/capita year. It has been calculated that based on population growth by 2050, the global average needs to be reduced to ~2 t CO_2/capita year in order to restrict warming to below 2°C above pre-industrial levels. The challenge is that most major Annex-1 countries are showing a reduction since 1990 to 2011 whereas most non-Annex-1 countries have shown significant increases as they develop economically[32]. In addition, when the trade of products is taken into account, the picture becomes more obscure. For example, the UK has shown a per capita CO_2 reduction since much of its manufacturing has moved off-shore, but it now imports a far greater share of consumer goods and machinery from Asia and the carbon footprint of these goods are presently not included in the UK national inventory.

As discussed, direct avoided GHG emissions are a coarse measure for the GEF impact, but do not tell the whole story, in particular not in the context of the GEF where catalytic impact is the order of the day. While total emissions might be a good proxy for mitigation opportunities, other factors also determine the potential of a country to contribute to the global environmental benefit. These are partially very tangible, like local engineering capacity and trade streams, but partially also intangible, like role model functions.

In order to tailor the GBI to the purpose of the GEF funding paradigm, the "allocation program logic" needs to be more clearly defined: What are the causal mechanisms that will allow GEF to maximize GEB impacts through the STAR? Once it is clear how the impact can be measured, such questions can be answered more easily, and the STAR can be designed towards maximum impact. It is important to note, however, that this was not part of the original objective of the STAR.

c) Clear and GEF-operational definition of GEBs and Pareto-optimal allocation

As explained above, the economically optimal resource allocation would be at the point where one dollar more would generate the exact same global benefit in every country (Pareto-optimum). The only way to determine this would be through extensive empirical analysis of marginal benefits in all countries. In addition, any such determination or any

[32]http://www.pbl.nl/sites/default/files/cms/publicaties/PBL_2012_Trends_in_global_CO2_emissions_500114022.pdf

allocation that targets the optimization regarding the impact on global environmental benefits would require a clear, comprehensive and workable definition of the global environmental benefit.

d) Credibility through independence

Governance-related indicators, according to the assessment of this panel as discussed above, are effective and useful in terms of reflecting the GEF Council's intentions. Another critique of the indicators, however, is that the confidentiality requirements, as well as the provenience of the data, reduce their transparency, and the fact that they are created by the World Bank might give reason to doubt their independence. These two aspects do not make it immediately apparent why the database and transparency of this indicator should be so much stronger than the other options that were considered by the Secretariat.[33] GtZ (2008) for example, detected that "parts of it still reflect the priorities of the Washington Consensus". These criticisms will consistently resurface with every review of the Allocation Framework and its governing indicators. A civil society source for governance-related indicators might be trusted better in this context.[34]

The CPIA-derived indicators are also determinants of the IDA allocations. This might be an advantage when considered from the viewpoint of the GEF as it might allow for larger IDA co-financing to become available if and when the GEF allocation becomes bigger. However, if a country ranks poorly on this indicator, then resources from both GEF and IDA will be reduced. That means that non-governmental actors (international agencies, private sector, civil society) who might be able to use GEF resources well, even in a context with poor governance, will also have access to fewer resources.

Due to these aspects, the GEF Secretariat should feel encouraged to continue its search for "better indicators" in the sense of equally or more robust and more widely accepted governance indicators. It is not to be expected that the outcomes, in terms of the allocation, would be changed drastically overall, but the STAR might be less exposed to criticism and more widely accepted.

e) Mitigation readiness

As discussed above, countries might show an increase in their carbon intensity for various reasons, some of which could justify higher resources rather than lower resources. General GDP growth would be easier and more reflective of the availability of resources for

[33] "In developing the STAR, the Secretariat reviewed a number of indices currently in use, including the Human Development Index, the Environmental Sustainability Index, the Environmental Performance Index, and the use of Millennium Development Goals indicators, with the aim of further strengthening the GPI. After careful analyses, it was concluded that that these indices were not sufficiently robust for a resource allocation model, neither from a transparency perspective nor from a consideration of data quality and availability." p.11 of PL-RA-01

[34] Although the process proposed in GIZ (2008) – to open the assessment process and increase its transparency – might also help,it is largely outside of the influence of the GEF Secretariat .

investments – typically, the mitigation potential is high in these situations, as lifestyles become more luxurious. Leapfrogging would enable these countries to achieve Western comfort standards without the dependence on associated carbon-intensive and inefficient energy technologies.

Historic performance on mitigation, as expressed in the carbon intensity adjustment factor, might not necessarily be the right measure. Depending on the definition / objective, there are many other potential parameters as listed (although this list has limited practical value).

- Higher existing GHG emissions (total / per capita).
- Higher expected GHG emissions in the future (total / per capita).
- Higher ability to reduce their own GHG emissions (because of capital stock, lock-in effect, innovation systems etc.).
- Higher influence on how GHGs are produced globally (such as for exporting nations, oil producing nations, nations with significant political influence).
- Lower ability to change the track that they are on without external help.

On the other hand, mitigation opportunities do not necessarily imply that a country is ready to actually implement some or all of the mitigation measures, no matter whether they would be economically beneficial or not. This might depend on political movements and habits like a tradition of energy subsidies.
Once a country is ready to implement mitigation measures, it might still not have the institutional framework to actually receive carbon finance and utilize it in a way that maximizes the global benefit. Current research is being undertaken on climate finance readiness, a concept linked to mitigation readiness which should be observed closely in future revisions of the STAR.

f) The potential for linking with NAMAs

Nationally appropriate mitigation actions (NAMAs) are policies, programmes and projects that developing countries voluntarily undertake as their contribution to the global efforts to reduce GHG emissions. Some developed countries have committed to supporting such meaningful mitigation actions in developing countries through financing, technology transfer and capacity building. The policy framework around NAMAs is still being developed through the UNFCCC process, but it is becoming evident that NAMAs could become valuable evidence of intent and action for use in future climate negotiations and agreements. They are becoming the climate finance vehicle of choice for developing countries that wish to voluntarily implement GHG mitigation actions in support of sustainable development[35].

Together with partners from in and around the UNFCCC, several countries are beginning to operationalize the concept in order to source and lever climate finance from both domestic

[35] For example, IISD and partners are undertaking research to support NAMAS in 13 developing countries using funding from donor countries and organizations
http://www.iisd.org/pdf/2012/brochure_cop_2012.pdf

and international sources and donors. Developing low-carbon, climate-resilient frameworks can help national governments align sustainable development priorities with mitigation and adaptation aspirations. NAMAs are becoming a central element of this process. They require a focused assessment to identify actions that support development objectives so that mitigation and adaptation efforts can complement development actions. As of 1 June 2013, the NAMA database showed 59 NAMAS had either been submitted to the UNFCCC registry, or had evolved into concepts, proposals, plans or implemented projects[36].

GEF Council needs to carefully consider whether future STAR allocations should support the NAMA process. Indeed, it is conceivable that a total revamp of the STAR CCM focal area in the future could revolve around the development and uptake of NAMAs.

In addition this question would also require for the STAR to take into account that the landscape of climate finance has changed drastically every year since the initiation of the RAF. Initially, there was almost no financing for Sustainable Forestry Management/ REDD. Now there is. A country like Kazakhstan or Indonesia can go to the CIFS and get a multiple of their GEF allocations much faster. In addition, there is the Green Climate Fund looming on the horizon.

g) Incentives through GEF allocations?

The expectations for the impact of the STAR have not been the same for all stakeholders. If there were any stakeholders that expected the STAR to have impact on general governance indicators or even climate mitigation impact of the GEF projects, the impression arises from the discussion of these indicators that there might be an exaggerated expectation for the impact of the STAR. Specifically, the relevance of those indicators that are under direct control of GEF focal points for the allocation is very small. In fact, only the PPI can be meaningfully influenced by GEF stakeholders. The GEF allocations are too small to expect any incentive for "better governance" from other non-GEF stakeholders.
If that is the case, potentially, the rationale for the indicators and the allocation formulas should be re-investigated and assessed as to whether this aspect of the GEF allocation forming incentives for "good government behaviour" has played any role in the design of the allocation formula. Potentially, simplifying the expectations that stakeholders have of the STAR might also lead to a simplification of the indicators and assessment formula, and empower stakeholders to plan for even longer periods of time (i.e. beyond one GEF replenishment period).
Increasing the PPI rating in GEF-5 was an attempt to increase the relative importance of the achievements and results from the country environment portfolios, particularly those of the LDCs and the SIDS. The increase of this weighting is supported by the panel. As the data base on GEF performance becomes stronger, the weights might be further increased. For GEF stakeholders, however, the PPI in particular could form incentives in itself, for example for better national coordination, and for national knowledge build-up and competence management in the field of environmental public goods.
If the STAR should provide an incentive for internal change, it might be useful to change the system completely, and move to a more needs-based system, that has more "if – then" relationships of the type: "if LULUCF emissions change rates are more than energy-related

[36] http://www.nama-database.org/index.php/NAMAs_and_feasibility_studies_in_the_database

emissions change rates then….". In particular, this approach could be used for the environmental part of the allocation formula.

Summary and Outlook

This technical and theoretical assessment of the indicators as used in STAR allocations has identified the following.

1) The indicators have been tested with time. The main method of evaluation is an empirical analysis of their direct impact on the transparency of programming and indirectly on GEF projects.
2) From a theoretical perspective, the indicators have the benefit of having some relationship with the intentions of the STAR, and with the aspects that are considered relevant for a country's role regarding climate mitigation and the UNFCCC.
3) However, this relationship gets watered down through the mathematical formulation approach and the perceived need to combine all variables into one country score. This practice limits the transparency of the index and makes it difficult to really analyze the impact of each variable that goes into the index with any scientific certainty. The time granted to the panel was too short to apply the differential mathematical and simulation runs that would have been necessary to conduct such an analysis with scientific rigour. It is recommended that this be done.
   The methodology that was applied here was therefore a mechanical deconstruction and a singular look at each of the indices, therefore updating the work of the Delphi panel for the Mid-Term Review of the RAF in 2009.
4) In doing so, the panel members found that a number of the earlier criticisms have not been addressed and therefore still apply. In particular, the fact that the GBI formula mixes "needs" (i.e. large emissions) with "past performance" is questionable, not only from a transparency and incentive point of view. In theory, a lot of consideration and analysis is required in terms of the weights applied to each.
5) The database has been updated with the move from the RAF to the STAR and uses the best data presently available, although this will need reviewing in future as databases evolve and improve.
6) A forest-related index was included based on changes in ground-cover over time. This has several limitations but remains the best option available at present to penalise countries that continue to deforest at increasing rates.
7) As recommended, also by the MTR-panel, the relative weight of the GEF Portfolio Performance Index (PPI) has been increased. In addition, the PPI has been strengthened by including more emphasis on terminal evaluation reports. However, for the PPI, the panel found two points to reconsider:
   a. Some countries still have a weak database for the PPI. If they have only one datapoint that point is decisive for their allocation, while if they have no datapoint, the average of the whole portfolio is used. If this data point happens to be an "Unsatisfactory" rating, the country gets penalized and treated worse than if it had not used any GEF funds at all. It is recommended to buffer this effect.

b. The weight of the PPI has reduced the weight of the Broad Framework Index (BFI), which is the only part of the STAR indicators that reflects general governance aspects like Rule of Law or corruption, from 20 to 15%. This might be an unintended effect or might comply with the intention of the Council, but this dilution should be noted. The panel would recommend to use the following weights:, CEPIA 60%, PPI 20%, BFI 20% (instead of the present 65%, 15%, 20%).

So, technically, the indicators are more or less sound and fulfil the objective of the STAR.

8) There are larger and more strategic considerations to be formulated when assessing whether the STAR and its indicators support the objective of the GEF optimally. One such aspect revolves around the adjustment factors contained in the GBI. The formula includes an adjustment for changes in historic trends in carbon intensity and forest cover respectively. They are formulated so as to reward countries with trends that lead to endogenous reduction of global environment damage (i.e. reduced carbon intensity, reduced deforestation rates).
   This means they are intending to "reward" already "good behavior". The question is whether they should rather help those who have not been able to make any progress in the past? The analysis of the outliers shows that detrimental trends were often not necessarily the "fault" of the current governments, and in some cases the consequence of external disruptions. The use of historic data penalizes today's country leaders for situations as far back as the 1980, when the world was very different in terms of energy consumption patterns, deforestation, and overall knowledge about climate change and greenhouse gas emissions.
9) Whether the STAR complies with the STAR's objective is shown in Table 1 that indicates, the STAR fulfils its own objectives satisfactorily. The overarching question of the evaluation however, is whether the STAR fulfils the GEF's objective by helping to maximize the global environmental benefits.
   a. Theoretically, a distribution maximizes common benefit if it is Pareto-optimal, i.e. if the marginal benefit is the same across all recipients. It is easy to assume that the STAR is not Pareto-optimal.
   b. In order to analyze this in depth, there would be a need to define what exactly are the global benefits. The GEB in the case of climate change is not the absence of carbon emissions from energy, industry or deforestation, but the stabilization of the global climate. In other words, the tiny incremental efforts that the GEF can support are not achieving this global benefit. They can only help countries understand the scale of the problem, and then leverage local resources such that nations, cities and communities can make some steps towards the global benefit of a "stable atmosphere". Limited by funding size and political constraints, the GEF can only support a small selection of the necessary steps. Measured against this absolute global benefit of a stable atmosphere, all indicators that can be used to measure the GEF's success can only be very weak proxies, as the "responsibility" of the GEF is very limited.
   c. There might be other ways in which the GEF can maximize its specific impact on the global benefit. For example, the need for new technology development or for global political consensus might be more important than any market

development initiative in a country, small or large. The GEF, as the currently largest body of pure grant funding is better suited for some funding purposes than for others. The current allocation system disregards completely these needs of the global community as well as the specific properties of the GEF. Combined with the dynamic field of climate finance, this is a situation that is fundamentally different from the other conventions that the GEF serves.

Key questions for GEF-6 are whether the STAR could be simplified and re-designed so as not to reward the polluters but, rather give higher allocation shares to countries that have a good record of reducing their GHG emissions in recent years. The allocation system should be based on the potential for reduction instead of the current level of emissions.
This review has looked mainly at the indicators and their scientific and technical merit. The panel feels that the following wider-ranging considerations might also be assessed in this review process:

- Consider a completely new design, for example, one that is more needs-based and includes climate-finance readiness. Ideally it would also avoid lumping all indicators into one but base allocations on a step-by-step-process in order to enhance transparency and incentives for the countries. An output-based finance disbursement system could be one option to consider (ex-post rather than the present ex-ante) where the more wealthy eligible countries could invest up-front and then, to incentivize performance, be refunded later based on delivery of project outputs. Or possibly a progress-based disbursement system could be developed where the measured outcomes could include items such as policy and institution building.
- Evaluate whether some countries with high and growing emissions receive too favourable an allocation due to having a high assumed "potential" to reduce their emissions, or conversely, if they receive too small an allocation given their mitigation potential.
- The STAR does not take into account the availability of other large sources of finance that are accessible only to larger countries, in particular the CIFs which are by themselves larger than a GEF replenishment for all focal areas.

Overview of the STAR process in the future
The future options for the GEF Council to determine include whether the STAR can be improved sufficiently to meet the desired objectives of a fair and transparent allocation system or whether a completely new approach should be developed. The GEF Strategy 20/20 document asks "Should the GEF review its current resource allocation system?" The STAR may have helped enhance the predictability of individual country allocations compared with the RAF, but there has been a lack of flexibility between focal area allocations that may have restricted the opportunity for GEF to deploy high impact and visible integrated programmes. With respect to climate change, the current system has tended to reward the emitters by giving more generous allocations to countries with relatively high GHG emissions because their assumed potential to reduce emissions has been deemed to be greater, and hence could lead to more GEBs.

Rewarding the ambitions of countries in order to reduce their emissions is difficult. It cannot be based on national future targets alone since these are often politically much easier to propose than to achieve in practice. Using carbon intensity or energy demand reduction

trends over the past few years may not always result in the anticipated reduction in GHG emissions. Perhaps for GEF-6 the STAR could be modified in order to better support countries with stringent GHG mitigation policies in place in order to reduce GHG emissions, as for example identified in countries with NAMAs in place. Then GEF funding could be used to help meet those specific actions as outlined in the NAMAs.

At present projects that focus entirely on climate change adaption[37] have been excluded from funding through the STAR allocation process but are funded through the Adaptation Fund, the Least Developed Countries Fund (LDCF) and the Special Climate Change Fund (SCCF) that GEF hosts on behalf of the UNFCCC. The LDCF aims to support the special needs of the LDCs with the priority of preparing and implementing national adaptation programmes of action (NAPAs). The SCCF was established to finance activities, programmes and measures relating to climate change that are complementary to those funded by the climate change focal area of the GEF Trust Fund and by bilateral and multilateral sources. Integrating funding for climate change adaptation with mitigation projects in the STAR allocation process, with the aim for developing countries to become more resilient and to reduce economic losses at country level resulting from climate change impacts, would be challenging to achieve successfully. The suggestion was made that the GEF could perhaps focus on mitigation support for the major emitters and on adaptation for the other smaller countries where low-carbon development pathways and co-benefits can arise from investments in clean energy. However, the CIFs also set out with this same objective and changed course later, so an alternative might be to mainstream both mitigation and adaptation through low-carbon development and/or green growth. This could mean GEF working with the finance ministries of recipient countries rather than with the environment ministries and even making allocations on the size of their national budgets. Whether adaptation should be included in the STAR process requires a separate policy debate so is not discussed further here.
Climate mitigation and adaptation goals will vary between LDCs, SIDS and MICs (middle income countries). Most LDCs and SIDS are aiming primarily for sustainable development and successful advancement often depends, at least in part, on improved energy access and mobility. Here GEF can assist in providing "leapfrog" technologies to provide the necessary energy services but avoiding the route of increased fossil fuel dependency. Therefore GEF can maintain its goal of supporting innovative projects by looking towards the long-term benefits that demonstration projects might bring.

Finally, if the revised GEF-6 Strategies currently under discussion do eventually include the proposed "GEF Signature Programs" on Food Security, Sustainable Cities, Forests, and Oceans and Seas, then further flexibility of the STAR may be needed to fully encompass these multi-focal areas and this will need additional detailed consideration.

[37] Turn Down the Heat, World Bank report, June 2013. ISBN: 978-1-4648-0056-6

ANNEX 2.C

Review of Land Degradation GPI

Michael Stocking and Douglas Taylor

1. What is the scientific and technical validity of the global environmental benefits indices?

SUMMARY FINDING: The three proxy indicators used for the STAR GBI$_{LD}$ are valid and should be retained. They address three important aspects of LD and the requirements for SLM: area of land affected by LD (a direct measure of LD); extent of drylands (a measure relevant to the Convention priorities – hence global importance); and population affected (impact of LD on livelihoods). The original data sources should be contacted to update the information

According to the consolidated paper of Council decisions on STAR dated 14 March 2013, the GBI for land degradation is designed to take into account three key factors:

a) The need for controlling and preventing land degradation in the context of production systems;
b) The challenge of combating desertification in dry-lands, including the need for adaptation to drought risks; and
c) The need to address the livelihood needs of vulnerable populations.

The core problem was that these factors are difficult to translate directly to measurable indicators. Proxy indicators were therefore chosen to capture the baseline impact of or outcome of allowing degradation to continue – that is, area of productive land affected; extent of drylands and their particular sensitivity to drought risks; and livelihoods/well-being of affected populations

These three factors remain as valid now as when they were first drafted in 2009. The question that needs to be asked, then, is how well do the proxy indicators perform. Empiricism rather than determinism conditioned the choice of indicator – i.e. what changes could be demonstrated to have some relation to LD/SLM interventions. The Policy Paper (PL/RA/01) also notes that proxy indicator choice was based upon "data availability"; it does not mention scientific relevance.

A proxy has the advantage of rolling up a number of processes in a 'black-box' modelling approach; the major disadvantage of such an empirical approach is that proof of validity (other than in statistical terms) is impossible and there is dubious validity in extending results in one area to another.

PROXY 1 – Land Area Affected (20% weighting – see Section B4 below)
The country estimates for land area affected, are based upon estimates of change in net primary productivity (NPP) as a proxy for land degradation through use of AVHRR derived NDVI data from July 1981 to December 2003. Measurement was by GIMMS radiometer (AVHRR) data collected by National Oceanic and Atmospheric Administration (NOAA) satellites which was stated to be (1) a proxy for "loss of ecosystem function and productivity" and (2) the most feasible proxy for LD trends by country. The NDVI data were published by Bai *et al* (2008) at ISRIC (World Soil Information Centre), through the Global Assessment of Land Degradation and Improvement (GLADA) under the GEF-supported LADA project (FAO/UNEP). Subsequently the proxy was challenged by Wessels (2009), defended by Bai *et al* (2009) and further improved

with mapping (but not the country statistics) updated by Bai *et al* (2011), including addition of data to 2006. Subsequent to the 2011 report no further published scientific challenge has emerged to warrant replacement of the proxy.

The actual measure is the "annual sum of NDVI" in order to represent "annual accumulated greenness". The NDVI value is then adjusted according to Rain-Use Efficiency. The approach of the LADA project was designed as a replacement for GLASOD (1991 – Global Assessment of Soil Degradation) using consistent information sources, rather than GLASOD's 'expert judgement' method. Field observation and verification from North China, Kenya and Bangladesh was undertaken.

Conclusion A much better approach than GLASOD; it is consistent globally. However, it needs to be brought up-to-date, also using higher resolution than the original 8 km. Some surprising results e.g. two small African mountainous countries, both highly degraded – Lesotho has 34% of country affected while Swaziland has 95%. This difference is not readily explicable and is somewhat counter-intuitive. Checking for such possible anomalies on updated data is strongly indicated.

PROXY 2 – Drylands (60% weighting)
Drylands are defined as the proportion of the total area of each country falling into the three classes of aridity index (arid, semi-arid and sub-humid) used by the UNCCD following the method of UNEP. The use of this proxy indicator by the GEF is aligned with UNCCD's core interests and directly reflects each country's opportunity regarding drylands. It is measured as the proportion of country land area within arid to sub-humid zones based on the World Resources Institute (WRI) Earth Trends database.

The weighting of 60% is justified in Annex 7 of PLA/RA/01 on the basis that "drylands are predisposed to desertification" and that desertification is "one of the most prevalent forms of LD". This is somewhat circular reasoning – desertified areas will, by definition, be 'dry', and drylands also will be dry and desertified. The more relevant – but non-scientific – justification for the additional weighting for drylands is that these are a priority for the UNCCD.

Conclusion. This index has the virtue of simplicity, being easy to compute and relevant to the Convention priorities. However, existing country dryland status is not scientifically related to delivery of GEBs – indeed, extreme dryland conditions will be technically the most difficult and costly to rehabilitate (see below on returns on investments in soil conservation). The GEB criterion could, however, be said to have been met in that drylands are the priority target of the relevant global Convention, the UNCCD. Since the data were previously sourced from the WRI, they should be renewed from this source.

PROXY 3 – Vulnerable population (20% weighting)
This proxy indicator calculated as percentage of country total population living rurally; said to be "a good proxy for rural poverty" and hence addresses the important issue that LD impacts on poverty. There is circularity in reasoning again here – large rural populations tend to degrade their land resources; degraded land resources have reduced population-carrying capacity; lowered carrying capacity leads to poverty; and poverty leads to large rural populations.

Nevertheless, in a somewhat simplistic sense, it is true that there is greater rural poverty in degraded areas, and this factor should be included in the calculation of the GBI_{LD}. Several studies in Africa show that poor rural populations 'mine' their land resources of nutrients in order to survive; they have no choice. A further aspect is that having a human development proxy indicator aligns the GBI with the UNCCD impact indicators which also emphasize poverty of rural populations.

Conclusion. Again, this is a simple index, easy to compute and having validated relevance to pressure of human populations to degrade their environment. However, again there is no scientific evidence that concentrating resources on areas with poor rural populations would deliver greater GEBs – it is a political decision that areas most degraded or under threat of degradation should receive more resources. Social and equity criteria would support an allocation based upon the greatest need being with large poor rural populations, subsisting on meagre land resources. An economic case, however, might deliver the opposite result; that investments will achieve greater returns in areas where there are fewer poor rural people. Since the original data source was the World Development Report, 2008, indicator values should be renewed from later reports in this series.

2. Overall, to what extent, the results provided by the specified indices are likely to reflect the importance of a country in terms of global environmental benefits potential for the given focal area?

SUMMARY FINDING: The use of the three proxy indicators provides results that are relatively equitable for calculation of a distribution of resources for LD/SLM investments that can potentially deliver GEBs. The question of weighting - see below in Section B4, Q7 - needs to be reopened in terms of giving very dry and desert countries relatively large allocations.

The results of the indices will depend firstly on their relevance for any particular country, and secondly on the algorithm used to combine the indices into a GBI score.

(1) Relevance to countries. A country that is severely degraded, has a significant proportion of its area in drylands, and has a large rural population dependent on terrestrial resources will score highly. These are the very countries prioritized by the UNCCD as most in need of attention and support. By implication, this is where there should be greatest potential for delivery of GEBs. [It is only 'potential' because these same countries often have least capability in delivering GEBs in terms of national and human resources]. Therefore, on this aspect, the results provided do reflect country importance.
(2) The algorithm – see B.4 'weighting' below

It can be (and has been) argued that investments in SLM (such as soil conservation) are better and more efficiently targeted at areas where degradation has not yet become serious, on the grounds that it is much more difficult and costly to restore degraded areas than to prevent land becoming degraded in the first place. This is a compelling argument that runs counter to the present STAR allocations.

Conclusion. In general, the results provide a simple baseline for calculating a relatively equitable distribution of resources for LD/SLM investments that can potentially deliver GEBs. There are some anomalies especially for countries that are almost wholly dryland and desert, where the outcome of STAR gives them fairly large amounts but the opportunities of

investments in deserts is small in terms of delivery of GEBs [This finding must be set alongside the question below on 'weighting' and the algorithm used to combine the three proxies – this determines the actual allocations.]

3. What new developments have taken place in terms of scientific understanding of environmental problems that need to be addressed in the GEB indices for upcoming replenishment period (GEF-6)?

SUMMARY FINDING: The main new development to arise relates to resilience to climate change and the coping strategies of local populations, coupled to improved knowledge of environmental status and services. This needs to be factored into GEB indices through, for example, includingan indicator of SLM practices ('best practices').

It is known that some societies adapt well to changing climatic conditions; and others not. This needs to be factored into GEB indices through perhaps an indicator of SLM practices ('best practices'). The more 'best practices' reported through a KM platform such as WOCAT, the more resources could be devoted to further investments that will build on the country-based initiatives and will maximize delivery of GEBs. This will also align the GEF LD-FA and its STAR with the UNCCD which is soon to include 'best practice' reporting under the Convention. It will also spur countries to adopt a positive view of SLM and encourage tracking of demonstrations/pilot practices to ascertain GEB delivery (primarily carbon but perhaps also water conservation and agricultural productivity). This positive view would act as a counterweight to the negative attribution of two of the other proxy indicators – land area affected and drylands. This new development may also be seen in terms of our new understanding of the importance of ecosystem goods and services provided by land, soil and vegetation, especially in areas managed under land use.

Conclusion. The present suite of proxy indicators is intended to reflect the potential opportunity for SLM of eligible countries. New knowledge is available regarding carbon fluxes, water budgets, agricultural productivity and the importance of ecosystem goods and services provided by land, soil and vegetation, especially in areas managed under land use. Against this background of improved scientific understanding, the original proxy indicators should be augmented to include the track record for 'best practice' (i.e. wise use that leads to SLM) practiced by affected populations in drylands.B.2. Alignment with GEF-5

4. To what extent are the chosen indicators for STAR aligned with the GEF-5 priorities for the respective focal area?

SUMMARY FINDING: Alignment is generally poor, except for Proxy 1: NDVI. The other two chosen indicators (% dryland, % rural population) are only very indirectly related to the four LD-FA Strategy objectives.

Four objectives contribute to the focal area goal and drive the development of the GEF-5 portfolio – along with comment on alignment with GEF-5 priorities:

a) Maintain or improve flows of agro-ecosystem services to sustain the livelihoods of local communities [this includes the maintenance of the functionality and cover of agro-ecosystems – the NDVI measure will capture part of this];
b) Generate sustainable flows of forest ecosystem services in arid, semi-arid and sub-humid zones, including sustaining livelihoods of forest-dependent people [same comment as a)];

c) Reduce pressures on natural resources from competing land uses in the wider landscape; [SLM and SFM should reduce pressures on land resources through improving and intensifying production on smaller areas, thus allowing degraded area to rehabilitate. But this is only a very indirect and untestable linkage] and
d) Increase capacity to apply adaptive management tools in SLM. [not taken up by the current proxy indicators but would if the suggestion above of adding an indicator based on 'best practices' were adopted]

Conclusion: Although the general objective of reducing land degradation measured by change in productivity (NPP, NDVI), is aligned with proxy indicator 1, the other two chosen indicators (% dryland, % rural population) are only very indirectly related to the four LD-FA Strategy objectives. However, it would be near-impossible to choose to have single and measurable indicators for each of the objectives. An additional indicator capturing the degree to which a country has developed and disseminated 'best practices' would enhance alignment with GEF strategies and with the Convention.

5. To what extent is reliable data available on chosen indicators?

SUMMARY FINDING: The data are readily available from current sources that are regularly updated

Updated data underlying the three proxy indicators comprising: area of land affected obtained from AVHRR derived NDVI data from July 1981 to December 2003, the proportion of the total area of each country falling into the three classes of aridity index (arid, semi-arid and sub-humid) and the percentage of country total population living rurally, can be obtained from the original data providers and remains valid as detailed further in the answer to Question 1 above.

6. Have different types of datasets and data sources emerged that provide more reliable data than used for STAR?

SUMMARY FINDING: The datasets used for LD STAR are reliable and have been refined; there are alternative sources for the datasets which should be investigated further

For Proxy 2 and 3 (drylands and vulnerable populations), there are no new datasets, because the measures themselves are relatively straightforward and standard information:
For Proxy 1 (land area affected by LD), NDVI will remain the basic measure for land cover and hence proxy for LD. However, there is continual refinement of data sets for NDVI and new sources coming on-line and readily available. No studies have yet been undertaken on whether or how these might prove more reliable – but the question should be asked and the following at least investigated:

(1) AVHRR FASIR data set from the NERC Earth Observatory website - http://neodc.nerc.ac.uk/?option=displaypage&Itemid=88&op=page&SubMenu=-1 . This is the Fourier-Adjusted, Sensor and Solar zenith angle corrected, Interpolated, Reconstructed (FASIR) adjusted Normalized Difference Vegetation Index (NDVI) data set and derived biophysical parameter fields generated to provide a 17-year, satellite record of monthly changes in the photosynthetic activity of terrestrial vegetation.
(2) **F**raction of **A**bsorbed **P**hotosynthetically **A**ctive **R**adiation (**FAPAR**) – described as a novel alternative to NDVI. Managed by the EU Joint Research Centre, Institute for

Environment and Sustainability. Has global reach as of 2012. Website: http://fapar.jrc.ec.europa.eu/Home.php

Conclusion. Chosen indicators are robust and easily measurable; and therefore it can be concluded they are reliable since they have been used for many other purposes, not just for STAR. Alternative sources of data to populate these indicators should be considered. It is beyond the scope of this assessment to recommend specific alternatives.

7. To what extent are the weights accorded to different global environmental indicators and its constituents appropriate?

SUMMARY FINDING: The current weighting is arbitrary and has no scientific basis; furthermore the current weighting should be reviewed to enable adjustment of the GBI_{LD} and consequent allocations to countries to rebalance opportunities to deliver GEBs.

The weighting is a compromise solution designed primarily to allocate resources to countries with substantial amounts of dryland – there is no valid scientific rationale for this. It is understood that the algorithm eventually chosen for the STAR is the one that gave resources to countries for which it was politically expedient to allocate funds for LD/SLM. Indeed, it has been argued that investments in semi-arid zones especially bring lowest returns because of the limited options for SLM and because the degradation processes are naturally far greater than in, say, humid areas.

To address this weakness the weighting should be reconsidered. Proxy 1 (land area affected) represents change i.e. progression to degraded status over the 23 year measurement period; countries can be expected to address land degradation through GEF support. Proxy 2 (extent of dryland), however, is the baseline geo-physical location of the country concerned. It is proposed to reduce this dryland signal and to increase the attention to degraded land in order to treat them equally. Likewise for Proxy 3, the vulnerable population is itself a significant actor for change and its signal should be emphasized.

Conclusion: For the reason that countries with mainly drylands tend to obtain superior allocation weighting whether or not they have the need for resources to deliver towards sustainable land management, compared to countries with a significant land degradation record, but lower proportion of dryland, the weighting should be adjusted so that Proxies 1 and 2 are treated equally. Secondly, poor rural populations vulnerable to land degradation also have the greatest need and present an opportunity, not just a threat. For these reasons the GEF should consider adjusting the weighting from 20%, 60%, 20% to 33%, 33%, 33%.

8. To what extent are the GEB indices – within the context of performance and socio-economic indices included in the composite country score indices – effective in directing the GEF resources to countries where there is greater realizable potential to generate global environmental benefits?

SUMMARY FINDING: For the LD focal area the GEB indices attempt to direct resources to the appropriate countries equitably; there is no means of testing for effectiveness, hindered by the lack of a link between resource allocation and GEB delivery. For the future, the actual outcome allocations need to be checked for possible anomalies arising from the weightings used in the GBI_{LD} algorithm.

As per the November 2009 Council decision, the STAR is to be implemented from the point of GEF-5 effectiveness. [GEF/C.38/9/Rev.1 – 1 July 2010] – meaning that the LD STAR should use the flexibility and virement rules to ensure maximum effectiveness of the use of resources to deliver GEBs.

Overall objective of an allocation system for the GEF has not changed since it was first introduced through the policy recommendations for the fourth replenishment, as "…a system for allocating resources to countries in a transparent and consistent manner based on global environmental priorities and country capacity, policies and practices relevant to successful implementation of GEF projects" (GEF/C.27/Inf.8/Rev.1, 2005).

Introduction of 'floors and ceilings' [GEF/C.36.6 – 9 October 2009] was based more on pragmatism than on scientific validity. The 'floor' for LD was set at $0.5 m, so that even a small country could develop a reasonably-sized project. The 'ceiling' was set at no more than 10% of funds for any one country. Inevitably, this distorts allocations globally, favoring especially small African countries – so the outcome is probably beneficial in directing resources to the many resource-poor small countries – the very countries that have limited current capability to tackle LD/SLM and which therefore may have least potential to generate GEBs.

In the simulation exercise with total resources for LD of $400 million for testing how the algorithm and additional modifications performed, African countries overall received a total allocation of $169 million (42%), of which $151 million (38%) was allocated to countries in Sub-Saharan Africa. These figures were similar to GEF-4 investments in the LD focal area. All countries with drylands (arid, semi-arid, sub-humid areas) within their national boundary were allocated a total of $348 million (87%). Of this total amount, $226 million (56%) was allocated to countries with more than half of their land area in drylands. These are sensible, pragmatic and defendable outcomes.

Nevertheless, within this bigger picture of allocation some anomalies have been observed, related to the question of weighting discussed under Question 7 above. Comparing similar sized African countries, one comprising almost entirely dryland (GBI_{LD} 1.61) adjacent to another which has a high percentage of humid degraded forest, yet has a low percentage of dryland (GBI_{LD} 0.89), the former attracts almost double the allocation in spite of the likelihood that the latter country can deliver more GEBs. Under a scenario where the weightings were adjusted from 20%, 60%, 20% to 33%, 33%, 33%, the resulting GBI_{LD} for the countries concerned was 1.12 and 1.21 respectively. A similar pattern emerged for other comparable sized countries across the recalculated GBI_{LD} dataset.

Conclusion: there is no real means of testing effectiveness other than to see whether there are any obvious anomalies in allocation under STAR; some have been observed during this assessment. In the end, however, allocations were moderated by creating a threshold floor allocation for some countries and a cap for others. Inclusion of LD-STAR funds in multi-focal projects was also permitted, even encouraged. These deviations from the straight STAR allocation outcomes were sensible in enabling available resources to be used effectively. Especially for small countries, it gave a sum of money that would enable an investment that could deliver GEBs (rather than being a trivial contribution). This enhances 'country-drivenness'.

9. What are the areas where STAR may be improved upon and how?

SUMMARY FINDING: Better integrate LD-GEF investments within existing financial flows to domestic agriculture-related support in order to tie GEB considerations to national actions and to catalyze large scale impact by investing in promising technologies and approaches.

(1) *Link STAR allocation to domestic/national funding of agriculture and agricultural support activities.* It has long been recognized that STAR allocations are insufficient for affected countries to implement the LD-FA in anything other than a piecemeal or pilot fashion. See GEF/C.39.Inf.10 para.7 arguing for set-asides to support enabling activities as well as an 'incentive' set-aside of US$26 million for eligible countries to link nationally-developed projects (presumably with domestic benefits) to GEF-funding for GEBs. This is an aspect that needs to be developed more explicitly, since there is very substantial country-level funding for agriculture, extension, agricultural research, soil conservation and other similar activities through bilateral aid and through the CGIAR, for example. An improvement to STAR would be deliberately to allocate resources according to the level of national funding with domestic benefits, enabling countries to tie GEB considerations to these national on-going activities.

(2) *Ability to link with the very wide range of SLM options, especially those that already provide positive demonstration of effectiveness.* STAR needs to provide flexibility to technologies and approaches that can be applied for large-scale impact [see GEF/C.39.Inf.10 para.7(c)]. It has been suggested above that an additional indicator of 'best practices' may help to address this issue, providing a measure of positive experiences in delivering GEBs through SLM practices and approaches. Flexibility would enhance GEF's catalytic role.

(3) *Effectiveness of STAR index-directed investments should be further analysed.* Currently allocation indices are not used (and are not capable of such use) to monitor outcomes of resource usage. For example, the RAF MTR identified the lack of a link between the indexes that provide the allocations and how the allocations are actually used. A policy dialogue with countries could be initiated to review the use of the current indices and whether a future GBI_{LD} could better enable delivery of GEBs against country priorities.

Conclusions

This report has concluded that the present indices used for the Land Degradation STAR GBI are valid, capable of being updated and constitute a simple yet robust component within the overall STAR allocation indices for directing relatively equitable allocations towards countries to enable them to address dryland opportunities.

The relative weighting of indices needs to be further investigated to address the imbalance that results in relatively large allocations to very dry and desertified countries at the expense of countries having more potential for delivering SLM GEBs.

An additional indicator capturing the degree to which a country has developed and disseminated 'best practices' would enhance alignment with GEF strategies and with the Convention.

The GBI_{LD} alignment was poor with GEF-5 LD focal area priorities, echoing the findings of the RAF MTR that resource allocations are not linked to utilization; for GEF-6 a policy dialogue

with countries could determine whether GBI_{LD} should be adjusted to more explicitly deliver support for country-based initiatives (especially multi-focal approaches) to maximize delivery of GEBs and to increase effectiveness of GEB indices.