

53rd GEF Council Meeting
November 28 – 30, 2017
Washington, D.C.

PROJECT PERFORMANCE AND PROGRESS TO IMPACT

(Prepared by the Independent Evaluation Office of the GEF)

October 2017

TABLE OF CONTENTS

Executive Summary.....	v
I. Methodology.....	1
1. Performance of completed projects.....	1
2. Progress to replenishment targets	2
3. Progress to Impact	3
4. Project Cycle Time Lags.....	4
5. Co-financing Commitments and Materialization.....	4
II. Findings.....	5
1. Performance of Completed Projects.....	5
2. Quality of Terminal Evaluations.....	24
3. GEF Activity Cycle.....	26
4. Progress Towards GEF 5 & 6 Targets	30
5. Progress toward impact	34
6. Mechanisms of Broader Adoption	37
Works Cited.....	38
Annexes.....	39
Annex 1: Rating criteria and scale.....	39
Annex 2. Progress to impact results corrected for differences in sampling probability	44
Annex 3. Progress towards GEF 5 Targets	46
Annex 4: Regression Models.....	51

FIGURES AND TABLES

Figures

Figure 1: Coverage of completed projects by GEF Phase	1
Figure 2: Project with Outcomes Rated in Satisfactory Range - by GEF Period.....	5
Figure 3: Projects with Outcomes Rated in.....	6
Figure 4: Projects with Outcomes Rated in Satisfactory Range-by Country Groups	6
Figure 5: Projects with Outcomes Rated in.....	7
Figure 6: Projects with Outcomes Rated in Satisfactory Range- by GEF Agency.....	7
Figure 7: Projects with Sustainability Rated in Likely Range - by GEF Period.....	9

Figure 8: Projects with Sustainability Rated in Likely.....	9
Figure 9: Projects with Sustainability Rated in Likely Range.....	9
Figure 10: Projects with Sustainability in the Likely.....	11
Figure 11: Projects with Sustainability rated in the Likely Range	11
Figure 12: Projects with Quality of Implementation Rated in Satisfactory Range - by GEF Period.....	12
Figure 13: Projects with implementation rated in.....	13
Figure 14: Projects with Implementation Rated in Satisfactory Range - by select country groups	13
Figure 15: Projects with Implementation Rated in the.....	13
Figure 16: Projects with Implementation Rated in Satisfactory Range-- by GEF Agency	13
<i>Figure 17: Project M&E Design rated in the Satisfactory Range - by GEF Period</i>	<i>15</i>
Figure 18: Percentage of Projects with M&E Design	15
Figure 19: Percentage of Projects with M&E Design in the Satisfactory Range	15
Figure 20: Percentage of Projects with M&E Design	16
Figure 21: Percentage of Projects with M&E Design in the Satisfactory Range	16
Figure 22: Project M&E Implementation Rated in Satisfactory Range - by GEF Period	16
Figure 23: Percentage of Projects with M&E.....	17
Figure 24: Percentage of Projects with M&E Implementation in the Satisfactory Range.....	17
Figure 25: Percentage of Projects with M&E.....	18
Figure 26: Percentage of Projects with M&E Implementation in the Satisfactory Range.....	18
Figure 27: Cofinancing Promised per dollar of GEF Grant - for approved GEF projects.....	19
Figure 28: Promised cofinancing per dollar of GEF.....	20
Figure 29: Promised cofinancing per dollar of GEF Funding: by select country groups	20
Figure 30: Materialized cofinancing for GEF Projects.....	22
Figure 31: Materialized cofinancing for completed projects vis-a-vis promised cofinancing by select country groups.....	22
Figure 32: Materialization of Co-financing by Focal Area.....	23
Figure 33: Materialization of Co-financing by Agency.....	23
Figure 34: Share of different sources of co-financing in commitments and in actual cofinancing - as percentage of the total cofinancing committed/materialized	24
Figure 35: Quality of Terminal Evaluation Reports by Year of TE completion and Project Size	25
Figure 36: Quality of Terminal Evaluations Rated in the Satisfactory Range	26
Figure 37: First PIF Submission to CEO Endorsement - percentage of FSPs endorsed vis-a-vis time taken in months	27
Figure 38: Percentage of PIF Submissions that obtained	28
Figure 39: Percentage of approved FSPs by months needed for PIF approval, GEF-6 year 1 versus year 2	28
Figure 40: Percentage of FSPs that were CEO Endorsed by time taken in months from PIF Approval.....	29
Figure 41: CEO Endorsement / Approval to.....	30
Figure 42: Percentage of Project Completed versus time elapsed in months after expected completion date at project start	30
Figure 43: Expected Adjusted Environmental Results for GEF-5 Projects against the GEF-5 Environmental Results Targets.....	32
Figure 44: Unadjusted and Adjusted Expected Environmental Results as percentage of GEF-6 Targets ..	33

Tables

Table 1: Percentage of projects achieving environmental stress reduction – by scale 34
Table 2: Incidence of broader adoption - by scale..... 36
Table 3: Use of broader adoption mechanisms - percentage of projects where a mechanism was used 37

EXECUTIVE SUMMARY

1. Overall, performance ratings of completed GEF projects show an improvement from GEF-3 to GEF-4 period. While it remains to be seen whether this uptick in ratings is stable, as only 41% of the approved GEF-4 projects have been covered so far, it may be said that the performance of the GEF-4 project is either higher or as high as that of the projects from the preceding periods. The key findings of the analyses are:
2. **The GEF has built a strong record in delivering short- and medium-term outcomes.** Of the 1,173 projects rated on outcomes, 81 percent rated in the satisfactory range. Of the OPS-6 cohort, outcomes of 577 were rated and 79 percent rated in satisfactory range. The ratings underscore the solid track record of GEF projects in delivering expected short to medium term results.
3. There are considerable risks to continuation of the benefits from more than a third of GEF **projects**. Of the 1,118 projects rated on sustainability of outcomes, 62 percent rated in the 'Likely' range. Thus, roughly four out of 10 projects face considerable risks to continuation of their benefits. Of the OPS-6 cohort, 545 were rated for sustainability and of these 63 percent (346 projects) rated in the 'Likely' range.
4. **GEF Agencies generally implement GEF supported projects in a satisfactory manner.** Of the 970 projects rated on quality of implementation, 79 percent were rated in the satisfactory range¹. Of the OPS-6 cohort, 547 were rated for quality of implementation, and of these 79 percent (432 projects) rated in the satisfactory range. Although there is an improving trend across the GEF periods, much of the gains took place during the GEF-1 period.
5. **Despite an improving trend, cumulative ratings on quality of M&E design and implementation remain in the unsatisfactory range for a substantial percentage of projects.** Of the 1108 projects that were rated for quality of M&E design, 61 percent (673 projects) rated in the satisfactory range. Of the OPS-6 cohort, 570 were rated for M&E design and 62 percent (353 projects) of those rated were rated in the satisfactory range. There is a steady trend of improvement in quality of M&E design ratings. This trend is consistent with the findings of the quality at entry review presented in Annual Performance Report 2011, which showed improved compliance with the M&E design expectations. Of the 1012 projects that were rated on quality of M&E plan implementation, 64 percent rated in the satisfactory range. Of the OPS-6 cohort, 546 were rated for M&E implementation and 62 percent (341 projects) were rated in the satisfactory range. There is an improving trend across the replenishment periods in which projects were approved. However, as was the case of M&E design, much of the improvement in ratings for M&E implementation has been achieved from the Pilot Phase to GEF-1.

¹ The terminal evaluation reviews are conducted on an annual basis as part of the work for GEF IEO's Annual Performance Reports. During some of the review cycles quality of implementation was not assessed. Consequently, a relatively higher percentage of completed projects have not been rated on quality of implementation.

- 6. Cofinancing commitments for GEF-6 projects exceed the target set by the GEF's Co-financing Policy (2014).** Against the co-financing policy mandated target of 6:1, co-financing commitments so far for GEF-6 projects have been mobilized at a rate of 8.8:1. Across the GEF periods – from GEF-1 to GEF-6 – there has been a steady increase in the co-financing ratio of the GEF portfolio. In terms of the co-financing ratio of the median full size project, steady increase is evident from the Pilot Phase onwards.
- 7. Promised co-financing successfully materializes during implementation for majority of projects.** Co-financing commitments were fully met for a majority (59 percent) of completed GEF projects. For one out of eight completed projects (13 percent), less than half of the promised co-financing materialized during implementation.
- 8. Most of the terminal evaluations submitted by the GEF Agencies meet the minimum quality expectations.** Of the 1184 terminal evaluations, quality of 83 percent is in the satisfactory range. Of the 581 terminal evaluations received after the close of OPS-5, 571 were rated on quality of terminal evaluation and 82 percent were rated in the satisfactory range.
- 9. Despite some efficiency gains during the GEF-6 period, progress in improving project cycle efficiency has been slow.** Of the 90 full-sized projects for which PIFs were submitted during the first year of GEF-6, 37 percent had been CEO Endorsed through 24 months from their submission. Although this is an improvement over the performance during GEF-5 (26 percent) and GEF-4 (21 percent), the percentage of the GEF-6 PIF submissions that were CEO Endorsed within 24 months of submission is still low. While project cycle for GEF-6 projects was less efficient than GEF-5 between PIF submission to PIF approval, it was more efficient for the PIF approval to CEO endorsement stage. Increase in time taken from PIF submission to PIF approval for GEF-6 projects seems to be driven by the shortfall in GEF-6 replenishment. A fuller picture for the GEF-6 proposals will emerge only after GEF-6 has run its course and sufficient time has elapsed to track progress of the GEF-6 PIFs.
- 10. GEF programming for GEF-5 and GEF-6 is consistent with the corporate environmental results targets for these replenishment periods.** Analysis of data on the targets promised in project proposals of approved GEF-5 and GEF-6 projects allows an assessment of the extent to which programming for these periods is consistent with the corporate environmental results targets for these periods. GEF is projected to exceed targets for 8 of the 13 corporate environmental results indicators for GEF-5 period, although there may be some shortfall for the remaining 5 indicators as level of programming is low for some of the focal area programs. For GEF-6, despite a shortfall in GEF resources, the aggregated results from approved PIFs exceed GEF-6 targets for 6 out of 10 environmental results indicators. When the shortfall is accounted for, expected results are likely to be higher than the targets for seven out of 10 indicators.
- 11. Majority of GEF projects are already contributing to environmental stress reduction and/or environmental status change at implementation completion.** At project completion, 59 percent of the GEF projects from OPS-6 cohort had already led to environmental stress reduction and/or environmental status change. Thirteen percent of the projects were achieving environmental stress reduction and/or status change at a large scale and 45 percent of projects

were achieving it at a local scale. Whether a completed project achieved environmental stress reduction and/or environmental change appears to be linked with the environmental challenge being addressed, country context, global versus regional focus, or the scale of GEF funding.

12. **Approaches and technologies promoted by majority of GEF projects were being adopted by other stakeholders at project completion.** At project completion, 61 percent of completed GEF projects were achieving broader adoption. Country context plays an important role, as a substantially higher percentage of projects implemented in major emerging economies were achieving broader adoption at higher scales than projects in other countries.

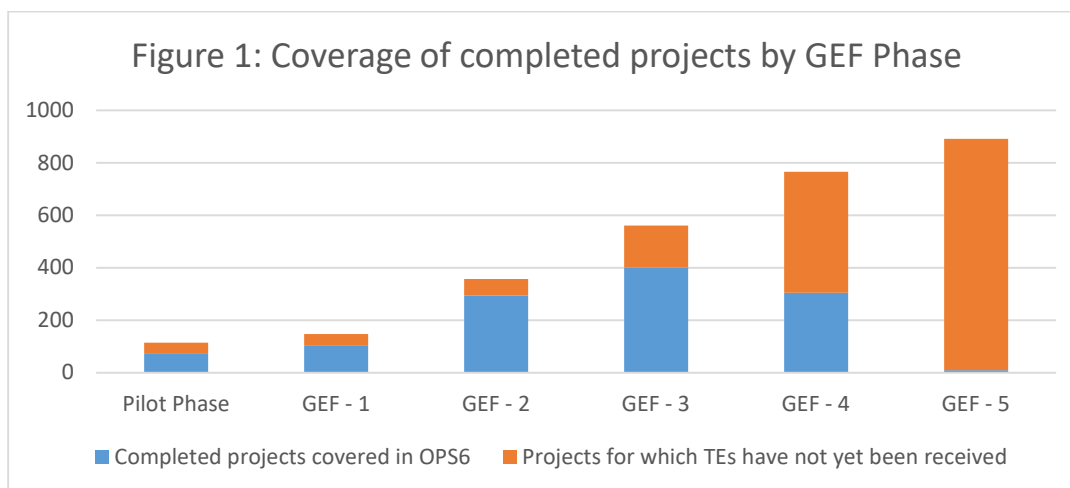
I. METHODOLOGY

1. Performance of completed projects

1. Cumulatively, through December 2016, terminal evaluations for 1184 completed projects have been received by the GEF IEO. These projects account for US \$ 5.4 billion in approved GEF funding and US \$ 22.7 billion in co-financing commitments. The analysis on outcomes, sustainability, implementation, materialization of co-financing, M&E, and quality of terminal evaluation, is based on data provided in these terminal evaluations. Of the 1184 terminal evaluations, 581 were received after close of the Fifth Overall Performance Study (OPS-5). These 581 projects account for US \$ 2.7 billion in GEF funding and US \$ 14.9 billion in co-financing commitments. From here on, these 581 projects are referred to as the OPS-6 cohort.

2. It generally takes a project six to 10 years to move from the PIF approval stage to implementation completion. Consequently, considerable time elapses before terminal evaluations for all or almost all projects approved during a replenishment period become available. Based on a comparison of the number of projects that were approved during a given replenishment period - excluding cancelled projects and projects for which terminal evaluations are not expected because of the small-scale of GEF funding (i.e. below \$ 0.5 million) – and the terminal evaluations received so far, the coverage of completed projects up to GEF-3 period is robust. For GEF-4, although 304 terminal evaluations are available these are only 41 percent of the projects (738 projects) from the period for which terminal evaluations are expected. As most GEF-5 projects are still under implementation, on nine terminal evaluations are available for projects from this period. It is too early for majority of GEF-6 projects to be under implementation let alone their being completed. When discussing the results of the completed projects all 1184 projects are covered. However, when data is presented based on replenishment periods, only data up to GEF-4 is presented.

Figure 1: Coverage of completed projects by GEF Phase



3. Sizable gap in coverage of projects from the GEF-4 replenishment period has methodological implications (Figure 1). Comparison of outcome ratings for the dataset used for OPS-6 analysis with those for OPS-5 analysis shows that the projects for which terminal evaluations are received after greater time lag tend to have lower outcome ratings than projects for which these evaluations are received earlier. For example, from OPS-5 to OPS-6 percentage of projects with outcomes in the satisfactory range declined from 81 percent (n=228 projects) to 79 percent (n=289 projects) for GEF-2 projects, and from 88 percent (n=176 projects) to 80 percent (n=399 projects) for GEF-3 projects. Some of this is related to size of the projects. A slightly higher percentage of medium sized projects are rated in the satisfactory range compared to full sized projects (84 percent versus 79 percent). Since medium sized projects also tend to have shorter duration, their terminal evaluations are usually received before the evaluations for full size projects. However, much of the decline in outcome ratings is driven by late receipt of terminal evaluations for projects that experience difficulties during project start up and implementation. Thus, it may be expected that as terminal evaluations for a higher percentage of GEF-4 projects become available the percentage of GEF-4 projects with outcomes rated in the satisfactory range may reduce.

4. Details on criteria used by the GEF IEO to assess outcomes, sustainability, implementation and quality of terminal evaluations are provided in the Guidelines on the Project and Program Cycle Policy (2017, GEF/C.52/Inf.06) and are also listed in Annex 1 of this paper. Independent evaluation offices of some the GEF Agencies such as the World Bank, UNDP, UNEP and IFAD, also provide performance ratings using criteria that is broadly consistent with that used by the GEF IEO. To avoid duplication of effort, and to encourage the independent evaluation offices of the Agencies to play a greater role in validation of terminal evaluations, beginning in 2009 with World Bank and UNEP, GEF IEO has been accepting the ratings provided by their evaluation offices. Of the 1184 projects, for 42 percent (501 projects) ratings provided by the Agency evaluation offices have been used. For the remainder, ratings provided by the GEF IEO have been used. For quality control, 218 terminal evaluations that had been validated by the Agency evaluation offices have also been validated by the GEF IEO. Analysis of the ratings by GEF IEO and Agency evaluation offices shows that on the net Agency evaluation offices tend to rate outcomes of 1.4 percent more projects in the satisfactory range than the GEF IEO: i.e. 85.8 percent versus 84.4 percent. This difference between the ratings from the two sources is not substantial, therefore, for projects covered from 2009 onwards, where available, ratings provided by the Agency evaluation offices have been used.

2. Progress to replenishment targets

5. The GEF-5 replenishment process established environmental results targets for the activities funded through the GEF-5 replenishment resources. To inform the process for the Sixth Replenishment of the GEF Trust Fund, GEF IEO took stock of progress towards the GEF-5 replenishment targets. The first analysis was prepared the December 2013 meeting of GEF-6 replenishment, and an updated analysis was prepared for its April 2014 meeting. These

analyses were primarily based on aggregation of the targets provided in the Project Identification Form (PIFs) of the approved project proposals. Most projects that were approved during the period (i.e. the GEF-5) projects are yet to be complete or have crossed the mid-term review milestone. Therefore, there is still little information on the results achieved on ground. However, 96 percent of the projects approved during GEF-5 have now been CEO Endorsed / Approved. Therefore, more detailed projections on expected results of the GEF-5 projects are now available. The documents for 686 projects funded partially or fully through GEF Trust Fund resources under GEF-5 submitted at CEO Endorsement / Approval for GEF-5 projects were reviewed to update the analysis for the projects approved in GEF-5. Enabling activities were excluded from the analysis as these are not expected to directly result into environmental results. To arrive at the projections for GEF-5 period, the aggregate of expected project results has been multiplied by a factor of 0.8 to account for cancellations and implementation failures.

6. For GEF-6 period data from GEF-6 PIF approvals maintained by the GEF Secretariat and presented in the GEF Corporate Scorecard has been used as basis for reporting. The projections for GEF-6 have been arrived at by adjusting for likely cancellations and implementation failures (multiplied by a factor of 0.8) and for the level of GEF-6 resources used so far vis-à-vis expected GEF-6 replenishment.

3. Progress to Impact

7. In preparation for OPS-6, assessment of progress to impact was mainstreamed in the reviews undertaken for Annual Performance Report 2015 and 2016. New terminal evaluations received for APR2015 and APR2016 were fully covered for assessment of progress of impact. For the remaining terminal evaluations of the OPS-6 cohort that were submitted after close of OPS-5 (i.e. submissions for APR2013 and APR2014), a representative sample of 50 percent was sampled. Thus, in all 426 completed projects were covered. After initial screening 11 targeted research and/or foundational activities – that are not expected to directly lead to environmental stress reduction and/or environmental status change – were removed from the analysis. Thus, in all progress to impact of 415 completed GEF projects was analyzed. The reviews to assess progress to impact were conducted using an instrument which, along with incidence of environmental stress reduction and/or environmental status change, and broader adoption, also recorded the design features and implementation experience of the reviewed project. The results were analyzed assigning probability weights so the results for the OPS-6 cohort are not skewed by the submissions for APR2015 and APR2016, which had 100 percent probability of being represented in the sample (compared to 50 percent for APR2013 and APR2014). The calculations were also made without correcting for the differences in probability of being sampled. There is not much material difference in the calculations using the two approaches. In this paper results that are not corrected for difference in sampling probability are presented in the main narrative. Probability adjusted figures are presented in the Annex 2.

8. Although progress to impact related analysis was also presented in OPS-5. The methodology for the OPS-6 assessment is different from that used for the OPS-5 analysis. Compared to OPS-5, the approach for OPS-6 uses a higher threshold for recording incidence of environmental stress reduction and/or environmental status change. Consequently, findings of the analysis undertaken for OPS-6 are not directly comparable to those presented in OPS-5.

4. Project Cycle Time Lags

9. Analysis on project cycle time lags focuses on stand-alone full sized projects. Stand-alone full size projects are endorsed by the GEF CEO following a two-step process. The first step involves submission of a PIF (Project Information Form) by an Agency and this step culminates in PIF Approval. The second step involves preparation of a detailed project proposal by the GEF Agency, submission of the proposal to the GEF Secretariat, and CEO Endorsement of the proposal. GEF has established an 18-month standard for full sized projects from its PIF Approval to CEO Endorsement (GEF/C.38/5/Rev.1). Project cycle time lags analysis focuses on the stages between PIF submission and project start. Time lag involved in project completion and reporting of data on project implementation and completion makes it difficult to assess time lags for these steps of the project cycle for projects that were recently completed. PMIS data through June 2017 has been used to determine the time lags.

10. Although it is important to cover medium sized projects and activities under programmatic approach framework, it was not feasible. Medium size projects were excluded because Council's approval of the single-step CEO Approval process for medium-sized projects, makes it difficult to assess the time taken in preparation of the MSP proposals that follow the single step process. Although, it is still possible to measure time lags for those that follow the two-step process, the number of observations are too low.

11. The number of child projects developed during GEF-6 under programmatic approach is still too small to allow a meaningful analysis. This constraint is accentuated because child projects developed within the programmatic approach are expected to meet their negotiated program commitment deadline given in their respective Program Framework Document (PFD) and not the 18-month standard applicable to stand alone full-sized projects.

5. Co-financing Commitments and Materialization

12. Analysis on co-financing trends draws from data from different sources. Analysis on trends in co-financing commitments is based on the PMIS data through June 2017. This data has been used to calculate co-financing ratio of the GEF project portfolio and median project co-financing ratio for different project types across GEF replenishment periods.

13. Analysis on materialization of co-financing is based on data provided in the 1184 terminal evaluations that have been submitted to the GEF IEO through December 2016. Of these, data on materialization of co-financing is available for 84 percent (994 projects).

14. Analysis on probability of materialization of co-financing commitments for different sources of co-financing is based on the survey of information provided in the project documents and in terminal evaluation reports. Data on sources of co-financing commitments and its materialized co-financing are from 323 projects from OPS-6 cohort for which this information was available.

II. FINDINGS

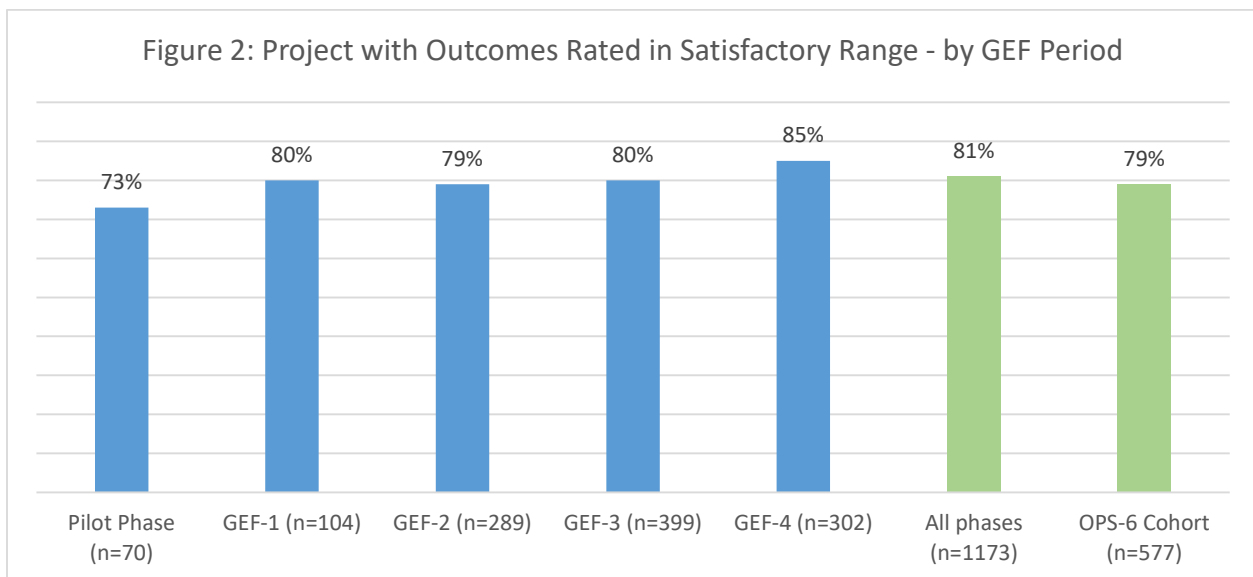
1. Performance of Completed Projects

Outcomes

15. OECD defines outcomes as ‘the likely or achieved short-term and medium-term effects of an intervention’s outputs’ (2002).² A GEF project is expected to deliver its expected outcomes by the end of its implementation. Terminal evaluations prepared by the GEF Agencies provide a record of the extent to which expected outcomes were delivered. Findings of these evaluations are then validated by the GEF IEO and/or the independent evaluation offices of the GEF Agencies. A six-point scale is used for rating level of project outcome achievements. Of these, the top three ratings comprise the ‘satisfactory range’ and the bottom three the ‘unsatisfactory range’³.

16. Of the 1184 completed GEF projects for which terminal evaluations have been submitted to the GEF IEO so far, 1173 have been rated for their outcome achievements.⁴ Of those rated, 81 percent rated in the satisfactory range. Of the 581 terminal evaluations that were received after close of OPS-5, outcomes of 577 were rated and 79 percent rated in satisfactory range (Figure 2). The ratings underscore the solid track record of GEF projects in delivering expected short to medium term results. Comparison across periods shows that most GEF projects continue to deliver their expected outcomes.

Figure 2: Project with Outcomes Rated in Satisfactory Range - by GEF Period



² *Glossary of Key Terms in Evaluation and Results Based Management*, OECD, 2002.

³ The ratings are: Highly Satisfactory, Satisfactory, and Moderately Satisfactory (all included in ‘satisfactory range’); and, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory (all included in ‘unsatisfactory range’).

⁴ For remainder the rating was not provided due to insufficient information provided in the terminal evaluations.

Of 304 GEF-4 projects for which terminal evaluations were submitted, outcomes of 302 projects were rated. Of those rated, 85 percent were rated in the satisfactory range. The *Policy Recommendations for the GEF-4 Replenishment* set a target of outcome ratings of 75 percent of projects in the satisfactory range for the projects approved during this period⁵. The GEF-4 projects have so far exceeded this expectation and are on track to meet the GEF-4 replenishment target.

Figure 3: Projects with Outcomes Rated in Satisfactory Range- by Region

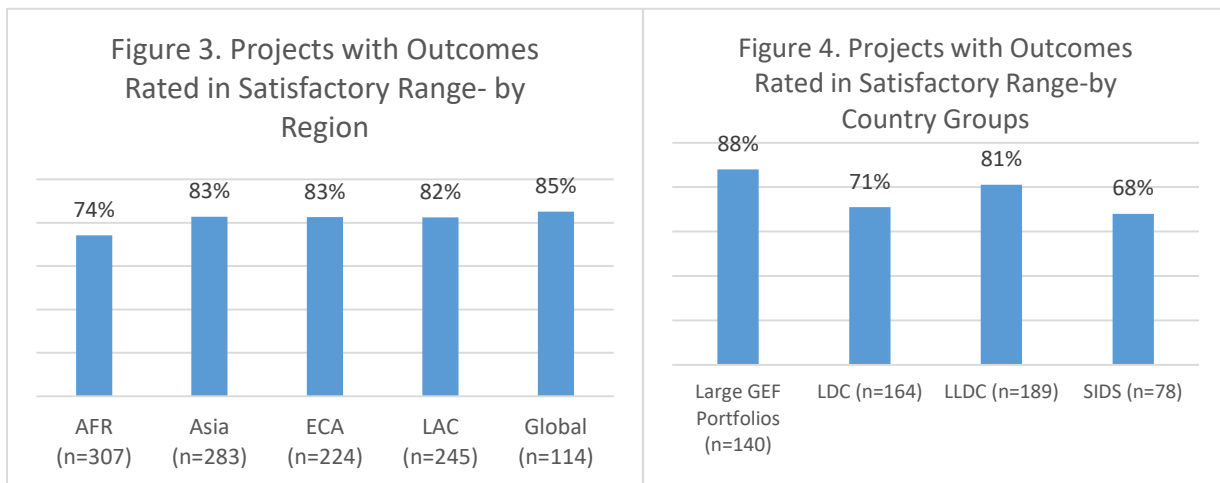


Figure 4: Projects with Outcomes Rated in Satisfactory Range-by Country Groups

17. Of the projects implemented in Africa, 74 percent were rated in the satisfactory range (Figure 3). This is significantly lower than 83 percent of projects in other regions, including global projects, that were rated in the satisfactory range. However, there is considerable difference in performance across African countries. While outcomes of 90 percent of the projects (n=29) implemented in North African countries were rated in the satisfactory range, those for 69 percent projects (n=74) implemented in East African countries and 62 percent of projects (n=26) in the West Sub Saharan countries were rated in the satisfactory range⁶.

18. Among select country groups where project performance was tracked, outcomes of a higher percentage of projects implemented in China, Brazil, India, Mexico and the Russian Federation, which account for the five largest country portfolios by GEF funding, (from here referred to as the countries with 'large GEF portfolios'), had outcomes in the satisfactory range.

⁵ In *Summary of Negotiations on the Fourth Replenishment of the GEF Trust Fund* (GEF/C.29/3).

⁶ The north African countries are Algeria, Egypt, Libya, Morocco, Tunisia; the East African countries are Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Mauritius, Rwanda, Seychelles, Somalia, South Sudan, Sudan, Tanzania, and Uganda; the West Sub Saharan countries are Benin, Cote d'Ivoire, Ghana, Guinea, Liberia, Nigeria, Sierra Leone, and Togo. These sub-regions correspond to a GEF constituency of member countries, and are each represented in the GEF Council.

Outcomes of projects in LDCs and SIDS were less likely to be rated in the satisfactory range (Figure 4).

19. For GEF focal areas, percentage of projects rated in the satisfactory range for outcomes ranges from 75 percent to 84 percent (Figure 5). Outcomes of 75 percent of the international waters focal area projects were rated in the satisfactory range, which is lower than other GEF projects at 90 to 95 percent confidence level depending on the model used. The difference between other focal areas and remaining projects is not statistically significant.

Figure 5: Projects with Outcomes Rated in Satisfactory Range - by Focal Area

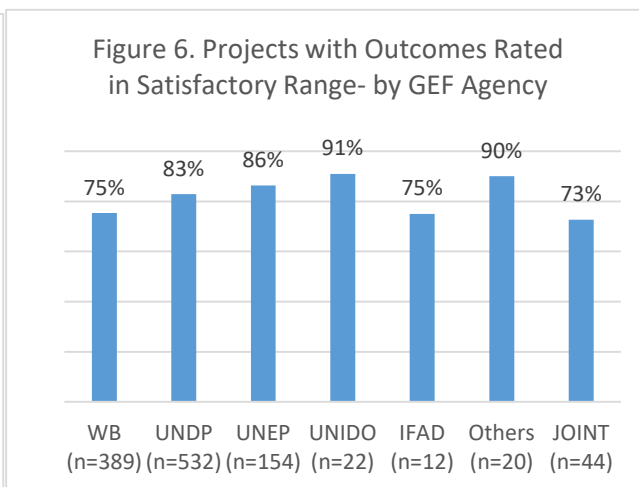
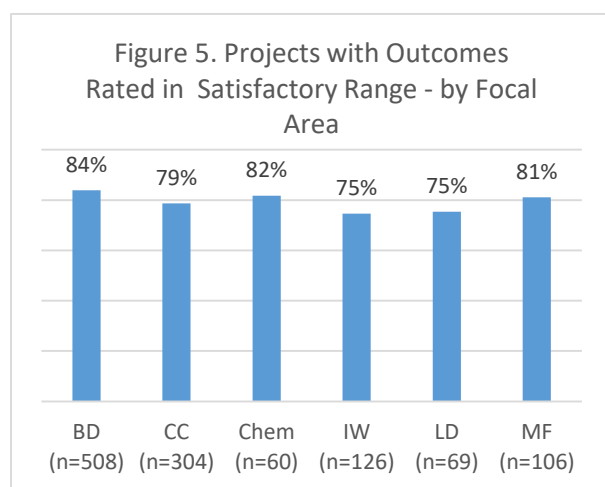


Figure 6: Projects with Outcomes Rated in Satisfactory Range- by GEF Agency

20. Compared to other GEF Agencies, outcomes of a higher percentage of projects implemented by UNEP were rated in the satisfactory range (Figure 6).⁷ On the other hand, a lower percentage of projects implemented by World Bank were rated in the satisfactory range. Difference in performance for other GEF Agencies is not statistically significant.

21. Level of outcome achievement may be determined by several factors. While some of them are tracked by the GEF, it is difficult to measure the extent to which they determine outcomes. Multiple linear regression model based analyses suggest that quality of implementation, quality of execution, and shortfall in materialization of co-financing, are among the key determinants of outcome ratings. Quality of implementation and quality of execution ratings positively affect outcome ratings. Materialization of less than 50 percent of promised co-financing negatively affects outcome ratings as several planned activities are

⁷ Difference is significant at a 90% level

dropped or scaled down. A statistically higher percentage of projects implemented in large GEF portfolios, and a significantly lower percentage of projects implemented in Africa, LDCs and SIDS, are rated in the satisfactory range for outcomes. However, when other variables such as quality of implementation, quality of execution, quality of M&E design, and materialization of co-financing, are controlled for, the relationship between whether a project was implemented in Africa or large economies and outcome ratings weakens and is not statistically significant. This shows that better outcome achievements may be achieved if implementing agencies accord greater attention to project preparation and to project implementation in Africa.

22. A review presented in APR 2014 analyzed the lessons reported in the terminal evaluations of 603 randomly selected⁸ completed GEF projects (APR 2014). The review identified several reasons that lead to lower level of results achievements. These include: overly ambitious objectives, inadequate budget for planned activities, weak intervention strategy, inadequate arrangements to facilitate follow up, inappropriate institutional arrangements, inadequate government and stakeholder support, poor M&E design, etc.

Sustainability

23. Consistent with the OECD (2002)⁹ definition of sustainability, the GEF M&E Policy (2002) defines sustainability as “*the likely ability of an intervention to continue to deliver benefits for an extended period of time after completion.*”¹⁰ GEF IEO rates sustainability on a four-point rating scale based on an assessment of the level of risk to continuation of project benefits at the point of project completion. It takes financial, sociopolitical, institutional & governance, and environment risks into account. The top two rating comprise the “likely” range and the bottom two the “unlikely” range.¹¹

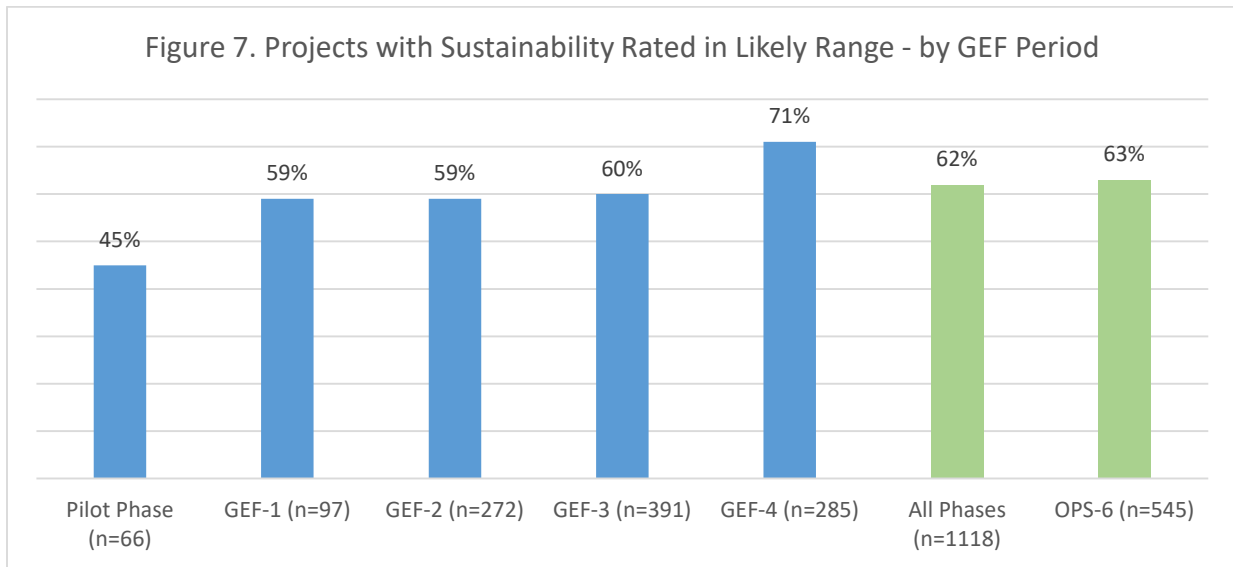
⁸ These were randomly selected from the pool of terminal evaluations that were available through December 2014. The analysis does not take into account the terminal evaluations received during 2015 and 2016.

⁹ *Glossary of Key Terms in Evaluation and Results Based Management*, OECD, 2002.

¹⁰ GEF Monitoring and Evaluation Policy, GEF IEO, 2010.

¹¹ The four-point rating scale used to rate sustainability is as follows: Likely, Moderately Likely (both ‘Likely’ range); and, Moderately Unlikely, Unlikely (both ‘Unlikely’ range).

Figure 7: Projects with Sustainability Rated in Likely Range - by GEF Period



24. Of the 1184 completed GEF projects for which terminal evaluations are available, 1118 have been rated on ‘sustainability’. Of the rated projects, 62 percent (689 projects) were rated in the ‘Likely’ range (Figure 7). This shows that roughly four out of 10 projects face considerable risks to continuation of their benefits. Of the 581 terminal evaluations that were received after close of OPS-5, 545 were rated for sustainability and of these 63 percent (346 projects) rated in the ‘Likely’ range. The trend across the GEF replenishment periods shows improvement in the sustainability ratings, although the figures for GEF-4 may regress towards the long-term average as more terminal evaluations of the GEF-4 projects become available.

Figure 8: Projects with Sustainability Rated in Likely Range- by GEF Region

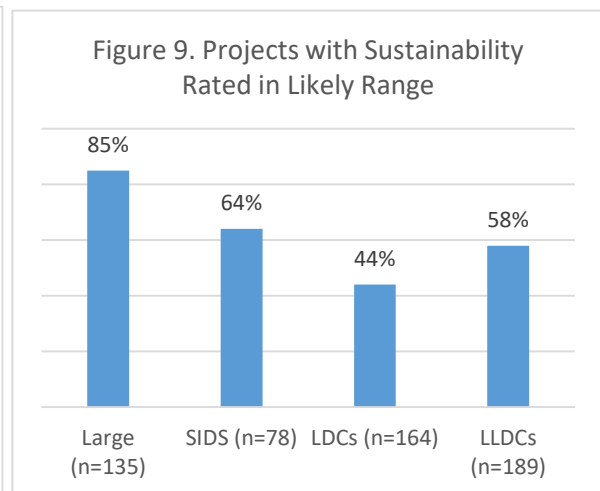
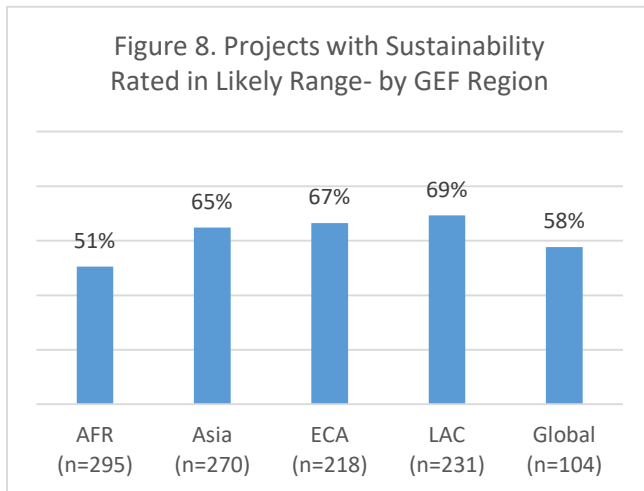


Figure 9: Projects with Sustainability Rated in Likely Range

25. Among the regions, a significantly lower percentage of projects in Africa rated in the 'Likely' range for sustainability (Figure 8). Within Africa too there is considerable variation in performance. While sustainability of 64 percent (n=28) of projects in north Africa¹² is rated in the 'likely' range, only 35 percent of projects (n=75) in west sub Saharan countries and east Africa are rated in the likely range¹³.

26. Among other select country groups, 85 percent of projects (n=135) in large GEF portfolio countries were rated in the likely range for sustainability (Figure 9). In comparison, 44 percent of projects in LDCs (n=154) and 55 percent of projects in SIDS (n=72), i.e. countries where there are considerable capacity and resource constraints, were rated in the 'Likely' range for sustainability.

27. Much of the sustainability related constraints are experienced in least developed countries, where financial resources and institutional capacities to ensure continuity may be limited. In comparison, in large emerging economies that account for top five GEF project portfolios outcomes of 85 percent of projects are rated in the 'likely' range.

Compared to projects from other focal areas, sustainability of a higher percentage of Climate Change projects (69 percent) was rated in the Likely range (Figure 10). The sustainability ratings of other focal areas are not statistically different from others. The sustainability ratings by GEF Agency does not show much difference among major Agencies (Figure 11). However, when the projects of Agencies with smaller portfolio of completed projects are pooled together (Others), a higher percentage of this pool of projects have sustainability ratings in the likely range. The relationship weakens and is not significant in several models when other variables are controlled for.

¹² This includes countries from the following GEF constituency: Algeria, Egypt, Libya, Morocco, Tunisia

¹³ This includes three GEF constituencies that consist of following countries: Benin, Cote d'Ivoire, Ghana, Guinea, Liberia, Nigeria, Sierra Leone, Togo; Burundi, Cameroon, Central African Republic, Congo, Congo DR, Equatorial Guinea, Gabon, Sao Tome and Principe; and, Burkina Faso, Cabo Verde, Chad, Guinea-Bissau, Mali, Mauritania, Niger, Senegal, The Gambia.

Figure 10: Projects with Sustainability in the Likely Range - by Focal Area

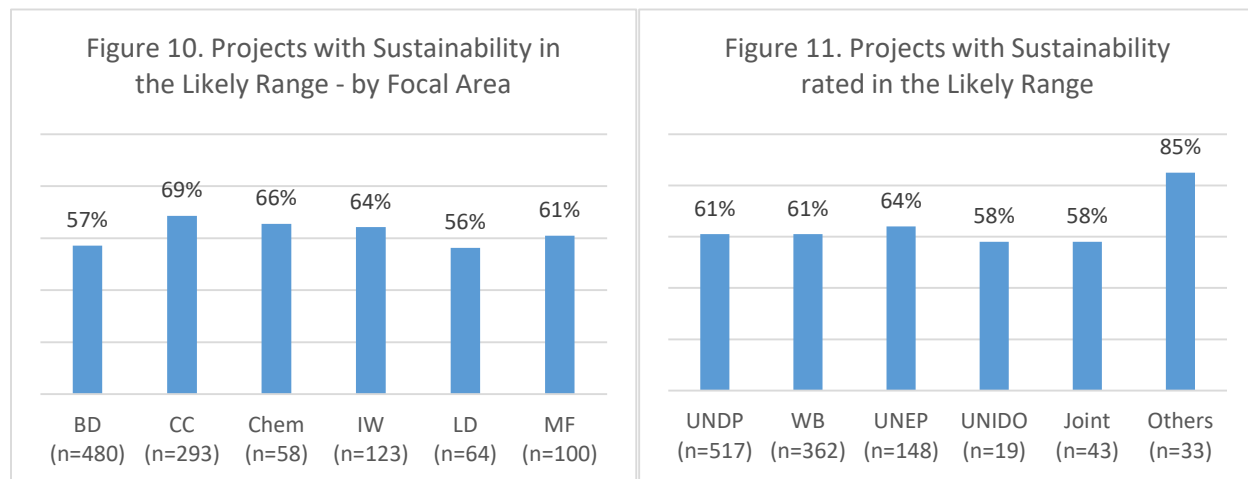


Figure 11: Projects with Sustainability rated in the Likely Range

28. Multiple linear regression based analysis shows that country context, quality of implementation and quality of execution influence project sustainability ratings. While both quality of implementation and execution have statistically significant effect on sustainability, quality of execution – which reflects capacities of the local partners – has greater coefficients and is less sensitive to changes in the regression model used for analysis.

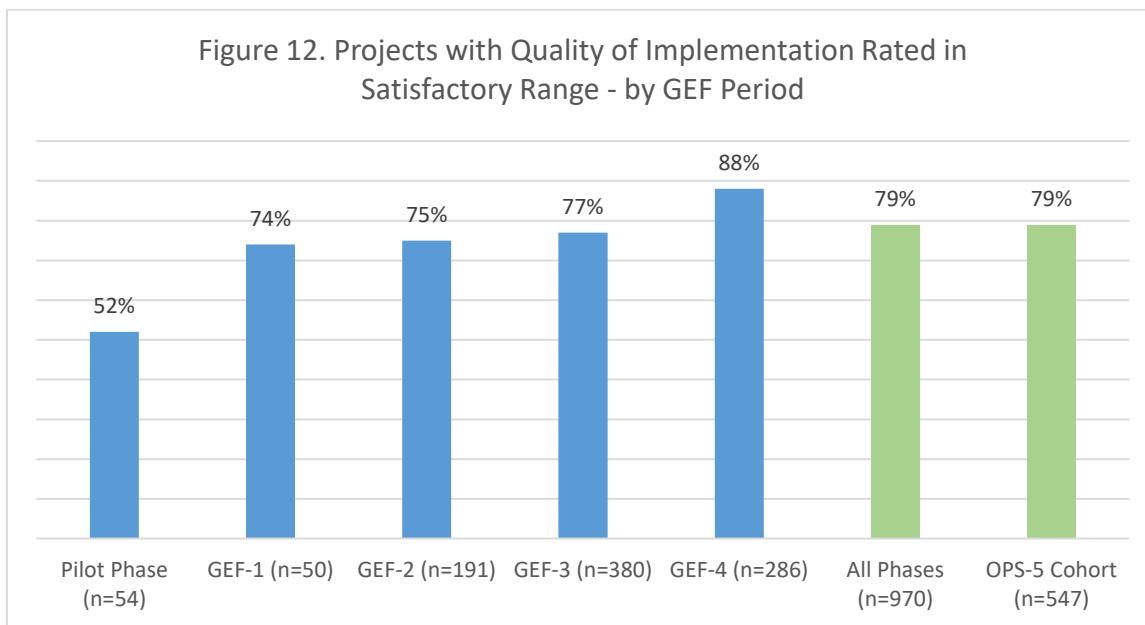
Quality of Implementation

29. Within the GEF Partnership, GEF Agencies are responsible for implementation of the projects and programs funded by the GEF. As part of their implementation related responsibilities, GEF Agencies are involved in project’s identification, concept preparation, appraisal, preparation of detailed proposal, approval and start-up, oversight, supervision, completion, and evaluation. GEF IEO assesses the extent to which a GEF Agency performed well in its role by reviewing the information provided in terminal evaluations and project implementation reports. In assessing implementation quality focus is on elements that are controllable for a given Agency along with the extent to which it identified and managed the risks well. GEF IEO uses a six-point scale to rate quality of project implementation. Of these, the top three ratings comprise the ‘satisfactory range’ and the bottom three the ‘unsatisfactory range’¹⁴.

¹⁴ The ratings are: Highly Satisfactory, Satisfactory, and Moderately Satisfactory (all included in ‘satisfactory range’); and, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory (all included in ‘unsatisfactory range’).

30. Of the 970 completed projects that were rated on quality of implementation, 79 percent (762 projects) were rated in the satisfactory range (Figure 12).¹⁵ Although there is an improving trend across the GEF periods, much of the gains took place during the GEF-1 period. Performance for GEF-5 period and GEF-4 is likely to regress closer to the long-term average when more projects from these periods are completed. Of the 581 project for which terminal evaluations were received after close of OPS5, 547 were rated for quality of implementation, and of these 79 percent (432 projects) rated in the satisfactory range.

Figure 12: Projects with Quality of Implementation Rated in Satisfactory Range - by GEF Period



31. A lower percentage of projects implemented in Africa, in SIDS, and in LDCs, are rated in the satisfactory range for quality of implementation (Figure 13 and Figure 14). The difference is statistically significant even when variables such as focal area and GEF Agency are controlled for. This suggests that GEF Agency capacities in these regions and country groups may be relatively weaker than in other regions and country groups.

¹⁵ The terminal evaluation reviews are conducted on an annual basis as part of the work for GEF IEO's Annual Performance Reports. During some of the review cycles quality of implementation was not assessed as a result of which 214 have not been rated on quality of implementation.

Figure 13: Projects with implementation rated in Satisfactory Range - by Region

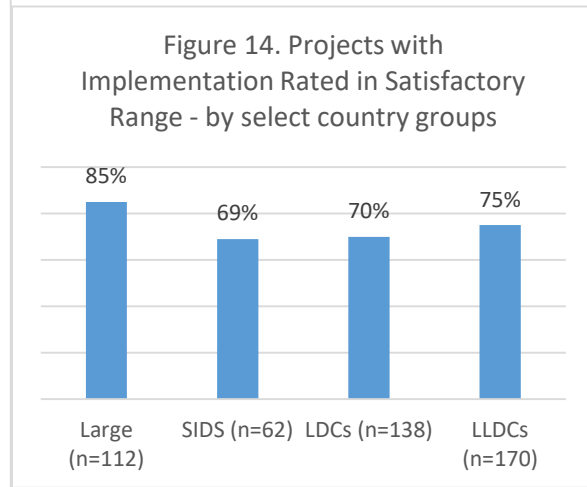
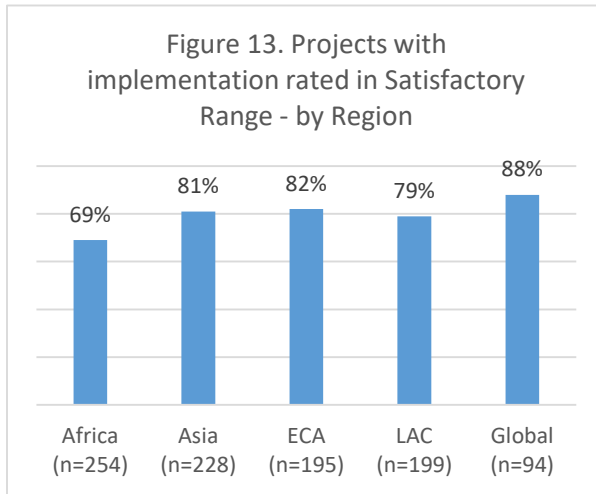


Figure 14: Projects with Implementation Rated in Satisfactory Range - by select country groups

32. Quality of implementation ratings of projects by focal area are closely bunched together (Figure 15). Seventy-three to 80 percent of projects are rated in the satisfactory range and differences across the focal areas are not statistically significant.

Figure 15: Projects with Implementation Rated in the Satisfactory Range - by Focal Area

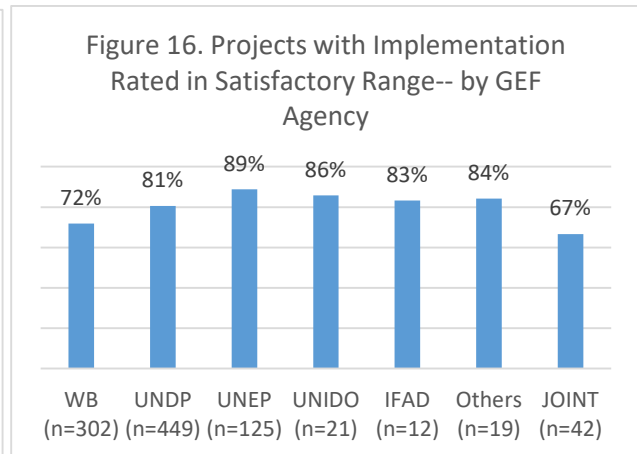
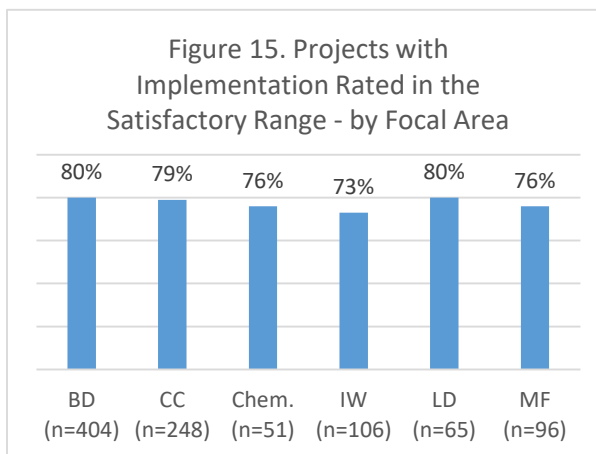


Figure 16: Projects with Implementation Rated in Satisfactory Range-- by GEF Agency

33. Among the GEF Agencies, a higher percentage of UNEP implemented projects were rated in the satisfactory range for quality of implementation, whereas a lower percentage of World Bank projects and jointly implemented projects were so rated (Figure 16). Lower ratings for World Bank implemented projects are driven by low ratings for the projects from the GEF-3 period: only 65 percent of the World Bank implemented projects from this period (n=126) were

rated in the satisfactory range. As explained in APR2013 and 2014 some of this drop may be due to stringent application of the rating criteria by the World Bank's Independent Evaluation Group for the project from this period¹⁶. Information from the online survey reported in Evaluation of Expansion of GEF Partnership (GEF IEO, 2017) and from quality of supervision reviews presented in APR2006 and APR2009 indicates that World Bank performs well in project implementation. Ratings for UNDP are close to the portfolio average. Within the UNDP portfolio, there was a substantial improvement from Pilot Phase (26%, n=23) to GEF-1 and beyond. For other Agencies and jointly implemented projects the observations are too few to draw inferences.

34. The analysis of lessons presented in APR2014 showed that quality of implementation may be poor because of inadequate oversight and technical support, inability to take corrective measures in a timely manner, high staff turnover, ineffective project governance structures, etc. GEF Agencies need to mitigate the gap in implementation services for regions and country groups with capacity constraints.

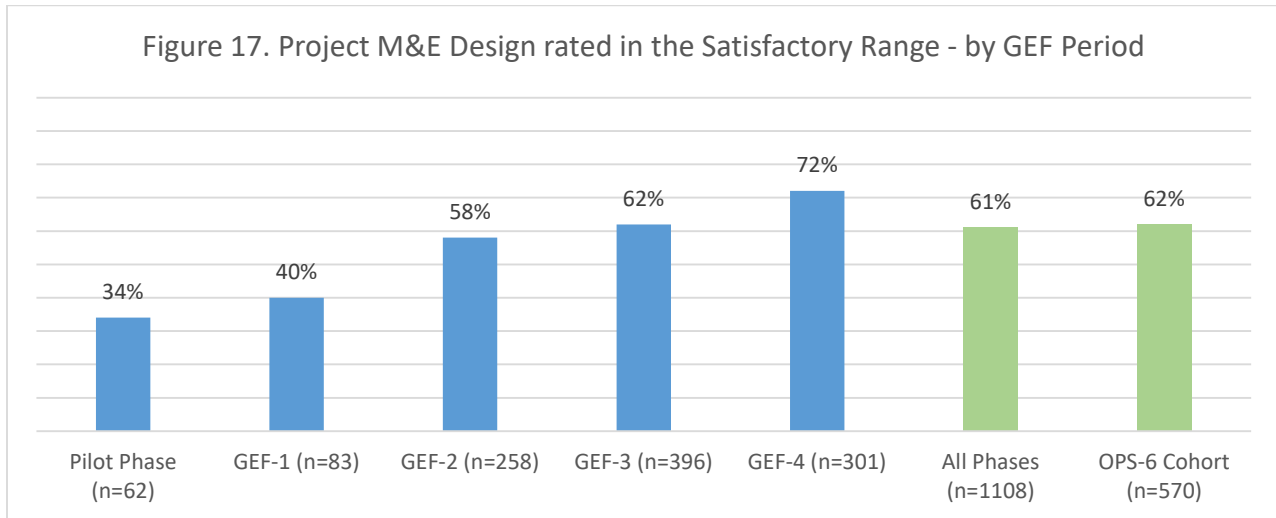
Project M&E Design and Implementation

35. The *GEF Monitoring and Evaluation Policy's* (GEF IEO 2006, 2010) 'minimum requirement 1' calls for fully developed and budgeted project M&E plan at CEO Endorsement. Its 'minimum requirement 2' calls for effective implementation of these plans. Tracking quality of M&E in GEF projects is important as GEF's ability to assess its results on ground and foster learning across the GEF Partnership depends on how well project M&E is designed and implemented. The GEF IEO rates quality of M&E design based on the information provided in the project documents submitted for CEO Endorsement (or Approval), whereas rating on quality of M&E plan implementation is provided based on the review of project implementation reports (PIRs), tracking tools, and information provided in the terminal evaluation. A six-point scale is used to rate quality of project M&E design and of M&E plan implementation. Of these, the top three ratings comprise the 'satisfactory range' and the bottom three the 'unsatisfactory range'¹⁷.

¹⁶ World Bank IEG appears to have applied more stringent criteria during validations that it conducted from 2009 to 2011 period. Since GEF IEO accepts ratings provided by the World Bank IEG, there was a drop in performance ratings for the projects from the GEF-3 period.

¹⁷ The ratings are: Highly Satisfactory, Satisfactory, and Moderately Satisfactory (all included in 'satisfactory range'); and, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory (all included in 'unsatisfactory range').

Figure 17: Project M&E Design rated in the Satisfactory Range - by GEF Period



36. Of the 1108 projects that were rated for quality of M&E design, 61 percent (673 projects) rated in the satisfactory range (Figure 17). There is steady improvement in quality of M&E design ratings across replenishment periods. This trend is consistent with the findings of the quality at entry review presented in Annual Performance Report 2011, which showed improved compliance with the M&E design expectations. Of the 570 projects of the OPS-6 cohort that were rated for quality of M&E design, 62 percent were rated in the satisfactory range.

Figure 18: Percentage of Projects with M&E Design in the Satisfactory Range

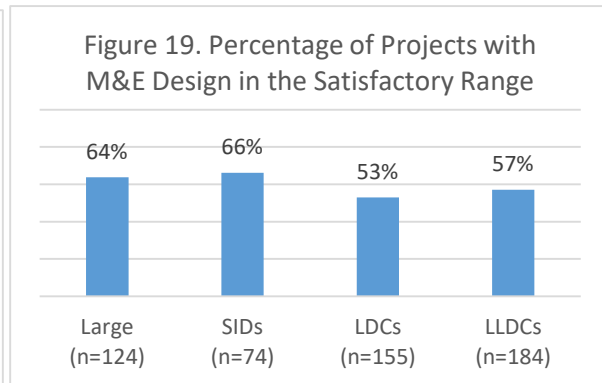
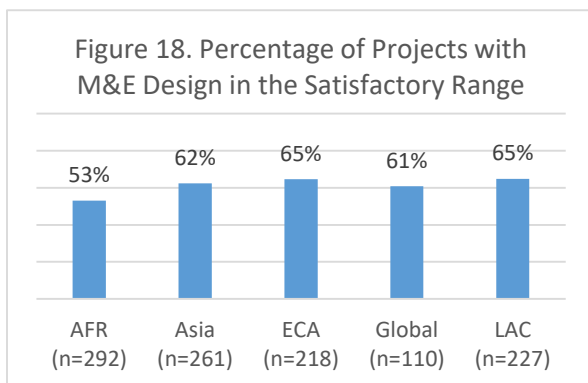


Figure 19: Percentage of Projects with M&E Design in the Satisfactory Range

37. A lower percentage of projects in Africa were rated in the satisfactory range for M&E design than projects in other regions (Figure 18). This difference stays even when other variables are controlled for. A lower percentage of projects in LDCs were rated in the satisfactory range (Figure 19). However, when other variables are controlled for, the difference in M&E design ratings of projects in LDCs and those in other countries is not significant.

Figure 20: Percentage of Projects with M&E Design in the Satisfactory Range

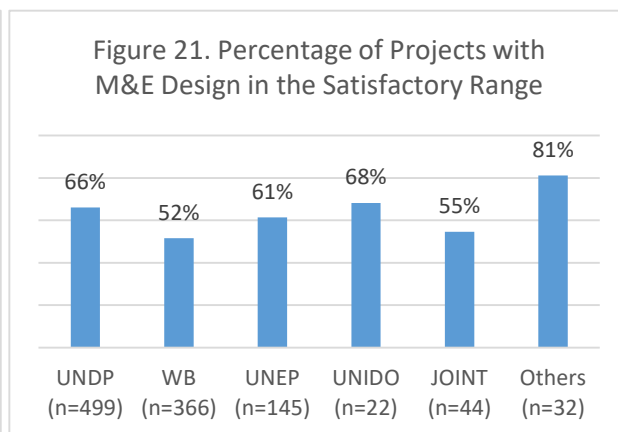
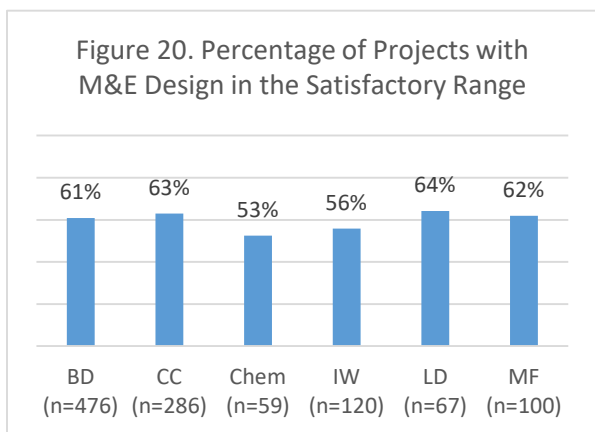
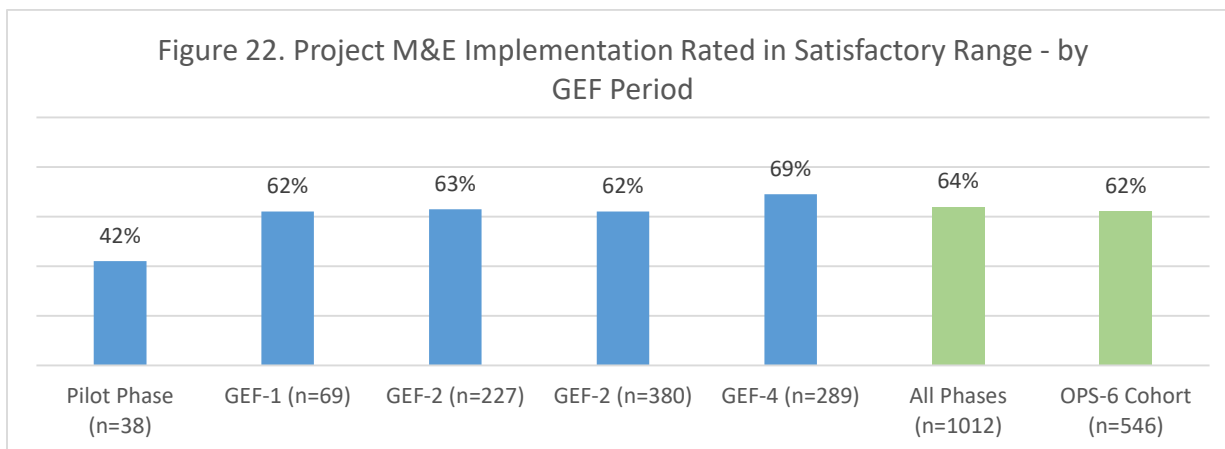


Figure 21: Percentage of Projects with M&E Design in the Satisfactory Range

38. A lower percentage of projects from the Chemicals focal area are rated in the satisfactory range for M&E design compared to projects from other focal areas (Figure 20). This difference stays significant when other variables are controlled for. Among the GEF Agencies, projects implemented by the World Bank tend to have lower M&E design ratings compared to those implemented by other Agencies and the difference is significant when other variables are controlled for (Figure 21).

Figure 22: Project M&E Implementation Rated in Satisfactory Range - by GEF Period



39. Of the 1012 projects that were rated on quality of M&E plan implementation, 64 percent rated in the satisfactory range (Figure 22). Much of the improvement in ratings was achieved from the Pilot Phase to GEF-1. Of the 546 projects from the OPS-6 cohort, 62 percent rated in the satisfactory range. After approval of the GEF Monitoring and Evaluation Policy in 2006 and its revision in 2010, there have been enhanced expectations on project M&E, especially inclusion of minimum standard 4 in the 2010 M&E policy, which calls for engagement of GEF Operational Focal Points in M&E activities and GEF wide adoption of tracking tools from GEF-4 onwards. This may mask the level of improvements in the quality of project M&E during the more recent periods.

Figure 23: Percentage of Projects with M&E Implementation in the Satisfactory Range

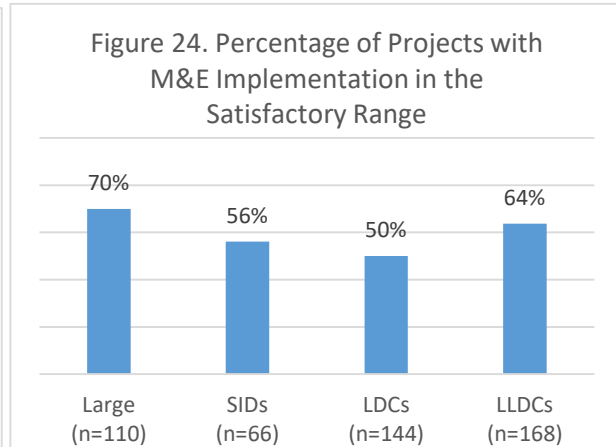
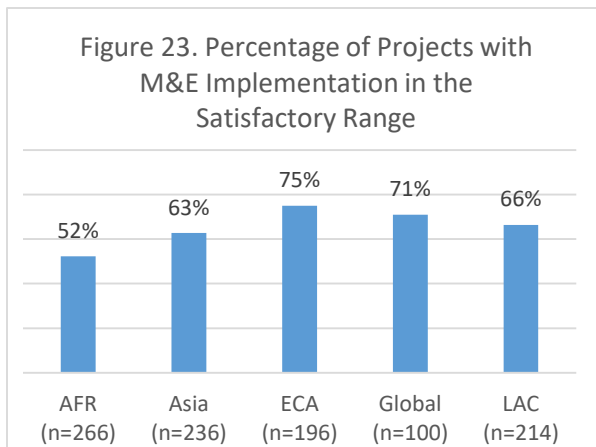


Figure 24: Percentage of Projects with M&E Implementation in the Satisfactory Range

40. Lower percentage of projects in Africa were rated in the satisfactory range for their quality of M&E during implementation (Figure 23). Compared to M&E design ratings, M&E implementation ratings improved the most for projects in ECA and global projects (10 percent each). Only half of the projects implemented in LDCs had M&E implementation rated in the satisfactory range (Figure 24). While M&E implementation ratings of projects in LLDCs and in countries with large portfolios showed some improvement vis-à-vis M&E design ratings, those in SIDS showed a 10 percent drop. This suggests that country context may affect the extent to which M&E plans are well implemented.

Figure 25: Percentage of Projects with M&E Implementation in the Satisfactory Range

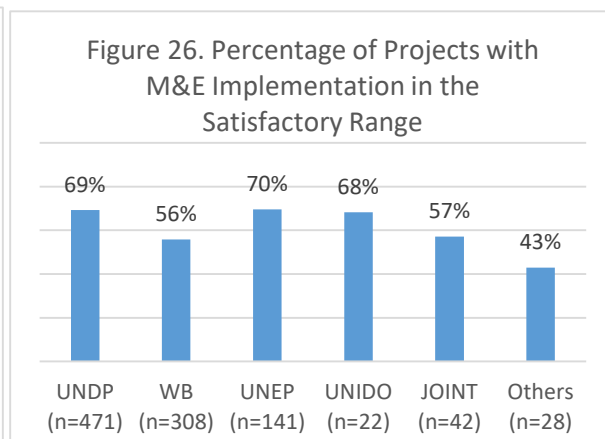
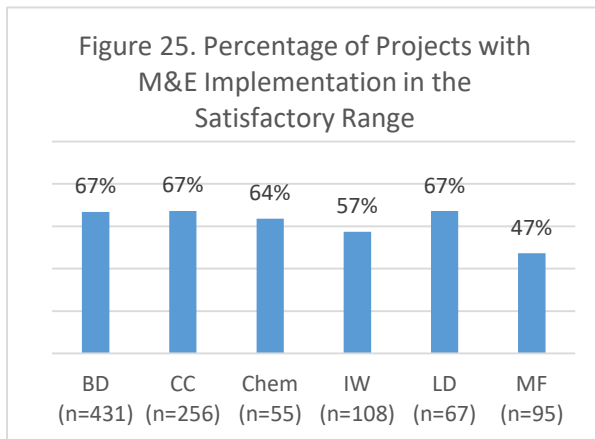


Figure 26: Percentage of Projects with M&E Implementation in the Satisfactory Range

41. Although percentage of multi-focal projects that are rated in the satisfactory range for M&E design is the same as that for projects of other focal areas, the percentage of multi-focal projects rated in the satisfactory range for M&E implementation is lower (Figure 20 and Figure 25). Compared to M&E design ratings, percentage of multi-focal projects rated in the satisfactory range for quality M&E implementation is lower by 15 percent. This suggests that M&E implementation for multi-focal projects may be more complicated than projects from other focal areas.

42. For most of the GEF Agencies the percentage of projects rated in the satisfactory range for quality of M&E implementation closely tracks the percentage rated in the satisfactory range for quality of M&E design. However, “Others,” which combines portfolios of Agencies that have small portfolios of completed GEF projects, is an anomaly. While 81 percent of the projects implemented by “Others” were rated in the satisfactory range, only 43 percent of these were rated in the satisfactory range for quality of M&E implementation (Figure 26). Reasons for the drop for this group are not well understood.

43. Multiple linear regression models indicate that quality of M&E design positively affects M&E implementation. Quality of M&E design in turn is affected by capacities of the GEF Agency and country context. It also shows that projects that were designed in more recent replenishment periods are more likely to be rated in the satisfactory range for quality of M&E design.

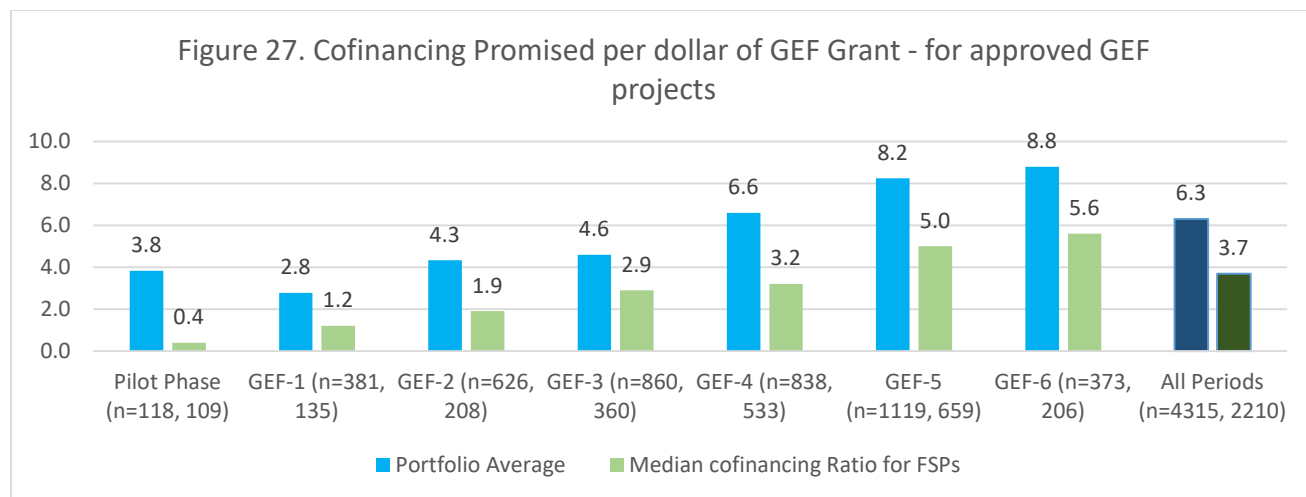
Co-financing

Co-financing is generally considered important for mobilizing resources for achievement of GEF objectives (OPS-5, 2014). Given that the GEF provides funding to “meet the agreed incremental

costs of measures to achieve agreed global environmental benefits” (GEF 2015)¹⁸, it needs to ensure that baseline costs are co-financed by other partners. OPS-5 noted the wide consensus across the preceding OPSs that co-financing is beneficial for GEF projects. Nonetheless, OPS-5 also noted skepticism on the extent to which co-financing helps in generating additional resources for achievement of global environmental benefits. A recent paper that analyzed the GEF project portfolio to assess effects of co-financing found that projects with higher co-financing ratios are correlated with higher outcome ratings.¹⁹ Given its importance, mobilization of co-financing is tracked as an indicator of GEF performance.

44. The new *Co-Financing Policy* (2014)²⁰ of GEF, which became operational during GEF-6, targets a 6:1 level of co-financing for the GEF portfolio. The promised co-financing mobilized for GEF-6 projects through June 2017 is 8.8:1, which exceeds the portfolio target (Figure 27). Across the GEF periods – from GEF-1 to GEF-6 – there has been a steady increase in the co-financing ratio at the portfolio level. In terms of the co-financing ratio of the median full size project, which corrects for the outside influence outliers may have, this improvement is evident from the Pilot Phase onwards. Steady improvement in the ratio for the median project shows the increase in the portfolio co-financing ratio is primarily due to increased effectiveness in seeking higher levels of co-financing for all or most projects. However, there may be variances across GEF periods as to what is reported as co-financing.

Figure 27: Cofinancing Promised per dollar of GEF Grant - for approved GEF projects



45. The GEF *Co-Financing Policy* (2014) also called for seeking “greater co-financing for upper middle income countries that are not SIDS.” The promised co-financing mobilized from the upper middle income countries for GEF-6 so far is 6.8:1, which lower than the GEF portfolio

¹⁸ Instrument for the Establishment of the Restructured Global Environment Facility, GEF, March 2015.

¹⁹ Kotchen and Negi (2014), Cofinancing in Environment and Development: Evidence from the Global Environment Facility, The World Bank Economic Review, 2016.

²⁰ Co-financing Policy (GEF/C.46/09), GEF Council Documents, May 2014.

average but higher than the portfolio target of 6:1. The co-financing ratio for the median full size project from this group, at 5.7:1, is slightly higher than that for the GEF portfolio of full size projects (at 5.6:1), which suggests that the difference in the co-financing ratio may be due to outliers. The GEF-6 period is yet to be complete so the ratios may change when the replenishment period ends.

46. Among the GEF regions, co-financing ratios for LAC are somewhat lower than that for other regions (figure 28). The GEF portfolio in LAC is dominated by biodiversity focal area projects, which generally generate lower levels of co-financing. Across periods, the co-financing ratio of ECA shows a drop during GEF-6. High co-financing ratio achieved by ECA during GEF-5 was driven by two World Bank implemented Climate Change projects in Russia and Turkey.²¹ The drop partly due to the ECA co-financing ratio reverting to its mean and partly due to non-approval of GEF-6 projects for Russia, where GEF projects have traditionally generated higher levels of co-financing.

47. The co-financing ratio for projects that are global in geographic scope has shown substantial increase from GEF-4 to GEF-6 (Figure 28). Some of this increase is also driven by the change type projects undertaken. During the earlier GEF periods most of global projects involved support for foundational activities, whereas during GEF-6 several global projects, especially those within the framework of IAPs, have focused on activities that generate higher level of co-financing commitments.

Figure 28: Promised cofinancing per dollar of GEF Funding: for approved GEF Projects by GEF Region

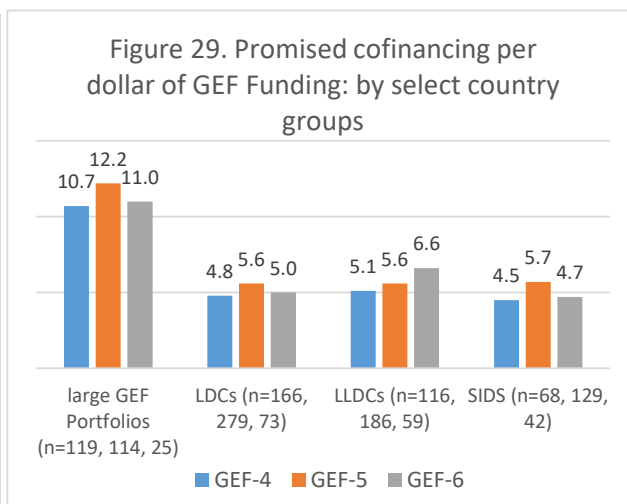
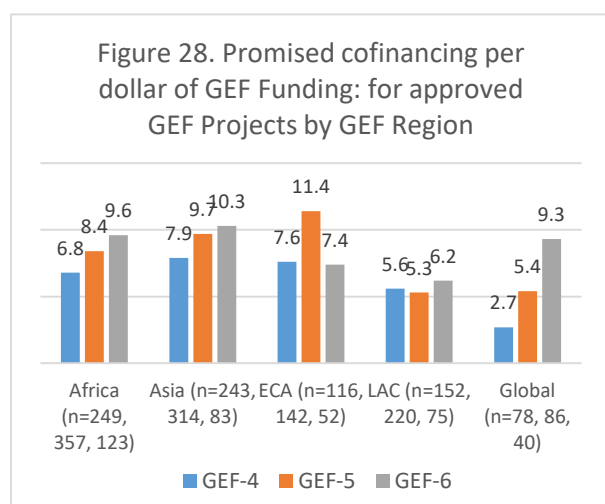


Figure 29: Promised cofinancing per dollar of GEF Funding: by select country groups

²¹ The projects are: *Small and Medium Enterprise Energy Efficiency Project*, GEF ID 4957, GEF grant \$ 3.6 m, co-financing at CEO Endorsement \$ 302 m; and, *Russia Energy Efficiency Financing (REEF) Project*, GEF ID 4427, GEF grant \$ 23 m, co-financing \$ 1,249 m.

48. The approved projects in countries with large GEF portfolios (China, India, Brazil and Mexico) mobilized promised co-financing at 11:1, with a median full size project raising co-financing at 6.6:1, for the GEF-6 period through June 2017 (Figure 29). Both the portfolio co-financing ratio and the median co-financing ratio for projects is higher for this group of countries than the GEF portfolio. Thus, co-financing is being mobilized in the large emerging economies at a higher rate, which is consistent with the higher capacity of these countries to provide co-financing. The co-financing ratio for countries with special circumstances such as LDCs, LLDCs and SIDS is lower than the GEF portfolio average.

49. Multiple linear regression models indicate that promised co-financing is determined by the type of GEF Agency, country context, environmental concern being tackled, size of GEF funding, and year of project approval. Projects that are implemented by the development banks generate higher levels of co-financing. Controlling for some of the key observable variables, compared to other Agencies development banks generate an additional \$ 5.2 of promised co-financing per dollar of GEF grant. Projects implemented in large GEF portfolio countries generate an additional \$ 2.6 of promised co-financing vis-à-vis other countries, whereas SIDS, LDCs and landlocked countries generate lower levels of co-financing. Projects that tackle climate change and international waters related concerns generate more co-financing per dollar of GEF funding, whereas projects that address biodiversity and chemicals generate less co-financing. Another key determinant is size of the GEF funding – greater the GEF funding for a project, higher the co-financing ratio. Much of the influence of the size of GEF funding is also because of the differences in the underlying activities – projects that involve less than US \$ 0.5 million are usually enabling activities, whereas those that involve US \$ 2.0 million or more are exclusively full size projects. When the analysis is restricted to full size projects, effect of the size of GEF funding on co-financing ratio reduces and is not statistically significant. After controlling for other variables, more recent projects generate greater co-financing ratios than projects that were approved earlier – when year of approval increase by a year, an additional \$ 0.35 is generated per dollar of GEF funding.

50. Materialization of the co-financing commitments during implementation is important as several critical outputs of a project may depend on it. Therefore, GEF IEO tracks reported materialization of co-financing. Of the 1181 completed projects for which terminal evaluations are available, co-financing data is available for 991 projects. Despite lack of information for a significant percentage of completed projects (16 percent), the availability is sufficient to draw some broad inferences.

51. On average, materialized co-financing is 126 percent of co-financing commitments. For 59 percent of the projects the co-financing fully materialized and for 69 percent at least 90 percent materialized. For 13 percent of the projects less than half of the promised co-financing materialized.

52. Performance of projects in Africa in meeting co-financing commitments is lower than other GEF regions (Figure 30). Global projects perform better than other projects in meeting at

least 90 percent of the co-financing commitments. However, these correlations weaken and are not statistically significant when other variables are controlled for.

53. For projects in countries with large GEF portfolios, co-financing commitment fully materialized for 72 percent of projects; and at least 90 percent for 83 percent of projects (Figure 31). This is higher than the performance of projects in other countries. In LDCs co-financing commitments are fully met for 49 percent of projects; and, at least 90 percent for 59 percent of projects. This performance is significantly lower than that of projects in other countries. Performance of projects in LLDCs and SIDS in meeting the co-financing commitments is in the same range as other projects in the GEF portfolio.

Figure 30: Materialized cofinancing for GEF Projects vis-a-vis promised cofinancing by Region

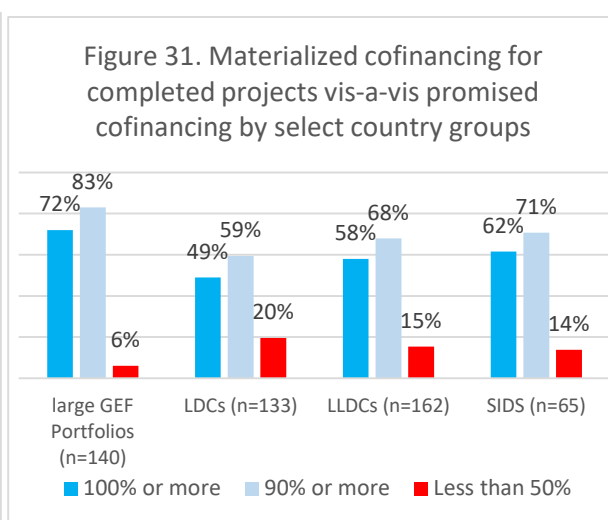
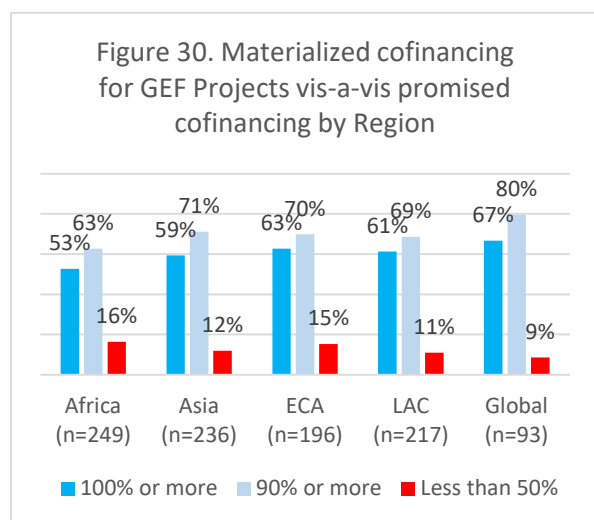


Figure 31: Materialized cofinancing for completed projects vis-a-vis promised cofinancing by select country groups

54. Although co-financing fully materializes for a lower percentage of climate change projects, the difference between projects from climate change focal area and other focal areas is not statistically significant (Figure 32). Jointly implemented projects are more likely to experience less than 50 percent materialization of expected co-financing, than those implemented by a single Agency (Figure 33). A statistically higher percentage of UNIDO implemented projects achieved at least 90 percent of expected co-financing than those implemented by other Agencies.

Figure 32: Materialization of Co-financing by Focal Area

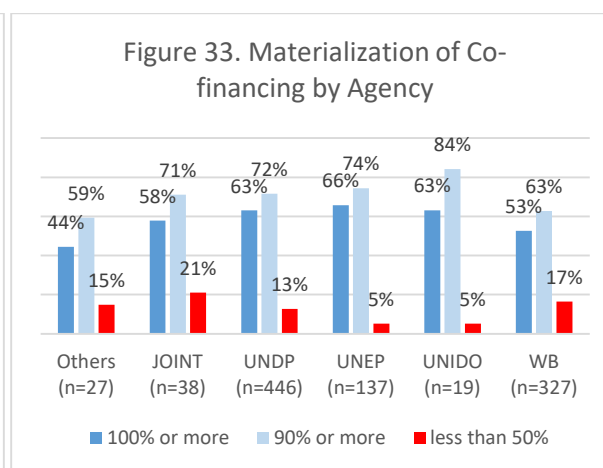
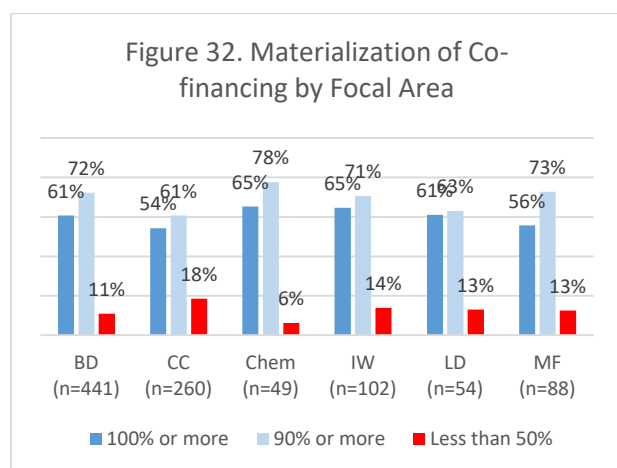


Figure 33: Materialization of Co-financing by Agency

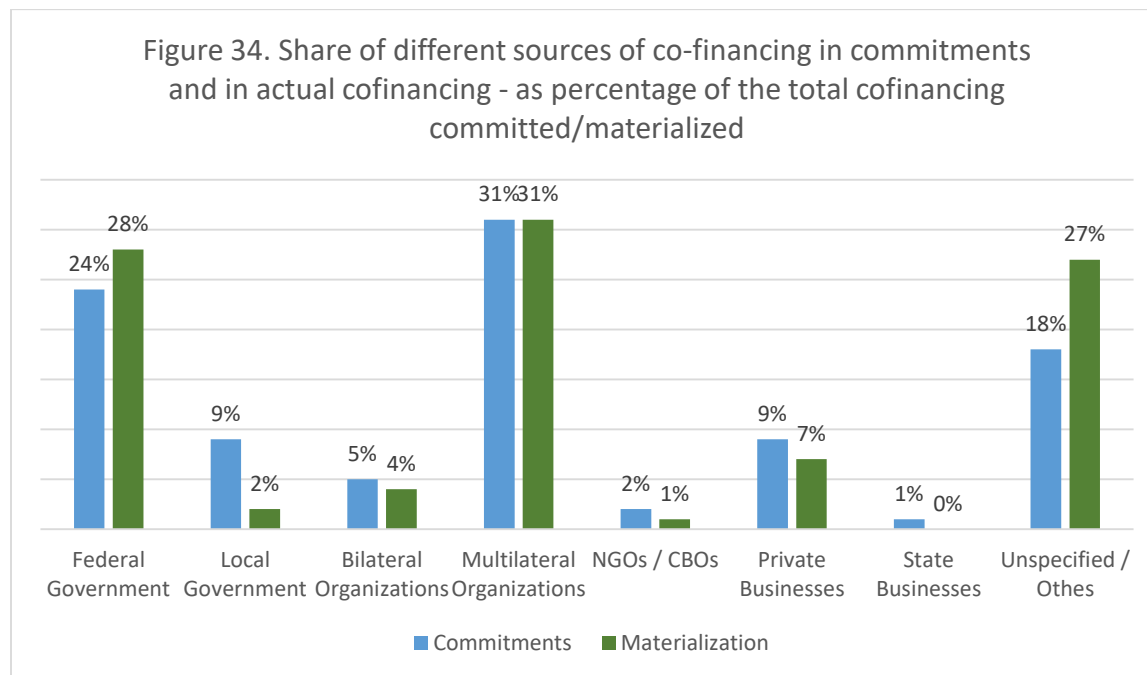
55. The multiple linear regression models used to analyze causal linkages of variables with materialization of expected co-financing do not explain the observed variations well. This indicates that factors that affect materialization of co-financing have not been adequately represented in these models. Nonetheless, two factors stand out. Whether a project is implemented in countries with a large GEF portfolio (positive correlation), and whether it is implemented by a development bank (negative correlation), seem to affect materialization of co-financing. Controlling for other observed factors, promised co-financing is about 15 percent more likely to fully materialize for projects implemented in countries with a large GEF portfolio than in other countries. Further, projects implemented in countries with a large GEF portfolio are also about 10 percent less likely to have less than 50 percent materialization of expected co-financing. On the other hand, projects implemented by development banks have about a 10 percent lower probability of full materialization. Thus, while projects implemented by development banks generate a substantially higher level of promised co-financing, this performance is mitigated by higher risks to full materialization of promised co-financing during implementation.

56. As part of the OPS6 study, data was collected at the level of project co-financier on promised and materialized co-financing levels. Analysis of data from 323 projects from the OPS6 cohort for which information on co-financing at CEO endorsement and completion stage is available is presented in the figure below.²² Figure 34 shows that some co-financing sources such as local government and private businesses do not represent as large a percentage of co-financing at completion as they do at endorsement. It also shows that share of other sources such as federal government increases from commitments made at CEO endorsement to actual materialization during implementation. The largest share of co-financing comes from multilateral organizations, who account for 31 percent of co-financing promised at

²² 426 projects were reviewed, but due to a lack of information at either CEO endorsement or materialization stage on co-financing at the level of co-financier type, analysis can be presented for 323 projects.

endorsement and of materialized co-financing. A larger share of co-financing at completion is reported to have been contributed by unspecified or other sources. This is largely due to the fact reporting of co-financing by co-financier type at completion is often less complete than at endorsement.

Figure 34: Share of different sources of co-financing in commitments and in actual cofinancing - as percentage of the total cofinancing committed/materialized



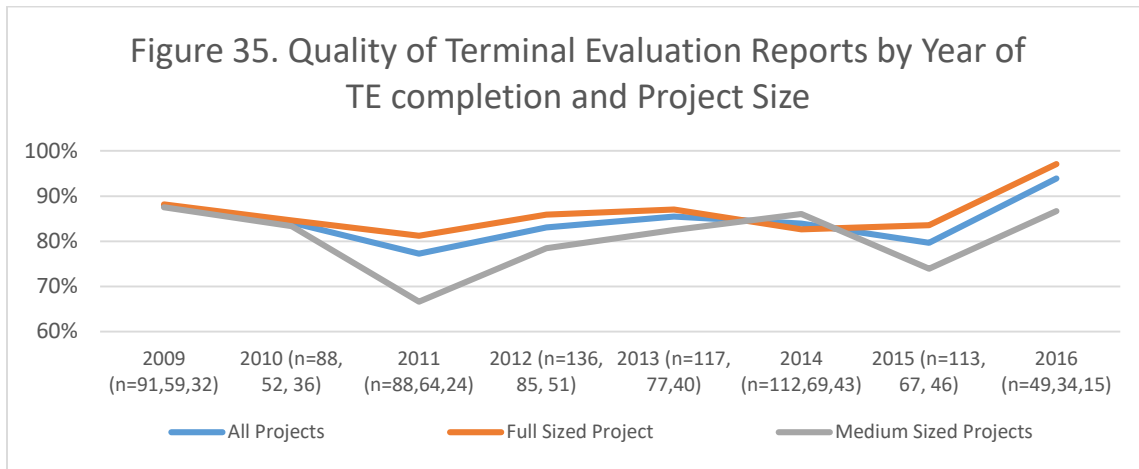
2. Quality of Terminal Evaluations

57. Terminal evaluations are an essential source of information on performance of GEF projects. ‘Minimum requirement 3’ of the GEF Monitoring and Evaluation Policy (2010) requires that the Agencies prepare terminal evaluations for each full-sized project and program at completion. Although Agencies are not required to submit a detailed evaluation for a medium sized project, it is expected that they submit a summarized report for these. Submission of high quality terminal evaluations is an important indicator of Agency performance. Therefore, the GEF IEO tracks quality of terminal evaluations that are submitted to it.

58. Of the 1,169 projects rated on quality of TE, 83 percent are rated in the satisfactory range. For the 581 projects received after the close of OPS5, 571 projects are rated on quality of TE, and 82 percent of those are rated in the satisfactory range. This shows that trends in quality of TEs are stable.

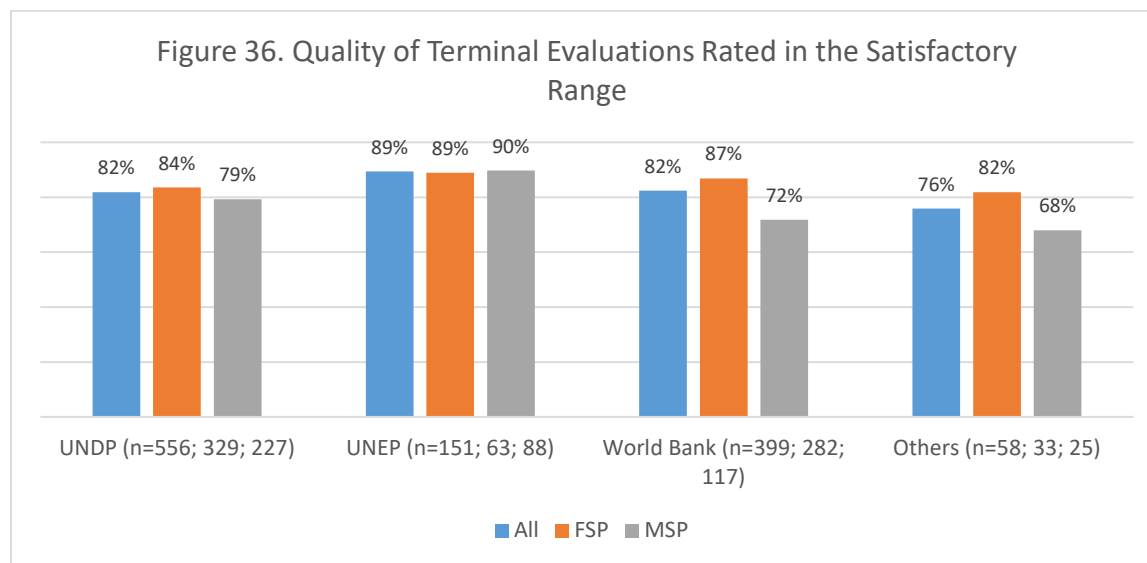
59. To assess the changes in quality of terminal evaluations, it is better to track quality by year of terminal evaluation completion than period of project approval because new guidance on terminal evaluation may be adopted even for projects that had been approved in past periods. Figure 35 shows that the trend of quality of terminal evaluations for the full-sized projects has been stable and has moved in a narrow band of 80 to 90 percent from 2009 to 2015.²³ Generally, a lower percentage of terminal evaluations of medium sized projects are rated in the satisfactory range for quality than those for the full-sized projects. The percentage rated in the satisfactory range also fluctuates more for the medium sized projects owing to the smaller number of observations than those for the full-sized projects.

Figure 35: Quality of Terminal Evaluation Reports by Year of TE completion and Project Size



²³ While a higher percentage of full size projects for which terminal evaluations were completed and submitted in 2016 are rated in the satisfactory range, it is still too early to make a conclusive statement for the terminal evaluations completed in 2016. Most of the terminal evaluations prepared in 2016 are likely to be submitted in 2017.

Figure 36: Quality of Terminal Evaluations Rated in the Satisfactory Range



60. Figure 36 provides a comparison among Agencies based on the percentage of projects that are rated in the satisfactory range for quality of terminal evaluation. Performance of Agencies in the same ball park for quality of terminal evaluations of full size projects. However, UNEP clearly performs better than other Agencies in ensuring quality of terminal evaluations for medium sized projects. While quality of terminal evaluations of UNEP implemented medium-sized projects is at par with the quality for full sized projects, terminal evaluations for World Bank implemented full sized projects is much higher than that for medium-sized projects. While a lower percentage of terminal evaluations of medium-sized projects implemented by other Agencies is in the satisfactory range, number of observations are too small to allow strong conclusions.

3. GEF Activity Cycle

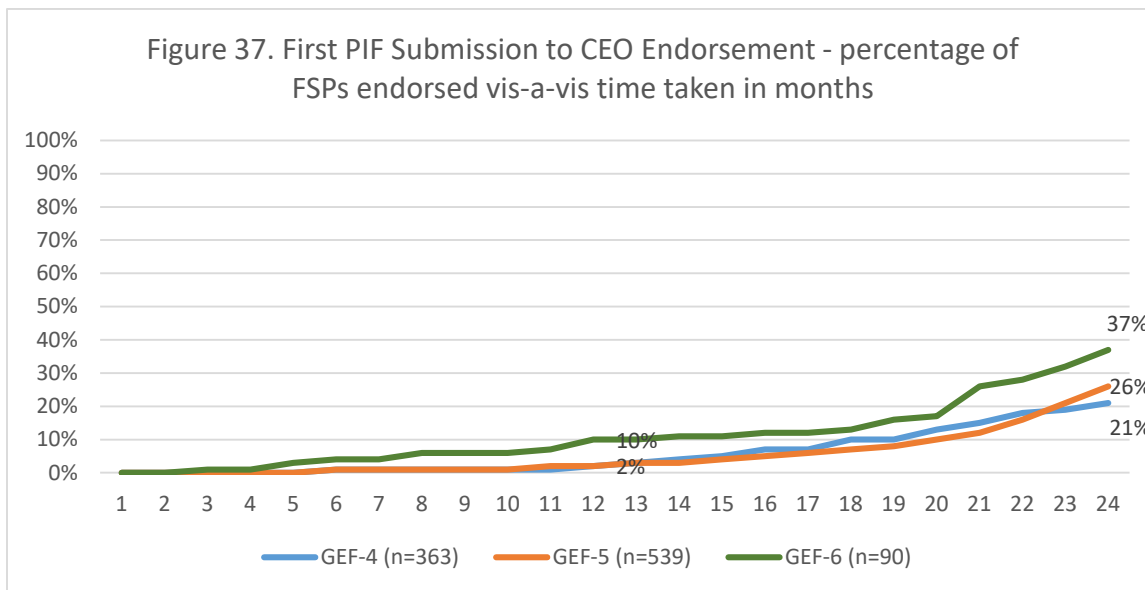
61. Efficiency of the GEF activity cycle is an important concern for the GEF Partnership. A delay in project preparation and implementation reduces GEF efficiency in producing global environmental benefits and may lead to frustration among the key stakeholders. Therefore, several evaluations conducted by the GEF IEO have addressed efficiency of the GEF activity cycle. Further, Secretariat biannually reports on GEF performance on some of the activity cycle indicators through the GEF Corporate Scorecard.

62. This analysis is focused on stand-alone full size projects. These projects are endorsed by the CEO based on a two-step process. The first step involves submission of a Project Information Form (PIF) by an Agency and it culminates after the PIF is approved. The second stage involves preparation of a detailed project proposal by the GEF Agency, submission of the proposal to the GEF Secretariat, and CEO Endorsement of the proposal. GEF has established an 18-month standard for full sized projects from its Project Information Form (PIF) Approval to CEO Endorsement (GEF/C.38/5/Rev.1), i.e. the second step of the process. While efficiency of activity cycle for projects prepared under programmatic framework is important to track, it

poses challenges because child projects need to meet their negotiated program commitment deadline given in their respective Program Framework Document (PFD), and not the 18-month standard applicable to stand alone FSPs. Similarly, with the advent of the single step CEO Approval process for MSPs, along with the continuation of the two-step CEO Approval process, it is difficult to assess efficiency of MSP project cycle.

63. Given that implementation of GEF-6 is in its fourth year, it is possible to track the PIF's submitted during the first year of GEF-6 for at least 24 months. Figure 37 compares performance of the GEF-6 period with that of GEF-5 and GEF-4. It shows that of the 90 PIFs for full-sized projects submitted during the first year of GEF-6, 37 percent had been CEO Endorsed through 24 months from their submission. This is a substantial improvement over the performance during GEF-5 (26 percent) and GEF-4 (21 percent), where a substantially lower percentage of PIF submissions had been CEO Endorsed by the end of 24 months. A fuller picture for the GEF-6 proposals will emerge only after GEF-6 has run its course and sufficient time has elapsed to track progress of the PIFs submitted in GEF-6.

Figure 37: First PIF Submission to CEO Endorsement - percentage of FSPs endorsed vis-a-vis time taken in months



64. Disaggregating performance between PIF submission and CEO Endorsement is important to assess performance during the PIF submission to PIF Approval and PIF Approval to CEO Endorsement stages separately. It allows us to pin-point where progress has taken place and where there has been little or no progress. Figure 38 compares the time taken from PIF submission to PIF Approval for GEF-4, GEF-5 and GEF-6. The data shows that a greater percentage of GEF-6 PIF submissions were approved at various time thresholds than PIFs during

the GEF-4 period. However, performance of the GEF-5 projects for this stage of the project cycle was substantially superior to both GEF-4 and GEF-6 period.

65. Funding short fall during the GEF-6 period may have had some influence at slow progress during GEF-6 through the PIF approval stage. This is illustrated by Figure 39, wherein performance of PIF submissions during the first year and second year of GEF-6 have been compared. It shows that the first-year submissions, which were relatively unaffected by the funding shortfall, moved faster through the PIF approval. In comparison submissions during the second year which faced effects of funding shortfall (in the third year of GEF-6) were slower at achieving PIF approval. This is in contrast to the performance during GEF-5 where the submissions during the second year moved faster through PIF approval than the submissions during the first year, although like GEF-6 there were more submissions during the second year.

Figure 38: Percentage of PIF Submissions that obtained PIF Approval by time taken in months

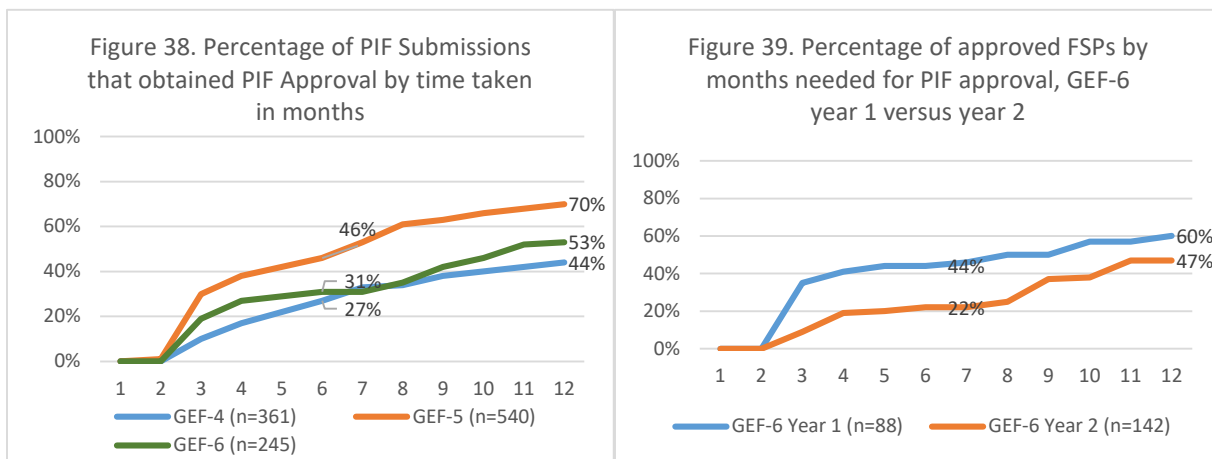
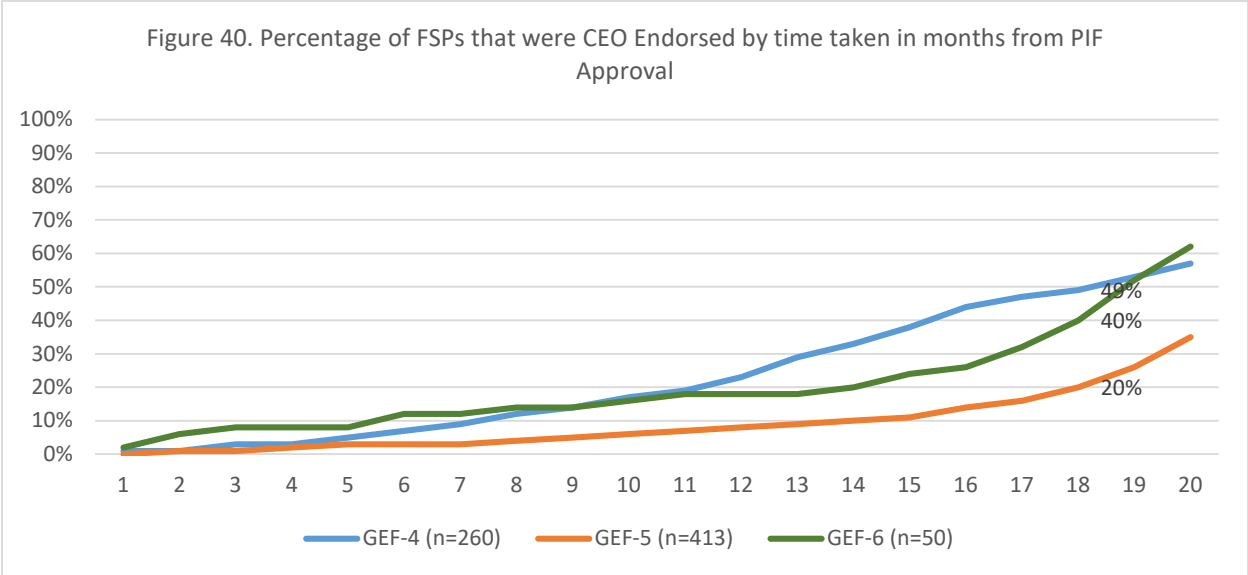


Figure 39: Percentage of approved FSPs by months needed for PIF approval, GEF-6 year 1 versus year 2

66. Figure 40 tracks progress of approved PIF up to 20 months after their approval. The data shows that 40 percent of GEF-6 FSP proposals had obtained CEO Endorsement through 18 months of their PIF approval. This performance is superior to that of the GEF-5 proposals. However, performance at this stage lags the performance of the GEF-4 proposals. The CEO Endorsement rate for GEF-6 proposals catches-up with that for the GEF-4 proposals by the 19th month and is ahead by the 20th month. Overall, the combined performance for the PIF submission to Approval stage, and PIF Approval to CEO Endorsement stage, is superior because lower performance during the first stage is adequately compensated for through improved performance during the second stage. The assessment of the GEF-6 performance is, however,

based on a small pool of PIF approvals. Its only after a couple of years past the GEF-6 period, it would be possible to assess the progress of the entire cohort of PIFs for full sized projects through the CEO endorsement stage.

Figure 40: Percentage of FSPs that were CEO Endorsed by time taken in months from PIF Approval



67. Time lag in data availability for completed projects makes it difficult to assess time taken for project start, implementation, and completion for projects approved during the GEF-5 and GEF-6 period. However, analysis of data on completed projects may provide some indication of the trends for the preceding periods. The data on time taken from CEO Endorsement to project start suggests an improvement in performance of projects that were approved during the more recent GEF periods (Figure 41). This progress is also evident in terms of extensions required for implementation completion – in general projects approved during the more recent periods are more likely to be completed closer to the closing date expected at project start than projects from the earlier periods (Figure 42). This said, the picture for projects approved during GEF-4 will become clearer only after more projects from this period are completed.

Figure 41: CEO Endorsement / Approval to Implementation Start - Time taken in months

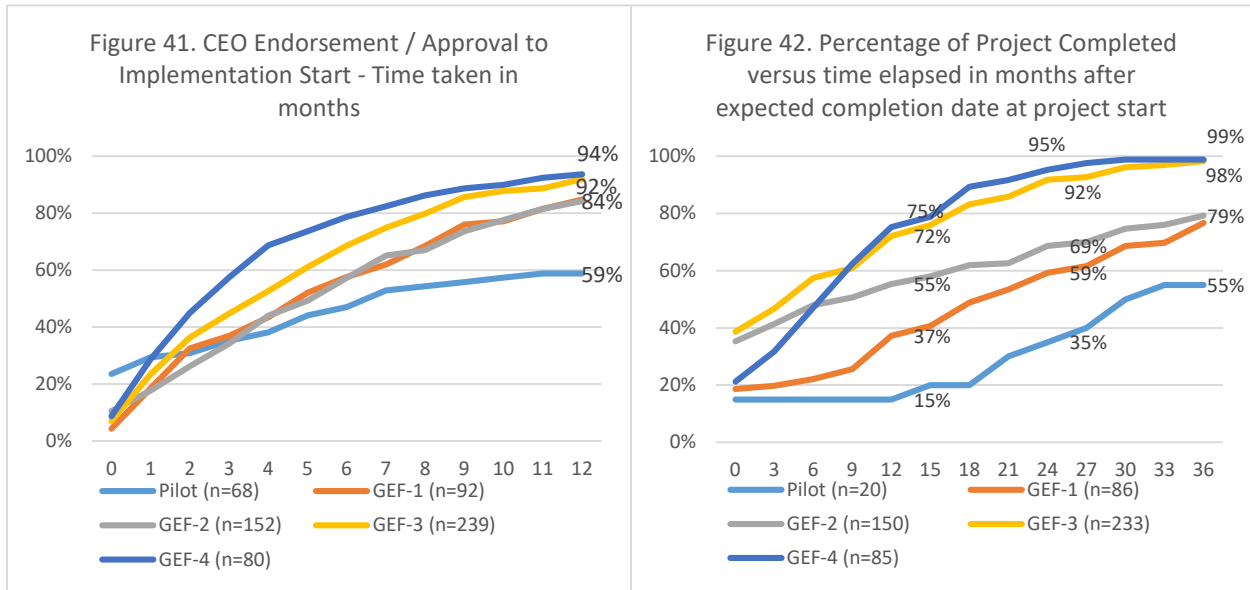


Figure 42: Percentage of Project Completed versus time elapsed in months after expected completion date at project start

4. Progress Towards GEF 5 & 6 Targets

68. **GEF programming for GEF-5 and GEF-6 is consistent with the corporate environmental results targets for these replenishment periods.** The results promised in the documents of projects approved in the GEF-5 period were reviewed. The data from CEO Endorsement / Approval documents shows that GEF is on track to meet most of its GEF-5 replenishment’s environmental results targets. Given that a year remains in completion of the GEF-6 period, and there is about a 15 percent shortfall in actual replenishment versus expected replenishment, the progress for the GEF-6 period is reasonable.

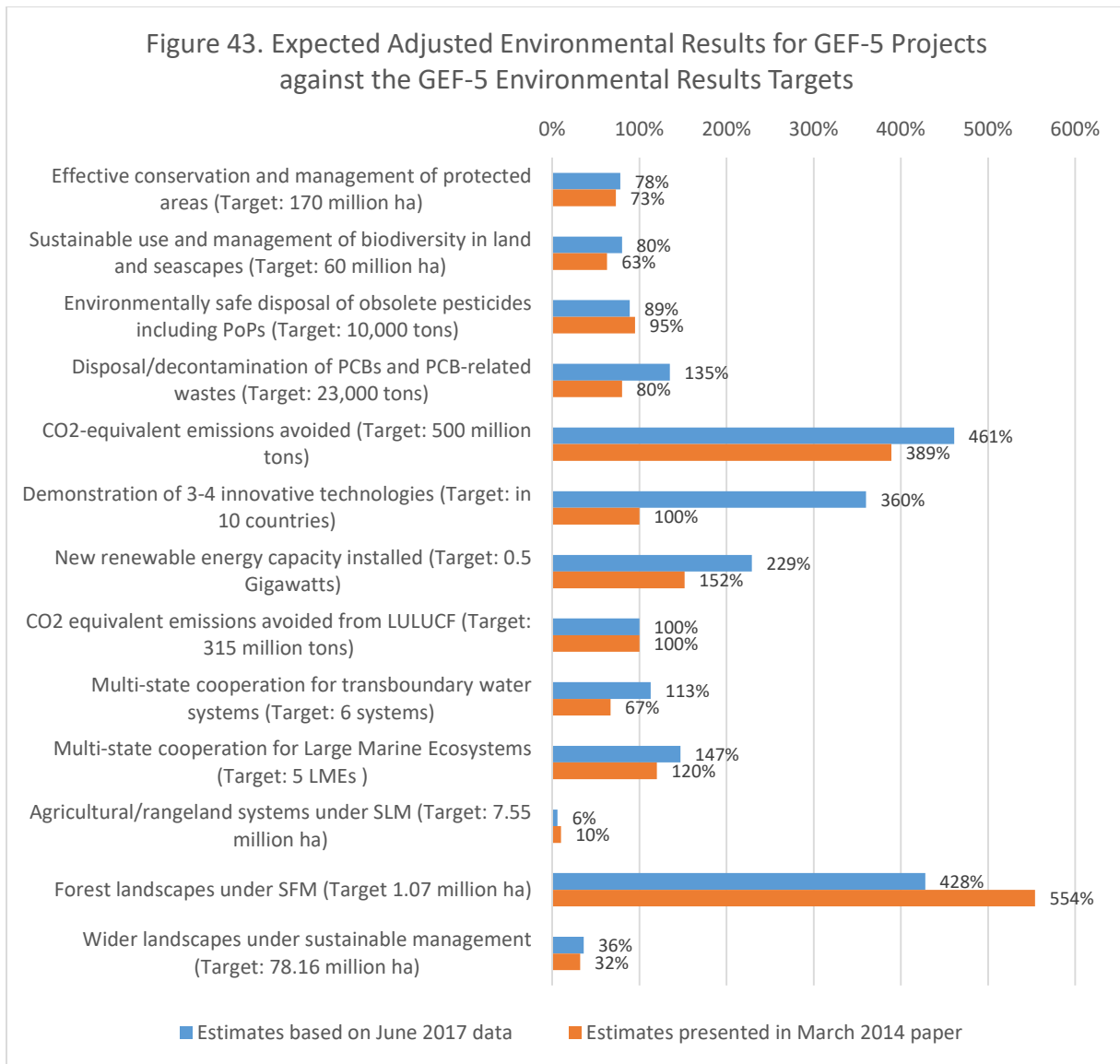
69. To inform the GEF-6 replenishment process, the GEF IEO prepared analyses on progress to environmental results targets for GEF-5 replenishment period. The first paper (GEF/R.6/Inf.09) was presented to the third meeting of the GEF-6 replenishment and update of the paper (GEF/R.6/Inf.13) was presented to the fourth meeting. The former accounted for GEF-5 PIF approvals through June 2013 and the latter updated it based on GEF-5 PIF approvals through December 2013. In preparation for the OPS6, GEF IEO undertook an assessment to update its analysis for the GEF-5 period.

70. The analysis for the GEF-5 period is based on data for 686 projects funded partially or fully through GEF Trust Fund resources.²⁴ The analysis is still primarily based on the aggregation of expected targets for the approved GEF-5 projects. However, unlike the analysis presented in the December 2013 and March 2014 papers, which were primarily based on the aggregated the targets listed in the approved PIFs, the analyses conducted for OPS-6 is based on the expected results indicated in the CEO Endorsement / Approval documents for 96 percent of projects. For the remainder the targets at PIF approval have been used. The aggregated project environmental results targets have been multiplied by 0.8 to account for the likelihood of cancellations and implementation failure.

71. Figure 43 presents the GEF IEO projects of the expected environmental results as a percentage of the targets committed to in the programming directions for GEF-5 for the full GEF-5 cohort and compares it with the progress estimates presented in the March 2014 paper. Of the 13 environmental indicators that could be tracked, and after adjustments for cancellations and implementation failures, GEF is on course to achieve or exceed its expected level of targets for eight indicators. Level of achievement is likely to be slightly lower than the target for three indicators, of which two pertain to chemicals and one to biodiversity conservation. Of the three indicators that are relevant to land degradation focal area, targets are unlikely to be met for two. Compared to the progress reported in the March 2014 paper, the estimate for the GEF-5 period prepared based on June 2017 data shows increased expectations for nine of the 13 indicators. For the remaining four indicators, there has been a decrease in expected benefits as the detailed proposals of some of the projects submitted for CEO Endorsement down scaled the level of expected results. More details are provided in Annex 3.

²⁴ As of December 31, 2015, 973 projects were partially or fully funded through GEF Trust fund resources under GEF-5. 258 Enabling Activity projects, 12 Small Grants Program projects, 3 targeted research projects, and 17 cancelled projects are excluded from analysis.

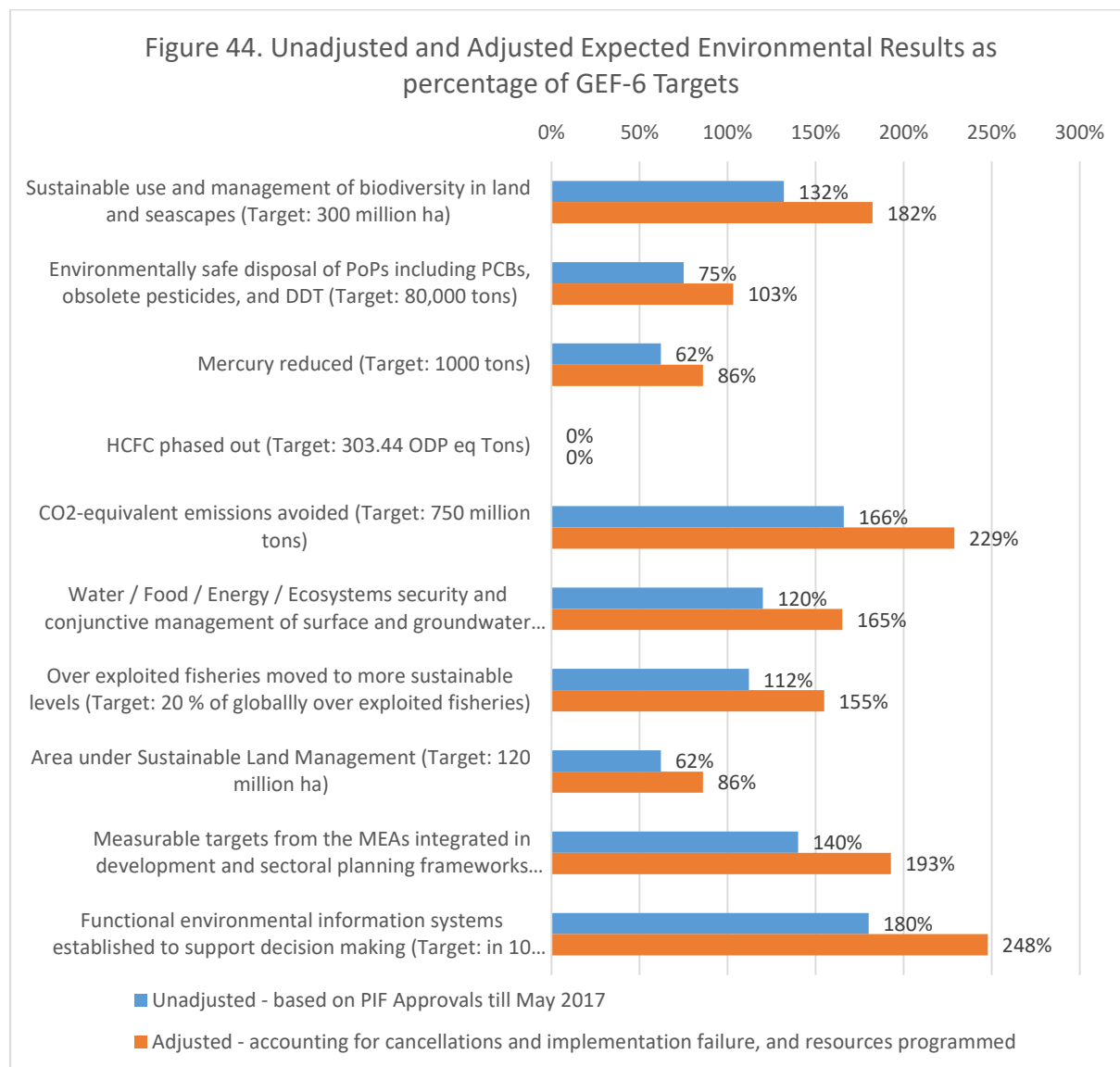
Figure 43: Expected Adjusted Environmental Results for GEF-5 Projects against the GEF-5 Environmental Results Targets



72. During GEF-6, the Secretariat mainstreamed recording of the expected environmental results of the projects in PMIS. It is now reporting on the aggregated targets provided in the approved PIFs for the GEF-6 period through the GEF Corporate Scorecard. Two adjustments have been made to the GEF-6 figures provided in the scorecard. The expected results have been multiplied by a factor of 0.8 to account for cancellations and implementation failures and by 1.72 to account for the level of programmed resources vis-à-vis the GEF-6 replenishment expected at the start of GEF-6. For GEF-6, the aggregated results from approved PIFs exceed GEF-6 targets for 6 out of 10 environmental results indicators (Figure 44). The only indicator for which there was no uptake relates to ozone-depleting substances phaseout, where GEF

involvement has been declining. When the level of fund utilization, and likely cancellations and implementation failure rate, is accounted for, adjusted expected results are commensurate with funding for seven of the 10 indicators. The detailed proposals for most of the approved GEF-6 projects are still under preparation. Further, it is likely that more projects will be approved during the last year of GEF-6 (2017-18). Once the period is complete, it will be possible to make more reliable projections.

Figure 44: Unadjusted and Adjusted Expected Environmental Results as percentage of GEF-6 Targets



5. Progress toward impact

73. It is often too early to assess the long-term impacts of a project at the point of its implementation

completion. Many environmental results take more than a decade to manifest. Similarly, achievement of some environmental results of GEF projects may also be contingent on future actions by other actors. Therefore, any assessment of impacts of GEF projects at project completion is likely to underestimate the number of projects with impacts, as well as the likely scale. Nonetheless, reviewing progress to impacts at project completion helps determine what has already been achieved and the extent to which long-term results are likely. Of the 584 terminal evaluations that were submitted to the IEO after the close of OPS5, 415 were reviewed to determine the extent to which projects had achieved environmental stress reduction and/or status change; and whether broader adoption of promoted approaches, initiatives, or technologies by other stakeholders was taking place, and through what mechanisms.²⁵ The probability adjusted figures for the OPS-6 cohort are presented in Annex 2. In this section, unadjusted figures have been presented because there is little material difference in the unadjusted and adjusted figures.

Environmental Stress Reduction and Status Change

74. Environmental stress reduction may be understood as biophysical changes that reflect reduction of threats emanating from human actions. Fifty-nine percent of the GEF projects achieved stress reduction and/or environmental status change at project completion (Table 1). Achievement of environmental stress reduction and/or environmental change appear to be linked with the environmental challenge being addressed, country context, global versus regional focus, or the scale of GEF funding. Thirteen percent of the projects were achieving environmental stress reduction and/or status change at a large scale—i.e., targeted the system level or national level—and 45 percent of projects were achieving it at a local scale. Forty-one percent of the projects had either not achieved any environmental stress reduction and/or environmental status change yet, or it was not possible to assess whether this had taken place.

Table 1: Percentage of projects achieving environmental stress reduction – by scale

Focal Area	Stress reduction taking place at...			No evidence or unable to assess
	Large scale	Local scale	Any scale	
Biodiversity (<i>n</i> = 147)	10	41	52	48

²⁵ Initially, 426 projects were sampled. After preliminary screening, 11 of these projects were dropped because they were focused on foundational activities and were not expected to deliver environmental stress reduction and status change, and broader adoption.

Climate change (<i>n</i> = 122)	20	49	70	30
Chemicals and waste (<i>n</i> = 25)	16	64	80	20
International waters (<i>n</i> = 38)	11	24	35	66
Land degradation (<i>n</i> = 35)	11	63	74	26
Multifocal area (<i>n</i> = 48)	6	42	48	52
All focal areas (<i>n</i> = 415)	13	45	59	41

75. Projects’ ability to achieve environmental stress reduction at implementation completion is affected by the environmental concern they tackle. For example, 80 percent of projects that focus on chemicals and waste, and 69 percent of those that focus on climate change, achieve stress reduction by implementation completion (Table 1). In comparison, only 35 percent of projects that address international waters–related concerns achieve stress reduction. This result is not surprising, as most of the GEF projects that address international waters focus more on strengthening the intergovernmental arrangements to address these issues, and there is a time lag before these efforts lead to actual stress reduction and/or environmental status change on the ground. Country circumstances also play a role, as stress reduction and/or environmental status change was achieved in 73 percent of the projects implemented in the five countries with large GEF portfolio, but only in 52 percent implemented in SIDS.

76. Compared to projects that are implemented in countries or that are regional in focus, global projects seem to be less likely to be achieving environmental stress reduction and/or status change. Only 21 percent of the global projects, compared to 62 percent of other projects, are reported to be achieving environmental stress reduction and/or status change at implementation completion. Much of the difference is because global projects have, in the past, given more attention to building capacities than to activities that target stress reduction. This variation is evident among the GEF Agencies as well. Projects implemented by UNEP, which accounts for a disproportionately higher percentage of global projects, are less likely to be achieving stress reduction at completion than those implemented by other Agencies (33 percent for UNEP versus 62 percent for other Agencies). Compared to 67 percent of FSPs, 44 percent of medium-size projects (MSPs) achieve environmental stress reduction and/or status change at completion; this difference is statistically significant.

Broader adoption and transformational change

77. The majority (61%) of GEF projects achieved broader adoption at project completion. Country context plays an important role, as projects implemented in major emerging economies are more likely to achieve broader adoption at higher scales than projects in other countries. Broader adoption is said to take place when governments and other stakeholders adopt, expand, and build on the initiatives that the GEF promotes, during program/project implementation or afterwards. GEF IEO's past work shows that broader adoption tends to occur through five mechanisms: sustaining, mainstreaming, replication, scaling-up, and market change (box 3.1). Broader adoption may take place through one or more mechanisms, that may operate simultaneously or sequentially. Broader adoption facilitates transformational change in the systems that the GEF targets. As outlined in the GEF 2020 Strategy, support leading to transformational change is one of the GEF's strategic priorities.

Table 2: Incidence of broader adoption - by scale

Broader adoption taking place (252)	61%
<i>At large scale (100)</i>	24%
<i>At local scale (152)</i>	37%
Broader adoption not taking place (163)	39%
<i>But some progress (108)</i>	26%
<i>No broader adoption or unable to assess (55)</i>	13%
Total (415)	100%

78. The data from review of terminal evaluations shows that 24 percent of completed projects achieved broader adoption at a large scale (Table 2). Among those that achieved broader adoption at a large scale, most of the promoted approaches, initiatives, and/or technologies were being adopted for a third whereas for the remaining third only some of these were being adopted. Thirty-seven percent of the projects achieved broader adoption at a local scale. For 26 percent of the projects, although broader adoption was not yet taking place, plans were in place to facilitate this in the future. Only 13 percent of projects showed no progress in terms of broader adoption, or it was difficult to ascertain their broader adoption status. Differences across focal areas in terms of likelihood of projects achieving broader adoption at completion are not as apparent as they were for environmental stress reduction and/or status change.

79. A higher percentage of projects implemented in the countries with large GEF portfolios (73 percent) achieved broader adoption at the point of completion than projects in other countries (59 percent).²⁶ Broader adoption at completion was also achieved by a higher percentage of projects that replicated an approach that had been piloted elsewhere (75 percent versus 58 percent) and projects that followed up on a preceding GEF project (75 percent versus 59 percent).

6. Mechanisms of Broader Adoption

80. Of the mechanisms for broader adoption, mainstreaming (38 percent of projects), sustaining progress (25 percent), and replication (23 percent) were observed more frequently than upscaling (11 percent of projects) and market change (8 percent) (Table 3). Although broader adoption took place for an optically higher percentage of full sized projects (63 percent) compared to medium sized projects (56 percent), the difference is not statistically significant. A slightly higher number of medium size projects reported mainstreaming than full size projects, however for all other mechanisms the opposite was true. Differences are not statistically significant between full sized and medium sized projects for any of the mechanisms. Projects can result in broader adoption through more than one mechanism. Thirty-four percent of projects achieved broader adoption through a single mechanism. Seventeen percent of projects achieved it through two mechanisms. Ten percent achieved broader adoption through more than two mechanisms.

Table 3: Use of broader adoption mechanisms - percentage of projects where a mechanism was used

Project Size	Full Size Projects (n=268)	Medium Size Project (n=147)	All Projects (n=415)
Sustaining Progress	27%	22%	25%
Mainstreaming	36%	41%	38%
Replication	24%	22%	23%
Upscaling	12%	8%	11%
Market Change	9%	7%	8%

²⁶ This difference is significant at a 90 percent confidence level but not at a 95 percent level.

WORKS CITED

OECD (Organization for Economic Co-operation and Development). *Glossary of Key Terms in Evaluation and Results Based Management*. 2002

GEF (Global Environment Facility). 2006. *Summary of Negotiations on the Fourth Replenishment of the GEF Trust Fund*. (GEF/C.29/3)

- 2010. *Streamlining the Project Cycle & Refining the Programmatic Approach* (GEF/C.38/5/Rev.1)

- 2013. *Verification Note on GEF Secretariat Data on the Realization of GEF-5 Targets*. (GEF/R.6/Inf.09)

-2014. *Updated Verification Note on the Realization of GEF-5 Targets* (GEF/R.6/Inf.13)

-2014. *Co-financing Policy* (GEF/C.46/09)

-2015. *Instrument for the Establishment of the Restructured Global Environment Facility*

-2017. *GEF Corporate Scorecard*

GEF IEO (Global Environment Facility Independent Evaluation Office). 2010. *Monitoring and Evaluation Policy*

-2014. *Annual Performance Report 2013*

-2014. *Co-financing* (Fifth Overall Performance Study (OPS5) Technical Document 21)

- 2014. *At the Crossroads for Higher Impact: Fifth Overall Performance Study (OPS-5) of the GEF*

-2015. *Annual Performance Report 2014*

-2016. *Annual Performance Report 2015*

-2017. *Evaluation of the Programmatic Approaches in the GEF*

Kotchen and Negi. *Cofinancing in Environment and Development: Evidence from the Global Environment Facility*. The World Bank Economic Review. 2016

ANNEXES

Annex 1: Rating criteria and scale

Outcomes²⁷

1. In the causal pathways of a project, its outputs are expected to lead to its intended outcomes. Although achievement of outcomes is not certain, most GEF projects may be expected to achieve the targeted outcomes at implementation completion. The evaluators should, therefore, assess the extent to which the expected outcomes were achieved and the extent to which its achievement was dependent on delivery of project outputs. They should also assess the factors that affected outcome achievement, e.g. project design, project's linkages with other activities, extent and materialization of co-financing, stakeholder involvement, etc. Where the project was developed within the framework of a program, the assessment should also report on the extent the project contributed to the program outcomes.

2. Outcome ratings will take into account the outcome achievements of the projects against its expected targets.²⁸ Project outcomes will be rated on three dimensions:

- (a) **Relevance:** Were the project outcomes congruent with the GEF focal areas/operational program strategies, country priorities, and mandates of the Agencies? Was the project design appropriate for delivering the expected outcomes?
- (b) **Effectiveness:** The extent to which the project's actual outcomes commensurate with the expected outcomes?
- (c) **Efficiency:** Was the project cost-effective? How does the project cost/time versus output/outcomes equation compare to that of similar projects?

Rating Scale for Outcomes: A six-point rating scale is used to assess overall outcomes:

- (a) Highly satisfactory (HS): Level of outcomes achieved clearly exceeds expectations and/or there were no short comings.
- (b) Satisfactory (S): Level of outcomes achieved was as expected and/or there were no or minor short comings.
- (c) Moderately Satisfactory (MS): Level of outcomes achieved more or less as expected and/or there were moderate short comings.

²⁷ Outcomes are "the likely or achieved short-term and medium-term effects of an intervention's outputs. Outputs are the products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes." *Glossary of key terms in evaluation and results based management. OECD, Development Assistance Committee.*

²⁸ Where measurement of outcome achievements is not realistic at the point of project completion, quality and level of outputs delivered may be used as a proxy to indicate outcome achievement.

- (d) Moderately Unsatisfactory (MU): Level of outcomes achieved somewhat lower than expected and/or there were significant shortcomings.
- (e) Unsatisfactory (U): Level of outcomes achieved substantially lower than expected and/or there were major shortcomings.
- (f) Highly Unsatisfactory (HU): Only a negligible level of outcomes achieved and/or there were severe shortcomings.

3. Unable to Assess (UA): The available information does not allow an assessment of the level of outcome achievements.

4. The calculation of the overall outcomes rating of projects will consider all the three criteria, of which relevance and effectiveness are critical. The rating on relevance will determine whether the overall outcome rating will be in the unsatisfactory range (MU to HU = unsatisfactory range). If the relevance rating is in the unsatisfactory range then the overall outcome will be in the unsatisfactory range as well. However, where the relevance rating is in the satisfactory range (HS to MS), the overall outcome rating could, depending on its effectiveness and efficiency rating, be either in the satisfactory range or in the unsatisfactory range.

5. The second constraint applied is that the overall outcome achievement rating may not be higher than the effectiveness rating.

6. During project implementation, the results framework of some projects may have been modified. In cases where modifications in the project impact, outcomes and outputs have not scaled down their overall scope, the evaluator should assess outcome achievements based on the revised results framework. In instances where the scope of the project objectives and outcomes has been scaled down, the magnitude of and necessity for downscaling is taken into account and despite achievement of results as per the revised results framework, where appropriate, a lower outcome effectiveness rating may be given.

Sustainability

7. The GEF Monitoring and Evaluation Policy (2010), minimum requirement 3, specifies that a terminal evaluation will assess the likelihood of sustainability²⁹ of outcomes at project termination and provide a rating. The assessment of sustainability will weigh risks to continuation of benefits from the project. The assessment should identify key risks and explain how these risks may affect continuation of benefits after the GEF project ends. The analysis should cover financial, socio-political, institutional, and environmental risks.

8. The overall sustainability of project outcomes will be assessed based on the likelihood and magnitude of the effect of risks to sustainability. Higher levels of risks and magnitudes of effect, imply lower likelihood of sustainability. The sustainability will be assessed taking into

²⁹ The GEF M&E Policy 2010 adopts the following definition of sustainability: the likely ability of an intervention to continue to deliver benefits for an extended period of time after completion; projects need to be environmentally as well as financially and socially sustainable.

account the risks related to financial, sociopolitical, institutional, and environmental sustainability of project outcomes. The evaluator may also take other risks into account that may affect sustainability. The overall sustainability will be assessed using a four-point scale.

- (a) Likely (L). There is little or no risks to sustainability.
- (b) Moderately Likely (ML). There are moderate risks to sustainability.
- (c) Moderately Unlikely (MU). There are significant risks to sustainability.
- (d) Unlikely (U). There are severe risks to sustainability.

9. Unable to Assess (UA). Unable to assess the expected incidence and magnitude of risks to sustainability.

Project Monitoring & Evaluation

10. The GEF M&E minimum requirement 1 calls for fully developed and budgeted project M&E plan at CEO Endorsement, and the minimum requirement 2 calls for implementation of these plans. The evaluators will include an assessment of the strengths and weaknesses of the project M&E plan and its implementation.

11. **M&E Design.** To assess the quality of the M&E plan, the evaluators will assess: Was the M&E plan at the point of CEO Endorsement practical and sufficient? Did it include baseline data? Did it: specify clear targets and appropriate (SMART³⁰) indicators to track environmental, gender, and socio economic results; a proper methodological approach; specify practical organization and logistics of the M&E activities including schedule and responsibilities for data collection; and, budget adequate funds for M&E activities?

12. **M&E Implementation.** The evaluators should assess: Whether the M&E system operated as per the M&E plan? Where necessary, whether the M&E plan was revised in a timely manner? Was information on specified indicators and relevant GEF focal area tracking tools gathered in a systematic manner? Whether appropriate methodological approaches have been used to analyze data? Were resources for M&E sufficient? How was the information from M&E system used during the project implementation?

13. Quality of M&E on these two dimensions will be assessed on a six-point scale:

- (a) Highly satisfactory (HS): There were no short comings and quality of M&E design / implementation exceeded expectations.
- (b) Satisfactory (S): There were no or minor short comings and quality of M&E design / implementation meets expectations.

³⁰ SMART: Specific, Measurable, Achievable/Attributable, Relevant/Realistic, and Time-bound, Timely, Trackable and Targeted.

- (c) Moderately Satisfactory (MS): There were some short comings and quality of M&E design/implementation more or less meets expectations.
- (d) Moderately Unsatisfactory (MU): There were significant shortcomings and quality of M&E design / implementation somewhat lower than expected.
- (e) Unsatisfactory (U): There were major short comings and quality of M&E design/implementation substantially lower than expected.
- (f) Highly Unsatisfactory (HU): There were severe short comings in M&E design/ implementation.

14. Unable to Assess (UA): The available information does not allow an assessment of the quality of M&E design / implementation.

Implementation

15. The assessment of the implementation of GEF full size projects will take into account the performance of the GEF Agencies in discharging their expected roles and responsibilities. The performance of the Agencies will be rated using a six-point scale (Highly Satisfactory to Highly Unsatisfactory). Within the GEF partnership, GEF Agencies are involved in activities related to a project's identification, concept preparation, appraisal, preparation of detailed proposal, approval and start-up, oversight, supervision, completion, and evaluation.³¹ To assess performance of the GEF Agencies, the evaluators will assess the extent to which the agency delivered effectively on these counts, with focus on elements that were controllable from the given GEF Agency's perspective. The evaluator will assess how well risks were identified and managed by the GEF Agency.

16. The quality of implementation will be rated on a six-point scale.

- (a) Highly satisfactory (HS): There were no short comings and quality of implementation exceeds expectations.
- (b) Satisfactory (S): There were no or minor short comings and quality of implementation meets expectations.
- (c) Moderately Satisfactory (MS): There were some short comings and quality of implementation is more or less meets expectations.
- (d) Moderately Unsatisfactory (MU): There were significant shortcomings and quality of implementation is somewhat lower than expected.

³¹ See GEF/C.41/06/Rev.01 and GEF/C.39/9

(e) Unsatisfactory (U): There were major short comings and quality of implementation substantially lower than expected.

(f) Highly Unsatisfactory (HU): There were severe short comings in quality of implementation.

17. Unable to Assess (UA): The available information does not allow an assessment of the quality of implementation / execution.

Annex 2. Progress to impact results corrected for differences in sampling probability

Annex 2.1. Percentage of projects achieving environmental stress reduction – by scale

Focal Area	Not adjusted: Stress reduction at...			Probability adjusted: Stress reduction at...		
	Large scale	Local scale	Any scale	Large scale	Local scale	Any scale
Biodiversity (<i>n</i> = 147)	10	41	52	11	43	54
Climate change (<i>n</i> = 122)	20	49	70	20	52	72
Chemicals and waste (<i>n</i> = 25)	16	64	80	18	62	79
International waters (<i>n</i> = 38)	11	24	35	9	25	35
Land degradation (<i>n</i> = 35)	11	63	74	12	60	72
Multifocal area (<i>n</i> = 48)	6	42	48	5	37	42
All focal areas (<i>n</i> = 415)	13	45	59	13	46	59

Annex 2.2 Incidence of broader adoption by scale

Broader adoption status	Gross figures	Probability adjusted figures
Broader adoption taking place (252)	61%	61%
<i>At large scale (100)</i>	24%	24%
<i>At local scale (152)</i>	37%	37%
Broader adoption not taking place (163)	39%	39%
<i>But some progress (108)</i>	26%	25%
<i>No broader adoption or unable to assess (55)</i>	13%	14%
All projects (415)	100%	100%

Annex 2.3 Use of broader adoption mechanisms - percentage of projects that use a mechanism

Mechanism	Gross figures			Probability Adjusted Figures		
	Full Size Projects (n=268)	Medium Size Project (n=147)	All Projects (n=415)	Full Size Projects (n=268)	Medium Size Project (n=147)	All Projects (n=415)
Sustaining	27%	22%	25%	26%	22%	24%
Mainstreaming	36%	41%	38%	37%	40%	38%
Replication	24%	22%	23%	24%	23%	24%
Upscaling	12%	8%	11%	13%	7%	11%
Market Change	9%	7%	8%	8%	6%	7%

Annex 3. Progress towards GEF 5 Targets

Annex 3.1 Strategic goals and targets from the May 2010 GEF-5 programming document, summation of project-level (excluding 15 cancelled projects) targets to date, and estimated percentage of replenishment targets that will be achieved from GEF-5 projects (December 31, 2015).

Focal area	Strategic goal	Targets under	GEF-5 project level targets to date (excluding 17 cancelled projects)	Percentage of replenishment target contained in project-level targets to date	Estimated % of replenishment target to be achieved from projects to date*
Biodiversity	Improved sustainability of protected area systems	Effective conservation and management of 170 million hectares of protected areas	63.33 million Ha of new protected areas; 101.45 million Ha of existing protected areas	97% of target**	78% of target
	Sustainably managed landscapes and seascapes that integrate biodiversity conservation increased	Sustainable use and management of biodiversity in 60 million hectares of production landscapes and seascapes	55.31 million Ha of production landscapes; 4.87 million Ha of production seascapes	100% of target	80% of target
Chemicals	Phased out and reduced releases of POPs, ODS, and other chemicals of global concern	10,000 tons of obsolete pesticides, including POPs, disposed of in an environmentally sound manner	11,146 tons of obsolete pesticides, including POPs, disposed of in an environmentally sound manner	111% of target	89% of target

		23,000 tons of PCBs and PCB-related wastes disposed of or decontaminated	33,560 tons of PCBs and PCB-related wastes disposed of or decontaminated	169% of target	135% of target
Climate Change	Slowed growth in GHG emissions to the atmosphere from demonstration and transfer of advanced low-carbon technologies and deployment and diffusion of technologies in energy efficiency, renewable energy, and sustainable transport and urban systems	500 million tons of CO2-equivalent emissions avoided	459MTCO2eq direct mitigation; 2,414 MtCO2eq indirect mitigation;	94% of target if only direct is included; 577% of target if including indirect	75% of target if only direct included; 461% of target if including indirect
		Demonstration of 3-4 innovative technologies in 10-15 countries	Demonstration of 16 innovative technologies in 24 countries	450% of country target***	360% of country target
		0.5 gigawatts of new renewable energy capacity installed	1.42 gigawatts of new renewable energy capacity installed	286% of target	229% of target
		315-675 million tons of CO2 equivalent emissions avoided from LULUCF	549 MtCO2eq emission reductions****	100% of target	100% of target
Intl. Waters	Catalyze multi-state cooperation to balance conflicting water uses in transboundary surface and groundwater basins while considering climatic variability and change	Multi-state cooperation results in: adoption/implementation of national/local reforms in 50% of States and demonstration results in at least 50% of States participating in 6-7 transboundary water systems	10 transboundary water systems targeted through 10 projects involving 48 different countries.	142% of measurable target(# of transboundary water system targeted)	113% of measurable target

	Catalyze multi-state cooperation to rebuild marine fisheries and reduce pollution of coasts and Large Marine Ecosystems (LMEs) while considering climatic variability and change	Multi-state cooperation results in: adoption/implementation of national/local reforms in 50% of States and demonstration results in at least 50% of States participating in 5-6 LMEs	11 LMEs targeted through 15 projects involving 66 countries.	183% of measurable target (# of LMEs targeted)	147% of measurable target
Land Degradation	Arrested or reversed current global trends in land degradation, specifically desertification and deforestation	Sustainable management of agriculture, range and forest landscapes, including drylands and affected transboundary areas: 100 million Ha in agriculture; 200,000 Ha of forest landscapes; 175 million Ha in wider production landscapes	7.55 million Ha of agricultural/rangeland systems under SLM; 1.07 million Ha of forest landscapes under SFM; 78.16 million Ha of wider production landscapes under sustainable management	8% of target 535% of target 45% of target	6% of target 428% of target 36% of target

* Estimated percentage of replenishment target to be achieved by projects assumes that 80% of project-level targets will be achieved

**Improved management of hectares of protected areas achieved indirectly by systemic improvement of the entire protected area system through increased financial resources and/or strengthened capacity were not counted when calculating the target achievement.

***Aggregate project-level target is derived from projects with CCM-1 (tech transfer) funding. See section 3 for detail on the types of innovative technology demonstrated in these GEF-5 projects.

**** Total project-level reductions from LULUCF includes both direct and indirect reductions

Annex 3.2 Progress towards GEF 6 Targets

Focal Area	Strategic goal	Targets under	Estimated % of replenishment target to be achieved from projects to date*
Biodiversity	Maintain globally significant biodiversity and the ecosystem goods and services it provides to society.	300 million hectares of landscapes and seascapes under improved biodiversity management.	132% of target
Chemicals	Promote the sound management of chemicals throughout their lifecycle to minimize adverse effects on the global environment and health of both women and men.	80,000 tons of Persistent Organic Pollutants including PCB, obsolete pesticides and DDT disposed of in an environmentally sound manner.	75% of target
		1000 tons of mercury reduced	62% of target
		303.44 ODP tons of HCFC phased out.	0% of target
Climate Change	To support developing countries and economies in transition to make transformational shifts towards a low-emission, resilient development path	750 millions tons CO ₂ equivalent avoided, both direct and indirect, over the investment or impact period of the projects.	166% of target
Intl. Waters	Promotion of collective management of	Water/Food/Energy/Ecosystems security and conjunctive management	120% of target

	transboundary water systems and implementation of the full range of policy, legal, and institutional reforms and investments contributing to sustainable use and maintenance of ecosystem services	of surface and groundwater enhanced in at least 10 freshwater basins	
		20% of globally over-exploited fisheries (by volume) moved to more sustainable levels	112% of target
Land Degradation	To contribute to arresting and reversing current global trends in land degradation, specifically desertification and deforestation	120 million hectares under Sustainable Land Management	62% of target
	Enhance capacity of countries to implement Multilateral Environmental Agreements (MEAs)	Development and sectoral planning frameworks integrate measurable targets drawn from the MEAs in at least 10 countries	140% of target
		Functional environmental information systems are established to support decision-making in at least 10 countries	180% of target

* Source, April 2017 GEF Scorecard. Based on 299 projects at the stage of Project Identification (PIF approval) in GEF-6, 124 projects of which were CEO endorsed/approved by April 30, 201

Annex 4: Regression Models

VARIABLES	Outcomes model 1	Outcomes model 2	Outcomes model 3	Outcomes model 4	Outcomes model 5	Outcomes model 6
Quality of Implementation	0.467***	0.467***	0.466***	0.466***	0.461***	0.463***
Quality of Execution	0.305***	0.304***	0.302***	0.305***	0.279***	0.291***
Large GEF Portfolios	0.00592					
LDC		-0.00840				
SIDS			-0.0599			
AFR				-0.00757		
<50% of Co-financing materialized					-0.0884***	
Co-financing fully realized						0.0411**
<i>Control Variables</i>						
Focal Area	yes	yes	yes	yes	yes	yes
GEF Grant	yes	yes	yes	yes	yes	yes
Year of Implementation Start	yes	yes	yes	yes	yes	yes
Project Preparation Grant Given	yes	yes	yes	yes	yes	yes
Observations	943	943	943	943	828	828
R-squared	0.491	0.491	0.492	0.491	0.486	0.483

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	Sustainability model 1	Sustainability model 2	Sustainability model 3	Sustainability model 4	Sustainability model 5	Sustainability model 6
Quality of Implementation	0.180***	0.183***	0.189***	0.177***	0.163***	0.163***
Quality of Execution	0.254***	0.250***	0.265***	0.256***	0.248***	0.256***
Large GEF Portfolios	0.230***					
LDC		-0.147***				
SIDS			0.0547			
AFR				-0.111***		
<50% of Co-financing materialized					-0.0872*	
Co-financing fully realized						0.0548
Climate Change	0.0784**	0.0857**	0.0823**	0.0783**	0.0972***	0.0956**
<i>Control Variables</i>						
GEF Grant	yes	yes	yes	yes	yes	yes
Year of Project	yes	yes	yes	yes	yes	yes
PPG	yes	yes	yes	yes	yes	yes
Observations	919	919	919	919	806	806
R-squared	0.146	0.135	0.125	0.134	0.113	0.113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	Implementation model 1	Implementation model 2	Implementation model 3	Implementation model 4	Implementation model 5	Implementation model 6	Implementation model 7	Implementation model 8
Large GEF Portfolio	0.0803*							
LDC		-0.129***						
SIDS			-0.0999*					
AFR				-0.131***				
>50% Co-financing materialized					-0.233***			
Co-financing fully realized						0.110***		
WB							-0.0751**	
UNEP								0.112***
<i>Control Variables</i>								
Focal Area	yes	yes	yes	yes	yes	yes	yes	yes
GEF Grant	yes	yes	yes	yes	yes	yes	yes	yes
Year of Project Start	yes	yes	yes	yes	yes	yes	yes	yes
PPG	yes	yes	yes	yes	yes	yes	yes	yes
Observations	970	970	970	970	846	846	970	970
R-squared	0.037	0.045	0.037	0.052	0.069	0.049	0.039	0.041

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	M&E Design model 1	M&E Design model 2	M&E Design model 3	M&E Design model 4	M&E Design model 5	M&E Design model 6
Large GEF Portfolios	0.0688					
LDC		-0.123***				
SIDS			0.0595			
AFR				-0.118***		
UNDP					0.0665**	
WB						-0.0917***
Chemicals	-0.125*	-0.138**	-0.118*	-0.144**	-0.109*	-0.137**
<i>Control Variables</i>						
GEF Grant	yes	yes	yes	yes	yes	yes
Year of Project Start	yes	yes	yes	yes	yes	yes
PPG	yes	yes	yes	yes	yes	yes
Observations	1,108	1,108	1,108	1,108	1,108	1,108
R-squared	0.044	0.049	0.043	0.053	0.046	0.049

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	M&E Implementation model 1	M&E Implementation model 2	M&E Implementation model 3	M&E Implementation model 4	M&E Implementation model 5	M&E Implementation model 6
Large GEF Portfolios	0.0995**					
LDC		-0.176***				
SIDS			-0.0929			
AFR				-0.156***		
WB					-0.0916***	
Other Agencies						-0.132*
MF	-0.211***	-0.213***	-0.210***	-0.201***	-0.204***	-0.203***
<i>Control Variables</i>						
GEF Grant	yes	yes	yes	yes	yes	yes
Year of Project Start	yes	yes	yes	yes	yes	yes
PPG	yes	yes	yes	yes	yes	yes
Observations	1,012	1,012	1,012	1,012	1,012	1,012
R-squared	0.028	0.040	0.026	0.044	0.031	0.027

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1