



GEF/ME/C.40/Inf.2
April 26, 2011

GEF Council
May 24-26, 2011
Washington, D.C.

ANNUAL PERFORMANCE REPORT 2010 (FULL REPORT)

(Prepared by the GEF Evaluation Office)

TABLE OF CONTENTS

1. BACKGROUND AND MAIN CONCLUSIONS	1
1.1 Background.....	1
1.2 Findings and Conclusions.....	2
2. SCOPE AND METHODOLOGY.....	10
2.1 Scope.....	10
3. OUTCOMES AND SUSTAINABILITY.....	11
3.1 Rating Approach.....	11
3.2 Outcomes	12
3.3 Sustainability.....	16
4. FACTORS AFFECTING ATTAINMENT OF PROJECT RESULTS.....	18
4.1 Materialization of Cofinancing.....	18
4.2 Delays in Project Completion.....	21
5. QUALITY OF PROJECT MONITORING	23
5.1 Rating Approach.....	23
5.2 Quality of M&E.....	23
6. QUALITY OF TERMINAL EVALUATIONS.....	25
6.1 Rating Approach.....	26
6.2 Findings.....	26
6.3 Performance by Quality Dimension.....	26
6.4 Delays in Terminal Evaluation Report Submission.....	30
6.5 Comparison of Ratings	32
7. MANAGEMENT ACTION RECORDS.....	33
7.1 Rating Approach.....	33
7.2 Findings.....	34
8. PERFORMANCE MATRIX	40
8.1 Rating Approach.....	40
8.2 Findings.....	41
ANNEX A	42
ANNEX B	48
ANNEX C	52

1. BACKGROUND AND MAIN CONCLUSIONS

1.1 BACKGROUND

1. This document is the seventh annual performance report (APR) that the Evaluation Office presents. It includes a detailed account of some aspects of project results, of processes that may affect these results, and of monitoring and evaluation (M&E) arrangements in completed GEF projects. For the fourth time, a performance matrix, which summarizes the performance of the GEF Implementing Agencies and the GEF Secretariat on various parameters tracked by the Office, is presented.

2. The APR 2010 continues the annual presentation of assessment of project outcomes, project sustainability, project completion delays, materialization of cofinancing, and quality of monitoring in completed projects.

3. In August 2010, 57 terminal evaluations were received by the Evaluation Office, however, nine projects were removed from the FY10 cohort because they were not accompanied by a TER from the Agency. These will be considered in the FY11 cohort.

4. In total 388 project terminal evaluations, which represent US\$ 1,763 million in GEF funding, have been submitted to the GEF Evaluation Office. These include terminal evaluations for 48 projects submitted to the Office in FY2010 that account for US\$177 million in GEF funding.

5. For the **assessment of project outcomes, project sustainability, and delays in project completion**, 313 projects, for which terminal evaluation reports were submitted by the GEF Agencies to the Evaluation Office since FY 2005, have been considered.

6. For reporting on **materialization of cofinancing** 313 projects for which terminal evaluation reports have been submitted since FY 2005 were considered. Of these, information on materialization of cofinancing has been reported for 257 projects (83 percent). The GEF has invested a total of US\$ 1,063 million in these 257 projects; the implementing agencies reported that a cofinancing of US\$ 2,463 million materialized during implementation. We have reported actual cofinancing amounts for 44 projects of the FY 2010 cohort, representing a total cofinancing amount of \$406 million (based on TE and Agency reports).

7. This year's **management action record (MAR)** tracks the level of adoption of 35 Council decisions based on 16 evaluation reports. The Evaluation Office was able to verify 35 of these decisions. Nine (9) of these decisions that achieved 'high' adoption ratings and 16 are no longer relevant and, therefore, will not be tracked in the next MAR.

8. The **performance matrix** provides a summary of the performance of the GEF Agencies and the GEF Secretariat on relevant parameters. The matrix is limited to the three implementing agencies and the Secretariat because the completed projects from the other

GEF agencies are still not significant. Most of the parameters included in the matrix are assessed on an annual basis by the Evaluation Office whereas others are tracked after two to three years. This year, ratings have been presented on six parameters.

9. The APR primarily involves review of the evidence presented in the terminal evaluation reports, with verification of performance ratings based primarily on desk reviews. The evaluation offices of several agencies have been conducting similar reviews for past couple of years. Since FY 2007 the UNEP's Evaluation Office has been providing performance ratings for all the completed GEF projects. Similarly, The World Bank's Independent Evaluation Group (IEG) conducts desk reviews of all the terminal evaluation reports produced by the management for the full size projects and conducts more intensive field verifications for a sample of these projects. In FY2009, for the first time, UNDP's evaluation office began to review terminal evaluation and did so again for FY2010.

10. The GEF Evaluation Office has tracked the ratings provided by the evaluation offices of the agencies and has found that these match well with its ratings. This is in contrast to the comparisons with ratings provided in the terminal evaluations which, when compared to the ratings provided by the GEF Evaluation Office, tend to be more optimistic. To reduce duplication of effort, the GEF Evaluation Office started accepting the ratings provided by the evaluation offices of UNEP and the World Bank. In FY2009, for the first time, UNDP's Evaluation Office also began to review terminal evaluations. Consistency between GEFEO ratings and ratings of the evaluation offices of the three Agencies has been established and therefore ratings from the three agencies were accepted for 21 projects. To ensure compatibility, the Office will continue to review a sample of terminal evaluations from all Agencies.

11. In the past, the GEF Evaluation Office has received terminal evaluation reviews from Agencies in a batch. This year for the first time this turned out to be a problem for UNDP's Evaluation Office due to workload. The GEF Evaluation Office expected 26 UNDP terminal evaluations and reviews but received 17. The delays were discussed with UNDP bilaterally as well as at the inter-agency meeting on preliminary findings of the APR2010 in April 2011 with an agreement reached to proceed with receipt of terminal evaluation reviews on a rolling basis from Agencies rather than in batches.

1.2 FINDINGS AND CONCLUSIONS

Conclusion 1: Outcome achievements of 92 percent of completed projects reviewed for FY2010 were rated in the satisfactory range.

12. The Evaluation Office rated the achievement of project outcomes on criteria of relevance, effectiveness, and efficiency. The key findings of this assessment are as follows:

- Of the 48 projects of the FY2010 cohort, 44, or 92 percent, were rated moderately satisfactory or above (Table 1.1). This is similar to the FY 2009 figure of 91 percent and the highest percentage to date.

- The long term average of projects with moderately satisfactory or above in outcome rating is 84 percent. The factors contributing to the higher than average rating of outcomes over the past two years are not clear. Given that the number and type of TE reports submitted varies from year to year, the jump in outcome achievements over the past two years should not be interpreted as a trend.
- Of the total \$177 million GEF investment in the rated projects of FY2010 cohort, 89 percent, or \$157 million, was allocated to projects that were rated moderately satisfactory or above.
- Considering the FY 2005 to FY 2010 cohorts, so far 308 projects have been rated on outcome achievements. Of these, outcome achievements of 84 percent have been rated moderately satisfactory or above. Of the total GEF investment in rated projects, 83 percent is in projects that were rated moderately satisfactory or above.

13. Within the FY 2010 cohort alone, all the UNEP and UNDP implemented projects received outcome ratings of moderately satisfactory or above. Of the World Bank implemented projects 82 percent were rated moderately satisfactory or above. For the first time in this APR, GEFEO has further disaggregated this information of projects by ratings in the satisfactory range for each agency (see Table 3.1 on page 13).

14. Within the ranges of satisfaction, in UNEP's portfolio of 9 projects 56 percent were rated moderately satisfactory and 44 percent satisfactory. Of the 17 UNDP implemented projects, 41 percent were rated moderately satisfactory, 41 percent satisfactory and 18 percent highly satisfactory. Within the World Bank implemented projects, 55 percent were rated moderately satisfactory and 27 percent satisfactory. Due to the blend of MSPs and FSPs in the mix of projects and the small size of each Agency's portfolio, this performance does not indicate a trend.

15. **Sustainability of project outcomes.** This is rated based on the level of risks to sustainability of outcomes on four dimensions: financial, sociopolitical, institutional and governance, and environmental. All of the 48 projects from the FY 2010 cohort were rated on overall sustainability of outcomes¹. The key findings of this assessment are:

- Sixty-three percent (30 projects) were rated moderately likely or above for overall sustainability of outcomes (Table 1.1).
- For the FY2010 cohort, of the four risks to sustainability dimensions, financial risks and environmental risks were more frequently found to pose a threat to outcome sustainability. Out of the 43 projects rated for financial sustainability, 7 (17 percent) were rated as moderately unsatisfactory or below, and out of the 34

¹ Twelve projects had no overall sustainability rating in the Agency TER. The overall rating was applied by GEFEO using the Office's TE review guidelines

projects rated for environmental sustainability, 8 (24 percent) were rated as moderately unsatisfactory or below.

- Out of the total GEF investment in FY2010, 73 percent (US\$130 million out of a total US\$ 177 million) went to 30 projects rated moderately likely or above in terms of the sustainability of their outcomes. This is a considerable increase from the FY2009 level of 64 percent and the 5-year average (2005-2009) of 60 percent.

Table 1.1: Outcomes and Sustainability of Outcomes

<i>Fiscal Year</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>
Number of TE submitted	41	66	41	62	55	48
Number of projects with outcomes rating	39	64	40	62	55	48
Percentage rated MS or above in outcomes rating	82%	84%	73%	81%	91%	92%
Number of projects with sustainability of outcomes rating	39	54	39	60	51	48
Percentage rated ML or above in sustainability of outcomes rating	49%	65%	59%	57%	71%	63%
Number of projects rated both on outcomes and sustainability of outcomes	39	54	39	60	51	48
Percentage of rated projects with MS/ML or above in both	44%	61%	51%	55%	67%	63%

16. The Evaluation Office assessed the extent to which projects that were rated moderately satisfactory or above in achievement of outcomes were also rated moderately likely or above in sustainability of outcomes. It found:

- Of the 48 rated projects, 63 percent were rated **both** moderately satisfactory or above in outcomes and moderately likely or above in sustainability.
- In terms of GEF investment, 73 percent was invested in projects that were rated **both** moderately satisfactory or above in outcomes and moderately likely or above in sustainability.

Conclusion 2: Materialization of cofinancing reported by the GEF Agencies, on average is higher than that expected at project approval.

17. The GEF Council views cofinancing to be an indicator of a project's sustainability², country ownership³, and mainstreaming⁴ of GEF activities in activities of the partner institutions, and way to mobilize additional resources for the global environment⁵. As stated in last year's APR, the Council has shown continued preference for a higher overall cofinancing ratio for the GEF project portfolio. The Secretariat publications also portray cofinancing as an indicator of the effectiveness of GEF in mobilizing additional resources for generation of global environmental benefits. A few documents clearly state that the high cofinancing ratio achieved is an indicator of GEF's "multiplier" effect in generating additional resources for generation of global environmental benefits⁶.

18. For the FY2010 cohort, the overall ratio of promised cofinancing to the GEF grant amount at approval was 2.0. The overall ratio of actual, or materialized, cofinancing to the GEF grant amount at approval was higher at 2.7. The overall percentage of actual cofinancing to promised cofinancing is 134 percent, indicating that actual cofinancing was 34 percent higher than expected. This percentage was slightly higher than the FY2009 level of 132 percent and higher than the 98 percent calculated for the period FY 2005 to FY 2008.

19. A detailed discussion on cofinancing was presented in the Annual Performance Report 2009. The analysis reported that although the quality of reporting on cofinancing both in project proposals and in terminal evaluations is improving, reporting on cofinancing by Agencies and in the TE reports is not always consistent with the agreed definition of cofinancing⁷. Further, usage of cofinancing ratio as a means to track adequacy of cofinancing for the GEF portfolio has its pitfalls as this ratio is prone to be skewed by outliers. For example, of the 313 projects for which terminal evaluations have been submitted since FY2005, agencies have reported data on materialized cofinancing for only 257. Furthermore the top 13 percent of projects out these 257 accounted for close to 50 percent of the total reported materialized cofinancing. A detailed analysis on cofinancing was not undertaken in APR2010.

Project Completion Delays

20. The Evaluation Office began tracking project completion delays (expected completion date and actual completion date) in FY 2005. Of the 281 projects for which this

² Highlights of Council Discussions, Agenda Item 9 – Work Program Submitted for Council Approval, December 1999.

³ Highlights of Council Discussions, Agenda Item 11 – Strategic Approach to Enhance Capacity Building, November 2003.

⁴ Discussions of the Council, Agenda Item 7 – Work Program, April/May 1997.

⁵ Joint Summary of Chairs: Highlights of Council Discussions – Agenda item 9, Cofinancing, October 2002.

⁶ GEF Contributions to Agenda 21, GEF, June 2000 (page 15); GEF... Dynamic Partnerships: Real Solutions, GEF, 2002 (page 6); GEF Annual Report 2005, GEF, 2005, (page 13).

⁷ The agreed definition, adopted in 2003 (GEF/C.20/6/Rev.1), points out several key conditions for a contribution to be considered cofinancing: (i) a contribution should be managed with the GEF allocation; (ii) should be part of the initial financial package; and, (iii) should be essential to achievement of GEF objectives.

data is available, 27 percent were completed after a delay of at least two years and 7 percent after a delay of three years or more. Of the projects in the FY 2010 cohort, 32 percent were completed after a delay of one to two years, 11 percent after a delay of two to three years, and three percent after a delay of more than three years. In comparison to the long-term distribution, the majority of projects of the FY2010 cohort tended to experience shorter delays in completion.

PROJECT MONITORING AND EVALUATION

Conclusion 3: Quality of M&E arrangements is fluctuating with an average of 68 percent of projects being rated moderately satisfactory or above since 2006.

21. Among the projects of FY 2010 cohort, 70 percent (30 out of 44 projects for which we have M&E assessments) were rated moderately satisfactory or above in terms of monitoring during implementation (Table 1.2). This represents an improvement over the 2009 level of 62 percent but a decrease from 72 percent in FY 2008. The fluctuation is present in previous years ratings as well. One reason for the discontinuous performance is that most of the projects in previous cohorts were designed before the adoption of the GEF M&E Policy (2006). As current and future cohorts increasingly include projects designed following the 2006 policy change, we would expect to see continuous improvement in ratings along this dimension. The GEF Council approved a revised M&E Policy in November 2010.

Table 1.2 Quality of monitoring during project implementation

Fiscal Year of TE submission	2004	2005	2006	2007	2008	2009	2010
Terminal evaluation reports submitted	42	41	66	41	62	55	48
Projects rated on M&E	29	32	46	33	50	39	44
Percentage of projects rated moderately satisfactory or above	55%	66%	78%	61%	72%	62%	70%

22. Quality of M&E at entry has also improved in relation to the FY 2009 cohort, from 69 percent to 73 percent. For M&E during implementation, however, 71 percent (22 out of 31 projects), of projects were in the satisfactory, a small decline from the 79 percent in the FY 2009 cohort.

23. Consistent with previous years, there is a strong correlation between quality of M&E arrangements at entry and actual quality of monitoring during implementation. Of the completed projects from the FY2010 cohort that were rated both on quality of monitoring during implementation and quality of M&E at entry, only 18 percent of those rated in the unsatisfactory range at entry were rated in the satisfactory range during implementation. In contrast, of those rated in the satisfactory range at entry, 79 percent were also rated in the satisfactory range during implementation.

Conclusion 4: The quality of 86 percent of the terminal evaluations submitted during FY2010 was rated moderately satisfactory or above.

24. Since 2004, when the Evaluation Office first started rating the quality of terminal evaluation reports, 355 terminal evaluations have been submitted and 351 rated. Of these, the quality of 87 percent of reports has been rated moderately satisfactory or above. Fifty five (55) percent of the reports achieved a higher rating of satisfactory or above. The figures for FY 2010 cohort are 86 percent and 55 percent respectively. This represents a decline from last year's cohort, in which 96 percent of the terminal evaluation reports were rated moderately satisfactory and 72 percent were rated satisfactory or above.

25. The primary reason for the drop this year is that World Bank Medium Sized Projects (MSPs) accounted for over a quarter of all projects in the FY2010 cohort. As the World Bank does not prepare independent terminal evaluation reports for MSPs, the quality ratings for the terminal evaluation documents provided for these projects tend to be low. In FY2009, terminal evaluation reports from World Bank MSPs accounted for less than 4 percent of the sample.

Conclusion 5: No improvement can be reported on the long term lags between Terminal Evaluation Report completion and submission to the GEF Evaluation office.

26. The GEF Agencies are required to submit terminal evaluation reports within 12 months of project completion. As specified in the terminal evaluation guidelines, the GEFEO allows for an additional two months after the terminal evaluation report is complete to submit reports to the GEF Evaluation Office.

27. The Evaluation Office began collecting data on the time lags between project closure and terminal evaluation report in FY 2008. In APR 2009, the Office concluded that long time lags between the completions and submission of terminal evaluation reports continued to be a concern, as did uncertainty regarding project status. For the FY2010 cohort, data was available on the time lags between project closure and terminal evaluation report submission for 41 projects (excluding resubmitted projects). Five (5) percent of these terminal evaluation reports were submitted within 12 months of closure. The majority of reports, 56% were submitted within one to two years following project closure. In the FY 2009 and FY 2008 cohorts, 47 percent of reports were submitted within 12 months of project closure. The sharp decline in this year's cohort is most likely due to the fact that for those instances in which the exact date of submission was unknown, this analysis imputes a date of August 2010.

28. For the FY 2010 cohort, nine percent of the terminal evaluation reports were submitted within two months of completion. For 34 percent, this lag was over two months to one year. And, for the majority, 57 percent, the time lag between report completion and submission was more than one year. Again, in comparison with previous years, particularly FY 2009, this represents a sharp decline. It must be noted once more that for those

instances where the exact date of submission was unknown, the Office imputes a submission date of August 2010, meaning that some lags have been magnified in the analysis.

29. Accurate data on status of project completion is difficult to arrive at and influences the data comparing time lags in terminal evaluation completion and submission. Despite significant improvements in quality of PMIS, the quality of information on project completion status remains weak. Status of projects is not updated in a regular and systematic manner. As a result, it is difficult to determine whether a project has been completed. Resolution of this concern requires collaborative efforts from the Secretariat, Agencies and the Evaluation Office.

30. Council decided in June 2010 that the GEF Evaluation Office, the Secretariat and the Agencies should work together in identifying and implementing measures to improve the quality of information available through PMIS on the status of projects through the project cycle, including agency compliance with deadlines for terminal evaluations. The Evaluation Office was requested to report on the progress made in the Annual Performance Report 2010. On this matter, the GEF Trustee is completing an independent review of GEF systems to be submitted to Council in November 2011. Based on the findings of the review, the Secretariat, in consultation with the Trustee, Evaluation Office and the Agencies will chart a way forward to upgrade the PMIS.

MANAGEMENT ACTION RECORD

31. The Management Action Records (MAR) keep track of the level of adoption of Council's decisions on the basis of evaluation findings and recommendations. This year's MAR tracks the level of adoption of 35 Council decisions, which were based on 16 GEF Evaluation Office documents, by presenting ratings from GEF management and verification of these ratings by the Evaluation Office. The Evaluation Office was able to verify progress on adoption of these 35 decisions.

32. Of the 35 Council Decisions for which adoption was rated, for nine (16 percent) it was rated as 'high'. Three of these decisions pertain to the *'Joint Evaluation of the Small Grants Programme'* and are related to the Council's request to the GEF to propose a level of management costs, processes for management systems and strengthening monitoring and evaluation. GEF has also adopted the Council decision based on *'GEF Annual Report on Impact 2007'* requesting the GEF Secretariat to put in place mechanics along the project cycle to ensure adequate monitoring of progress towards institutional continuity. Similarly GEF action on the decision concerning *'GEF Annual Report on Impact 2009'*, that GEF-5 proposals should include further investment and capacity development to assist countries with economies in transition to address the remaining threats to the ozone layer has been rated "high".

33. Exploration within the GEF partnership modalities to review the significant gap of available resources for combating land degradation to support key challenges is addressed in the new System for Transparent Allocation of Resources (STAR), which

ensures that countries can access funding for land degradation. The level of adoption of the associated decision was rated as ‘high’.

34. Of the 35 Council Decisions for which the level of adoption was rated, 46 percent were rated to be ‘no longer relevant’, eight percent as “substantial’ and 11 percent as ‘medium’. For two decisions (6 percent) the level of adoption was rated to be negligible. The adoption of the Council decision based on the ‘*Country Portfolio Evaluation Report (2009)*’ that the GEF conduct a survey to assess “exceptional situations concerning limited access to GEF partner International Financial Institutions” was rated as negligible because GEF has not yet conducted this survey. The adoption of the Council decision based on the *Annual Country Portfolio Report (2010)* that GEF “provide specific M&E training to the national focal point mechanisms through the country Support Program” was rated negligible because training sessions on M&E have not yet been developed that can be incorporated in the Expanded constituency workshops. This will be incorporated for workshops in 2012.

35. Since its start the MAR has tracked the level of adoption for 98 GEF Council decisions based on 27 evaluations. Of these, 88 have graduated. Of the graduated decisions, 37 percent have been graduated because their adoption was rated ‘high’. These include the nine for which adoption was rated high in FY2010. The remainder had been graduated because they ceased to be relevant. In most instances this has happened due to higher level GEF policy shifts that rendered the earlier decisions irrelevant in the emerging context.

PERFORMANCE MATRIX

36. The performance matrix provides a summary of the performance of the GEF Agencies and GEF Secretariat on six parameters, covering key areas such as results, processes affecting results and M&E. Several of the parameters included in the matrix are assessed by the Evaluation Office on an annual basis. Since performance ratings on these parameters fluctuate from year to year, running averages of two to four years, depending on the parameter, are used in the matrix (see Annex D). Ratings for six parameters are included in the performance matrix this year.

37. The project outcome achievement figures included in the matrix are based on the four year running averages. Of the projects for which terminal evaluations have been submitted since FY 2006, the Evaluation Office rated outcome achievements (parameter 1) of 83 percent to be moderately satisfactory or above. The outcome achievements of 86 percent of the World Bank projects, 90 percent of UNEP projects and 81 percent of UNDP projects have been rated in the satisfactory range (see table 11.1).

RECOMMENDATION

38. The GEF Evaluation Office and the independent evaluation offices of GEF Agencies should strengthen their collaboration on the review of terminal evaluations to

ensure a more streamlined process which will lead to reduction of delays in submission of terminal evaluations and improve the information concerning project status.

ISSUES FOR THE FUTURE

39. The GEF Evaluation Office will assess the reporting systems of new Agencies that enter the GEF partnership and closely support their terminal evaluation review process.

40. The Office will seek ways to improve the use of data on completed projects to find innovative, new and potentially interesting correlations and perspectives.

2. SCOPE AND METHODOLOGY

2.1 SCOPE

41. In the APR, each year the Evaluation Office presents an assessment of the results of completed GEF projects, an analysis of the processes that affect accomplishment of results, and the findings of its oversight of project monitoring and evaluation activities across the portfolio. Through the APR the Evaluation Office provides feedback to the GEF Council, other GEF institutions, and stakeholders to help improve the performance of GEF projects. Some issues are addressed in the APR annually, some biennially; others are addressed whenever such a need is felt.

42. The APR 2010 includes the following:

- **An overview of the extent to which GEF projects are achieving their objectives (chapter 3).** This overview consists of an assessment of the extent to which the completed projects, for which terminal evaluation reports were submitted from FY 2005 to FY 2010, achieved expected outcomes and the risks to sustainability of the achieved outcomes. These issues are covered in the APR annually.
- **Presentation of assessments on some of the factors that affect attainment of project results (chapter 4).** This chapter reports on the extent to which cofinancing promised at the point of project endorsement has materialized; and, on delays in project completion. The assessment on materialization of cofinancing is based on figures reported by the respective GEF Agencies.
- **An assessment of quality of project monitoring in completed projects (Chapter 5).** This chapter presents an assessment of quality of project monitoring in completed projects for which terminal evaluation reports have been submitted to the Office since FY 2008.

- **An assessment of the quality of terminal evaluation reports submitted by the GEF Agencies to the Evaluation Office (chapter 6).** This chapter provides information on the quality of terminal evaluation reports by agency and an assessment of trends in the quality of terminal evaluation reports. It also includes a comparison of the ratings on outcome achievements for assessed projects given by the Evaluation Office with those provided by others including the evaluation offices of the Implementing Agencies, in the terminal evaluations, and in project implementation reports submitted by the Implementing Agencies.
- **A presentation of findings on management action records (chapter 7).** As part of this annual assessment, the Evaluation Office reviews and follows up on the implementation status of evaluation recommendations that have been accepted by management and/or the GEF Council.
- **A presentation of the performance matrix (chapter 8).** This assessment was first presented in APR 2007. Some of the indicators previously reported on have been fine tuned so as to focus on key performance parameters.

3. OUTCOMES AND SUSTAINABILITY

43. This chapter discusses verified ratings on outcomes and sustainability for 313 projects, for which terminal evaluation reports have been submitted from FY 2005 to FY 2010. These include 168 full size projects, including two enabling activities, and 145 medium size projects. Given the small number of enabling activities for which terminal evaluations are required, they are reported as either full size projects or medium size projects based on the level of GEF funding. Altogether, the GEF has invested approximately US\$1.4 billion dollars in these 313 projects.

44. Of the 313 projects, 48 terminal evaluations were submitted during FY2010. These include 23 full-size and 25 medium-size projects. The GEF has invested a total of US\$176.8 million in these 48 projects.

3.1 RATING APPROACH

45. Of the 48 projects for which terminal evaluations were submitted during FY2010, the GEF Evaluation Office accepted the outcome ratings provided by the evaluation offices of the respective agencies for 21 projects. For the remaining 27 projects, the GEF Evaluation Office rated project outcomes based on level of achievement of project objectives and expected outcomes in terms of **relevance** on a two point scale and **effectiveness** and **efficiency** on a six-point scale. The rating approach followed is identical to that followed for the 2009APR.

46. The assessment on relevance primarily focuses on determining whether the anticipated outcomes are relevant to the GEF mandate for generating global environmental benefits. Only **satisfactory** or **unsatisfactory** ratings were assigned for

relevance. Relevance ratings are considered critical: if the relevance of outcomes rating is unsatisfactory then the overall outcome rating cannot be higher than unsatisfactory. Among the other criteria, effectiveness is an important criterion; the overall rating on achievement of outcomes cannot be higher than the rating on effectiveness.

47. During project implementation, the results framework of some projects had been modified. This poses a challenge because assessing actual outcomes for all projects based on original outcome expectations may discourage adaptive management. To address this challenge, for projects where modifications and improvements are made in the project objectives, outcomes and outputs, without a down scaling of their overall scope, the Office assesses outcome achievements based on the revised results framework. In instances where the scope of the project objectives, outcomes and outputs has been downscaled, the expected outcomes and/or the original objectives of the project are used as a yard stick for performance assessment.

48. Among the 313 terminal evaluation reports reviewed in the period FY2005-FY2010, five did not provide sufficient information to allow the Evaluation Office to rate outcome achievements.

49. The GEF Evaluation Office rates sustainability of outcomes based on an assessment of four key risk dimensions: financial, sociopolitical, institutional framework and governance, and environmental. Based on the evidence presented in the terminal evaluation reports, **risks to sustainability of outcomes** were assessed on each of these dimensions. All risk dimensions were regarded as critical; overall ratings cannot be higher than the rating for a dimension that had been rated the lowest.

3.2 OUTCOMES

50. Ninety-two percent of the projects of the FY2010 cohort have been rated moderately satisfactory or above in terms of their outcome achievements. This percentage is higher than the long-term average of 84 percent (table 3.1). However, since the percentage of projects with outcomes rated in the satisfactory range differs considerably from year to year, this improved performance does not indicate a trend.

Table 3.1 Distribution of GEF Projects: by Outcome ratings

Outcome Rating	FY2005	FY2006	FY2007	FY2008	FY2009	FY2010	Total Average
<i>Highly Satisfactory</i>	3%	6%	3%	5%	4%	6%	4%
Satisfactory	54%	44%	35%	52%	56%	35%	46%
Moderately Satisfactory	26%	34%	35%	24%	31%	50%	33%
Moderately Satisfactory or above	82%	84%	73%	81%	91%	92%	84%
Moderately unsatisfactory	10%	14%	8%	13%	9%	4%	10%

Unsatisfactory	8%	2%	18%	5%	0%	4%	5%
Highly Unsatisfactory	0%	0%	3%	2%	0%	0%	1%
Moderately unsatisfactory or below	18%	16%	27%	19%	9%	8%	16%
<i>Factor</i>	<i>Number</i>						<i>Total</i>
Terminal evaluation reports submitted	41	66	41	62	55	48	313
Projects rated on outcomes	39	64	40	62	55	48	308

Table 3.2 Distribution of GEF Investment (in US\$ million): by Outcome Ratings

<i>Outcome Rating</i>	<i>FY2005</i>	<i>FY2006</i>	<i>FY2007</i>	<i>FY2008</i>	<i>FY2009</i>	<i>FY2010</i>	<i>Total Average</i>
Highly Satisfactory	0%	6%	5%	8%	3%	3%	4%
Satisfactory	64%	30%	18%	48%	56%	48%	45%
Moderately Satisfactory	20%	53%	46%	12%	33%	38%	34%
Moderately Satisfactory or above	84%	88%	69%	74%	92%	89%	83%
Moderately unsatisfactory	15%	11%	14%	13%	8%	8%	11%
Unsatisfactory	1%	1%	12%	10%	0%	3%	5%
Highly Unsatisfactory	0%	0%	5%	3%	0%	0%	1%
Moderately unsatisfactory or below	16%	12%	31%	26%	8%	11%	17%
<i>Investment</i>	<i>Millions (\$)</i>						<i>Total</i>
Total GEF Investment in reviewed projects	258.3	248.8	199.3	275.3	207.8	176.8	1,373.33
Total GEF Investment in rated projects	248.3	254.3	198.3	275.3	207.8	176.8	1,367.83

51. The GEF has invested US\$1,367 million in projects that have received ratings on their outcome achievements. Of this investment, 83 percent is invested in projects whose outcome achievements were rated moderately satisfactory or above. Of the total GEF investment in the FY 2010 cohort, 89 percent, or \$157 million, was invested in projects whose outcome achievements were rated moderately satisfactory or above (table 3.2). Chapter 4 presents a detailed discussion on drivers of project outcome achievements.

52. Table 3.3 provides information on the outcome ratings of different categories of GEF projects based on Implementing Agency; the agency type for project execution; focal areas; project size; geographical scope; geographical region; and, important country groupings, for all projects rated since FY 2005. Outcomes of 81 percent of the FSPs and

87 percent of the MSPs were rated moderately satisfactory or above. The outcome ratings of World Bank implemented projects tend to be slightly higher than other agencies. Compared to other categories a significantly lower percentage of national projects implemented in African countries are rated moderately satisfactory or above.

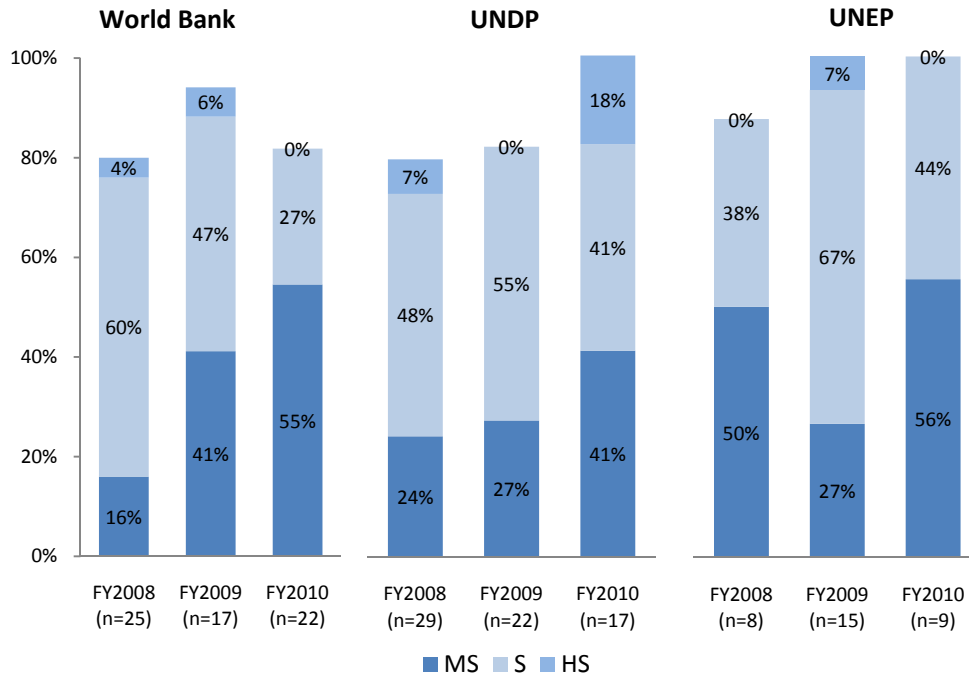
Table 3.3 Outcome performance: by categories of projects

<i>Category</i>		<i>Number reviewed</i>	<i>Number rated</i>	<i>% Rated MS or above</i>
All projects		313	308	84%
GEF Agency	World Bank	134	132	86%
	UNDP	123	122	82%
	UNEP	54	53	85%
Executing Agency	Government or parastatal agency	162	160	83%
	Nongovernmental organization or foundation	68	67	85%
	Bilateral or multilateral agency	62	60	83%
	Other, including private sector organization	21	21	90%
Focal Area	Biodiversity	163	162	84%
	Climate change	81	79	84%
	International waters	37	37	86%
	Other	32	30	83%
Size	FSP	168	16	81%
	MSP	145	14	87%
Scope	National (single-country project)	221	21	84%
	Regional or global	92	90	83%
Location	Africa	67	66	76%
	Asia	67	65	85%

Europe and Central Asia	57	57	82%
Middle East and North Africa	15	15	93%
Latin America and the Caribbean	75	74	88%

53. Figure 3.1 below shows the breakdown of projects by ratings in the satisfactory range for each agency. Within the FY 2010 cohort alone, all the UNEP and UNDP implemented projects received outcome ratings of moderately satisfactory or above. Of the World Bank implemented projects 82 percent were rated moderately satisfactory or above. In UNEP's portfolio of 9 projects, 56 percent were rated moderately satisfactory and 44 percent satisfactory. Of the 17 UNDP implemented projects, 41 percent were rated moderately satisfactory, 41 percent satisfactory and 18 percent highly satisfactory. Within the World Bank implemented projects, 55 percent were rated moderately satisfactory and 27 percent satisfactory. Due to the blend of MSPs and FSPs in the mix of projects and the small size of each Agency's portfolio, this performance does not indicate a trend.

Figure 3.1 Outcome rated MS or above, by Agency (FY2008-FY2010)



3.3 SUSTAINABILITY

54. Of the 313 terminal evaluation reports submitted from FY 2005 to FY 2010, sustainability ratings have been provided for 291 (93 percent) of the projects. Since FY2005, of the 291 projects rated on sustainability, 167 projects (57 percent) were rated moderately likely or above. For the FY2010 cohort of 48 projects, outcome sustainability ratings have been provided for all projects. For 12 out of 48 projects, overall sustainability ratings were missing (as they had not been given by the Agencies, whose ratings were in some cases accepted by the GEF EO), for which they were rated according to the GEF EO guidelines. Of the FY2010 cohort, 30 out of the 48 projects (63 percent) were rated moderately likely or above on overall sustainability. This is higher than the long-term average of 57 percent.

55. In terms of GEF investments, for 291 rated projects since FY 2005, an amount of US\$ 0.79 out of US\$ 1.37 billion of investments (58 percent) was in 177 projects with sustainability rated moderately likely or above. For the FY2010 cohort only, an amount of US\$ 130 out of US\$ 177 million went to 30 projects rated moderately likely or above (73 percent).

56. When both the outcome ratings and the sustainability ratings are taken together into account, of the 291 projects rated since FY 2005, 165 projects (57 percent) were rated both moderately satisfactory or above in terms of outcomes and moderately likely or above in terms of sustainability. Of the FY 2010 cohort only, 30 projects (63 percent) had been so rated – the same amount of projects with sustainability ratings of moderately likely or above.

57. Regarding GEF investment, an amount of US\$ 0.77 out of US\$ 1.77 (43 percent) was in 175 projects that were rated both MS or above in terms of outcome achievements and moderately likely or above in terms of sustainability. For the FY 2010 Cohort only, US\$ 130 million out of US\$ 177 million went to projects rated moderately satisfactory or above for outcomes and moderately likely or above for sustainability (73 percent) – the same amount of projects with sustainability ratings of moderately likely or above.

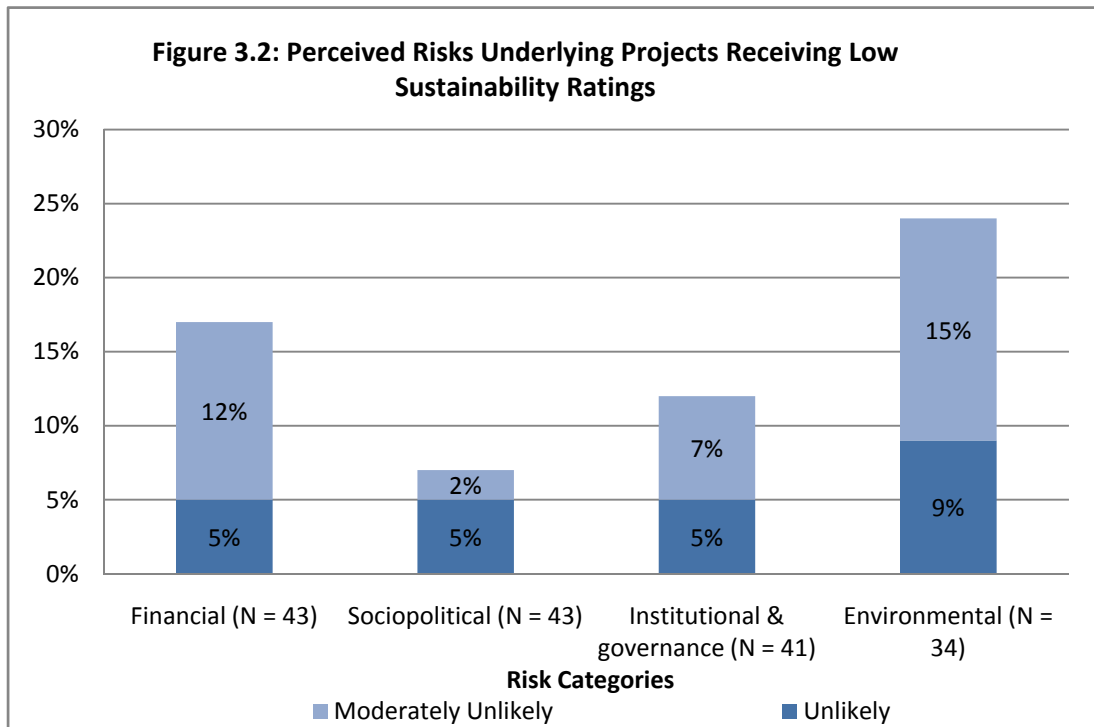
Table 3.4 Sustainability and Outcome Achievements

Category	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	Total
By Project							
Sustainability of Outcomes rated ML or above	49%	65%	59%	57%	71%	63%	57%
Outcomes rated MS or above and Sustainability of Outcomes rated ML or above	44%	61%	51%	55%	67%	63%	57%

By GEF Investment							
Sustainability of Outcomes rated ML or above	65%	60%	55%	58%	64%	73%	58%
Outcomes rated MS or above and Sustainability of Outcomes rated ML or above	60%	56%	44%	56%	63%	73%	43%

58. For the FY 2010 cohort, among the four sustainability dimensions, the outcomes of five percent were unlikely and of 12 percent were moderately unlikely to be sustained due to financial risks. The outcomes achievements of five percent were unlikely and seven percent were moderately unlikely to be sustained due to institutional and governance risks. Out of 43 projects rated for socio-political risks, five percent were unlikely and two percent was moderately unlikely to be sustained. Out of 34 projects rated for environmental sustainability, nine percent were unlikely and 15 percent were moderately unlikely to be sustained (figure 3.2).

59. For the FY2010 cohort, of the four risks to sustainability dimensions, financial risks and environmental risks were more frequently found to pose threat to outcome sustainability. This contrasts partially with FY2009 cohort, in which, besides the persisting financial risks, institutional risks also constituted the most frequent threat to sustainability. According to the last APR, “separately, financial risks and institutional risks posed a threat to the sustainability of outcomes for 19 percent of projects”. In the FY2010 cohort, out of the 43 projects rated for financial sustainability, 17 percent were rated as moderately unlikely or below, and out of the 34 projects rated for environmental sustainability, 24 percent were rated as moderately unlikely or below.



4. FACTORS AFFECTING ATTAINMENT OF PROJECT RESULTS

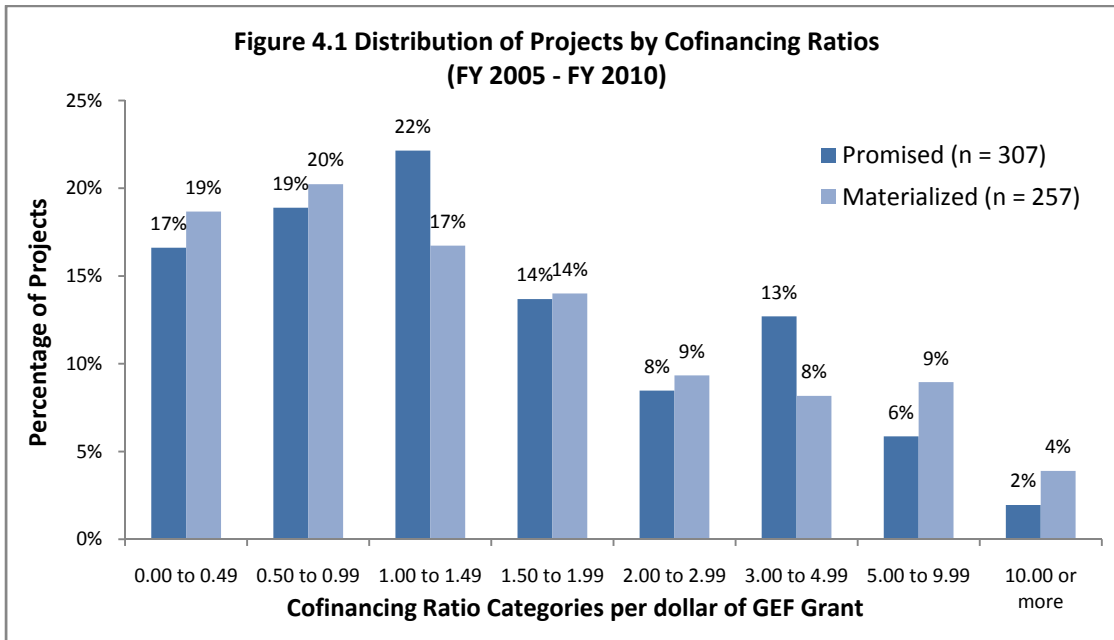
60. Project outcomes may be affected by factors such as project design, quality of project implementation and execution, the operational context in which projects are implemented and executed, and exogenous factors beyond the control of project management. Given the wide range of variables that may affect project outcomes and their interactions, it is difficult to isolate variables and determine their specific effects on a project's results. However, associations among variables and results can be determined. The Evaluation Office has been reporting on some of these variables such as cofinancing and delays in project completion annually. This chapter includes a brief discussion on materialization of cofinancing reported by the implementing agencies (World Bank, UNDP, and UNEP) and on implementation delays. In addition, two other chapters included in this report – quality of supervision and agency fees – cover factors that may affect attainment of project results.

4.1 MATERIALIZATION OF COFINANCING

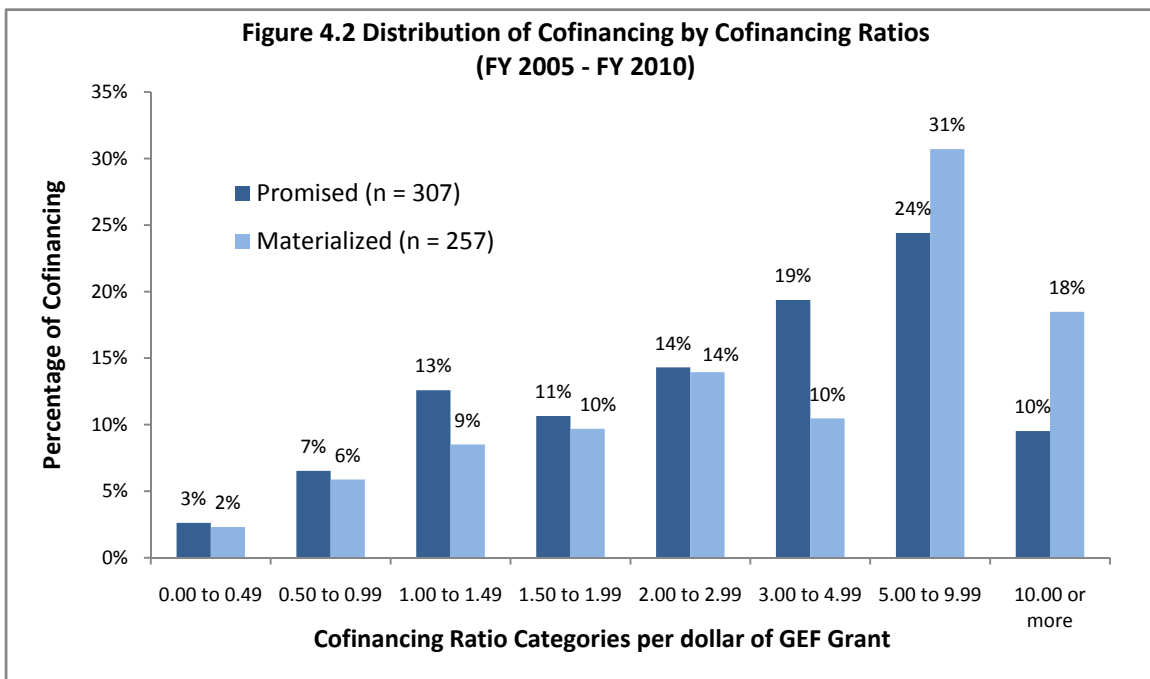
61. The Office reports on materialization of cofinancing in completed projects every year based on the information presented in the terminal evaluation reports and communicated by the agencies through other project reports. However, issues such as the level of global environmental benefits that accrue from cofinancing, the processes through which these benefits accrue, and the costs incurred by the GEF partnership in mobilizing cofinancing, are not assessed in this APR. An assessment of the GEF approach to cofinancing was presented in APR2009.

62. The figures reported by the Office in this section are based on the cofinancing materialization figures reported by the agencies. The level of cofinancing differs considerably from project to project. Since the focus of reporting by the Office had been on averages, overall figures for the portfolio tended to be skewed by a few projects that entailed high level of cofinancing. In this APR, although the Office continues to report the average figures for the portfolio, the focus has shifted to reporting frequency distribution of the levels of cofinancing achievements because findings on this indicator are less likely to be skewed by outliers. The analysis presented in this section is based on the information available on 313 completed projects for which terminal evaluations were submitted since FY 2005. Of these, for 257 projects the information on cofinancing has been reported by the agencies.

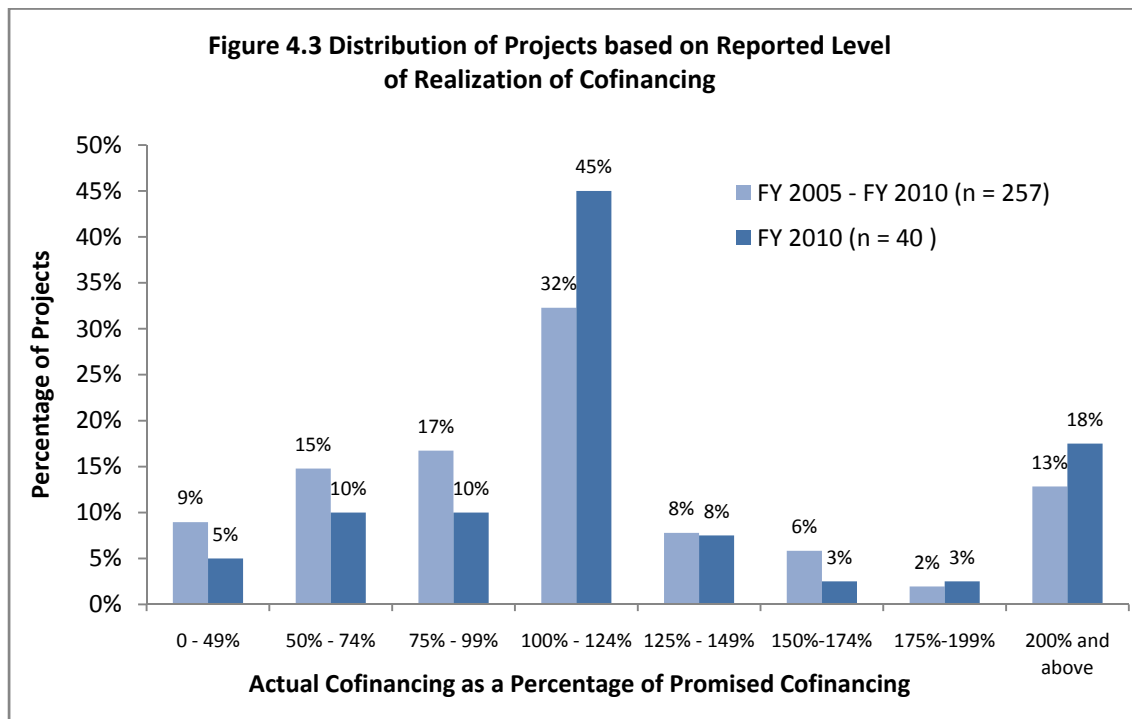
63. Figure 4.1 shows the distribution of projects based on the ratio of cofinancing promised and reported to have materialized vis-à-vis the approved GEF grant. For 72 percent of the projects less than two dollars of cofinancing was promised per dollar of GEF grant. Cofinancing of five dollars or more per GEF dollar was promised for 8 percent of the projects. The distribution of projects based on the ratio of materialized cofinancing vis-à-vis approved GEF grant shows a similar pattern. Agencies reported that for 70 percent of projects the ratio of materialized cofinancing vis-à-vis GEF grant was less than two; and for 13 percent it was five dollars or more.



64. Figure 4.2 presents percentage of cofinancing in each of the ratio categories based on the amount of cofinancing promised or materialized for projects in that category. The two graphs together clearly show that even though projects with cofinancing of five dollars or more are relatively fewer in number they account for over 50% of total actual cofinancing.



65. Figure 4.3 presents a comparison of the frequency distribution of the long-term average and the FY 2010 cohort based on the level of realization of cofinancing reported by the agencies vis-à-vis cofinancing promised at inception. Of the 257 projects for which data is available, 24 percent were reported to have realized less than 75 percent of the promised cofinancing. For the FY 2010 cohort this figure was 15 percent. Likewise, compared to the long-term average, a higher percentage of the FY 2010 projects was reported to have met or exceeded the level of expected cofinancing.



66. Table 4.1 presents key statistics on materialization of cofinancing by agency during different terminal evaluation submission periods. The projects of the FY2009 and FY2010 cohorts in general had greater level of materialization of cofinancing vis-à-vis other cohorts. For FY 2010, on average the projects implemented by UNDP and UNEP are reported to have achieved considerably higher than expected level of cofinancing.

Table 4.1 Materialization of Cofinancing: By Agency⁸

<i>Period</i>	<i>World Bank</i>	<i>UNDP</i>	<i>UNEP</i>	<i>All Agencies</i>
<i>Pre OPS4 Period (FY 2002 to FY 2004)</i>				
Number of projects for which Cofinancing data is available	31	11	6	48
Promised cofinancing per dollar of approved GEF grant	6.5	3.6	1.1	5.9

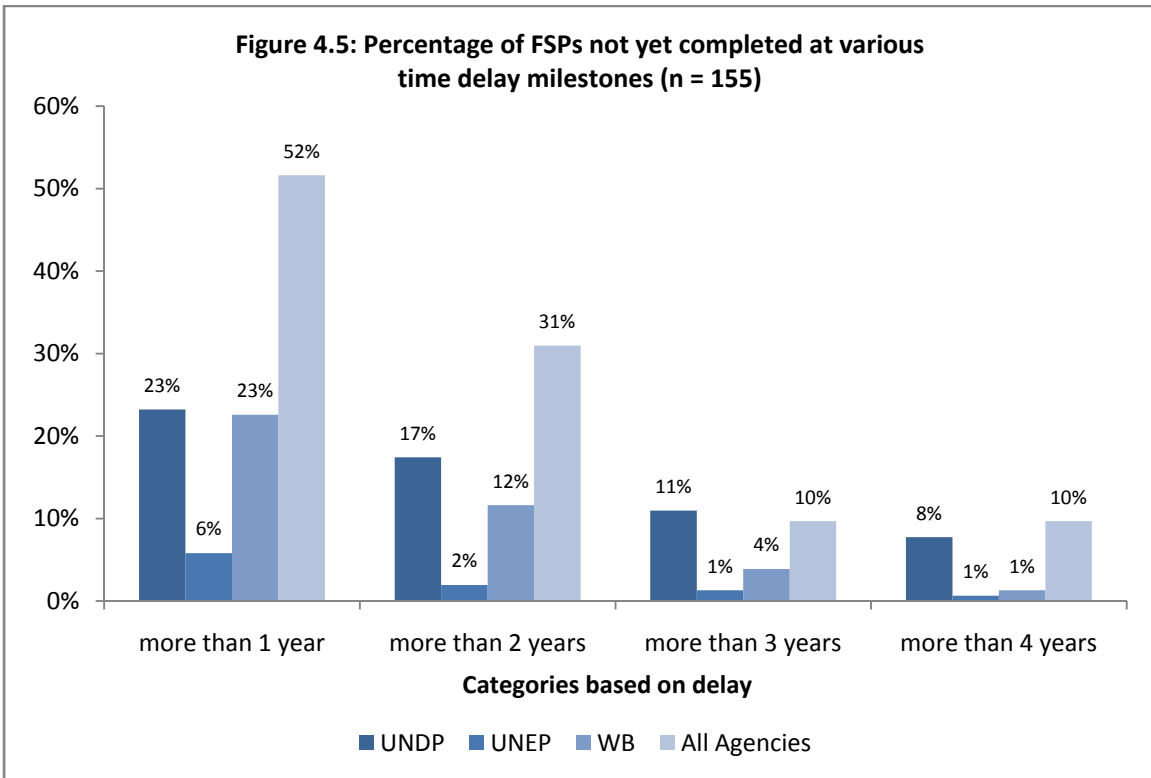
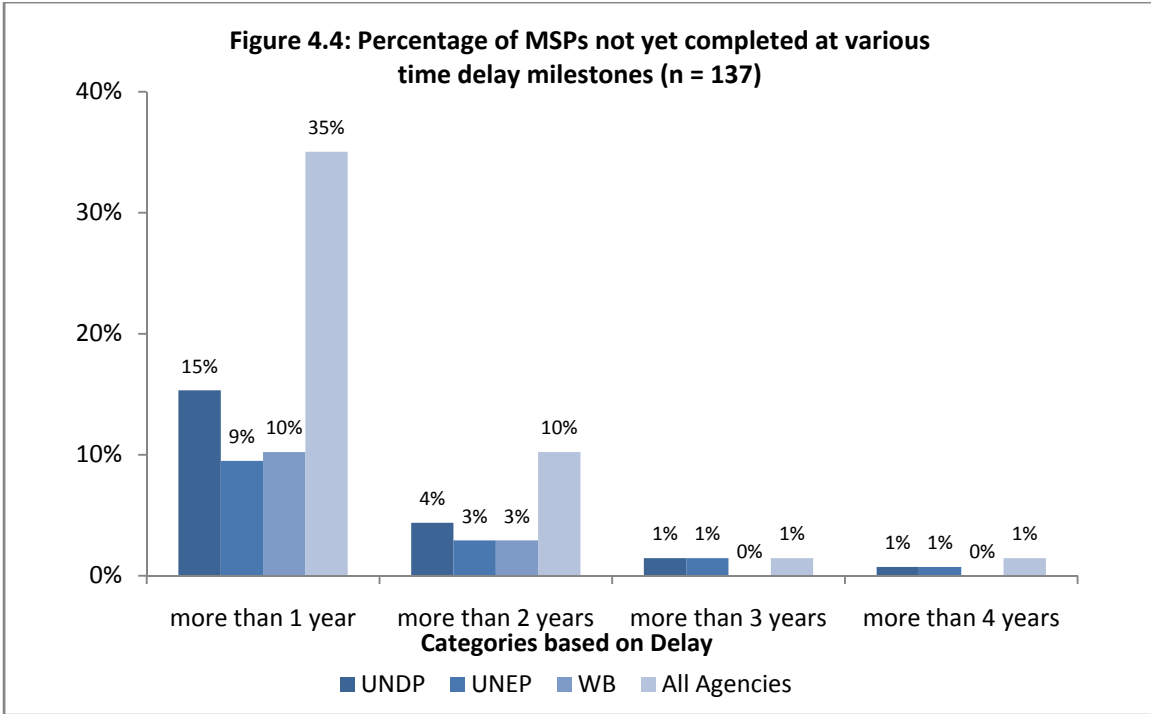
⁸ Joint projects have been attributed to the lead implementing agency.

Reported materialized cofinancing per dollar approved of GEF grant	6	2.8	1	5.4
Materialized cofinancing per dollar of promised cofinancing	93%	78%	87%	92%
<i>OPS 4 Period (FY 2005 to FY 2008)</i>				
Number of projects for which Cofinancing data is available	80	59	23	162
Promised cofinancing per dollar of approved GEF grant	2.4	1.6	1.5	2.1
Reported materialized cofinancing per dollar approved of GEF grant	2.3	1.6	1.5	2
Materialized cofinancing per dollar of promised cofinancing	98%	99%	101%	98%
<i>FY 2009</i>				
Number of projects for which Cofinancing data is available	17	22	15	55
Promised cofinancing per dollar of approved GEF grant	2.3	2.9	0.9	2.3
Reported materialized cofinancing per dollar approved of GEF grant	2.3	4.8	1.2	3
Materialized cofinancing per dollar of promised cofinancing	101%	163%	141%	132%
<i>FY 2010</i>				
Number of projects for which Cofinancing data is available	20	16	4	40
Promised cofinancing per dollar of approved GEF grant	1.9	3.7	0.7	2.0
Reported materialized cofinancing per dollar approved of GEF grant	1.9	6.5	1.3	2.7
Materialized cofinancing per dollar of promised cofinancing	102%	176%	172%	134%

4.2 DELAYS IN PROJECT COMPLETION

67. The Evaluation Office tracks the time difference between expected completion date at project start and actual operational completion of the project to evaluate the extent to which projects supported by GEF are being completed in a timely manner. The data on this indicator is available for projects that have been submitted since FY2005. Important causes of delay have been reported in detail in APR 2007 and, therefore, will not be discussed in this APR.

68. In earlier APRs the central tendencies in completion delays have been reported primarily in terms of averages. In this year's APR the focus of the reporting will be on frequency distributions to facilitate easy assessment of the extent of the delay. Figure 4.4 presents the distribution of MSPs in categories based on time milestones. The figure shows that implementation of 10 percent of MSPs was completed after a delay of more than two years.



69. Figure 4.5 presents the distribution of FSPs in terms of completion delays. Overall a greater percentage of FSPs are completed with delays and experience longer delays: more than 31 percent of FSPs are completed with a delay of more than two years. This is

understandable because FSPs are generally expected to be under implementation for a longer period of time and are more complex in terms of project design. The delay pattern across the World Bank and UNDP is quite similar in that more FSPs experience delays than MSPs. However, the pattern is somewhat different for UNEP, with a higher percentage of UNEP MSPs than FSPs experiencing delays of up to 3 years.

5. QUALITY OF PROJECT MONITORING

70. A project's monitoring and evaluation (M&E) system provides information early on about progress towards achievement of its intended results. It also helps in identification of issues that warrant corrective measures in order to facilitate progress. The Evaluation Office reports on quality of project monitoring in completed projects annually.

5.1 RATING APPROACH

71. Quality of project monitoring in completed projects was assessed on a six-point scale. All the 307 terminal evaluation reports submitted since FY 2004 were considered. Of these for 230 the Evaluation Office has provided or adopted ratings on quality of monitoring during implementation. To rate quality of monitoring in completed projects, it was assessed whether:

- an M&E system was in place and facilitated timely tracking of results and progress toward project objectives by collecting information on chosen indicators continually throughout the project implementation period
- annual project reports were complete and accurate, with well-justified ratings
- the information provided by the M&E system was used for project management
- the parties responsible for M&E activities were properly trained to ensure that correct procedures were followed and quality was maintained in data collection

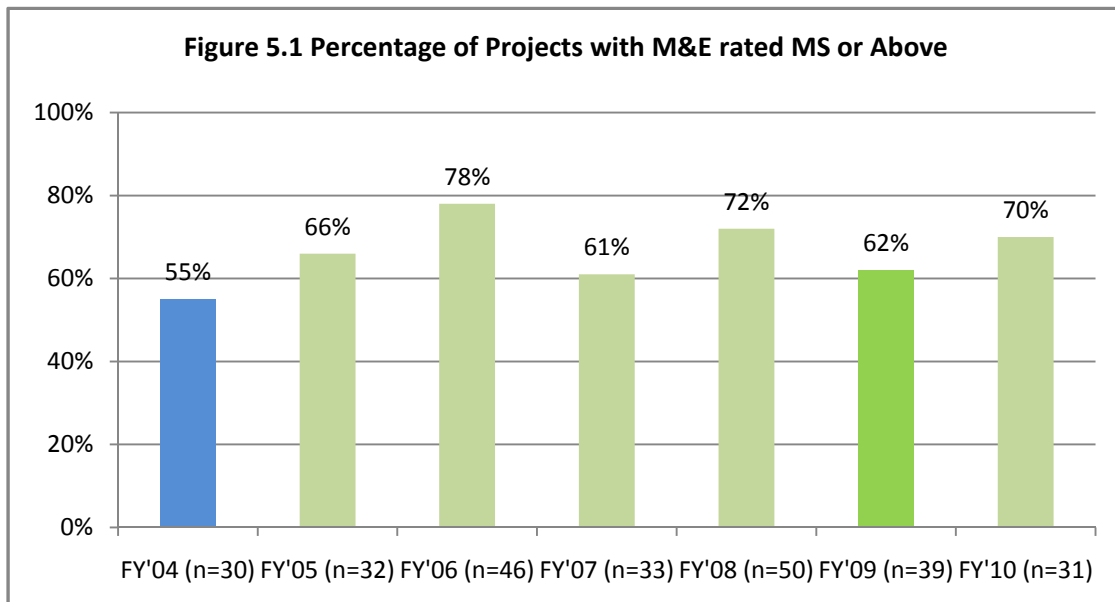
5.2 QUALITY OF M&E

72. Out of 212 rated projects on M&E since FY2006, 68 percent (145 projects) were moderately satisfactory or above. Out of 44 rated projects on M&E in the FY 2010 cohort, 70 percent (31 projects) were moderately satisfactory or above. The result for FY 2010 is not significantly different from the long-term average, but represents an improvement in relation to FY 2009, with ratings in the satisfactory range increasing from 62 percent to 70 percent. Quality at entry has also improved in relation to the last cohort, from 69 percent to 73 percent. For M&E during implementation, however, despite the satisfactory range assessment of 22 out of 31 projects (71 percent), the percentage declines in relation to the last cohort (79 percent).

73. Considering the trend since FY 2005, M&E ratings in the satisfactory range have fluctuated between the sixties and the seventies in percentage terms. In the FY 2010 cohort, 70 percent of projects were rated so. M&E ratings in the mostly satisfactory range have fluctuated between the sixties and the seventies in percentage terms with 72 percent in FY 2008, 61 percent in FY 2007, 78 percent in FY 2006 and 66 percent in 2005. This represents an improvement over the 2009 level of 62 percent but a decrease from 72 percent in FY 2008. The fluctuation is present in previous year's ratings as well (Figure 5.1). One reason for the discontinuous performance is that most of the projects in previous cohorts were designed before the adoption of the GEF M&E Policy (2006). As current and future cohorts increasingly include projects designed following the 2006 policy change, we would expect to see continuous improvement in ratings along this dimension. The GEF Council approved a revised M&E Policy in November 2010.

74. Overall M&E ratings are directly related to the M&E ratings at entry. Of the projects for which quality of M&E arrangements at entry is rated in this range, quality during implementation is also rated in the satisfactory range for 79 percent since FY 2004, and 72 percent for FY 2010 only. In comparison, of projects rated in the unsatisfactory range for quality of M&E arrangements at entry, the quality of monitoring was rated in satisfactory range for only 36 percent since FY 2004, and 18 percent for FY 2010 only. This is consistent with the conclusions presented in the earlier APRs – APR 2006, APR 2007, APR 2008 and APR 2009 – that for the projects that have weak M&E arrangements at entry it is less likely that sufficient corrections will be made to improve their quality of monitoring during implementation.

75. Following predictions of APR 2009, as projects designed after adoption of the 2006 *Monitoring and Evaluation Policy* of GEF are being completed, some improvement in performance both in terms of arrangements at entry and actual implementation of monitoring may be expected.



6. QUALITY OF TERMINAL EVALUATIONS

76. Terminal evaluations provide an assessment of project accomplishments and shortcomings. They are the building blocks for the assessment of performance of completed projects. Their effectiveness as a learning tool for the GEF partnership may be compromised if the information they provide is inaccurate, incomplete, or biased. The Evaluation Office reviews terminal evaluations to provide verified ratings on project performance and on the quality of terminal evaluation reports. By assessing the quality of the terminal evaluation reports, the Evaluation Office identifies the areas where the reports could be improved.

77. To date, 397 terminal evaluation reports have been submitted to the GEF Evaluation Office. This is the seventh year the Evaluation Office has rated the quality of these reports; 351 have been rated thus far. In FY 2010, 48 terminal evaluation reports were submitted by the GEF Agencies, and 44 were rated on quality, including 19 for which the ratings provided by the independent evaluation offices of the respective agencies have been adopted.

78. A major obstacle has been the uncertainty in the number of projects completed and terminal evaluations expected during the annual reporting cycle. Accurate data on status of project completion is difficult to arrive at and influences the data comparing time lags in terminal evaluation completion and submission. Despite significant improvements in quality of PMIS, the quality of information on project completion status remains weak. Status of projects is not updated in a regular and systematic manner. As a result, it is difficult to determine whether a project has been completed. Resolution of this concern requires collaborative efforts from the Secretariat, Agencies and the Evaluation Office.

79. Council decided in June 2010 that the GEF Evaluation Office, the Secretariat and the Agencies should work together in identifying and implementing measures to improve the quality of information available through PMIS on the status of projects through the project cycle, including agency compliance with deadlines for terminal evaluations. The Evaluation Office was requested to report on the progress made in the Annual Performance Report 2010. On this matter, the GEF Trustee is completing an independent review of GEF systems to be submitted to Council in November 2011. Based on the findings of the review, the Secretariat, in consultation with the Trustee, Evaluation Office and the Agencies will chart a way forward to upgrade the PMIS.

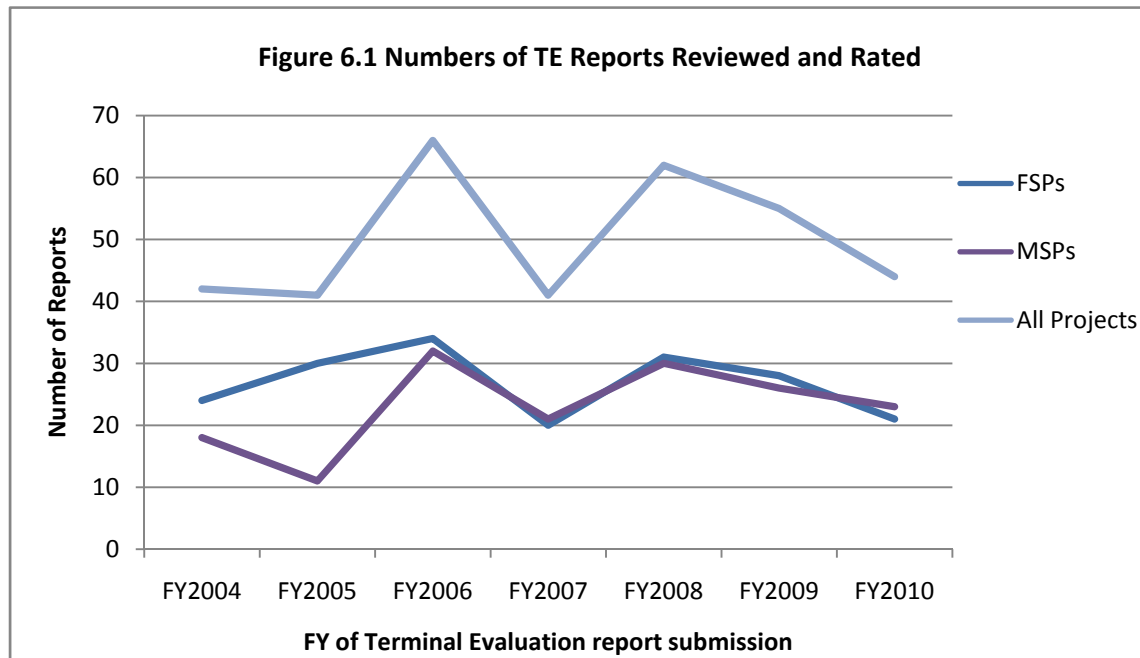
80. The overall quality of 86 percent of the terminal evaluation reports of the FY2010 cohort was rated moderately satisfactory or above.

81. During the annual review cycle for APR2007, the Evaluation Office initiated field verification of terminal evaluations to confirm the achievements results of completed projects. So far 14 verifications have been conducted including one terminal evaluation that was field verified for the FY2009 cycle. No field verifications were conducted for

the FY2010 cohort. Although in some instances the ratings provided after field verifications vary from those provided through desk reviews, generally they confirm the findings of the desk reviews.

6.1 RATING APPROACH

82. The approach adopted by the Evaluation Office to assess the quality of terminal evaluation reports submitted from FY 2005 to FY 2010 has remained the same.



83. The reports submitted since FY 2005 (figure 6.1) have been assessed by the Evaluation Office based on the following criteria:

- Did the report present an assessment of relevant outcomes and achievement of project objectives in the context of the focal area program indicators, if applicable?
- Was the report consistent, the evidence complete and convincing, and the ratings substantiated when used?
- Did the report present a sound assessment of sustainability of outcomes?
- Were the lessons and recommendations supported by the evidence presented?
- Did the report include the actual project costs (total and per activity) and actual cofinancing used?
- Did the report include an assessment of the quality of the project M&E system and its use in project management?

84. Performance on each of these criteria is rated on a six-point scale. The overall rating is a weighted average of these ratings: the first two criteria are given a weight of 0.3 each, and the remainder a weight of 0.1 each.

85. The Evaluation Office also tracks consistency between the its own verified project outcome ratings and those provided by (i) the last project implementation report (PIR) that was submitted to the GEF Secretariat, (ii) the terminal evaluation report, and (iii) the evaluation offices of the implementing agencies. Since not all ratings have been provided on the six point scale used by the Office, to make comparisons possible ratings are converted to a binary scale.

6.2 FINDINGS

86. Of the 48 terminal evaluation reports submitted during the FY2010, the EO was able to rate the quality of 44. Eighty-six percent (38 reports) of these were rated moderately satisfactory or above (see table 6.1). 86 percent of the FSPs and 92 percent of MSPs in the FY 2010 cohort were rated moderately satisfactory or above. Compared to previous years, this year's ratings are lower (see table 6.2). In FY 2009, 96 percent of the terminal evaluation reports were rated moderately satisfactory or above and in FY 2008, 92 percent.

87. The primary reason for the drop this year is that World Bank Medium Sized Projects (MSPs) accounted for over a quarter of all projects in the FY2010 cohort. As the World Bank does not prepare independent terminal evaluation reports for MSPs, the quality ratings for the terminal evaluation documents provided for these projects tend to be low. In FY2009, terminal evaluation reports from World Bank MSPs accounted for less than 4 percent of the sample.

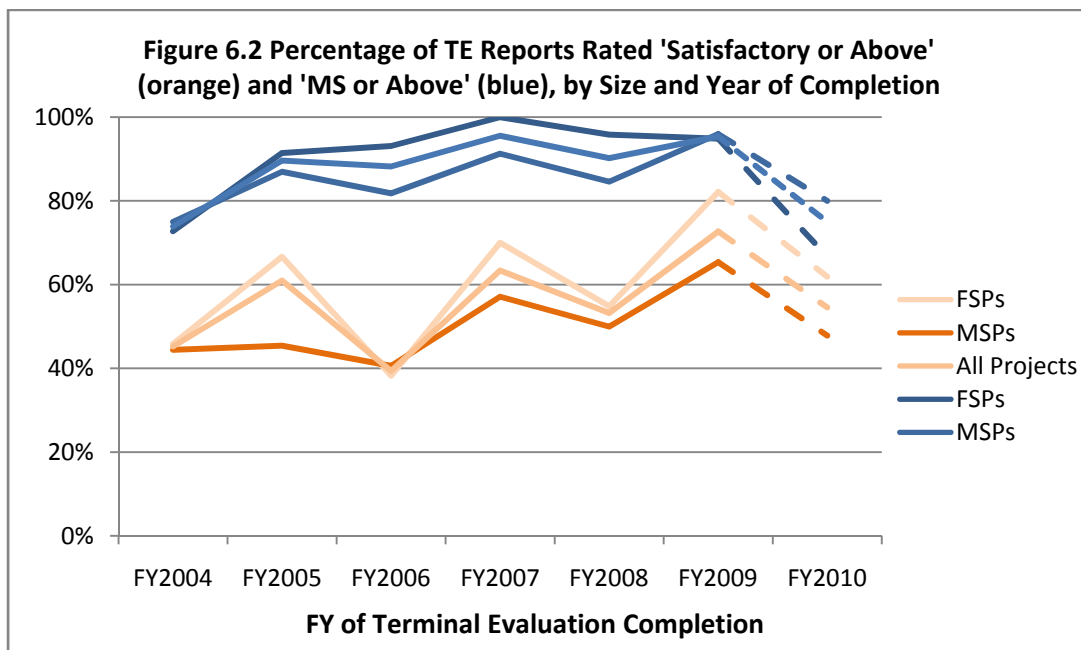
88. Figure 6.2 present the trends in terminal evaluation report quality ratings based on the year when the terminal evaluation reports were completed (instead of the year of submission). Figure 6.2 presents the percentage of reports that were rated moderately satisfactory or above in terms of quality *and* the percentage of reports that meet a higher quality standard of satisfactory or above quality rating. Since all reports for some cohorts, especially the most recent one, have not yet been submitted, ratings for as-of-yet-unsubmitted reports are not known. The dotted lines in the figures indicate the period for which it is estimated that a significant percentage of the terminal evaluations completed in that year have not yet been submitted. Both sets of lines show identical trend in improving TE report quality. The annual variations are more accentuated when the yard stick of 'satisfactory or above' quality ratings is used.

Table 6.1 Terminal Evaluation Reports Rated MS or above, by Project Size and Agency

FY 2010 Agency	FSPs		MSPs		All Projects	
	Number Rated	Percent Rated MS or above	Number Rated	Percent Rated MS or above	Number Rated	Percent Rated MS or above
UNDP	8	75%	9	100%	17	88%
UNEP	4	100%	1	100%	5	100%
World Bank	9	89%	13	77%	22	82%
All Agencies	21	86%	23	87%	44	86%

Table 6.2 Percentage of Terminal Evaluation Reports Rated MS or Above, by Year of Submission and Agency

Agency	FY2004	FY2005	FY2006	FY2007	FY2008	FY2009	FY2010	All years
UNDP	54%	91%	87%	94%	93%	95%	88%	88%
UNEP	67%	50%	67%	100%	100%	100%	100%	84%
World Bank (WB)	80%	100%	86%	94%	88%	94%	82%	88%
All Agencies	69%	88%	83%	95%	92%	96%	86%	87%



6.3 PERFORMANCE BY QUALITY DIMENSION

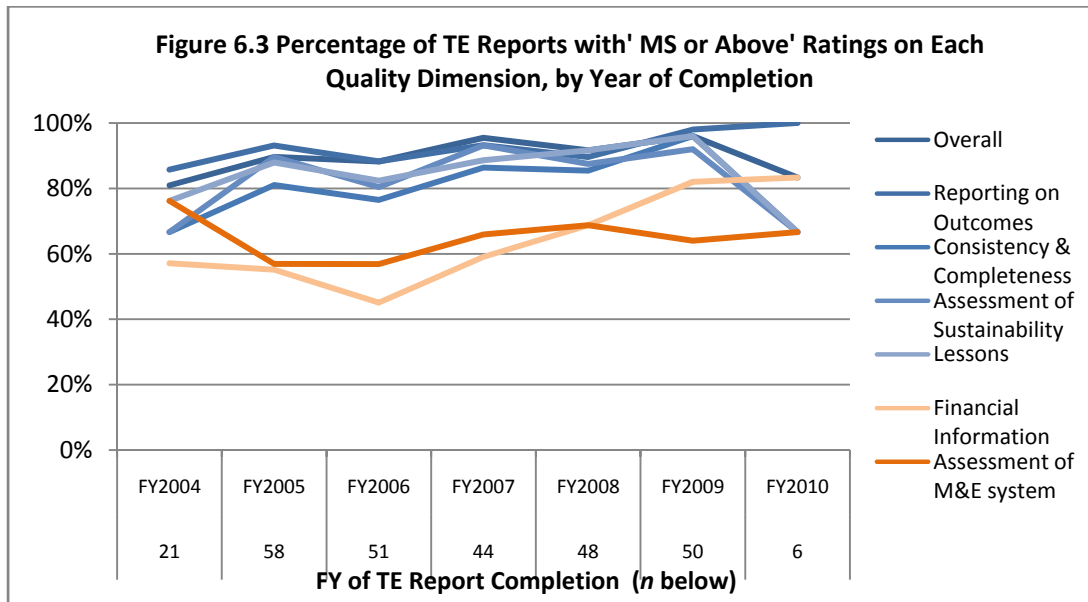
89. The Evaluation Office rates quality of a terminal evaluation report on six dimensions. Figure 6.3 shows the trends in percentage of terminal evaluation reports rated moderately satisfactory or above on individual quality dimensions. Figure 6.4 presents the trends on same quality dimensions using a more stringent yardstick – terminal evaluation reports rated satisfactory or above. Both figures together show that ratings on quality dimensions such as financial information and assessment of M&E remain lower than that on other dimensions.

90. Often, different types of monitoring that may be relevant to GEF projects are not distinguished in the terminal evaluation reports. For example, GEF projects may have three different types of monitoring:

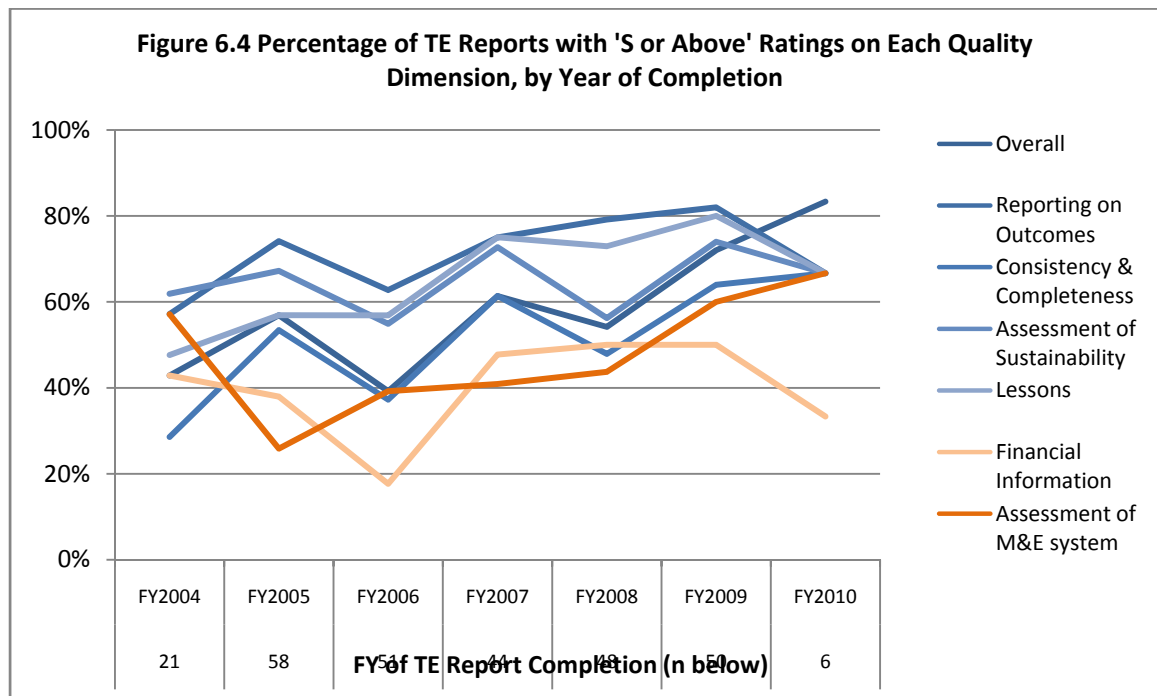
- Monitoring of project inputs, outputs, and processes, that would facilitate better tracking of project implementation progress;
- Monitoring of results that will facilitate evaluation of project outcome and impacts; and,
- Monitoring of environmental trends and baselines in trans-boundary areas as a component of the GEF project.

91. Among the three types of monitoring listed above, the second is different from the third as the former tracks changes on the outcome or impact indicators of a GEF project whereas the latter tracks changes in the environmental and socio-economic indicators without these changes being directly linked to the activities taken up by the GEF project. The reporting on implementation of monitoring and evaluation arrangements often misses these differences.

92. The reporting on implementation of monitoring and evaluation arrangements often misses these differences.



93. Figures 6.3 also show that there has been some improvement in the quality of reporting on financial issues. Most reports now present information on utilization of GEF grant and include information on amount of cofinancing that materialized. However, on most instances the information provided is at the aggregate level and is not available at the component or activity level. The reporting on cofinancing often does not clarify whether the cofinancing was managed by the project implementation unit or was managed by other organizations.

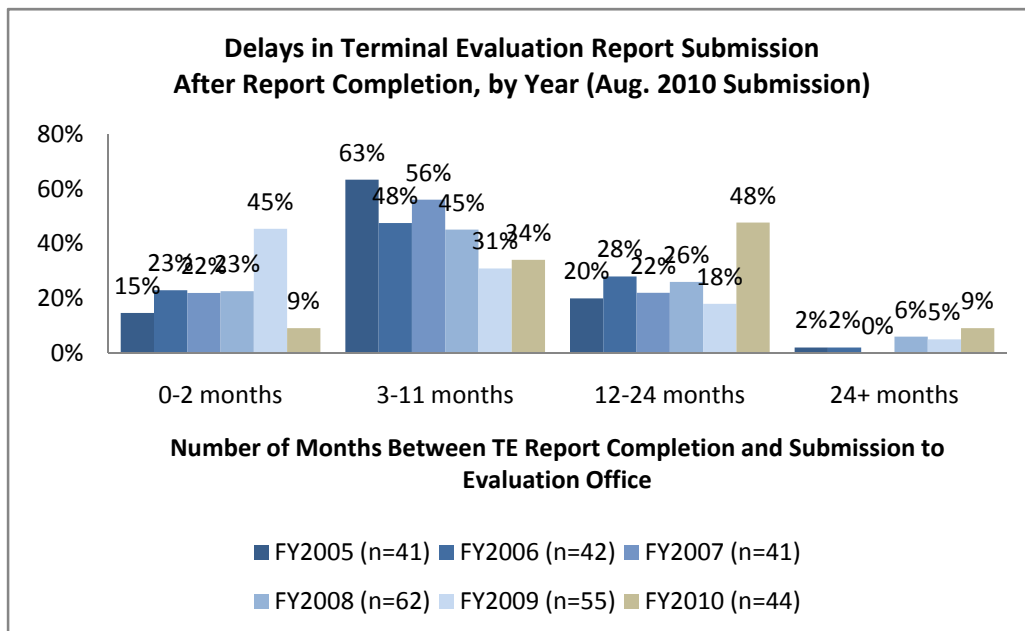
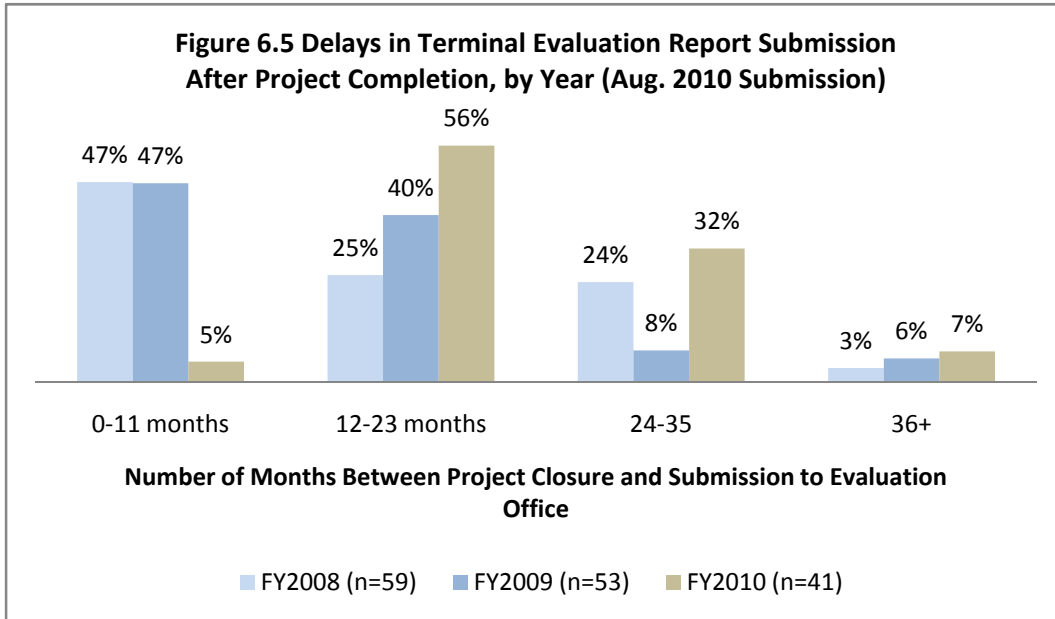


6.4 DELAYS IN TERMINAL EVALUATION REPORT SUBMISSION

94. The GEF Agencies are required to submit terminal evaluation reports within 12 months of project completion. As specified in the terminal evaluation guidelines, Agencies then have a 2-month window within which to submit the completed reports to the GEF Evaluation Office.

95. The Evaluation Office began collecting data on the time lags between project closure and terminal evaluation report in FY 2008. For the FY2010 cohort, data was available on the time lags between project closure and terminal evaluation report submission for 41 projects (excluding resubmitted projects). Figure 6.5 below shows that for the FY 2010 cohort, only 5 percent of terminal evaluation reports were submitted within 12 months of closure. The majority of reports, 56% were submitted within one to two years following project closure. In the FY 2009 and FY 2008 cohorts, 47 percent of reports were submitted within 12 months of project closure. The sharp decline in this year's cohort is most likely due to the fact that for those instances in which the exact date of submission was unknown, this analysis imputes a date of August 2010.

96. Figure 6.6 shows trends in the lags between terminal evaluation report completion and submission to the GEF Evaluation Office over the period FY 2005 to FY 2010. For the FY 2010 cohort, only 9 percent of the terminal evaluation reports were submitted within two months of completion. For 34 percent, this lag was over two months to one year. And, for the majority, 57 percent, the time lag between report completion and submission was more than one year. Again, in comparison with previous years, particularly FY 2009, this represents a sharp decline. It must be noted once more that for those instances where the exact date of submission was unknown, the Office imputes a submission date of August 2010, meaning that some lags have been magnified in the analysis.



97. Accurate data on status of project completion and terminal evaluation report submission is difficult to arrive at and influences the data comparing time lags in terminal evaluation completion and submission. Despite significant improvements in the quality of the PMIS, the quality of information on project completion and terminal evaluation status remains weak. The status of projects is not updated in a regular and systematic manner, making it difficult to determine when a project has been completed or when a terminal evaluation report has been submitted.

98. A spot check of 12 projects, comparing information in the terminal evaluation report with the PMIS revealed at least 5 instances where the PMIS contained no information or inaccurate information. Resolution of this concern requires collaborative efforts from the Secretariat, Agencies and the Evaluation Office. In last year's APR, a Council Decision on Agenda Item 9 stated that the GEF Evaluation Office, the Secretariat and the Agencies should work together in identifying and implementing measures to improve the quality of information available through PMIS on the status of projects through the project cycle, including agency compliance with deadlines for terminal evaluations. The Evaluation Office was requested to report on the progress made in the Annual Performance Report 2010. On this matter, the GEF Trustee is completing an independent review of GEF systems to be submitted to Council in November 2011. Based on the findings of the review, the Secretariat, in consultation with the Trustee, Evaluation Office and the Agencies will chart a way forward to upgrade the PMIS.

6.5 COMPARISON OF RATINGS

99. The Office compares its verified project outcome ratings, with those provided in:

- the reviews conducted by the evaluation offices of implementing agencies
- the terminal evaluation reports
- the last PIR submitted by the relevant GEF agency before project completion

100. Disconnects between the outcome ratings given by the GEF Evaluation Office and by others could potentially go in two directions: positive – when GEF Evaluation Office rates the outcome achievements of a project as satisfactory but others rate them as unsatisfactory; or, negative – when the Evaluation Office rates them to be unsatisfactory but the others rate it to be satisfactory. Among the three major implementing agencies, the evaluation office of the World Bank provides ratings on outcome achievements of all FSPs – it does not provide such ratings for MSPs. The evaluation offices of UNEP and UNDP provide ratings on outcome achievements for both FSPs and MSPs. While the evaluation offices of World Bank and UNEP have been providing ratings on outcome achievements, in FY2009, the UNDP Evaluation Office – which until then been providing ratings on only quality of terminal evaluation report – started providing its ratings on outcome achievements.

101. Of the 48 projects reviewed in FY 2010, 27 were rated by the GEF Evaluation Office. Of these, 13 also had outcome ratings provided by the implementing agencies (UNDP, UNEP, World Bank). Comparison shows that agency ratings and Evaluation

office ratings were equivalent for 77 percent of the projects (10 out of 13 projects). There was a difference in ratings for only three projects.

102. Of the three projects where there was a difference in ratings between Agencies and the Evaluation Office, in two cases the Office downgraded the Agency rating from satisfactory to moderately satisfactory, and in one case the Office upgraded the Agency rating from moderately satisfactory to satisfactory. The percent of projects rated moderately satisfactory or above was the same for Agencies and for the Evaluation Office. When considering the Agency ratings alone, 92 percent (12 out of 13) were rated moderately satisfactory or above on outcomes by Agencies. When considering the Evaluation Office ratings alone, also 92 percent of these 13 projects were rated moderately satisfactory or above by the Office.

7. MANAGEMENT ACTION RECORDS

103. The GEF's Management Action Records (MARs) track the level of adoption of Council's decisions that are based on findings and recommendations of evaluations presented by the GEF Evaluation Office. The MARs seek to increase GEF management accountability regarding Council decisions on monitoring and evaluation issues. The GEF Council approved the format and procedures for the MAR at its November 2005 meeting and requested that the GEF Evaluation Office prepare an updated MAR to be presented to the Council for review and follow up on an annual basis.

7.1 RATING APPROACH

104. The rating categories for the progress of adoption of Council decisions were agreed upon in a consultative process of the Evaluation Office, the GEF Secretariat, and the GEF Agencies and are as follows:

- **High:** Fully adopted and fully incorporated into policy, strategy or operations.
- **Substantial:** Decision largely adopted but not fully incorporated into policy, strategy or operations as yet.
- **Medium:** Adopted in some operational and policy work, but not to a significant degree in key areas.
- **Negligible:** No evidence or plan for adoption, or plan and actions for adoption are in a very preliminary stage.
- **N/A:** Non-applicable
- **Not possible to verify yet:** verification will have to wait until more data is available or proposals have been further developed.
- **No longer relevant:** Decision has lost relevance because another programming, strategic change or policy report has taken precedence.

105. The MAR presents ratings of GEF management and the verification of these ratings by the GEF Evaluation Office. They track management actions on Council

decisions based on 16 GEF Evaluation Office documents. These include 12 evaluations that were already presented in last year’s MAR:

- Annual Performance Report 2006 (GEF/ME/C.31/1, May 2007)
- Joint Evaluation of the Small Grants Programme – Executive Version (GEF/ME/C.32/2, October 2007)
- GEF Annual Report on Impact 2007 – Executive Version (GEF/ME/C.32/4, October 2007)
- Annual Country Portfolio Evaluation Report 2009 (GEF/ME/C.35/1, June 2009)
- GEF Annual Report on Impact 2009 (GEF/ME/C.36/2, November 2009)
- Role of Local Benefits in Global Environmental Programs Part One: Nature and Conclusions of the Study (GEF/ME/C.27/4, October 2005)
- Evaluation of Incremental Cost Assessment (GEF/ME/C.30/2, Nov 2006)
- Evaluation of the GEF Activity Cycle and Modalities (GEF/ME/C.30/6, Nov 2006)
- Evaluation of the GEF Support to Biosafety (GEF/ME/C.28/Inf.1, May 2006)
- Annual Performance Report 2005 (GEF/ME/C.28/2/Rev.1, May 2006)
- Annual Country Portfolio Evaluation Report 2008 (GEF/ME/C.33/4, Mar 2008)
- Mid-term Review of the Resource Allocation Framework (GEF/ME/C.34/2, Oct 2008)

106. The remaining four documents are new evaluations presented to Council in 2010:

- Annual Country Portfolio Evaluation Report 2010 (GEF/ME/C.38/2 June 2010)
- GEF Annual Performance Report 2009 (GEF/ME/C.38/4 June 2010)
- Review of the Earth Fund (GEF/ME/C.39/2 and GEF/ME/C.39/1 October 2010)
- Evaluation of the Strategic Priority for Adaptation (GEF/ME/C.39/4 October 2010)

7.2 FINDINGS

107. This year, the MAR tracks management actions on 35 Council decisions (see Table 10.1). The GEF Evaluation Office was able to verify all 35 decisions. Twenty-five have been graduated for having either achieved a “high” adoption rating (9) or are considered “no longer relevant” (16), reasons they will not be tracked in the next MAR.

Table 7.1: Ratings of GEF Progress in Adopting Council Decisions

Management ratings	GEF EO ratings						Total Sum of Mgmt ratings
	High	Substantial	Medium	Negligible	Not possible to verify yet	No longer relevant	
High	9	1	1	1	0	1	13
Substantial	0	1	0	0	0	0	1

Medium	0	1	3	1	0	1	6
Negligible	0	0	0	1	0	0	1
Not possible to verify yet	0	0	0	0	0	0	0
No longer relevant	0	0	0	0	0	0	0
Total Sum of GEF EO ratings	9	3	4	3	0	2	21 ⁹

DECISIONS WITH ADOPTION RATED AT A SUBSTANTIAL OR HIGH LEVEL

108. Of the 35 Council decisions that were tracked, for 34 percent (12) the Evaluation Office rated adoption of the decision by the management to be “high” (9) or “substantial” (3). For the MAR prepared for APR2009, the level of adoption was so rated for 65 percent of decisions. Lower levels of adoption ratings for the decisions tracked by MAR for APR FY2010 vis-à-vis those tracked for MAR 2009 do not automatically indicate that the management has become less responsive in terms of adopting the Council decisions. Rather, much of the decline in adoption ratings in this APR FY2010 refer to 16 out of the 35 Council decisions (46 percent) that were considered “no longer relevant.” However, the decline can also imply that, on average for the decisions that were tracked for MAR 2010, the management had a longer duration to adopt decisions compared to those tracked for MAR 2009, MAR 2008 and so on.

109. Nine decisions have been upgraded by GEF EO to “high” from five “substantial”, three “medium”, and one “negligible” in the previous assessment, for these issues have continue to improve during the past year. Referring to the Joint Evaluation of the Small Grants Programme (SGP), five Council decisions have been graduated for having been significantly addressed in the past year. They are decisions in which the Council requested the SGP Steering Committee to; firstly, propose a level of management costs on the basis of services rendered and cost-efficiency rather than on the basis of a stated percentage; secondly, start a process to change SGP’s central management system suitable for the new phase of growth and to address the risks of growing complexity; thirdly, further strengthen M&E; fourthly, propose a revision of the current criteria for access to SGP resources to maintain cost efficiency, and; lastly, further develop a graduation policy for the SGP country programs which takes into account the identified risks to GEF achievements and cost-effectiveness, especially in SIDS and LDCs.

110. Beyond these four decisions graduated in the scope of SGP, other five Council decisions were also graduated as high on level of adoption. Firstly, referring to the *GEF*

⁹ GEF EO did not seek management comment on 14 Council decisions that were deemed ‘no longer relevant’ due to the GEF-5 policies superseding decisions.

Annual Performance 2006, the decision that UNEP should develop a systemic approach to supervision of its GEF portfolio has been followed and therefore graduated. Secondly, on the *GEF Annual Report on Impact 2007*, the decision requesting the GEFSEC to incorporate its recommendations into project preparation and to ensure adequate monitoring of progress towards institutional continuity have been put in place across the project cycle and will have the effectiveness of these mechanisms assessed by GEF EO in future evaluations. Thirdly, on the *Annual Country Portfolio Evaluation Report 2009*, the decision to explore within the GEF partnership modalities to address the significant gap of available resources for combating land degradation to support key challenges facing countries like Egypt, Syria, and Cameroon was graduated since the new System for Transparent Allocation of Resources (STAR) makes available resources for land degradation. Finally, on the *GEF Annual Report on Impact 2009*, the decision that GEF-5 strategy proposals, prepared by the Secretariat, should include further investment and capacity development to assist countries with economies in transition to address the remaining threats to the ozone layer was also graduated, since GEF-5 strategies include further funding for investments in economies in transition to support meeting their reporting obligations under the Montreal protocol.

DECISIONS THAT HAVE SHOWN NO CHANGE IN RATINGS

111. Only one of the Council decisions that was included in the previous MAR had rating on level of adoption that had not improved, which signals that most issues have been either addressed in the past year and progress on the ground has taken place, or that the referred issues have lost relevance. This decision with the unchanged rating is part of the *GEF Annual Performance Report 2006*, in which all GEF agencies would need to ensure that terminal evaluation reports include adequate information on sustainability of outcomes, quality of M&E systems and reporting on co-financing. The rating was maintained as “substantial” because although Agencies have made progress by putting in place systems to review and provide feedback on the quality of terminal evaluations, the extent of TE improvement on the specific aspects indicated in this recommendation will be reported in APRs to come.

COMPARISON BETWEEN THE EVALUATION OFFICE AND MANAGEMENT RATINGS

112. This year the GEF Management and the GEF Evaluation Office agree in the rating on the level of adoption of 14 out of 21 (66 percent) Council decisions (not considering the 14 decisions deemed ‘no longer relevant’ for which GEF EO did not seek Management comment due to the GEF-5 policies superseding decisions) as compared to 56 percent in the preceding year. While an important reason for the differences in the rating is because of the bias that could be expected when performance is self-rated, an equally important reason is a difference in interpretation of Council decisions. In the last APR, although the level of agreement was lower in relation to the previous year, much of the disagreement was due to the management being more conservative in rating level of adoption for some of the Council decision (Table 10.1).

113. This year, however, the situation is opposite. The GEF Management ratings for level of adoption are higher than the Evaluation Office's for four out of seven decisions in which GEF Management and GEF EO diverged. Two out of these seven were considered "no longer relevant" by GEF EO because another programming or policy report has taken precedence. Only one of the seven decisions was upgraded by GEF EO.

114. The single decision in which the GEF Evaluation Office upgraded the rating from the GEF Management was in the Review of the Earth Fund, for which GEF EO rated "substantial" as opposed to Management's "medium", noting that preparation of a new Private Sector Strategy is underway by the Secretariat, with a paper to be presented to Council in May 2011.

115. For the four decisions in which GEF Management ratings for level of adoption were higher than the GEF Evaluation Office's, the divergence was largely due to EO's different interpretation in measures taken to respond to the decision. Firstly, in the downgrading of SGP's efforts for strengthening country program oversight from "high" to "substantial", the EO explained that "it is not clear how effective [these measures] will be." Secondly, in the *Annual Country Portfolio Evaluation Report 2009*, the downgrading from "high" to "negligible" of the response to the decision to conduct a survey of countries in exceptional situations concerning limited access to GEF partner International Financial Institutions, like Syria, was because "the issue goes beyond availability of resources for countries like Syria." Thirdly, in the *GEF Annual Report on Impact 2009*, the downgrading from "high" to "medium" on the decision to incorporate lessons from the positive private sector engagement in the Ozone Layer Depletion focal area into its efforts to engage the private sector, where possible and as appropriate, in other focal areas, was due to GEF EO's interpretation that a new strategic document is a first step that still needs to have its effects assessed by GEF EO in the future.

116. Lastly, regarding the *GEF Annual Performance Report 2009*, a downgrading from "medium" to "negligible" was made due to the "slow progress" in the decision on the Agenda Item 9 that the GEF EO, the Secretariat and the Agencies should work together to improve quality of information available through PMIS on the status of projects through the project cycle, including agency compliance with deadlines for terminal evaluations. The Evaluation Office was requested to report on the progress made in the Annual Performance Report 2010. On this matter, despite significant improvements in quality of PMIS, the quality of information on project completion status remains weak. Status of projects is not updated in a regular and systematic manner. As a result, it is difficult to determine whether a project has been completed. The GEF Trustee is completing an independent review of GEF systems to be submitted to Council in November 2011. Based on the findings of the review, the Secretariat, in consultation with the Trustee, Evaluation Office and the Agencies will chart a way forward to upgrade the PMIS.

117. Two decisions were considered to have lost relevance. In relation to the decision that UNDP and UNEP should involve social and institutional expertise in project

supervision where appropriate. The GEF Management had rated progress “high” and GEF EO “substantial” in the last MAR, this year GEF EO considers that this recommendation is now superseded by the GEF5 program documents, which specify that the GEF Secretariat will develop a Gender and Social Policy. Also, in relation to the *Annual Country Portfolio Evaluation Report 2010*, the decision that GEF Agencies should systematically involve Operational Focal Points in M&E activities by sharing information with them in a timely manner, which was rated as “substantial” by GEF Management, is considered “no longer relevant” due to GEF EO’s interpretation that the revised GEF M&E policy now sets the minimum requirement for GEF Agencies to systematically involve OFPs in M&E activities by sharing information with them in a timely manner, and the GEF EO will review implementation of policy during GEF-5, OPS5 and through cited examples. A continuous review is also taking place through the Country Portfolio Evaluation work program.

GRADUATED DECISIONS

118. Since its start, the MAR has followed the adoption of 98 GEF Council decisions based on recommendations of 27 evaluations. In general, the GEF has been very responsive to Council decisions, which have led to the ongoing reform process. The GEF Evaluation Office graduates decisions for which either a “high” adoption rating has been achieved or that are considered “no longer relevant.” A total of 88 (90 percent) Council decisions, including nine that attained the “high” rating this year, have been graduated since the first MAR was presented in the 2005 APR (table 10.2).

Table 7.2: Summary of Council Decisions graduated from MAR

MAR	Fully adopted	No longer relevant					TOTAL
	High	Substantial	Medium	Negligible	Not possible to verify yet	Not Applicable	
2005	5	15	7	3	-	-	30
2006	5	1	-	-	-	-	6
2007	7	8	-	-	2	-	17
2008	5	-	-	-	-	-	5
2009	5	-	-	-	-	-	5
2010	9	3	4	3	-	2	21
TOTAL	36	27	11	6	2	2	84

119. This year, a total of 16 Council decisions were considered to have lost relevance, and therefore graduated, because another programming, strategic change or policy report contained in GEF's replenish documents have taken precedence. Graduation of decisions that are no longer relevant allows the GEF partnership to focus on issues that are more reflective of the present concerns of the Council.

120. Out of these 16 decisions that were considered to have lost relevance, 2 of them were so deemed after GEF EO assessed the ratings provided by the GEF Management. As mentioned before, in relation to the decision that UNDP and UNEP should involve social and institutional expertise in project supervision where appropriate. While the GEF Management had rated "high" for having considered that UNEP and UNDP have addressed this recommendation, GEF EO considered it now superseded by the GEF5 Program Document – which specifies that the GEF Secretariat develop a Gender and Social Policy – and therefore "no longer relevant." Also, in relation to the Annual Country Portfolio Evaluation Report 2010, the decision that GEF Agencies should systematically involve Operational Focal Points in M&E activities by sharing information with them in a timely manner, which was rated as "substantial" by GEF Management, is considered "no longer relevant" due to GEF EO's interpretation that the revised GEF M&E policy now sets the minimum requirement for GEF Agencies to systematically involve OFPs in M&E activities by sharing information with them in a timely manner, and the GEF EO will review implementation of policy during GEF-5, OPS5 and through cited examples. A continuous review is also taking place through the Country Portfolio Evaluation work program.

121. Besides this two graduated decisions, the GEF Evaluation Office also graduated 14 other Council decisions as "no longer relevant":

- Four were related to the recommendations of the evaluation on 'Role of Local Benefits in Global Environmental Programs – Part One: Nature and Conclusions of the Study (GEF/ME/C.27/Inf.4, October 2005);
- Two decisions were related to the 'Evaluation of Incremental Cost Assessment (GEF/ME/C.30/2, November 2006);
- Three decisions related to the 'Evaluation of GEF Activity Cycle and Modalities (GEF/ME/C.30/6, November 2006);
- One was related to the 'Evaluation of the GEF Support to Biosafety (GEF/ME/C.28/Inf.1, May 2006);
- Two were related to the 'Annual Country Portfolio Evaluation Report 2008 (GEF/ME/C.33/4, March 2008);
- One was related to 'GEF Annual Performance Report 2005 (GEF/ME/C.28/2/Rev.1, May 2006), and;
- One was related to the 'Mid-term Review of the Resource Allocation Framework (GEF/ME/C.34/2, October 2008).

122. Although these decisions are deemed "no longer relevant" due to a superseding policy, levels of adoption will be assessed in the context of a review of GEF5 program

and policy in the Fifth overall performance study (OPS5). For the graduated decisions deemed “no longer relevant”, progress on the development, implementation and reporting, including the use of tracking tools, GEF-5 strategies, and additional follow up on levels of adoption, would all be assessed in the context of OPS5.

123. A complete version of the 2009 MAR is available at the GEF EO website (www.gefeo.org).

8. PERFORMANCE MATRIX

124. The performance matrix provides a summary of the performance of the GEF Implementing Agencies and GEF Secretariat on a variety of parameters (see Table 11.1). Although several of these parameters are assessed by the Evaluation Office on an annual basis, to mitigate fluctuations in performance ratings due to differences in project mix or other idiosyncratic factors, the values presented in the matrix are, depending on the parameter, running averages of two to four years. Information has been provided for six parameters in the performance matrix. In the future, if data for more years become available, it will be possible to track improvement in performance on a greater number of dimensions.

8.1 RATING APPROACH

125. Reporting methodology varies by parameter:

- Four performance parameters—project outcomes, implementation completion delays, materialization of cofinancing (per dollar of approved GEF financing and as percentage of promised cofinancing), and quality of monitoring and evaluation during project implementation—are being reported as five-year running averages, as systemic changes in performance on these parameters are expected to be gradual.
- The figures for the two other parameters – reported for performance on quality assurance of project M&E arrangement at entry and on quality of project terminal evaluation reports – are two-year running averages, as meaningful changes can be attained in the short run.
- Changes in performance are also likely to be gradual for a second set of parameters: quality of supervision and adaptive management, and quality of project M&E arrangements. Moreover, assessment of performance on these parameters requires intensive thematic appraisals, which was not undertaken for this APR. For the sake of efficiency, the Evaluation Office will take up such appraisals as part of the APR after a two-year interval.

Table 8.1: GEF Agency and Institutions Performance Matrix (in percentages)

Parameter	UNDP	UNEP	World Bank	GEF Secretariat	Overall GEF Performance
Results					
1. Project Outcomes: percentage of completed projects with outcomes rated moderately satisfactory or above	81	90	86	—	86
Efficiency					
2. Implementation completion delays: average delay in completion of projects in months	19	18	13	—	17
3. Materialization of Co-financing:					
a. Reported materialization of cofinancing per dollar of approved GEF financing	6.5	1.2	1.9	—	2.7
b. Reported materialization of cofinancing as percentage of promised cofinancing	176	172	102	—	134
Quality of M&E					
4. Quality assurance of project M&E arrangements at entry	72	63	83	—	73
5. Quality of project M&E during implementation	60	62	74	—	65
6. Quality of project terminal evaluation	92	100	88	—	93

8.2 FINDINGS

126. Information has been provided on six parameters included in the performance matrix. The findings on outcome achievements are highlighted here. Based on the review of terminal evaluation reports submitted to the GEF Evaluation Office from FY 2006 to FY 2010, the EO rated outcome (parameter 1) in 86 percent of the projects to be moderately satisfactory and above. Among the Implementing Agencies, outcome of 81 percent of UNDP-, 90 percent of UNEP, and 86 percent of World Bank-implemented projects were rated in the satisfactory range. UNDP and UNEP displayed an important improvement in outcome ratings, while the World Bank has been downgraded one percent only which is statistically insignificant.

127. The quality of terminal evaluations from UNEP maintained were at 100 percent in the satisfactory range, while UNDP and the World Bank have been downgraded two percent in relation to the previous assessment.

ANNEX A. TERMINAL EVALUATION REPORT REVIEW GUIDELINES

128. The assessments in the terminal evaluation reviews will be based largely on the information presented in the terminal evaluation report. If insufficient information is presented in a terminal evaluation report to assess a specific issue such as, for example, quality of the project's monitoring and evaluation system or a specific aspect of sustainability, then the preparer of the terminal evaluation reviews will briefly indicate so in that section and elaborate more if appropriate in the section of the review that addresses quality of report. If the review's preparer possesses other first-hand information such as, for example, from a field visit to the project, and this information is relevant to the terminal evaluation reviews, then it should be included in the reviews only under the heading "Additional independent information available to the reviewer." The preparer of the terminal evaluation review will take into account all the independent relevant information when verifying ratings.

A.1 CRITERIA FOR OUTCOME RATINGS

129. Based on the information provided in the terminal evaluation report, the terminal evaluation review will make an assessment of the extent to which the project's major relevant objectives were achieved or are expected to be achieved, relevance of the project results, and the project's cost-effectiveness¹⁰. The ratings on the outcomes of the project will be based on performance on the following criteria¹¹:

- **Relevance.** Were project outcomes consistent with the focal area/operational program strategies and country priorities? Explain.
- **Effectiveness.** Are project outcomes commensurate with the expected outcomes (as described in the project document) and the problems the project was intended to address (the original or modified project objectives)?
- **Efficiency.** Include an assessment of outcomes and impacts in relation to inputs, costs, and implementation times based on the following questions: Was the project cost-effective? How does the project's cost-time versus outcomes equation compare to that of similar projects? Was the project implementation delayed due to any bureaucratic, administrative, or political problems and did that affect cost-effectiveness?

¹⁰ *Objectives* are the intended physical, financial, institutional, social, environmental, or other development results to which a project or program is expected to contribute (OECD DAC 2002).

¹¹ *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 that result from a development intervention; these may also include changes resulting from the intervention that are relevant to the achievement of outcomes (OECD DAC 2002). For the GEF, environmental outcomes are the main focus.

130. An overall rating will be provided according to the achievement and shortcomings in the three criteria ranging from highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, highly unsatisfactory, and unable to assess.

131. The reviewer of the terminal evaluation will provide a rating under each of the three criteria (relevance, effectiveness, and efficiency). Relevance of outcomes will be rated on a binary scale: a 'satisfactory' or an 'unsatisfactory' rating will be provided. If an 'unsatisfactory' rating has been provided on this criterion, the overall outcome achievement rating may not be higher than "unsatisfactory". Effectiveness and Efficiency will be rated as following:

- **Highly satisfactory.** The project had no shortcomings.
- **Satisfactory.** The project had minor shortcomings.
- **Moderately satisfactory.** The project had moderate shortcomings.
- **Moderately unsatisfactory.** The project had significant shortcomings.
- **Unsatisfactory.** The project had major shortcomings.
- **Highly unsatisfactory.** The project had severe shortcomings.
- **Unable to assess.** The reviewer was unable to assess outcomes on this dimension.

132. The calculation of the overall outcomes score of projects will consider all three criteria, of which relevance criterion will be applied first - the overall outcome achievement rating may not be higher than "unsatisfactory". The second constraint that is applied is that the overall outcome achievement rating may not be higher than the "effectiveness" rating. The third constraint that is applied is that the overall rating may not be higher than the average score of effectiveness and efficiency criteria calculated using the following formula:

$$\text{Outcomes} = (b + c) \div 2$$

133. In case the average score is lower than the score obtained after application of the first two constraints, then the average score will be the overall score. The score will then be converted into an overall rating with mid values being rounded upwards.

A.2 CRITERIA FOR SUSTAINABILITY RATINGS

134. Sustainability will be understood as the likelihood of continuation of project benefits after completion of project implementation (GEF 2000). To assess sustainability, the terminal evaluation reviewer will identify and assess the key risks that could undermine continuation of benefits at the time of the evaluation. Some of these risks might include the absence of or inadequate financial resources, an enabling legal framework, commitment from key stakeholders, and enabling economy. The following four types of risk factors will be assessed by the terminal evaluation reviewer to rate the likelihood of sustainability of project outcomes: financial, sociopolitical, institutional frameworks and governance, and environmental.

135. The following questions provide guidance to assess if the factors are met:

- **Financial resources.** What is the likelihood that financial resources will be available to continue the activities that result in the continuation of benefits (income-generating activities, and trends that may indicate that it is likely that in future there will be adequate financial resources for sustaining project outcomes)?

- **Sociopolitical.** Are there any social or political risks that can undermine the longevity of project outcomes? What is the risk that the level of stakeholder ownership is insufficient to allow for project outcomes/benefits to be sustained? Do the various key stakeholders see it in their interest that the project benefits continue to flow? Is there sufficient public/stakeholder awareness in support of the long-term objectives of the project?
- **Institutional framework and governance.** Do the legal frameworks, policies, and governance structures and processes pose any threat to the continuation of project benefits? While assessing this parameter, consider if the required systems for accountability and transparency, and the required technical know-how, are in place.
- **Environmental.** Are there any environmental risks that can undermine the future flow of project environmental benefits? The terminal evaluation should assess whether certain activities in the project area will pose a threat to the sustainability of project outcomes. For example, construction of dam in a protected area could inundate a sizable area and thereby neutralize the biodiversity-related gains made by the project.

136. The reviewer will provide a rating under each of the four criteria (financial resources, sociopolitical, institutional, and environmental) as follows:

- **Likely.** There are no risks to sustainability of outcomes.
- **Moderately likely.** There are moderate risks to sustainability of outcomes.
- **Moderately unlikely.** There are significant risks to sustainability of outcomes.
- **Unlikely.** There are severe risks to sustainability of outcomes.
- **Unable to assess.** Unable to assess risks on this dimension.
- **Not applicable.** Risks on this dimension are not applicable to the project.

137. A number rating 1–4 will be provided in each category according to the achievement and shortcomings with likely = 4, moderately likely = 3, moderately unlikely = 2, unlikely = 1, and not applicable = NA. A rating of unable to assess will be used if the reviewer is unable to assess any aspect of sustainability. In such instances, it may not be possible to assess the overall sustainability.

138. All the risk dimensions of sustainability are critical. Therefore, the overall rating will not be higher than the rating of the dimension with the lowest rating. For example, if the project has an unlikely rating in either of the dimensions, then its overall rating cannot be higher than unlikely, regardless of whether higher ratings in other dimensions of sustainability produce a higher average.

A.3 CRITERIA FOR ASSESSMENT OF QUALITY OF PROJECT M&E SYSTEMS

139. GEF projects are required to develop M&E plans by the time of work program inclusion, to appropriately budget M&E plans, and to fully carry out the M&E plan during implementation. Project managers are also expected to use the information generated by the M&E system during project implementation to improve and adapt the project to changing situations. Given the long-term nature of many GEF projects, projects are also encouraged to include long-term monitoring plans that measure results (such as environmental results) after project completion. Terminal evaluation reviews will include an assessment of the achievement and shortcomings of M&E systems.

140. **M&E design.** Project should have a sound M&E plan to monitor results and track progress in achieving project objectives. An M&E plan should include a baseline (including data, methodology, and so on), appropriate indicators and data analysis systems, and evaluation studies at specific times to assess results. The time frame for various M&E activities and standards for outputs should have been specified. The questions to guide this assessment include: In retrospect, was the M&E plan at entry practicable and sufficient (sufficient and practical indicators identified; timely baseline; targets created; effective use of data collection; analysis systems including studies and reports; practical organization and logistics in terms of what, who, and when for M&E activities)?

141. **M&E plan implementation.** The M&E system was in place and allowed the timely tracking of results and progress toward project objectives throughout the project. Annual project reports were complete, accurate, and with well-justified ratings. The information provided by the M&E system was used to improve and adapt project performance. An M&E system should be in place with proper training for parties responsible for M&E activities to ensure that data will continue to be collected and used after project closure. The questions to guide this assessment include: Did the project M&E system operate throughout the project? How was M&E information used during the project? Did it allow for tracking of progress toward project objectives? Did the project provide proper training for parties responsible for M&E activities to ensure data will continue to be collected and used after project closure?

142. **Other questions.** This includes questions on funding and whether the M&E system was a good practice.

- Was sufficient funding provided for M&E — in the budget included in the project document?
- Was sufficient and timely funding provided – for M&E during project implementation?
- Can the project M&E system be considered – a good practice?

143. A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately

satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = UA. The reviewer of the terminal evaluation will provide a rating under each of the criterion (M&E design, and M&E plan implementation) as follows:

- **Highly satisfactory.** There were no shortcomings in that criterion of the project M&E system.
- **Satisfactory.** There were minor shortcomings in that criterion of the project M&E system.
- **Moderately satisfactory.** There were moderate shortcomings in that criterion of the project M&E system.
- **Moderately unsatisfactory.** There were significant shortcomings in that criterion of the project M&E system.
- **Unsatisfactory.** There were major shortcomings in that criterion of the project M&E system.
- **Highly unsatisfactory.** There was no project M&E system.

144. The rating for M&E during implementation will be the overall rating of the M&E system: Rating on the Quality of the Project Monitoring and Evaluation System = b

A.4 CRITERIA FOR ASSESSMENT OF QUALITY OF TERMINAL EVALUATION REPORTS

145. The ratings on quality of terminal evaluation reports will be assessed using the following criteria:

- The report presents an assessment of all relevant outcomes and achievement of project objectives in the context of the focal area program indicators if applicable.
- The report was consistent, the evidence presented was complete and convincing, and ratings were well substantiated.
- The report presented a sound assessment of sustainability of outcomes.
- The lessons and recommendations are supported by the evidence presented and are relevant to the portfolio and future projects.
- The report included the actual project costs (totals, per activity and per source) and actual cofinancing used.
- The report included an assessment of the quality of the M&E plan at entry, the M&E system used during implementation, and whether the information generated by the M&E system was used for project management.

146. A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = UA. Each criterion to assess the quality of the terminal evaluation will be rated as follows:

- **Highly satisfactory.** There were no shortcomings in the terminal evaluation on this criterion.
- **Satisfactory.** There were minor shortcomings in the terminal evaluation on this criterion.
- **Moderately satisfactory.** There were moderate shortcomings in the terminal evaluation on this criterion.
- **Moderately unsatisfactory.** There were significant shortcomings in the terminal evaluation on this criterion.
- **Unsatisfactory.** There were major shortcomings in the terminal evaluation on this criterion.
- **Highly unsatisfactory.** There were severe shortcomings in the terminal evaluation on this criterion.

147. The first two criteria (of all relevant outcomes and achievement of project objectives and report consistency and substantiation of claims with proper evidence) are more important and have therefore been assigned a greater weight. The quality of the terminal evaluation reports will be calculated by the following formula: Quality of the Terminal Evaluation Report = $0.3 \times (a + b) + 0.1 \times (c + d + e + f)$

148. The total number will be rounded and converted to the scale of highly satisfactory to highly unsatisfactory.

A.5 ASSESSMENT OF PROCESSES AFFECTING ATTAINMENT OF PROJECT OUTCOMES AND SUSTAINABILITY

149. This section of the terminal evaluation review will summarize the factors or processes related to implementation delays and cofinancing that may have affected attainment of project results. This section will summarize the description in the terminal evaluation on key causal linkages of these factors:

- Cofinancing and project outcomes and sustainability. If there was a difference in the level of expected cofinancing and actual cofinancing, what were the reasons for it? To what extent did materialization of cofinancing affect project outcomes and/or sustainability? What were the causal linkages of these effects?
- Delays and project outcomes and sustainability. If there were delays, what were the reasons for them? To what extent did the delay affect project outcomes and/or sustainability? What were the causal linkages of these effects?
- Country Ownership and sustainability. Assess the extent to which country ownership has affected project outcomes and sustainability? Describe the ways in which it affected outcomes and sustainability highlighting the causal links.

ANNEX B. LIST OF TERMINAL EVALUATIONS SUBMITTED IN FY2010

150. **This annex lists the projects for which terminal evaluation reviews were conducted in FY 2010.** Corresponding lists for previous reports are found in APR 2004, annex D; APR 2005, annex F; APR 2006, annex B; APR 2007, annex B; APR 2008, annex B; and APR 2009, annex B.

GEF ID	Project Name	Country	Focal	Size	Agency
267 Resubmitted	Energy Efficiency Improvements and Greenhouse Gas Reductions	Regional	CC	FSP	UNDP
325	Coal Bed Methane Capture and Commercial Utilization	India	CC	FSP	UNDP
444	Energy and Water Sector Reform and Development	Cape Verde	CC	FSP	WB
464	Global Environmental Citizenship (GEC)	Regional	MF	FSP	UNDP/UNEP
646	Market Development for Solar Water Heaters	Morocco	CC	FSP	UNDP
771	Amazon Region Protected Areas Program (ARPA)	Brazil	BD	FSP	WB
777	Northern Savanna Biodiversity Conservation (NSBC) Project	Ghana	BD	FSP	WB
876	Partnership for Natural Ecosystem Management Program (PAGEN)	Burkina Faso	BD	FSP	WB
885	Reversing Environmental Degradation Trends in the South China Sea and Gulf of Thailand	Regional	IW	FSP	UNEP
948	Vilnius Heat Demand Management Project	Lithuania	CC	FSP	WB
974	Environmental Protection and Sustainable Integrated Management of the Guarani Aquifer	Regional	IW	FSP	WB
977	Conservation and Sustainable Use of Traditional	Zimbabwe	BD	MSP	UNDP

	Medicinal Plants				
1031	Biodiversity Conservation and Sustainable Use of the Marine Resources at Con Dao National Park	Vietnam	BD	MSP	UNDP
1034	Strengthening Romania's Protected Area System by Demonstrating Best Practices for Management of Small Protected Areas in Macin Mountains National Park	Romania	BD	MSP	UNDP
1045	Biodiversity Protection in North Vidzeme Biosphere Reserve	Latvia	BD	FSP	UNDP
1067 Resubmitted	Integrated Coastal and Marine Biodiversity Management	Gambia	BD	MSP	WB
1079	Off-grid Rural Electrification for Development (PERZA)	Nicaragua	CC	FSP	UNDP/WB
1084	Caribbean: Mainstreaming Adaptation to Climate Change	Regional	CC	FSP	WB
1096	Energy Management and Performance Related Energy Savings Scheme (EMPRESS)	Regional	CC	FSP	UNEP
1098	Conservation of Globally Significant Wetlands	Republic of Korea	BD	FSP	UNDP
1109	Senegal River Basin Water and Environmental Management Program	Regional	IW	FSP	WB/UNDP
1155	Introduction of Climate Friendly Measures in Transport	Mexico	CC	FSP	WB
1196	Transformation of the Rural Photovoltaics (PV) Market	Tanzania	CC	FSP	UNDP
1216	Building Scientific and Technical Capacity for Effective Management and Sustainable Use of Dryland Biodiversity in West African Biosphere Reserves	Regional	BD	FSP	UNEP

1265	Polish Energy Efficiency Motors Programme	Poland	CC	FSP	UNDP
1296	The Green Corridor	Vietnam	BD	MSP	WB
1330	Sustainable Land Management in the Zambian Miombo Woodland Ecosystem	Zambia	MF	MSP	WB
1471	Improving Management of NGO and Privately Owned Nature Reserves and High Biodiversity Islands in Seychelles	Seychelles	BD	MSP	WB
1527	Conservation and Restoration of the Globally Significant Biodiversity of the Tisza River Floodplain through Integrated Floodplain Management	Hungary	BD	MSP	UNDP
1591	Regional Program of Action and Demonstration of Sustainable Alternatives to DDT for Malaria Vector Control in Mexico and Central America	Regional	IW	FSP	UNEP
1666	Development and Implementation of a Sustainable Resource Management Plan for Marsabit Mountain and its associated Watersheds	Kenya	LD	MSP	UNEP
1679	Strengthening Romania's Protected Area System by Demonstrating Government-NGO Partnership in Romania's Maramures Nature Park	Romania	BD	MSP	UNDP
1705	Conservation of Biological Diversity of Carpathian Mountain Grasslands in the Czech Republic through Targeted Application of New EU Funding Mechanisms	Czech Republic	BD	MSP	UNDP
1769	Integrated Management of Peatlands for Biodiversity and Climate Change: The Potential of Managing Peatlands for Carbon Accumulation While Protecting Biodiversity	Global	MF	MSP	UNEP
1782	Richtersveld Community Biodiversity Conservation Project	South Africa	BD	MSP	WB
1859	Conservation of the Eg-Uur Watershed	Mongolia	BD	MSP	WB/IFC

1876	Naya Biological Corridor in the Munchique-Pinche Sector	Colombia	BD	MSP	WB
1929	Participatory Community-based Conservation in the Anjozorobe Forest Corridor	Madagascar	BD	MSP	UNDP
2077	Lambusango Forest Conservation, Sulawesi	Indonesia	BD	MSP	WB
2118 / 3185	Total Sector Methyl Bromide Phase Out in Countries with Economies in Transition	Regional	OD	FSP	UNEP/UNDP
2151	Novel Forms of Livestock & Wildlife Integration Adjacent to Protected Areas in Africa	Tanzania	BD	MSP	WB
2183	Community-based Integrated Natural Resources Management Project in Okyeman	Ghana	MF	MSP	WB
2490	Renewable Energy from Agricultural Wastes	Moldova	CC	MSP	WB
2594	DHEKUANA NONOODO: Sustainable Use and Conservation of Biodiversity Resources of Dhekuana Indigenous Lands	Venezuela	BD	MSP	WB
2630	Lake Balaton Integrated Vulnerability Assessment, Early Warning and Adaptation Strategies	Hungary	CC	MSP	UNDP
2665	Southern Cone Development Marketplace (Environment Window)	Regional	MF	MSP	WB
2667	Community Micro Hydro for Sustainable Livelihood	Bhutan	CC	MSP	UNDP
2856	Knowledge Base for Lessons Learned and Best Practices in the Management of Coral Reefs	Global	BD	MSP	UNEP

Note: WB = World Bank; FSP = Full Size Project; MSP = Medium Size Project; BD = Biodiversity; CC = Climate Change; IW = International Waters; MF = Multi Focal Area; LD = Land Degradation; OD = Ozone Depletion Substances.

ANNEX C. METHODOLOGICAL NOTES ON THE PERFORMANCE MATRIX

151. This annex briefly describes the considerations taken into account for each of the performance matrix's 6 parameters.

C.1 PROJECT OUTCOMES

152. The figures on project outcomes are five-year moving averages based on the terminal evaluation reports submitted in the preceding years, including the fiscal year for which the APR is being presented; the figures presented in this year's APR are based on the terminal evaluation reports submitted during FYs 2006, 2007, 2008, 2009 and 2010. The aggregate figures are weighted averages, with each project considered to have equal weight.

C.2 PROJECT IMPLEMENTATION COMPLETION DELAYS

153. The information presented in the terminal evaluation reports is the primary source for this parameter. The figures for implementation completion delays are five-year averages and are based on the information provided in the terminal evaluation reports. The figures presented in this year's APR are based on the terminal evaluation reports submitted during FYs 2006, 2007, 2008, 2009 and 2010.

C.3 MATERIALIZATION OF COFINANCING

154. The analysis is based on the information provided by the Agencies in the terminal evaluation reports or through other communications. These figures have not been verified.

C.4 QUALITY ASSURANCE OF PROJECT M&E ARRANGEMENTS AT ENTRY

155. An assessment of quality assurance of project M&E arrangements at entry was carried out in the 2005 APR. It was based on a review of the M&E plans of the project appraisal documents that were endorsed by the GEF Chief Executive Officer in that fiscal year. In FY 2008, the Evaluation Office updated the ratings on this parameter based on the findings of a follow-up assessment.

C.5 QUALITY OF PROJECT M&E DURING IMPLEMENTATION

156. Figures on quality of project M&E during implementation are based on review of the terminal evaluation reports submitted to the Evaluation Office. The figures need to be five-year running averages of the percentage of projects rated moderately satisfactory or above in M&E during implementation. The figures reported in the matrix are a weighted average, with each project having an equal weight, of the data from the review of the reports submitted during FYs 2006, 2007, 2008, 2009 and 2010.

C.6 QUALITY OF PROJECT TERMINAL EVALUATION

157. Figures on quality of terminal evaluation reports are based on the ratings provided by the Evaluation Office after their review. For this parameter, two-year running averages are used, with each project having an equal weight. The figures presented in the matrix pertain to FYs 2009 and 2010.