

A GEF  
ANNUAL  
REPORT

# GEF Annual Performance Report 2012





GLOBAL ENVIRONMENT FACILITY  
EVALUATION OFFICE

# GEF Annual Performance Report 2012

August 2013

EVALUATION REPORT NO. 83

*This report was presented to the GEF Council in June 2013.*

© 2013 Global Environment Facility Evaluation Office  
1818 H Street, NW  
Washington, DC 20433  
Internet: [www.gefeo.org](http://www.gefeo.org)  
Email: [gefevaluation@thegef.org](mailto:gefevaluation@thegef.org)

All rights reserved.

The findings, interpretations, and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the GEF Council or the governments they represent.

The GEF Evaluation Office does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of the GEF concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

### **Rights and Permissions**

The material in this work is copyrighted. Copying and/or transmitting portions or all of this work without permission may be a violation of applicable law. The GEF encourages dissemination of its work and will normally grant permission promptly.

ISBN-10: 1-933992-59-X

ISBN-13: 978-1-933992-59-4

### **Credits**

*Director of the GEF Evaluation Office:* Robert D. van den Berg

*Team Leader/Task Manager:* Neeraj Kumar Negi, Senior Evaluation Officer, GEF Evaluation Office

*Editing and design:* Nita Congress

*Cover photo:* Gulmarg, India, by Neeraj Kheda

Evaluation Report No. 83

A FREE PUBLICATION

# Contents

FOREWORD .....	vii
ACKNOWLEDGMENTS .....	viii
ABBREVIATIONS .....	ix
<b>1. BACKGROUND AND MAIN CONCLUSIONS.....</b>	<b>1</b>
1.1 Background .....	1
1.2 Findings and Conclusions .....	2
1.3 Management Action Record Findings.....	6
1.4 Progress on Ongoing Performance Evaluation Work.....	7
<b>2. SCOPE AND METHODOLOGY .....</b>	<b>9</b>
2.1 Scope .....	9
2.2 APR 2012 Cohort.....	9
2.3 Methodology.....	10
<b>3. OUTCOMES AND SUSTAINABILITY OF OUTCOMES .....</b>	<b>19</b>
3.1 Rating Scale.....	19
3.2 Outcomes .....	20
3.3 Sustainability .....	25
<b>4. FACTORS AFFECTING ATTAINMENT OF PROJECT RESULTS ....</b>	<b>27</b>
4.1 Quality of Implementation and Execution.....	27
4.2 Cofinancing and Realization of Promised Cofinancing.....	28
4.3 Factors Attributed to Higher and Lower Project Performance.....	31
4.4 Trends in Project Extensions .....	33
<b>5. QUALITY OF M&amp;E DESIGN AND IMPLEMENTATION .....</b>	<b>37</b>
5.1 Rating Scale .....	37
5.2 Findings.....	37

<b>6. QUALITY OF TERMINAL EVALUATION REPORTS .....</b>	<b>41</b>
6.1 Findings.....	41
6.2 Comparison of Ratings from GEF Evaluation Office and GEF Agency Evaluation Offices.....	43
<b>7. MANAGEMENT ACTION RECORD .....</b>	<b>45</b>
7.1 Rating Approach.....	45
7.2 Findings.....	46
<b>8. PERFORMANCE MATRIX .....</b>	<b>50</b>
8.1 Performance Indicators.....	50
8.2 Findings.....	52
<b>ANNEXES</b>	
A. Projects Included in APR 2012 Cohort .....	55
B. Terminal Evaluation Report Review Guidelines .....	58
C. Notes on Methodology and Analysis.....	64
D. GEF Regions.....	66
<b>BIBLIOGRAPHY .....</b>	<b>67</b>
<b>BOXES</b>	
1.1 OPS Terminology Used in This Report .....	2
<b>FIGURES</b>	
2.1 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by Focal Area .....	12
2.2 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by Region .....	12
2.3 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by GEF Agency..	12
2.4 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by GEF Phase.....	12
3.1 Percentage of GEF Projects and Funding in Projects with Overall Outcome Ratings of Moderately Satisfactory or Above, by Year .....	21
3.2 Percentage of Rated Projects in GEF Replenishment Phase Cohorts with Overall Outcome Ratings of Moderately Satisfactory or Above .....	22
3.3 Trends in Project Performance by GEF Agency and APR Year Grouping.....	24
3.4 Trends in Project Performance by Focal Area and Region.....	24
3.5 Perceived Risks Underlying Projects with Sustainability Ratings of Moderately Unlikely or Below, APR 2005–12 Cohort.....	26
3.6 Percentage of GEF Projects and Funding in Projects with Outcomes Rated Moderately Satisfactory or Above and Sustainability Rated Moderately Likely or Above, by APR Year.....	26
4.1 Median and Total Ratio of Promised Cofinancing to GEF Funding, by Year.....	29
4.2 Trends in the Ratio of Total Promised Cofinancing to Total GEF Grant, by GEF Agency and Four-Year APR Groupings .....	30

4.3	Distribution among GEF projects by GEF Agencies and Four-Year APR Groupings of the Percentage of Promised Cofinancing Realized.....	30
4.4	Results of Analysis of Factors Attributed to High and Low Project Performance .....	32
4.5	Summary Statistics on One- and Two-Year Project Extensions, by GEF Agency and Project Size, within the APR 2005–12 Cohort.....	36
5.1	Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by Project Size and Year.....	39
5.2	Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by GEF Phase .....	39
5.3	Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by GEF Agency.....	40
6.1	Percentage of Terminal Evaluation Reports with Overall Quality Rated Moderately Satisfactory or Above by Project Size and GEF Agency.....	42
6.2	Quality of Terminal Evaluation Reporting on Individual Dimensions .....	43

## TABLES

1.1	Percentage of GEF Projects and Funding in GEF Projects with Overall Outcome Ratings in the Satisfactory Range, APR 2005–12 Cohorts.....	3
2.1	Composition of the APR 2005–11 and APR 2012 Cohorts.....	11
2.2	Sources of Terminal Evaluation Review Ratings Used in APR 2012.....	15
3.1	Distribution of GEF Projects by Outcome Rating .....	20
3.2	Distribution of GEF Funding in Projects by Overall Outcome Ratings.....	20
3.3	Percentage of Projects and Funding in Projects with Overall Outcome Ratings of Moderately Satisfactory or Above, by GEF Replenishment Phase.....	21
3.4	Overall Project Outcome Ratings by APR Cohort and Various Project Characteristics.....	23
3.5	Percentage of GEF Projects and Funding in Projects with Sustainability Ratings of Moderately Likely or above, by Year .....	25
4.1	Quality of Project Implementation and Execution, by Year.....	28
4.2	Quality of Implementation, by GEF Agency and Year .....	28
4.3	Promised and Realized Cofinancing for APR 2005–08, 2009–12, and 2005–12 Cohorts.....	29
4.4	Project Extensions by Project Size, GEF Agency, and APR Cohort Grouping .....	35
5.1	Quality of M&E Design, by Project Size and Year.....	38
5.2	Quality of M&E Implementation, by Project Size and Year.....	38
6.1	Percentage of Terminal Evaluation Reports Rated Moderately Satisfactory or Above, or Satisfactory or Above, by Project Size, GEF Agency, and Year of Report Completion.....	42
6.2	Comparison of Overall Outcome Ratings from GEF Agency Independent Evaluation Offices and from the GEF Evaluation Office for All Jointly Rated Projects, APR 2005–12 .....	44
7.1	GEF Management and GEF Evaluation Office Ratings of Council Decisions Tracked in MAR 2012.....	47
7.2	Summary of Council Decisions Graduated from the MAR.....	49
8.1	Performance Matrix for GEF Agencies and the GEF Overall.....	51





# Foreword

---

**T**he Global Environment Facility (GEF) Evaluation Office is pleased to present its ninth Annual Performance Report (APR). The report presents independent assessments of GEF activities on key performance parameters: project outcomes and sustainability, factors affecting attainment of project results, and quality of monitoring and evaluation arrangements.

The preliminary findings of this report were shared with the Secretariat and the GEF Agencies in an interagency meeting held in Washington, D.C., in April 2013. Draft versions of this report were also shared with the Secretariat and the Agencies, and their comments have been addressed in this report.

The APR 2012 was prepared as an input to the Fifth Overall Performance Study (OPS5). The

report was presented as an information document to the GEF Council during its June 2013 meeting. Although the report does not contain any recommendations, its findings and conclusions have informed the OPS5 recommendations on performance-related issues.

I would like to thank all of those involved for their support and criticism. The Evaluation Office remains fully responsible for the contents of this report.



Rob D. van den Berg  
Director, GEF Evaluation Office

# Acknowledgments

---

**N**eeraj Kumar Negi, Senior Evaluation Officer with the Global Environment Facility's (GEF's) Evaluation Office, is the leader of the Office's Performance Evaluation team, and he served as Task Team Leader for the Annual Performance Report 2012. The report was prepared by Joshua Schneck, Consultant.

The terminal evaluation review process was coordinated by Joshua Schneck. The terminal evaluation reviews were prepared by Sandra Romboli,

Evaluation Officer, and Anoop Agarwal and Sunpreet Kaur, Consultants.

The GEF's annual performance reports, including this year's, incorporate important contributions from the evaluation offices of the GEF Agencies, especially independent assessments of the terminal evaluations prepared by these offices. The GEF Evaluation Office appreciates the time and input provided by the GEF Secretariat and the Agencies during preparation of this report.

# Abbreviations

---

APR	Annual Performance Report	NPF	National Portfolio Formulation Exercise
CEO	Chief Executive Officer	OPS	overall performance study
FSP	full-size project	PMIS	Project Management Information System
FY	fiscal year	SCCF	Special Climate Change Fund
GEF	Global Environment Facility	SIDS	small island developing states
IDB	Inter-American Development Bank	STAR	System for Transparent Allocation of Resources
M&E	monitoring and evaluation	UNDP	United Nations Development Programme
LDCF	Least Developed Countries Fund	UNEP	United Nations Environment Programme
MAR	management action record	UNIDO	United Nations Industrial Development Organization
MSP	medium-size project		

All dollar amounts are U.S. dollars unless otherwise indicated.



# 1. Background and Main Conclusions

## 1.1 Background

The Annual Performance Report (APR), prepared by the Evaluation Office of the Global Environment Facility (GEF), provides a detailed overview of the performance of GEF activities and processes, key factors affecting performance, and the quality of monitoring and evaluation (M&E) systems within the GEF partnership. The APR provides GEF Council members, countries, Agencies, and other stakeholders information on the degree to which GEF activities are meeting their objectives and areas for further improvement.

APR 2012, the ninth APR produced by the GEF Evaluation Office, contains an assessment of 78 completed projects that are being covered for the first time. These projects account for \$289.5 million in GEF funding. The cohort consists of projects for which terminal evaluation reports have been submitted to the GEF Evaluation Office for the period October 1, 2011, to September 30, 2012.<sup>1</sup> To assess any trends, the performance of completed projects that have been reported on in earlier APRs is included as well. This year's APR is also being prepared as an input to the Fifth Overall

Performance Study (OPS5) being conducted by the Evaluation Office.

As in past years, APR 2012 reports on project outcomes, sustainability of project outcomes, quality of project implementation and execution, trends in cofinancing, trends in project completion extensions, quality of project M&E systems, and quality of terminal evaluation reports.

Findings presented are based primarily on the evidence found in terminal evaluation reports prepared by GEF Agencies at the time of project completion. Verification of performance ratings is largely based on desk review. The evaluation offices of the United Nations Development Programme (UNDP), the United Nations Environment Programme (UNEP), and the World Bank have been conducting desk reviews for verification of the project performance and ratings assessments provided in their respective Agency's terminal evaluations. The GEF Evaluation Office has started adopting the ratings from the Agency evaluation offices as past reviews have shown them to be fairly consistent with those provided by the GEF Evaluation Office. Where the evaluation offices of these Agencies have undertaken independent reviews of terminal evaluations, their ratings have been adopted. In other instances, ratings provided by the GEF Evaluation Office are reported.

This year's management action record (MAR) tracks the level of adoption of 21 separate decisions of the GEF Council: 10 that were part of MAR 2011, and 11 new decisions introduced during the

---

<sup>1</sup> A small number of recently completed projects for which terminal evaluations were submitted to the GEF Evaluation Office before the September 30 cutoff are not included in the APR 2012 cohort because the respective evaluation offices of the GEF Agencies were still undertaking independent reviews of the terminal evaluations.

two Council meetings held in fiscal year (FY) 2012. In addition to the decisions that pertain to the GEF Council, the Evaluation Office has started tracking the decisions of the Least Developed Countries Fund and Special Climate Change Fund (LDCF/SCCF) Council. One decision from that council's November 2011 meeting is tracked in MAR 2012.

The [performance matrix](#) presented in chapter 8 provides a summary of GEF Agency performance on key indicators. Of the 10 indicators presented in the matrix, based on the additional information on the APR 2012 cohort, values on five of the indicators have been updated.

## 1.2 Findings and Conclusions

**CONCLUSION 1: Eighty-seven percent of projects within the APR 2012 cohort have overall outcome ratings in the satisfactory range. While not necessarily indicative of a trend, the percentage of projects with outcome ratings in this range has risen between OPS cohorts.**

To date, overall outcomes of 486 completed projects have been rated, based on the extent to which project objectives were achieved; the relevance of project results to GEF strategies and goals, and to country priorities; and the efficiency with which project outcomes were achieved. Key findings of this assessment follow:

- Outcome ratings on GEF projects have, on average, risen over the past eight years, such that 86 percent of projects in the OPS5 cohort (see box 1.1) have ratings in the satisfactory range compared with 80 percent of projects in the OPS4 cohort (table 1.1). OPS5 cohort projects with outcome ratings in the satisfactory range account for 83 percent of GEF funding.
- A substantial improvement in the overall outcome ratings of UNEP and UNDP projects is seen between four-year OPS cohorts. Ninety-five percent of UNEP projects and 88 percent of UNDP projects within the OPS5 cohort have

### BOX 1.1 OPS Terminology Used in This Report

APR 2012 coincides with the release of OPS5 by the GEF Evaluation Office. To facilitate comparability between APR 2012 and the OPS5 reports, APR 2012 uses the terms "OPS4" and "OPS5" to refer to two distinct four-year APR cohorts of reviewed projects:

- OPS4 covers the APR 2005–08 cohorts
- OPS5 covers the APR 2009–12 cohorts

outcome ratings in the satisfactory range, compared to 74 percent and 78 percent of projects, respectively, in the OPS4 cohort.

- Two areas that continue to underperform relative to the larger GEF portfolio are projects in African states and projects in small island developing states (SIDS). Seventy-seven percent of African projects have outcome ratings in the satisfactory range, versus 85 percent for non-African projects. Similarly, 74 percent of SIDS projects have outcome ratings in the satisfactory range versus 84 percent for non-SIDS projects.
- A small rise in the percentage of GEF projects with overall outcome ratings in the satisfactory range is seen between projects in the GEF-2 (1999–2002) and GEF-3 (2003–06) replenishment period cohorts. The difference is statistically significant at a 90 percent confidence level.

**CONCLUSION 2: Sixty-six percent of projects in the APR 2012 cohort have sustainability ratings of moderately likely or above—similar to the long-term average. Financial risks continue to present the biggest threat to sustainability.**

Seventy-six of 78 projects within the APR 2012 cohort, and 468 projects within the APR 2005–12 cohort, were rated on likelihood of sustainability of outcomes. Key findings of this assessment follow:

- Roughly two-thirds of GEF projects and funding in projects in the APR 2012 cohort have

**TABLE 1.1 Percentage of GEF Projects and Funding in GEF Projects with Overall Outcome Ratings in the Satisfactory Range, APR 2005–12 Cohorts**

Criterion	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
% of projects with outcomes rated moderately satisfactory or above	82	84	73	81	91	91	80	87	84
Number of rated projects	39	64	40	62	55	46	102	78	486
% of GEF funding in projects with outcomes rated moderately satisfactory or above	84	88	69	74	92	88	79	80	81
Total GEF funding in rated projects (million \$)	255.3	254.3	198.3	275.3	207.8	158.6	414.3	289.5	2,053.4

sustainability of outcome ratings of moderately likely or above—just above the eight-year averages.

- Financial risks present the most common threat to project sustainability, with outcomes of 29 percent of projects in the APR 2005–12 cohort either unlikely or moderately unlikely to be sustained due to financial risks (out of 405 rated projects). Threats to project sustainability arising from institutional or governance risks are not far behind, with outcomes of 21 percent of projects either unlikely or moderately unlikely to be sustained due to institutional or governance factors (out of 407 rated projects).
- Within the APR 2005–12 cohort, just over half of GEF projects and funding have both outcome ratings in the satisfactory range and sustainability of outcome ratings of moderately likely or above. Percentages for the APR 2012 cohort are slightly higher than the long-term average, although the difference is not statistically significant.

**CONCLUSION 3: More than 80 percent of rated projects were assessed to have been implemented and executed in a satisfactory manner. Overall, jointly implemented projects have lower quality of implementation ratings than those implemented by a single Agency.**

The Evaluation Office has been tracking the quality of project implementation and execution of completed projects from FY 2008 onwards. Key findings from this assessment follow:

- Eighty-six percent of projects and funding within the APR 2012 cohort (out of 76 rated projects) have quality of implementation and quality of execution ratings in the satisfactory range.
- Projects under joint implementation, which comprise some 3.5 percent of GEF projects (17 projects) within the APR 2005–12 cohort, have lower quality of implementation ratings than those implemented by a single GEF Agency—63 percent versus 83 percent, respectively. This disparity probably reflects the increased complexity in jointly implemented projects and suggests that these projects do not receive the same degree of implementation support as they would warrant.

**CONCLUSION 4: There has been a significant increasing trend in the percentage of promised cofinancing realized.**

APR 2009 (GEF EO 2010b) concluded that the GEF benefits from mobilization of cofinancing through efficiency gains, risk reduction, synergies, and greater flexibility in terms of the types of projects it may undertake. Given these benefits, cofinancing has been a major performance indicator for the GEF. Some key findings from this year's assessment of trends in cofinancing follow:

- The amount of total promised cofinancing to the total GEF grant has increased 40 percent between OPS cohorts, from \$2.00 of promised cofinancing per dollar of GEF grant in the OPS4

cohort to \$2.80 of promised cofinancing per dollar of GEF grant in the OPS5 cohort.

- The amount of realized (actual) to promised cofinancing has increased 55 percent between OPS cohorts, from just over 90 percent of promised cofinancing realized in the OPS4 cohort to more than 140 percent of promised cofinancing realized in the OPS5 cohort.
- The increase in the median amount of realized to promised cofinancing between OPS cohorts is more modest—from 100 percent to 110 percent—indicating that a few outlying projects are responsible for generating large amounts of additional cofinancing.

**CONCLUSION 5: High quality of project management and a high level of support from government and nongovernmental stakeholders appear to be important determinants of high outcome achievements. Poor quality of project design and management, on the other hand, lead to low outcome achievements.**

---

To provide additional insights into the kinds of factors attributed to higher and lower project performance—i.e., projects with overall outcomes of moderately satisfactory or above, and those below this threshold—the GEF Evaluation Office conducted an in-depth desk review of the terminal evaluations in the OPS4 and OPS5 cohorts, looking for evidence within the evaluation narratives. Key findings include the following:

- Seventy-one percent of the 223 assessed terminal evaluations of projects with overall outcome ratings in the satisfactory range report that high quality of project management led to the project's overall high outcome achievements.
- Fifty-six percent of assessed terminal evaluations of projects with overall outcome ratings in the satisfactory range cite strong nonstate stakeholder support as positively contributing to the project's overall outcome rating.

- Poor project design is the factor most often cited as hindering project performance among the 81 assessed projects with overall outcome ratings below moderately satisfactory.
- Poor project management and low country support are the second and third most frequently cited factors attributed to poor performance in projects with overall outcome ratings below the satisfactory range (cited in 65 percent and 36 percent, respectively, of assessed projects).

Some evidence is found in assessed terminal evaluations that strong project management can sometimes overcome weaknesses in project design. Thirty-one, or 19 percent, of the 223 assessed projects with outcome ratings in the satisfactory range had significant weaknesses in design, according to the terminal evaluations, but succeeded in large part in meeting project expectations due to timely corrective actions taken by project management.

**CONCLUSION 6: Ratings on quality of M&E design and M&E implementation continue to be low.**

---

Despite changes in M&E policy designed to improve the quality of M&E systems,<sup>2</sup> ratings of M&E systems provided in terminal evaluations since APR 2006 continue to show gaps in M&E arrangements. Key findings of this assessment follow:

- Sixty-six percent, or two-thirds, of rated projects (out of 421 projects) have M&E design ratings in the satisfactory range, and ratings have remained essentially flat between OPS cohorts.<sup>3</sup>

---

<sup>2</sup> These changes include the adoption of the 2006 M&E Policy, and subsequent adoption of a revised M&E Policy in November 2010.

<sup>3</sup> Ratings on M&E design and implementation are not available for APR 2005, so the four-year OPS4 cohort (APR 2005–08) includes data from FY 2006–08 only.



- Sixty-eight percent of rated projects (out of 390 projects) have M&E implementation ratings in the satisfactory range. Ratings between OPS cohorts have declined slightly, from 71 percent in the OPS4 cohort to 66 percent in the OPS5 cohort. The difference is not statistically significant.
- Among rated projects, a greater proportion (74 percent) of projects approved during the GEF-3 replenishment period have M&E implementation ratings of moderately satisfactory or above compared to projects approved during the GEF-2 replenishment period (64 percent). The difference is statistically significant at a 90 percent confidence level.
- Among rated projects, a higher proportion of medium-size projects (MSPs) have M&E implementation ratings in the satisfactory range compared to full-size projects (FSPs): 73 percent versus 64 percent, respectively. Reasons for this difference are not well understood.
- Significant shifts in the M&E implementation ratings of two GEF Agencies are found between OPS cohorts. The percentage of UNDP projects with M&E implementation ratings in the satisfactory range has risen from 58 percent of projects in the OPS4 cohort to 75 percent of projects in the OPS5 cohort. The percentage of World Bank projects with M&E implementation ratings in the satisfactory range has declined from 80 percent of projects in the OPS4 cohort to 57 percent in the OPS5 cohort. Differences in ratings are statistically significant at a 95 percent confidence level.

**CONCLUSION 7: There has been a slight decline in the percentage of projects with project extensions between OPS cohorts.**

While project extensions—defined as time taken to complete project activities beyond that anticipated in project design documents—are not a strong

predictor of project outcomes, they do indicate that project activities were not completed in the time frame anticipated. In some situations, inability to complete the project in the planned time frame may lead to cost overruns, scaling down of activities, or greater time lag in achievement of outcomes. In other situations, extensions may allow project management to complete planned activities and outputs, thereby facilitating achievement of project outcomes.

Key findings from this year's assessment of trends in project extensions follow:

- Between OPS cohorts, there has been a slight decline in the percentage of projects with project extensions, from 81 percent of projects in the OPS4 cohort to 78 percent of projects in the OPS5 cohort. The difference is not statistically significant.
- Among projects with project extensions, the median lengths of extension are 18 months for FSPs and 12 months for MSPs.
- GEF Agencies differ substantially with regard to trends in project extensions.<sup>4</sup> Even when accounting for differences in the project size composition of GEF Agency portfolios, World Bank projects typically experience fewer and shorter project extensions than UNDP and UNEP projects.

**CONCLUSION 8: Eighty-six percent of terminal evaluations submitted in FY 2012 are rated in the satisfactory range for overall quality of reporting—in line with the long-term average.**

The GEF Evaluation Office has been reporting on the quality of terminal evaluations since APR 2004. To date, 527 terminal evaluations have been rated

<sup>4</sup>There is currently insufficient information on project extensions to report on trends for GEF Agencies other than UNDP, UNEP, and the World Bank.

for overall quality of reporting. Key findings of this analysis follow:

- Eighty-six percent of assessed terminal evaluations (out of 527) have ratings of moderately satisfactory or above for overall quality of reporting.
- The quality of terminal evaluations of MSPs has typically lagged that of FSPs. Using the threshold of satisfactory or above, only 46 percent of MSPs are rated as such compared with 59 percent of FSPs. The difference is statistically significant at a 95 percent confidence level.
- The quality of UNDP evaluations from 2005 onwards is higher than that of earlier years. At the same time, the percentage of UNDP terminal evaluations with overall ratings of satisfactory or above is 44 percent, compared with 63 percent for UNEP evaluations, and 61 percent for World Bank evaluations. This difference is statistically significant at a 95 percent confidence level.
- In general, reporting on project financing and M&E systems has not been as strong as reporting on other factors. The performance of terminal evaluations along these two dimensions has improved within the FY 2012 cohort. However, this cohort is not yet complete, and ratings may change as more terminal evaluations from this year become available in subsequent APRs.

### 1.3 Management Action Record Findings

The MAR tracks the level of adoption by the GEF Secretariat and/or the GEF Agencies of GEF Council decisions that have been made on the basis of GEF Evaluation Office recommendations. In addition, the Evaluation Office has begun tracking decisions of the LDCF/SCCF Council. One decision from the LDCF/SCCF Council's November 2011 meeting is included in MAR 2012.

Of the 21 separate GEF Council decisions tracked in MAR 2012, the Evaluation Office was able to verify management's actions on 14. None of the tracked decisions will be graduated this year, either because there has been insufficient time for management to act on Council decisions, or the Evaluation Office was unable to verify that a high level of adoption of the relevant Council decisions has occurred. All 21 decisions are still considered by the Evaluation Office to be relevant, and will be tracked in next year's MAR.

Five of the 10 GEF Council decisions tracked in previous MARs and in MAR 2012 have been rated by the Evaluation Office as having a substantial level of adoption. For the majority of newly tracked decisions, it is not yet possible to verify the level of adoption by management.

Management and the Evaluation Office are in agreement on the level of adoption for 8 of the 21 tracked decisions in MAR 2012; although for 7 tracked decisions, the Evaluation Office was unable to verify ratings either because insufficient information is available at this time, or proposals needed more time to be developed. Excluding the seven decisions where the Evaluation Office was unable to verify ratings, the level of agreement between management and the Evaluation Office is 57 percent—in line with that for MAR 2011 (58 percent) and MAR 2010 (66 percent). At the same time, in all cases where ratings have been provided by both management and the Evaluation Office and the ratings do not match, ratings by the GEF Evaluation Office are lower than those provided by management; in one case, substantially lower.

The largest gap between ratings provided by management and the GEF Evaluation Office is found in assessing the level of adoption of the GEF Council's request, based upon the Annual Country Portfolio Evaluation Report of 2012, that the Secretariat reduce the burden of monitoring requirements of multifocal area projects to a level comparable to that of single focal area projects. While the GEF Secretariat rates adoption of this decision as substantial,

the GEF Evaluation Office has assessed the actions taken thus far in response as negligible. The Office finds “no evidence that tracking tools burdens for MFAs [multifocal areas] have been reduced.” This finding is supported by UNDP and UNEP commentary included in the MAR management response as separate responses from these Agencies.

Since the commencement of the MAR in June 2006, the Evaluation Office has tracked the adoption of 111 Council decisions based on the recommendations of 32 evaluations. Overall, GEF management has been highly responsive to Council decisions, allowing for an ongoing reform process. To date, 86 (77 percent) tracked decisions have been graduated, including 65 for which a high or substantial level of adoption was reached at the time the decision was graduated.

Regarding adoption of the LDCF/SCCF Council decision, which is based on the Evaluation of the Special Climate Change Fund (GEF EO 2012a), both the Evaluation Office and the Secretariat are in agreement that, overall, a substantial level of adoption of the Council’s recommendations has occurred. This is particularly the case with respect to the LDCF/SCCF Council’s request that the Secretariat prepare proposals to ensure “transparency of the project pre-selection process and dissemination of good practices through existing channels.” At the same time, the Evaluation Office finds that additional work is needed by the Secretariat to fulfill the Council’s request that proposals be prepared to ensure greater visibility of the SCCF. This decision will be tracked in MAR 2013.

## 1.4 Progress on Ongoing Performance Evaluation Work

### NATIONAL PORTFOLIO FORMULATION EXERCISE MIDTERM EVALUATION

A midterm evaluation of the National Portfolio Formulation Exercise (NPFE) was initiated during

FY 2013. The evaluation will provide an assessment of NPFE activities undertaken and determine the overall relevance and effectiveness of the initiative, using a formative approach with a focus on learning.

During GEF-5 (2010–14), it was agreed that voluntary NPFEs would be encouraged as a tool to help interested recipient countries in establishing or strengthening national processes and mechanisms for GEF programming. NPFEs are expected to enhance country ownership in determining programming priorities in a given GEF replenishment period. They are also meant to set forth country priorities for the use of GEF resources in a transparent manner for the benefit of all GEF stakeholders—including the anticipated demand for resources, both from countries’ national allocations under the System for Transparent Allocation of Resources (STAR) and outside these allocations (GEF 2010). Another aim of the NPFE process is to strengthen country capacity to coordinate ministries and other involved stakeholders from both the private and public sectors.

The GEF Secretariat has been providing grants since 2010 for up to \$30,000 to support the costs of these exercises, mainly consisting of broad consultation meetings with key stakeholders. The expected output is a national portfolio formulation document that summarizes each country’s GEF programming priorities. To date, 42 countries have participated in the exercise—with or without GEF funding. More than half of these have been implemented in Africa (53 percent).

The midterm evaluation is currently ongoing and is in its data gathering phase. Several countries are being visited in order to interview key stakeholders that took part in the NPFE consultations. An online survey is being used to reach other stakeholders and to increase the coverage and outreach of this evaluation. A blog has been established on the GEF Evaluation Office website to elicit a discussion of this type of formative/learning evaluation approach. The NPFE midterm evaluation is expected to be finalized during the fall of 2013.

## STAR MIDTERM EVALUATION

During FY 2013, the Evaluation Office initiated a midterm evaluation of STAR performance. The evaluation aims to assess

- the extent to which the STAR's design facilitates allocation and utilization of scarce GEF resources to enhance global environmental benefits,
- the extent to which the STAR promotes transparency and predictability in allocation of GEF resources and strengthens country-driven approaches,
- the level of flexibility that has been provided by the STAR in allocation and utilization of GEF resources,
- the efficiency and effectiveness of the STAR implementation process, and
- the extent to which the Resource Allocation Framework midterm review has been followed

up on in the STAR through relevant Council decisions and general lessons learned.

The approach paper of the evaluation has been prepared. It outlines a variety of methodological approaches that the evaluation team will use to respond to the key questions of the evaluation. The team will use a mix of quantitative and qualitative tools and methods, including desk review of relevant documents; assessment of the appropriateness, adequacy, and scientific validity of resource allocation indexes by an expert panel; portfolio review and statistical modeling to assess the STAR's effect on the resource flows and the nature of the GEF portfolio; survey of key stakeholders to gather information on STAR design and implementation; and an online survey of a wider set of stakeholders. Various activities of the evaluation—such as portfolio analysis, desk review of other resource allocation frameworks, online survey, fieldwork, and panel review of STAR design—are presently under way. This evaluation will be completed in time to be an input to OPS5.

## 2. Scope and Methodology

### 2.1 Scope

The APR provides a detailed overview of the performance of GEF projects and funding, as well as analysis of some key factors affecting performance and M&E systems. APR 2012 includes the following:

- **An overview of the extent to which GEF projects and funding are achieving desired outcomes (chapter 3).** The assessment provided covers 486 completed projects within the APR 2005–12 cohort for which ratings on overall project outcomes are available. Also presented here are ratings on the sustainability of project outcomes and an assessment of the risks to project sustainability.
- **Analysis of factors affecting project outcomes (chapter 4).** Factors covered include quality of project implementation and execution, realization of cofinancing, and trends in project extensions. Also included are findings from a GEF Evaluation Office assessment identifying factors associated with higher and lower outcome achievements.
- **Quality of M&E design and implementation (chapter 5).** Ratings on quality of M&E design and M&E implementation are presented. Ratings are available from FY 2006 onwards.
- **Assessment of the quality of terminal evaluation reports submitted by the GEF Agencies to the GEF Evaluation Office (chapter 6).**

Trends in the overall quality of reporting, as well as trends in reporting along individual performance dimensions, are presented, based on the year in which terminal evaluation reports were completed.

- **Presentation of the MAR (chapter 7).** The MAR, which assesses the degree to which relevant GEF Council decisions based on GEF Evaluation Office recommendations have been adopted by GEF management, is presented. In addition, the Evaluation Office has started tracking decisions of the LDCF/SCCF Council. Twenty-one separate GEF Council decisions are tracked in MAR 2012: 10 that were part of MAR 2011, and 11 decisions that appear for the first time in MAR 2012. A single decision from the LDCF/SCCF Council is tracked.
- **Presentation of the performance matrix (chapter 8).** The performance matrix, which has been reported on since APR 2007, provides a summary of GEF Agency performance on key indicators. Ten indicators are tracked in the matrix included in APR 2012. Based on the additional information on the APR 2012 cohort, values on five of the indicators have been updated.

### 2.2 APR 2012 Cohort

The assessment of performance presented in the APR is primarily based on evidence provided in terminal evaluation reports. Seventy-eight projects,

totaling \$289.5 million in GEF funding, for which terminal evaluation reports have been submitted to the Evaluation Office from the period October 1, 2011, to September 30, 2012, are covered for the first time.<sup>1</sup> A complete listing of the 78 projects comprising the APR 2012 cohort is found in [annex A](#). To assess any trends in performance, the performance of cohorts reported on in prior APR years is included as well.

Table 2.1 and figures 2.1–2.4 present a side-by-side overview of the APR 2012 and APR 2005–11 cohorts in terms of focal area and regional composition,<sup>2</sup> GEF Agency representation, and GEF phase. In general, the composition of the APR 2012 cohort is similar to that of the larger APR 2005–11 cohort, with some key differences. Compared with the APR 2005–11 cohort, the APR 2012 cohort is distinguished by the following:

- A lower share of climate projects (18 percent in APR 2012 versus 25 percent in APR 2005–11), although a similar level of funding, and an increased share of land degradation projects (10 percent versus 3 percent) and multifocal projects (12 percent versus 6 percent)
- Less funding in Europe and Central Asia (13 percent versus 22 percent), and Asia (16 percent versus 24 percent), and additional funding in global projects (23 percent versus 9 percent)
- Heavy representation of UNDP among GEF Agencies in the 2012 cohort, with UNDP responsible for implementation of 65 percent of projects and 44 percent of funding; a relatively

small percentage of projects is implemented by the World Bank (8 percent), but these account for 22 percent of GEF funding in APR 2012

- Three projects in APR 2012 implemented by the Inter-American Development Bank (IDB), three projects implemented by the United Nations Industrial Development Organization (UNIDO), and three projects under joint implementation by IDB–World Bank, UNDP–World Bank, and UNDP–UNEP
- The majority of projects (63 percent) within the APR 2012 cohort are from GEF-3, while 47 percent of APR 2005–11 projects are from GEF-2; GEF-4 projects also make up a larger percentage of the current APR cohort—nearly one-quarter of APR 2012 projects, although only 7 percent of GEF funding

The median length of projects in the APR 2005–12 cohort is 61 months, or just over 5 years.

## 2.3 Methodology

Reporting on project outcomes and sustainability, factors affecting outcomes, quality of M&E, and quality of terminal evaluations—discussed in chapters 3, 4, 5, and 6, respectively—are based on analysis of the ratings and information provided in terminal evaluations which have been first reviewed by the GEF Evaluation Office and/or the evaluation offices of GEF Agencies. GEF activities under the Small Grants Programme, as well as enabling activities with GEF funding below \$0.5 million, are not required to submit terminal evaluations and are not covered in this report.<sup>3</sup>

<sup>1</sup> A small number of recently completed projects for which terminal evaluations were submitted to the GEF Evaluation Office before the September 30, 2012, cutoff date are not included in the APR 2012 cohort because the respective evaluation office of the relevant GEF Agency was still undertaking independent review of the terminal evaluations.

<sup>2</sup> For a description of the GEF regions used in this report, see [annex D](#).

<sup>3</sup> The GEF classifies projects based on the size of the associated GEF grant; whether GEF funding supports country activities related to the conventions on biodiversity, climate change, and persistent organic pollutants; and implementation approach. These categories are FSPs, MSPs, enabling activities, and programmatic approaches. For a complete description, see the GEF [website](#).



TABLE 2.1 Composition of the APR 2005–11 and APR 2012 Cohorts

		APR 2005–11				APR 2012			
Criterion		Projects (#)	Projects (%)	Funding (million \$)	Funding (%)	Projects (#)	Projects (%)	Funding (million \$)	Funding (%)
Total projects and funding		413	—	1,769.4	—	78	—	289.5	—
Projects and funding with outcome ratings		408	—	1,763.9	—	78	—	289.5	—
Focal area composition <sup>a</sup>	Biodiversity	205	50	811.2	46	37	47	127.3	44
	Climate change	102	25	436.5	25	14	18	73.7	26
	International waters	51	13	323.3	18	7	9	55.5	19
	Land degradation	13	3	16.7	1	8	10	11.3	4
	Multifocal	23	6	57.4	3	9	12	16.4	6
	Other	14	3	118.9	7	3	4	5.1	2
Regional composition <sup>a</sup>	Africa	89	22	331.1	19	13	17	49	17
	Asia	99	24	428.8	24	14	18	47.6	16
	Europe & Central Asia	89	22	393.1	22	19	24	37.2	13
	Latin America and Caribbean	92	23	457.4	26	22	28	90	31
	Global	39	10	151.5	9	10	13	65.7	23
Lead GEF Agency <sup>a</sup>	UNDP	177	43	623.4	35	51	65	126.9	44
	UNEP	56	14	157.1	9	12	15	49	17
	World Bank	159	39	889.6	50	6	8	64.5	22
	Other	2	<1	12.1	1	6	8	11.6	4
	Joint	14	3	81.7	5	3	4	37.5	13
GEF phase <sup>a</sup>	Pilot	12	3	98.1	6	0	0	0	0
	GEF-1	65	16	516.5	29	2	3	40.5	14
	GEF-2	193	47	837.9	47	9	12	51	18
	GEF-3	127	31	296.4	17	49	63	179	62
	GEF-4	11	3	15	1	18	23	19.1	7

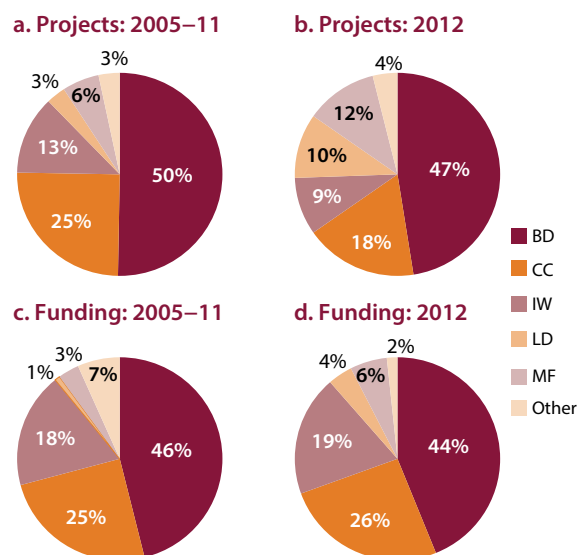
a. Describes only the 486 projects (408 in APR 2005–11 and 78 in APR 2012) with outcome ratings, as these are the projects on which performance is primarily compared in the analysis below.

Among the 491 projects contained in the APR 2005–12 cohort are two enabling activities that have met the threshold for review. For analysis, these have been grouped with FSPs based on the size of the associated GEF funding.

All of the terminal evaluations used for analysis and reporting in APRs are first reviewed to verify that ratings are properly substantiated and, where needed, to provide additional or revised

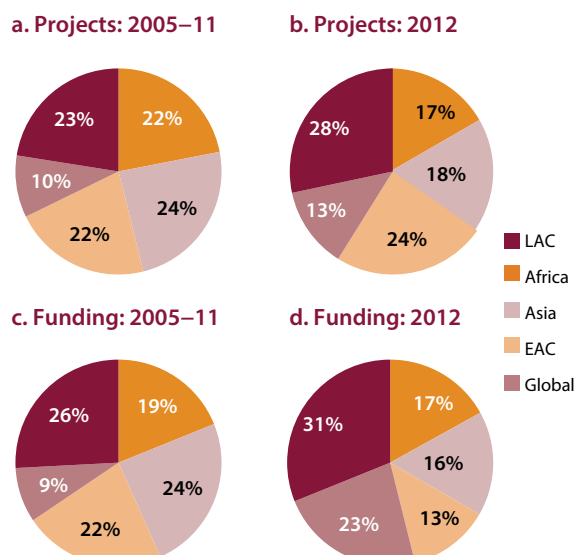
ratings (such as for quality of terminal evaluations). For earlier APR years, this oversight was performed entirely by the GEF Evaluation Office. Beginning in 2009, the Office began accepting ratings from the independent evaluation offices of the World Bank, UNEP, and—subsequently—UNDP. This approach, which reduces duplicative work, follows the GEF Evaluation Office finding that ratings from these three evaluation offices are largely consistent with

**FIGURE 2.1 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by Focal Area**



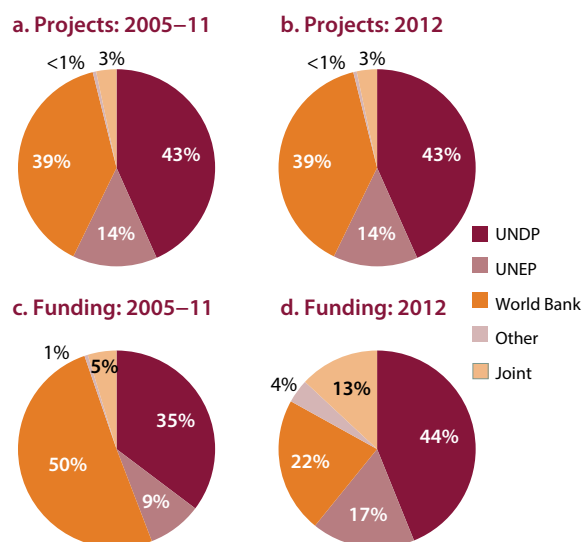
NOTE: BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal.

**FIGURE 2.2 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by Region**

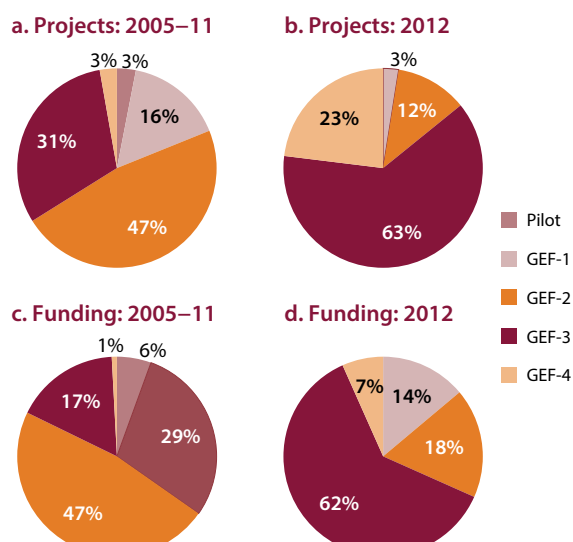


NOTE: EAC = Europe and Central Asia; LAC = Latin America and the Caribbean.

**FIGURE 2.3 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by GEF Agency**



**FIGURE 2.4 Distribution of Projects and Funding in APR 2005–11 and APR 2012 Cohorts, by GEF Phase**





those provided by the GEF Evaluation Office (GEF EO 2009b). The Office will consider accepting the ratings provided by the evaluation offices of the other GEF Agencies when there is a sufficient record of ratings on which to compare consistency and when the ratings from the two offices are found to be consistent.

## RATINGS APPROACH

The principle dimensions of project performance on which ratings are first provided in terminal evaluations, and in subsequent GEF Evaluation Office or GEF Agency evaluation office reviews of terminal evaluations, are described here in brief, and in full in [annex B](#).

- **Project outcomes.** Projects are evaluated on the extent to which project objectives, as stated in the project's design documents approved by the GEF Council and/or the GEF Chief Executive Officer (CEO),<sup>4</sup> were achieved or are expected to be achieved; the relevance of project results to GEF strategies and goals and to country priorities; and the efficiency, including cost-effectiveness, with which project outcomes and impacts were achieved. A six-point rating scale, from highly satisfactory to highly unsatisfactory, is used.
- **Sustainability of project outcomes.** Projects are evaluated on the likelihood that project benefits will continue after implementation. To arrive at an overall sustainability rating, evaluators are asked to identify and assess key risks to sustainability of project benefits, including financial risks, sociopolitical risks, institutional/governance risks, and environmental risks. A four-point rating scale, from likely to be sustained to unlikely to be sustained, is used.
- **Quality of implementation and quality of execution.** Since FY 2008, the Evaluation Office has been assessing the quality of project implementation and the quality of project execution. Quality of implementation primarily covers quality of project design, as well as quality of supervision and assistance provided by the GEF Agency to its executing agency throughout project implementation. Quality of execution primarily covers the effectiveness of the executing agency in performing its roles and responsibilities. In both instances, the focus is on factors that are largely within the control of the respective Implementing/executing agency. A six-point rating scale, from highly satisfactory to highly unsatisfactory, is used.
- **Quality of M&E systems.** M&E facilitates adaptive management during project implementation, and assessment of project outcomes and impacts after project completion. The quality of project M&E arrangements is evaluated in two ways: (1) assessment of the project's M&E design, including whether indicators used are SMART,<sup>5</sup> whether relevant baselines are established, and whether M&E activities are properly budgeted; and (2) the degree and quality of M&E during implementation. A six-point rating scale, from highly satisfactory to highly unsatisfactory, is used for quality of M&E design and quality of M&E implementation.
- **Quality of terminal evaluation reports.** Terminal evaluations, which are the primary source of information on which project performance is assessed, are themselves assessed for quality, consistency, coverage, and quality of lessons and recommendations; and to evaluate the degree to which project ratings provided in terminal evaluations are properly substantiated. A six-point rating scale from highly satisfactory to highly

<sup>4</sup> All GEF FSPs require approval by the GEF Council and endorsement by the GEF CEO prior to funding; MSPs require only the GEF CEO's approval.

<sup>5</sup> SMART indicators are Specific, Measurable, Achievable and Attributable, Relevant and Realistic, and Time-bound, Timely, Trackable and Targeted. See GEF EO (2010c) for a complete description.

unsatisfactory is used to indicate the quality of terminal evaluations.

## PROCEDURE FOR GEF EVALUATION OFFICE REVIEW OF TERMINAL EVALUATIONS

When terminal evaluations are reviewed by the GEF Evaluation Office prior to inclusion in the APR, as well as for oversight purposes, the procedure is as follows. Using a set of detailed guidelines to ensure that uniform criteria are applied (see [annex B](#) for these guidelines), Evaluation Office reviewers assess the degree to which project ratings provided in the terminal evaluations are properly substantiated, and address the objectives and outcomes set forth in the project design documents approved by the GEF Council and/or the GEF CEO. In the process of drafting a terminal evaluation review, a peer reviewer with substantial experience in assessing terminal evaluations provides feedback on the report. This feedback is incorporated into subsequent versions of the report.

When a primary reviewer proposes downgrading project outcome ratings from the satisfactory range to the unsatisfactory range, a senior evaluation officer in the GEF Evaluation Office also examines the review to ensure that the proposed rating is justified.

In cases where a terminal evaluation report provides insufficient information to make an assessment or to verify the report's ratings on any of the performance dimensions, the Evaluation Office rates the project as "unable to assess," and excludes it from further analysis on the respective dimension.

Reviews are then shared with the GEF Agencies and, after their feedback is taken into consideration, finalized.

## SOURCE OF RATINGS REPORTED IN APR 2012

As noted above, prior to FY 2009, the GEF Evaluation Office reviewed all terminal evaluations

reported on in APRs and verified the ratings provided therein. Beginning in FY 2009, the Evaluation Office began accepting ratings from the independent evaluation offices of UNEP, the World Bank, and subsequently UNDP. Because the procedure used by the GEF Agencies for arriving at overall ratings in terminal evaluations is not identical to that used by the GEF Evaluation Office, comparability between ratings from APR 2009 and later cohorts and earlier APR cohorts is of some concern.

The GEF Evaluation Office has been tracking the consistency between ratings provided by itself and the evaluation offices of the GEF Agencies. This is accomplished through random sampling and GEF Evaluation Office review of a portion of terminal evaluations included in the APR for which ratings have been provided by Agency evaluation offices. To date, ratings provided by the Agency evaluation offices are largely consistent with those provided by the GEF Evaluation Office. A small—4 percent—increase in the percentage of projects with overall outcome ratings of moderately satisfactory or above is found among sampled reviews from Agency evaluation offices, compared with those from the GEF Evaluation Office (see [chapter 6](#) for a complete breakdown of sampled reviews). This difference is not statistically significant, however. Moreover, adjusting for a possible bias would not lead to significant changes in the findings presented in APRs from 2009 onwards. The Office will continue to track the consistency of ratings going forward.

For projects implemented by GEF Agencies other than UNDP, UNEP, and the World Bank, the GEF Evaluation Office currently provides final project ratings. Additionally, where ratings are not provided by the independent evaluation offices of UNDP, UNEP and the World Bank, the Office provides final ratings. Examples of these projects include all projects under joint implementation; MSPs implemented by the World Bank, for which the Bank's Independent Evaluation Group does not provide review; and projects where independent

review of terminal evaluations is not received in a timely manner.

Table 2.2 lists the sources of terminal evaluation review ratings used for analysis and reporting in APR 2012.

## COFINANCING AND MATERIALIZATION OF COFINANCING

The reporting in [section 4.2](#) on cofinancing and materialization of cofinancing is based on information in project design documents, as well as information provided by GEF Agencies on completed projects both through terminal evaluation reports and other project reports. Reporting covers APR cohorts from 2005 to 2012, for which information on the amount of promised cofinancing is available for all 491 projects, and information on actual (realized) cofinancing is available for 426 projects.

## FACTORS ATTRIBUTED TO HIGHER AND LOWER PROJECT PERFORMANCE

[Section 4.3](#) presents an analysis of factors cited in OPS4 and OPS5 cohort terminal evaluations as important contributors to project outcome ratings. The methodology used to identify these factors is as follows.

Among the 281 terminal evaluations that comprise the OPS5 cohort, evaluations were first sorted between those with overall outcome ratings of moderately satisfactory and above (239 evaluations) and those with overall outcome ratings below this threshold (41 evaluations). To this latter group were added 40 evaluations with overall outcome ratings below moderately satisfactory from the OPS4 cohort. Within these two groups, terminal evaluations were then reviewed to determine whether the respective narratives specifically identify factors having a direct impact on project outcomes or that were important contributors to project outcomes. That is, for projects with overall outcome ratings of moderately satisfactory or above, did the terminal evaluation narrative identify factors that were reported to have had a direct effect, or important indirect effect, on overall outcome achievements? Similarly, for projects with overall outcome ratings below moderately satisfactory, did the terminal evaluation narrative identify factors that directly hindered, or made an important indirect contribution that hindered, the project's overall outcome achievements?

Of the 239 projects in the OPS5 cohort with overall outcome ratings of moderately satisfactory or above, 223 terminal evaluations reported factors that led to high outcome achievements. Of the 81 projects in the OPS4 and OPS5 cohorts with overall outcome ratings below moderately

TABLE 2.2 Sources of Terminal Evaluation Review Ratings Used in APR 2012

Source	Project	Total
UNDP Evaluation Office	51 UNDP projects	51
UNEP Evaluation Office	11 UNEP projects	11
World Bank Independent Evaluation Group	3 World Bank projects	3
GEF Evaluation Office	3 joint implementation projects	13
	3 IDB projects	
	3 UNIDO projects	
	1 UNEP project (GEF ID 1776)	
	3 World Bank projects (GEF IDs 112, 1081, 1221)	
Total		78

satisfactory, all terminal evaluations reported on factors that led to lower outcome achievements.

Factors contributing to outcome ratings were then grouped into non-overlapping categories. For factors positively contributing to overall outcome ratings of moderately satisfactory or above, the following four categories emerged:

- **Project design**—projects for which the project’s design is reported in the terminal evaluation as positively contributing to project outcome achievements. Design factors cited included having a sound logical framework, generation of project outputs that directly enhanced local livelihoods, projects closely tailored to the circumstances of the project site(s), and project design that established and/or facilitated strong communication between project actors and stakeholders.
- **Project management**—projects where project management is reported in the terminal evaluation as positively contributing to project outcome achievements. Management strengths cited include the capacity and commitment of management; quality of supervision provided, including strong technical inputs; and adaptive management.
- **High country support**—projects where strong country support is mentioned in the terminal evaluation as positively contributing to project outcome ratings. Projects evidencing strong country support include those where national agencies/ministries with a role in project execution are seen as actively driving the project forward, and/or project support is provided in the form of additional cofinancing or through supporting legislation or policy.
- **Stakeholder support (nongovernmental actors)**—projects where strong support from nonstate stakeholders is mentioned in the terminal evaluation as positively contributing to project outcome ratings. Such stakeholders include private sector actors, nongovernmental organizations, academia, and others.

For projects with overall outcome ratings below moderately satisfactory, the following four categories emerged for factors that directly or indirectly led to lower outcome achievements:

- **Project design**—projects for which the project’s design is mentioned in the terminal evaluation as hindering the project’s outcome ratings. Design factors cited included significant problems in the project’s logical framework, failure to tailor the project adequately to the local context, failure to adequately budget project activities, overly ambitious project goals, and poor choices in executing arrangements.
- **Project management**—projects for which poor management is mentioned in the terminal evaluation as hindering the project’s outcome ratings. This category includes problems related to both project implementation and execution, such as insufficient capacity of the executing agency, poor supervision by the GEF Agency, insufficient technical inputs, poor coordination with project partners, financial mismanagement, major issues with procurement, and high staff turnover.
- **Low country support**—projects for which weak support/commitment from the state (or some levels or sectors of state administration) is reported in the terminal evaluation as hindering the project’s outcome ratings. Evidence cited includes excessive delays regarding permitting of project activities, failure to advance legislation or policy critical to the success of the project, and development plans that conflict with the project.
- **Exogenous factors**—projects for which exogenous factors are reported to have hindered the project’s outcome achievements. Exogenous factors cited include political instability, natural disasters, economic crises, and changes in foreign exchange markets.

While the eight categories defined above are non-overlapping in terms of what kinds of factors each respective category covers, individual projects can and often do cite more than one factor as contributing to the project's overall outcome ratings. Percentages of projects in the combined categories reported in [section 4.3](#) are therefore greater than 100 percent.

## PROJECT EXTENSIONS

The reporting in [section 4.4](#) on trends in project extensions is based on information in the GEF Project Management Information System (PMIS), as well as in project terminal evaluations. Project extensions are defined as time taken from the start of the project to complete project activities beyond that anticipated in project approval documents. These exclude any delays that may occur prior to the start of project activities. Reporting covers APR cohorts from 2005 to 2012, for which information on project extensions is available for 466 projects.

## MANAGEMENT ACTION RECORD ASSESSMENT

At the request of the GEF Council, the GEF Evaluation Office tracks the level of adoption by relevant actors within the GEF partnership (here referred to broadly as GEF management) of GEF Council decisions that have been made on the basis of GEF Evaluation Office recommendations. The MAR is updated annually and reported on in the APR. The procedure for compiling the MAR is as follows. The GEF Evaluation Office produces a working document containing all of the relevant GEF Council decisions being tracked for the current MAR. This includes all Council decisions from the prior year MAR that continue to be tracked because the level of adoption is not yet sufficient to warrant graduation. Decisions are graduated from the MAR when a high level of adoption has been achieved or the decision is no longer relevant. For

decisions that continue to be tracked, a full record of prior GEF management actions and ratings, as well as of GEF Evaluation Office ratings, is provided in the working document. The working document also includes all relevant Council decisions that have been adopted at the GEF Council meetings in the preceding calendar year.

Following distribution of the working document to GEF management, management provides self-assessment and ratings on the level of adoption of each tracked Council decision. Once management completes its self-assessment and ratings on the level of adoption of tracked Council decisions, it shares these with the GEF Evaluation Office. The Evaluation Office then provides its own assessment and ratings on adoption. The completed MAR is then published and reported in the APR.

## PERFORMANCE MATRIX

The performance matrix, first presented in APR 2007 (GEF EO 2008b), provides a summary of the performance of three GEF Agencies and the GEF Secretariat on relevant parameters. Performance on five indicators—project outcomes, materialization of cofinancing, project extensions, M&E implementation quality, and quality of terminal evaluations—is assessed annually by the Evaluation Office. Performance on three other indicators—quality of supervision and adaptive management, realism of risk assessment, and quality of project M&E arrangements—is assessed every two to four years through special appraisals. Independence of terminal evaluations and review of terminal evaluations is appraised through assessment of the process followed in conducting terminal evaluations through field verifications and based on interviews with relevant staff and consultants of the partner Agencies. Performance on one parameter included in the performance matrix—project preparation elapsed time—is the subject of an ongoing study by the Evaluation Office, and the study's findings are not yet available.

## REVIEW OF FINDINGS

The preliminary findings of this report were presented to and discussed with the GEF Secretariat and GEF Agencies during an interagency meeting held in Washington, D.C., on April 11, 2013. GEF Evaluation Office reviews of project

terminal evaluation reports have been shared with GEF Agencies for comment, and their feedback has been incorporated into this final report. The analysis presented herein also incorporates feedback received from both the GEF Secretariat and the GEF Agencies at the interagency meeting.



## 3. Outcomes and Sustainability of Outcomes

This chapter presents verified ratings on outcomes for GEF projects. To date, outcomes of 491 completed projects have been assessed, which account for \$2.06 billion in GEF funding. Of these, the GEF Evaluation Office has provided or adopted outcome ratings on 486 projects, including all 78 projects in the APR 2012 cohort. An additional 408 rated projects are found in the APR 2005–11 cohort. Together, these 486 projects account for \$2.05 billion in GEF funding.

Also presented are ratings on likelihood of sustainability of outcomes and an assessment of the perceived risks to project sustainability.

### 3.1 Rating Scale

As described in [chapter 2](#), project outcomes are rated based on the extent to which project objectives were achieved, the relevance of project results to GEF strategies and goals and to country priorities, and the efficiency with which project outcomes were achieved. A six-point rating scale is used to assess overall outcomes, with the following categories:

- **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.

For likelihood of sustainability of outcomes, and overall assessment on the likelihood of project benefits continuing after project closure, a four-point rating scale is used, with the following categories:

- **Likely.** There are no risks to the sustainability of project outcomes.
- **Moderately likely.** There are moderate risks to the sustainability of project outcomes.
- **Moderately unlikely.** There are significant risks to the sustainability of project outcomes.
- **Unlikely.** There are severe risks to the sustainability of project outcomes.

### METHODOLOGICAL NOTE

It is not uncommon for the results frameworks of projects to be modified during project implementation. This presents a challenge to project evaluation in that assessing project outcomes based on original outcome expectations may discourage adaptive management. To address this concern, for projects

where modifications were made to project objectives, outcomes, and outputs without a downscaling of the project's overall scope, the Evaluation Office assesses outcome achievements based on the revised results framework. In instances where the scope of project objectives, outcomes, and outputs were downscaled, the project's original outcomes and/or objectives are used to measure project performance.

## 3.2 Outcomes

Tables 3.1 and 3.2 and figure 3.1 present overall outcome ratings on GEF projects and funding in the APR 2005–12 cohorts. For the APR 2012

cohort, 87 percent of projects have overall outcome ratings in the satisfactory range (i.e., projects with overall outcome ratings of moderately satisfactory or above), which is a little higher than the eight-year average of 84 percent. Similarly, 80 percent of funding is invested in projects with outcomes rated in the satisfactory range, which is in line with the long-term average. While not necessarily indicative of a trend, the percentage of projects with outcomes rated in the satisfactory range in the OPS5 cohort (APR 2009–12) is 85 percent compared with 80 percent for the previous four-year OPS4 cohort (APR 2005–08). This difference in the proportion of projects with outcomes rated in the satisfactory

TABLE 3.1 Distribution of GEF Projects by Outcome Ratings

Outcome rating	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
Percentage distribution									
Highly satisfactory	3	6	3	5	4	9	4	6	5
Satisfactory	54	44	35	52	56	28	38	41	44
Moderately satisfactory	26	34	35	24	31	54	38	37	35
<b>Moderately satisfactory or above</b>	<b>82</b>	<b>84</b>	<b>73</b>	<b>81</b>	<b>91</b>	<b>91</b>	<b>80</b>	<b>87</b>	<b>83</b>
Moderately unsatisfactory	10	14	8	13	9	4	15	13	11
Unsatisfactory	8	2	18	5	0	4	5	3	5
Highly unsatisfactory	0	0	3	2	0	0	0	0	<1
Number									
Projects rated on outcomes	39	64	40	62	55	46	102	78	486

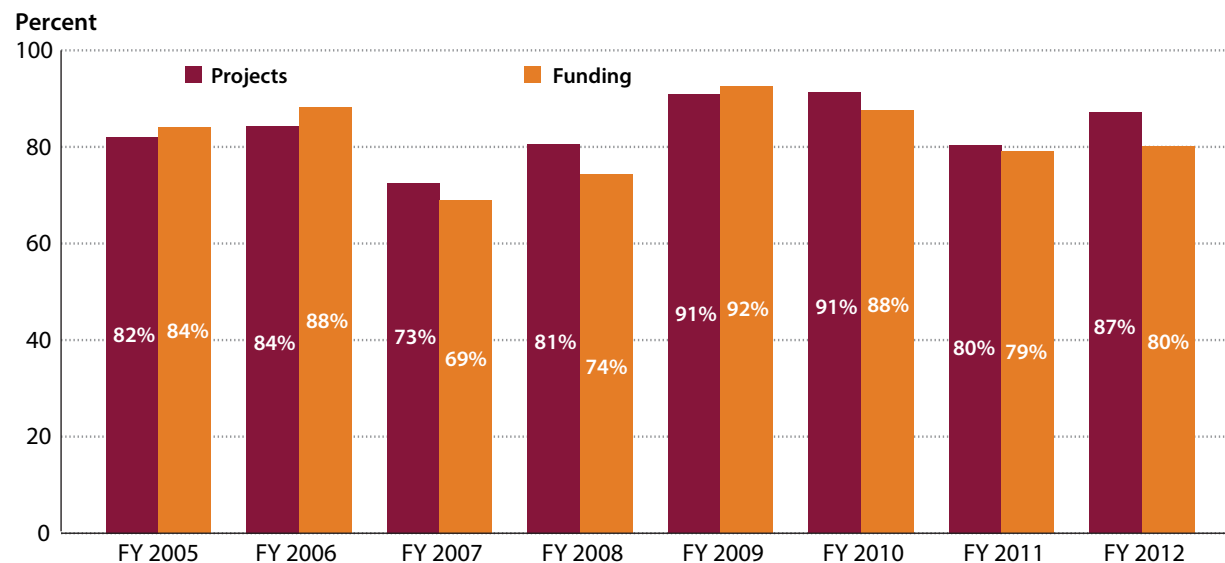
TABLE 3.2 Distribution of GEF Funding in Projects by Overall Outcome Ratings

Outcome rating/criteria	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
Highly satisfactory	<1	6	5	8	3	2	6	2	4
Satisfactory	64	30	18	55	56	44	34	34	42
Moderately satisfactory	20	53	46	12	33	41	39	40	35
<b>% of GEF funding in projects with outcomes rated moderately satisfactory or higher</b>	<b>84</b>	<b>88</b>	<b>69</b>	<b>74</b>	<b>92</b>	<b>88</b>	<b>79</b>	<b>80</b>	<b>81</b>
Moderately unsatisfactory	15	11	14	13	8	9	16	20	14
Unsatisfactory	1	1	12	10	0	4	4	4	5
Highly unsatisfactory	0	0	5	3	0	0	0	0	1
Total GEF funding in rated projects (million \$)	255.3	254.3	198.3	275.3	207.8	158.6	414.3	289.5	2,053.4

NOTE: Details may not sum to 100 percent due to rounding.



**FIGURE 3.1 Percentage of GEF Projects and Funding in Projects with Overall Outcome Ratings of Moderately Satisfactory or Above, by APR Year**



range between OPS cohorts is statistically significant at a 90 percent confidence level. In short, outcome ratings on GEF projects have, on average, risen over the past eight years such that more than 80 percent of projects, and of funding in projects, in the OPS5 cohort has overall outcome ratings in the satisfactory range.

Overall outcome ratings can also be assessed by GEF replenishment phase, as shown in table 3.3 and figure 3.2. Because GEF phase cohorts are not

complete, and a very limited number of ratings are available for the pilot, GEF-1, and GEF-4 replenishment phases, care must be taken in assessing any trends in outcome ratings by GEF phase at this time. That said, a small rise in the percentage of GEF projects with overall outcome ratings of moderately satisfactory or above is seen between the GEF-2 and GEF-3 phase cohorts. The difference is statistically significant at a 90 percent confidence level.

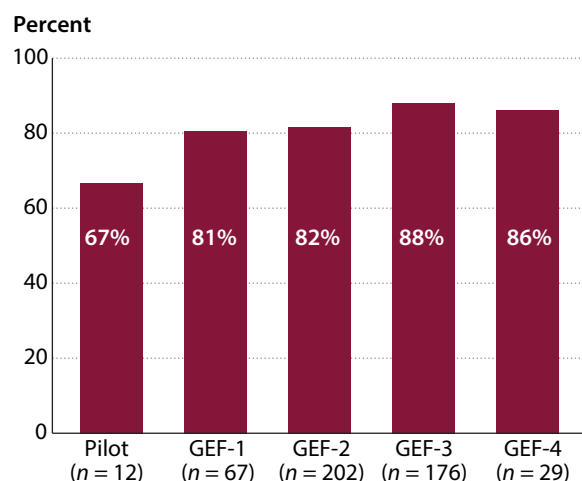
**TABLE 3.3 Percentage of Projects and Funding in Projects with Overall Outcome Ratings of Moderately Satisfactory or Above, by GEF Replenishment Phase**

Criterion	Pilot	GEF-1	GEF-2	GEF-3	GEF-4	GEF-5
Number of approved projects <sup>a</sup>	106	142	344	490	660	384
% of approved projects that are completed and covered in APRs	19	58	65	36	4	0
% of approved projects that are completed, covered in APRs, and with outcome ratings	11	47	59	36	4	0
% of rated projects with overall outcomes of moderately satisfactory or above	67	81	82	88	86	—
% of funding in projects with overall outcomes of moderately satisfactory or above <sup>b</sup>	58	83	79	89	72	—

a. As of April 30, 2013. Excludes Small Grant Programme projects and projects involving less than \$0.5 million.

b. Percentage covers only funding in projects with ratings for overall outcomes.

**FIGURE 3.2 Percentage of Rated Projects in GEF Replenishment Phase Cohorts with Overall Outcome Ratings of Moderately Satisfactory or Above**



NOTE: The difference in the shares of GEF-2 and GEF-3 projects with outcome ratings in the satisfactory range is statistically significant at a 90 percent confidence level.

While no projects stemming from the GEF-5 replenishment period, and just 29 GEF-4 replenishment period projects, are found in current APR year cohorts, the 86 percent proportion of GEF-4 projects with outcome ratings of moderately satisfactory or above is higher than the 80 percent target for GEF-5 projects and the 75 percent target for GEF-4 projects established at the respective replenishment negotiations (GEF Assembly 2006; GEF Secretariat 2010). Assuming the current level of project performance continues, GEF projects overall appear to be on track to meet the targets for their respective GEF replenishment periods.

Overall outcomes can be further assessed by looking at key project traits including the responsible GEF Agency; executing agency; focal area, size, and scope; and where the project was implemented (table 3.4). Because the number of projects within yearly APR cohorts in these groupings is often small, they are presented here in two four-year cohorts: APR 2005–08 (OPS4), and APR 2009–12 (OPS5).

Figure 3.3 shows overall outcome ratings for projects in the OPS4 and OPS5 cohorts, by GEF Agency. Overall outcome ratings have risen quite dramatically for UNEP, with from 74 percent to 95 percent of projects having overall outcome ratings of moderately satisfactory or above. A less striking but still pronounced increase occurred for UNDP projects, with 88 percent of projects in the OPS5 cohort having ratings of moderately satisfactory or above compared with 78 percent in the OPS4 cohort. Projects implemented by the World Bank show a slight decline in overall outcome ratings between the two cohorts—from 85 percent to 79 percent of projects with ratings of moderately satisfactory or above. While the increase in outcome ratings for UNEP, UNDP, and all projects is statistically significant at a 90 percent confidence level, the difference in performance between cohorts for World Bank projects is not statistically significant at this level of confidence.

GEF Agencies other than UNDP, UNEP, and the World Bank are not represented in the APR 2005–08 cohort, but are among the Agencies in the 2009–12 cohort, implementing eight projects. For these eight projects—four of which are implemented by UNIDO, three by IDB, and one by the Asian Development Bank—seven have overall outcome ratings of moderately satisfactory or above. This is similar to the figures for the overall GEF portfolio.

A separate category, not shown in figure 3.3, includes projects under joint implementation by two or more GEF Agencies. There are 17 projects under joint implementation in the APR 2005–12 cohort—3 within the OPS4 cohort, and 11 within the OPS5 cohort. Thirteen of these jointly implemented projects, or 76 percent, have overall outcome ratings of moderately satisfactory or above. Although this is below the eight-year average of 84 percent, the difference is not statistically significant. Projects under joint implementation also have lower ratings for quality of implementation, for which an association with lower outcome ratings has been found in the APR 2005–12 cohort (see [chapter 4](#)).

TABLE 3.4 Overall Project Outcome Ratings by APR Cohort and Various Project Characteristics

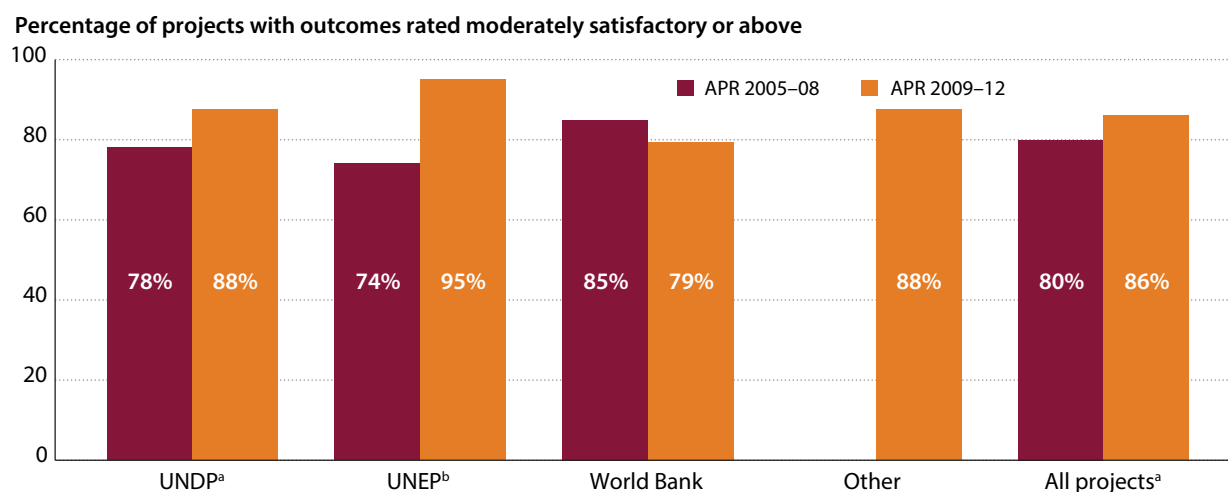
Characteristic		APR 2005–08		APR 2009–12		APR 2005–12	
		Number of rated projects	% of projects with outcomes rated MS or above	Number of rated projects	% of projects with outcomes rated MS or above	Number of rated projects	% of projects with outcomes rated MS or above
GEF Agency	UNDP	82	78	146	88	228	84
	UNEP	27	74	41	95	68	87
	World Bank	93	85	72	79	165	82
	Other	0	—	8	88	8	88
	Joint	3	67	14	79	17	76
Executing agency	Government or parastatal agency	108	82	159	84	267	84
	NGO or foundation	53	79	48	90	101	84
	Bilateral or multilateral agency	35	71	65	89	100	83
	Other, inc. private sector orgs.	9	100	9	78	18	89
Focal area	Biodiversity	116	81	126	87	242	84
	Climate change	49	84	67	82	116	83
	International waters	23	78	35	89	58	84
	Land degradation	4	50	17	94	21	86
	Multifocal	9	67	23	87	32	81
	Other	4	100	13	77	17	82
Region	Africa	45	73	57	81	102	77
	Asia	56	84	57	88	113	86
	Europe and Central Asia	36	78	72	88	108	84
	Latin America and the Caribbean	51	84	63	87	114	86
	Global	17	82	32	88	49	86
Country characteristic <sup>a</sup>	Fragile state	12	67	17	88	29	79
	SIDS	14	71	13	77	27	74
	Least developed country	22	77	23	83	45	80
	Landlocked	25	84	43	93	68	90
Size	FSP	114	78	160	85	274	82
	MSP	91	84	121	88	212	86
Scope	National (single-country project)	147	83	204	85	351	84
	Regional	41	71	45	89	86	80
	Global	17	82	32	88	49	86
GEF phase	Pilot	11	73	1	0	12	67
	GEF-1	52	81	15	80	67	81
	GEF-2	125	81	77	83	202	82
	GEF-3	17	82	159	89	176	88
	GEF-4	0	—	29	86	29	86
	All projects	205	80	281	86	486	84

NOTE: — = not available; MS = moderately satisfactory; NGO = nongovernmental organization. The difference in the shares of African and non-African projects with outcome ratings of MS or above is statistically significant at a 95 percent confidence level. The difference in the shares of SIDS and non-SIDS projects with outcome ratings of MS or higher is statistically significant at a 90 percent confidence level, as is the difference in the shares of GEF-2 and GEF-3 projects with overall outcome ratings of MS or higher.

a. For regional and global projects, includes only those projects in which all participating countries were members of the relevant group.

b. FSPs include two enabling activities based on size of the GEF grant.

**FIGURE 3.3 Trends in Project Performance by GEF Agency and APR Year Grouping**



NOTE: Projects under joint implementation are not included in individual Agency percentages.

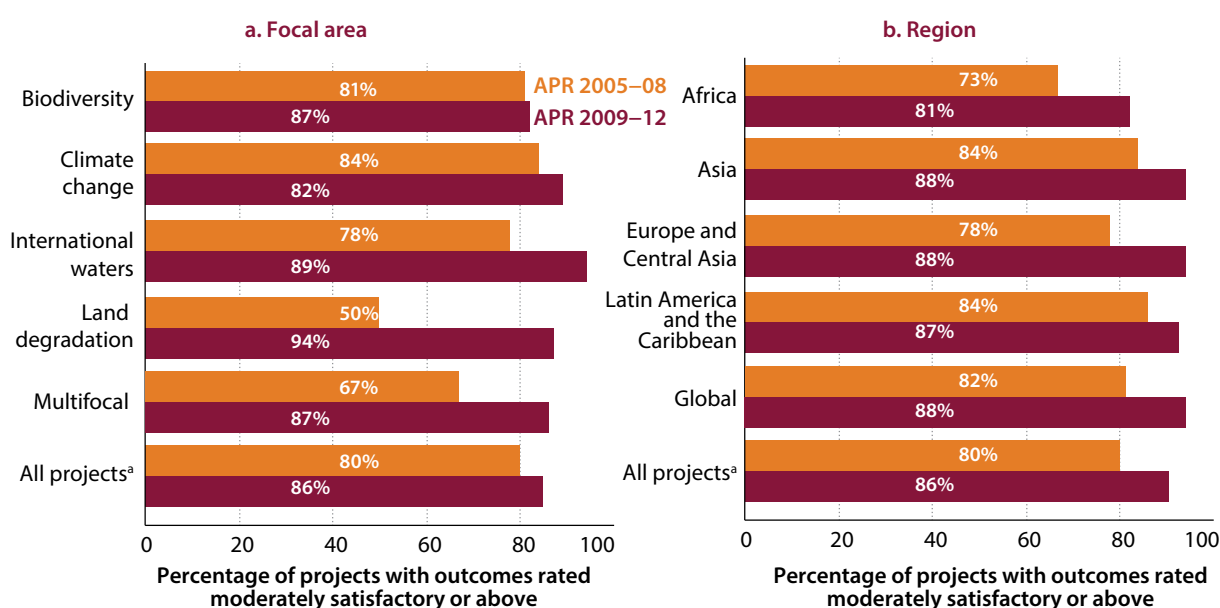
a. The difference in the shares between APR groupings of projects rated moderately satisfactory or above is statistically significant at a 90 percent confidence level.

b. The difference in the shares between APR groupings of projects rated moderately satisfactory or above is statistically significant at a 95 percent confidence level.

Figure 3.4 shows overall outcome ratings on projects in the two OPS cohorts by GEF focal area and region. Although a fair amount of variability

is seen among the focal areas and within the two OPS cohorts, much of this can be attributed to the small number of projects in each focal area/

**FIGURE 3.4 Trends in Project Performance by Focal Area and Region**



a. The difference in the shares between APR groupings of projects rated moderately satisfactory or above is statistically significant at a 90 percent confidence level.

four-year cohort. For example, the focal area exhibiting the biggest swing in overall outcome ratings—land degradation—has only four projects in the OPS4 cohort. None of the differences in four-year outcome ratings in focal areas or regions is statistically significant. Among regions, projects in Africa have performed, on average, below projects in other regions, with 77 percent of African projects having overall outcome ratings of moderately satisfactory or above for the APR 2005–12 cohort, versus 85 percent for non-African projects. The difference is statistically significant at a 95 percent confidence level.

Other project groupings not shown in figures 3.3 and 3.4 but presented in table 3.4 are those based on type of executing agency, country characteristics, the size and scope of the project, and the GEF replenishment phase in which projects originate. Among these groupings, projects implemented in SIDS have performed on average below projects in other countries. For the eight-year APR 2005–12 cohort, 74 percent of projects implemented in SIDS have overall outcome ratings of moderately satisfactory or above compared with 84 percent for non-SIDS projects. This difference is statistically significant at a 90 percent confidence

level. With the exception of differences between African and non-African projects described above, none of the variances in outcome ratings between other project groupings were found to be statistically significant.

### 3.3 Sustainability

Of the 491 projects in the APR 2005–12 cohort, 468 have been rated on sustainability of outcomes, which assesses the likelihood of project benefits continuing after project closure. Table 3.5 presents ratings on sustainability of project outcomes. Of projects with sustainability ratings in the APR 2012 cohort, 66 percent have ratings of moderately likely or above. This is a little higher than the eight-year average of 61 percent. Similar numbers are found when assessing sustainability ratings by GEF funding. For the APR 2012 cohort, the percentage of GEF funding in projects with sustainability ratings of moderately likely or above is 65 percent, which is just above the eight-year average of 63 percent.

To provide some insights into the perceived threats to project sustainability, key risks to the continuation of project benefits following project closure—including financial risks, sociopolitical

**TABLE 3.5 Percentage of GEF Projects and Funding in Projects with Sustainability Ratings of Moderately Likely or above, by Year**

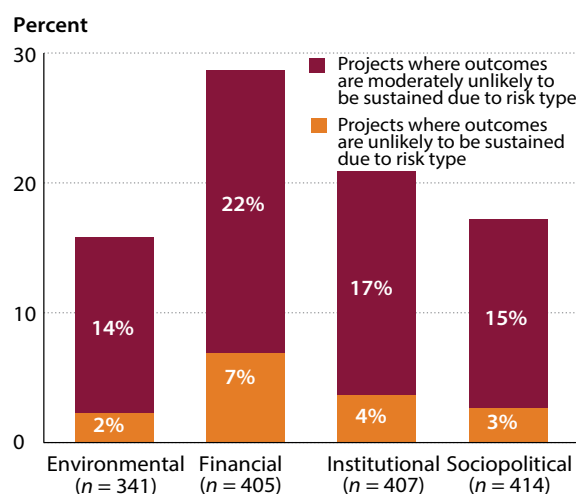
Criterion	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
% of projects with sustainability ratings of ML or above	49	65	59	57	71	63	58	66	61
% of projects with outcomes rated MS or above and sustainability rated ML or above	44	61	51	55	67	63	55	59	57
Number of rated projects	39	54	39	60	55	46	99	76	468
% of GEF funding in projects with sustainability ratings of ML or above	65	60	55	58	66	75	60	65	63
% of GEF funding in projects with outcomes rated MS or above and sustainability rated ML or above	60	56	44	56	65	75	55	61	58
Total GEF funding in rated projects (million \$)	255.3	218.3	182.1	251.4	207.8	158.6	411.6	258.4	1,943.5

NOTE: ML = moderately likely; MS = moderately satisfactory.

risks, institutional/governance risks, and environmental risks—are identified in terminal evaluation reviews. Figure 3.5 presents the findings from this assessment of risks to sustainability for the APR 2005–12 cohort. As shown in the figure, financial risks present the most common perceived threat to project sustainability, with 29 percent of project outcomes either unlikely or moderately unlikely to be sustained due to financial risks (out of 405 rated projects). Threats to project sustainability arising from institutional or governance risks are not far behind, with 21 percent of project outcomes either unlikely or moderately unlikely to be sustained for institutional or governance reasons (out of 407 rated projects).

Figure 3.6 and the shaded rows in table 3.5 present information on the percentage of projects that have both overall outcome ratings of moderately satisfactory or above and sustainability ratings of moderately likely or above. Fifty-nine percent of projects and 61 percent of GEF funding within the APR 2012 cohort meet this threshold, compared with 57 percent and 58 percent,

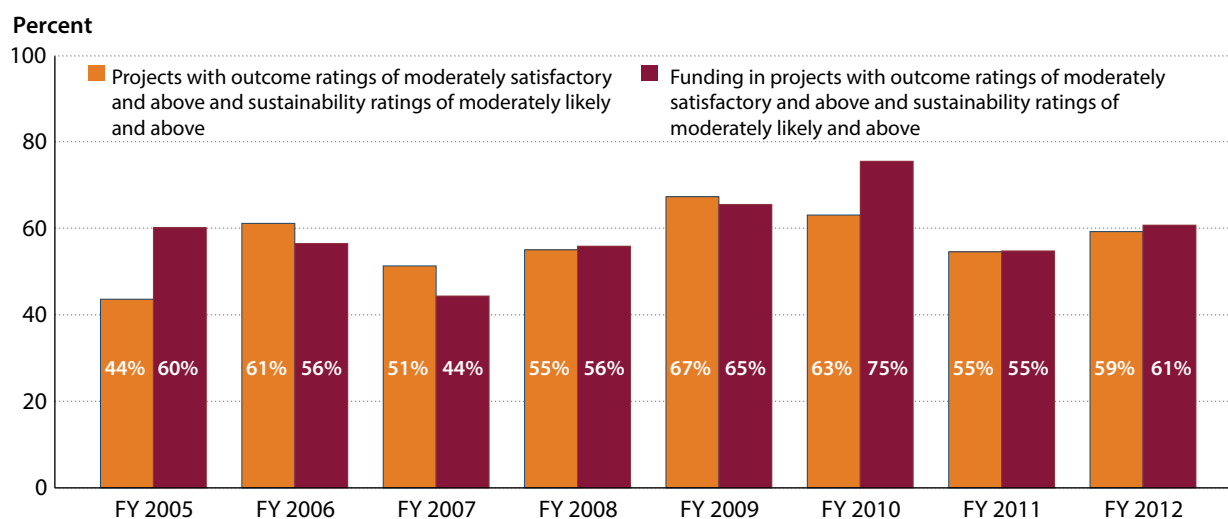
**FIGURE 3.5 Perceived Risks Underlying Projects with Sustainability Ratings of Moderately Unlikely or Below, APR 2005–12 Cohort**



NOTE: Details may not sum to 100 percent due to rounding.

respectively, in the eight-year APR cohort. In short, a little over half of GEF projects and of GEF funding are meeting both commonly used thresholds for positive outcomes and sustainability ratings within the APR 2005–12 cohort.

**FIGURE 3.6 Percentage of GEF Projects and Funding in Projects with Outcomes Rated Moderately Satisfactory or Above and Sustainability Rated Moderately Likely or Above, by APR Year**



## 4. Factors Affecting Attainment of Project Results

**M**any factors may affect project outcomes, from project design and quality of project implementation and execution, to the operational context in which projects take place, to exogenous factors beyond the control of project management. Given the range and complexity of these factors, it is difficult to isolate variables and determine their specific effects on project outcomes. At the same time, associations between factors and project outcomes, and between factors themselves, can be determined.

This chapter reports on three factors for which strong associations to project outcomes have been found in the APR 2005–12 cohort: quality of project implementation, quality of project execution, and realization of promised cofinancing (see [annex C](#) for the methodology and results of this analysis). In addition to reporting on ratings for these factors, the GEF Evaluation Office conducted a desk review of terminal evaluations within the APR 2009–12 cohort to identify in more detail factors associated with higher and lower performing projects—i.e., projects with overall outcome ratings of moderately satisfactory and above, and those with outcome ratings below this threshold. The results of this analysis are presented here. Lastly, trends in project completion extensions are reported.

### 4.1 Quality of Implementation and Execution

From FY 2008 onwards, the Evaluation Office has assessed quality of project implementation and

execution. As noted in chapter 2, **quality of implementation** covers the quality of project design, as well as the quality of supervision and assistance provided by GEF Implementing Agencies to executing agencies throughout project implementation. **Quality of execution** primarily covers the effectiveness of executing agencies in performing their roles and responsibilities. In both instances, the focus is on factors that are largely within the control of the respective agency.

Table 4.1 presents ratings on quality of project implementation and execution. For both criteria, the percentage of projects with ratings of moderately satisfactory and above exceeds 80 percent for all cohorts except APR 2008, where the percentage of projects with quality of implementation ratings of moderately satisfactory or above was 72 percent. Five-year averages for quality of implementation and execution are 82 percent and 84 percent, respectively.

Table 4.2 looks at quality of project implementation by GEF Agency and APR year. A fair amount of variation can be seen in the ratings from year to year, due in part to the small number of projects in individual APR year cohorts for any given agency. The percentage of UNEP projects within the five-year 2008–12 cohort with quality of implementation ratings of moderately satisfactory or above (80 percent) is slightly below that for UNDP and World Bank projects. The difference is not statistically significant, however. What is significant is the share of projects under joint implementation with

TABLE 4.1 Quality of Project Implementation and Execution, by Year

Criterion	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
% of projects with quality of implementation rated moderately satisfactory or above	72	85	86	81	86	82
Number of rated projects	60	55	43	101	76	335
% of projects with quality of execution rated moderately satisfactory or above	83	87	86	81	86	84
Number of rated projects	59	54	43	98	76	330

TABLE 4.2 Quality of Implementation, by GEF Agency and Year

Criterion	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
% of UNDP projects with quality of implementation rated moderately satisfactory or above	68	77	93	88	86	83
Number of rated projects	28	22	15	58	51	174
% of UNEP projects with quality of implementation rated moderately satisfactory or above	71	87	67	80	82	80
Number of rated projects	7	15	6	5	11	44
% of World Bank projects with quality of implementation rated moderately satisfactory or above	78	94	89	76	100	84
Number of rated projects	23	17	19	29	6	94
% of jointly implemented projects with quality of implementation rated moderately satisfactory or above	50	—	67	50	100	63 <sup>a</sup>
Number of rated projects	2	0	3	8	3	16
% of all projects with quality of implementation rated moderately satisfactory or above	72	85	86	81	86	82
Number of rated projects	60	55	43	101	76	335

a. The difference in the share of jointly and nonjointly implemented projects with quality of implementation ratings of moderately satisfactory or above is significant at a 95 percent confidence level.

quality of implementation ratings of moderately satisfactory or above. Only 63 percent of rated projects (10 of 16 projects) under joint implementation are so rated, compared with 83 percent of nonjointly implemented projects. This finding suggests that jointly implemented projects do not receive the same degree or quality of implementation support as nonjointly implemented projects. The difference is statistically significant at a 95 percent confidence level.

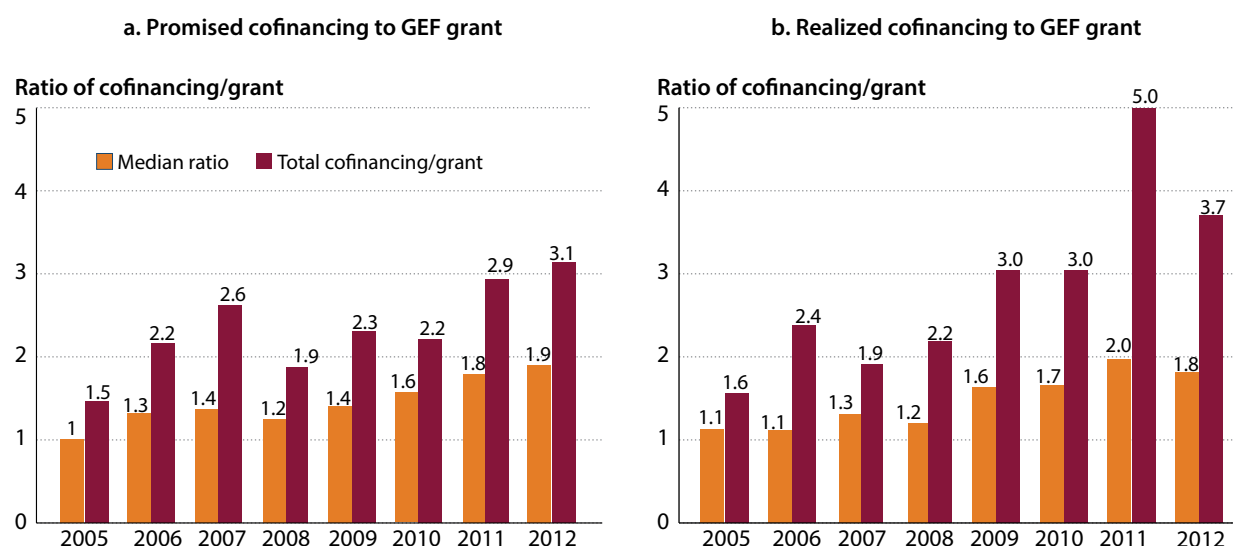
## 4.2 Cofinancing and Realization of Promised Cofinancing

APR 2009 concluded that the GEF gains from mobilization of cofinancing through efficiency gains, risk reduction, synergies, and greater flexibility in terms of the types of projects it may undertake. Given these benefits, cofinancing has been a key performance indicator for the GEF.

Figure 4.1 displays both the median and total ratio of promised cofinancing to GEF grant, as well as the median and total ratio of actual cofinancing



FIGURE 4.1 Median and Total Ratio of Promised Cofinancing to GEF Funding, by APR Year



NOTE: Data on promised cofinancing available for 491 projects in APR 2005–12 cohort; data on actual financing available for 426 projects in APR 2005–12 cohort.

to GEF grant by year.<sup>1</sup> The figure clearly shows a general increasing trend in the level of promised

<sup>1</sup> Total refers to the total amount of promised cofinancing over the total amount of GEF funding for an APR year cohort.

and realized cofinancing to GEF funding among APR cohorts from 2005 to 2012. When assessed in four-year APR cohorts, as shown in table 4.3, the change in cofinancing is considerable. The amount of total promised cofinancing to the total GEF grant has risen from \$2.00 of promised cofinancing

TABLE 4.3 Promised and Realized Cofinancing for APR 2005–08, 2009–12, and 2005–12 Cohorts

Criterion	APR 2005–08	APR 2009–12	APR 2005–12
Total projects with data on promised cofinancing	210	281	491
Total GEF funding (million \$)	988.7	1,070.3	2,058.9
Total promised cofinancing (million \$)	1,970.1	2,952.9	4,923
Median ratio promised cofinancing to GEF grant	1.2	1.6	1.4
Ratio of total promised cofinancing to total GEF grant	2.0	2.8	2.4
Total projects with data on actual (realized) cofinancing	162	264	426
Total realized cofinancing (million \$) <sup>a</sup>	1,425.6	4,008.3	5,433.8
Median ratio of realized cofinancing to GEF grant	1.2	1.8	1.6
Ratio of total realized cofinancing to total GEF grant <sup>b</sup>	2.0	4.0	3.2
Median ratio of realized to promised cofinancing <sup>b</sup>	1.0	1.1	1.0
Ratio of total realized to total promised cofinancing <sup>b</sup>	0.9	1.4	1.3

a. Total realized cofinancing is likely higher than reported figure as data are missing for 65 projects in the APR 2005–12 cohort.

b. Ratios include only projects for which data on realized cofinancing are available.

per dollar of GEF grant for the OPS4 cohort to \$2.80 of promised cofinancing per dollar of GEF grant for the OPS5 cohort—an increase of 40 percent. An even more dramatic rise is seen in the total amount of realized cofinancing to the total GEF grant between OPS cohorts. This metric has risen from \$2 of realized cofinancing per dollar of GEF grant in the OPS4 cohort to \$4 in the OPS5 cohort—a 100 percent increase.

Perhaps more important than the absolute amount of promised or realized cofinancing within APR year cohorts is the percentage of promised cofinancing realized, as this gives an indication of the degree to which project financing needs anticipated in project design documents have been met. As shown in the bottom half of table 4.3, there has been a substantial increase in the percentage of promised cofinancing realized from FY 2005 to FY 2012. For the OPS4 cohort, a little over 90 percent of promised cofinancing materialized. For the OPS5 cohort, more than 140 percent of promised cofinancing materialized—an increase of about 55 percent. At the same time, the increase in the median ratio of actual to promised cofinancing is far less dramatic—from 1.0 to 1.1—indicating that a few outlying projects

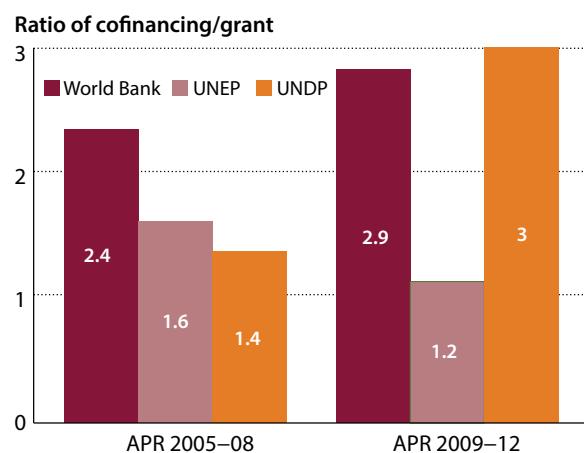
are responsible for generating large amounts of additional cofinancing.

Trends in cofinancing can also be distinguished by GEF Agency, as shown in figure 4.2. The amount of promised cofinancing to GEF funding has more than doubled for UNDP projects, rising from \$1.40 in cofinancing per dollar of GEF funding for projects within the APR 2005–08 cohort, to \$3.00 of cofinancing per dollar of GEF funding for projects within the APR 2009–12 cohort. The ratio has risen for World Bank projects, although less dramatically, and fallen slightly for UNEP projects. Considering all projects within the APR 2005–12 cohort, the ratio of total promised cofinancing to total GEF grant is higher for World Bank and UNDP projects compared with UNEP projects, at 2.7, 2.4, and 1.3, respectively.

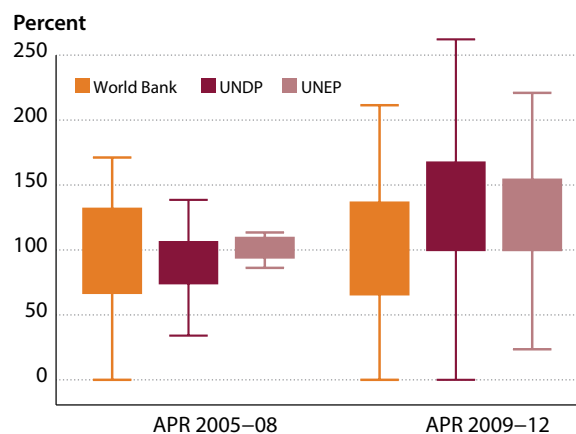
Figure 4.3 shows distribution among projects by GEF Agencies of the percentage of promised cofinancing that materialized. While the median value is at or close to 100 percent for all three GEF Agencies in both four-year APR groupings,<sup>2</sup> some

<sup>2</sup>There are currently insufficient data to report cofinancing percentages for GEF Agencies other than UNDP, UNEP, and the World Bank.

**FIGURE 4.2 Trends in the Ratio of Total Promised Cofinancing to Total GEF Grant, by GEF Agency and Four-Year APR Groupings**



**FIGURE 4.3 Distribution among GEF projects by GEF Agencies and Four-Year APR Groupings of the Percentage of Promised Cofinancing Realized**



movement is seen in materialized cofinancing for UNDP and UNEP projects within the APR 2005–12 cohort. For both these Agencies, the percentage of projects realizing more than 100 percent of promised cofinancing has risen to the point where 75 percent of all projects in the APR 2009–12 grouping realized at least 100 percent of promised cofinancing, and 25 percent of projects realized at least 150 percent of promised cofinancing. For World Bank projects, the numbers have remained fairly stable, with the inner quartile (25th to 75th percentile) of projects in the APR 2005–12 cohort realizing between 67 percent and 134 percent of promised cofinancing.

### 4.3 Factors Attributed to Higher and Lower Project Performance

To provide additional insights into the kinds of factors attributed to higher and lower project performance—i.e., projects with overall outcomes of moderately satisfactory or above and those below this threshold—the GEF Evaluation Office conducted a desk review of the 281 terminal evaluations within the OPS5 cohort, looking for evidence within the evaluations’ narratives. A similar analysis looking at factors associated with lower performing projects was performed on 40 terminal evaluations in the OPS4 cohort, and reported on in APR 2008. To provide greater comparability between the APR 2012 and APR 2008 studies, for the APR 2012, the 40 OPS4 terminal evaluations with overall outcome ratings below moderately satisfactory were combined with those meeting the same threshold in the OPS5 cohort. These 81 terminal evaluations of lower performing projects were then assessed together as a group (see [chapter 2](#) for a complete description of the methodology).

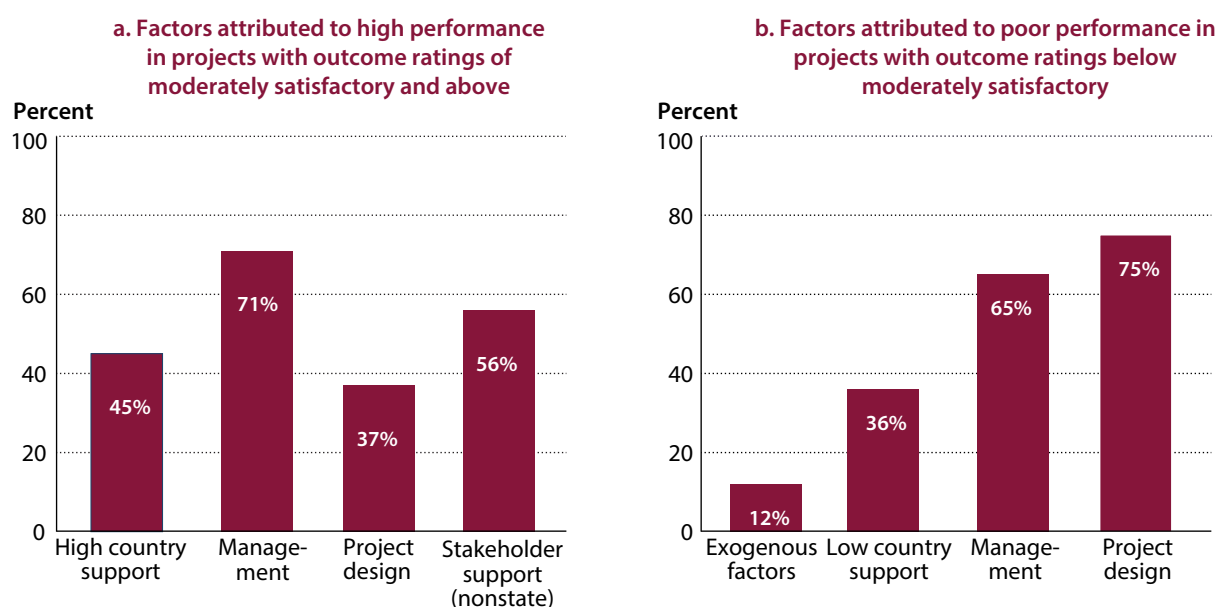
The results, shown in figure 4.4, suggest that project outcomes of higher performing projects are highly reflective of the quality of project management, and also frequently benefit from high

stakeholder and country support. Seventy-one percent (159 of 223) of assessed terminal evaluations of projects with overall outcome ratings of moderately satisfactory or above cite project management as positively contributing to the project’s overall outcome rating. Management strengths described in terminal evaluations include the capacity and commitment of management; the quality of supervision provided, including strong technical inputs; and adaptive management. Also noteworthy, roughly half of assessed terminal evaluations with outcome ratings in the satisfactory range attributed project achievements to high levels of support received from nonstate stakeholders and country actors. Evidence cited included high levels of cofinancing provided; or the emergence of project actors from the private sector or nongovernmental organizations, or within national agencies/ministries who actively drove the project forward.

For projects with overall outcome ratings below the satisfactory range, the two most frequently cited factors in assessed terminal evaluations are weaknesses in project design and management. Seventy-five percent of the evaluations (61 of 81) attributed low project performance to design shortcomings, which included significant problems in the project’s logical framework, failure to tailor the project adequately to the local context, failure to adequately budget project activities, overly ambitious project goals; and poor choices in executing arrangements. Similarly, weak management—evidence of which included poor supervision by the GEF Agency, poor coordination with project partners, financial mismanagement, major issues with procurement, and high staff turnover—was identified as a factor that limited project performance in 65 percent (53 of 81) of assessed terminal evaluations with outcome ratings below the satisfactory range.

Among the factors associated with lower performing projects, the percentages and categories are largely consistent between the APR 2012 and APR 2008 studies. In the APR 2012 study, a fourth

FIGURE 4.4 Results of Analysis of Factors Attributed to High and Low Project Performance



NOTE: Sample includes 223 terminal evaluations of projects in the OPS5 cohort with overall outcome ratings of moderately satisfactory or above and that identified factors that directly or indirectly contributed to project outcome achievements (a). Sample also includes 81 terminal evaluations of projects in the OPS4 and OPS5 cohorts with overall outcome ratings below moderately satisfactory and that identified factors that directly or indirectly hindered project outcome achievements (b). Factor categories are non-exclusive (individual project evaluations can cite more than one factor).

category emerged—low country support—including projects for which weak support/commitment from the country (or some levels or sectors of country administration) is reported in the terminal evaluation as hindering the project's outcome achievements. Evidence cited includes excessive delays regarding permitting of project activities, failure to advance legislation or policy critical to the success of the project, and development plans that conflict with the project. Thirty-six percent (29 of 81) of assessed terminal evaluations with outcome ratings below the satisfactory range attributed a portion of the project's limited success to this factor.

Some evidence is found in assessed terminal evaluations that strong project management can sometimes overcome weaknesses in project design. Thirty-one, or 19 percent, of the 223 assessed projects with overall outcome ratings of moderately satisfactory or above had important weaknesses in

design, according to the terminal evaluations, but succeeded in large part in meeting project expectations due to strong project management. Further analysis is needed to understand under what conditions strong project management can or cannot overcome weaknesses in design, and how this is accomplished.

A few examples from the study help to illustrate the identified factors more clearly:

- **Strong management.** The UNDP-implemented Conservation of Globally Significant Biodiversity in the Landscape of Bulgaria's Rhodope Mountains project (GEF ID 1042) achieved most of its intended outcomes despite starting with a design that, according to the project's terminal evaluation, was "too complex," with "too many activities" (110 in all) and that did not consider the failure to establish nature parks—a key component of the project—as a possibility. Evidence

of strong management included adaptive management following a critical midterm evaluation, efficient coordination of subcontracts, effective project monitoring, and strong trust built between the management team and local stakeholders through continuous consultation.

- **Poor design.** The World Bank–implemented Vilnius Heat Demand Management project (GEF ID 948), which sought to reduce greenhouse gas emissions from the residential building sector of the city through a demand-side management program, suffered from several design issues identified in the terminal evaluation. These included design assumptions that two of the project’s executing agencies would closely coordinate their efforts—an assumption that proved to be false; splitting of the GEF grant into two subgrants, which prohibited reallocation of GEF funds between project components during project execution; and insufficient consultation with homeowner associations regarding demand for the project’s outputs.
- **Strong nonstate stakeholder support.** During execution of the UNDP–implemented Biodiversity Conservation in the Sierra Gorda Biosphere Reserve in Mexico project (GEF ID 887), project managers sought the participation and involvement of various stakeholders, many of whom collaborated with the project on a volunteer basis. Because of these partnerships, which included domestic private sector organizations as well as international donor institutions, the project was able to triple the amount of projected cofinancing realized, as well as obtain pro bono advice from experts. These facilitated strong results and enhanced project efficiency.
- **Poor project management.** The World Bank’s Rural Environment Project (GEF ID 1535), which sought to improve biodiversity conservation and introduce sustainable natural resource

management in two mountainous areas of Azerbaijan, was understaffed in the early years of the project’s execution; in particular, it lacked a qualified procurement specialist. Additionally, it experienced severe delays in the production of key project outputs; and high staff turnover in the project management team, which disrupted communication between the Bank and the local ministry of environment. As a result, investments in park infrastructure and equipment called for in the project design were not made, and no national park or protected area staff benefited from the training programs implemented by the project.

- **Low country support.** Insufficient country ownership and support limited the achievements of the World Bank–implemented Biodiversity Conservation in the Azov-Black Sea Ecological Corridor project in Ukraine (GEF ID 412). The project, which sought to conserve biodiversity within the Azov-Black Sea coastal corridor by strengthening the protected area network and mainstreaming biodiversity conservation into the surrounding agricultural areas, faced numerous obstacles. These included a two-year delay in the provision of the national cofinancing agreed upon at appraisal, inaction on the part of the national executing agency, lack of interagency coordination, and repeated changes in project management (it had five project directors in two years). The project was ultimately canceled after 16 percent of GEF funding was spent, and few desired outputs and objectives were achieved.

## 4.4 Trends in Project Extensions

Project extensions—defined as the time taken to complete project activities beyond that anticipated in project approval documents—can be incurred for reasons both within and outside

of management's control,<sup>3</sup> and are not a strong predictor of project outcomes within the APR 2005–12 cohort. That is, no statistically significant difference is found in the proportion of projects with outcome ratings of moderately satisfactory or above between projects that did or did not have project extensions. The same holds true when projects are sorted on the basis of those having extensions of more than one year, or even of two years.<sup>4</sup> Moreover, project extensions may allow for the realization of intended project outputs, and may be a consequence of good adaptive management.

At the same time, project extensions likely mean that the intended return on GEF funding—project outputs and environmental outcomes—have not materialized within the time frame anticipated in project approval documents. When a trend in project extensions appears over time, it may signal that project time frames or strategies are unrealistic given the conditions in which projects take place. Project extensions are therefore one aspect of project performance that is tracked in the APR.

Table 4.4 presents summary statistics on project extensions for projects in four-year APR cohorts, and the eight-year APR 2005–12 cohort, where data are available. Overall, 80 percent of assessed projects in the APR 2005–12 cohort have project extensions, and the percentages within the four-year APR cohorts differ by only 3 percent. A small difference is also seen between FSPs and MSPs, with 81 percent of the FSPs in the APR 2005–12 cohort having project extensions versus 78 percent of the MSPs.

---

<sup>3</sup>This definition excludes any delays that may occur prior to the start of project activities.

<sup>4</sup>A very small (0.2 point) difference is found in the mean outcome rating between projects with and without extensions of more than one and two years when using a 6-point rating scale for outcomes.

More distinctions are found when assessing project extensions by GEF Agency.<sup>5</sup> The percentage of UNDP projects in the APR 2005–12 cohort with project extensions is 87 percent, versus 79 percent for UNEP, and 71 percent for World Bank projects.<sup>6</sup> The percentage of UNDP and UNEP projects with project extensions has declined between the four-year APR cohorts: from 93 percent to 83 percent for UNDP, and from 82 percent to 77 percent for UNEP. For World Bank projects, the percentage of projects with extensions are essentially unchanged between the four-year APR cohorts.

Because GEF Agencies differ with respect to the proportion of FSPs and MSPs in their respective portfolios, comparisons between Agency trends in project extensions need to be separately assessed for FSPs and MSPs. As table 4.4 indicates, even when accounting for these differences, the trends in project extensions among UNDP, UNEP, and World Bank projects is largely consistent with the numbers for Agencies' overall portfolios. That is, World Bank projects typically experience fewer and shorter project extensions than UNDP and UNEP projects.

Using two thresholds—the percentage of projects with extensions greater than one year, and the percentage of projects with extensions of greater than two years—illustrates the same point more clearly. As shown in figure 4.5, more than half of all full-size UNDP and UNEP projects have project extensions beyond one year, versus 43 percent for World Bank projects. For MSPs, the numbers are 35 percent for both UNDP and UNEP, and 28 percent for the World Bank. Similarly, 38 percent of

---

<sup>5</sup>There is currently insufficient information on project extensions to report on GEF Agencies other than UNDP, UNEP, and the World Bank.

<sup>6</sup>The difference in the proportion of UNDP and World Bank projects with project extensions is statistically significant at a 95 percent confidence level. Differences in the proportion of projects with project extensions between other GEF Agencies is not statistically significant.

TABLE 4.4 Project Extensions by Project Size, GEF Agency, and APR Cohort Grouping

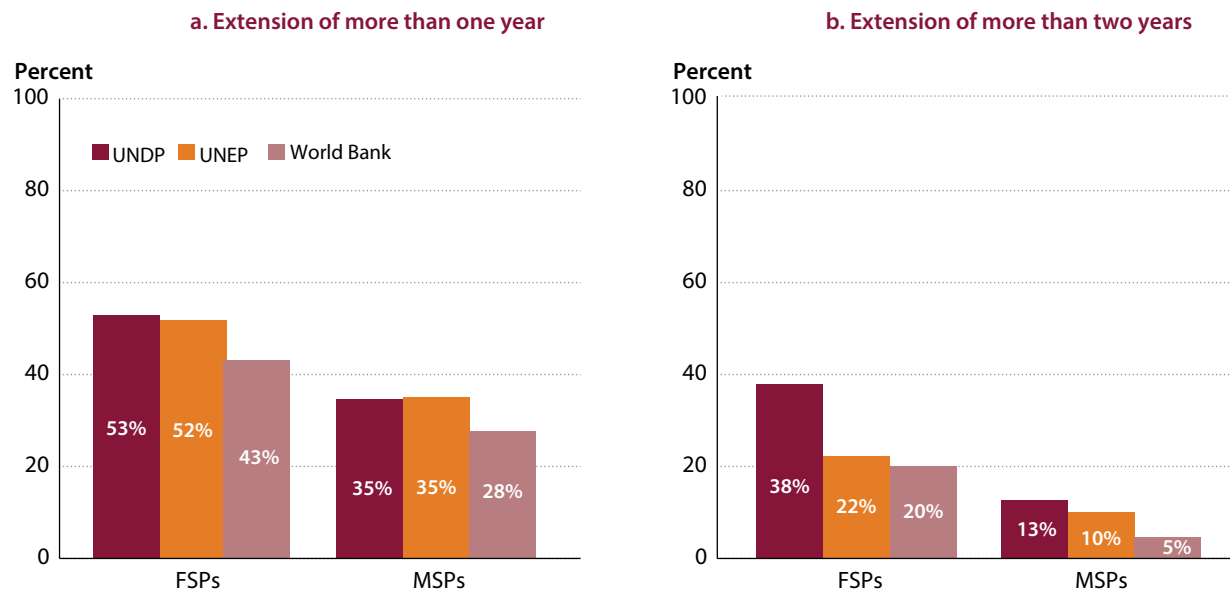
Criterion			APR 2005–08	APR 2009–12	APR 2005–12
Number of projects with data on project extensions			198	268	466
Percentage of projects with project extensions	All projects		81	78	80
	FSPs		81	80	81
	MSPs		81	76	78
	UNDP		93	83	87
	UNEP		82	77	79
	World Bank		71	72	71
Median length of project extension (months) <sup>a</sup>	All projects		14	14	14
	FSPs	All FSPs	23	18	18
		UNDP	26	17	20
		UNEP	22	20	21
		World Bank	23	18	18
	MSPs	All MSPs	11.5	12	12
		UNDP	14	12	12
		UNEP	6	12	9.5
		World Bank	10	13	12
	Percentage of projects with extensions of > 1 year	All projects		42	40
FSPs		All FSPs	50	46	48
		UNDP	63	48	53
		UNEP	45	56	52
		World Bank	43	43	43
MSPs		All MSPs	34	31	32
		UNDP	47	28	35
		UNEP	35	35	35
		World Bank	21	38	28
Percentage of projects with extensions of > 2 years		All projects		23	18
	FSPs	All FSPs	35	23	28
		UNDP	51	31	38
		UNEP	18	25	22
		World Bank	26	12	20
	MSPs	All MSPs	9	10	9
		UNDP	12	13	13
		UNEP	6	13	10
		World Bank	8	0	5

NOTE: FSPs include two enabling activities based on size of the GEF grant.

a. Includes only those projects with project completion extensions.



**FIGURE 4.5 Summary Statistics on One- and Two-Year Project Extensions, by GEF Agency and Project Size, within the APR 2005–12 Cohort**



full-size UNDP projects have project extensions of greater than two years, compared with 22 percent and 20 percent for UNEP and the World Bank,

respectively. For MSPs, the percentages are 13 percent, 10 percent, and 5 percent for UNDP, UNEP, and the World Bank, respectively.



## 5. Quality of M&E Design and Implementation

**P**roject M&E systems provide real-time information to managers on the progress made in achieving intended results and facilitate adaptive management. Effective M&E systems also allow for the evaluation of project impacts and sustainability following project closure. They are therefore among the key project performance indicators tracked and reported on by the GEF Evaluation Office in the APR.

### 5.1 Rating Scale

As discussed in the [methodology section of chapter 2](#), M&E systems are assessed in terminal evaluations on two principle dimensions: (1) the design of a project's M&E system, and (2) the implementation of a project's M&E system. A six-point rating scale is used to assess overall M&E design and M&E implementation, with the following categories:

- **Highly satisfactory.** The project had no shortcomings in M&E design/implementation.
- **Satisfactory.** The project had minor shortcomings in M&E design/implementation.
- **Moderately satisfactory.** The project had moderate shortcomings in M&E design/implementation.
- **Moderately unsatisfactory.** The project had significant shortcomings in M&E design/implementation.
- **Unsatisfactory.** The project had major shortcomings in M&E design/implementation.

Among projects that have been rated on both M&E design and implementation by the GEF Evaluation Office or GEF Agency evaluation offices, strong associations are found between the two ratings. That is, projects with M&E design ratings of moderately satisfactory or above are more likely than not to have M&E implementation ratings of moderately satisfactory or above as well, and vice versa (see [annex C](#) for the full methodology and results of this analysis). At the same time, project M&E systems can be, and often are, modified and improved upon during project implementation.

### 5.2 Findings

Table 5.1 shows the percentage of rated projects with quality of M&E design ratings of moderately satisfactory or above. Only 66 percent of rated projects ( $n = 421$ ) have M&E design ratings of moderately satisfactory or above. Also noteworthy, M&E design ratings between four-year APR cohorts are essentially flat:<sup>1</sup> 67 percent of projects within the APR 2005–08 cohort and 65 percent of projects within the APR 2009–12 cohort have M&E design ratings of moderately satisfactory or

<sup>1</sup> Ratings for M&E design are not available in APR year cohorts prior to FY 2006, so here the four-year APR 2005–08 cohort includes ratings from only three years.

TABLE 5.1 Quality of M&amp;E Design, by Project Size and Year

Project size	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
Percentage of projects with M&E design rated moderately satisfactory or above								
All projects	59	68	72	72	70	65	57	66
FSPs	44	50	77	57	67	67	64	63
MSPs	75	85	67	88	72	62	47	69
Number of rated projects								
All projects	49	40	61	54	46	94	77	421
FSPs	25	20	31	28	21	55	47	227
MSPs	24	20	30	26	25	39	30	194

NOTE: FSPs include two enabling activities.

above. In short, only two-thirds of rated GEF projects are meeting the commonly used threshold for satisfactory M&E design, and the percentages have remained fairly stable for the past seven APR years.

Some differentiation is found between MSPs and FSPs, with a higher percentage of MSPs at or above the moderately satisfactory threshold compared with FSPs. The difference is not statistically significant.

Ratings on the quality of M&E implementation are presented in table 5.2 and figure 5.1. The proportion of projects with M&E implementation ratings of moderately satisfactory or above largely tracks, and is similar to, ratings on M&E design. Of the 390 projects for which ratings are available, only 68 percent of projects have M&E

implementation ratings of moderately satisfactory or above. Between four-year APR cohorts, the percentage of projects with M&E implementation ratings of moderately satisfactory or above has declined slightly, from 71 percent in the APR 2005–08 cohort to 66 percent in the APR 2009–12 cohort.<sup>2</sup> The decline in ratings between OPS cohorts is not statistically significant, however.

As with ratings on M&E design, ratings on M&E implementation can be distinguished by project size. Among rated projects, a higher proportion

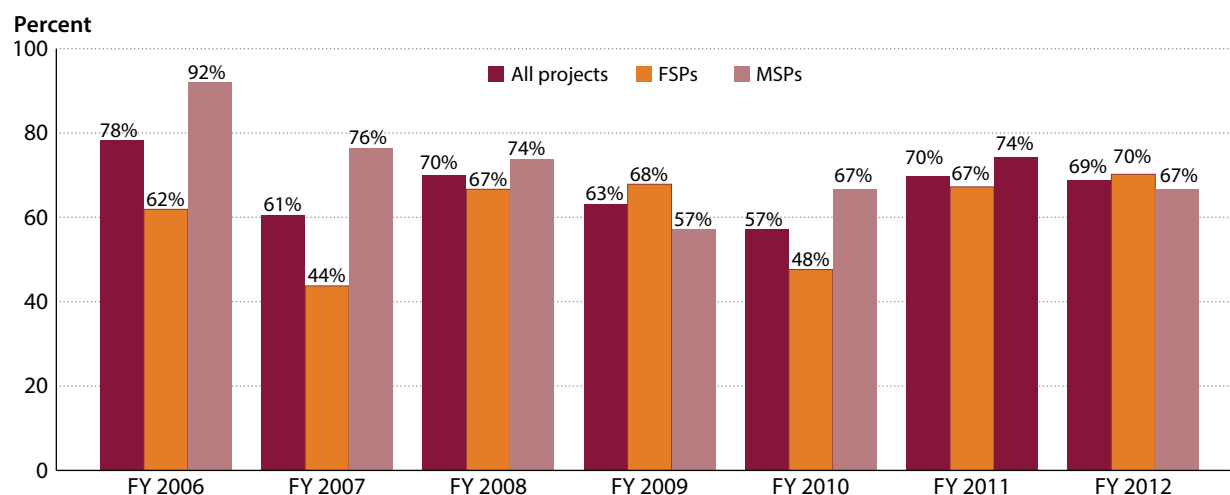
<sup>2</sup> Ratings for M&E implementation are not available in APR year cohorts prior to FY 2006, so here the four-year APR 2005–08 cohort includes ratings from only three years.

TABLE 5.2 Quality of M&amp;E Implementation, by Project Size and Year

Project size	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	All cohorts
Percentage of projects with M&E implementation rated moderately satisfactory or above								
All projects	78	61	70	63	57	70	69	68
FSPs	62	44	67	68	48	67	70	64
MSPs	92	76	74	57	67	74	67	73
Number of rated projects								
All projects	46	33	50	49	42	93	77	390
FSPs	21	16	27	28	21	58	47	218
MSPs	25	17	23	21	21	35	30	172

NOTE: FSPs include two enabling activities. The difference in the overall share of FSPs and non-FSPs with quality of M&E implementation ratings of moderately satisfactory or above is significant to a 95 percent confidence level.

**FIGURE 5.1 Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by Project Size and APR Year**

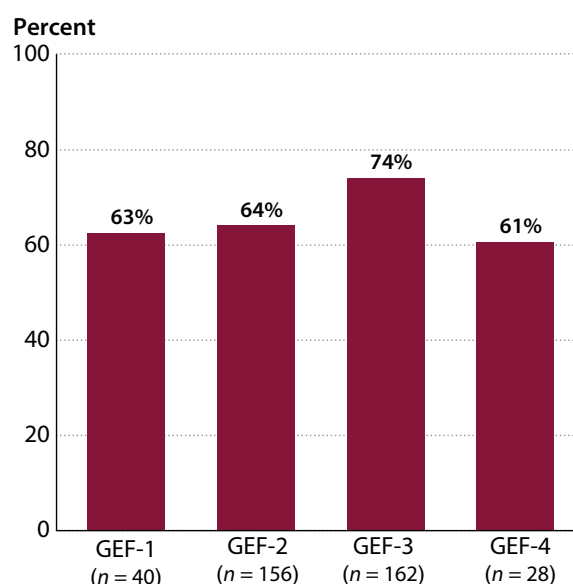


of MSPs have M&E implementation ratings of moderately satisfactory or above compared to FSPs: 73 percent versus 64 percent, respectively. Whether this is due to the increased complexity or more stringent M&E requirements for FSPs, or to some other factors, is not known. The difference is statistically significant at a 95 percent confidence level.

Figure 5.2 shows M&E implementation ratings by GEF replenishment phase. Because GEF phase cohorts are not complete and a very limited number of ratings are available for GEF-1 and GEF-4, care must be taken in assessing any trends in M&E implementation ratings by GEF phase at this time. That said, among rated projects, a greater proportion (74 percent) of projects authorized during the GEF-3 replenishment period have M&E implementation ratings of moderately satisfactory or above compared to projects authorized during GEF-2 (64 percent). The difference is statistically significant at a 90 percent confidence level.

Between the OPS4 and OPS5 cohorts, significant shifts in the M&E implementation ratings of two GEF Agencies are found. As shown in figure 5.3, the percentage of UNDP projects with M&E implementation ratings of moderately

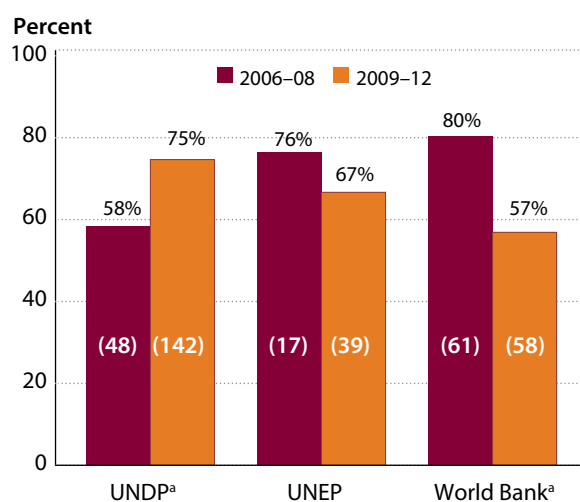
**FIGURE 5.2 Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by GEF Phase**



NOTE: GEF phase cohorts are not complete, and a very limited number of ratings are available for GEF-1 and GEF-4.

satisfactory or above has risen from 58 percent of projects in the OPS4 cohort (again, ratings are not available for FY 2005), to 75 percent of projects in the OPS5 cohort. The difference is statistically

**FIGURE 5.3 Percentage of Projects with M&E Implementation Ratings of Moderately Satisfactory or Above, by GEF Agency**



a. The difference in the percentage of projects with quality of M&E implementation ratings of moderately satisfactory or above between APR year groupings is significant to a 95 percent confidence level.

significant at a 95 percent confidence level. In contrast, M&E implementation ratings between OPS cohorts have declined for both UNEP and World Bank projects. For World Bank projects, the decline from 80 percent to 57 percent of projects with M&E implementation ratings of moderately satisfactory or above is statistically significant at a 95 percent confidence level. The decline in M&E implementation ratings for UNEP projects is not statistically significant.

## 6. Quality of Terminal Evaluation Reports

**T**erminal evaluation reports provide one of the principle ways by which the GEF Council, management, GEF Agencies, GEF Evaluation Office, and other stakeholders are able to assess the performance of GEF projects. This assessment facilitates continued learning and adaptation throughout the GEF partnership. The integrity and quality of terminal evaluations is therefore essential to the validity of any findings that may arise from analysis of terminal evaluations.

The GEF Evaluation Office has been reporting on the quality of terminal evaluations since APR 2004. To date, 566 terminal evaluations have been submitted to the Office. Of these, 527 have been rated by either the GEF Evaluation Office or GEF Agency evaluation offices. Year of terminal evaluation completion is used for analysis rather than APR year, as year of terminal evaluation does a better job of capturing when the actual work of reporting took place.

As noted in chapter 2 and described in full in [annex B](#), terminal evaluations are assessed and rated by the GEF Evaluation Office and GEF Agency evaluation offices 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?

- 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?

Performance on each of these criteria is rated on a six-point scale, from highly satisfactory to highly unsatisfactory. The overall rating for the terminal evaluation is a weighted average of the six subratings, with the first two subratings receiving more weight than the other four (see [annex B](#)).

### 6.1 Findings

Table 6.1 and figure 6.1 present overall ratings on terminal evaluation reports by project size, GEF Agency, and year of terminal evaluation completion. While a fair amount of annual variability in the ratings is apparent, in most years, the percentage of terminal evaluations with ratings of moderately satisfactory or above exceeds 80 percent. Overall, 86 percent of rated terminal evaluations have ratings of moderately satisfactory or above.<sup>1</sup>

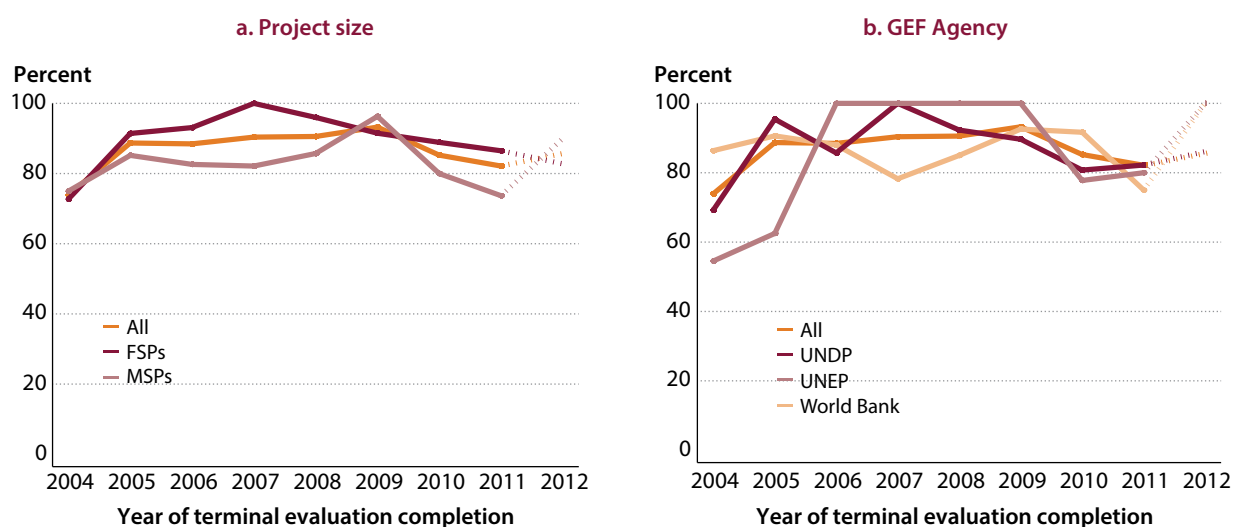
<sup>1</sup> Note that the 2011 and 2012 cohorts are not yet complete.

**TABLE 6.1 Percentage of Terminal Evaluation Reports Rated Moderately Satisfactory or Above, or Satisfactory or Above, by Project Size, GEF Agency, and Year of Report Completion**

Project size and Agency	2004 and earlier	2005	2006	2007	2008	2009	2010	2011	2012	All years
<b>Percentage of reports rated moderately satisfactory or above</b>										
All projects	72	89	87	90	91	93	85	82	86	86
FSPs	71	91	93	100	96	91	89	86	83	89
MSPs	72	85	83	82	86	96	80	74	90	83
UNDP	75	95	86	100	92	90	81	82	86	86
UNEP	50	63	100	100	100	100	78	80	100	84
World Bank	83	91	88	78	85	93	92	75	100	87
<b>Percentage of reports rated satisfactory or above</b>										
All projects	43	53	40	60	55	73	61	38	49	53
FSPs	43	60	45	67	68	72	75	46	48	59
MSPs	44	44	35	54	43	74	40	21	50	46
UNDP	25	55	33	50	54	59	54	31	42	44
UNEP	40	25	33	71	57	88	67	60	86	63
World Bank	50	59	48	65	55	81	67	75	100	61
Number	67	62	53	52	53	74	61	56	49	527

NOTE: The difference in the share of terminal evaluations with overall ratings of moderately satisfactory or above between MSPs and FSPs is significant to a 95 percent confidence level. The difference in the share of terminal evaluations with overall ratings of satisfactory or above between MSPs and FSPs is significant to a 95 percent confidence level. The difference in the share of terminal evaluations with overall ratings of satisfactory or above between UNDP and non-UNDP evaluations is significant to a 95 percent confidence level.

**FIGURE 6.1 Percentage of Terminal Evaluation Reports with Overall Quality Rated Moderately Satisfactory or Above by Project Size and GEF Agency**



NOTE: Cohorts for 2011 and 2012 are not yet complete. Dotted lines indicate that trend lines are provisional and may change as additional ratings of terminal evaluations become available in subsequent APRs.

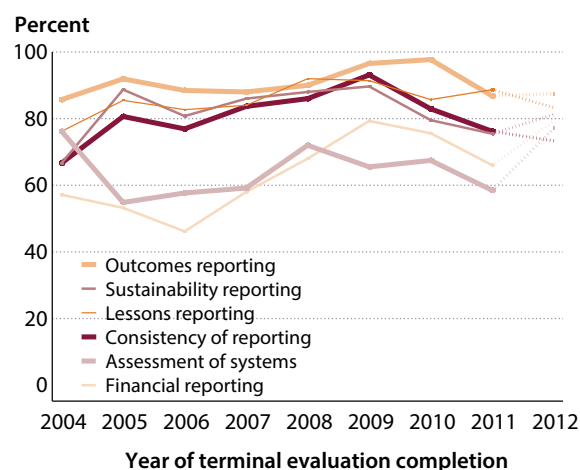
The quality of terminal evaluations of MSPs has typically lagged that of FSPs, with 83 percent of assessed terminal evaluations for MSPs rated moderately satisfactory or above compared with 89 percent of FSPs. This difference is statistically significant to a 95 percent confidence level, and becomes more pronounced when a more stringent yardstick of satisfactory and above is used (as shown in the lower half of table 6.1). Only 46 percent of rated MSP evaluations compared with 59 percent of rated FSP evaluations meet the threshold of satisfactory and above.

Little distinction is seen in overall reporting quality among GEF Agencies when using the moderately satisfactory or above threshold. Differences in the overall quality of terminal evaluations among GEF Agencies become more visible when using the satisfactory and above threshold. The percentage of assessed UNDP terminal evaluations with overall ratings of satisfactory or above is 44 percent, compared with 63 percent for UNEP evaluations, and 61 percent for World Bank evaluations. This difference is statistically significant at a 95 percent confidence level.

As noted above, overall ratings on terminal evaluations are based on an assessment of the quality in terminal evaluation reporting along six criteria. Figure 6.2 shows how reporting on these six criteria has fared over the 2004–12 cohort in terms of ratings. In general, reporting on most dimensions has been strong, with more than 80 percent of terminal evaluations rated as moderately satisfactory or above for reporting on outcomes, consistency, sustainability, and lessons and recommendations. Reporting on project financing and M&E systems has not been as strong, with only 67 percent and 66 percent, respectively, of rated terminal evaluations within the 2004–12 cohort receiving ratings of moderately satisfactory or above.

The performance of terminal evaluations along these two dimensions has improved within the FY 2012 cohort. However, as the dotted lines in

**FIGURE 6.2 Quality of Terminal Evaluation Reporting on Individual Dimensions**



NOTE: Dotted lines indicate that the FY 2012 cohort is not yet complete.

figure 6.2 indicate, this cohort is not yet complete, and ratings may change as more terminal evaluations from this year become available in subsequent APRs.

## 6.2 Comparison of Ratings from GEF Evaluation Office and GEF Agency Evaluation Offices

As discussed in chapter 2, a number of GEF Agencies have independent evaluation offices that provide oversight and review ratings provided in their Agency’s respective terminal evaluations. Beginning in 2009, the GEF Evaluation Office began accepting ratings from the independent evaluation offices of the World Bank, UNEP, and—subsequently—UNDP. This approach, which reduces duplicative work, follows the GEF Evaluation Office finding that ratings from these three evaluation offices are largely consistent with those provided by the GEF Evaluation Office (GEF EO 2009b).

The GEF Evaluation Office continues to track consistency between its ratings and those provided by Agencies’ independent evaluation offices. To do so, the Office reviews a random sample of

terminal evaluations that have also been reviewed by Agency evaluation offices.

Table 6.2 shows how the ratings on overall outcomes compare for all projects where two sets of ratings are available (127 projects). Overall, ratings provided by the GEF Agencies continue to be largely consistent with those provided by the GEF Evaluation Office. Among the sampled reviews, a small (4 percent) difference in the percentage of

projects with overall outcome ratings of moderately satisfactory or above is found between ratings from Agency evaluation offices and those from the GEF Evaluation Office. This difference is not statistically significant. Moreover, adjusting for a possible bias would not lead to significant changes in the findings presented in APRs from 2009 onward. The GEF Evaluation Office will continue to track the consistency of ratings going forward.

**TABLE 6.2 Comparison of Overall Outcome Ratings from GEF Agency Independent Evaluation Offices and from the GEF Evaluation Office for All Jointly Rated Projects, APR 2005–12**

GEF Agency	Number of projects with ratings from both Agency and GEF Evaluation Office	% of projects rated moderately satisfactory or above by Agency	% of projects rated moderately satisfactory or above by GEF Evaluation Office	Difference in ratings between Agency and GEF Evaluation Office (%)
ADB	1	100	100	0
UNDP	24	88	83	5
UNEP	37	96	89	7
UNIDO	3	67	67	0
World Bank	62	89	84	5
Total	127	89	85	4

NOTE: ADB = Asian Development Bank.



## 7. Management Action Record

**T**he GEF management action record tracks the level of adoption, by the GEF Secretariat and/or the GEF Agencies (together here referred to as GEF management), of GEF Council decisions that have been made on the basis of GEF Evaluation Office recommendations. The MAR serves two purposes:

(1) to provide Council with a record of its decision on the follow-up of evaluation reports, the proposed management actions, and the actual status of these actions; and (2) to increase the accountability of GEF management regarding Council decisions on monitoring and evaluation issues (GEF EO 2005).

The format and procedures for the MAR were approved by the GEF Council at its November 2005 meeting. They call for the MAR to be updated and presented to the Council for review and follow-up on an annual basis.

MAR 2012 tracks 21 separate GEF Council decisions: 10 that were part of MAR 2011 and 11 new decisions. In addition, this year the Evaluation Office has also started tracking adoption of the decisions of the LDCF/SCCF Council. One decision from the LDCF/SCCF Council's November 2011 meeting is tracked in MAR 2012.

### 7.1 Rating Approach

For each tracked GEF Council and LDCF/SCCF Council decision, self-ratings are provided by GEF

management on the level of adoption, along with commentary as necessary. Ratings and commentary on tracked decisions are also provided by the GEF Evaluation Office for verification. The rating categories for the progress of adoption of Council decisions were agreed upon through a consultative process of the Evaluation Office, the GEF Secretariat, and the GEF Agencies. Adoption categories are as follows:

- **High**—fully adopted and fully incorporated into policy, strategy, or operations
- **Substantial**—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
- **Not possible to verify yet**—verification will have to wait until more data are available or proposals have been further developed
- **N.A.**—not applicable or no rating provided (see commentary)

MAR 2012 tracks management actions on GEF Council and LDCF/SCCF Council decisions based on 12 GEF Evaluation Office documents. Seven of these evaluations were included in MAR 2011:

- Annual Performance Report 2006 (GEF/ME/C.31/1, May 2007)
- Joint Evaluation of the Small Grants Program—Executive Version (GEF/ME/C.32/2, October 2007)
- Annual Country Portfolio Evaluation Report 2009 (GEF/ME/C.35/1, June 2009)
- Annual Report on Impact 2009 (GEF/ME/C.36/2, November 2009)
- Annual Performance Report 2009 (GEF/ME/C.38/4, June 2010)
- Evaluation of the GEF Strategic Priority for Adaptation (GEF/ME/C.39/4, October 2010)
- Annual Thematic Evaluations Report 2011 (GEF/ME/C.41/02, October 2011)

Five additional evaluations are the source of 12 new tracked Council decisions:

- Evaluation of the Special Climate Change Fund (GEF/LDCF.SCCF.11/ME/02, October 2011)
- Annual Performance Report 2011 (GEF/ME/C.42/01, May 2012)
- Annual Country Portfolio Evaluation Report 2012 (GEF/ME/C.42/03, May 2012)
- Annual Thematic Evaluations Report 2012 (GEF/ME/C.43/02, October 2012)
- GEF Annual Impact Report 2012 (GEF/ME/C.43/04, October 2012)

## 7.2 Findings

Of the 21 GEF Council decisions tracked in MAR 2012, the Evaluation Office was able to verify management's actions on 14. None of the tracked decisions will be graduated this year, either because there has been insufficient time for management to act on Council decisions, or the Evaluation Office

has been unable to verify that a high level of adoption of the relevant Council decision has occurred. All 21 decisions are considered by the Evaluation Office to still be relevant and will be tracked in next year's MAR.

Five of the 10 decisions that were tracked in previous MARs and in MAR 2012 have been rated by the Evaluation Office as having a substantial level of adoption (table 7.1). In two cases, management is finalizing new policy guidelines based upon Council recommendations; in another two, minor issues are still being addressed; and for the fifth case, there are too few observations to justify a high adoption rating at this time. For the other five previously tracked decisions, adoption has been slow; and in one case, management has not acted upon Council's request (see [below](#)). For the majority of newly tracked decisions, it is not yet possible to verify the level of adoption by management.

### GEF COUNCIL DECISIONS WITH ADOPTION RATED AT A HIGH OR SUBSTANTIAL LEVEL

An example of progress made in adopting Council recommendations includes the GEF Council decision based on the Evaluation of the Strategic Priority for Adaptation. The Council's request to the Secretariat that screening tools to identify and reduce climate risks to the GEF portfolio be developed has been acted on through development of the Climate Risk Screening Tool and Adaptation Monitoring and Assessment Tool. Further work to integrate climate resilience considerations across all focal areas and improve GEF-6 (2014–18) focal area strategies in this regard is ongoing.

Adoption of 10 of the tracked MAR 2012 decisions was rated by management as substantial or high. For one of these—a decision by the Council, based on review of the 2009 GEF Annual Impact Report, that the Secretariat should incorporate lessons from the GEF's positive experience working

TABLE 7.1 GEF Management and GEF Evaluation Office Ratings of Council Decisions Tracked in MAR 2012

Management rating	GEF Evaluation Office rating						Sum of management ratings
	High	Substantial	Medium	Negligible	Not possible to verify yet	Not applicable/not rated	
High	0	1	0	0	1	0	2
Substantial	0	4	2	1	0	0	7
Medium	0	0	4	0	5	0	9
Negligible	0	0	0	0	0	0	0
Not possible to verify yet	0	0	0	0	0	0	0
Not applicable/not rated	0	0	0	2	1	0	3
<b>Sum of Office ratings</b>	<b>0</b>	<b>5</b>	<b>6</b>	<b>3</b>	<b>7</b>	<b>0</b>	<b>21</b>

NOTE: Highlighted fields show agreement between the ratings of management and the GEF Evaluation Office; fields to the right of the diagonal represent higher ratings by management than by the Evaluation Office (except in the case of not possible to verify yet).

with the private sector in the ozone layer depletion focal area into other focal areas, where appropriate—the Evaluation Office is presently undertaking a review of GEF engagement with the private sector and has withheld rating the adoption of this decision until the findings of this review are complete.

Three decisions rated by management as having substantial adoption were rated lower by the GEF Evaluation Office. Differences between management and GEF Evaluation Office ratings for MAR 2012 pertain to four decisions; these are discussed [below](#).

## DECISIONS WITH NO CHANGE IN RATING

The GEF Evaluation Office ratings for 8 of the 10 MAR 2012 decisions that were also included in MAR 2011 remained unchanged. For five of these decisions, lack of movement from the MAR 2011 ratings is not reflective of a lack of progress being made to address Council recommendations. For example, the Council decision based on the Joint Evaluation of the Small Grants Programme—that country program oversight needs to be strengthened—has seen continued responsive action taken by management. Efforts include regular

coordination and consultation meetings with the Central Program Management Team, plans by UNDP for risk-based audits in 2013, and work on improving and streamlining the Small Grants Programme’s monitoring system as part of the design of GEF-6.

Another example where progress has been made despite no change in the ratings is in the Council decision based on the Evaluation of the Strategic Priority for Adaptation. The Council’s request that the Evaluation Office, the GEF Scientific and Technical Advisory Panel, and the Adaptation Task Force provide guidelines for Strategic Priority for Adaptation projects to learn from the outcomes and impacts of these projects has been acted on, with revised guidelines for terminal evaluations applying to such projects nearly finalized.

Adoption of three Council decisions tracked in MAR 2011 has been slow; in one case, it is not clear that actions taken by management are adequately addressing the Council’s concerns. In the latter case, the Council decided in June 2007, based on review of the 2006 APR, that special attention is required to ensure continued and improved supervision by GEF Agencies during project implementation, and that adequate funding should be provided for this supervision from project fees. While a new fee structure was developed and approved by

the Council in June 2012, project fees for MSPs and FSPs were reduced from their previous level.<sup>1</sup> The GEF Secretariat and GEF Agencies have worked on measures to streamline the project cycle, some of which were approved by the GEF Council in November 2012. However, there is little information on how these activities have resulted in greater resources being made available for project supervision, especially considering that overall project fees have declined.

The Council's decision based on the 2009 APR that management and the GEF Evaluation Office should work together to improve the quality of information available through the GEF PMIS on the status of projects has been acted upon to some degree. However, a recent review of PMIS data undertaken by the Evaluation Office shows that concerns related to the poor quality of information on project status still remain. In particular, while new features have been added to the PMIS, relatively little attention from the Secretariat has focused on the quality of information provided.

Lastly, the GEF Secretariat has not acted on a June 2009 Council decision requesting the Secretariat conduct a survey of countries in exceptional situations concerning limited access to GEF partner international financial institutions.

## COMPARISON OF EVALUATION OFFICE AND MANAGEMENT RATINGS

Management and the Evaluation Office are in agreement on the level of adoption for only 8 of the 22 tracked decisions in MAR 2012; for 7 tracked decisions, the Evaluation Office was unable to

verify ratings either because insufficient information was available at the time, or the proposals needed more time to be developed. Excluding these seven decisions for which the Office was unable to verify ratings, the level of agreement between management and the Evaluation Office is 57 percent—in line with the level for MAR 2011 (58 percent) and MAR 2010 (66 percent). In all cases where ratings provided by management and the Evaluation Office do not match, those from the Evaluation Office are lower than those provided by management—and in one case, substantially lower.

The largest gap between ratings provided by management and the GEF Evaluation Office is for assessing the level of adoption of the Council's request, based on the 2012 Annual Country Portfolio Evaluation Report, that the Secretariat reduce the burden of monitoring requirements for multifocal area projects to a level comparable to that of single focal area projects. While the GEF Secretariat rates adoption of this decision as substantial, the Evaluation Office has assessed the actions taken thus far in response as negligible. The Office finds "no evidence that tracking tools burdens for MFAs [multifocal areas] have been reduced." This finding is supported by UNDP and UNEP commentary included in the MAR management response as separate responses from these Agencies.

## GRADUATED DECISIONS

Since the commencement of the MAR in June 2006, the Evaluation Office has tracked the adoption of 111 Council decisions based on the recommendations of 32 evaluations. Overall, the GEF has been highly responsive to Council decisions, allowing for an ongoing reform process. Evidence of this reform process is seen in the high or substantial level of adoption reached on 65 of the decisions at the time of their graduation. The Evaluation Office graduates decisions for which a high level of adoption rating has been achieved or those that are considered no longer relevant. To date, 86 (77 percent)

<sup>1</sup> Project fees for projects up to \$10 million in GEF funding were reduced from 10 percent to 9.5 percent of GEF funding, while project fees for grants above \$10 million were reduced from 10 percent to 9 percent. No changes were made to the fee structure for Programmatic Approach grants or grants awarded under the Small Grants Programme.

of tracked decisions have been graduated. Table 7.2 provides a summary of Council decisions graduated from the MAR.

## DECISIONS OF THE LDCF/SCCF COUNCIL

As discussed above, this year the Evaluation Office has started tracking decisions of the LDCF/SCCF Council in the MAR. MAR 2012 tracks the level of adoption of a single decision with three subcomponents from the LDCF/SCCF Council's November 2011 meeting, based on the Evaluation of the Special Climate Change Fund. Both the Evaluation Office and the Secretariat are in agreement that, overall, a substantial level of adoption of the Council's recommendations has occurred, particularly with respect to the LDCF/SCCF Council's request that the Secretariat prepare proposals to ensure "transparency of the project pre-selection process and dissemination of good practices through existing channels." The Secretariat developed

a document detailing the preselection process and criteria for SCCF-funded projects which was circulated during the 12th LDCF/SCCF Council meeting. These guidelines were included in the "Updated Operational Guidelines for the Special Climate Change Fund for Adaptation and Technology Transfer," approved by the LDCF/SCCF Council in November 2012 (GEF 2012). Regarding the LDCF/SCCF Council's request that proposals be prepared to ensure "visibility of the fund by requiring projects to identify their funding source," the Evaluation Office finds that additional work is needed by the Secretariat to fulfill the Council's request, and that the Secretariat may wish to consider adopting measures such as a separate logo to enhance the fund's visibility.

This LDCF/SCCF Council decision will be included in MAR 2013, as the level of adoption is not yet sufficient to warrant its graduation, and the decision is still relevant to the SCCF.

A complete version of MAR 2012 is available at the GEF Evaluation Office [website](#).

TABLE 7.2 Summary of Council Decisions Graduated from the MAR

MAR	Fully adopted			No longer relevant			Total
	High	Substantial	Medium	Negligible	Not possible to verify yet	N.A.	
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
2011	2	—	—	—	—	—	2
Total	38	27	11	6	2	2	86

NOTE: — = not available; N.A. = not applicable.

## 8. Performance Matrix

This chapter presents a summary, in table form (see table 8.1), of the performance of GEF Agencies across a range of parameters including results, processes affecting results, and M&E.<sup>1</sup> Some of the parameters included in the performance matrix, such as outcome ratings and cofinancing, are covered in the preceding chapters, while others are only reported here. Values presented are two- and four-year averages depending upon the parameter, or, in the case of Parameters 6 and 8, assessments of oversight processes and M&E arrangements updated as needed (see below). Ten parameters are covered, for which information is available on nine.

### 8.1 Performance Indicators

The 10 performance indicators and associated reporting methodology used are as follows:

- Overall outcome ratings, cofinancing, project extensions, and quality of M&E implementation (Parameters 1, 3a, 3b, 4, and 9) are four-year averages (APR 2009–12). For averages on outcome ratings, project extensions, and quality of M&E implementation, each project is given equal weight. Averages on cofinancing are four-year averages of total materialized cofinancing

in a given APR year cohort to the total GEF grant in a given APR year cohort, and percentage of total promised cofinancing materialized in a given APR year cohort. Percentages and values on individual GEF Agencies exclude projects under joint implementation.

- Quality of supervision and adaptive management (Parameter 2) and realism of risk assessment (Parameter 7) are findings from a 2009 follow-up assessment of project supervision, and candor and realism in project supervision reporting, first conducted in FY 2006. Forty-seven projects under implementation during FY 2007 and 2008 were sampled for this review (see [APR 2009](#) for complete details on the methodology used). A follow-up study is anticipated for APR 2013.
- Parameter 4, average time required to prepare projects, is the subject of an ongoing assessment, and will be reported on in OPS5.
- Parameter 5, average length of project extensions, is a four-year average (APR 2009–12) of the time taken to complete project activities beyond that anticipated in project approval documents. The averages include all projects with and without project extensions for which data on project extensions are available. Data for individual GEF Agencies exclude projects under joint implementation.
- Parameter 6, which assesses the independence and integrity of the process followed by GEF

<sup>1</sup> There is currently insufficient information to report on the individual performance of GEF Agencies other than UNDP, UNEP, and the World Bank.

TABLE 8.1 Performance Matrix for GEF Agencies and the GEF Overall

Parameter	UNDP	UNEP	World Bank	Overall GEF performance
<b>Results</b>				
1. Project outcomes: percentage of completed projects with outcomes rated moderately satisfactory or above (FY 2009–12)	88	95	79	86
<b>Processes affecting results</b>				
2. Quality of supervision and adaptive management: percentage of projects rated moderately satisfactory or above (FY 2007–08)	92	73	86	85
<b>Reported cofinancing<sup>a</sup></b>				
3a. Reported materialization of cofinancing per dollar of approved GEF financing (FY 2009–12)	5.8	1.7	3.0	4.0
3b. Reported materialization of cofinancing as percentage of promised cofinancing	190	145	106	144
<b>Efficiency</b>				
4. Project preparation elapsed time: average number of months required to prepare projects	—	—	—	—
5. Average length of project extensions (months; FY 2009–12) <sup>b</sup>	16	14	12	15
<b>Quality of M&amp;E</b>				
6. Independence of terminal evaluations and review of terminal evaluations (where applicable) (FSPs/MSPs)	HS/HS	HS/HS	HS/n.a. <sup>c</sup>	S
7. Realism of risk assessment (robustness of project at-risk systems): percentage of projects rated moderately satisfactory or above in candor and realism in supervision reporting (FY 2007–08)	77	73	80	77
8. Quality assurance of project M&E arrangements at entry: percentage of projects compliant with critical parameters	88	92	100	80
9. Percentage of projects with M&E implementation ratings of moderately satisfactory or above (FY 2009–12)	75	67	57	66
10. Percentage of terminal evaluations rated moderately satisfactory or above (FY 2011–12)	83	92	83	84

NOTE: — = not available; HS = highly satisfactory; S = satisfactory.

a. Ratios include only projects for which data on realized cofinancing are available.

b. Average includes all projects with and without extensions.

c. Not applicable, because the World Bank's Independent Evaluation Group does not conduct independent review for MSPs.

Agencies in conducting terminal evaluations and independent review of terminal evaluations (where applicable), comprises findings from an assessment last updated in FY 2011. Ratings were provided on a six-point scale from highly unsatisfactory to highly satisfactory, and separately assessed for FSP and MSP evaluations. The following six dimensions were evaluated in arriving at overall ratings: (1) the extent to which the drafting of the terms of reference is independent of the project management

team, (2) the extent to which the recruitment of the evaluator was independent of the project management team, (3) the extent to which the Agency recruited the appropriate evaluator for the project, (4) the extent to which the M&E system provides access to timely and reliable information, (5) the extent to which there was any undue pressure from management on the evaluators regarding the evaluation process (e.g., in terms of site selection, selection of informants, confidentiality during interviews,



information disclosure, and ratings), and (6) the extent to which the evaluation was subjected to an independent review process.

- Parameter 8 assesses the extent to which projects' M&E design, as specified in the final version of an Agency's respective project approval document, meets critical parameters, as specified in the GEF's 2010 M&E Policy. Values shown are different from the M&E design ratings presented in chapter 4, as the ratings here are from a set of projects currently under implementation. Percentages shown are for 80 FSPs randomly sampled from the full FY 2011 cohort of 137 approved FSPs. For a complete description of the methodology used, see [APR 2011](#).
- Parameter 10, percentage of terminal evaluations rated moderately satisfactory or above, is a two-year average of terminal evaluation completion and includes FY 2011–12.

## 8.2 Findings

For the OPS5 cohort (APR 2009–12), outcome achievements on 281 completed projects were assessed in terminal evaluations. Of these, 86 percent were rated in the satisfactory range. Within this four-year cohort, 88 percent of 146 UNDP projects, 95 percent of 41 UNEP projects, and 79 percent of 72 World Bank projects had overall outcome ratings in the satisfactory range.

For the OPS5 cohort, there were reportedly \$4 of cofinancing realized per \$1 of GEF funding (based upon 264 projects for which data on actual cofinancing are available). Among Agencies, UNDP realized nearly \$6 in cofinancing per \$1 of GEF funding. For UNEP and the World Bank, the cofinancing realized per dollar of GEF funding is \$1.70 and \$3.00, respectively. Overall, GEF projects in the OPS5 cohort realized 144 percent of promised cofinancing. By GEF Agency, UNDP realized 190 percent of promised cofinancing, UNEP realized 145 percent, and the World Bank realized

106 percent. Figures are based on information provided by the Agencies in terminal evaluation reports or through other communications, and have not been verified.

Projects within the OPS5 cohort had on average a 15-month extension. While not indicative of project performance (see [chapter 4](#)), this does suggest that, in general, project time frames may be unrealistic given the conditions in which projects take place. By Agency and among the same cohort of projects, full-size UNDP projects received on average a 20-month extension, and full-size UNEP and World Bank projects received 18- and 13-month extensions on average, respectively. For MSPs, there is less distinction among Agencies in terms of project extensions. UNDP and UNEP MSPs received 12-month extensions on average, and World Bank MSPs received 11-month extensions on average.

The independence and integrity of the process followed by GEF Agencies in conducting terminal evaluations and independent review of terminal evaluations (Parameter 6) is satisfactory for the GEF overall; and highly satisfactory for UNDP, UNEP, and the World Bank, according to the most recent assessment conducted in FY 2011. The Independent Evaluation Group of the World Bank does not review MSP evaluations, and thus a rating of not applicable was assessed for the World Bank's independent review of MSPs.

Findings from the most recent realism of risk assessment, undertaken for APR 2009, show that of the 47 sampled GEF projects under implementation during FY 2007 and 2008, 77 percent were rated in the satisfactory range for candor and realism of risk reporting in project monitoring. By GEF Agency, 77 percent of sampled UNDP projects, 73 percent of sampled UNEP projects, and 80 percent of sampled World Bank projects were rated in the satisfactory range for realism of risk reporting.

Findings from the most recent assessment of project M&E arrangements at entry, undertaken in FY 2011, suggest that 80 percent of GEF projects



at the point of entry (based upon the final version of project approval documents submitted for GEF CEO endorsement) are compliant with critical M&E parameters called for in the 2010 GEF M&E Policy. By Agency, the percentage of sampled projects rated in the satisfactory range on this parameter was 88 percent for UNDP, 92 percent for UNEP, and 100 percent for the World Bank.

Only 66 percent of GEF projects in the OPS5 cohort have M&E implementation ratings in the satisfactory range. By GEF Agency, the percentage of projects with M&E implementation ratings in the satisfactory range is 75 percent for UNDP,

67 percent for UNEP, and 57 percent for the World Bank. Ratings of M&E systems provided in terminal evaluations since APR 2006 continue to show gaps in performance relative to other performance metrics.

For the APR 2011 and APR 2012 cohort, more than 80 percent of terminal evaluations are rated in the satisfactory range for overall quality of reporting. By GEF Agency, 83 percent of UNDP terminal evaluations, 92 percent of UNEP terminal evaluations, and 83 percent of World Bank terminal evaluations meet the threshold of moderately satisfactory or above.



# Annex A. Projects Included in APR 2012 Cohort

GEF ID	Project name	GEF Agency	Type	Focal area
87	Protected Areas Management Project	WB	FSP	BD
112	Photovoltaic Market Transformation Initiative (IFC)	WB/IFC	FSP	CC
503	Paraguayan Wildlands Protection Initiative	UNDP	FSP	BD
668	Coastal and Wetland Biodiversity Management at Cox's Bazar and Hakakuki Haor	UNDP	FSP	BD
776	Conservation and Sustainable Use of Medicinal Plants in Arid and Semi-arid Ecosystems	UNDP	FSP	BD
834	Promoting Biodiversity Conservation and Sustainable Use in the Frontier Forests of Northwestern Mato Grosso	UNDP	FSP	BD
843	Removal of Barriers to Rural Electrification with Renewable Energy	UNDP	FSP	CC
886	Implementation of Strategic Action Program for the Bermejo River Binational Basin: Phase II	UNEP	FSP	IW
963	Environmental Protection and Maritime Transport Pollution Control in the Gulf of Honduras	IDB	FSP	IW
1022	Integrated Ecosystem Management of Transboundary Areas between Niger and Nigeria Phase I: Strengthening of Legal and Institutional Frameworks for Collaboration and Pilot Demonstrations of IEM	UNEP	FSP	MF
1029	Renewable Energy Technology Development and Application Project (RETDAP)	UNDP	MSP	CC
1036	Conservation of "Tugai Forest" and Strengthening Protected Areas System in the Amu Darya Delta of Karakalpakstan	UNDP	MSP	BD
1043	Establishing Conservation Areas Landscape Management (CALM) in the Northern Plains	UNDP	FSP	BD
1081	Lima Urban Transport	WB	FSP	CC
1092	Integrated Ecosystem Management in Indigenous Communities	WB/IDB	FSP	BD
1093	Reversing Land and Water Degradation Trends in the Niger River Basin	WB/UNDP	FSP	IW
1097	Development of a Wetland Site and Flyway Network for Conservation of the Siberian Crane and Other Migratory Waterbirds in Asia	UNEP	FSP	BD
1100	Community-based Conservation of Biological Diversity in the Mountain Landscapes of Mongolia's Altai Sayan Ecoregion	UNDP	FSP	BD
1104	Conservation of the Montane Forest Protected Area System in Rwanda	UNDP	FSP	BD
1128	Biodiversity Management in the Coastal Area of China's South Sea	UNDP	FSP	BD
1137	Promoting the Use of Renewable Energy Resources for Local Energy Supply	UNDP	FSP	CC
1148	In-Situ Conservation of Kazakhstan's Mountain Agrobiodiversity	UNDP	FSP	BD
1177	Biodiversity Conservation in the Russian Portion of the Altai-Sayan Ecoregion	UNDP	FSP	BD
1221	Coastal and Biodiversity Management Project	WB	FSP	BD

GEF ID	Project name	GEF Agency	Type	Focal area
1246	Partnerships for Marine Protected Areas in Mauritius	UNDP	MSP	BD
1254	Integrating Watershed and Coastal Area Management (IWCAM) in the Small Island Developing States of the Caribbean	UNEP/UNDP	FSP	IW
1281	Solar and Wind Energy Resource Assessment	UNEP	FSP	CC
1308	Strategic Planning and Design for the Environmental Protection and Sustainable Development of Mexico	UNDP	MSP	MF
1338	South Africa Wind Energy Programme (SAWEP), Phase I	UNDP	FSP	CC
1343	Demonstrations of Integrated Ecosystem and Watershed Management in the Caatinga, Phase I	UNDP	FSP	MF
1353	Nature Conservation and Flood Control in the Yangtze River Basin	UNEP	FSP	MF
1399	Capacity Building for Implementation of Malaysia's National Biosafety Framework	UNDP	FSP	BD
1515	Consolidation of Ecosystem Management and Biodiversity Conservation of the Bay Islands	IDB	FSP	BD
1520	Development of a National Implementation Plan in India as a First Step to Implement the Stockholm Convention on Persistent Organic Pollutants (POPs)	UNIDO	FSP	POPs
1531	Coral Reef Targeted Research and Capacity Building for Management	WB	FSP	IW
1557	Removing Barriers to the Reconstruction of Public Lighting (PL) Systems in Slovakia	UNDP	MSP	CC
1612	Second National Communication of Brazil to the UNFCCC	UNDP	FSP	CC
1713	Improved Management and Conservation Practices for the Cocos Island Marine Conservation Area	UNDP	MSP	BD
1725	Biodiversity Conservation in Altos de Cantillana	UNDP	MSP	BD
1776	Strengthening the Network of Training Centers for Protected Area Management through Demonstration of a Tested Approach	UNEP	MSP	BD
1854	Biodiversity Conservation and Sustainable Development in the Gissar Mountains of Tajikistan	UNDP	MSP	BD
1899	Regional Programme on Electrical Energy Efficiency in Industrial and Commercial Service Sectors in Central America	UNDP	FSP	CC
2068	Integrating Protected Area and Landscape Management in the Golden Stream Watershed	UNDP	MSP	BD
2104	Catalyzing Sustainability of the Wetland Protected Areas System in Belarusian Polesie through Increased Management Efficiency and Realigned Land Use Practices	UNDP	FSP	BD
2107	Removing Barriers to Energy Efficiency Improvements in the State Sector in Belarus	UNDP	FSP	CC
2178	Promoting Sustainable Transport in Latin America (NESTLAC)	UNEP	MSP	CC
2193	Enabling Sustainable Dryland Management Through Mobile Pastoral Custodianship	UNDP	MSP	LD
2257	Demonstration of Fuel Cell Bus Commercialization in China, Phase 2	UNDP	FSP	CC
2440 <sup>a</sup>	Sustainable Land Management in Drought Prone Areas of Nicaragua	UNDP	FSP	LD
2492	Strengthening the Protected Area Network (SPAN)	UNDP	FSP	BD
2509	Sustainable Land Management for Combating Desertification (Phase I)	UNDP	FSP	LD
2538	Assessment of Risk Management Instruments for Financing Renewable Energy	UNEP	MSP	CC
2589	Institutionalizing Payments for Ecosystem Services	UNDP	FSP	BD
2654	Consolidation of the Protected Area System (SINAP II)—Third Tranche	WB	FSP	BD

GEF ID	Project name	GEF Agency	Type	Focal area
2686	Integrated Management of the Montecristo Trinational Protected Area	IDB	FSP	BD
2715	Disposal of PCB Wastes in Romania	UNIDO	MSP	POPs
2730	Conservation of Globally Important Biodiversity in High Nature Value Semi-natural Grasslands through Support for the Traditional Local Economy	UNDP	MSP	BD
2796	Building the Partnership to Track Progress at the Global Level in Achieving the 2010 Biodiversity Target (Phase I)	UNEP	FSP	BD
2800	Developing Institutional and Legal Capacity to Optimize Information and Monitoring System for Global Environmental Management in Armenia	UNDP	MSP	MF
2836	Conservation and Sustainable use of Biodiversity in the Kazakhstani Sector of the Altai-Sayan Mountain Ecoregion	UNDP	FSP	BD
2848	Improved Conservation and Governance for Kenya Coastal Forest Protected Area System	UNDP	MSP	BD
2863	Ensuring Impacts from SLM—Development of a Global Indicator System	UNDP	MSP	LD
2915	CPP Namibia: Adapting to Climate Change through the Improvement of Traditional Crops and Livestock Farming (SPA)	UNDP	MSP	CC
3011	Introduction of BAT and BEP methodology to demonstrate reduction or elimination of unintentionally produced POPs releases from the industry in Vietnam	UNIDO	MSP	POPs
3037	Conservation and Use of Crop Genetic Diversity to Control Pests and Diseases in Support of Sustainable Agriculture (Phase 1)	UNEP	FSP	BD
3062	Strengthening Institutional Capacities for Coordinating Multi-Sectoral Environmental Policies and Programmes	UNDP	MSP	MF
3068	Mainstreaming the Multilateral Environmental Agreements into the Country's Environmental Legislation	UNDP	MSP	MF
3069	Strengthening Capacity to Integrate Environment and Natural Resource Management for Global Environmental Benefits	UNDP	MSP	MF
3163	Strengthening Capacity to Implement the Global Environmental Conventions in Namibia	UNDP	MSP	MF
3235	CACILM Rangeland Ecosystem Management-under CACILM Partnership Framework, Phase 1	UNDP	MSP	LD
3237	Demonstrating Local Responses to Combating Land Degradation and Improving Sustainable Land Management in SW Tajikistan-under CACILM Partnership Framework, Phase 1	UNDP	MSP	LD
3309	Participatory Planning and Implementation in the Management of Shantou Intertidal Wetland	UNEP	MSP	IW
3310	Environmental Learning and Stakeholder Involvement as Tools for Global Environmental Benefits and Poverty Reduction	UNDP	MSP	MF
3355	CPP Namibia: Enhancing Institutional and Human Resource Capacity Through Local Level Coordination of Integrated Rangeland Management and Support (CALLC)	UNDP	MSP	LD
3557	Catalyzing Financial Sustainability of Georgia's Protected Area System	UNDP	MSP	BD
3620	The Caspian Sea: Restoring Depleted Fisheries and Consolidation of a Permanent Regional Environmental Governance Framework	UNDP	FSP	IW
3706	CBPF: Emergency Biodiversity Conservation Measures for the Recovery and Reconstruction of Wenchuan Earthquake Hit Regions in Sichuan Province	UNDP	MSP	BD
3811	International Commission on Land Use Change and Ecosystems	UNEP	MSP	BD

NOTE: BD = biodiversity; CC = climate change; IW = international waters; MF = multifocal; OD = ozone depletion; POPs = persistent organic pollutants; IFC = International Finance Corporation; WB = World Bank.

a. The Food and Agriculture Organization of the United Nations and the International Fund for Agricultural Development were part of the project steering committee for GEF Project 2440, implemented by UNDP.

# Annex B. Terminal Evaluation Report Review Guidelines

The assessments in the terminal evaluation reviews are 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 the 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 takes into account all the independent relevant information when verifying ratings.

## B.1 Criteria for Outcome Ratings

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,<sup>1</sup> relevance of the project results, and the project’s cost-effectiveness.

<sup>1</sup> Objectives are the intended physical, financial, institutional, social, environmental, or other develop-

The ratings on the outcomes of the project will be based on performance on the following criteria:<sup>2</sup>

- **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 (that is, 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?

ment results to which a project or program is expected to contribute (OECD DAC 2002).

<sup>2</sup> *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.

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.

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 follows:

- **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.

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

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

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 midvalues rounded upward.

## B.2 Impacts

Has the project achieved impacts, or is it likely that outcomes will lead to the expected impacts? Impacts are understood to include positive and negative, primary and secondary, long-term effects produced by a development intervention. They could be produced directly or indirectly and could be intended or unintended. The terminal evaluation review's preparer will take note of any mention of impacts, especially global environmental benefits, in the terminal evaluation report including the likelihood that the project outcomes will contribute to their achievement. Negative impacts mentioned in the terminal evaluation report should be noted and recorded in Section 2 of the terminal evaluation review template in the subsection on "Issues that require follow-up." Although project impacts will be described, they will not be rated.

## B.3 Criteria for Sustainability Ratings

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 framework and governance, and environmental.

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 as 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 a dam in a protected area could inundate a sizable area and thereby neutralize the biodiversity-related gains made by the project.

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.

A number rating of 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 = N.A. 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.

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.

## B.4 Criteria for Assessment of Quality of Project M&E Systems

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 plans 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.

- **M&E design.** Projects 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), SMART (specific, measurable, achievable, realistic, and timely) 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)?
- **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?

- **Other questions.** These include 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?

A number rating of 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 three criteria (M&E design, M&E plan implementation, and M&E properly budgeted and funded) 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.

The rating for M&E during implementation will be the overall rating of the M&E system:

**Rating on the Quality of the Project  
M&E System = b**

## B.5 Criteria for Assessment of Quality of Terminal Evaluation Reports

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.

A number rating of 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 report 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.

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 =  
 $0.3 \times (a + b) + 0.1 \times (c + d + e + f)$**

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

## B.6 Assessment of Processes Affecting Attainment of Project Outcomes and Sustainability

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 C. Notes on Methodology and Analysis

## C.1 Factors Associated with Project Outcomes

In chapter 4, it is reported that strong associations are found within the APR 2005–12 cohort between (1) overall outcome ratings and quality of implementation ratings, (2) overall outcome ratings and quality of execution ratings, and (3) overall outcome ratings and the realization of promised cofinancing. Shown in tables C.1–C.3 are the results of a GEF Evaluation Office analysis of ratings from APR 2005–12 projects (where available) supporting these claims.

In each table, “unsatisfactory outcomes” are projects with overall outcome ratings below moderately satisfactory. “Satisfactory outcomes” are projects with overall outcome ratings of moderately satisfactory or above. The same threshold is used to sort projects on the basis of quality of implementation ratings and quality of execution ratings. For realization of promised cofinancing, projects were sorted on the basis of whether or not the amount of materialized cofinancing was at least equal to the amount of promised cofinancing.

Shown in each two-way table are the actual counts of projects meeting both individual criteria and the expected number of projects (shown in parentheses), assuming no association between

criteria. Also shown below each table are Fisher’s chi-square statistic and results of a chi-square test. As the results indicate, associations between all three criteria are strong and statistically significant at a 95 percent confidence level.

## C.2 Associations between Performance Indicators

In chapter 5, it is reported that there is a strong association between M&E design and M&E implementation ratings. Shown in table C.4 are the results of a GEF Evaluation Office analysis of ratings from APR 2006–12 projects ( $n = 328$ ) supporting these claims. “Unsatisfactory M&E design” includes all projects with M&E design ratings below moderately satisfactory. “Satisfactory M&E design” includes all projects with M&E design ratings of moderately satisfactory or above. The same threshold is used to sort projects on the basis of M&E implementation ratings.

As before, the two-way table shows actual counts of projects meeting both individual criteria, and the expected number of projects (shown in parentheses), assuming no association between M&E design and M&E implementation ratings. As the results indicate, there is a strong association between the two ratings, statistically significant at a 95 percent confidence level.

TABLE C.1 Relationship between Overall Outcomes and Quality of Implementation Ratings

Outcomes	Quality of implementation		Total
	Unsatisfactory	Satisfactory	
Unsatisfactory	35 (9)	15 (41)	50
Satisfactory	26 (52)	259 (233)	285
Total	61	274	335

NOTE:  $\chi^2(1) = 105.8$ ;  $p$ -value = 0.000. Each cell shows actual and (expected) frequency of projects.

TABLE C.2 Relationship between Overall Outcomes and Quality of Execution Ratings

Outcomes	Quality of execution		Total
	Unsatisfactory	Satisfactory	
Unsatisfactory	33 (8)	15 (40)	48
Satisfactory	20 (45)	262 (237)	282
Total	53	277	330

NOTE:  $\chi^2(1) = 115.7$ ;  $p$ -value = 0.000. Each cell shows actual and (expected) frequency of projects.

TABLE C.3 Relationship between Overall Outcome Ratings and Realization of Promised Cofinancing

Outcomes	Actual cofinancing < promised cofinancing	Actual cofinancing $\geq$ promised cofinancing	Total
Unsatisfactory	34 (26)	45 (53)	79
Satisfactory	126 (134)	281 (273)	407
Total	160	326	486

NOTE:  $\chi^2(1) = 4.4$ ;  $p$ -value = 0.037. Each cell shows actual and (expected) frequency of projects.

TABLE C.4 Relationship between M&amp;E Design and M&amp;E Implementation

M&E design	M&E implementation		Total
	Unsatisfactory	Satisfactory	
Unsatisfactory	40 (10)	13 (44)	53
Satisfactory	19 (50)	256 (226)	275
Total	59	269	328

NOTE:  $\chi^2(1) = 141.6$ ;  $p$ -value = 0.000. Each cell shows actual and (expected) frequency of projects.

## Annex D. GEF Regions

The analysis presented in chapters 2 and 3 includes ratings on the basis of the region in which GEF project activities take place. Four regions are defined; following are the countries included in each region.

- **Africa.** Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Democratic Republic of Congo, Republic of Congo, Côte d'Ivoire, Djibouti, Arab Republic of Egypt, Eritrea, Ethiopia, Gabon, The Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mayotte, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, São Tomé and Príncipe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Republic of Yemen, Zambia, Zimbabwe
- **Asia.** Afghanistan, American Samoa, Bangladesh, Bhutan, Cambodia, China, Fiji, India, Indonesia, Kiribati, Democratic People's Republic of Korea, Republic of Korea, Lao People's Democratic Republic, Malaysia, Maldives, Marshall Islands, Federated States of Micronesia, Mongolia, Myanmar, Nepal, Palau, Pakistan, Papua New Guinea, Philippines, Samoa, Solomon Islands, Sri Lanka, Thailand, Timor-Leste, Tuvalu, Tonga, Vanuatu, Vietnam
- **Europe and Central Asia.** Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Georgia, Iran, Iraq, Jordan, Kazakhstan, Kosovo, Kyrgyz Republic, Latvia, Lebanon, Lithuania, former Yugoslav Republic of Macedonia, Moldova, Montenegro, Romania, Russian Federation, Serbia, Syrian Arab Republic, Tajikistan, Turkey, Turkmenistan, Ukraine, Uzbekistan, West Bank and Gaza
- **Latin America and the Caribbean.** Antigua and Barbuda, Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, St. Lucia, St. Vincent and the Grenadines, Suriname, Uruguay, República Bolivariana de Venezuela

# Bibliography

- GEF Assembly (Global Environment Facility Assembly). 2006. "[Summary of Negotiations on the Fourth Replenishment of the GEF Trust Fund](#)." GEF/A.3/6.
- GEF (Global Environment Facility). 2010. "[Policies and Procedures for the Execution of Selected GEF Activities—National Portfolio Formulation Exercises and Convention Reports—with Direct Access by Recipient Countries](#)." GEF/C.38/6/Rev.1.
- . 2012. "[Updated Operational Guidelines for the Special Climate Change Fund for Adaptation and Technology Transfer](#)." GEF/LDCF.SCCF.13/05.
- GEF EO (Global Environment Facility Evaluation Office). 2005. "[Procedures and Format of the GEF Management Action Record](#)." GEF/ME/C.27/3.
- . 2008a. [GEF Annual Performance Report 2006](#). Evaluation Report No. 38.
- . 2008b. [GEF Annual Performance Report 2007](#). Evaluation Report No. 40.
- . 2008c. [Joint Evaluation of the GEF Small Grants Programme](#). Evaluation Report No. 39.
- . 2009a. [GEF Annual Country Portfolio Evaluation Report 2009](#). Evaluation Report No. 50.
- . 2009b. [GEF Annual Performance Report 2008](#). Evaluation Report No. 49.
- . 2010a. [GEF Annual Impact Report 2009](#). Evaluation Report No. 55.
- . 2010b. [GEF Annual Performance Report 2009](#). Evaluation Report No. 57.
- . 2010c. [The GEF Monitoring and Evaluation Policy 2010](#). Evaluation Document No. 4.
- . 2011a. [Evaluation of the GEF Strategic Priority for Adaptation](#). Evaluation Report No. 61.
- . 2011b. [GEF Annual Thematic Evaluations Report 2011](#). Evaluation Report No. 69.
- . 2012a. [Evaluation of the Special Climate Change Fund](#). Evaluation Report No. 73.
- . 2012b. [GEF Annual Country Portfolio Evaluation Report 2012](#). Evaluation Report No. 74.
- . 2012c. [GEF Annual Impact Report 2012](#). Evaluation Report No. 76.
- . 2012d. [GEF Annual Performance Report 2011](#). Evaluation Report No. 80.
- . 2012e. [GEF Annual Thematic Evaluations Report 2012](#). GEF/ME/C.43/02.
- GEF Secretariat (Global Environment Facility Secretariat). 2010. "[GEF-5 Programming Document](#)." GEF/R.5/25/CRP.1.
- . 2012. "[Updated Operational Guidelines for the Least Developed Countries Fund](#)." GEF/LDCF.SCCF.13/04.





# Recent GEF Evaluation Office Publications

## Evaluation Reports

82	Evaluación de la cartera de proyectos del FMAM en Cuba (1992–2011), Volumens 1 y 2	2013
81	Avaliação de Portfólio de Projetos do GEF: Brasil (1991–2011), Volumes 1 e 2	2013
80	GEF Annual Performance Report 2011	2013
79	OPS5: First Report: Cumulative Evidence on the Challenging Pathways to Impact	2013
78	Evaluation of the GEF Focal Area Strategies	2013
77	GEF Country Portfolio Study: Timor-Leste (2004–2011)	2013
76	GEF Annual Impact Report 2012	2013
75	The GEF in the South China Sea and Adjacent Areas	2013
74	GEF Annual Country Portfolio Evaluation Report 2012	2012
73	Evaluation of the Special Climate Change Fund	2012
72	GEF Beneficiary Countries of the OECS (1992–2011) (Antigua and Barbuda, Dominica, Grenada, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines), Volumes 1 and 2	2012
71	Evaluación de la cartera de proyectos del FMAM en Nicaragua (1996–2010), Volumens 1 y 2	2012
70	Evaluation of GEF National Capacity Self-Assessments	2012
69	Annual Thematic Evaluation Report 2011	2012
68	GEF Annual Impact Report 2011	2012
67	Estudio de la cartera de proyectos del FMAM en El Salvador (1994–2010), Volumens 1 y 2	2012
66	GEF Country Portfolio Study: Jamaica (1994–2010), Volumes 1 and 2	2012
65	GEF Annual Performance Report 2010	2011
64	GEF Annual Country Portfolio Evaluation Report 2011	2011
63	GEF Annual Impact Report 2010	2011
62	Review of the Global Environment Facility Earth Fund	2011
61	Evaluation of the GEF Strategic Priority for Adaptation	2011
60	GEF Country Portfolio Evaluation: Turkey (1992–2009)	2011
59	GEF Country Portfolio Evaluation: Moldova (1994–2009)	2011
58	GEF Annual Country Portfolio Evaluation Report 2010	2010
57	GEF Annual Performance Report 2009	2010
56	GEF Impact Evaluation of the Phaseout of Ozone-Depleting Substances in Countries with Economies in Transition, Volumes 1 and 2	2010
55	GEF Annual Impact Report 2009	2010
54	OPS4: Progress Toward Impact—Fourth Overall Performance Study of the GEF, Full Report	2010
53	OPS4: Progress Toward Impact—Fourth Overall Performance Study of the GEF, Executive Version	2010

## Evaluation Documents

ED-4	The GEF Monitoring and Evaluation Policy 2010	2010
ED-3	Guidelines for GEF Agencies in Conducting Terminal Evaluations	2008
ED-2	GEF Evaluation Office Ethical Guidelines	2008

## Learning Products

LP-3	The Journey to Rio+20: Gathering Evidence on Expectations for the GEF	2012
LP-2	Climate Change and the GEF	2010
LP-1	Biodiversity and the GEF	2010

To see all GEF Evaluation Office publications, please visit our [webpage](#).



Global Environment Facility  
Evaluation Office  
1818 H Street, NW  
Washington, DC 20433  
USA  
[www.gefeo.org](http://www.gefeo.org)