Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

GEF Project Performance and Progress to Impact

NOVEMBER 2018

FULL REPORT

Global Environment Facility
Independent Evaluation Office

GEF Project Performance and Progress to Impact

November 2018

Evaluation Report No. 121

This report was presented to the Council in November 2017.

1818 H Street, NW, Washington, DC 20433
Internet: www.gefieo.org/; email: gefevaluation@thegef.org

ISBN: 978-1-64233-004-5

Cover: Train tracks to Lower Usuthu Smallholder Irrigation Project in Swaziland; photo by Malac Kabir/GEF IEO

Contents

Tables

Foreword

The independent overall performance studies of the Global Environment Facility (GEF) are undertaken to inform the GEF replenishment by providing evaluative evidence of achievements and results. This evaluation, which was undertaken as an input for the preparation of the GEF Sixth Comprehensive Performance Study (OPS6), assesses performance of the GEF portfolio in terms of outcomes, sustainability, implementation, monitoring and evaluation, cofinancing, quality of terminal evaluations, time lags in project cycle, and progress to replenishment targets. It also addresses progress to impact achieved by the completed projects at the point of their completion, including environmental stress reduction and/or environmental status change, and broader adoption.

The analysis of completed projects is based primarily on information provided in the terminal evaluations for 1,184 completed projects. This includes the OPS6 cohort of 581 projects for which terminal evaluations were received after close of the Fifth Overall Performance Study. Depending on the evidence provided in the terminal evaluations and other sources of information, project performance was assessed. When there was an information gap that precluded assessment of a project's performance on a parameter, that project was not rated on the given parameter.

For project cycle time lags, data in the Project Management Information System (PMIS) submitted to the GEF Secretariat on project information forms were used. For trends in promised cofinancing, PMIS data on approved GEF projects WERE used.

Juha I. Uitto
Director, GEF Independent Evaluation Office

Acknowledgments

This evaluation was led by Neeraj Kumar Negi, Senior Evaluation Officer in the Independent Evaluation Office of the Global Environment Facility (GEF IEO), who was also the lead author of the report. Molly Watts Sohn (Evaluation Analyst, GEF IEO) coauthored the report.

The terminal evaluation review for fiscal year 2016 and the review of progress toward impact of the Sixth Comprehensive Performance Study (OPS6) cohort was coordinated by Molly Watts Sohn. Terminal evaluations and performance reviews were prepared by Malac Kabir and Molly Watts Sohn (IEO staff) and by Spandana Battula, Matteo Borzoni, Mathias Einberger, Ritu Kanotra, Caroline Laroche, Punjanit Leagnavar, Chenhao Liu, Mireia Duran Mate, Laura Reynolds Nissley, Maria Elisa Passeri, and Ekaterina Anastasia Verkhovsky (consultants). Analysis on progress toward GEF-5 and GEF-6 targets was conducted by Ritu Kanotra. Analysis of the GEF activity cycle was prepared by Mathias Einberger.

The evaluation benefited from guidance and oversight provided by Juha Uitto, Director of the IEO; quality control was provided by Geeta Batra, IEO Chief Evaluation Officer. The study team was supported by Evelyn Chihuguyu, IEO Program Assistant. We would also like to acknowledge Charles Hagner for editing the report and Nita Congress for design and layout assistance.

The GEF IEO is grateful to all of these individuals and institutions for their contributions. Final responsibility for this report remains firmly with the Office.

Abbreviations

APR	annual performance report
CEO	Chief Executive Officer
FSP	full-size project
GEF	Global Environment Facility
IEO	Independent Evaluation Office
LDC	least developed country
LLDC	landlocked developing country
M&E	monitoring and evaluation
MSP	medium-size project
OPS	overall performance study
PIF	project identification form
PMIS	Project Management Information System
SIDS	small island developing states
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme

The GEF replenishment periods are as follows: pilot phase: 1991–94; GEF-1: 1995–98; GEF-2: 1999–2002; GEF-3: 2003–06; GEF-4: 2006–10; GEF-5: 2010–14; GEF-6: 2014–18; GEF-7: 2018–22.

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

Executive summary

Overall, performance ratings of completed Global Environment Facility (GEF) projects show an improvement from GEF-3 to GEF-4. While it remains to be seen whether this uptick in ratings is stable, as only 41 percent of approved GEF-4 projects have been covered so far, it may be said that the performance of GEF-4 projects is either higher or as high as that of projects from the preceding periods. The key findings of the analyses follow.

The GEF has built a strong record in delivering short- and medium-term outcomes. Of the 1,173 projects rated on outcomes, 81 percent rated in the satisfactory range. Of the cohort included in the Sixth Comprehensive Evaluation of the GEF (OPS6), outcomes of 577 projects were rated, and 79 percent rated in satisfactory range. The ratings underscore the solid track record of GEF projects in delivering expected short- to medium-term results.

There are considerable risks to continuation of the benefits from more than a third of GEF projects. Of the 1,118 projects rated on sustainability of outcomes, 62 percent rated in the likely range. Thus, roughly 4 of 10 projects face considerable risks to continuation of their benefits. Of the OPS6 cohort, 545 projects were rated for sustainability, of which 63 percent (346 projects) rated in the likely range.

GEF Agencies generally implement GEF-supported projects in a satisfactory manner. Of the 970 projects rated on quality of implementation, 79 percent were rated in the satisfactory range.[1] Of the OPS6 cohort, 547 projects were rated for quality of implementation, of which 79 percent (432 projects) rated in the satisfactory range. Although there is an improving trend across the GEF periods, much of the gains took place during the GEF-1 period.

Despite an improving trend, cumulative ratings on the quality of the design and implementation of the monitoring and evaluation (M&E) remain in the unsatisfactory range for a substantial percentage of projects. Of the 1,108 projects that were rated for quality of M&E design, 61 percent (673 projects) rated in the satisfactory range. Of the OPS6 cohort, 570 were rated for M&E design, and 62 percent (353 projects) rated in the satisfactory range. There is a steady trend of improvement in quality of M&E design ratings. This trend is consistent with the findings of the quality-at-entry review presented in *Annual Performance Report* 2011, which showed improved compliance with the M&E design expectations. Of the 1,012 projects

[1] The terminal evaluation reviews are conducted annually as part of the work for the GEF Independent Evaluation Office's annual performance reports. During some of the review cycles, quality of implementation was not assessed. Consequently, a relatively higher percentage of completed projects have not been rated on quality of implementation.

that were rated on quality of M&E plan implementation, 64 percent rated in the satisfactory range. Of the OPS6 cohort, 546 were rated for M&E implementation, and 62 percent (341 projects) rated in the satisfactory range. There is an improving trend across the replenishment periods in which projects were approved. However, as was the case of M&E design, much of the improvement in ratings for M&E implementation has been achieved from the pilot phase to GEF-1.

Cofinancing commitments for GEF-6 projects exceed the target set by the GEF's cofinancing policy (2014). Against the target of 6:1 mandated by the cofinancing policy, cofinancing commitments for GEF-6 projects have been mobilized at a rate of 8.8:1 so far. Across the GEF periods—from GEF-1 to GEF-6—the cofinancing ratio of the GEF portfolio has increased steadily. In terms of the cofinancing ratio of the median full-size project (FSP), steady increase is evident from the pilot phase onward.

The promised cofinancing successfully materializes during implementation for the majority of projects. Cofinancing commitments were fully met for a majority (59 percent) of completed GEF projects. For one of eight completed projects (13 percent), less than half of the promised cofinancing materialized during implementation.

Most of the terminal evaluations submitted by the GEF Agencies meet the minimum quality expectations. Quality is in the satisfactory range for 83 percent of the 1,184 terminal evaluations. Of the 581 terminal evaluations received after the close of the Fifth Overall Performance Study, 571 were rated on quality of terminal evaluation, and 82 percent rated in the satisfactory range.

Despite some efficiency gains during the GEF-6 period, progress in improving project cycle efficiency has been slow. Of the 90 FSPs for which project identification forms (PIFs) were submitted during the first year of GEF-6, 37 percent had been CEO endorsed within 24 months of submission. Although this is an improvement over the performance during GEF-5 (26 percent) and GEF-4 (21 percent), the percentage of the GEF-6 PIF submissions that were CEO endorsed within 24 months of submission is still low. While the project cycle from PIF submission to PIF approval for GEF-6 projects was less efficient than for GEF-5 projects, it was more efficient for the PIF approval to CEO endorsement stages. The increase in time taken from PIF submission to PIF approval for GEF-6 projects seems to have been driven by the shortfall in GEF-6 replenishment. A fuller picture for the GEF-6 proposals will emerge only after GEF-6 has run its course and sufficient time has elapsed to track progress of the GEF-6 PIFs.

GEF programming for GEF-5 and GEF-6 is consistent with the corporate environmental results targets for these replenishment periods. Analysis of data on the targets promised in proposals for approved GEF-5 and GEF-6 projects allows an assessment of the extent to which programming is consistent with the corporate environmental results targets for these periods. The GEF is projected to exceed targets for 8 of the 13 corporate environmental results indicators for the GEF-5 period, although there may be some shortfall for the remaining five indicators, as level of programming is low for some of the focal area programs. For GEF-6, despite a shortfall in GEF resources, the aggregated results from approved PIFs exceed GEF-6 targets for 6 of 10 environmental results indicators. When the shortfall is accounted for, expected results are likely to be higher than the targets for 7 of 10 indicators.

The majority of GEF projects are already contributing to environmental stress reduction and/or status change at implementation completion. At project completion, 59 percent of the GEF projects

from the OPS6 cohort had already led to environmental stress reduction and/or status change. Thirteen percent of the projects were achieving environmental stress reduction and/or status change at a large scale, and 45 percent of projects were achieving it at a local scale. Whether a completed project achieved environmental stress reduction and/or status change appears to be linked with the environmental challenge being addressed, country context, global versus regional focus, or scale of GEF funding.

At project completion, stakeholders were adopting the approaches and technologies promoted by the majority of GEF projects. At project completion, 61 percent of completed GEF projects were achieving broader adoption. Country context plays an important role, as the percentage of implemented projects that were achieving broader adoption was substantially higher in the major emerging economies that account for the top five GEF project portfolios than in other countries.

1: Methodology

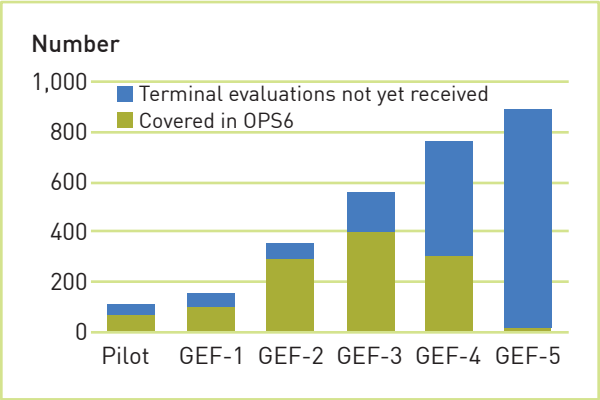
1.1 Performance of completed projects

Cumulatively, through December 2016, terminal evaluations for 1,184 completed projects have been received by the Global Environment Facility Independent Evaluation Office (GEF IEO). These projects account for $5.4 billion in approved GEF funding and $22.7 billion in cofinancing commitments. The analysis of outcomes, sustainability, implementation, materialization of cofinancing, monitoring and evaluation (M&E), and quality of terminal evaluation is based on data provided in these terminal evaluations. Of the 1,184 terminal evaluations, 581 projects were received after the close of the Fifth Overall Performance Study (OPS5). These 581 projects account for $2.7 billion in GEF funding and $14.9 billion in cofinancing commitments. From here on, these 581 projects are referred to as the OPS6 cohort.

It generally takes a project 6 to 10 years to move from the approval of the project identification form (PIF) to implementation completion. Consequently, considerable time elapses before terminal evaluations for all or almost all projects approved during a replenishment period become available. Based on a comparison of the number of projects that were approved during a given replenishment period—excluding canceled projects and projects for which terminal evaluations are not expected because of the small scale of GEF funding (that is, below $0.5 million)—and the terminal evaluations received so far, the coverage of completed projects up to GEF-3 is robust. For GEF-4, although 304 terminal evaluations are available, these represent only 41 percent of the projects (738 projects) from the period for which terminal evaluations are expected. As most GEF-5 projects are still under implementation, only nine terminal evaluations are available for projects from this period. It is too early for the majority of GEF-6 projects to be under implementation, let alone to be completed. When discussing the results of the completed projects, all 1,184 projects are covered. However, when data are presented based on replenishment periods, only data up to GEF-4 are presented.

The sizable gap in coverage of projects from GEF-4 has methodological implications (figure 1.1). Comparing outcome ratings for the data set used for the OPS6 analysis with those used for the OPS5 analysis shows that the projects for which terminal evaluations are received after a greater time lag tend to have lower outcome ratings than projects for which terminal evaluations are received earlier. For example, from OPS5 to OPS6, the percentage of projects with outcomes in the satisfactory range declined from 81 percent (*n* = 228 projects) to 79 percent (n = 289 projects) for GEF-2 projects, and from 88 percent (*n* = 176 projects) to 80 percent (*n* = 399 projects) for GEF-3 projects. Some of this is related to the size of the projects. A slightly higher percentage of medium-size projects (MSPs) are rated in the satisfactory range compared to full-size projects

FIGURE 1.1 OPS6 coverage, by GEF replenishment period of project approval



SOURCES: GEF IEO terminal evaluation review data set and GEF PMIS.

(FSPs) (84 percent versus 79 percent). Since MSPs also tend to have a shorter duration, their terminal evaluations are usually received before the evaluations for FSPs. However, much of the decline in outcome ratings is driven by late receipt of terminal evaluations for projects that experience difficulties during startup and implementation. Thus, it may be expected that, as terminal evaluations for a higher percentage of GEF-4 projects become available, the percentage of GEF-4 projects with outcomes rated in the satisfactory range may decline.

Details on criteria used by the GEF IEO to assess outcomes, sustainability, implementation, and quality of terminal evaluations are provided in "Guidelines on the Project and Program Cycle Policy" (GEF 2017b) and are also listed in annex A. Independent evaluation offices of some GEF Agencies, such as the World Bank, the United Nations Development Programme (UNDP), the United Nations Environment Programme (UNEP), and the International Fund for Agricultural Development, also provide performance ratings using criteria that are broadly consistent with those used by the GEF IEO. To keep from duplicating effort, and to encourage the Agencies' independent evaluation offices to play a greater role in the validation of terminal evaluations, beginning in 2009 with the World Bank and UNEP, the GEF IEO has been accepting the ratings provided by their evaluation offices. For 501 of the 1,184 projects (42 percent), ratings provided by the Agency evaluation offices have been used. For the remainder, ratings provided by the GEF IEO have been used. For quality control, 218 terminal evaluations that had been validated by Agency evaluation offices were also validated by the GEF IEO. Analysis of the ratings by the GEF IEO and Agency evaluation offices shows that, on the net, Agency evaluation offices tend to rate outcomes of 1.4 percent more projects in the satisfactory range than the GEF IEO—that is, 85.8 percent versus 84.4 percent. This difference between the ratings is not substantial. Therefore, for projects covered from 2009 onward, where available, ratings provided by the Agency evaluation offices have been used.

1.2 Progress to replenishment targets

The GEF-5 replenishment process established environmental results targets for the activities funded through the GEF-5 replenishment resources. To inform the process for the Sixth Replenishment of the GEF Trust Fund, the GEF IEO took stock of progress toward the GEF-5 replenishment targets. The first analysis was prepared for the December 2013 meeting of GEF-6 replenishment, and an updated analysis was prepared for its April 2014 meeting. These analyses were based primarily on an aggregation of the targets provided in the PIFs for the approved proposals. Most projects that were approved during the period (that is, GEF-5) have yet to be completed or to cross the midterm review milestone. Therefore, there is still little information on results achieved on the ground. However, 96 percent of the projects approved during GEF-5 have now been endorsed

or approved by the Chief Executive Officer (CEO). Therefore, more detailed projections of expected results are now available. To update the analysis of projects approved in GEF-5, documents submitted at CEO endorsement or approval for 686 projects that were funded partially or fully through GEF Trust Fund resources were reviewed. Enabling activities were excluded from the analysis, as these are not expected to result in environmental results directly. To arrive at the projections for the GEF-5 period, the aggregate of expected project results has been multiplied by a factor of 0.8 to account for cancellations and implementation failures.

For the GEF-6 period, data from GEF-6 PIF approvals maintained by the GEF Secretariat and presented in the GEF Corporate Scorecard have been used as the basis for reporting (GEF 2017a). The projections for GEF-6 have been arrived at by adjusting for likely cancellations and implementation failures (multiplied by a factor of 0.8) and for the level of GEF-6 resources used so far vis-à-vis expected GEF-6 replenishment.

1.3 Progress to impact

In preparation for OPS6, an assessment of progress to impact was mainstreamed in the reviews undertaken for the 2015 and 2016 annual performance reports (APRs). New terminal evaluations received for APR 2015 and APR 2016 were fully covered for assessment of progress to impact. For the remaining terminal evaluations of the OPS6 cohort that were submitted after the close of OPS5 (that is, submissions for APR 2013 and APR 2014), a representative sample of 50 percent was sampled. Thus, 426 completed projects in all were covered. After initial screening, 11 targeted research and/or foundational activities that are not expected to lead directly to environmental stress reduction and/or status change were removed from the analysis. Thus, progress to impact of 415 completed GEF projects was analyzed. The reviews to assess progress to impact were conducted using an instrument that, along with incidence of environmental stress reduction and/or status change, and broader adoption, also recorded the design features and implementation experience of the reviewed project. Probability weights were assigned when analyzing the results, so the results for the OPS6 cohort are not skewed by the submissions for APR 2015 and APR 2016, which had 100 percent probability of being represented in the sample (compared to 50 percent for APR 2013 and APR 2014). Also, the calculations were made without correcting for the differences in probability of being sampled. There is not much material difference in the calculations using the two approaches. In this paper, results that are not corrected for difference in sampling probability are presented in the main narrative. Probability-adjusted figures are presented in annex B.

Although progress to impact–related analysis was also presented in OPS5 (GEF IEO 2014b), the methodology for the OPS6 assessment is different from that used for the OPS5 analysis. The approach for OPS6 uses a higher threshold than OPS5 for recording incidence of environmental stress reduction and/or status change. Consequently, findings of the analysis undertaken for OPS6 are not directly comparable to those presented in OPS5.

1.4 Project cycle time lags

Analysis of project cycle time lags focuses on stand-alone FSPs, which are endorsed by the GEF CEO following a two-step process. The first step involves submission of a PIF by an Agency. This step culminates in PIF approval. The second step involves preparation of a detailed project proposal by the GEF Agency, submission of the proposal to the GEF Secretariat, and CEO endorsement of the

proposal. The GEF has established 18 months as the standard amount of time required for FSPs to move from PIF approval to CEO endorsement (GEF 2010a). Analysis of project cycle time lags focuses on the stages between PIF submission and project start. The time lag between project completion and the reporting of data on project implementation and completion makes it difficult to assess project cycle time lags for projects that were completed recently. Data in the Project Management Information System (PMIS) through June 2017 were used to determine the time lags.

Although it is important to cover MSPs and activities under the programmatic-approach framework, doing so was not feasible. MSPs were excluded because the Council's approval of the single-step CEO approval process for MSPs makes it difficult to assess the time required to prepare proposals that follow the single-step process. Although it is still possible to measure time lags for those that follow the two-step process, the number of observations are too low.

The number of child projects developed during GEF-6 under the programmatic approach is still too small to allow meaningful analysis. This constraint is accentuated because child projects developed within the programmatic approach are expected to meet their negotiated commitment deadline given in their respective framework document and not the 18-month standard applicable to stand-alone FSPs.

1.5 Cofinancing commitments and materialization

The analysis of cofinancing trends draws on data from different sources. The analysis of trends in cofinancing commitments is based on data in the PMIS through June 2017. These data have been used to calculate the cofinancing ratio of the GEF project portfolio and the median project cofinancing ratio for different project types across GEF replenishment periods.

The analysis of the materialization of cofinancing is based on data provided in the 1,184 terminal evaluations that have been submitted to the GEF IEO through December 2016. Of these, data on materialization of cofinancing are available for 84 percent (994 projects).

The analysis of the probability of materialization of cofinancing commitments for different sources of cofinancing is based on the survey of information provided in the project documents and in the terminal evaluation reports. Data on sources of cofinancing commitments and its materialized cofinancing are from 323 projects from OPS6 cohort for which this information was available.

2: Findings

2.1 Performance of completed projects

OUTCOMES

The Organisation for Economic Co-operation and Development defines outcomes as "the likely or achieved short-term and medium-term effects of an intervention's outputs" (OECD 2002). A GEF project is expected to deliver its expected outcomes by the end of its implementation. Terminal evaluations prepared by the GEF Agencies provide a record of the extent to which expected outcomes were delivered. Findings of these evaluations are then validated by the GEF IEO and/or the independent evaluation offices of the GEF Agencies. A six-point scale is used to rate the level of outcome achievement. Of these, the top three ratings compose the "satisfactory range," and the bottom three the "unsatisfactory range."[1]

Of the 1,184 completed GEF projects for which terminal evaluations have been submitted to the GEF IEO so far, 1,173 have been rated for their outcome achievements.[2] Of those rated, 81 percent rated in the satisfactory range. Of the 581 terminal evaluations that were received after the close of OPS5, outcomes of 577 were rated, and 79 percent rated in the satisfactory range (figure 2.1). The ratings underscore the solid track record of GEF projects in delivering expected short- to medium-term results. Comparison across periods shows that most GEF projects continue to deliver their expected outcomes.

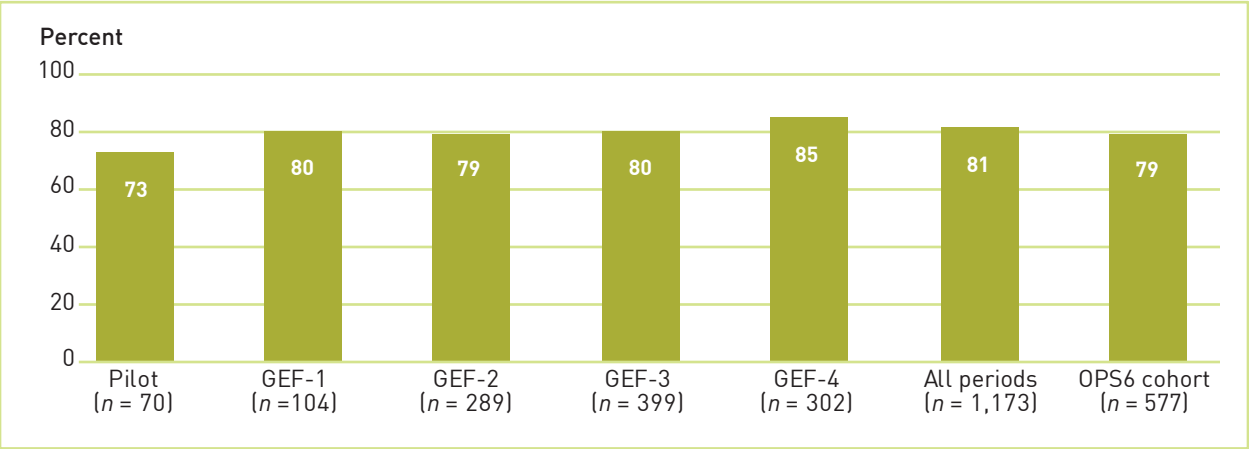
Of 304 GEF-4 projects for which terminal evaluations were submitted, outcomes of 302 projects were rated. Of those rated, 85 percent were rated in the satisfactory range. The policy recommendations for the GEF-4 replenishment set a target for outcome ratings of 75 percent of projects in the satisfactory range for the projects approved during this period (GEF 2006). So far, the GEF-4 projects have exceeded this expectation and are on track to meet the GEF-4 replenishment target.

Of projects implemented in Africa, 74 percent rated in the satisfactory range (figure 2.2a). This is significantly lower than 83 percent of projects in other regions, including global projects, which rated in the satisfactory range. However, there is considerable difference in performance across African countries. While outcomes of 90 percent of projects ($n = 29$) implemented in North African countries rated in the satisfactory range, outcomes for 69 percent of projects ($n = 74$) implemented in East African countries and 62 percent

[1] The ratings are highly satisfactory, satisfactory, and moderately satisfactory (the "satisfactory range") and moderately unsatisfactory, unsatisfactory, and highly unsatisfactory (the "unsatisfactory range").

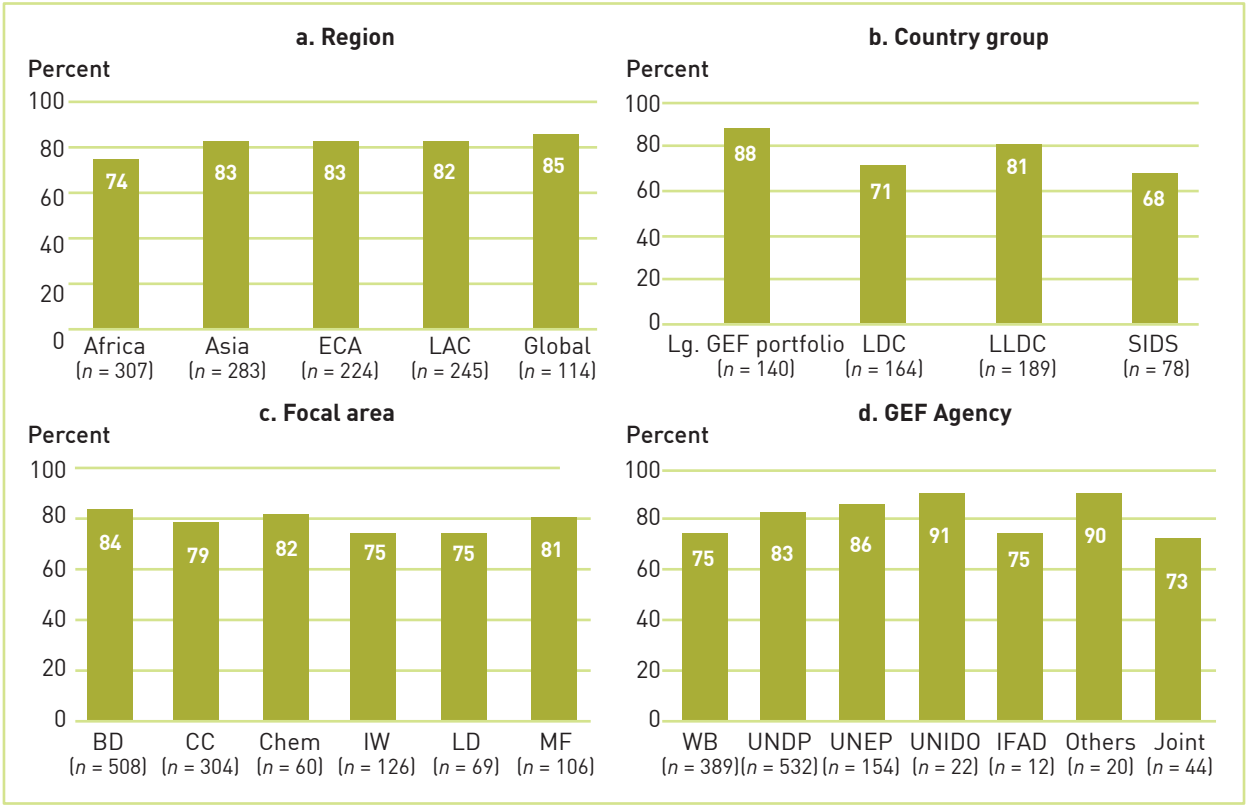
[2] For the remainder, ratings were not provided due to insufficient information provided in the terminal evaluations.

FIGURE 2.1 Percentage of projects with outcomes rated in the satisfactory range, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.2 Percentage of projects with outcomes rated in the satisfactory range, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization; IFAD = International Fund for Agricultural Development.

of projects (*n* = 26) in the West Sub-Saharan countries rated in the satisfactory range.[3]

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

Outcomes of projects in least developed countries (LDCs) and small island developing states (SIDS) were less likely to be rated in the satisfactory range (figure 2.2b).

For GEF focal areas, the percentage of projects rated in the satisfactory range for outcomes ranges from 75 percent to 84 percent (figure 2.2c). Outcomes of 75 percent of the international waters focal area projects were rated in the satisfactory range, which is lower than other GEF projects at 90 to 95 percent confidence level depending on the model used. The difference between other focal areas and remaining projects is not statistically significant.

Compared to other GEF Agencies, a higher percentage of projects implemented by UNEP were rated in the satisfactory range for outcomes. (The rating is by project and not by individual outcome). (figure 2.2d).[4] On the other hand, a lower percentage of projects implemented by the World Bank were rated in the satisfactory range. The difference in performance for other GEF Agencies is not statistically significant.

Level of outcome achievement may be determined by several factors. While some of them are tracked by the GEF, it is difficult to measure the extent to which they determine outcomes. Multiple linear regression suggests that quality of implementation, quality of execution, and shortfall in materialization of cofinancing are among the key determinants of outcome ratings (see annex D). Quality of implementation and quality of execution positively affect outcome ratings. Materialization of less than 50 percent of promised cofinancing negatively affects outcome ratings, as several planned activities are dropped or scaled down. Outcomes of a statistically higher percentage of projects implemented in large GEF portfolios and a significantly lower percentage of projects implemented in Africa, LDCs, and SIDS are rated in the satisfactory range. However, when variables such as quality of implementation, quality of execution, quality of M&E design, and materialization of cofinancing are controlled for, the relationship between whether a project was implemented in Africa or large economies and outcome ratings weakens and is not statistically significant. This shows that better outcome achievements may be achieved if implementing agencies accord greater attention to project preparation and to project implementation in Africa.

A review presented in APR 2014 analyzed the lessons reported in the terminal evaluations of 603 randomly selected completed GEF projects (GEF

[3] The North African countries are Algeria, the Arab Republic of Egypt, Libya, Morocco, and Tunisia. The East African countries are Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Mauritius, Rwanda, Seychelles, Somalia, South Sudan, Sudan, Tanzania, and Uganda. The West Sub-Saharan countries are Benin, Côte d'Ivoire, Ghana, Guinea, Liberia, Nigeria, Sierra Leone, and Togo. These subregions correspond to a GEF constituency of member countries, and each is represented in the GEF Council.

[4] The difference is significant at a 90 percent confidence level.

IEO 2015).[5] The review identified several reasons that lead to lower levels of results achievements. These include overly ambitious objectives, an inadequate budget for planned activities, a weak intervention strategy, inadequate arrangements to facilitate follow-up, inappropriate institutional arrangements, inadequate government and stakeholder support, poor M&E design, and so forth.

SUSTAINABILITY

Consistent with the Organisation for Economic Co-operation and Development's definition of sustainability (OECD 2002), the GEF's M&E policy defines sustainability as "the likely ability of an intervention to continue to deliver benefits for an extended period of time after completion" (GEF IEO 2010). The GEF IEO rates sustainability on a four-point scale based on an assessment of the level of risk to continuation of project benefits at the point of project completion. It takes financial, sociopolitical, institutional and governance, and environment risks into account. The top two ratings compose the "likely" range, and the bottom two the "unlikely" range.[6]

Of the 1,184 completed GEF projects for which terminal evaluations are available, 1,118 have been rated on sustainability. Of the rated projects, 62 percent (689 projects) rated in the likely range (figure 2.3). This shows that roughly 4 of 10 projects face considerable risks to continuation of their benefits. Of the 581 terminal evaluations that were received after the close of OPS5, 545 were rated for sustainability, and of these, 63 percent (346 projects) rated in the likely range. The trend across the GEF replenishment periods shows improvement in the sustainability ratings, although the figures for GEF-4 may regress toward the long-term average, as more terminal evaluations of the GEF-4 projects become available.

Within Africa, there is considerable variation in performance. While the sustainability of 64 percent of projects in North Africa (n = 28) is rated as likely, only 35 percent of projects in Sub-Saharan countries excluding Eastern and Southern Africa (n = 76) are so rated.[7]

Among other select country groups, 85 percent of projects (n = 135) in large GEF portfolio countries rated in the likely range for sustainability (figure 2.4b). In comparison, 44 percent of projects in LDCs (n = 154) and 55 percent of projects in SIDS (n = 72)—that is, countries where there are considerable constraints on capacity and resources—rated in the likely range for sustainability.

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

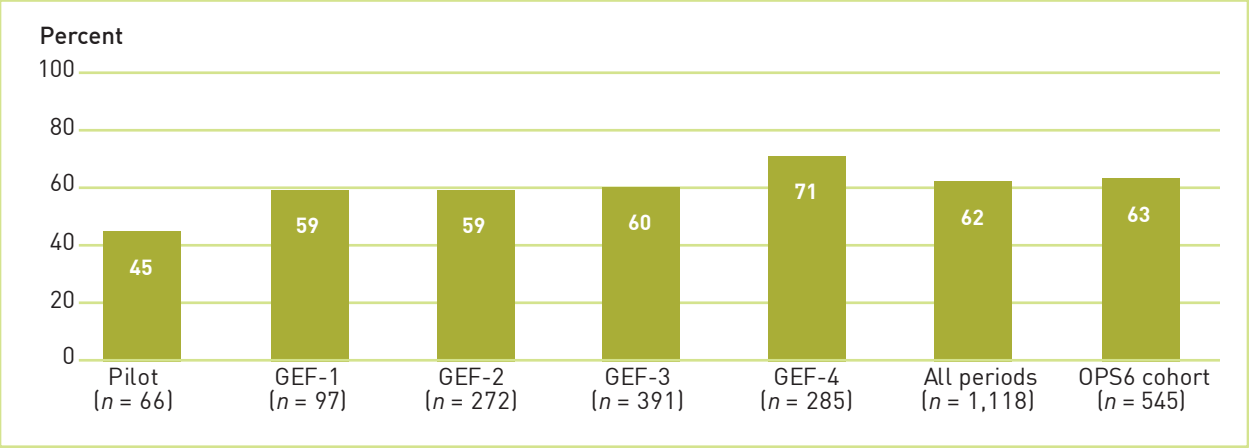
Compared to projects from other focal areas, sustainability of a higher percentage of climate

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

[6] The four-point scale used to rate sustainability is as follows: likely and moderately likely (both included in "likely" range) and moderately unlikely and unlikely (both included in "unlikely" range).

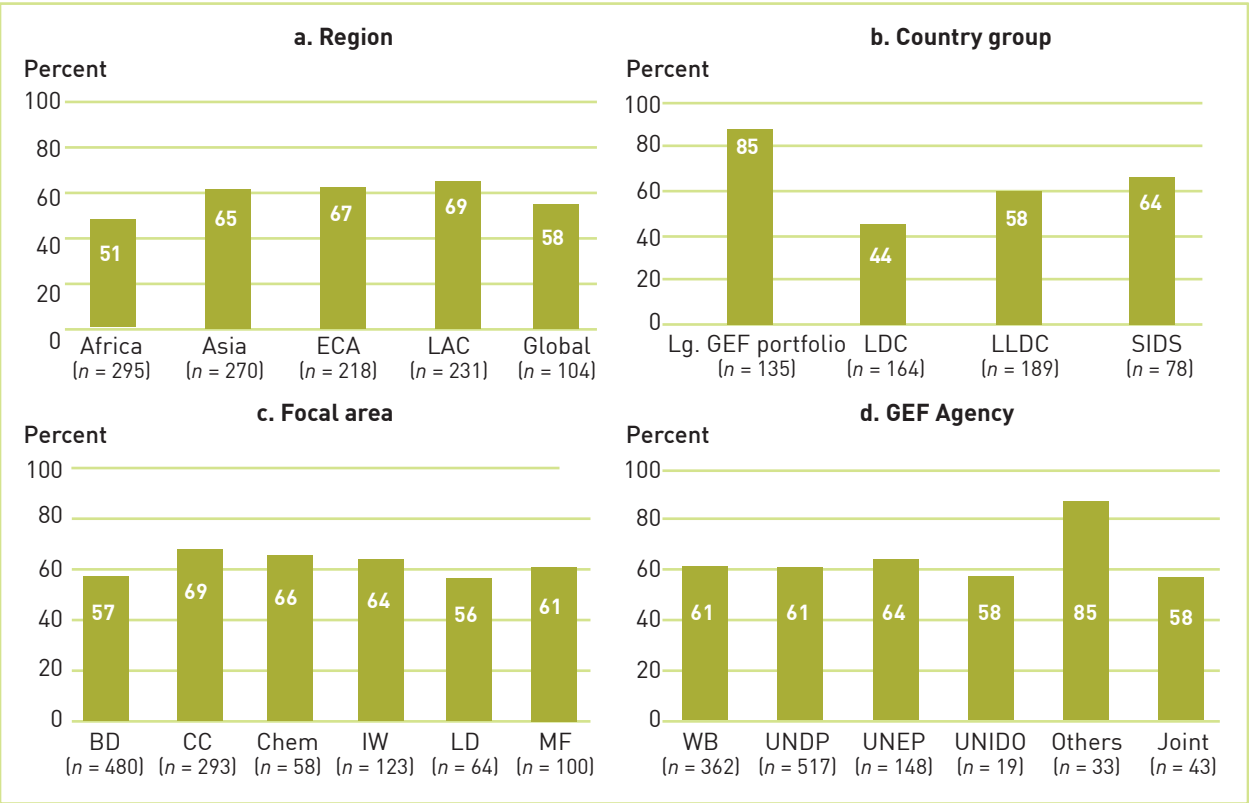
[7] This includes three GEF constituencies that consist of the following countries: Benin, Côte d'Ivoire, Ghana, Guinea, Liberia, Nigeria, Sierra Leone, and Togo; Burundi, Cameroon, Central African Republic, Democratic Republic of Congo, Republic of Congo, Equatorial Guinea, Gabon, and São Tomé and Principe; and Burkina Faso, Cabo Verde, Chad, Guinea-Bissau, Mali, Mauritania, Niger, Senegal, and The Gambia.

FIGURE 2.3 Percentage of projects with sustainability rated in the likely range, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.4 Percentage of projects with sustainability rated in the likely range, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization.

change projects (69 percent) rated in the likely range (figure 2.4c). The sustainability ratings of other focal areas are not statistically different from each other. The sustainability ratings by GEF Agency do not show much difference among major Agencies (figure 2.4d). However, when the projects of Agencies with smaller portfolios of completed projects are pooled together (Others), a higher percentage has sustainability ratings in the likely range. The relationship weakens and is not significant in several models when other variables are controlled for.

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

QUALITY OF IMPLEMENTATION

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

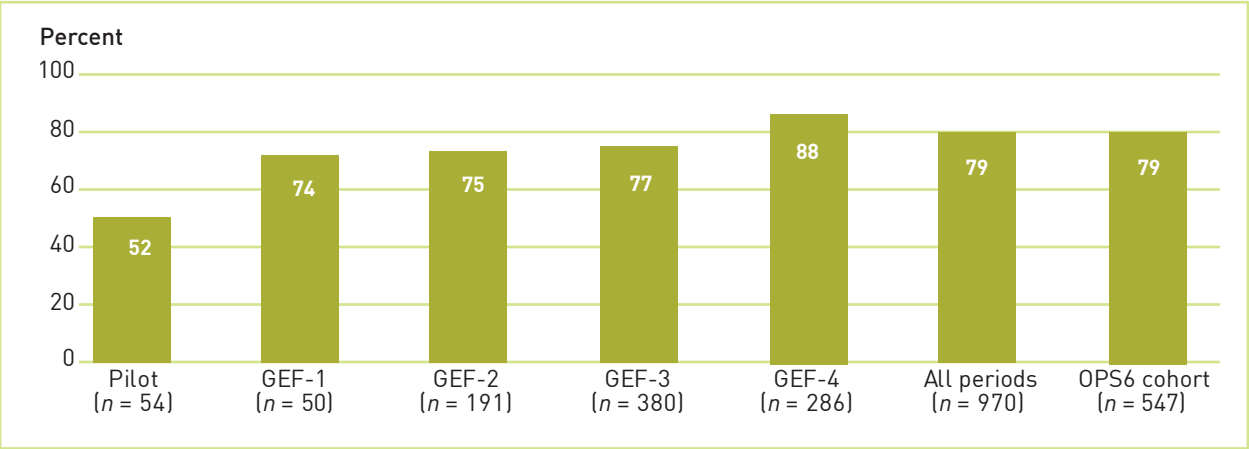
Of the 970 completed projects that were rated on quality of implementation, 79 percent (762 projects) rated in the satisfactory range (figure 2.5). Although there is an improving trend across the GEF periods, much of the gains took place during the GEF-1 period. Performance for GEF-5 and GEF-4 is likely to regress closer to the long-term average when more projects are completed. Of the 581 projects for which terminal evaluations were received after the close of OPS5, 547 were rated for quality of implementation, and of these, 79 percent (432 projects) rated in the satisfactory range.

A lower percentage of projects implemented in Africa, SIDS, and LDCs rated in the satisfactory range for quality of implementation (figure 2.6a and b). The difference is statistically significant even when controlling for variables such as focal area and GEF Agency. This suggests that GEF Agency capacities in these regions and country groups may be relatively weaker than in other regions and country groups.

Quality of implementation ratings of projects by focal area are closely bunched together (figure 2.6c). Seventy-three to 80 percent of projects rated in the satisfactory range, and differences across the focal areas are not statistically significant.

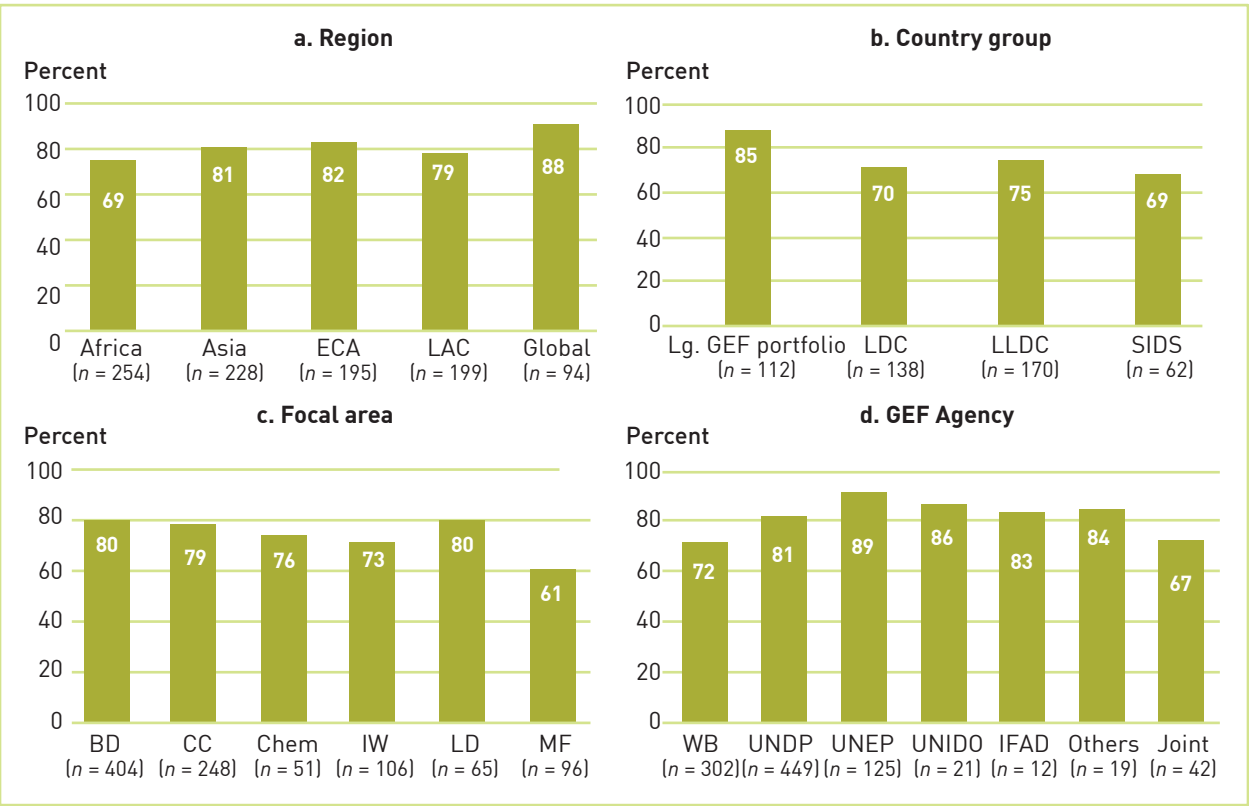
Among the GEF Agencies, a higher percentage of UNEP-implemented projects rated in the satisfactory range for quality of implementation, whereas a lower percentage of World Bank projects and jointly implemented projects were so rated (figure 2.6d). Lower ratings for World Bank–implemented projects are driven by low ratings for the projects from the GEF-3 period: Only 65 percent of World Bank–implemented projects from this period (n = 126) rated in the satisfactory range. As explained in APR 2013 and APR 2014, some of this drop may be due to stringent application of the rating criteria by the World Bank's Independent

FIGURE 2.5 Percentage of projects with quality of implementation rated in the satisfactory range, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.6 Percentage of projects with quality of implementation rated in the satisfactory range, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization; IFAD = International Fund for Agricultural Development.

Evaluation Group for the projects from this period (GEF IEO 2014a, 2015).[8] Information from the online survey reported in "Evaluation of the Expansion of the GEF Partnership: First Phase" (GEF IEO 2016) and from quality of supervision reviews presented in APR 2006 and APR 2009 indicates that the World Bank performs well in project implementation. Ratings for UNDP are close to the portfolio average. Within the UNDP portfolio, there was a substantial improvement from the pilot phase (26 percent, $n = 23$) to GEF-1 and beyond. For other Agencies and jointly implemented projects, the observations are too few to draw inferences.

The analysis of lessons presented in APR 2014 showed that quality of implementation may be poor because of inadequate oversight and technical support, an inability to take corrective measures in a timely manner, high staff turnover, ineffective project governance structures, and so forth. GEF Agencies need to mitigate the gap in implementation services for regions and country groups with capacity constraints.

PROJECT M&E DESIGN AND IMPLEMENTATION

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

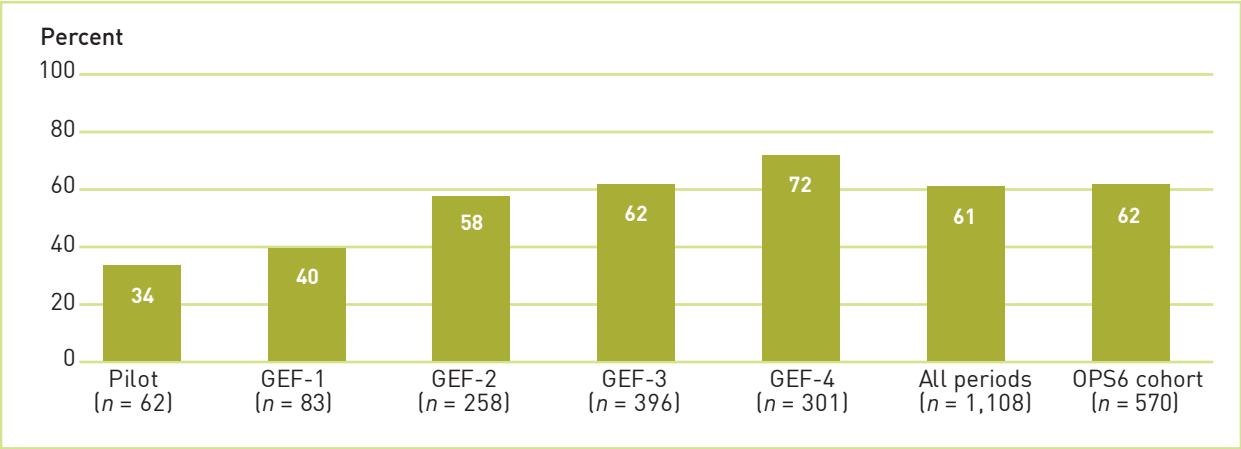
Of the 1,108 projects that were rated for quality of M&E design, 61 percent (673 projects) rated in the satisfactory range (figure 2.7). Ratings for quality of M&E design improved steadily across replenishment periods. This trend is consistent with the findings of the quality-at-entry review presented in APR 2011, which showed improved compliance with the M&E design expectations. Of the 570 projects of the OPS6 cohort that were rated for quality of M&E design, 62 percent rated in the satisfactory range.

A lower percentage of projects in Africa rated in the satisfactory range for M&E design than projects in other regions (figure 2.8a). This difference stays even when other variables are controlled for. A lower percentage of projects in LDCs rated in the satisfactory range (figure 2.8b). However, when other variables are controlled for, the difference in ratings of M&E design for projects in LDCs and those in other countries is not significant.

The percentage of projects from the chemicals focal area that rated in the satisfactory range for M&E design is lower than that of projects from other focal areas (figure 2.8c). This difference stays significant when other variables are controlled for. Among the GEF Agencies, projects implemented by the World Bank tend to have lower M&E design ratings than projects implemented by

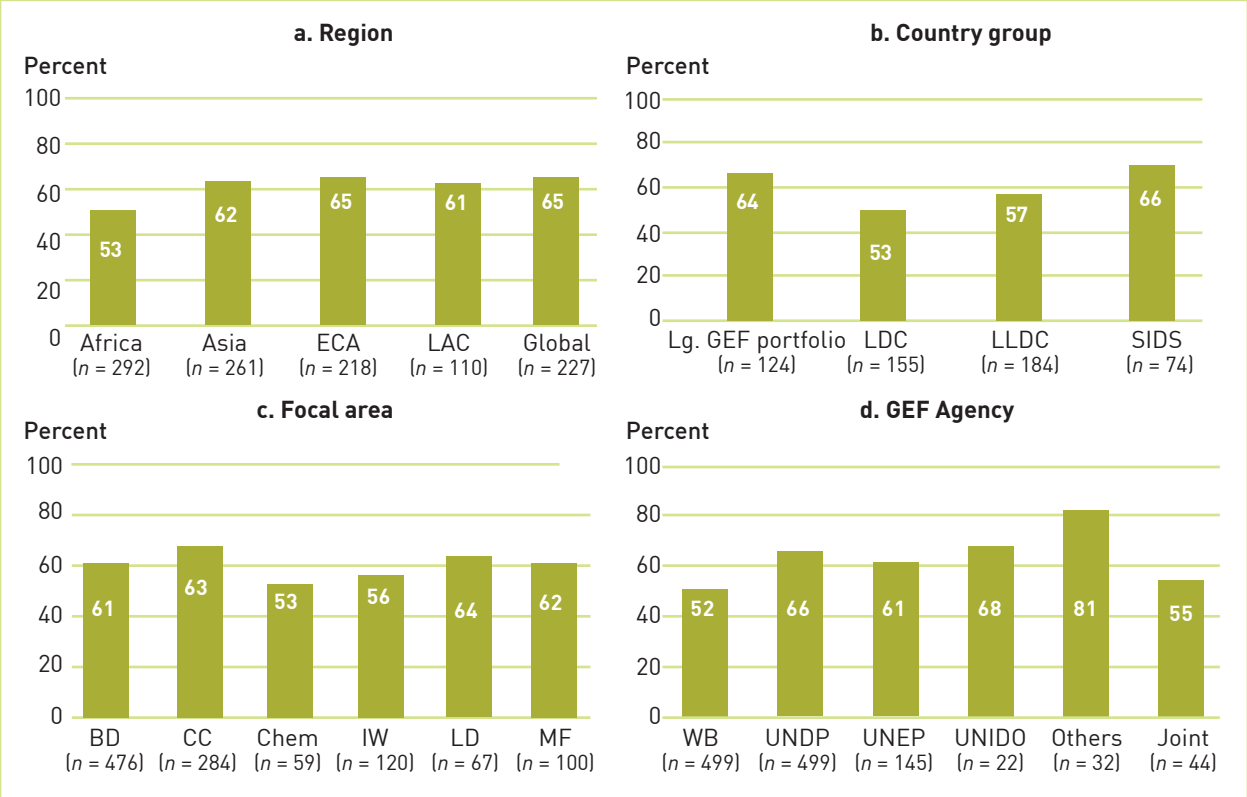
[8] The World Bank's Independent Evaluation Group appears to have applied more stringent criteria during validations that it conducted from 2009 to 2011. Since the GEF IEO accepts ratings provided by the Independent Evaluation Group, there was a drop in performance ratings for the projects from GEF-3.

FIGURE 2.7 Percentage of projects with quality of M&E design rated in the satisfactory range, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.8 Percentage of projects with quality of M&E design rated in the satisfactory range, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization.

other Agencies. The difference is significant when other variables are controlled for (figure 2.8d).

Of the 1,012 projects that were rated on quality of M&E plan implementation, 64 percent rated in the satisfactory range (figure 2.9). Much of the improvement in ratings was achieved from the pilot phase to GEF-1. Of the 546 projects from the OPS6 cohort, 62 percent rated in the satisfactory range. After approval of the GEF's M&E policy in 2006 and its revision in 2010—especially the inclusion of Minimum Standard 4 in the 2010 M&E policy, which calls for engagement of GEF Operational Focal Points in M&E activities and GEF-wide adoption of tracking tools from GEF-4 onward—there have been enhanced expectations for project M&E. This may mask the level of improvements in the quality of project M&E during the more recent periods.

A lower percentage of projects in Africa rated in the satisfactory range for their quality of M&E during implementation (figure 2.10a). Compared to M&E design ratings, M&E implementation ratings improved the most for projects in Europe and Central Asia and global projects (10 percent each). Only half of the projects implemented in LDCs had M&E implementation ratings in the satisfactory range (figure 2.10b). While M&E implementation ratings of projects in landlocked developing countries (LLDCs) and in countries with large portfolios showed some improvement vis-à-vis M&E design ratings, those in SIDS showed a 10 percent drop. This suggests that country context may affect how well M&E plans are implemented.

Although the percentage of multifocal projects that rated in the satisfactory range for M&E design is the same as that for projects of other focal areas, the percentage of multifocal projects rated in the satisfactory range for M&E implementation is lower (figures 2.8c and 2.10c). Compared to M&E design ratings, the percentage of multifocal projects rated in the satisfactory range for quality M&E implementation is 15 percent lower. This suggests that M&E implementation for multifocal projects may be more complicated than projects from other focal areas.

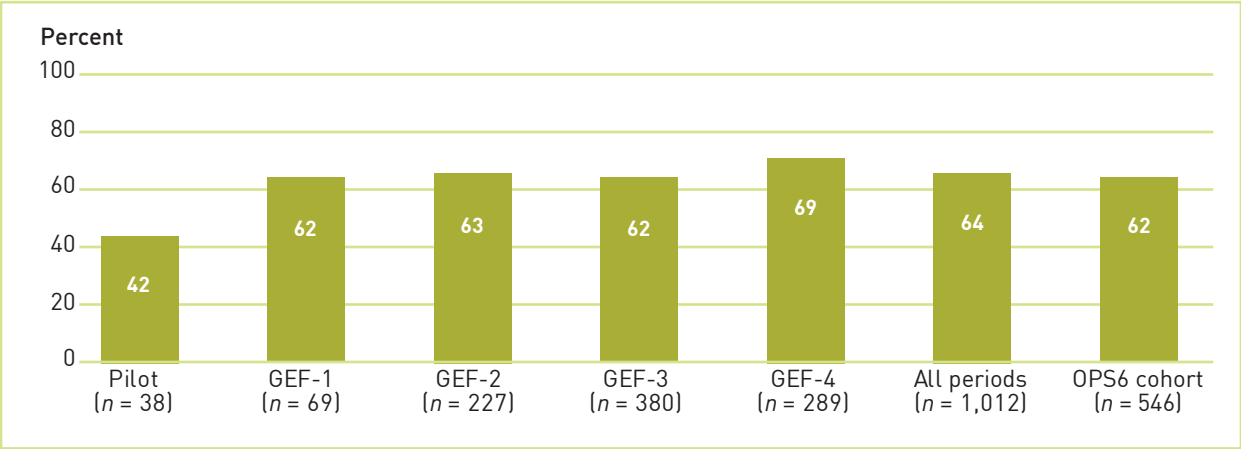
For most of the GEF Agencies, the percentage of projects rated in the satisfactory range for quality of M&E implementation closely tracks the percentage rated in the satisfactory range for quality of M&E design. However, "Others," which combines portfolios of Agencies that have small portfolios of completed GEF projects, is an anomaly. While 81 percent of the projects implemented by "Others" rated in the satisfactory range for M&E design, only 43 percent rated in the satisfactory range for quality of M&E implementation (figure 2.10d). Reasons for the drop are not well understood.

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

COFINANCING

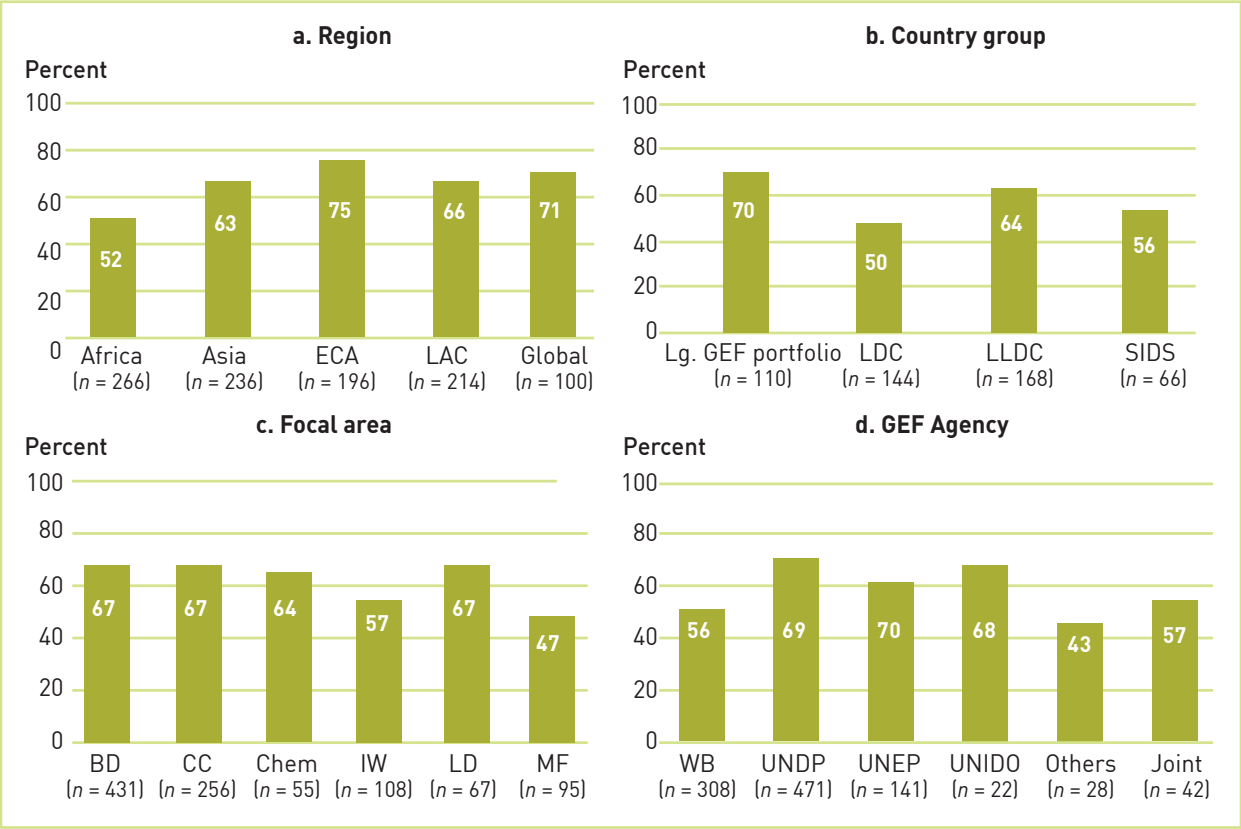
Cofinancing is generally considered important for mobilizing resources to achieve GEF objectives (GEF IEO 2013a). Given that the GEF provides funding to "meet the agreed incremental costs of measures to achieve agreed global environmental benefits" (GEF 2015a, 12), it needs to ensure that baseline costs are cofinanced by other partners. OPS5 noted the wide consensus across the preceding OPSs that cofinancing is beneficial to GEF projects (GEF IEO 2014b). Nonetheless, OPS5 also noted skepticism about the extent to which cofinancing helps generate additional resources

FIGURE 2.9 Percentage of projects with quality of M&E implementation rated in the satisfactory range, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.10 Percentage of projects with quality of M&E implementation rated in the satisfactory range, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization.

to achieve global environmental benefits. A recent paper that assessed the effects of cofinancing by analyzing the GEF project portfolio found that projects with higher cofinancing ratios are correlated with higher outcome ratings (Kotchen and Negi 2016). Given its importance, mobilization of cofinancing is tracked as an indicator of GEF performance.

The GEF's new cofinancing policy (GEF 2014), which became operational during GEF-6, targets a 6:1 level of cofinancing for the GEF portfolio. The promised cofinancing mobilized for GEF-6 projects through June 2017 is 8.8:1, exceeding the portfolio target (figure 2.11). Across the GEF periods—from GEF-1 to GEF-6—there has been a steady increase in the cofinancing ratio at the portfolio level. In terms of the cofinancing ratio of the median FSP, which corrects for the outsize influence outliers may have, this improvement is evident from the pilot phase onward. Steady improvement in the ratio for the median project shows that the increase in the portfolio cofinancing ratio is primarily due to increased effectiveness in seeking higher levels of cofinancing for all or most projects. However, there may be variances across GEF periods as to what is reported as cofinancing.

The GEF's cofinancing policy also called for seeking "greater co-financing in upper middle income countries that are not SIDS" (GEF 2014, 8). The promised cofinancing mobilized from the upper middle-income countries for GEF-6 so far is 6.8:1. This is lower than the GEF portfolio average but higher than the portfolio target of 6:1. The cofinancing ratio for the median FSP from this group, 5.7:1, is slightly higher than that for the GEF's portfolio of FSPs (5.6:1), suggesting that the difference in the cofinancing ratio may be due to outliers. The GEF-6 period is yet to be complete, so the ratios may change when the replenishment period ends.

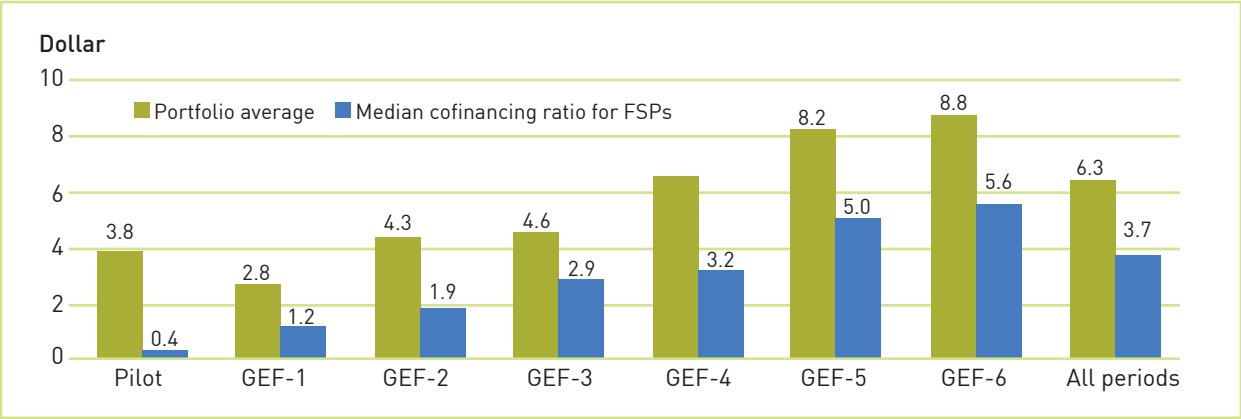
Among the GEF regions, cofinancing ratios for Latin America and the Caribbean are somewhat lower than those for other regions (figure 2.12a). The GEF portfolio in this region is dominated by biodiversity focal area projects, which generally generate lower levels of cofinancing. Across periods, the cofinancing ratio for Europe and Central Asia shows a drop during GEF-6. The high cofinancing ratio achieved by Europe and Central Asia during GEF-5 was driven by two World Bank–implemented climate change projects in Russia and Turkey.[9] The drop is partly due to the Europe and Central Asia cofinancing ratio reverting to its mean and partly due to non-approval of GEF-6 projects for Russia, where GEF projects have traditionally generated higher levels of cofinancing.

The cofinancing ratio for projects that are global in geographic scope has shown substantial increase from GEF-4 to GEF-6 (figure 2.12a). Some of this increase is also driven by the change in type of projects undertaken. During the earlier GEF periods, most global projects involved support for foundational activities. During GEF-6, several global projects, especially those within the framework of integrated approach pilots, have focused on activities that generate a higher level of cofinancing commitments.

Approved projects in countries with large GEF portfolios (China, India, Brazil, and Mexico) mobilized promised cofinancing at 11:1, with a median FSP raising cofinancing at 6.6:1, for the GEF-6 period through June 2017 (figure 2.12b). Both the portfolio cofinancing ratio and the median cofinancing ratio for projects are higher for this

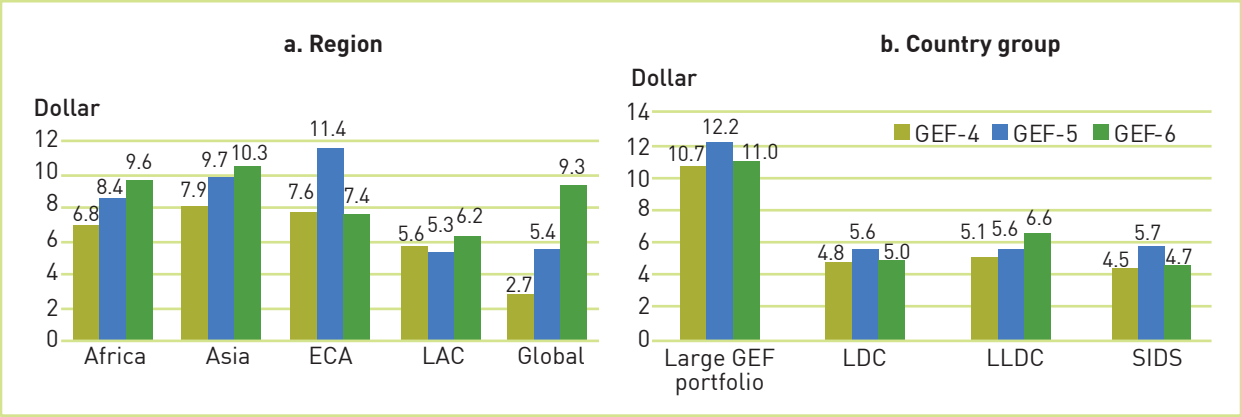
[9] The projects are the Small and Medium Enterprise Energy Efficiency Project (GEF ID 4957), $3.6 million GEF grant and $302 million cofinancing at CEO endorsement; and the Russia Energy Efficiency Financing Project (GEF ID 4427), $23 million GEF grant and $1.249 billion cofinancing.

FIGURE 2.11 Promised cofinancing per dollar of GEF funding for approved GEF projects, by GEF replenishment period



SOURCE: GEF IEO terminal evaluation review data set.

FIGURE 2.12 Promised cofinancing per dollar of GEF funding for approved GEF projects, by region, country group, and GEF replenishment periods



SOURCE: GEF IEO terminal evaluation review data set.

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

group of countries than the GEF portfolio. Thus, cofinancing is being mobilized in the large emerging economies at a higher rate, which is consistent with the higher capacity of these countries to provide cofinancing. The cofinancing ratio for countries with special circumstances, such as LDCs, LLDCs, and SIDS, is lower than the GEF portfolio average.

Multiple linear regression indicates that promised cofinancing is determined by the type of GEF Agency, country context, environmental concern being tackled, size of GEF funding, and year of project approval. Projects that are implemented by the development banks generate higher levels of cofinancing. Controlling for some of the key observable variables, compared to other

Agencies, development banks generate an additional $5.20 of promised cofinancing per dollar of the GEF grant. Projects implemented in large GEF portfolio countries generate an additional $2.60 of promised cofinancing vis-à-vis other countries, whereas SIDS, LDCs, and LLDCs generate lower levels of cofinancing. Projects that tackle climate change and international waters–related concerns generate more cofinancing per dollar of GEF funding, whereas projects that address biodiversity and chemicals generate less cofinancing. Another key determinant is the size of the GEF funding—the greater the GEF funding for a project, the higher the cofinancing ratio. Much of the influence of the size of GEF funding stems from differences in the underlying activities; projects that involve less than $0.5 million are usually enabling activities, whereas those that involve $2.0 million or more are exclusively FSPs. When the analysis is restricted to FSPs, the effect of the size of GEF funding reduces and is not statistically significant. After controlling for other variables, more recent projects generate greater cofinancing ratios than projects that were approved earlier. When the year of approval increases by a year, an additional $0.35 per dollar of GEF funding is generated.

Materialization of the cofinancing commitments during implementation is important, as several critical outputs of a project may depend on it. Therefore, the GEF IEO tracks reported materialization of cofinancing. Of the 1,181 completed projects for which terminal evaluations are available, cofinancing data are available for 991 projects. Despite the lack of information for a significant percentage of completed projects (16 percent), the availability is sufficient to draw some broad inferences.

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

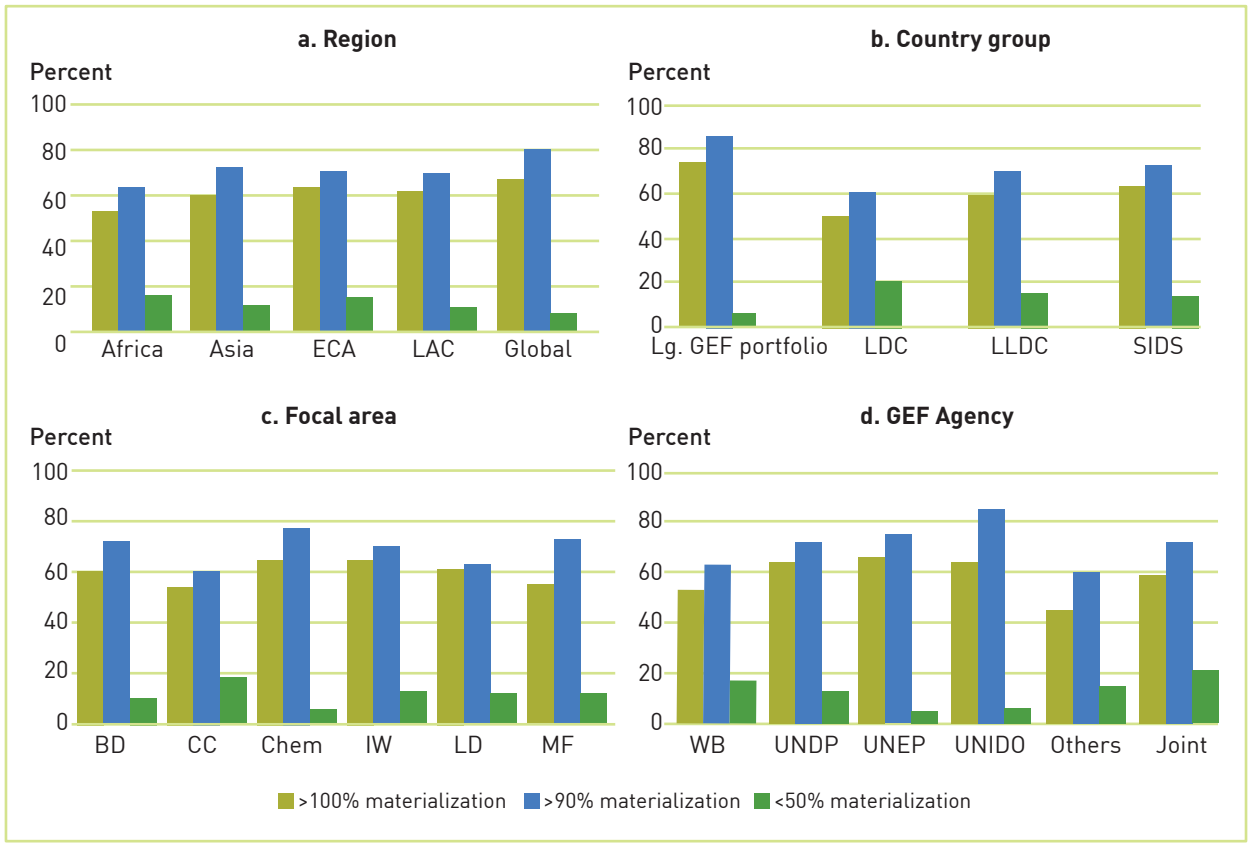
Projects in Africa met cofinancing commitments at a lower rate than other GEF regions (figure 2.13a). Global projects perform better than other projects in meeting at least 90 percent of the cofinancing commitments. However, these correlations weaken and are not statistically significant when other variables are controlled for.

In countries with large GEF portfolios, the cofinancing commitment materialized fully for 72 percent of projects, and at least 90 percent of cofinancing materialized for 83 percent of projects (figure 2.13b). This performance is higher than that of projects in other countries. In LDCs, cofinancing commitments were met fully for 49 percent of projects, and at least 90 percent of cofinancing materialized for 59 percent of projects. This performance is significantly lower than that of projects in other countries. Projects in LLDCs and SIDS met cofinancing commitments at the same rate as other projects in the GEF portfolio.

Although cofinancing materializes fully for a lower percentage of climate change projects, the difference between projects from the climate change focal area and those from other focal areas is not statistically significant (figure 2.13c). Jointly implemented projects are more likely to experience less than 50 percent materialization of expected cofinancing than those implemented by a single Agency (figure 2.13d). A statistically higher percentage of projects implemented by the United Nations Industrial Development Organization achieved at least 90 percent of expected cofinancing than those implemented by other Agencies.

The multiple linear regression models used to analyze causal linkages of variables with the materialization of expected cofinancing do not explain the observed variations well. This indicates

FIGURE 2.13 Materialized cofinancing per dollar of GEF funding for approved GEF projects, by region, country group, focal area, and GEF Agency



SOURCE: GEF IEO terminal evaluation review data set.

NOTE: ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; BD = biodiversity; CC = climate change; IW = international waters; LD = land degradation; MF = multifocal; WB = World Bank; UNIDO = United Nations Industrial Development Organization.

that factors that affect the materialization of cofinancing have not been adequately represented in these models. Nonetheless, two factors stand out. Whether a project is implemented in countries with a large GEF portfolio (positive correlation), and whether it is implemented by a development bank (negative correlation), seem to affect the materialization of cofinancing. Controlling for other observed factors, promised cofinancing is about 15 percent more likely to materialize fully for projects implemented in countries with a large GEF portfolio than for those in other countries.

Further, projects implemented in countries with a large GEF portfolio are also about 10 percent less likely to have less than 50 percent materialization of expected cofinancing. On the other hand, projects implemented by development banks have about a 10 percent lower probability of full materialization. Thus, while projects implemented by development banks generate a substantially higher level of promised cofinancing, this performance is mitigated by higher risks to full materialization of promised cofinancing during implementation.

As part of the OPS6 study, data on promised and materialized cofinancing levels were collected at the level of project cofinancer. Figure 2.14 presents data from 323 projects from the OPS6 cohort for which information is available on cofinancing at the CEO endorsement and the completion stages.[10] The figure shows that some cofinancing sources, such as local government and private businesses, represent a smaller percentage of cofinancing at completion than at endorsement. Figure 2.14 also shows that the share of other sources, such as the federal government, increases from commitments made at CEO endorsement to actual materialization during implementation. The largest share of cofinancing comes from multilateral organizations, which account for 31 percent both of cofinancing promised at endorsement and of materialized cofinancing. A larger share of cofinancing at completion is reported to have been contributed by unspecified or other sources. This is largely due to the fact that reporting of cofinancing by cofinancer type is often less complete at completion than at endorsement.

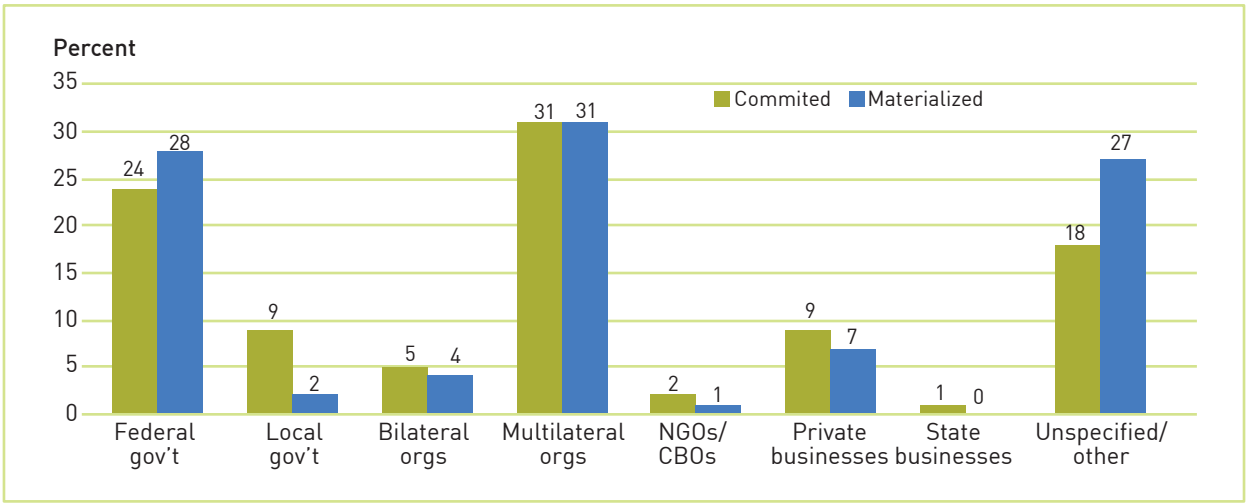
2.2 Quality of terminal evaluations

Terminal evaluations are an essential source of information about the performance of GEF projects. Minimum Requirement 3 of the GEF's M&E policy (GEF IEO 2010) requires that the Agencies prepare a terminal evaluation for each FSP and program at completion. Although Agencies are not required to submit a detailed evaluation for an MSP, it is expected that they submit a summarized report. Quality of terminal evaluations is an important indicator of Agency performance. Therefore, the GEF IEO tracks the quality of the terminal evaluations that are submitted to it.

Of the 1,169 projects rated on the quality of the terminal evaluation, 83 percent rated in the satisfactory range. For the 581 projects received after the close of OPS5, 571 projects are rated, and 82 percent of those rated in the satisfactory range.

[10] A total of 426 projects were reviewed, but due to a lack of information on cofinancing at the level of cofinancer type at either the CEO endorsement or the materialization stage, analysis can be presented for 323 projects.

FIGURE 2.14 **Cofinancing sources as shares of total cofinancing committed/materialized**



SOURCE: GEF IEO terminal evaluation review data set.

This shows that trends in the quality of terminal evaluations are stable.

To assess changes in the quality of terminal evaluations, it is better to track quality by year of terminal evaluation completion than by period of project approval, because new guidance on terminal evaluations may be adopted even for projects that had been approved in past periods. Figure 2.15 shows that the trend in the quality of terminal evaluations for FSPs has been stable and has moved in a narrow band of 80 to 90 percent from 2009 to 2015.[11] Generally, a lower percentage of terminal evaluations for MSPs rated in the satisfactory range than terminal evaluations for FSPs. The percentage rated in the satisfactory range also fluctuates more for MSPs, owing to the smaller number of observations for MSPs than FSPs.

Figure 2.16 compares percentages of Agency projects that rated in the satisfactory range for quality of terminal evaluation. The quality of Agencies' terminal evaluations is in the same ballpark for FSPs. However, UNEP clearly performs better than other Agencies in ensuring the quality of terminal evaluations for MSPs. While the quality of terminal evaluations for UNEP-implemented MSPs is at par with the quality for FSPs, the quality of terminal evaluations for World Bank–implemented FSPs is much higher than that for MSPs. While a lower percentage of terminal evaluations for MSPs implemented by other Agencies is in the satisfactory range, the number of observations is too small to allow strong conclusions.

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

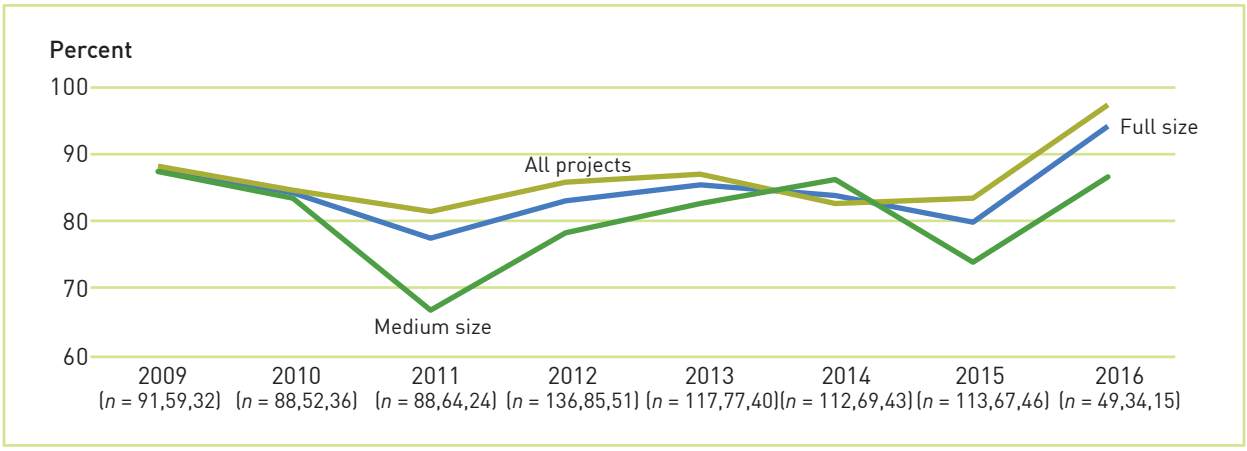
2.3 GEF activity cycle

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

This analysis is focused on stand-alone FSPs, which are endorsed by the CEO based on a two-step process. The first step involves submission by an Agency of a PIF, and it culminates after the PIF is approved. The second stage involves preparation by the GEF Agency of a detailed proposal, submission of the proposal to the GEF Secretariat, and CEO endorsement of the proposal. The GEF has established 18 months as the standard amount of time required for FSPs to move from PIF approval to CEO endorsement (GEF 2010a)—that is, the second step of the process. While it is important to track the efficiency of the activity cycle for projects prepared under the programmatic framework, doing so poses challenges, because child projects need to meet the negotiated commitment deadlines given in their framework documents, and not the 18-month standard applicable to stand-alone FSPs. Similarly, with the advent of the single-step CEO approval process for MSPs, along with the continuation of the two-step CEO approval process, it is difficult to assess the efficiency of the MSP project cycle.

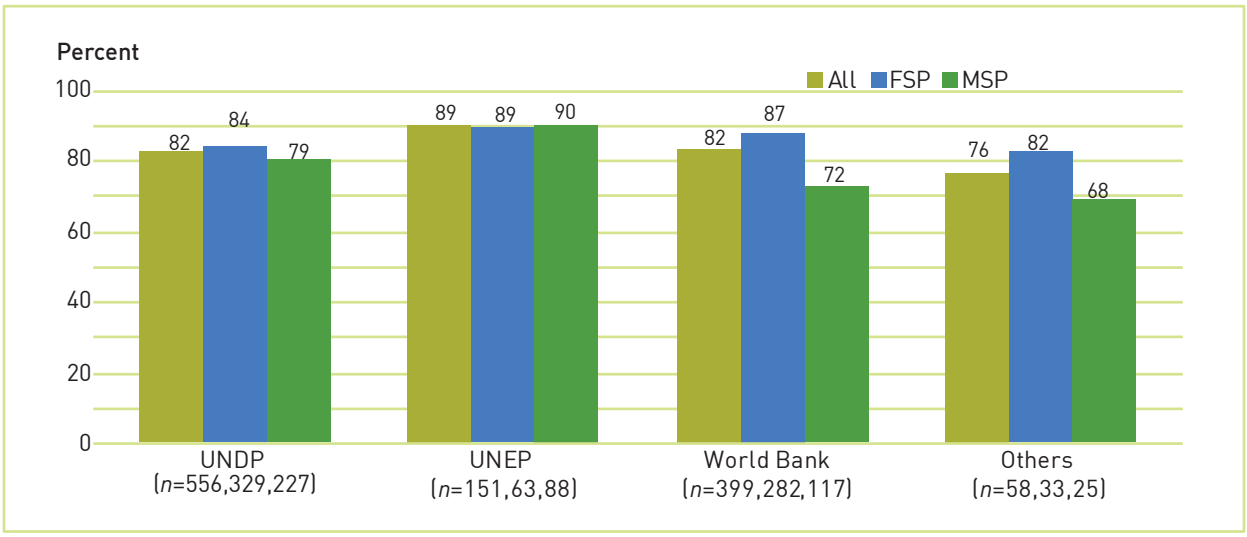
Given that implementation of GEF-6 is in its fourth year, it is possible to track the PIFs submitted during the first year of GEF-6 for at least 24 months. Figure 2.17 compares the performance of GEF-6 with that of GEF-5 and GEF-4. Of the 90

FIGURE 2.15 Percentage of terminal evaluation reports with quality rated in the satisfactory range, by year of terminal evaluation completion and project modality



SOURCE: GEF IEO terminal evaluation review data set.

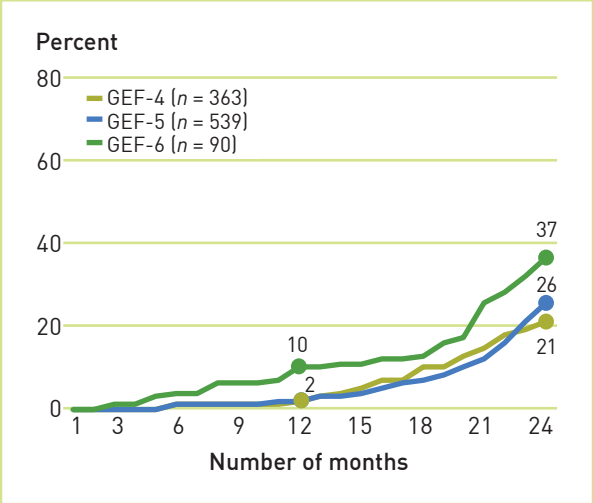
FIGURE 2.16 Percentage of terminal evaluation reports with quality rated in the satisfactory range, by GEF Agency and project modality



SOURCE: GEF IEO terminal evaluation review data set.

PIFs for FSPs submitted during the first year of GEF-6, 37 percent had been CEO endorsed within 24 months of submission. This is a substantial improvement over the performance during GEF-5 (26 percent) and GEF-4 (21 percent), during which a substantially lower percentage of PIF submissions had been CEO endorsed after 24 months. A fuller picture for the GEF-6 proposals will emerge only after GEF-6 has run its course and sufficient time has elapsed to track the progress of the PIFs submitted.

FIGURE 2.17 **Percentage of FSPs endorsed by months between first PIF submission to CEO endorsement**



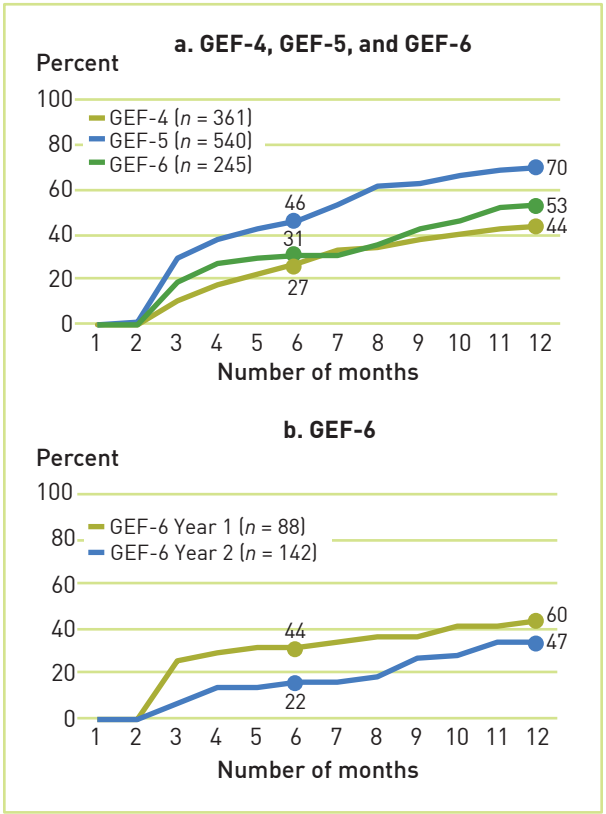
SOURCE: GEF IEO terminal evaluation review data set.

In order to assess performance during the PIF submission to PIF approval stage separately from the PIF approval to CEO endorsement stage, it is important to disaggregate performance between PIF submission and CEO endorsement. Doing this allows us to pinpoint where progress has been made and where there has been little or no progress. Figure 2.18a compares the time taken from PIF submission to PIF approval for GEF-4, GEF-5, and GEF-6. The data show that, at various time thresholds, the percentage of PIF submissions approved was greater during GEF-6 than GEF-4. However, the performance of the GEF-5 projects for this stage of the project cycle was substantially superior to both GEF-4 and GEF-6 period.

A funding shortfall during GEF-6 may have had some influence on the slow progress through the PIF approval stage. This is illustrated by figure 2.18b, which compares the performance of PIF submissions during the first and second years of GEF-6. It shows that the first-year submissions, which were relatively unaffected by the funding shortfall, moved faster through PIF approval. Submissions during the second year, which faced the effects of the funding shortfall (in the third year of GEF-6), achieved PIF approval more slowly. This is in contrast to the performance during GEF-5, in which submissions during the second year moved faster through PIF approval than submissions during the first year, although, like GEF-6, there were more submissions during the second year.

Figure 2.19 tracks the progress of approved PIFs up to 20 months after their approval. The data show that 40 percent of GEF-6 FSP proposals had obtained CEO endorsement within 18 months of PIF approval. This performance is superior to that of GEF-5 proposals. However, performance at this stage lags that of GEF-4 proposals. The CEO

FIGURE 2.18 **Percentage of PIF submissions approved, by time taken in months**

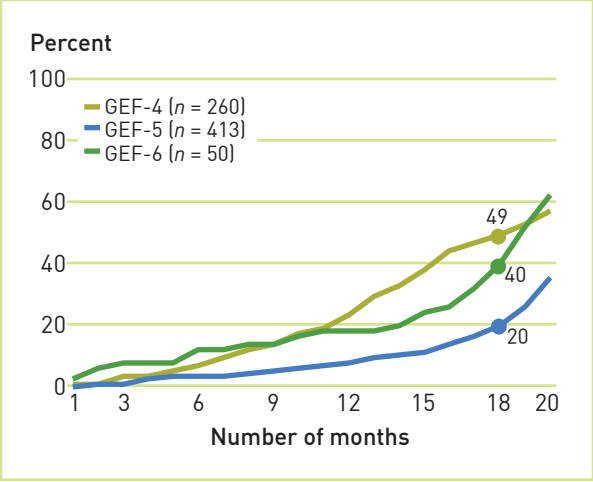


SOURCE: GEF IEO terminal evaluation review data set.

endorsement rate for GEF-6 proposals catches up with that for the GEF-4 proposals by the 19th month and is ahead by the 20th month. Overall, the combined performance for the PIF submission to approval stage, and for the PIF approval to CEO endorsement stage, is superior, because lower performance during the first stage is adequately compensated for through improved performance during the second stage, and also because the combined performance for the GEF-6 period is limited to the submissions during the first year. The assessment of GEF-6 performance is based on a small pool of PIF approvals. Therefore, only after a couple of years past the GEF-6 period will it be possible to assess the progress of the entire cohort of PIFs for FSPs through the CEO endorsement stage.

The time lag in data availability for completed projects makes it difficult to assess the time taken for project start, implementation, and completion for projects approved during the GEF-5 and GEF-6 periods. However, data from completed projects may indicate trends for the preceding periods. The time taken from CEO endorsement to project start suggests an improvement in performance for projects that were approved during the more recent GEF periods (figure 2.20a). This progress is also evident in terms of extensions required for implementation completion. In general, projects approved during the more recent periods are more likely to be completed closer to the closing date expected at project start than projects from the earlier periods (figure 20b). That said, the picture for projects approved during GEF-4 will become clear only after more projects are completed.

FIGURE 2.19 **Percentage of FSPs that obtained CEO endorsement, by time taken in months from PIF approval**



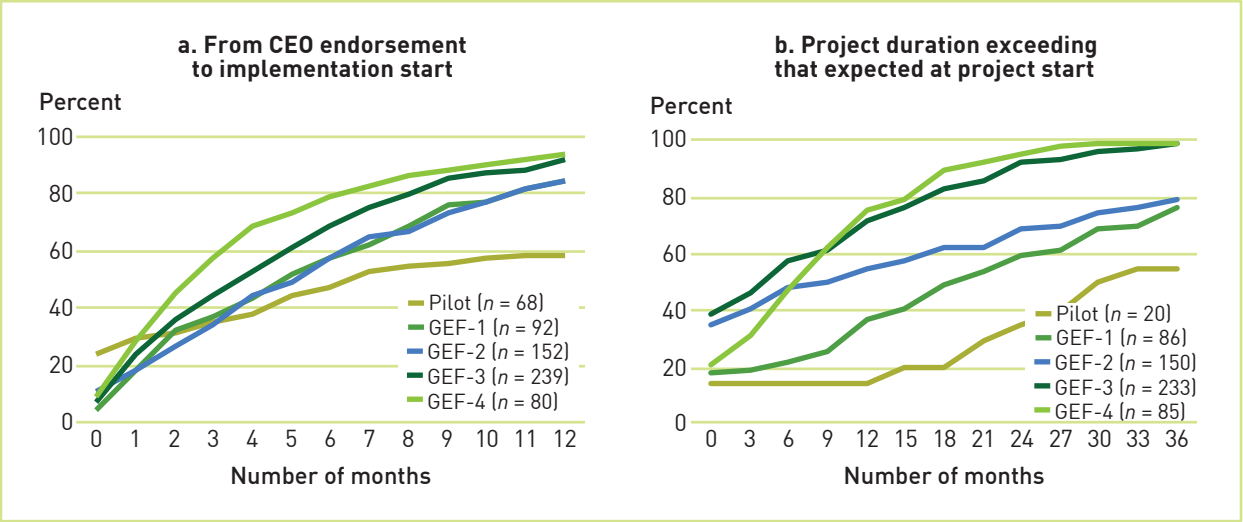
SOURCES: GEF IEO terminal evaluation review data set and GEF PMIS.

2.4 Progress toward GEF-5 and GEF-6 targets

GEF programming for GEF-5 and GEF-6 is consistent with the corporate environmental results targets for these replenishment periods. The results promised in documents of projects approved in the GEF-5 period were reviewed. The data from CEO endorsement/approval documents show that the GEF is on track to meet most of the GEF-5 replenishment's environmental results targets. Given that a year remains before the GEF-6 period is complete, and that there is about a 15 percent shortfall in actual replenishment versus expected replenishment, the progress for the GEF-6 period is reasonable.

To inform the GEF-6 replenishment process, the GEF IEO analyzed the progress to environmental results targets for the GEF-5 replenishment period. The first paper (GEF IEO 2013b) was presented to the third meeting of the GEF-6 replenishment, and an update of the paper (GEF 2013c) was presented to the fourth meeting. The former accounted for GEF-5 PIF approvals through June 2013, and the latter updated it based on GEF-5 PIF approvals through December 2013. In preparation for the OPS6, the GEF IEO undertook

FIGURE 2.20 Status of FSPs after CEO endorsement, by time taken in months (%)



SOURCES: GEF IEO terminal evaluation review data set and GEF PMIS.

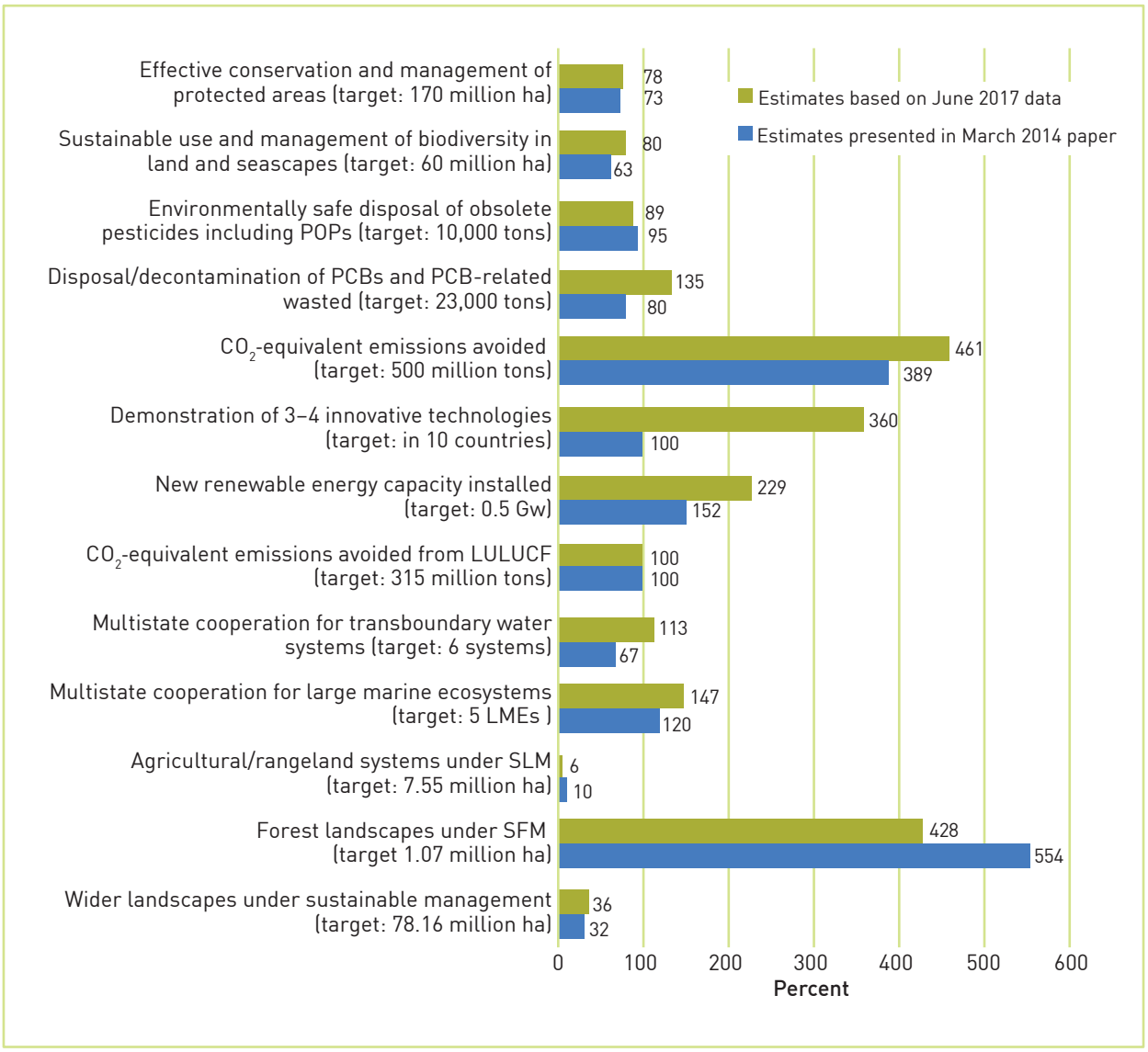
an assessment to update its analysis of the GEF-5 period.

The analysis of the GEF-5 period is based on data for 686 projects funded partially or fully through GEF Trust Fund resources.[12] The analysis is still primarily based on the aggregation of expected targets for the approved GEF-5 projects. However, unlike the analysis presented in the December 2013 and March 2014 papers, which were primarily based on the aggregated targets listed in the approved PIFs, the analyses conducted for OPS6 are based on the expected results indicated in the CEO endorsement/approval documents for 96 percent of projects. For the remainder, the targets at PIF approval have been used. The aggregated environmental results targets have been multiplied by 0.8 to account for the likelihood of cancellations and implementation failure.

Figure 2.21 presents the GEF IEO's projections of the expected environmental results as a percentage of the targets committed to in the programming directions for the full GEF-5 cohort and compares it with the progress estimates presented in the March 2014 paper. Of the 13 environmental indicators that could be tracked, and after adjustments for cancellations and implementation failures, the GEF is on course to achieve or exceed its expected targets for 8 indicators. Level of achievement is likely to be slightly lower than the target for three indicators, of which two pertain to chemicals and one to biodiversity conservation. Of the three indicators that are relevant to the land degradation focal area, targets are unlikely to be met for two. Compared to the progress reported in the March 2014 paper, the estimate for the GEF-5 period, which was prepared based on June 2017 data, shows increased expectations for 9 of the 13 indicators. For the

[12] As of December 31, 2015, 973 projects were partially or fully funded through GEF Trust Fund resources under GEF-5. A total of 258 enabling activity projects, 12 Small Grants Programme projects, 3 targeted research projects, and 17 canceled projects are excluded from the analysis.

FIGURE 2.21: **Expected adjusted environmental results for GEF-5 projects as a percentage of GEF-5 targets**

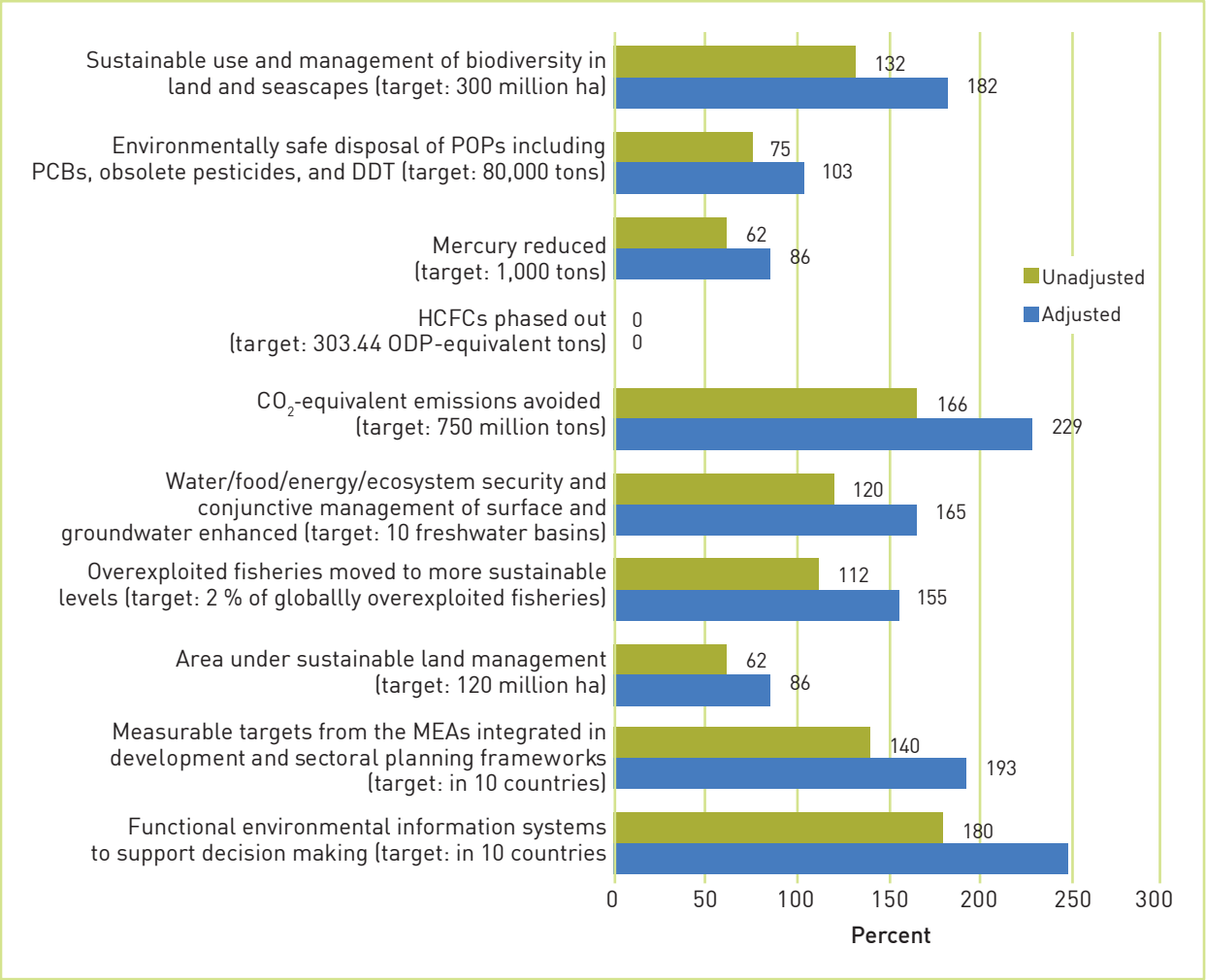


SOURCES: GEF IEO terminal evaluation review data set and GEF PMIS.

remaining 4 indicators, there has been a decrease in expected benefits, as the detailed proposals of some projects submitted for CEO endorsement downscaled the level of expected results. More details are provided in annex C.

During GEF-6, the Secretariat mainstreamed recording of the projects' expected environmental results in the PMIS. It is now reporting on the aggregated targets provided in the approved PIFs for the GEF-6 period through the GEF Corporate Scorecard (GEF 2017a). Two adjustments have been made to the GEF-6 figures provided in the scorecard. The expected results have been multiplied by a factor of 0.8, to account for cancellations and implementation failures, and by 1.72,

FIGURE 2.22 **Unadjusted and adjusted expected environmental results as percentage of GEF-6 targets**



SOURCES: GEF IEO terminal evaluation review data set and GEF PMIS.

NOTE: Unadjusted results are based on PIF approvals received to May 2017; adjusted results account for cancellations and implementation failure and for resources programmed.

to account for the level of programmed resources vis-à-vis the GEF-6 replenishment expected at the start of GEF-6. For GEF-6, the aggregated results from approved PIFs exceed GEF-6 targets for 6 of 10 environmental results indicators (figure 2.22). The only indicator for which there was no uptake relates to the phaseout of ozone-depleting substances, where GEF involvement has been declining. When the level of fund utilization and the likely cancellations and implementation failure rate are accounted for, the adjusted expected results are commensurate with funding for 7 of the 10 indicators. The detailed proposals for most of the approved GEF-6 projects are still under preparation. Further, it is likely that more projects will be approved during the last year of GEF-6 (2017–18). Once the period is complete, it will be possible to make more reliable projections.

2.5 Progress toward impact

It is often too early to assess the long-term impacts of a project at the point of its implementation completion. Many environmental results take more than a decade to manifest. Similarly, achievement of some of the environmental results of GEF projects may also be contingent on future actions by other actors. Therefore, any assessment of the impacts of GEF projects at project completion is likely to underestimate the number of projects with impacts as well as the likely scale. Nonetheless, reviewing progress toward impact at project completion helps determine what has already been achieved and the extent to which long-term results are likely. Of the 584 terminal evaluations that were submitted to the IEO after the close of OPS5, 415 were reviewed to determine the extent to which the projects had achieved environmental stress reduction and/or status change and whether broader adoption of the promoted approaches, initiatives, or technologies by other stakeholders was taking place, and through what mechanisms.[13] The probability-adjusted figures for the OPS6 cohort are presented in annex B. In this section, unadjusted figures have been presented because there is little material difference in the unadjusted and adjusted figures.

ENVIRONMENTAL STRESS REDUCTION AND STATUS CHANGE

Environmental stress reduction may be understood as biophysical changes that reflect the reduction of threats emanating from human actions. Fifty-nine percent of the GEF projects achieved stress reduction and/or environmental status change at project completion (table 2.1). Achievement of environmental stress reduction and/or environmental change appears to be linked with the environmental challenge being addressed, country context, global versus regional focus, and scale of GEF funding. Thirteen percent of the projects were achieving environmental stress reduction and/or status change at a large scale—that is, targeted at the system or national level—and 45 percent of projects were achieving it at a local scale. Forty-one percent of the projects either had not achieved any environmental stress reduction and/or environmental status change yet or it was not possible to assess whether this had taken place.

Projects' ability to achieve environmental stress reduction at implementation completion is affected by the environmental concern they tackle. For example, 80 percent of projects that focus on chemicals and waste, and 69 percent of those that focus on climate change, achieve stress reduction by implementation completion (table 2.1). In comparison, only 35 percent of projects that address international waters–related concerns achieve stress reduction. This result is not surprising, as most of the GEF projects that address international waters focus more on strengthening intergovernmental arrangements to address these issues, and there is a time lag before these efforts lead to actual stress reduction and/or environmental status change on the ground. Country circumstances also play a role, as stress reduction and/or environmental status change was achieved in 73 percent of the projects implemented in the five countries with large GEF portfolios but in only 52 percent of projects implemented in SIDS.

Compared to projects that are implemented in countries or that are regional in focus, global projects seem to be less likely to achieve environmental stress reduction and/or status change. Only 21 percent of the global projects, compared

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

TABLE 2.1 Percentage of projects achieving environmental stress reduction, by focal area and scale

Focal area	Stress reduction taking place at...			No evidence or unable to assess
	Large scale	Local scale	Any scale	
Biodiversity (*n* = 147)	10	41	52	48
Climate change (*n* = 122)	20	49	70	30
Chemicals and waste (*n* = 25)	16	64	80	20
International waters (*n* = 38)	11	24	35	66
Land degradation (*n* = 35)	11	63	74	26
Multifocal area (*n* = 48)	6	42	48	52
All focal areas (*n* = 415)	13	45	59	41

SOURCE: GEF IEO terminal evaluation review data set.

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

BROADER ADOPTION AND TRANSFORMATIONAL CHANGE

The majority (61 percent) of GEF projects achieved broader adoption at project completion. Country context plays an important role, as projects implemented in major emerging economies are more likely to achieve broader adoption at higher scales than projects in other countries. Broader adoption is said to take place when governments and other stakeholders adopt, expand, and build on the initiatives that the GEF promotes during program/project implementation or afterward. The GEF IEO's past work shows that broader adoption tends to occur through five mechanisms: sustaining, mainstreaming, replication, scaling-up, and market change. Broader adoption may take place through one or more of these mechanisms that may operate simultaneously or sequentially. Broader adoption facilitates transformational change in the systems that the GEF targets. As outlined in the GEF's 2020 Strategy (GEF 2015b), support leading to transformational change is one of the GEF's strategic priorities.

Data from the review of terminal evaluations show that 24 percent of completed projects achieved broader adoption at a large scale (table 2.2). Among those that achieved broader adoption at a large scale, most of the promoted approaches, initiatives, and/or technologies were being adopted for a third, whereas for the remaining two-thirds, only some of these were being adopted. Thirty-seven percent of the projects achieved broader adoption at a local scale. For 26 percent of the projects, although broader adoption was not yet taking place, plans were in place to facilitate this in the future. For only 13 percent of the projects, either no progress was reported in terms of broader adoption or it was difficult to

TABLE 2.2 Incidence of broader adoption at project completion, by scale

Broader adoption status and scale	Incidence (%)
Taking place (*n* = 252)	61
At large scale (*n* = 100)	24
At local scale (*n* = 152)	37
Not taking place (*n* = 163)	39
But some progress (*n* = 108)	26
No broader adoption or unable to assess (*n* = 55)	13
Total (*n* = 415)	100

SOURCE: GEF IEO terminal evaluation review data set.

TABLE 2.3 Percentage of projects using broader adoption mechanisms, by type of mechanism and project modality

Mechanism	FSPs (*n* = 268)	MSPs (*n* = 147)	Total (*n* = 415)
Sustaining	27	22	25
Mainstreaming	36	41	38
Replication	24	22	23
Scaling-up	12	8	11
Market change	9	7	8

SOURCE: GEF IEO terminal evaluation review data set.

ascertain their adoption status. Differences across focal areas in terms of likelihood of projects achieving broader adoption at completion are not as apparent as they were for environmental-stress reduction and/or status change.

A higher percentage of projects implemented in the countries with large GEF portfolios (73 percent) achieved broader adoption at the point of completion than projects in other countries (59 percent).[14] Projects that replicated an approach that had been piloted elsewhere (75 percent versus 58 percent) and projects that followed up on a preceding GEF project (75 percent versus 59 percent) also achieved broader adoption at completion.

[14] This difference is significant at a 90 percent confidence level but not at a 95 percent level.

2.6 Mechanisms of broader adoption

Of the mechanisms for broader adoption, mainstreaming (38 percent of projects), sustaining progress (25 percent), and replication (23 percent) were observed more frequently than scaling-up (11 percent of projects) and market change (8 percent) (table 2.3). Although broader adoption took place for an optically higher percentage of FSPs (63 percent) compared to MSPs (56 percent), the difference is not statistically significant. A slightly higher number of MSPs reported mainstreaming than FSPs, but for all other mechanisms, the opposite was true. Differences between FSPs and MSPs are not statistically significant for any of the mechanisms. Projects can result in broader adoption through more than one mechanism.

Thirty-four percent of projects achieved broader adoption through a single mechanism. Seventeen percent of projects achieved it through two mechanisms. Ten percent achieved broader adoption through more than two mechanisms.

Annex A: Rating criteria and scale

A.1 Outcomes

In the causal pathways of a project, its outputs are expected to lead to its intended outcomes.[1] Although achievement of outcomes is not certain, most GEF projects may be expected to achieve the targeted outcomes at implementation completion. The evaluators should, therefore, assess the extent to which the expected outcomes were achieved and how much its achievement depended on the delivery of project outputs. They should also assess the factors that affected outcome achievement—for example, project design, a project's linkages with other activities, the extent and materialization of cofinancing, stakeholder involvement, and so forth. Where the project was developed within the framework of a program, the assessment should also report on how much the project contributed to the program outcomes.

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

- **Relevance:** Were the project outcomes congruent with the GEF focal areas/operational program strategies, country priorities, and mandates of the Agencies? Was the project design appropriate for delivering the expected outcomes?
- **Effectiveness:** To what extent were the project's actual outcomes commensurate with the expected outcomes?
- **Efficiency:** Was the project cost-effective? How did its cost/time versus output/outcomes equation compare to that of similar projects?

Rating scale for outcomes: The following six-point rating scale is used to assess overall outcomes:

- **Highly satisfactory (HS):** Level of outcomes achieved clearly exceeded expectations, and/or there were no shortcomings.
- **Satisfactory (S):** Level of outcomes achieved was as expected, and/or there were no or minor shortcomings.
- **Moderately satisfactory (MS):** Level of outcomes achieved was more or less as expected, and/or there were moderate shortcomings
- **Moderately unsatisfactory (MU):** Level of outcomes achieved was somewhat lower

[1] Outcomes are "the likely or achieved short-term and medium-term effects of an intervention's outputs. Outputs are the products, capital goods, and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes" (OECD 2002).

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

than expected, and/or there were significant shortcomings.

- **Unsatisfactory (U):** Level of outcomes achieved was substantially lower than expected, and/or there were major shortcomings.
- **Highly unsatisfactory (HU):** Only a negligible level of outcomes was achieved, and/or there were severe shortcomings.
- **Unable to assess (UA):** The available information did not allow an assessment of the level of outcome achievements.

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

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

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

A.2 Sustainability

Minimum Requirement 3 of the GEF's 2010 M&E policy specifies that a terminal evaluation will assess the likelihood of sustainability of outcomes at project termination and provide a rating.[3] The assessment of sustainability will weigh risks to continuation of benefits from the project. The assessment should identify key risks and explain how these risks may affect continuation of benefits after the GEF project ends. The analysis should cover financial, sociopolitical, institutional, and environmental risks.

The overall sustainability of project outcomes will be assessed based on the likelihood and magnitude of the effect of risks to sustainability. Higher levels of risks and magnitudes of effect imply lower likelihood of sustainability. The sustainability will be assessed taking into account the risks related to financial, sociopolitical, institutional, and environmental sustainability of project outcomes. The evaluator may also take other risks into account that may affect sustainability. The overall sustainability will be assessed using the following four-point scale:

- **Likely (L):** There are little or no risks to sustainability.
- **Moderately likely (ML):** There are moderate risks to sustainability.

[3] The GEF Monitoring and Evaluation Policy 2010 adopts the following definition of sustainability: "the likely ability of an intervention to continue to deliver benefits for an extended period of time after completion; projects need to be environmentally as well as financially and socially sustainable" (GEF IEO 2010, 27).

- **Moderately unlikely (MU):** There are significant risks to sustainability.
- **Unlikely (U):** There are severe risks to sustainability.
- **Unable to assess (UA):** The available information did not allow an assessment of the expected incidence and magnitude of risks to sustainability.

A.3 Project monitoring and evaluation

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

M&E design. To assess the quality of the M&E plan, the evaluators will assess the following: At the point of CEO endorsement, was the M&E plan practical and sufficient? Did it include baseline data? Did it specify clear targets and appropriate (SMART) indicators to track environmental, gender, and socioeconomic results?[4] Did it include a proper methodological approach? Did it have a practical organization and include logistics for the M&E activities, including a schedule and responsibilities for data collection? Did it budget adequate funds for M&E activities?

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

Quality of M&E on these two dimensions will be assessed on the following six-point scale:

- **Highly satisfactory (HS):** There were no shortcomings, and the quality of M&E design/implementation exceeded expectations.
- **Satisfactory (S):** There were no or minor shortcomings, and the quality of M&E design/implementation met expectations.
- **Moderately satisfactory (MS):** There were some shortcomings, and the quality of M&E design/implementation more or less met expectations.
- **Moderately unsatisfactory (MU):** There were significant shortcomings, and the quality of M&E design/implementation was somewhat lower than expected.
- **Unsatisfactory (U):** There were major shortcomings, and the quality of M&E design/implementation was substantially lower than expected
- **Highly unsatisfactory (HU):** There were severe shortcomings in M&E design/implementation.
- **Unable to assess (UA):** The available information did not allow an assessment of the quality of M&E design/implementation.

A.4 Implementation

The assessment of the implementation of GEF FSPs will take into account the performance of the GEF Agencies in discharging their expected

[4] SMART: Specific, measurable, achievable/attributable, relevant/realistic, and time-bound, timely, trackable, and targeted.

roles and responsibilities. The performance of the Agencies will be rated using a six-point scale (Highly Satisfactory to Highly Unsatisfactory). Within the GEF partnership, GEF Agencies are involved in activities related to a project's identification, concept preparation, appraisal, preparation of detailed proposal, approval and startup, oversight, supervision, completion, and evaluation.[5] To assess performance of the GEF Agencies, the evaluators will assess the extent to which the Agency delivered effectively on these counts, with focus on elements that were controllable from the given GEF Agency's perspective. The evaluator will assess how well risks were identified and managed by the GEF Agency.

The quality of implementation will be rated on the following six-point scale.

- **Highly satisfactory (HS):** There were no shortcomings, and quality of implementation exceeded expectations.
- **Satisfactory (S):** There were no or minor shortcomings, and quality of implementation met expectations.
- **Moderately satisfactory (MS):** There were some shortcomings, and quality of implementation more or less met expectations.
- **Moderately unsatisfactory (MU):** There were significant shortcomings, and quality of implementation was somewhat lower than expected.
- **Unsatisfactory (U):** There were major shortcomings, and quality of implementation was substantially lower than expected.
- **Highly unsatisfactory (HU):** There were severe shortcomings in quality of implementation.
- **Unable to assess (UA):** The available information did not allow an assessment of the quality of implementation/execution.

[5] See GEF (2010b) and (2011).

Annex B: Progress to impact results corrected for sampling probability differences

TABLE B.1 Percentage of projects achieving environmental stress reduction, by scale

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

SOURCE: GEF IEO terminal evaluation review data set.

TABLE B.2 Incidence of broader adoption, by scale

Broader adoption status and scale	Incidence (%)	
	Not adjusted	Probability adjusted
Broader adoption taking place (*n* = 252)	61	61
At large scale (*n* = 100)	24	24
At local scale (*n* = 152)	37	37
Broader adoption not taking place (*n* = 163)	39	39
But some progress (*n* = 108)	26	25
No broader adoption or unable to assess (*n* = 55)	13	14
Total (*n* = 415)	100	100

SOURCE: GEF IEO terminal evaluation review data set.

TABLE B.3 Percentage of projects using broader adoption mechanisms, by type of mechanism and project modality

Mechanism	FSPs (*n* = 268)	MSPs (*n* = 147)	Total (*n* = 415)	FSPs (*n* = 268)	MSPs (*n* = 147)	Total (*n* = 415)
	Not adjusted			Probability adjusted		
Sustaining	27	22	25	26	22	24
Mainstreaming	36	41	38	37	40	38
Replication	24	22	23	24	23	24
Scaling-up	12	8	11	13	7	11
Market change	9	7	8	8	6	7

SOURCE: GEF IEO terminal evaluation review data set.

Annex C: Progress toward targets

TABLE C.1 Progress toward GEF-5 targets

Focal area	Strategic goal	Targets	Project-level targets achieved to date[a]	% of target contained in project-level targets to date	Estimated % of target to be achieved from projects to date[b]
Biodiversity	Improved sustainability of protected area systems	Effective conservation and management of 170 million ha of protected areas	63.33 million ha of new protected areas; 101.45 million ha of existing protected areas	97%[c]	78%
	Sustainably managed landscapes and seascapes that integrate biodiversity conservation increased	Sustainable use and management of biodiversity in 60 million ha of production landscapes and seascapes	55.31 million ha of production landscapes; 4.87 million ha of production seascapes	100%	80%
Climate change	Slowed growth in greenhouse gas emissions to the atmosphere from demonstration and transfer of advanced low-carbon technologies and deployment and diffusion of technologies in energy efficiency, renewable energy, and sustainable transport and urban systems	500 Mt of CO_2-equivalent emissions avoided	459 Mt CO_2-equivalent direct mitigation; 2,414 Mt CO_2-equivalent indirect mitigation	94% if only direct included; 577% if including indirect	75% if only direct included; 461% if including indirect
		Demonstration of 3–4 innovative technologies in 10–15 countries	Demonstration of 16 innovative technologies in 24 countries	450% of country target[d]	360% of country target
		0.5 Gw of new renewable energy capacity installed	1.42 Gw of new renewable energy capacity installed	286%	229%
		315–675 Mt of CO_2-equivalent emissions avoided from LULUCF	549 Mt CO_2-equivalent emission reductions[e]	100%	100%
International waters	Catalyze multistate cooperation to balance conflicting water uses in transboundary surface and groundwater basins while considering climatic variability and change	Multistate cooperation results in adoption/implementation of national/local reforms in 50% of states and demonstration results in at least 50% of states participating in 6–7 transboundary water systems	10 transboundary water systems targeted through 10 projects involving 48 different countries	142% of measurable target (# of transboundary water systems targeted)	113% of measurable target

Focal area	Strategic goal	Targets	Project-level targets achieved to date[a]	% of target contained in project-level targets to date	Estimated % of target to be achieved from projects to date[b]
International waters	Catalyze multistate cooperation to rebuild marine fisheries and reduce pollution of coasts and large marine ecosystems (LMEs) while considering climatic variability and change	Multistate cooperation results in adoption/ implementation of national/local reforms in 50% of states and demonstration results in at least 50% of states participating in 5–6 LMEs	11 LMEs targeted through 15 projects involving 66 countries	183% of measurable target (# of LMEs targeted)	147% of measurable target
Land degradation	Arrested or reversed current global trends in land degradation, specifically desertification and deforestation	Sustainable management of agriculture, range, and forest landscapes, including drylands and affected transboundary areas: 100 million ha in agriculture; 200,000 ha of forest landscapes; 175 million ha in wider production landscapes	7.55 million ha of agricultural/ rangeland systems under sustainable land management; 1.07 million ha of forest landscapes under sustainable forest management; 78.16 million ha of wider production landscapes under sustainable management	8% 535% 45%	6% 428% 36%
Chemicals	Phased-out and reduced releases of persistent organic pollutants (POPs), ozone-depleting substances, and other chemicals of global concern	10,000 tons of obsolete pesticides, including POPs, disposed of in an environmentally sound manner	11,146 tons of obsolete pesticides, including POPs, disposed of in an environmentally sound manner	111%	89%
		23,000 tons of PCBs and PCB-related wastes disposed of or decontaminated	33,560 tons of PCBs and PCB-related wastes disposed of or decontaminated	169%	135%

SOURCE: GEF Secretariat 2010.

NOTE: Summation of project-level (excluding 17 canceled projects) targets to date, and estimated percentage of replenishment targets that will be achieved from GEF-5 projects (December 31, 2015).

a. Excludes 17 canceled projects.

b. Assumes 80 percent of project-level targets will be achieved.

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

d. The aggregate project-level target is derived from projects with CCM-1 (technology transfer) funding.

e. Total project-level reductions from land use, land use change, and forestry (LULUCF) include both direct and indirect reductions.

TABLE C.2 Progress toward GEF-6 targets

Focal area	Strategic goal	Programmatic targets	Estimated % of projects approved to date[a]
Biodiversity	Maintain globally significant biodiversity and the ecosystem goods and services it provides to society	300 million ha of landscapes and seascapes under improved biodiversity management	132
Climate change	To support developing countries and economies in transition to make transformational shifts toward a low-emission, resilient development path	750 Mt CO_2-equivalent emissions avoided, both direct and indirect, over the investment or impact period of the projects	166
International waters	Promotion of collective management of transboundary water systems and implementation of the full range of policy, legal, and institutional reforms and investments contributing to sustainable use and maintenance of ecosystem services	Water/food/energy/ecosystems security and conjunctive management of surface and groundwater enhanced in at least 10 freshwater basins	120
		20% of globally overexploited fisheries (by volume) moved to more sustainable levels	112
Land degradation	To contribute to arresting and reversing current global trends in land degradation, specifically desertification and deforestation	120 million ha under sustainable land management	62
	Enhance capacity of countries to implement multilateral environmental agreements (MEAs)	Development and sectoral planning frameworks integrate measurable targets drawn from the MEAs in at least 10 countries	140
		Functional environmental information systems are established to support decision making in at least 10 countries	180
Chemicals	Promote the sound management of chemicals throughout their life cycle to minimize adverse effects on the global environment and health of both women and men	80,000 tons of persistent organic pollutants, including PCBs, obsolete pesticides, and DDT, disposed of in an environmentally sound manner	75
		1,000 tons of mercury reduced	62
		303.44 ozone-depleting potential tons of HCFCs phased out	0

SOURCE: GEF 2017a.

NOTE: Based on 299 projects at the PIF approval stage in GEF-6, 124 of which were CEO endorsed/approved by April 30, 2017.

Annex D: Regression models

TABLE D.1 **Outcomes**

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Quality of implementation	0.467***	0.467***	0.466***	0.466***	0.461***	0.463***
Quality of execution	0.305***	0.304***	0.302***	0.305***	0.279***	0.291***
Large GEF portfolios	0.00592					
LDC		−0.00840				
SIDS			−0.0599			
Africa				−0.00757		
<50% of cofinancing materialized					−0.0884***	
Cofinancing fully realized						0.0411**
Control variables						
Focal area	Yes	Yes	Yes	Yes	Yes	Yes
GEF grant	Yes	Yes	Yes	Yes	Yes	Yes
Year of implementation start	Yes	Yes	Yes	Yes	Yes	Yes
Project preparation grant given	Yes	Yes	Yes	Yes	Yes	Yes
Observations	943	943	943	943	828	828
R-squared	0.491	0.491	0.492	0.491	0.486	0.483

TABLE D.2 Sustainability

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Quality of implementation	0.180***	0.183***	0.189***	0.177***	0.163***	0.163***
Quality of execution	0.254***	0.250***	0.265***	0.256***	0.248***	0.256***
Large GEF portfolios	0.230***					
LDC		−0.147***				
SIDS			0.0547			
Africa				−0.111***		
<50% of cofinancing materialized					−0.0872*	
Cofinancing fully realized						0.0548
Climate change	0.0784**	0.0857**	0.0823**	0.0783**	0.0972***	0.0956**
Control variables						
GEF grant	Yes	Yes	Yes	Yes	Yes	Yes
Year of project	Yes	Yes	Yes	Yes	Yes	Yes
Project preparation grant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	919	919	919	919	806	806
R-squared	0.146	0.135	0.125	0.134	0.113	0.113

TABLE D.3 Implementation

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Large GEF portfolio	0.0803*							
LDC		−0.129***						
SIDS			−0.0999*					
Africa				−0.131***				
>50% cofinancing materialized					−0.233***			
Cofinancing fully realized						0.110***		
World Bank							−0.0751**	
UNEP								0.112***
Control variables								
Focal area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GEF grant	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Year of project start	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Project preparation grant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	970	970	970	970	846	846	970	970
R-squared	0.037	0.045	0.037	0.052	0.069	0.049	0.039	0.041

NOTE: *** p < 0.01, ** p < 0.05, * p < 0.1.

TABLE D.4 M&E design

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Large GEF portfolios	0.0688					
LDC		−0.123***				
SIDS			0.0595			
Africa				−0.118***		
UNDP					0.0665**	
World Bank						−0.0917***
Chemicals	−0.125*	−0.138**	−0.118*	−0.144**	−0.109*	−0.137**
Control variables						
GEF grant	Yes	Yes	Yes	Yes	Yes	Yes
Year of project start	Yes	Yes	Yes	Yes	Yes	Yes
Project preparation grant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,108	1,108	1,108	1,108	1,108	1,108
R-squared	0.044	0.049	0.043	0.053	0.046	0.049

NOTE: *** p < 0.01, ** p < 0.05, * p < 0.1.

TABLE D.5 M&E implementation

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Large GEF portfolios	0.0995**					
LDC		−0.176***				
SIDS			−0.0929			
Africa				−0.156***		
World Bank					−0.0916***	
Other Agencies						−0.132*
Multifocal	−0.211***	−0.213***	−0.210***	−0.201***	−0.204***	−0.203***
Control variables						
GEF grant	Yes	Yes	Yes	Yes	Yes	Yes
Year of project start	Yes	Yes	Yes	Yes	Yes	Yes
Project preparation grant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,012	1,012	1,012	1,012	1,012	1,012
R-squared	0.028	0.040	0.026	0.044	0.031	0.027

NOTE: *** p < 0.01, ** p < 0.05, * p < 0.1.

References

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

——. 2010a. "Streamlining the Project Cycle and Refining the Programmatic Approach." GEF/C.38/5/Rev.1. GEF, Washington, DC.

——. 2010b. "Rules and Guidelines for Agency Fees and Project Management Costs." GEF/C.39/9. GEF, Washington, DC.

——. 2011. "GEF Minimum Fiduciary Standards: Separation of Implementation and Execution Functions in GEF Partner Agencies." GEF/C.41/06/Rev.01. GEF, Washington, DC.

——. 2014. "Co-financing Policy." GEF/C.46/09. GEF, Washington, DC.

——. 2015a. *Instrument for the Establishment of the Restructured Global Environment Facility*. Washington, DC: GEF.

——. 2015b. *GEF2020: Strategy for the GEF*. Washington, DC: GEF.

——. 2017a. "GEF Corporate Scorecard, April 2017." GEF, Washington, DC.

——. 2017b. "Guidelines on the Project and Program Cycle Policy." GEF/C.52/Inf.06. GEF, Washington, DC.

GEF IEO (Global Environment Facility Independent Evaluation Office; formerly Global Environment Facility Evaluation Office). 2010. *The GEF Monitoring and Evaluation Policy 2010*. Evaluation Document No. 4. Washington, DC: GEF IEO.

——. 2013a. "Co-financing." OPS5 Technical Document No. 21. GEF IEO, Washington, DC.

——. 2013b. "Verification Note on GEF Secretariat Data on the Realization of GEF-5 Targets." GEF/R.6/Inf.09. GEF, Washington, DC.

——. 2013c. "Updated Verification Note on the Realization of GEF-5 Targets." GEF/R.6/Inf.13. GEF, Washington, DC.

——. 2014a. *GEF Annual Performance Report 2013*. Evaluation Report No. 92. Washington, DC: GEF IEO.

——. 2014b. *OPS5 Final Report: At the Crossroads for Higher Impact*. Evaluation Report No. 86. Washington, DC: GEF IEO.

——. 2015. *GEF Annual Performance Report 2014*. Evaluation Report No. 102. Washington, DC: GEF IEO.

——. 2016. "Evaluation of the Expansion of the GEF Partnership: First Phase." GEF/ME/C.50/06. GEF IEO, Washington, DC.

GEF Secretariat (Global Environment Facility Secretariat). 2010. "GEF-5 Programming Document." GEF/R.5/31. GEF, Washington, DC.

Kotchen, Matthew J., and Neeraj Kumar Negi. 2016. "Cofinancing in Environment and Development: Evidence from the Global Environment Facility." *The World Bank Economic Review*.

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

Recent GEF IEO Publications

Evaluation Reports

120	Land Degradation Focal Area Study	2018
119	Evaluation of GEF Engagement with Indigenous Peoples	2018
118	Evaluation of Gender Mainstreaming in the GEF	2018
117	Program Evaluation of the Special Climate Change Fund	2018
116	Review of the GEF Policy on Agency Minimum Standards on Environmental and Social Safeguards	2018
115	Chemicals and Waste Focal Area Study	2018
114	International Waters Focal Area Study	2018
113	Evaluation of Programmatic Approaches in the GEF	2018
112	Review of Results-Based Management in the GEF	2017
111	Evaluation of GEF Engagement with the Private Sector, Volumes 1 and 2	2017
110	OPS6: The GEF in the Changing Environmental Finance Landscape – Final Report	2017
109	GEF Annual Performance Report 2015	2017
108	Evaluation of the GEF–Civil Society Organization Network, Volumes 1 and 2	2016
107	GEF Country Portfolio Evaluation: Tajikistan (1999–2015), Volumes 1 and 2	2016
106	Program Evaluation of the Least Developed Countries Fund	2016
105	FEM Evaluation du Portefeuille de Pays : Maroc (1997–2015)	2016
104	Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems—Full Report	2016
103	Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems—Highlights	2015
102	GEF Annual Performance Report 2014	2015
101	GEF Country Portfolio Evaluation: Eritrea (1992–2012)	2015
100	GEF Country Portfolio Evaluation: Tanzania (1992–2012)	2015
99	GEF Country Portfolio Study: Sierra Leone (1998–2013)	2015
98	GEF Country Portfolio Evaluation: Vanuatu and SPREP (1991–2012)	2015
97	Joint GEF-UNDP Evaluation of the Small Grants Programme	2015
96	Joint GEF–Sri Lanka Country Portfolio Evaluation (1991–2012)	2015
95	GEF Annual Country Portfolio Evaluation Report 2014	2015
94	Midterm Evaluation of the System of Transparent Allocation of Resources	2014
93	Midterm Evaluation of the National Portfolio Formulation Exercise	2014
92	GEF Annual Performance Report 2013	2014
91	GEF Annual Impact Report 2013	2014
90	Impact Evaluation on Climate Change Mitigation: GEF Support to Market Change in China, India, Mexico and Russia	2014
89	Report of the Second Professional Peer Review of the GEF Evaluation Function	2014
88	OPS5: Final Report: At the Crossroads for Higher Impact—Summary	2014
87	GEF Annual Country Portfolio Evaluation Report 2013	2014
86	OPS5: Final Report: At the Crossroads for Higher Impact	2014

To see all GEF Independent Evaluation Office publications, please visit our webpage.

Independent Evaluation Office
Global Environment Facility
1818 H Street, NW Washington, DC 20433, USA
www.gefieo.org
/gefieo_tweets
/gefieo