



Independent
Evaluation Office
GLOBAL ENVIRONMENT FACILITY



Guidelines for Conducting Program Evaluation

October 2023

© 2023 Global Environment Facility Independent Evaluation Office
1818 H Street, NW, Washington, DC 20433
Internet: www.gefio.org/; email: gefevaluation@thegef.org

Reproduction permitted provided source is acknowledged. Please cite the work as follows: Global Environment Facility Independent Evaluation Office (GEF IEO), *Guidelines for Conducting Terminal Evaluations of Full-Size Projects*, Washington, DC: GEF IEO, 2023.

ISBN: 978-1-64233-052-6

Task Team Leader: Neeraj Kumar Negi, nnegi1@TheGEF.org
Chief Evaluation Officer: Geeta Batra
GEF IEO Director: Juha Uitto

Editing and design: Nita Congress

Contents

- About These Guidelines iii**

- 1. Introduction 1**

- 2. Roles and Responsibilities 3**
 - 2.1 Lead GEF Agency 3
 - 2.2 Evaluators 4
 - 2.3 GEF Operational Focal Points 5
 - 2.4 Other GEF Agencies 5

- 3. Report Content 6**
 - 3.1 Introduction and Background 6
 - 3.1.1 Program Information and Description 6
 - 3.1.2 Theory of Change 6
 - 3.1.3 Scope of Terminal Evaluation 7
 - 3.2 Methodology 7
 - 3.3 Program Performance 8
 - 3.3.1 Additionality 8
 - 3.3.2 Relevance 9
 - 3.3.3 Coherence 9
 - 3.3.4 Integration 9
 - 3.3.5 Effectiveness 10
 - 3.3.6 Efficiency 11
 - 3.3.7 Sustainability 11
 - 3.3.8 Knowledge Management 11
 - 3.4 Program Implementation 12

3.5 Program Monitoring and Evaluation	13
3.6 Program Cofinancing	14
3.7 Alignment with GEF Policies.	14
3.8 Lessons and Recommendations.	15
3.9 Performance Ratings	16
4. Report Submission and Use	17
4.1 Submission of Evaluation Report	17
4.2 Use of Evaluation Report.	17
Annex: Ratings	19
A.1 Outcome Ratings	19
A.2 Sustainability Ratings	20
A.3 Program Implementation Ratings	22
A.4 Program M&E Ratings	23
References and Resources	27

About These Guidelines

- i. Summary.** The Global Environment Facility (GEF) [Evaluation Policy](#) (GEF IEO 2019) requires that the lead GEF Agency will evaluate a program at the end of its implementation to provide a comprehensive and systematic account of its performance and the factors that affect performance. Program evaluations are expected to support accountability, learning, transparency, knowledge sharing, and decision making in the GEF. The purpose of these guidelines is to support GEF Agencies in conducting program evaluation, and to facilitate consistency and completeness in their reporting.
- ii. Applicability.** Program evaluation reports are required for all programs funded through the GEF Trust Fund and, where applicable, other GEF-administered trust funds.¹ These guidelines are applicable to the actors across the GEF partnership that have a role in the conduct of program evaluation. These include the lead GEF Agency of a program, evaluators, other GEF Agencies involved in implementation, and the GEF operational focal points of the countries in which a program is implemented. These guidelines are aimed at programs that were approved from GEF-6 onwards, but may also be used to evaluate programs that were approved earlier. Even though a program evaluation will cover the child projects of that program, Agencies should follow the [GEF Guidelines for Conducting Terminal Evaluations of Full-Size Projects](#) to prepare a terminal evaluation for each child project.

¹ In August 2006, the GEF Council decided that all GEF operational policies, procedures, and governance structures are to be applied to the Least Developed Countries Fund (LDCF) and the Special Climate Change Fund (SCCF), unless the LDCF/SCCF Council decides otherwise (GEF 2006). Accordingly, the LDCF and the SCCF currently apply the [GEF Evaluation Policy](#).

1. Introduction

- 1. Definition of the program modality.** The Global Environment Facility (GEF) [Project and Program Cycle Policy](#) defines the term “program” as “a longer-term and strategic arrangement of individual yet interlinked projects that aim at achieving large-scale impacts on the global environment” (GEF 2018, 4). The GEF uses the program modality as one of its main means of providing support to recipient countries.¹ GEF programs often consist of a coordinating project that provides the overall direction and support for program activities and several decentralized child projects that are tied together by a shared vision and framework. Although the GEF has used the program modality since its inception, total funding through the modality was low up to GEF-3. In GEF-4, the GEF Council endorsed the objectives and basic principles for programmatic approaches (GEF 2008). Thereafter, the share of GEF funding provided through programs increased substantially.
- 2. Evolution of the program modality.** The nature of GEF programs has changed over time because of a shift toward the use of **integrated approaches**. Although the GEF started providing support for the use of integrated approaches through its Operational Program on Integrated Ecosystem Management in 1999, use of the approach was limited through GEF-5. Use increased in GEF-6 when the GEF approved a set of integrated approach pilot programs. These pilots aimed to achieve transformational change at scale to address drivers of global environmental degradation (GEF IEO 2022). The pilots placed strong emphasis on knowledge management facilitated through a coordinating project. Other programs approved during GEF-6 also emphasized integrated approaches, system transformation, and addressing challenges at scale. Reliance on integrated approaches deepened in GEF-7 through support for **impact programs** (GEF IEO 2022).
- 3. Definition of evaluation.** The [GEF Evaluation Policy](#) defines evaluation as “Systematic and impartial assessment of planned, ongoing, or completed activities, projects, programs in specific focal areas or sectors; policies; strategies and their implementation; or other topics relevant to the GEF partnership and organization” (GEF IEO 2019, 3).
- 4. Definition of program evaluation.** Within the context of the GEF program cycle, “program evaluation” means evaluation of a GEF-supported program. A program evaluation is expected to provide a comprehensive and systematic account of a

¹ Other modalities include full-size projects, medium-size projects, and enabling activities.

program's performance, including aspects related to its design, implementation, achievement of objectives and targets, and factors that affected performance. The evaluation is expected to discuss the strengths and weaknesses of a program and lessons that may be drawn from its implementation. Program evaluation is expected to support accountability, learning, transparency, knowledge sharing, and decision making in the GEF.

5. **Purpose of guidelines.** These guidelines have been prepared within the framework of the [GEF Evaluation Policy](#) and draw upon the [GEF Evaluation Office Ethical Guidelines](#) of the GEF Independent Evaluation Office (IEO). They also draw on international norms and standards such as [Applying Evaluation Criteria Thoughtfully](#) (OECD 2021), [Norms and Standards for Evaluation](#) (UNEG 2016), and [Ethical Guidelines for Evaluation](#) (UNEG 2020). The guidelines are aimed at facilitating evaluation of GEF programs, especially those approved from GEF-6 onwards. They clarify the [roles and responsibilities](#) of the relevant actors, and the [content](#) of the evaluation report, and report's [submission and use](#). They also explain the [rating scales](#) used to assess program performance on select criteria.

2. Roles and Responsibilities

2.1 Lead GEF Agency

6. The lead GEF Agency² will perform the following:
 - a. **Evaluate a GEF program close to or at the end of its implementation.**³ The earliest a program evaluation may be initiated is when implementation of at least 60 percent of its child projects is complete and at least 80 percent of the GEF funding for the program has been utilized.⁴ The latest a program may be evaluated is up to six months after the end of program implementation.
 - b. **Develop specific terms of reference for a program’s evaluation and provide guidance, documentation, and support to the evaluation team.** The documentation submitted at program framework document approval, and endorsement/approval of coordinating project and child projects, will be useful to the evaluators. Project implementation reports and midterm reviews of program activities will also be useful.
 - c. **Archive information gathered** during program implementation and evaluation and, at the request of the GEF IEO, make this information available for follow-up studies.
 - d. **Ensure appropriate evaluator expertise.** Ensure that the evaluation team is composed of evaluators with expertise in the thematic area addressed by the program, integrated approaches and other important elements of the program’s strategy, fiduciary oversight, environmental and social safeguards, and gender-related concerns.
 - e. **Ensure that the program evaluation team is independent, unbiased, and free of conflicts of interest.** As required by Minimum Requirement 3 of the

² As of this writing, there are 18 GEF Agencies accredited by the GEF that may have direct access to GEF resources. Any of these Agencies may be involved in a program as its lead Agency.

³ End of implementation is understood as completion of the last program activity—including activities of child projects and the coordinating project but excluding program evaluation—supported through GEF funding.

⁴ The coordinating projects of some programs—especially those approved in GEF-6—are designed for a shorter duration than may be necessary to meet the threshold conditions for starting program evaluation. In these cases, the lead Agency may undertake program evaluation at the time the coordinating project is closed, even if the threshold conditions in terms of program progress have not been met. The evaluation report will clearly explain the reasons for conducting the evaluation earlier than required.

[GEF Evaluation Policy](#), the evaluators should not have been involved in the design, approval, implementation, or execution of the program or its child projects.

- f. Seek and address feedback from other GEF Agencies** involved in implementation of coordinating projects and/or child projects when developing the program evaluation's terms of reference and drafting the evaluation report (see [paragraph 10](#)).
 - g. Share the draft evaluation report with the GEF operational focal points (OFPs)** in the countries covered by the program (i.e., the countries in which the program activities were implemented) for comment and transmit the final report to them (see [paragraph 9](#)).
 - h. Submit the program evaluation report**, in English, through the GEF Portal within 6 months of the evaluation's completion and within 12 months of the end of program implementation.
 - i. Share the final evaluation report** with the relevant stakeholders and ensure that the report is publicly available.
- 7. Approaches to program evaluation.** The credibility of program evaluation relies on its independence from program implementation. GEF Agencies will use one of two broad approaches to ensure the independence of program evaluation: (1) the evaluation unit of an Agency commissions and/or manages the program evaluation; or (2) the Agency's operations unit manages and/or commissions the program evaluation, and its evaluation unit validates the information and performance ratings provided in the evaluation report.

2.2 Evaluators

- 8.** The evaluators conducting the program evaluation will observe the following guidelines.
 - a.** Evaluators will **be knowledgeable of**, or become familiar with, GEF programs and strategies, and with relevant GEF policies such as those on program approval, the program cycle, monitoring and evaluation (M&E), cofinancing, fiduciary standards, gender, stakeholder engagement, and environmental and social safeguards.
 - b.** Evaluators will take the **perspectives of all relevant stakeholders** into account. They will gather information on program performance and results from multiple sources to facilitate triangulation. They will seek the necessary

contextual information to assess the significance and relevance of observed performance and results.

- c. Evaluators will present a **balanced account** that is clear, candid, and consistent with the evidence.
- d. Evaluators will apply the **evaluation criteria** discussed in these guidelines (see [paragraph 16](#)). They will also evaluate performance in other areas that are important to the GEF. Where necessary, they may cover other topics to enhance the utility of the evaluation.
- e. Evaluators will use the **rating scales** provided in the [annex](#) to these guidelines.
- f. Evaluators will **abide by** the [GEF Evaluation Office Ethical Guidelines](#).

2.3 GEF Operational Focal Points

- 9. Minimum Requirement 4 of the [GEF Evaluation Policy](#) requires the involvement of the relevant OFPs in the program evaluation process. Where applicable, the OFPs will receive a draft evaluation report from the lead Agency for comment, as described in [paragraph 1.h](#). The OFPs are encouraged to provide inputs, while respecting the independent nature of evaluation.

2.4 Other GEF Agencies

- 10. Other GEF Agencies involved in implementation of the program will facilitate conduct of the evaluation by providing necessary support to the evaluation team. They will provide documentation that may be a useful source of information on the implementation and results of the child projects. The Agencies will also provide data on utilization of GEF resources and materialization of cofinancing for child projects. When requested, they will facilitate the evaluation team in accessing child project sites for verification, and help the team establish contact with key informants and project beneficiaries. They will provide feedback on the draft terms of reference and the draft evaluation report, as described in [paragraph 1.f](#).

3. Report Content

3.1 Introduction and Background

3.1.1 Program Information and Description

Outcome: an intended or achieved short- or medium-term effect of a project or program's outputs (GEF IEO 2019).

11. The evaluation report will provide general information on the program such as its name, list of coordinating and child projects and their respective GEF project IDs; countries covered by the program; the lead GEF Agency and other Agencies involved in implementation; the program implementation team, the program evaluation team; the program milestone dates (e.g., start date, expected end date, actual implementation completion date, or completion date anticipated when the evaluation was conducted); and GEF financing, and promised and materialized cofinancing. The report will list program objectives and expected [outcomes](#), the program's strategy, a brief description of program activities, coordinating project and child projects, and the institutional arrangements for implementation.

3.1.2 Theory of Change

Impact: positive or negative, primary or secondary long-term effect produced by a project or program, directly or indirectly, intended or unintended (GEF IEO 2019).

12. **Definition.** A program's theory of change describes the causal relationships through which a program would deliver its expected [impact](#), along with the assumptions about the conditions necessary to make the causal relationships work; it provides a basis for evaluating the program. Where available, the program evaluation will use the theory of change presented in the program framework documents as a basis for evaluation. If necessary, the evaluators will update the theory in consultation with the key stakeholders to address gaps and reflect changes made to the program during implementation. Where the program framework documents do not explicitly provide a theory of change, the evaluators will develop one based on the information in the program documents and in consultation with the key stakeholders.
13. **Coverage and consistency.** The program evaluation report will describe the program's theory of change including the outputs, outcomes, intermediate states, and intended long-term impacts of the program; the causal pathways for the long-term impacts; mechanisms through which these impacts are to be achieved; and implicit and explicit assumptions. The theory should illustrate how child projects, along with the coordinating project, contribute to achieving the program's intended outcomes and long-term impacts, including any planned interactions among the child projects that are expected to contribute to these

ends. The theory of change should outline synergies the program intends to tap to achieve the intended systemic change, and the trade-offs involved in key choices made in program design. Where applicable, the theory of change should indicate how the given environmental challenges have been addressed through integration.

3.1.3 Scope of Terminal Evaluation

14. The evaluation report will define its scope in terms of the program's objectives, theory of change, supported activities (e.g., coordinating project and child projects), the context in which the program was designed and implemented, and program M&E data. The scope will also be determined based on the questions the evaluation seeks to answer, the boundaries of the system covered, the features of the program and its components, geographical coverage, period under review, key users, and likely uses of the evaluation.

3.2 Methodology

15. **Coverage.** The evaluation report will describe the evaluation methodology. It will discuss evaluation questions, analytical framework, and criteria, with reference to the program's theory of change; information sources; methods; and limitations. The theory of change (see [3.1.2 Theory of Change](#)) will provide a basis for the type of evidence needed to assess whether the theory is valid and whether the actual program [results](#) and causal relationships are consistent with the theory.
16. **Criteria.** The Organisation for Economic Co-operation and Development (OECD) defines six main criteria for evaluation (OECD 2021): relevance (see [3.3.2 Relevance](#)), coherence (see [3.3.3 Coherence](#)), effectiveness (see [3.3.5 Effectiveness](#)), efficiency (see [3.3.6 Efficiency](#)), impact, and sustainability (see [3.3.7 Sustainability](#)). These criteria will be applied to evaluate GEF programs. The evaluation will also cover other areas of significance to the GEF such as additionality (see [3.3.1 Additionality](#)); program M&E (see [3.5 Program Monitoring and Evaluation](#)); performance of GEF Agencies (see [3.4 Program Implementation](#)); materialization of cofinancing (see [3.6 Program Cofinancing](#)); application of environmental and social safeguards, gender equality, and stakeholder engagement including involvement of the private sector (see [3.7 Alignment with GEF Policies](#)); integration (see [3.3.4 Integration](#)); and knowledge management (see [3.3.8 Knowledge Management](#)). These last two considerations—integration and knowledge management—have been key features of GEF programs approved from GEF-6 onwards. Similarly, several programs target transformative change at scale; this may be covered in the evaluation report using applicable approaches (Uitto and

Results: include intervention outputs, outcomes, and progress toward longer-term impact including global environmental benefits; they should be discernible/measurable (GEF IEO 2019).

Batra 2022). Additional themes and topics may be covered to enhance the utility of the evaluation.

17. **Information sources.** The evaluation report will draw from different information sources to facilitate triangulation. Evaluations that draw upon desk reviews will provide information on the documents reviewed, the rules for document inclusion and exclusion, and the instrument(s) used for review. For interviews, the evaluation report will provide information on interviewees, including their names, positions, and affiliations; the interview date(s); and the interviewee selection process. For field verifications, information such as the basis for site selection, which sites were visited and when, and who visited the sites will be provided. If surveys (including online surveys) are conducted, the report will discuss the sample frame, sampling approach, response rates, and period during which the survey was conducted.
18. **Program M&E.** The program's M&E system is expected to gather information on environmental stress reduction (greenhouse gas emissions reduction, reduction of waste discharge, etc.) and environmental status change (change in population of endangered species, forest stock, pollutants in the waterbody, water retention in degraded lands, etc.). The M&E system is also an important source of information on program implementation, M&E plan implementation, and adaptive management. The evaluation should assess the quality of information gathered through the program M&E system at both the program and child project levels and, where its quality is adequate, make use of this information.

3.3 Program Performance

3.3.1 Additionality

19. To determine [additionality](#), the evaluation will compare the benefits of GEF support to a scenario without GEF support. It will identify specific areas where GEF support has contributed additional results and the nature of these additional results.
20. The evaluation will assess the extent to which benefits delivered using a programmatic approach are quantitatively and qualitatively different from a nonprogrammatic baseline approach—that is, are the benefits of a program different from a comparable cluster of stand-alone projects with similar objectives, inputs, and activities? A comparable GEF stand-alone investment may provide a basis on which to assess the benefits of the program modality.

Additionality: the extent to which GEF support to a project has contributed to environmental and socioeconomic impacts; legal, institutional, and financial results; socioeconomic benefits; learning; and innovation (GEF IEO 2020).

3.3.2 Relevance

21. The evaluation will assess a program's relevance by considering the extent to which the program's objectives and design are responsive to the environmental and development priorities of the participating countries and are congruent with GEF strategic priorities and objectives.

3.3.3 Coherence

22. **External coherence.** The report will discuss the external coherence of the program by considering its compatibility with other programs and activities being implemented in the participating countries, targeted sectors, and institutions. It will discuss the extent to which the program involves socioeconomic trade-offs. It will also assess whether the program is well targeted (in terms of countries, regions, and/or communities covered) to generate global environmental benefits.
23. **Internal coherence.** The report will discuss the internal coherence of the program. It will assess the extent to which program design tackled the drivers of the environmental challenges it sought to address. It will assess the extent to which the theory of change, governance structure, activities, child projects, and M&E system are aligned with program objectives. It will also discuss whether and how the program design incorporated lessons from past GEF activities and other relevant non-GEF interventions.

3.3.4 Integration

24. GEF programs approved from GEF-6 onwards use integrated approaches to address drivers of environmental degradation. The evaluation will assess how the program design applied integrated approaches, the extent to which [integration](#) was achieved, and the results. It will also weigh and discuss the extent to which the use of an integrated approach—as opposed to a simpler segregated approach or other program design choices—was appropriate given the context, objectives, and costs involved.
25. Individual GEF programs may apply systems thinking in different ways. For example, the Sustainable Cities Impact Program uses the circular economy approach, and the Commodities Integrated Approach Pilot is designed to provide support across a commodity's supply chain. An integrated program includes feedback loops to facilitate quick adaptation (GEF STAP 2018). The specified results indicators cover the major pathways through which the different drivers of environmental degradation are mitigated. The indicators are tracked regularly and used to improve program effectiveness. A program should have applied these

Integration implies the use of systems thinking. It involves specifying system boundaries, addressing multiple drivers of environmental degradation simultaneously, addressing relationships among the system elements across scales, addressing key risks and vulnerabilities, considering system resilience, and establishing a feedback loop that facilitates timely course correction (GEF STAP 2018).

principles to generate global environmental benefits. The evaluation report will discuss the level of integration achieved, highlighting areas where integration was evident and effective, as well as areas where it was not evident and/or may be improved.

3.3.5 Effectiveness

26. The evaluation report will discuss the extent to which program objectives and targets—as stated in the program framework documents—have been achieved. It will note if there were any changes in the program’s design and/or expected results after the start of implementation.
27. **Achievement.** Consistent with Minimum Requirement 3 of the [GEF Evaluation Policy](#), the report will present an assessment of achievement of program results. The report will describe the relevant system boundaries, the drivers of environmental degradation, and the system transformation outcomes targeted by the program. It will present an assessment of the extent to which progress has been made in achieving the targeted outcomes and long-term impacts of the program framework document. If the document provides targets for expected program [outputs](#), their level of achievement will also be covered.
28. The report will aggregate data on expected and actual contributions of the program and its child projects to GEF corporate results targets (core indicators). The targets provided in the program framework document will provide a basis for comparison with actual results. For child projects that may still be under implementation, progress made up to the point of evaluation will be reported.
29. **Adoption.** The evaluation report should provide an account of processes such as mainstreaming, replication, scale-up, and market change that indicate progress toward environmental and social changes at scales beyond that directly addressed by the program activities. The report should point out the specific elements of the program that are being adopted beyond GEF support. It should also discuss the extent to which experiences from one child project have influenced activities in others and, where applicable, the role played by the coordinating project in the process.
30. **Factors affecting achievement.** The report will discuss factors that affected delivery of outputs and outcomes, and progress to impacts, such as program design, implementation, linkages with other interventions, materialization of cofinancing, stakeholder involvement, and macroeconomic factors. If there are substantial differences in the performance of child projects, the report should discuss factors that may explain the variance.

Output: product or service that results from the completion of activities implemented within a project or program (GEF IEO 2019).

- 31. Unintended effects.** The evaluators should document the unintended impacts—both positive and negative—of the program and assess their overall scope and implications. In cases where these impacts are undesirable from an environmental or socioeconomic perspective, the report should suggest corrective actions, along with preventive/mitigative measures that may be useful in future interventions.

3.3.6 Efficiency

- 32.** The criterion of efficiency is aimed at assessing the extent to which the inputs are converted into results in an [economic](#) and [timely](#) way (GEF IEO 2019; OECD 2021). The report will discuss the extent to which the program was efficient in achieving its results. It will discuss areas of efficiency and inefficiency, and factors that affected efficiency.

3.3.7 Sustainability

- 33.** Minimum Requirement 3 of the [GEF Evaluation Policy](#) specifies that a program evaluation will assess [sustainability](#). The assessment will weigh risks to the continuation of net benefits from the program given their probability and severity. The report will discuss key risks and explain how these may affect the continuation of net benefits over a reasonable time frame. The report will also discuss the extent to which the flow of program benefits is resilient to shocks. This analysis should cover financial, sociopolitical, institutional, and environmental risks and shocks, with other likely risks discussed as applicable. The report should address the variation in levels of risk faced by the different program implementation sites and the reasons for these differences. It should discuss the measures that were undertaken, and/or may be needed, to mitigate risks.

3.3.8 Knowledge Management

- 34.** Programs approved from GEF-6 onwards give considerable attention to knowledge management. The coordinating projects of these programs are expected to facilitate knowledge exchange among child projects, GEF Agencies, executing agencies, and other stakeholders. These programs are often expected to develop knowledge platforms that aid data capture, transformation, and sharing.
- 35. Coverage.** The report will discuss the extent to which, and how, the knowledge platforms created by the program are contributing to the generation of global environmental benefits. It will discuss the effectiveness of the information technology system put in place to facilitate capture, storage, retrieval, and use of knowledge; as well as the types of knowledge products planned and delivered by

Economic: the cost-effective conversion of inputs such as funds, expertise, personnel, equipment, etc., into results compared to feasible alternatives.

Timely: the extent to which project activities were started and completed within a reasonable time frame.

Sustainability: the continuation/likely continuation of positive effects from an intervention after it has come to an end, and its potential for scale-up and/or replication; interventions need to be environmentally as well as institutionally, financially, politically, culturally, and socially sustainable (GEF IEO 2019).

the program and their utility. The report will highlight strong and weak areas of performance in knowledge management. It will also discuss good practices from the program that may be relevant for other GEF activities, along with areas where knowledge management in future programs and projects may be improved.

3.4 Program Implementation

36. **Coverage.** The evaluation report will discuss how well the program was implemented by the GEF Agencies. It will describe the program's implementation experience, challenges that were faced at the program and project levels at different stages of the program cycle, how these challenges were addressed, and circumstances that aided or hindered implementation. It will describe how information from program M&E was used to steer the program and its use of adaptive management. The report will also document instances where opportunities for adaptive management were lost.
37. The report will discuss the performance of the lead GEF Agency and other partner Agencies, providing an account of how the Agencies performed in their respective roles. In cases where the lead Agency is involved in child project implementation, its role in this capacity will be discussed along with those of other Agencies involved in child project implementation; its role as lead Agency will be assessed separately. The report will also discuss issues that may need follow-up from the lead Agency and other GEF Agencies.
38. **Lead GEF Agency.** The report will discuss how well the lead Agency led the program. The operational units of the lead Agency are responsible for overall delivery of program outputs and outcomes. The key functions of a lead Agency may include—but are not restricted to—development of the program framework document, coordination among child projects, program-level monitoring and reporting, identification and management of program-level risks, application of GEF fiduciary standards (see [paragraph 48](#)) in implementation of the coordinating project (and the child projects implemented by it), knowledge management across the program, midterm review, and program evaluation. The evaluation report will discuss how well the lead Agency performed its responsibilities, the challenges it faced, and the measures it took to respond to these challenges.
39. **Other GEF Agencies.** The report will discuss how well the Agencies involved in the implementation of child projects performed their responsibilities as these relate to overall program progress and performance. It will present an assessment of how well these Agencies collaborated with the lead GEF Agency and among themselves. It will cover topics such as choice of executing agencies,

timeliness of child project start-up and implementation, supervision during implementation, how well relevant risks were identified and managed for child projects, application of GEF fiduciary standards (see [paragraph 48](#)), implementation of M&E plans for child projects and collection of data on core indicators and/or tracking tools, and support for program midterm review and final evaluation. A table that summarizes information by child project and responsible Agency may be useful.

3.5 Program Monitoring and Evaluation

40. Minimum Requirement 1 of the [GEF Evaluation Policy](#) calls for a fully developed and budgeted program M&E plan during program preparation; Minimum Requirement 2 calls for implementation of this plan during program implementation. The evaluation will assess strengths and weaknesses of the program M&E plan and its implementation. The focus will be on arrangements to track and report on program implementation and its results. Most coordinating projects include activities to coordinate program M&E. M&E-related discussions included in the terminal evaluation of the coordinating project will be a useful source of information on program M&E for the program evaluation.
41. **M&E design.** The report will present an assessment of the program's M&E plan. The assessment will consider the extent to which the M&E plan was practical and well thought out. For example, did it clearly identify the M&E activities at the program level and how these will draw on the M&E activities to be implemented at the child project level? The report will also assess if the program's M&E plan adequately addressed the program's theory of change, GEF M&E requirements, incorporated applicable core indicators and tracking tools, and provided baseline information. It will discuss whether the indicators specified to track environmental, gender, socioeconomic, and other results are appropriate ([SMART](#)) and consistent across child projects.
42. **M&E implementation.** The evaluation report will present an assessment of M&E implementation. It will discuss the extent to which the M&E system operated as planned. If changes were made to the M&E plan, it will note these changes, the reasons for them, and whether the changes were timely. The report will describe the extent to which information on the specified indicators was gathered systematically, in a timely fashion, and following a robust methodology. It will discuss how data from the child projects and coordinating project were aggregated and whether and how these data were used for decision making. It will discuss whether the resources allocated for M&E were sufficient and used prudently.

SMART: specific; measurable; achievable and attributable; relevant and realistic; and time-bound, timely, trackable, and targeted.

Finally, it will discuss how the information from the M&E system was used in decision making and adaptive management.

3.6 Program Cofinancing

Cofinancing: financing that is additional to GEF financing and that supports implementation of a GEF-financed activity and achievement of its objectives (GEF 2018).

43. The use of [cofinancing](#) in GEF activities ensures that the GEF finances only the incremental (or additional) costs of generating global environmental benefits. Cofinancing may also enhance the scale of the supported activities, support sustainability, and strengthen partnerships. Timely materialization of cofinancing is crucial to ensure that sufficient resources are available for planned activities and that these activities are completed on time.
44. The evaluation report will provide information on cofinancing sources, cofinancing commitments, and activities supported through cofinancing. It will report on the type of cofinancing mobilized: cash or in kind, grant, loan, or equity; managed by the GEF Agency implementing the program or in parallel by other partner organizations. It will provide information on the extent to which expected cofinancing materialized and was timely. In cases where materialization of cofinancing deviated from commitments, or where there was a delay in materialization, the report will discuss the reasons for these, and the effects on program implementation and results.

3.7 Alignment with GEF Policies

45. **Environmental and social safeguards.** It is expected that GEF activities will not cause any harm to the environment or to any stakeholder and that, where applicable, Agencies will take measures to prevent and/or mitigate any adverse effects. The evaluation report will discuss the extent to which appropriate environmental and social safeguards were applied, in accordance with the GEF [Policy on Environmental and Social Safeguards](#), especially in the design and implementation of child projects. It will note how project activities affected marginalized communities including, where applicable, indigenous peoples.
46. **Gender equality.** In accordance with the GEF [Policy on Gender Equality](#), the report will discuss the extent to which gender considerations were addressed in designing and implementing the program, including its child projects. The report will discuss whether a gender analysis was conducted by the child projects, the extent to which the program was implemented so as to ensure gender-equitable participation and benefits, and whether gender-disaggregated data on beneficiaries were gathered. If the GEF program under evaluation or any of its

activities disadvantaged or may disadvantage women, this should be documented and reported. The evaluators should also determine the extent to which relevant gender-related concerns were tracked through the program's M&E system, including for the individual child projects.

- 47. Stakeholder engagement.** The report will provide an account of stakeholder engagement in various program activities including its design and implementation. The GEF [Policy on Stakeholder Engagement](#) defines stakeholder engagement as a process involving stakeholder identification and analysis; planning of stakeholder engagement; consultations with, and participation of, stakeholders in monitoring, evaluation, and learning through implementation; addressing their grievances; and ongoing reporting to [stakeholders](#). Considering the program objectives, the report will discuss the extent to which, and how, the program identified and engaged with the relevant stakeholders including the private sector, civil society organizations, government agencies, and program beneficiaries. The report will discuss the role stakeholder engagement has played in determining program ownership, building partnerships, and drawing on stakeholder expertise and capacities in program design and implementation. The report should discuss the extent to which, and how, stakeholder engagement has affected implementation, results, and risks to program sustainability.
- 48. Fiduciary standards.** The report will discuss the extent to which fiduciary standards were applied during program implementation. The GEF fiduciary standards are aimed at ensuring “accountability, integrity, transparency, efficiency and effectiveness in the use of GEF financing” by the GEF Agencies (GEF 2019b). The standards cover areas including, but not limited to, program appraisal standards, procurement processes, financial management and control, financial audits, financial disclosure, conflict of interest, code of ethics, whistle-blower protections, and safeguards against anti-money laundering and financing of terrorism. The evaluators will take stock of the extent to which program implementation was consistent with GEF fiduciary standards as set out in the [Policy on Minimum Fiduciary Standards](#), relying on Agency responses to evaluator questions and information gathered through review of terminal evaluations, midterm reviews, and child project implementation reports. Where necessary, the program evaluation report will point out areas for improvement.

Stakeholder: an individual or group that has an interest in a GEF-financed activity or is likely to be affected by it, such as local communities, indigenous peoples, civil society organizations, and private sector entities (GEF 2017b).

3.8 Lessons and Recommendations

- 49. Lessons.** Evaluators should provide a few well-formulated lessons based on program experience, including experiences from individual child projects. Wherever possible, an evaluation report should include examples of good practices in

design and implementation that have led to effective stakeholder engagement, successful broader adoption of GEF initiatives by stakeholders, and large-scale environmental impacts. The evaluators should describe aspects of program performance that worked well, along with reasons for this. They should discuss the conditions under which these good practices may or may not be replicated.

- 50. Recommendations.** Recommendations should be well formulated and targeted. They should clearly state the need for action; the action to be taken, along with its likely consequences vis-à-vis the status quo and other courses of action; the specific actor(s) to carry out the recommended action; and the time frame.

3.9 Performance Ratings

- 51.** Program performance will be rated on criteria for outcomes, sustainability, implementation, and M&E ([table 1](#)). The ratings will be on an even-number Likert scale, as described in the [annex](#). These ratings may be provided in a table in the main body of the report or as an annex to the report. They may also be embedded in the corresponding report sections covering each topic.

Table 1: Performance Ratings

Criterion	Dimension	Rating scale
Outcomes	Relevance, coherence, effectiveness, efficiency	Six points: highly satisfactory to highly unsatisfactory
Sustainability	Financial, sociopolitical, institutional, environmental	Six points: highly likely to highly unlikely
Implementation: lead GEF Agency	Performance of the Agency leading the program in its role as program lead	Six points: highly satisfactory to highly unsatisfactory
Implementation: other Agencies	Performance of Agencies implementing child projects of the program	Six points: highly satisfactory to highly unsatisfactory
M&E design	Quality of design of program M&E plan	Six points: highly satisfactory to highly unsatisfactory
M&E implementation	Quality of program M&E during implementation	Six points: highly satisfactory to highly unsatisfactory

4. Report Submission and Use

4.1 Submission of Evaluation Report

52. GEF Agencies will submit a final program evaluation report through the GEF Portal no later than six months after completion of the program evaluation and no later than one year after completion of program activities. If preparation of the program evaluation was delayed, the evaluation report will explain the reasons for the delay. In cases where the program evaluation has been conducted by an Agency's operations unit, the lead GEF Agency will also submit a validation report from its evaluation unit up to one year after completion of the program evaluation.
53. The GEF IEO will track and report on Agency submission of completed program evaluations. The IEO will also follow up with the respective Agencies on any errors noted in their submissions.

4.2 Use of Evaluation Report

54. **GEF IEO.** The information provided by the program evaluations is used by the GEF IEO to report to the GEF Council on GEF portfolio performance. It is also an input to the comprehensive evaluations prepared by the GEF IEO. The IEO synthesizes information presented in the program evaluations to assess performance of GEF investments through programs. It validates a sample of the program evaluation reports, and shares the evaluation reports publicly through its website.
55. **Other GEF entities.** The GEF Secretariat and GEF Agencies may use program evaluations to inform their future work and for knowledge sharing. The evaluations may also be used by the GEF Scientific and Technical Advisory Panel to develop advisory products for the GEF partnership.

Annex: Ratings

1. The evaluators will rate program performance on outcomes, sustainability, implementation, execution, M&E design, and M&E implementation. The rated dimensions are described here along with a description of the level of performance for a specific rating. In most instances, actual performance may not fully correspond to any of the rating descriptions. Therefore, a rating will be assigned based on the description that best fits the evidence. Where available evidence is insufficient to rate performance, performance will be rated as unable to assess.

A.1 Outcome Ratings

2. The overall rating of program outcomes will be based on the following criteria.
 - a. **Relevance.** The evaluators will assess the extent to which the program outcomes aligned with the GEF focal areas/operational program strategies, country priorities, beneficiary needs, and the mandates of the GEF Agency and its executing partners. The evaluators will assess if the program is well targeted and the extent to which the program design is appropriate for delivering the expected outcomes.
 - b. **Coherence.** The evaluators will assess the extent to which the program is compatible with other relevant projects and programs in the recipient country or countries—that is, the extent to which it supported and was supported by other relevant activities in the program context. They will assess internal coherence by determining the extent to which there is alignment among the program’s theory of change, governance structure, activities, and M&E system. They will assess alignment with GEF policies and guidelines, and integration of lessons from similar interventions into the design.
 - c. **Effectiveness.** The evaluators will consider the extent to which program outcome achievements were commensurate with the ex ante targets. They will weigh the extent to which the program made the expected level of contributions to global environmental benefits. They will consider overall progress in achieving the long-term objectives. They will also consider the unintended consequences of the program and the extent to which these add to, or negate, program benefits.
 - d. **Efficiency.** This criterion is focused on the extent to which the program was cost-effective in delivering its intended results. The evaluators will consider

the program's cost/time versus outputs/outcomes equation, and, where feasible, compare the program's costs and benefits to alternatives. They will also consider the extent to which program activities were completed in a timely manner.

3. The program outcome rating will be based on the extent to which the expected outcomes were achieved (effectiveness), and the extent to which the program was relevant, coherent, and efficient. Although the evaluators will consider performance on these four criteria, the overall rating need not be a simple average of the criteria because a criterion may be more or less important depending on the type of program and its operational context. A six-point rating scale is used to assess outcome. The top three ratings comprise the satisfactory range and the bottom three (excluding unable to assess) the unsatisfactory range.
 - a. **Highly satisfactory.** The outcomes exceed targets and are highly relevant, coherent, and cost-effective.
 - b. **Satisfactory.** The level of outcomes achieved meets targets. The outcomes are relevant, coherent, and cost-effective.
 - c. **Moderately satisfactory.** The level of outcomes achieved was generally close to the targets. The majority of the targets were met or almost met, but some were not. The outcomes are generally relevant, coherent, and cost-effective.
 - d. **Moderately unsatisfactory.** Overall, the level of outcomes achieved is lower than the targets, although some outcomes were substantially achieved. The outcomes are generally relevant but not sufficient given the costs or, alternatively, are generally cost-effective but not adequately relevant and coherent.
 - e. **Unsatisfactory.** The expected outcomes were not achieved, or achievement was substantially lower than expected, and/or the achieved outcomes are not relevant or coherent. Alternatively, the outcome was cost-ineffective compared to alternatives.
 - f. **Highly unsatisfactory.** A negligible level of outcomes was achieved and/or the program had substantial negative consequences that outweigh its benefits.
 - g. **Unable to assess.** The available information does not allow assessment of the level of outcome achievement.

A.2 Sustainability Ratings

4. The rating for likelihood of sustainability will be based on the probability of occurrence of a risk and the magnitude/severity of its effects on continuation of net

benefits and achievement of long-term program objectives should it materialize. The assessment also considers the resilience of the program benefit stream to the likely risks. It will assess the likelihood of continuation over a time frame reasonable for the given program. At the time of the evaluation, a program may not face the consequences of the risk materializing, or the risk may be just beginning to materialize. The assessment should be based on the evidence of risks available at the time of evaluation. Most risks may be categorized as financial, sociopolitical, institutional, or environmental.

- a. Financial risks.** The evaluators will assess both the likelihood that financial resources will be available to continue the activities that sustain program benefits and the risks associated with their availability. They should, for example, look at support for income-generating activities that promote environmentally friendly behavior, regular government budget allocations for activities supported by the GEF program, and trends suggesting the future adequacy of financial resources for sustaining program outcomes. In cases where a program supported a business model, evaluators should consider risks that undermine the financial viability of the model.
- b. Sociopolitical risks.** The evaluators will assess the extent to which social or political risks may undermine the longevity of program outcomes. They will assess the extent to which the level of stakeholder ownership is insufficient to allow for program outcomes/benefits to be sustained, as well as the extent to which the interests of key stakeholders are aligned to support continuation of program benefits flow. For example, there may be a risk of policy reversal that will negate the progress made by the program, or enabling conditions that will alternatively lead to supportive policies. Evaluators will also assess the extent to which there is sufficient knowledge and public/stakeholder awareness of the program's long-term objectives and the progress made so far by the program in achieving these objectives.
- c. Institutional framework and governance risks.** The evaluators will assess if the legal framework, policies, governance structures, and processes pose any threat to the continuation of program benefits. In assessing these risks, the evaluators will consider if the requisite systems for accountability and transparency, and the necessary technical and institutional know-how, are in place.
- d. Environmental risks.** The evaluators will assess if there are any environmental risks that can undermine the future flow of program benefits. The evaluators should assess whether certain activities in the program area will pose a threat to the sustainability of program outcomes. For example, program outcomes may be especially vulnerable to climate change risks.

Similarly, biodiversity-related gains made by a program targeting marine protected areas may be affected by an increase in pollutant accumulation.

5. Additional risks that do not fall into the above categories also need to be considered in determining the overall sustainability rating. Once the probability of incidence of all relevant risks and the magnitude of their effect/severity have been taken into account, the following six-point scale for the overall likelihood of sustainability should be used.
 - a. **Highly likely.** There is negligible risk to continuation of benefits and based on the progress made so far it is expected that the long-term objectives of the program will be achieved.
 - b. **Likely.** Either there is negligible risk to continuation of benefits or there are some risks, but the magnitude of their effect is too small and/or the probability that they will materialize is too small. Overall, it is likely that the net benefits of the program will continue.
 - c. **Moderately likely.** There are some risks to sustainability, and they may have some effect on continuation of benefits if they materialize. However, probability of materialization of these risks is low. Net benefits are more likely to continue than abate.
 - d. **Moderately unlikely.** There are significant risks to sustainability. The effect on continuation of benefits would be substantial if these risks materialize and the probability of materialization of these risks is significant. Overall, net benefits of the program are likely to abate.
 - e. **Unlikely.** Because of the high risks it is unlikely that net benefits of the program will continue to accrue, and the progress made so far is likely to be lost. It is unlikely that the program will achieve its long-term objectives.
 - f. **Highly unlikely.** It is expected that the program will not achieve its long-term objectives. Major risks have either already materialized and halted accrual of net benefits or have high probability of materializing soon and will halt accrual of net benefits when they materialize.
 - g. **Unable to assess.** Unable to assess the expected incidence and magnitude of risks to sustainability.

A.3 Program Implementation Ratings

6. Program implementation ratings will be based on an assessment of the extent to which the GEF lead Agency and the GEF Agencies involved in implementation

of child projects fulfilled their respective roles satisfactorily. This assessment excludes the performance of those agencies that execute activities on the ground under the supervision of GEF Agencies.

7. Two ratings will be provided for program implementation: for the lead GEF Agency of the program in its role as program lead, and, collectively, for the GEF Agencies involved in child project implementation. That is, a single rating is provided for implementation of all child projects; the rationale being that each child project will be rated in its respective terminal evaluation. [Table A.1](#) provides a description of the scale used to rate the performance of the lead Agency and other Agencies.

A.4 Program M&E Ratings

8. The quality of M&E arrangements will be rated at the program level in terms of both design and implementation. The evaluation will cover the coordinating project's M&E activities focused on program M&E. The evaluation will also cover the extent to which M&E arrangements in child projects were aligned with, and fed into, program M&E. The quality of M&E on these two dimensions will be assessed separately on a six-point scale ([table A.2](#)).
 - a. **Design.** The assessment will consider the practicality and sufficiency of the M&E plan developed for the program. It will examine the extent to which the program and child project documents provide a baseline. It will consider the extent to which the program M&E plan specifies (1) applicable core indicators and clear targets, including appropriate (SMART—specific, measurable, attributable, realistic, and time-bound) indicators to track environmental, gender, and socioeconomic results; (2) a proper methodological approach for data collection; (3) logistics of the M&E activities, including a schedule and responsibilities for data collection; and (4) a budget with adequate funds for M&E activities.
 - b. **Implementation.** The assessment will consider the extent to which the program's M&E system operates as per the M&E plan. It will examine whether there was a need to revise the plan, and if so, it will consider if revisions were carried out in a timely manner. It will consider whether sufficient resources were provided for M&E. It will determine if the data on indicators—including core indicators, tracking tools, and other specified indicators—were gathered in a systematic manner. It will consider the extent to which appropriate methodological approaches were used, and the extent to which program management made use of the information from M&E during program implementation.

Table A.1: Program Implementation Rating Scale

Rating	Lead program Agency	Agencies implementing child projects
Highly satisfactory	Performance of the lead implementing GEF Agency was exemplary and exceeds expectations. The program framework document was robust, and the program was well led. The Agency maintained strong oversight and ensured that relevant GEF policies were implemented. The Agency identified and addressed emerging concerns in a timely manner. Program M&E was robust, and its activities were implemented in a timely manner. The Agency ensured that program implementation was on track.	Performance of the GEF Agencies implementing the child projects was exemplary and exceeds expectations. The preparation of child projects was robust and well aligned with program objectives. Child projects were well implemented. The Agencies collaborated well with the lead Agency and among themselves. They identified and addressed emerging concerns in the child projects in a timely manner. The M&E plans of child projects were robust, and their activities were implemented in a timely manner. The Agencies ensured that implementation of all or almost all child projects was completed, or is on track to be completed, on time.
Satisfactory	Performance of the lead Agency did not have any major weaknesses, and expectations were met. Program preparation and implementation were robust, and relevant GEF policies were applied. The Agency supervised the program well—it identified and addressed emerging concerns in a timely manner, although some weaknesses remained. The M&E plan was robust, and its activities were implemented in a timely manner. The Agency ensured that program implementation was on track.	Performance of the GEF Agencies implementing the child projects did not have any major weaknesses, and expectations were met. The preparation and implementation of child projects was robust, and relevant GEF policies were applied. The Agencies supervised the child projects well—they identified and addressed emerging concerns in a timely manner, although some weaknesses may have remained. The M&E plans of the child projects were robust, and M&E activities were implemented in a timely manner. The Agencies ensured that implementation of about two-thirds or more of the child projects was completed, or is on track to be completed, on time.
Moderately satisfactory	Performance of the lead Agency had some weaknesses, but overall expectations were met. Program preparation and implementation were adequate and relevant. GEF policies were applied, although there were some areas of weak implementation. The Agency supervised the program adequately—it identified and addressed emerging concerns, although some remained unaddressed or were inadequately addressed. The M&E plan was solid, although some activities were dropped or were implemented with delays. Program implementation had minor delays, and a few activities were dropped.	Performance of the GEF Agencies implementing the child projects had some weaknesses, but overall expectations were met. The preparation and implementation of child projects was adequate and relevant. GEF policies were applied, although there were some areas of weak implementation. The Agencies supervised implementation of the child projects adequately—they identified and addressed emerging concerns, although some concerns may have remained inadequately addressed. The M&E plans of the child projects were solid. At least half of the projects were completed, or are on track to be completed, on time.

(continued)

Table A.1: Program Implementation Rating Scale (continued)

Rating	Lead program Agency	Agencies implementing child projects
Moderately unsatisfactory	<p>Performance of the lead Agency did not meet expectations overall, although there were several areas of solid performance. Program preparation and implementation had weaknesses, although these were not too severe. Program supervision was somewhat weak. Although most emerging concerns were identified, many remained unaddressed or inadequately addressed. M&E implementation was somewhat weak. Program implementation was delayed, and a few activities were dropped or reduced in scale.</p>	<p>Performance of the GEF Agencies implementing the child projects did not meet expectations overall, although there were several areas of solid performance. The preparation and implementation of child projects had weaknesses, although these were not too severe. Supervision of child project implementation was somewhat weak. Most emerging concerns were identified, but several remained unaddressed or inadequately addressed. M&E plan and/or implementation in child projects was somewhat weak. Implementation of at least half of the child projects experienced delays. Consequently, a few activities were dropped or reduced in scale.</p>
Unsatisfactory	<p>There were major shortcomings in implementation, and the lead Agency did not meet the expected level of performance. Program preparation and implementation had severe weaknesses. Adherence to fiduciary standards was somewhat weak. Program supervision was weak—emerging concerns were not identified in time and remained unaddressed or inadequately addressed. M&E implementation was weak, and M&E activities were not implemented in a timely manner. Program implementation was delayed, and several activities were dropped or reduced in scale.</p>	<p>There were major shortcomings in implementation, and the GEF Agencies that implemented the child projects did not meet the expected level of performance. Child project preparation and implementation had severe weaknesses. Adherence to fiduciary standards was somewhat weak. Supervision of child projects was weak—emerging concerns were not identified in time and remained unaddressed or inadequately addressed. M&E implementation was weak, and M&E activities were not implemented in a timely manner. Implementation of at least two-thirds of the child projects was delayed, and several activities were dropped or reduced in scale.</p>
Highly unsatisfactory	<p>There were severe shortcomings in the quality of implementation. The lead Agency did not perform its fiduciary responsibilities satisfactorily. Supervision was poor. Emerging concerns were not identified in time, including those that should have been obvious. Although instances of mismanagement were discovered, corrective actions were not undertaken. Program activities were very poorly implemented, and several had to be dropped.</p>	<p>There were severe shortcomings in the quality of implementation of child projects. The Agencies did not discharge their fiduciary responsibilities satisfactorily. Supervision of implementation of a substantial number of child projects was poor. Emerging concerns were not identified in time, including those that should have been obvious. Although instances of mismanagement were discovered, corrective actions were not undertaken. All or almost all child projects were very poorly implemented, experienced delays, and had activities dropped.</p>
Unable to assess	<p>The available information does not allow assessment of the performance of the lead Agency.</p>	<p>The available information does not allow assessment of the performance of the Agencies that implemented child projects.</p>

Table A.2: Program M&E Rating Scale

Rating	M&E plan	M&E implementation
Highly satisfactory	The program M&E plan is robust and does not have any weaknesses. Baseline data are complete. The specified indicators are appropriate, and arrangements for M&E plan implementation are adequate. Overall, the M&E plan exceeds expectations and is exemplary.	M&E plan implementation was excellent. Weaknesses in M&E were addressed in a timely manner. M&E activities were conducted in a timely manner, and data from M&E were used to improve program implementation. The quality of M&E plan implementation exceeds expectations and is exemplary.
Satisfactory	The M&E plan is robust and has no or only minor weaknesses. Baseline data are provided or their collection is planned at program start. The specified indicators are appropriate, and arrangements for M&E plan implementation are adequate. The plan meets expectations.	M&E plan implementation was generally robust. Weaknesses in M&E were addressed in a timely manner. M&E activities were conducted in a timely manner, and data from M&E were used in improving program implementation. Overall, M&E implementation meets expectations.
Moderately satisfactory	The M&E plan is solid overall. It does not have any major weaknesses, although there are some minor weaknesses. The specified indicators are generally appropriate, and arrangements for M&E plan implementation are adequate. The plan meets expectations.	M&E plan implementation was generally robust. Weaknesses in M&E were generally addressed, although some remained. Some M&E activities were delayed. M&E data were used for reporting but had little use in improving program implementation. Overall, M&E implementation meets expectations with some areas of low performance.
Moderately unsatisfactory	The M&E plan is weak overall, although it has strengths in some areas. The specified indicators are generally appropriate but additional indicators are required to adequately capture program results. There are gaps in arrangements for M&E plan implementation. The plan needs several improvements to meet expectations.	M&E plan implementation was weak, although there were strengths in some areas. Most M&E activities were completed, but some were dropped or delayed. M&E data were generally not reported in a timely manner, and use of M&E data to improve program implementation was limited. Overall, M&E implementation does not meet expectations, although there are some areas of adequate performance.
Unsatisfactory	The M&E plan has severe shortcomings. No baseline data are provided nor is there any indication that these would be collected at program start. Indicators do not adequately address program outcomes and other results; relevant indicators have not been specified for several results. There are gaps in arrangements for M&E plan implementation. Either no budget or an inadequate budget has been provided for M&E.	M&E plan implementation was flawed and/or did not address severe weaknesses in the original plan. Several M&E activities were either dropped or were incomplete. The data collection methodology was not sound. M&E data were not reported in a timely manner, and there is little evidence to suggest that the data were used to improve program implementation. M&E implementation does not meet expectations.
Highly unsatisfactory	No M&E plan was prepared.	No, or negligible, M&E activity was implemented other than conduct of the program evaluation.
Unable to assess	Unable to assess because program documents are not available.	Unable to assess because the terminal evaluation does not cover M&E implementation adequately.

References and Resources

GEF (Global Environment Facility). 2006. "[Governance of the Climate Change Funds.](#)" GEF/C.29/05. GEF, Washington, DC.

———. 2008. "[Joint Summary of the Chairs—GEF Council Meeting. April 22–25, 2008.](#)" GEF/C.33/JointSummary. GEF, Washington, DC.

———. 2010. "[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.

———. 2012. "[Clarifying the Responsibilities of the GEF's Key Actors with Respect to the Use of GEF Resources.](#)" GEF/C.42/04. GEF, Washington, DC.

———. 2017a. "[Policy on Gender Equality.](#)" GEF/C.53/04. GEF, Washington, DC.

———. 2017b. "[Policy on Stakeholder Engagement.](#)" GEF/C.53/05/Rev.01. GEF, Washington, DC.

———. 2018. "[Project and Program Cycle Policy.](#)" OP/PL/01. GEF, Washington, DC.

———. 2019a. [Instrument for the Establishment of the Restructured Global Environment Facility.](#) Washington, DC: GEF.

———. 2019b. "[Policy on Minimum Fiduciary Standards.](#)" GA/PL/02. GEF, Washington, DC.

———. 2019c. "[Policy on Environmental and Social Safeguards.](#)" SD/PL/03. GEF, Washington, DC.

———. 2020. "[Guidelines on the Project and Program Cycle Policy.](#)" GEF/C.59/Inf.03. GEF, Washington, DC.

GEF IEO (Global Environment Facility Independent Evaluation Office). 2007. "[GEF Evaluation Office Ethical Guidelines.](#)" Evaluation Document No. 2. GEF IEO, Washington, DC.

———. 2019. "[The GEF Evaluation Policy.](#)" GEF IEO, Washington, DC.

———. 2020. [An Evaluative Approach to Assessing the GEF's Additionality.](#) Evaluation Report No. 139. Washington, DC: GEF IEO.

———. 2022. [GEF Integrated Approach to Address the Drivers of Environmental Degradation.](#) Evaluation Report No. 154. Washington, DC: GEF IEO.

- . 2023. "GEF Guidelines for Conducting Terminal Evaluations of Full-Size Projects." GEF IEO, Washington, DC.
- GEF STAP (Global Environment Facility Scientific and Technical Advisory Panel). 2018. "[Integration: To Solve Complex Environmental Problems.](#)" GEF STAP, Washington, DC.
- . 2021. "[The Circular Economy and Climate Mitigation: A STAP Advisory Document.](#)" GEF/STAP/C.60/Inf.02. GEF STAP, Washington, DC.
- OECD (Organisation for Economic Co-operation and Development). 2021. "[Applying Evaluation Criteria Thoughtfully.](#)" OECD, Paris.
- OECD DAC (Organisation for Economic Co-operation and Development Development Assistance Committee). 2002. "[Glossary of Key Terms in Evaluation and Results-Based Management.](#)" Working Party on Aid Evaluation. OECD DAC, Paris.
- Uitto, Juha I., and Geeta Batra, eds. 2022. *[Transformational Change for People and the Planet: Evaluating Environment and Development.](#)* Cham, Switzerland: Springer.
- UNEG (United Nations Evaluation Group). 2016. "[Norms and Standards for Evaluation.](#)" UNEG, New York.
- . 2020. "[Ethical Guidelines for Evaluation.](#)" UNEG, New York.

The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) was established by the GEF Council in July 2003. The Office is independent from GEF policy making and its delivery and management of assistance.

The Office undertakes independent evaluations at the strategic level. These evaluations typically focus on cross-cutting themes, such as focal area-wide topics or integrated approaches to delivering global environmental benefits. The IEO presents a GEF-wide annual performance report and also undertakes institutional evaluations, such as assessing GEF governance, policies, and strategies. The Office's work culminates in a quadrennial comprehensive evaluation of the GEF.

The Office cooperates with professional evaluation networks on developing evaluation approaches, setting standards, and delivering training—particularly with regard to environmental evaluation and evaluation at the interface of environment and socioeconomic development. We also collaborate with the broader global environmental community to ensure that we stay on the cutting edge of emerging and innovative methodologies.

To date, the Office has produced over 150 evaluation reports; explore these on our website: www.gefio.org/evaluations.

Independent Evaluation Office, Global Environment Facility

1818 H Street, NW Washington, DC 20433, USA

www.gefio.org X @gefio YouTube @gefio



**Independent
Evaluation Office**
GLOBAL ENVIRONMENT FACILITY

