GREEN CLIMATE FUND

Evaluation Operational Procedures and Guidelines for Accredited Entity-led Evaluations

Version 1

March 2023

EVALUATION OPERATIONAL PROCEDURES AND GUIDELINES FOR ACCREDITED ENTITY-LED EVALUATIONS

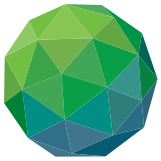
These Guidelines are a living document and will be updated to reflect feedback, lessons learned, as well as policy and operational changes. The GCF Secretariat encourage readers to use the online/electronic version of this document. Please provide any comments or feedback on this document at: evaluation_policy@gcfund.org.

All rights reserved
© Green Climate Fund
Published in March 2023

Green Climate Fund
Songdo Business District
175 Art Center-daero
Yeonsu-gu, Incheon 22004
Republic of Korea

+82 (0)32-458-6039
eda@gcfund.org
greenclimate.fund

Cover page photo: © Green Climate Fund



GREEN
CLIMATE
FUND

Evaluation Operational Procedures and Guidelines for Accredited Entity-led Evaluations

Version 1

March 2023

CONTENTS

Acronyms V

1. Introduction 1

1.1. What the document is for 1

1.2. How to use the document 1

1.3. How the document will be updated 2

2. Accredited Entity-led Evaluations in the GCF 5

2.1. Why evaluate 5

2.2. When to evaluate 7

2.3. What to evaluate 9

3. Planning, Scoping, and Commissioning AE-led Evaluations 13

3.1. Planning an AE-led evaluation 13

3.1.1. Planning during the Funding Proposal development stage 13

3.1.2. Planning for an evaluation during project/programme implementation 16

3.2. Scoping an AE-led evaluation 17

3.3. Commissioning an evaluation 21

3.3.1. Deciding on who will perform the evaluations 21

3.3.2. Commissioning evaluators external to the project/programme and the Accredited Entities 21

3.4. Summary of roles and responsibilities 23

4. Managing the Design and Implementation of AE-led Evaluations 25

4.1. Evaluation inception phase: Overseeing the design of an AE-led evaluation 25

4.2. Evaluation implementation phase: Supervision of evaluation activities prior to drafting the evaluation report 28

4.3. Evaluation implementation phase: Supervision of evaluation activities during the drafting stage 29

4.4. Roles and responsibilities 30

5. Reviewing and Submitting an Evaluation Report to the GCF Secretariat 33

5.1. Accredited Entity review of evaluation reports 33

5.2. Preparation and submission of management response and action plan 36

5.3. Summary of roles and responsibilities 37

Annex 1: Evaluation Matrix 39

1.a Suggested template 39

1.b Evaluation matrix example 40

Annex 2: GCF Secretariat Evaluation Quality Assurance and Review Framework for AE-led Evaluations 47

Annex 3: Technical Methodological Guidance 51

3.a Evaluating reduced GHG emissions/increased carbon sequestration 51

3.b Evaluating adaptation beneficiaries 52

3.c Evaluating the paradigm shift 54

3.d Evaluating the enabling environment 58

Annex 4: Management Response and Action Plan Template 65

Annex 5: Indicative Terms of Reference for Interim and Final Evaluations 69

ACRONYMS

ACRONYM	DEFINITION
AE	Accredited Entity
APR	Annual Performance Report
B.BM	Between Board Meetings
CDM	Clean Development Mechanism
CSO	Civil Society Organization
EE	Executing Entity
FAA	Funded Activity Agreement
FP	Funding Proposal
GCF	Green Climate Fund
GHG	Greenhouse Gas
IEU	Independent Evaluation Unit
IRMF	Integrated Results Management Framework
LOGFRAME	Logical Framework
LORTA	Learning-Oriented Real-Time Assessment
M&E	Monitoring and Evaluation
MAF	Monitoring and Accountability Framework
NDA	National Designated Authority
NGO	Non-governmental Organization
OECD-DAC	Organisation for Economic Co-operation and Development's Development Assistance Committee
PAP	Proposal Approval Process
SAP	Simplified Approval Process
TOC	Theory of Change
TOR	Terms of Reference

1. INTRODUCTION

1.1. WHAT THE DOCUMENT IS FOR

These Evaluation Operational Procedures and Guidelines ('Guidelines') aim to promote a shared understanding of the requirements of the GCF Evaluation Policy as they relate to the delivery of interim evaluations and final evaluations of funded activities.[1] They should be read alongside the **Monitoring and Accountability Framework**, **GCF Evaluation Policy**, **GCF Evaluation Standards** and the **Integrated Results Management Framework**.[2]

The Secretariat developed these Guidelines as mandated by the GCF Evaluation Policy[3] and **B.BM 2021/07**[4], and in consultation with the GCF Independent Evaluation Unit (IEU). Written from a commissioner's point of view, Accredited Entities[5] are strongly encouraged to use these Guidelines when evaluating all funded activities[6] including public sector and private sector adaptation, mitigation, or cross-cutting projects/programmes that were approved from the 32nd Board meeting and onwards.[7]

These Guidelines map out what is required and recommended at each stage of an Accredited Entities-led evaluation ('AE-led evaluation')[8] from planning right through to submission to the GCF to deliver a robust evaluation. They offer a guide for Accredited Entities and other stakeholders[9] when delivering on their evaluation-related roles and responsibilities at various stages of the GCF's Programming Cycle.[10] These Guidelines are not intended to provide exhaustive guidance to independent evaluators on options for design and methodologies for evaluating climate change projects/programmes.

1.2. HOW TO USE THE DOCUMENT

The document has two parts:

i. The **main text**, which presents the policy basis for these Guidelines and is structured chronologically from planning an evaluation to submitting an evaluation report, and

ii. **Annexes**, which contain additional technical guidance and examples.

---

1 *Paragraph 22, GCF Evaluation Policy.*

2 *Additionally, some Accredited Entities follow their established evaluation policy, guidance and processes.*

3 *Paragraph 33, GCF Evaluation Policy.*

4 *Paragraph (b), GCF Evaluation Policy.*

5 *The steps set out in these guidelines can also be applied by the Secretariat on Secretariat-led evaluations.*

6 *These guidelines recognize that Accredited Entities may or may not have a structurally independent evaluation unit, and this organizational particularity, as well as the relevant stipulations in the Accreditation Master Agreement and Funded Activity Agreement, will define how these guidelines are used and the extent to which the recommended steps herein are to be applied.*

7 *As per paragraph 66 of the GCF Evaluation Policy, this Policy will become effective at the end of 2021 or a year after Board approval, whichever is later. The Policy came into effect in May 2022. There will be no retroactive application of the Policy to Funding Proposals approved prior to B.32. However, these guidelines may be used as an overarching guidance by Funding Proposals that pre-date the effectiveness of this Policy.*

8 *Paragraph 22, GCF Evaluation Policy.*

9 *Including national designated authorities (NDAs), civil society organizations, the private sector, and beneficiaries (see paragraph 46, GCF Evaluation Policy).*

10 *Relevant roles and responsibilities are set out in detail in Section VIII of the GCF Evaluation Policy and the programming cycle is described throughout the Programming Manual.*

Readers may refer to the relevant sections and annexes depending on their needs at specific stages of the evaluation or read the document in its entirety. The Guidelines are structured as follows:

Section 2:	Accredited Entity-led Evaluations in the GCF
Section 3:	Planning, scoping and commissioning AE-led evaluations
Section 4:	Managing the design and implementation of AE-led evaluations
Section 5:	Reviewing and submitting an evaluation report to the GCF Secretariat

1.3. HOW THE DOCUMENT WILL BE UPDATED

Co-learning with stakeholders is the cornerstone of evaluative practice in the GCF. As a result of feedback and consultations, this document will be revisited on a periodic basis and revised as needed to reflect learning from practical applications, the changing needs in a constantly evolving programming and operating context and any amendments to GCF strategic documents or frameworks that may have an impact on the conduct of interim evaluations and final evaluations. GCF stakeholders will be notified in the event of a revision.

2. ACCREDITED ENTITY-LED EVALUATIONS IN THE GCF

An **evaluation** is an assessment of an intervention or group of interventions to determine what works and what does not, for whom, how much, why, and under what circumstances.

Appendix I, GCF Evaluation Policy

2.1. WHY EVALUATE

The GCF Evaluation Policy stipulates that evaluations in the GCF 'contribute to the overall performance of the Fund through developing accountability and creating a culture of learning built from producing and using impartial, credible, independent and trusted evidence and evidence-informed strategies and implementation'.[11] Under the principles of relevance, use and participation, all GCF evaluations must be suitable for decision-making, accountability and learning.[12]

The Monitoring and Accountability Framework (MAF) requires Accredited Entities to submit an **interim evaluation and a final evaluation**, collectively defined as AE-led evaluation, for each funded activity. The GCF Evaluation Policy expounds on this MAF requirement[13] and provides the overarching framework to AE-led evaluations including the evaluation principles and criteria and roles and responsibilities.[14] Two associated documents to the GCF Evaluation Policy, the **GCF Evaluation Standards** and these Guidelines,[15] guide the translation of the policy requirements into practical applications and together 'ensure that the Fund is able to inform its overall results, successes and unintended consequences in a credible and measurable manner'.[16]
The evaluation principles (**Figure 1**) and standards (**Figure 2**) underpin all GCF evaluations.

What are evaluation principles, criteria, and standards?

- Evaluation principles are essential requirements of the evaluation process and must be observed when designing and implementing an evaluation as per the GCF Evaluation Policy.
- Evaluation criteria are the basis for assessing a project/programme performance. They serve as a framing device for an evaluation.
- Evaluation standards identify how the quality of an evaluation is going to be judged.

[11] *Paragraph 8, GCF Evaluation Policy.*

[12] *Paragraph 15 (b), GCF Evaluation Policy.*

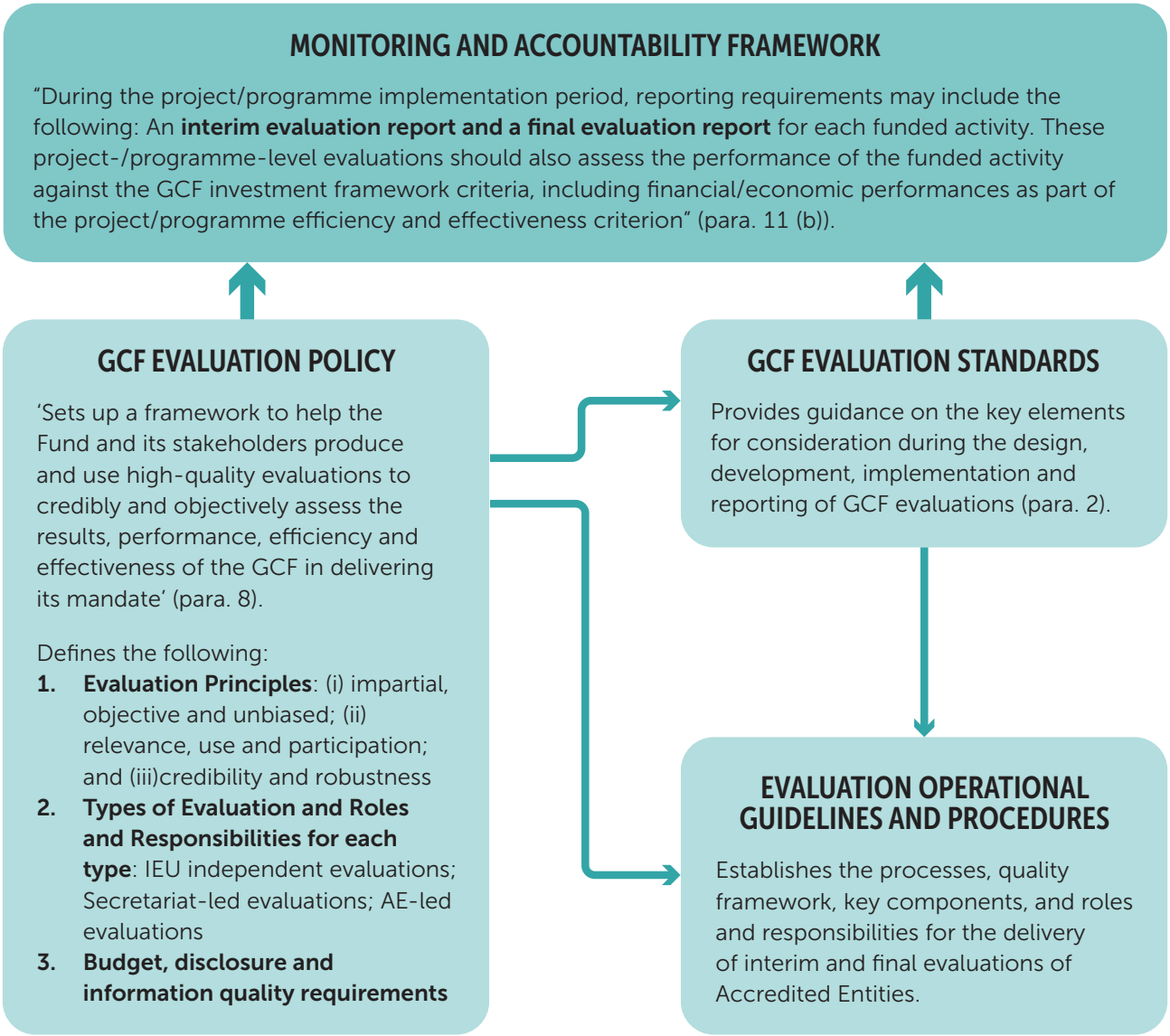
[13] *Paragraph 22, GCF Evaluation Policy; and Paragraph 11(b), MAF.*

[14] *Paragraph 7, GCF Evaluation Policy.*

[15] *Standards apply to evaluations undertaken by the GCF Independent Evaluation Unit, GCF Secretariat and Accredited Entities. These Guidelines only apply to interim and final evaluations from the Accredited Entities, although the guidance may also be useful for Secretariat evaluations.*

[16] *Paragraph 50, GCF Evaluation Policy.*

FIGURE 1. MAF,
GCF EVALUATION POLICY, EVALUATION STANDARDS AND THESE GUIDELINES

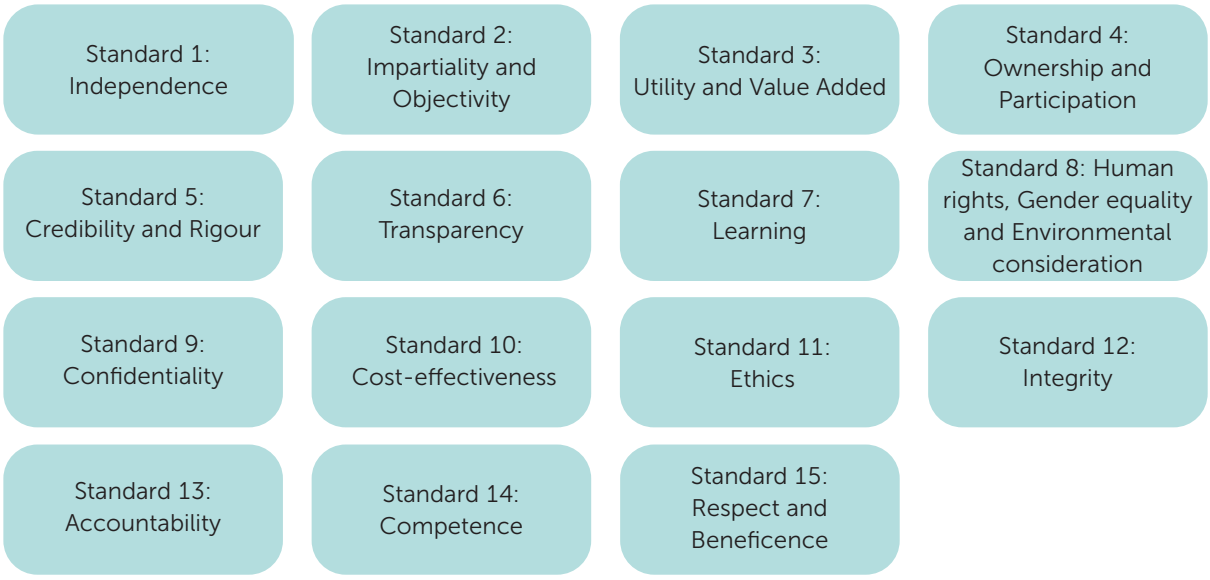


The **GCF Evaluation Standards** provide guidance on the expected quality of an evaluation. The 15 standards – recapped in **Figure 2** – have been developed to 'enable the production of state-of-the-art evaluations with high-quality evidence and recommendations'.[17] Similar to these guidelines, the Standards provide guidance on the design, development, implementation and reporting. By adhering to these standards GCF stakeholders can ensure a degree of consistency across the evaluations conducted.[18]

[17] *Paragraph 2, GCF Evaluation Standards.*

[18] *See Green Climate Fund Evaluation Standards for full details. Appendix 1 and Appendix 2 provide further guidance for applying these standards.*

FIGURE 2. GCF EVALUATION STANDARDS



As an accountability tool, AE-led evaluations allow Accredited Entities to comply with the reporting requirements of the GCF and demonstrate the effectiveness of a funded activity.

As a mechanism for learning and adaptive management, these evaluations inform decisions to improve project/programme implementation and/or future programming directions. As per the **GCF Programming Manual**, evaluations, in particular the interim evaluations, may 'recommend course-correcting adaptive management measures, including budget reallocations, partial or full project cancellations, restructuring of projects and project extensions, as applicable'.[19]

2.2. WHEN TO EVALUATE

> **Interim evaluation** – Evaluation performed **midway** through the implementation of an intervention to assess progress towards and likelihood of achievement of outcomes and impacts. It usually has a strong formative focus.
>
> **Final evaluation** – Evaluation that is **near or at the end of an intervention** to provide evaluative evidence covering the entire intervention. It measures the overall impact, effectiveness, efficiency, sustainability, replicability, and lessons learned of a project.
>
> *Appendix I, GCF Evaluation Policy*

The precise timing for undertaking interim and final evaluations for a particular project/programme is determined by the delivery timelines agreed as part of the Funded Activity Agreement (FAA).[20] More broadly and by the GCF's definition, an interim evaluation will be conducted midway through the implementation of an intervention, whereas a final evaluation will be undertaken near or at the end of project implementation, although there may also be instances where a final evaluation is undertaken two to three years after the end of a project/programme. Any deviations from these timings are discussed and assessed on a case-by-case basis as part of

[19] *P. 63, GCF Programming Manual.*

[20] *Paragraph 22, GCF Evaluation Policy.*

the term sheet negotiations, which will then be reflected in the FAA. The interim and final evaluations have different purposes as elaborated in **Table 1**, and this, amongst other things, highlight the importance of timing. The annual performance reports, generated using monitoring data, serve as one of the information sources of these two periodic evaluations.

TABLE 1. COMPARING INTERIM EVALUATION AND FINAL EVALUATION

	INTERIM EVALUATION	FINAL EVALUATION
TIMING	• To be performed and submitted at the midway point of project/programme implementation or as may be specified in the FAA.	• Near (three to six months) the end of a project/programme, or after project completion depending on the agreements contained in the FAA (usually within three months of submitting the final Annual Performance Report [APR]).
FOCUS	• Formative – get a good picture of current implementation performance and developing evidence that can be fed back into the programme/project to help steer its ongoing implementation towards the desired results. These can be used for improvement.	• Summative – assess the overall performance of the programme/project to better understand the results and why they came about.
PURPOSE	• Develop understanding of likelihood that the project/programme is on track to achieve its intended results or if there are corrective measures that should be considered to inform and to strengthen adaptive management. • Based on the corrective measures required and where needed, inform restructuring/FAA amendment process by providing a solid evidence base for the changes to the design and/or results. • Learning for improvement of implementation and management through unbiased, systematic feedback to identify lessons – from both successes and failures. • Create a platform for dialogue amongst key stakeholders.	• Assess the outputs, adaptation and/or mitigation outcomes and paradigm shift contributions [and enabling environment][21] achieved by the project/programme against the intended results set out in the FAA at the end of the implementation period. • To develop a strong understanding of overall impact, effectiveness, sustainability, and replicability. • Generate insights that explain why and how the intervention results have been achieved or not. • Learning for future programming through unbiased systematic feedback to identify lessons – from both successes and failures – that can be fed back into the GCF portfolio.

[21] *In the IRMF, it is the extent to which GCF has promoted a shift towards low-emission and climate-resilient development pathways in the context of sustainable development and made a significant and ambitious contribution to the global efforts towards attaining the goals set by the international community to combat climate change. The Governing Instrument for the GCF sets out the Fund's mandate to promote a paradigm shift.*

2.3. WHAT TO EVALUATE

AE-led evaluations examine the quality of the design, implementation arrangements and processes, and the climate results of a project/programme. The evaluations shall include the entire funding proposal package approved by the Board as opposed to only the GCF-funded component, or as may be specified in the accreditation master agreement or funded activity agreement.

As per the **Programming Manual**, AE-led evaluations aim to create a body of evidence that informs decision-making and investments. AE-led evaluations assess the extent to which results and impacts are achieved relative to projections that are made as part of the Funding Proposal and help to determine the contribution of the GCF to promoting a paradigm shift. They foster learning from experience and communication amongst stakeholders and deepen the understanding of project/programme context and achievements.

When looking at these elements, Accredited Entities shall apply all the GCF evaluation criteria (defined in **Table 2**) to frame the interim and final evaluations.[22] However, the importance of each evaluation criterion varies between an interim evaluation and a final evaluation. AE-led evaluations should apply proportionality and prioritize properly when identifying evaluation questions for each criterion to ensure that evaluation resources are used efficiently. For example, AE-led evaluations may want to specify less evaluation questions under impact criterion at mid-term because some higher-level results could take time to materialize, and they may not be ready for assessment at that point in time. In contrast, at the end of a project there should be adequate resources for assessing effectiveness, impact and sustainability of the project in its entirety.

TABLE 2. GCF EVALUATION CRITERIA DEFINITIONS AND INDICATIVE IMPORTANCE IN INTERIM AND FINAL EVALUATIONS

GCF EVALUATION CRITERIA	DEFINITION
RELEVANCE, EFFECTIVENESS, EFFICIENCY, IMPACT AND SUSTAINABILITY OF PROJECTS AND PROGRAMMES	The criteria are, as defined by OECD-DAC:[23] • Relevance – the extent to which the project/programme objectives and design respond to beneficiaries', global, country, and partners'/institutions' needs, policies, and priorities, and continue to do so if circumstances change. For GCF-funded projects/programmes, it is important to assess the responsiveness of a project/programme to climate-related needs, and how direct and indirect beneficiaries were identified and targeted are important elements under this criterion. • Effectiveness – the extent to which the project/programme achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups. • Efficiency – the extent to which the project delivers, or is likely to deliver, results in an economic and timely way. For GCF projects/programmes, the degree to which resources have been optimized to maximize the impact of resources spent on climate adaptation and mitigation results. • Impact – the extent to which the project has generated or is expected to generate significant, positive or negative, intended or unintended, higher-level effects. • Sustainability – the extent to which the net benefits of the project/programme continue or are likely to continue.

[22] *Paragraph 17, GCF Evaluation Policy.*

[23] *See Evaluation Criteria - OECD.*

GCF EVALUATION CRITERIA	DEFINITION
COHERENCE IN CLIMATE FINANCE DELIVERY WITH OTHER MULTILATERAL ENTITIES	• The extent to which the project/programme (i) is compatible with and/or complements (and not duplicate) other climate finance funding[24] and (ii) avoids creating parallel systems through creating synergies in funding mechanisms and interventions.
GENDER EQUITY	• The extent to which the project/programme level gender action plans have been implemented and its activities and outcomes are monitored as well as the extent to which key stakeholders have been consulted and their inputs were considered and used to inform design and implementation. • The extent to which the project/programme has addressed vulnerability to ensure equal participation of men, women, children and youth, indigenous people, people with disability and other marginalized groups as well as benefits from the GCF's investments. • The degree to which the project/programme demonstrated understanding on how the impacts of climate change are differentiated by gender and other markers of difference, the ways that behavioural changes, gender and other markers of difference can play in delivering a paradigm shift, and the role that women, indigenous people and other groups play in responding to climate change challenges both as agents but also for accountability and decision-making.
COUNTRY OWNERSHIP OF PROJECTS AND PROGRAMMERS	• The extent to which the project is consistent with the existing national climate strategy, policies, capacity to deliver of implementing entities, intermediaries or executing entities, and engagement with civil society organizations and other relevant stakeholders. This criterion is strongly related to the Relevance criterion.
INNOVATIVENESS IN RESULT AREAS	• The extent to which the project/programme creates opportunities for targeting innovative solutions, new market segments, developing or adopting new technologies, business models, modal shifts and/or processes that are appropriate to a given context and which can bring about a paradigm shift.[25]
REPLICATION AND SCALABILITY	• The extent to which the activities are replicated in other countries or other locations within the same country. • The extent to which the project/programme has expanded its scale and impact without increasing the total costs of implementation.
UNEXPECTED RESULTS, BOTH POSITIVE AND NEGATIVE	• The degree to which unforeseen and unintended results, both positive and negative, and which were not anticipated as part of the planning or articulated in the original theory of change, were captured and from which relevant lessons were identified and utilized to adaptively manage the project/programme and improve project/programme implementation and decision-making.

[24] *Local, national or transnational financing – drawn from public, private and alternative sources of financing – that seeks to support mitigation and adaptation actions that will address climate change (see Introduction to Climate Finance | UNFCCC).*

[25] *The rubric developed by the GCF's Independent Evaluation Unit for the Independent Evaluation of the Green Climate Fund's Investments in Small Island Developing States could provide further guidance to evaluators when assessing this criterion. The rubric dimensions (pp. 72-73) contain the following: (i) Type of innovation (e.g., technology, policy, business model); (ii) Scale of innovation (e.g., central or peripheral to the project design); (iii) Context of innovation (e.g., new to the world or the country); (iv) Intensity (e.g., incremental, radical or instrumental innovation).*

3. PLANNING, SCOPING, AND COMMISSIONING AE-LED EVALUATIONS

3.1. PLANNING AN AE-LED EVALUATION

3.1.1. PLANNING DURING THE FUNDING PROPOSAL DEVELOPMENT STAGE

Planning an interim evaluation and/or a final evaluation starts as early as the Funding Proposal (FP) development stage. As part of the FP package, Accredited Entities are required to provide budgetary information in the Monitoring and Evaluation (M&E) Plan in the annex of a Proposal Approval Process (PAP) or a Simplified Approval Process (SAP) proposal for these two evaluations. The evaluation budgets are then reflected in the detailed project budget and Accredited Entity fees of a PAP or SAP proposal depending on whether the cost is for delivering the evaluations or commissioning additional independent evaluative data collection (see **Table 3** for details).

Table 3 sets out the relevant policy requirements as well as the recommended steps to ensure meeting these requirements during the FP development.

TABLE 3. POLICY REQUIREMENTS AND RECOMMENDED STEPS DURING FP DEVELOPMENT

RELEVANT POLICY AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
GCF EVALUATION POLICY PARAGRAPH 41 IRMF PARAGRAPH 14 AND 30 EVALUATION STANDARD 10	Ensure that financial support for interim and final evaluations is budgeted adequately and allocated, and available in a timely manner. Note that the IRMF requires an independent assessment of the paradigm shift and enabling environment using scorecards.[26] The scorecards would be on top of the usual evaluative activities budgeted for in an evaluation that will impact on the required number of days from evaluators. See Annex 3 for an indication of what is required from the evaluations.	**Consider the potential requirements of the interim and final evaluation.** Some guiding questions may include: • How complex is the operating context, and will evaluation activities require specific types of expertise and logistical arrangements that would be costly?[27] • How in-depth should the evaluations be, considering what is known and unknown during the FP stage? • Are the baseline data and their methodologies well established, and will there be a need to revisit them at mid-term and final evaluations? • Are there any assumptions in the theory of change that will need to be revisited at mid-term and end of project? • How will the scorecard assessments for the paradigm shift potential and enabling environment be built into the evaluations? • How much learning potential is there – e.g., is the project/programme innovative – which would warrant a bigger (or smaller) budget allocation for the evaluative activities?
GCF EVALUATION POLICY PARAGRAPH 41	Interim and final evaluations are covered by AE fees.	**Specify a budget line for interim and final evaluations in the Accredited Entity Fees Annex of the Funding Proposal.[28]** The Accredited Entity Fees Policy specifies eligible expenses under categories such as the 'supervision of mid-term review' and 'supervision of terminal evaluation'. However, to meet the requirements of the GCF Evaluation Policy, Accredited Entities should budget beyond supervision of evaluations and think about how much budget is required to commission an external team to deliver the interim and final evaluations. It is ideal to include a separate budget line (i.e., 'Commissioning of independent evaluators for interim and final evaluations') for this purpose in the Funding Proposal's AE fees calculation. This amount should intuitively and clearly link to the budget in the Evaluation Plan table in Annex 11 of the Funding Proposal of a standard PAP (or equivalent annex in a SAP proposal). Accredited Entities may opt to stick to the known budget lines identified in the AE Fees Policy. Where this is the case, please include an explanatory note that the commissioning of evaluators is included in the budget.

[26] *Please refer to Integrated Results Measurement Framework.*

[27] *For example, complexities can include the diversity of context across population and communities, number of countries where the project/programme will be implemented, cascading uncertainties caused by climate change, coverage of mitigation and adaptation results areas, public versus private sector projects or a combination of financing modalities, etc.*

[28] ***Annex 12** of a standard proposal approval process.*

RELEVANT POLICY AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
GCF EVALUATION POLICY PARAGRAPH 41 AND PARAGRAPH 58(C) AND (D)	Include a budget line for the **generation and collection of evaluative data** for projects/programmes. This may include support for independent data collection (baseline and end-line data), real-time data collection systems and independent analyses to produce high quality evaluations, including impact assessments and evaluations. The budget line 'should be separate from the project management budget (which covers some costs of management results), an overhead line and the AE fee (which covers the costs of interim and final evaluations)'.	**Identify additional evaluative activities required to deliver the interim and final evaluations**, e.g., additional independent data collection, analyses and systems required outside of the typical monitoring activities such as for evaluating the paradigm shift and enabling environment, establishing an independent baseline, annual and end-line data in parallel to monitoring data collected through the M&E system. **Specify a budget line in Annex 4** (Project Budget) of the Funding Proposal of a standard proposal approval process for the additional evaluative activities. Ensure that the breakdown in Annex 11 (M&E Plan) can be intuitively linked with specific rows/columns in Annex 4 (Project Budget) of a standard proposal approval process. Budget line may follow the policy terminologies such as 'independent baseline and end-line data', 'real-time data collection systems', 'independent analysis', or 'impact assessments'. **Note that this budget line is different and separate from typical monitoring activities that generate, manage, analyse, and store monitoring data.** See Box 1 for a sample of costs that may be charged against the project budget or AE fees.
GCF EVALUATION POLICY PARAGRAPH 41	Ensure that the overall evaluation budgets included **within project budgets** (i.e., not AE fees) range from two to five per cent of the project budget.	**Check the adequacy of the budget allocation for evaluation** within the project budget. Consider ways to improve the allocation if it is not within the two to five per cent of the project budget as set out in Annex 4 of the FP package. The intent of this policy clause is to ensure adequacy of evaluation budgets. It is the responsibility of the Accredited Entity to demonstrate this adequacy, and this may include presenting a comprehensive M&E plan that links routine monitoring activities and periodic evaluative activities. See Box 1 for a sample of costs that may be charged against the project budget or Accredited Entity fees.

3.1.2. PLANNING FOR AN EVALUATION DURING PROJECT/ PROGRAMME IMPLEMENTATION

More detailed planning should be done before the launch of an interim evaluation or a final evaluation. The GCF Evaluation Policy does not have any specific requirements related to this stage. However, additional steps such as those specified in **Table 4** can be undertaken to ensure that planned AE-led evaluations will be able to adhere to policy requirements.

The main objective of planning for an evaluation during project/programme implementation is to set a realistic evaluation timeline, consider the cost requirements at a more granular level and prepare the right staffing to successfully commission, manage, complete, and deliver an evaluation.

Please note that Table 4 provides recommended steps only as there is no strict policy provision around these steps.

TABLE 4. RECOMMENDED STEPS FOR PLANNING AE-LED EVALUATIONS DURING IMPLEMENTATION

RELEVANT STANDARD	RECOMMENDED STEPS
	Set the evaluation timeline • **Check the contractual deadline for the evaluation report.** The Funded Activity Agreement (FAA) should be the main point of reference for when evaluation reports should be submitted to the Secretariat. Knowing the submission date will help ascertain the key dates (i.e., terms of reference preparations, commencement date, etc.) for the evaluation. It is important that there is sufficient time for preparations and undertaking the evaluation activities. • **Consider the procurement timeline.** Ideally, potential vendors will have at least a month after the terms of reference are published to send questions to the commissioner, respond to the request for proposals and deliver a good quality submission.
EVALUATION STANDARD 10	**Revisit the approved budget in the Funding Proposal.** The Funding Proposal will only contain a high-level, aggregate budget. At this planning stage, it will be extremely useful for Accredited Entities to think through the budget in a more detailed manner (see box 1), also considering new information and lessons that have come to light in the early implementation.
EVALUATION STANDARD 1, STANDARD 5, STANDARD 6, STANDARD 13	**Identify management structure of the evaluation**, including roles and responsibilities. Some guiding questions when considering the most appropriate management structure include: • Will there be an evaluation manager who will be tasked to draft the terms of reference, get internal approval for the terms of reference, the selection of a team of evaluators, managing the team of evaluators, etc.? If not, who will be responsible for the delivery of the evaluation and managing all the intermediate steps leading to the submission? • Who will be responsible for quality assurance? Is there a need to constitute an evaluation reference group or a quality assurance expert panel? • Will there be a need for additional support for compiling data and documents and logistics from the Accredited Entity's end? • Does the selected management structure protect the independence of the evaluators and promote clear lines of responsibility?
EVALUATION STANDARD 6	**Create an evaluation workplan.** Accredited Entities may prepare an evaluation workplan to detail the necessary steps and their schedule to deliver high-quality evaluations in a timely manner. Workplans may and can change so it is best to consider them as a live document. The external evaluator, when commissioned, will also provide inputs to this workplan.

BOX 1. COST CONSIDERATIONS FOR DETAILED BUDGETING FOR AN INTERIM OR A FINAL EVALUATION

Costs to be incurred during an evaluation may include professional fees, transportation costs, communication costs, translation costs, catering, venue hires, contingency costs and others that are essential to deliver evaluation activities.[29] Accredited Entities may also want to build in contingencies to the budget, in cases where evaluation reports require additional data collection or rewriting work to troubleshoot the quality of work of the evaluators.

An indicative suggestion under which these activities could be provided for (i.e., project budget (PB) or AE fees (AEF)) is included on the list.

Evaluation Inception Phase

- Inception workshop to prioritize and revise the evaluation questions (AEF)
- Desk-based review of project/programme documents (AEF)
- Meetings and interviews to inform the drafting of the inception report (AEF)
- Drafting and finalization of the inception report (AEF)
- Meetings to present/discuss the inception report and approach to the evaluation manager and project managers (AEF)
- Quality assurance of draft and final inception reports (AEF)

Evaluation Implementation Phase

- In-depth desk-based review of project/programme documents (AEF)
- Data collection – interviews, group discussions, household surveys (AEF and/or PB)
- Pilot and pre-testing of data collection instruments (PB)
- Data storage – if there is real time, independent data collection (PB)
- Field visits and validation of reported results (AEF and/or PB)
- Analysis – data cleaning and editing, encoding, coding, triangulation, quality assurance, etc. (AEF and/or PB)
- Drafting, revisions and finalization of the evaluation report (AEF)
- Quality assurance of the draft and final evaluation reports (AEF)
- Workshop to prioritize recommendations with the commissioner (AEF)
- Workshop of the draft evaluation findings and recommendations with the commissioner (AEF)
- Submission to the GCF Secretariat (or optionally, together with a presentation to the GCF Secretariat) (AEF)
- Review of GCF Secretariat's feedback and revision of the report (AEF)
- Drafting and submission of the Management Response and Action Plan (AEF)
- Submission of the revised evaluation report to the GCF Secretariat (AEF)

3.2. SCOPING AN AE-LED EVALUATION

Determining the scope of an evaluation is important to draft the terms of reference (see **Annex 4** for suggested content of a terms of reference) that could set out clearly what commissioners are looking for. Identifying the precise needs and requirements from an evaluation in the terms of reference will help getting high-quality responses from the market.

[29] *Please see Section 5 for further details on the inception and implementation phases of an evaluation.*

TABLE 5. POLICY REQUIREMENTS AND RECOMMENDED STEPS WHEN SCOPING AN AE-LED EVALUATION

RELEVANT POLICY AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
EVALUATION STANDARD 3, STANDARD 4, STANDARD 6		**Check quality of the theory of change and logical framework in the Funding Proposal** The theory of change and logical framework will serve as the backbone for the interim and final evaluations. If the results and indicators, as well as the pathways to change for achieving such results, are not clearly stated, then Accredited Entities should consider how to get them to a standard that would be useful for a meaningful evaluation. One way of doing this is by including in the scope of the evaluation the review and reconstruction of the theory of change as well as relevant methodologies such as for GHG emissions reduced and adaptation beneficiaries during the evaluation inception phase. A thorough analysis of the logical framework is expected to be a part of the interim and final evaluations. Both these activities will need to be participatory, including the project team, Executing Entity (EE), National Designated Authorities (NDAs) and other delivery partners. See box 2 for quality parameters when interrogating the theory of change and logical framework.
EVALUATION STANDARD 3, STANDARD 4, STANDARD 7, STANDARD 13		**Identify relevant stakeholders' accountability and learning needs.** The identification process could be in the form of a workshop or interview with the project team and ideally other stakeholders such as the EEs, Implementing Partners and NDAs. Some guiding questions may include but are not limited to: • What does the FAA require the AE to deliver as part of the interim or final evaluation? Are there additional conditions attached to the interim or final evaluation, for example? • What are the knowledge and information gaps that need to be filled to address ongoing implementation challenges and help the project/programme achieve the intended results? (If delivering an interim evaluation.) • What does the project/programme team need to learn to improve implementation at mid-term or for future programming at end of implementation?
GCF EVALUATION POLICY PARAGRAPH 17 EVALUATION STANDARD 3	Use the GCF evaluation criteria as a framework to organize the questions and ensure that the list is comprehensive but focussed. Note that the GCF evaluation criteria shall be applied to all AE-led evaluations.	**List evaluation questions based on the accountability and learning needs.** It is normal that the first set of evaluation questions is long. However, prioritization is vital to a streamlined and focussed evaluation and the efficient use of evaluation resources. Sample evaluation questions per evaluation criterion are in Annex 5.

RELEVANT POLICY AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
GCF EVALUATION POLICY PARAGRAPH 61 EVALUATION STANDARD 5	Ensure that high-quality data are stored and will be made available to evaluations.	**Identify additional data collection, analytical and storage needs.** Check the quality and quantity of monitoring data vis-à-vis the information requirements of the evaluation questions to identify information gaps. These gaps can be filled by commissioning additional data collection as part of the terms of reference. Additional data will have implications on analytical requirements and how the additional information will be stored, which would then require additional budget. Accredited Entities should consider how monitoring data and findings will be verified (e.g., desk-based, field, etc.) to broaden and strengthen the evidence base of evaluations.
GCF EVALUATION POLICY PARAGRAPH 5(A) EVALUATION STANDARD 11	Ethical considerations should be considered as early as possible in the process.	**Consider the ethical requirements of the evaluation** vis-à-vis the list of evaluation questions. At this scoping stage, knowing the potential ethical implications (e.g., if the evaluation activities put the safety of beneficiaries or the evaluators at risk) of the evaluation questions that users deem relevant can guide in prioritizing the evaluation questions as well as stipulating exact ethical requirements in the terms of reference.
EVALUATION STANDARD 10, STANDARD 4, STANDARD 5		**Revisit the revised budget from the planning stage.** The number of evaluation questions of interest and the additional data collection and verification required to address them will impact on the evaluation budget. Spreading a limited budget across too many evaluation questions is highly discouraged. Estimate the number of staff-days required to deliver the evaluation based on the required depth to answer the priority evaluation questions and the additional data collection work to be commissioned. Consider various data collection and analytical options that could deliver the most robust evidence in the most cost-effective manner.
GCF EVALUATION POLICY PARAGRAPH 15 (A)-(D), PARAGRAPH 17 ALL EVALUATION STANDARDS	AE-led evaluations shall uphold the GCF evaluation principles and apply the GCF evaluation criteria.	**Draft the Terms of Reference (TOR) and collate supporting documents** The TOR will, at a minimum, need to contain a project/programme description, scope and objectives, evaluation questions, methodological guidance/preferences (while retaining flexibility), deliverables, team requirements, timeframe, management structure, and other essential requirements that would clearly set out the expectations for the evaluators. The TOR must reflect the requirement to apply the evaluation principles and criteria. The TOR should be detailed but at the same time retain flexibility on the evaluation methodology. Drafting a TOR, if the recommended steps in this table are to be implemented, could take as long as four to six weeks. This timeframe could be shorter or longer depending on the amount of required information that is known or unknown to the Accredited Entity.

RELEVANT POLICY AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
GCF EVALUATION POLICY PARAGRAPH 15 (A)-(D), PARAGRAPH 17 ALL EVALUATION STANDARDS		Accredited Entities may share, on a completely optional basis, the draft terms of reference to the GCF Secretariat for review. The Secretariat's feedback will be limited to sense-checking and will be non-binding and will take no more than a week within receipt of the draft terms of reference. Annex 5 presents a TOR template for commissioning interim and final evaluations.

BOX 2. CONSIDERATIONS WHEN CHECKING LOGICAL FRAMEWORK AND THEORY OF CHANGE QUALITY

Accredited entities and the commissioned evaluators should ensure that the following (non-exhaustive) lines of enquiries are considered when checking the quality of the logical framework and theory of change:

Logical framework

- Are the outcome indicators consistent with the definition and methodology specified in the IRMF? Why/why not?
- Are the results specified at the right level (i.e., outputs do not measure activities, outcomes do not measure outputs, etc.), and co-benefits specified properly?
- If reporting core indicator 1, does the logframe disaggregate Greenhouse Gas (GHG) reporting by technology or result area?
- If reporting on core indicator 2, does the logframe clearly identify beneficiaries and provide disaggregation of the number of people supported by gender and other potentially important levels of disaggregation?
- Are the targets in relation to the baseline appropriate? If the project/programme has overachieved or is overachieving, is it because of efficiency and effectiveness or underestimation during the design stage?
- Is the AE undertaking (or has it undertaken) the means of verification specified in the FAA logframe? Did the expected data sources materialize?
- Are assumptions/risks identified and internalized within the logframe? Are they being monitored?
- Is there a logical progression in the results achieved, e.g., reporting states that outcomes have been overachieved when the outputs have not been met, and if not, has this been explained?

Theory of change

- Does the theory of change logically present the pathway(s) to the changes that the project/programme intends to achieve?
- Is it clear how the activities and outputs bring about the adaptation, mitigation or cross-cutting outcomes that relate to GCF result areas as well as the intended co-benefits, and how these outcomes could contribute to the desired contribution to the paradigm shift?
- Is the project/programme implementation consistent with the theory of change? Why/why not?
- What revisions to the theory of change may be required to make it more accurate and reflective of implementation realities?
- Which of the risks, barriers and assumptions hold/do not hold? How are they affecting the achievement of results?

3.3. COMMISSIONING AN EVALUATION

3.3.1. DECIDING ON WHO WILL PERFORM THE EVALUATIONS

The GCF evaluation principles of impartial, objective, and unbiased[30] should guide Accredited Entities when looking to commission an evaluation. These principles require independence, which means that evaluations are free from external influence and bias in their design, selection, frameworks, data collection, analysis, findings, conclusions, and recommendations. Selecting evaluators that have no prior involvement in the design of the project/programme under evaluation or its implementation and that will have no possibility of future contracts for the project/programme under evaluation is key to ensuring the independence of an evaluation.

The GCF Evaluation Policy recognizes varying institutional arrangements for independent evaluations across Accredited Entities.[31] Some Accredited Entities will have a structurally independent evaluation office and others will not. The mandate of these structurally independent evaluation offices also varies as they relate to project/programme evaluations: Some undertake these evaluations or perform a quality assurance function only, while others do not have the mandate to have any direct involvement.

In cases where an independent evaluation office of an Accredited Entity is only allowed to quality assure project/programme evaluations or has no involvement in these evaluations, commissioning external evaluators is recommended. The same applies to Accredited Entities without a structurally independent evaluation office.

As such, who does what in terms of delivering the evaluation activities will depend on the broader evaluation architecture within each Accredited Entity.

3.3.2. COMMISSIONING EVALUATORS EXTERNAL TO THE PROJECT/ PROGRAMME AND THE ACCREDITED ENTITIES

Evaluation is a well-established area in international development and policy. There are freelance evaluation experts and/or firms specializing in delivering independent evaluations. The designated evaluation manager can start linking with the market through national evaluation associations that usually have a broad range of individual and firm memberships.

Procurement process and timing. The selection process for external evaluators (if this is the feasible route to take) will be defined by each Accredited Entity's internal procurement process. To get the best results, it is important that the selection is based on an open and competitive process that provides sufficient response time to potential candidates. It will also be useful if the relevant staff, for example the evaluation manager, can identify beforehand some websites where the request for proposals can be published widely.

Requests for proposals should be launched as early as possible as previously mentioned. Whilst the timeframes will be defined by internal systems and bandwidth to process bids in a timely manner, tenders should ideally be advertised for at least a

[30] *Paragraph 15 (a), GCF Evaluation Policy.*

[31] *Paragraph 43-44, GCF Evaluation Policy.*

month. Considering internal approval processes, the entire selection process could take as long as 8 to 12 weeks.

Request for proposal structure. Accredited Entities can explore two options in the design of a request for proposal. The choice between them will depend on Accredited Entities' needs and preferences.

1. Launch two distinct calls for the interim evaluation and the final evaluation. This option can provide Accredited Entities with a fresh perspective at two distinct implementation stages.
2. Commission one evaluation partner for the duration of the project/programme to undertake both the interim evaluation and the final evaluation, as well as all the required additional independent collection of evaluative data and information. This option offers continuity to the Accredited Entity and may present opportunities for cost savings through bulk discounts from the bundle of evaluation services to be provided. However, evaluators should be able to demonstrate impartiality when they assess the quality of recommendations from the interim evaluation during the final evaluation.

Individual evaluator or a team of experts. Commissioners also have the option to procure a sole evaluator, a newly constituted team of evaluators or a firm or partnership of firms to supply the evaluation services. This choice will depend on the complexity and size of the evaluation and the requirements of the TOR, and AEs' internal processes and experience. The GCF Secretariat Evaluation Quality Assurance and Review Framework in **Annex 2** reflects the requirements of the GCF Evaluation Policy. This framework could help guide the selection process. Regardless of this decision, the potential service provider should be independent of project/programme management and implementation. This means that the evaluator or the members of a team should not have any prior involvement in the design and day-to-day implementation of the project/programme. AEs may require evaluators to sign non-disclosure forms on conflict of interest due to past or future engagement with the project and other potential sources of conflict of interest.

An evaluation of a cross-cutting GCF project/programme will most likely require a combination of expertise in evaluation, GHG accounting, adaptation, sector (i.e., forestry/ecosystems/agriculture/infrastructure etc.), gender and social inclusion, and qualitative and quantitative approaches depending on the evaluation design. As it is not possible for only one expert to have all the requisite expertise, a team of evaluators (newly constituted or a firm) instead of a sole expert will be ideal.

Local expertise. Importantly, it is good practice and consistent with the GCF Evaluation Standard (Standard 14) to include local expertise as an integral part of the evaluation team. Local experts bring to the table significant contextual, institutional, and cultural knowledge, which is critical to producing a nuanced and respectful (Standard 15) evaluation.

BOX 3. TEAM SELECTION CHECKLIST

- Is there proven expertise in undertaking/leading evaluations?
- Is the proposed team experienced in designing, using, and evaluating M&E tools such as theory of change and logical framework?
- Is the proposed team completely independent of project/programme management and implementation, and if not, has potential conflicts of interest been declared and mitigation procedures proposed?
- Is there a good balance of male/female and local/international experts on the team?
- Is the team able to deliver the proposed approach? E.g., if the proposal commits to a mixed methods approach, there must be expertise in qualitative and quantitative skills on data collection, analysis and reporting within the team.
- Can the team interrogate gender and social inclusion considerations beyond data disaggregation in reported results?
- Can the proposed evaluators demonstrate excellent writing skills (i.e., a copy of past evaluations can be requested)? If possible, ensure that references have been checked.
- Is there technical expertise in the sector of interest, in mitigation and/or adaptation results, and contextual knowledge of the country/region?

3.4. SUMMARY OF ROLES AND RESPONSIBILITIES

STAKEHOLDER	PRIMARY GCF EVALUATION POLICY RESPONSIBILITIES
ACCREDITED ENTITIES	• Adequately plan for and scope an evaluation • Allocate adequate time and budget for interim and final evaluations • Consult relevant stakeholders to ascertain accountability and learning needs which the evaluation must address • Establish a robust and independent quality assurance process • Draft a TOR that applies the GCF Evaluation Principles and Evaluation Standards • Manage the request for proposal if procuring external evaluators • Procure providers of evaluation services if AEs do not have an independent evaluation office, or if the independent evaluation office is not structurally allowed to undertake project/programme-level evaluations
EXECUTING ENTITIES	• Participate in consultations to shape the TOR
GCF SECRETARIAT	• Where an Accredited Entity opted to share the TOR with the Secretariat, review of draft TOR and provide non-binding feedback in a timely manner
BENEFICIARIES/ PROJECT STAKEHOLDERS	• Where possible, participate in consultations to shape the TOR
NATIONAL GOVERNMENTS AND AGENCIES	• Where possible, participate in consultations to shape the TOR

4. MANAGING THE DESIGN AND IMPLEMENTATION OF AE-LED EVALUATIONS

Once an evaluation has been commissioned (i.e., internally if through a structurally independent evaluation office or externally through an open competitive procurement exercise), commissioners will begin managing the evaluation.

There is no single approach to managing a project/programme evaluation.[32] Typically, commissioners ensure that evaluations are systematic and manageable by dividing tasks and activities into clear phases. One common approach is to split the evaluation activities into an evaluation-specific inception phase and an implementation phase.[33] The inception phase can include all the design and preparatory work for an evaluation, and the implementation phase is where the evaluation activities identified during inception are to be undertaken. Regardless of who undertakes the evaluation (i.e., a structurally independent evaluation office or external evaluators), the recommended steps in the following sections may provide guidance to commissioning Accredited Entities on how to manage project/programme evaluations.

4.1. EVALUATION INCEPTION PHASE: OVERSEEING THE DESIGN OF AN AE-LED EVALUATION

The evaluation inception phase is a valuable investment of time in systematically preparing the evaluation. This is the phase when the Accredited Entity based upon their needs and the requirements of the GCF agree with the evaluators on the scope, conduct and other elements of the evaluation. Therefore, it is important that there is adequate time allotted for the inception phase. Depending on the size of the project/programme and the evaluation requirements, the inception phase will, at the minimum, take two working weeks, but preferably more. The inception phase for a final evaluation

[32] *Some Accredited Entities follow their own established evaluation guidelines and procedures.*

[33] *These phases are specific to the interim and final evaluations and not to be confused with the project/programme inception and implementation phases.*

could be longer if it involves complex methodologies for assessing the climate results at the outcome level.

The design of an evaluation, often contained in an evaluation inception report (see also **Annex 5** for a suggested outline), is a critical first deliverable of the evaluators. It is during this stage that commissioners and evaluators agree on the final scope, evaluation questions, methodology and more. Hence, commissioners should ensure that ample time for discussions is allocated for this stage.

Please note that table 6 provides recommended steps only for commissioners, as there is no strict policy provision for the design of an evaluation.

TABLE 6. RECOMMENDED STEPS FOR OVERSEEING THE DESIGN OF AN EVALUATION

RELEVANT STANDARD	RECOMMENDED STEPS
	Hold a kick-off meeting. The first step after the evaluators' contract has been put in place is to hold a kick-off meeting. This meeting, usually called and convened by the evaluation manager and held either virtually or in-person, is where the evaluators and the project team discuss the TOR, further clarify expectations from the evaluation and agree on the next steps. The meeting could be held at least a week after the appointment and onboarding of the evaluators. Typical issues for a kick-off meeting include: 1) logistical and administrative requirements to help facilitate the conduct of the evaluation (communication, schedules, roles and responsibilities, contract, invoicing, reporting, approval of reports, dissemination, communication of evaluation results, etc.) and 2) technical specifics of the evaluation based on the TOR and the winning bidder's proposal including but not limited to the purpose and scope of the evaluation as well as the requirements of GCF Evaluation Policy. If the evaluation is being undertaken for several co-financiers (e.g., other climate funds, regional banks, private sector, etc.), it will also be useful to remind the evaluators of the relevant evaluation requirements as set forth in the relevant policies. The kick-off is a good opportunity to discuss resourcing priorities to ensure successful completion of the evaluation as outlined in the TOR.
EVALUATION STANDARD 5	**Compile background, project/programme, and other relevant documentation.** Availability of and access to relevant background documentation is an important pre-requisite for a successful evaluation. Therefore, the compilation and sharing of background documents should be one of the first tasks when preparing an evaluation. When done ahead of the kick-off, this step can aid an efficient start for the evaluator. Relevant documentation will include the FAA, the FP, APRs, and any other internal documentation such as baseline reports, market studies, technical and financial reports, and others. All relevant GCF policy documents, such as the GCF Evaluation Policy and IRMF, should be provided so that the evaluator understands the requirements. During the evaluation, the evaluator will be responsible for requesting and gathering additional documents.
EVALUATION STANDARD 4	**Generate and share stakeholder list.** The Accredited Entity shall provide the evaluator with a list of stakeholders that should be contacted for inputs (e.g., for interview or as part of focus groups). The first email linking the evaluators with the stakeholder may be done by the Accredited Entity for introductions and to seek support on the evaluation activities. After this, the Accredited Entity should not be copied into the communications to avoid influencing the outcome of the data collection.

RELEVANT STANDARD	RECOMMENDED STEPS
ALL EVALUATION STANDARDS	**Review the Evaluation Inception Report.** Accredited Entities will need to review and sign off the inception report, subject to satisfactory completion. An inception report is usually completed within four to six weeks of evaluation kick-of, but it may also take longer. Note that Accredited Entities are not required to submit the inception report to the GCF. However, this is a very important report as it will set out the approach to how the evaluation questions will be answered and a robust evaluation report will be delivered. Hence, it is important that the review of the inception report check how the proposed evaluation approach responds to the requirements of the TOR. Depending on the selected management structure, Accredited Entities may want to request the evaluation reference group or quality assurance panel to review the report. See a suggested inception report template in Annex 5.
	Hold an inception meeting. This meeting will take place once the evaluator has submitted the draft inception report and the Accredited Entity has reviewed this report. The meeting will offer an opportunity to review the contents of the inception report in greater detail and allow for exchange and discussion on any differences of view between the commissioner and the evaluator.

Broad methodological considerations that should be included in discussions during the inception phase are in **Box 4**. In addition to a detailed description of the evaluation approach and methodology, an inception report should also systematically capture these considerations against each evaluation question in an evaluation matrix (see **Annex 1**).

BOX 4. METHODOLOGICAL CONSIDERATIONS FOR COMMISSIONERS

- **Data collection methods:** The main methods for data collection and instruments should be designed during the inception. To meet the policy requirements of rigour, evaluations will ideally employ both qualitative and quantitative methods in a mixed methods approach. There should be a strong emphasis on data disaggregation. During inception, evaluators should also consider the quality and availability of baseline data. Methods (see table 7) should maximize the participation of relevant stakeholders. Accredited Entities may also want to consider how to integrate surveys for Learning-Oriented Real-Time Impact Assessment (LORTA) with other evaluative needs if participating in the IEU's LORTA programme.
- **Limitations:** Evaluations should be clear on the potential limitations in the availability of data as well as any limitations affecting the data analysis. For example, there may not be any reliable data available, baseline data may not have been collected, sample groups may not be fully representative, or contextual factors such as security may limit outreach to important stakeholders. These constraints should be noted during inception, monitored during implementation, and explained in the evaluation report, including how they were mitigated.
- **Data collection instruments:** Instruments for data collection (e.g., interview protocols or questionaries) that are suggested by the evaluator should be pre-tested before deployment.
- **Data analysis:** As part of the inception, the evaluator should be clear on how different sources will be compared and critically assessed for reliability (triangulation), and then how the combined information will be used as the basis for conclusions, leading to recommendations and lessons learned. Analysis of quantitative and qualitative data requires different approaches, and the evaluator should make clear how this will be undertaken, and how discrepancies and inconsistencies will be clarified to arrive at well balanced conclusions drawn from accurate data and information.

TABLE 7. EXAMPLES OF DATA COLLECTION METHODS AND INFORMATION SOURCES

	METHOD	SOURCE OF INFORMATION
QUANTITATIVE	Statistical analysis of national and publicly available data	• Official government statistics (national and subnational levels) • Other statistical data from Civil Society Organizations (CSOs)s and Non-governmental Organizations (NGOs)
	Statistical analysis of project data	• Funding Proposals • Baseline data • Monitoring data • Annual Performance Report • Data from other interventions in the same geography
	Surveys	• Questionnaires that are distributed to a pre-defined sample of potential respondents
	Physical measurements	• Relevant measurements of results
QUALITATIVE	Focus group discussions	• Facilitated discussions guided by the evaluation questions with different stakeholder groups
	Structured observation in the field	• Site visits and systematic collection of observational data in a structured way
	Semi-structured interviews	• Semi-structured interview questions to be devised to guide interviews with key project stakeholders
	Desk review of documents	• Internal project documents, Funding Proposal, gender action plan, environmental and social safeguards report, research papers (published and grey), other evaluations

4.2. EVALUATION IMPLEMENTATION PHASE: SUPERVISION OF EVALUATION ACTIVITIES PRIOR TO DRAFTING THE EVALUATION REPORT

The role of Accredited Entities. Following satisfactory completion of the inception phase, the evaluator will undertake the evaluation in accordance with the methodology in the terms of reference and agreed through the inception report based around the evaluation questions to be answered. The role of Accredited Entities is to facilitate the smooth implementation of the evaluation through timely and comprehensive responses to the evaluators' queries and provision for data and documentation. Such facilitation is important to the success of the evaluation.

The specific roles and responsibilities within the evaluation team will have been set out clearly as part of the inception phase. Broadly speaking, it falls under the purview of the evaluator to ensure that all evaluation tasks are completed to a high standard based on the parameters set by the TOR and further refined during inception.

Regular touchpoints with the evaluators. It is good practice to build into the evaluation processes (through the inception report) regular check-ins at a pre-defined frequency, for example, every two or four weeks depending on the need. These touchpoints will be standing meetings for the duration of the evaluation to allow for any issues and interesting findings to be picked up and discussed.

BOX 5. SUGGESTED AGENDA POINTS FOR CHECK-INS WITH EVALUATORS

- **Access to the correct stakeholders:** Are the evaluators able to access the stakeholders to achieve the required rigour, inclusivity, and participation?
- **Methodological discussion:** Are the selected methods, both quantitative and qualitative, appropriate or are there other methods that would yield better and more credible results? Does the combination of methods allow for a triangulation and validation of results?
- **Ethical considerations:** Are the evaluators adhering to the ethical commitments set out in the inception report?
- **Discussion on timing:** Is the evaluation proceeding as planned and on time or are there likely to be delays and why?

To protect an evaluation's independence during implementation:

- Regular meetings should be structured as check-ins on the evaluators' progress and not to influence emerging evaluation results.
- Accredited Entity staff, GCF staff and other co-financiers **must not** participate in the data collection process as interviewers or observers in an interview or group discussion to ensure that there is no undue influence on the evaluation. All interviews must be confidential.

Evaluation de-brief prior to drafting. The evaluators must ensure that there is sufficient time allocated for analysis of the evaluative data, as well as ample time for discussion of initial findings, conclusions and preliminary recommendations emanating from the evaluation. The outcome of this process will likely take the form of a summary presentation of findings, conclusions, and lessons at a full debriefing workshop. This should be scheduled by the evaluator after completion of the evaluation, and once ample time has elapsed for analysis. This will allow for the Accredited Entity to view the early results of the evaluation, and then discuss in more detail the information that will be included in the evaluation report. It also provides the opportunity to identify recommendations and consider possible prioritization prior to full drafting of the evaluation report. It is recommended that at least half a day be allocated for this debriefing workshop.

4.3. EVALUATION IMPLEMENTATION PHASE: SUPERVISION OF EVALUATION ACTIVITIES DURING THE DRAFTING STAGE

Responding to clarifications and requests for additional evidence. During the drafting stage, Accredited Entities should remain available to respond to queries that might arise whilst evaluators are collating, analysing, and drafting their findings. It is normal for evaluators to identify information gaps at this stage, which would require them to get back to stakeholders for clarification.

Securing early stakeholder buy-in on recommendations. Recommendations are to be independently formulated by the independent evaluators. An early presentation of recommendations to Accredited Entities by the evaluators, for instance, for prioritization purposes, facilitate ownership and uptake of the recommendations. Such a process will feed into the timely production of management response and action plan (see Section 5), a requirement of the Evaluation Policy.

4.4. ROLES AND RESPONSIBILITIES

STAKEHOLDER	PRIMARY GCF EVALUATION POLICY RESPONSIBILITIES
ACCREDITED ENTITIES	• Compile background documents and stakeholder list • Schedule kick-off meeting • Review inception report through internal quality assurance mechanism • Sign off on inception report • Participate in check-in meetings • Participate in debriefing workshop and provide comments to the inception report • Participate as resource person(s) for interviews or group discussions (but not as interviewers)
EXECUTING ENTITIES	• Assist evaluators in the smooth implementation of evaluation • Participate as resource person(s) for interviews or group discussions (but not as interviewers) • Participate in debriefing workshop and provide comments to the inception report
GCF SECRETARIAT	• Participate in evaluation activities, e.g., as a resource for interview or survey, depending on the project/programme evaluation approach
GCF INDEPENDENT EVALUATION UNIT	• Receive a copy of the interim and final evaluations
INDEPENDENT EVALUATORS	• Draft and deliver Inception Report • Undertake evaluative data collection activities in line with the agreed methodology in the Inception Report • Design and deliver interim evaluations • Design and deliver final evaluations • Schedule and facilitate check-in meetings • Schedule and facilitate debriefing workshop • Ensure participation of all relevant stakeholders throughout the evaluation process
BENEFICIARIES/ PROJECT STAKEHOLDERS	• Participate in evaluation activities, e.g., as interviewees or survey respondents, as determined by project/programme evaluation approach
NATIONAL GOVERNMENTS AND AGENCIES	• Participation in evaluation activities, e.g., as interviewees or survey respondents, as determined by project/programme evaluation approach

5. REVIEWING AND SUBMITTING AN EVALUATION REPORT TO THE GCF SECRETARIAT

5.1. ACCREDITED ENTITY REVIEW OF EVALUATION REPORTS

A formal and thorough review of evaluation reports prior to submission to the GCF Secretariat is pertinent to the delivery of high-quality and robust evaluation. Hence, Accredited Entities should ensure that evaluators set aside adequate time for review and revisions ahead of the submission date and after the feedback has been received from the GCF. The review process also provides a useful opportunity for dialogue between the Accredited Entity and evaluators on the validity and accuracy of the results.

It is the responsibility of Accredited Entities to review the report in its entirety. **Annex 2** presents the GCF Secretariat's Quality Assurance and Review Framework for evaluation reports. The framework demonstrates what the Secretariat will be looking for when reviewing the reports. However, Accredited Entities may also use its internal quality assurance framework and processes and find ways to integrate the additional requirements of the GCF.

Review process. Ahead of the review, Accredited Entities may set up a structured review process or follow existing processes (e.g., request a structurally independent evaluation office to perform its quality assurance function if allowed) to ensure its credibility and transparency. The review should ensure that the report is of a high quality with reliable and credible findings.

The Accredited Entity may also constitute and request an evaluation reference group or quality assurance panel to concurrently review the report, if so constituted. Additional reviewers, outside of the Accredited Entity, will help incorporate different perspectives on the evaluation.

Accredited Entities may refer to the GCF's quality assurance framework in **Annex 2**, which sets out the elements that the Secretariat consider essential to a high-quality evaluation report. A sample outline for an evaluation report is presented in **Annex 5** to guide Accredited Entities when reviewing the comprehensiveness of a report.

Evaluation managers typically oversee the review process and consolidate the comments, clarifications and/or requests for further information from various stakeholders and share them with the evaluation team through a separate log of comments that specify the relevant section, page number, comments and actions from the evaluation team. Should there be any disagreement between the evaluation team and the project team or the Accredited Entity, the management response mechanism shall be used to respond to the points under contention.[34] The Accredited Entity must not force the evaluation team to change any parts of the report due to these disagreements.

BOX 6. HIGH-LEVEL QUALITY PARAMETERS TO CHECK IN AN EVALUATION REPORT

- Overall content, structure and flow
- Precision of language used on findings
- Validity of the findings and any potential evidence gaps
- Assessment of mitigation, adaptation, enabling environment and paradigm shift results (see Annex 3)
- Triangulation of results and clarity of methodological limitations
- Balance of perspectives
- Sensitivities in the key messages from the evaluation

To protect an evaluation's independence when reviewing the report:

- As a general principle, comments and feedback should be limited to correcting factual errors, inaccuracies and the quality of evidence that underlies the findings, conclusions and recommendations.
- Accredited Entities **must refrain** from pressuring the evaluators to change their interpretations and conclusions on project/programme performance and results unless based on erroneous information and/or singular and unvalidated opinion.

[34] *Some Accredited Entities will opt to use an established dispute resolution mechanism to handle any disagreements related to the evaluation.*

TABLE 8. POLICY REQUIREMENTS AND RECOMMENDED STEPS WHEN REVIEWING AND SUBMITTING AN EVALUATION REPORT TO THE GCF

RELEVANT POLICY CLAUSE AND STANDARD	POLICY REQUIREMENT	RECOMMENDED STEPS
MAF 11(B) GCF EVALUATION POLICY PARAGRAPH 15 EVALUATION STANDARD 2, STANDARD 4, STANDARD 5, STANDARD 8, STANDARD 11, STANDARD 12	The project-/programme-level evaluations should assess the performance of the funded activity against the GCF investment framework criteria, including financial/economic performances as part of the project/programme efficiency and effectiveness criterion. The evaluation report must demonstrate that it upheld the GCF evaluation principles and standards and applied the evaluation criteria.	**Review the scope of the evaluation and all the requirements of the TOR**, including ethical requirements, to ensure that there are no gaps, and that findings and conclusions are comprehensive. In addition to the GCF's quality assurance review framework (Annex 2), Accredited Entities should use the TOR, the approved Inception Report and particularly the evaluation matrix. It is good practice to **share the report with beneficiaries**, to give those with local expertise/knowledge and relevant civil society organizations the opportunity to review and comment (commercial and confidentiality sensitivities notwithstanding). Where required, Accredited Entities should consider translation of key findings into local languages to improve the accessibility of the report. **Closely examine the findings, lessons, and recommendations.** These are key elements of the report. Ensure that the report provides adequate, robust analysis of whether the project/programme has achieved the intended results, and why or why not. A clear analysis as part of the report that links findings, conclusions and recommendations is the basis for using the evaluation for management response. **Compile the feedback and send to the evaluator** for discussion and integration into the report as appropriate. The evaluator should maintain a clear audit trail of stakeholders' comments and responses.
MAF, 11(B) EVALUATION STANDARD 1 GCF EVALUATION POLICY PARAGRAPH 31	An interim evaluation report and a final evaluation report for each funded activity must be submitted by the Accredited Entity. The Secretariat is responsible for receiving, reviewing, and assuring the quality of interim and final evaluations. The IEU shall also receive a copy of the evaluation reports (the mechanism of which shall be agreed by the Secretariat and the IEU).	**Submit the interim evaluation report or final evaluation report** via the Portfolio Performance Management System (AEs will be informed in case this changes). Evaluation reports must be submitted ahead of the FAA deadline to allow for review and revisions. Final evaluation reports must be submitted by the FAA deadline. It is recommended that Accredited Entities build time into the evaluation plan and budget for the evaluator to provide support beyond just completing the evaluation report before submission to the GCF. There is likely to be comments and feedback from the GCF Secretariat that will need to be addressed by the evaluator and/or warrant a revised evaluation report. The Secretariat will thoroughly review the evaluation report once and provide feedback to the Accredited Entity within a three-week turnaround time, based on the Quality Assurance and Review Framework in Annex 2. Following this, the Secretariat will review the Accredited Entities' second submission based on its response to the Secretariat's comments as well as the accompanying management response and action plan within a week of receipt of the revised evaluation report and management response and action plan.

BOX 7. CONSIDERATIONS FOR FORMULATING RECOMMENDATIONS

- Are the recommendations anchored in the conclusions set out in the report?
- Are the findings and recommendations clearly distinguished?
- Do they include proposed improvements, changes, any actions to remedy issues in the performance (formative and summative recommendations) or recommendations to capitalize on and leverage the strengths of the project/programme?
- Are the recommendations specific and have a clear indication of:
- To whom do the recommendations apply?
- Who is responsible for implementing the recommendations?
- What are the costs, benefits, and challenges associated with implementing the recommendations?
- What is the timeline for the implementation (immediate, short-term, medium-term to longer-term) of the recommendations?
- Are they focussed on actions within the reasonable scope/purview of the Accredited Entity?
- Are they sensitive to how they will impact different groups and stakeholders?
- Are they formulated in easy-to-understand language?
- Are they ordered based on priority?

All AE-led evaluations will be made public through the GCF website as the policy requires. The publication will be in accordance with the requirements of the GCF Information Disclosure Policy. The Secretariat will create a mechanism to systematically share evaluation reports on a publicly available database. With this in mind, AEs should consider highlighting sensitive information in the report that may be in violation of the relevant information disclosure policy to signify to the Secretariat what information should not be published, confine such information in a confidential annex or submit a redactable version for website upload.

5.2. PREPARATION AND SUBMISSION OF MANAGEMENT RESPONSE AND ACTION PLAN

The GCF Evaluation Policy[35] requires Accredited Entities to prepare and submit a management response and action plan to address recommendations from AE-led evaluations. A management response and action plan is a formal mechanism to ensure that the evaluation findings, conclusions and recommendations are utilized for an adaptive management of projects, mid-course corrections, iterative learning and learning across the GCF portfolio through considered action plans and follow-up actions.

The management response and action plan will be submitted to the Secretariat 30 days after the receipt of the Secretariat's feedback on the evaluation report. Requests to extend this period will be considered on a case-by-case basis and will be negotiated as part of the funded activities agreement. The Secretariat will have the opportunity to review the management response and action plan and will raise any follow-up queries within a week after the receipt of the management response and action plan.

[35] *Paragraph 42, GCF Evaluation Policy.*

Using the template in **Annex 4**, Accredited Entities will (1) include a response to the Secretariat assessment that will typically use the framework in **Annex 2**, and (2) set out which of the recommendations are accepted and how they will subsequently move forward. To respond to the Secretariat assessment, Accredited Entities may wish to enlist the assistance of their evaluators for input and include this as a formal deliverable.

- Where there is only partial agreement or disagreement, Accredited Entities will provide a justification for any partial acceptance or rejection of Secretariat feedback and evaluation recommendations.
- For evaluation recommendations that have been accepted, or partially accepted, key follow-up actions will be identified, with a time frame specified and the responsibility for implementation clearly established. A priority level will be assigned to each recommendation and follow up action.

The management response and action plan will be included as part of the final version of the interim evaluation or final evaluation submitted to the GCF. Accredited Entities shall report on its progress in implementing the identified actions through the annual progress reporting cycle. Where the recommendations lead to an amendment of the funded activity agreement, the monitoring will be built into the restructuring process and checked by the relevant portfolio manager and restructuring specialist in the GCF Secretariat.

5.3. SUMMARY OF ROLES AND RESPONSIBILITIES

STAKEHOLDER	PRIMARY GCF EVALUATION POLICY RESPONSIBILITIES
ACCREDITED ENTITIES (AES)	• Ensure that the evaluation report is in line with GCF requirements and is of a high quality through a review of the report and liaising with the evaluator to complete the report. • Identify confidential information and submit the report to the GCF Secretariat in a timely manner as per the FAA. • Prepare a management response and action plan in response to the recommendations contained in the report and the Secretariat's feedback.
BENEFICIARIES/PROJECT STAKEHOLDERS	• Review the evaluation report and provide feedback prior to submission to the GCF Secretariat.
GCF SECRETARIAT	• Receive and review the evaluation report in compliance with the Evaluation Policy. • Provide feedback on the evaluation report to the Accredited Entities within three weeks to help finalize the report and the AE's management response. • Manage public disclosure of the evaluations, ensuring compliance with the Evaluation Policy and Information Disclosure Policy, and consistency with the relevant provisions of the Accredited Master Agreement.
GCF INDEPENDENT EVALUATION UNIT	• Receive evaluation reports and relevant data.
INDEPENDENT EVALUATORS	• Draft the report in line with the TOR, including adequately responding to the evaluation question. • Ensure the data informs the findings, which will lead to logical conclusions and strong recommendations and conclusions, as well as lessons. • Consult with the Accredited Entity to ensure that the evaluation report findings, conclusions and recommendations are well understood to avoid any misinterpretation. • Revise the report as required.

ANNEX 1: EVALUATION MATRIX

The evaluation matrix forms a central analytical framework for an evaluation. It sets out how evaluation questions per evaluation criterion will be addressed. It breaks down the main questions into sub-questions and maps out how data will be collected and analysed to answer these questions.

The evaluation matrix also serves as an organizing tool to help plan the conduct of the evaluation, indicating where secondary data will be used and where primary data will need to be collected. It guides analysis, ensures that all data collected are analysed and triangulated, which would then support the identification of evidence gaps. As such, the evaluation matrix ensures that the evaluation design is robust, credible, and transparent.

The evaluators commissioned to undertake an interim evaluation, or a final evaluation, are encouraged to complete an evaluation matrix and include it as part of the inception report. Evaluation matrices should be updated during the evaluation and again included as an annex of the final evaluation report.

1.A SUGGESTED TEMPLATE

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE	DATA SOURCE AND COLLECTION METHODS	ANALYTICAL PROCEDURE	STRENGTH OF EVIDENCE
See Section 2.3 for the definition of the GCF evaluation criteria.	*Ensure each evaluation question is specific.*	*For each evaluation question, specify a limited and focused number of sub-questions that will be essential to answer the evaluation question. This is a good way of breaking down a question to manageable parts for a more systematic way of addressing the question.*	*The indicators and measures determine how performance or progress is judged for each sub-question. Indicators should be realistic in terms of data collection within the scope of the evaluation. Measures/ indicators should be clear and measurable (either qualitatively or quantitatively) and correspond to the evaluation question or sub-question.*	*Identify sources of the required information, such as databases, studies, subject area experts, programme officials, models, etc.*	*Identify the method(s) to be used to collect the required information or data. Examples include case studies, focus groups, perception survey, beneficiary survey, benchmarking to best practices, use of existing databases, etc.*	*Documents how all the collected data is analysed to ensure they can answer the evaluation questions. This can include regression analysis, statistics, qualitative analysis. Analytical methods should be appropriate to use for the given data that is collected.*	*Based on the available data and their quality, indicate the strength of evidence base for each evaluation question (i.e., strong, fair, weak).* *Elaborate the limitations if applicable.*
Relevance							
Effectiveness							

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE	DATA SOURCE AND COLLECTION METHODS	ANALYTICAL PROCEDURE	STRENGTH OF EVIDENCE
Efficiency							
Impact							
Sustainability							
Coherence in climate finance delivery with other multilateral entities							
Gender equity							
Country ownership of projects and programmes							
Innovativeness in result areas							
Replication							
Scalability							
Unexpected results, both positive and negative							

1.B EVALUATION MATRIX EXAMPLE

This completed example aims to demonstrate the typical content of an evaluation matrix. It uses a fictitious cross-cutting project, a Mass Transit System. A brief description is provided in what follows. The content of this example matrix is **indicative and non-exhaustive**. Accredited Entities should always adapt the contents to the purpose of their evaluation and may choose to expand more on some criteria than on others.

PROJECT TITLE	MASS TRANSIT SYSTEM
PROJECT BACKGROUND	The project will deliver 30 km of fully segregated low-emission and climate resilient bus rapid transit (BRT) infrastructure including cycle lanes, a bike sharing system, last-mile connectivity with e-pedicabs, and improved pedestrian facilities benefitting 1.5 million residents within city x. Biogas for the project's zero-GHG emission biomethane buses will be produced from cattle waste. The project includes restructuring the public transport network and a fleet-scrapping programme, as well as a compensation mechanism. It shifts passengers towards public and Non-Motorized Transport (NMT) and implements a BRT system powered completely by biomethane. The BRT detailed design caters for a projected increase in the city temperature and intense heatwaves and events of intense precipitation along the BRT route and makes the public transport system less vulnerable to climate risks. It will benefit the city's population through increasing access to climate-resilient, low-carbon, reliable and safe public transport. Other benefits for the population include improved air quality, time savings, reduced vehicle operating costs, and universal access for women, children, and the disabled at all stations and in buses, including segregated areas for women.

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE AND COLLECTION METHODS	ANALYSIS PROCEDURE/ TRIANGULATION	DATA AVAILABILITY/ RELIABILITY
Relevance **Note:** If undertaking an interim evaluation, the question should be about the continuing relevance of the project/ programme, which means revisiting the assessments made by the Accredited Entity at entry. It is not enough to assess how the Accredited Entity assessed the needs and consistency with government policies and priorities during the design stage. During the final evaluation, the question should centre around the project/ programme's relevance upon completion.	How responsive is the mass transit system to the needs of the population?	To what extent does the mass transit system address the barriers to safe and reliable transport systems in city x?	Percentage of beneficiaries who feel that the mass transit system is addressing their needs.	Beneficiary survey covering a representative sample of inhabitants of the area covered by the intervention (quantitative).	Quantitative survey data analysis, disaggregation by different population groups.	Strong.
	To what extent is the mass transit system relevant to the government's policies and priorities?	To what extent is the mass transit system in line with the government's policies and priorities? How does the intervention fit in or complement other interventions in the area?	Extent to which the mass transit system aligns with government policies and priorities. Extent to which the mass transit system complements other interventions. (On the scale of 1 = not at all; 2 = moderately; 3 = fully)	Key informant interviews with government officials (qualitative).	Qualitative analysis of interview transcripts, triangulation of information given by different respondents and respondent types.	Strong.

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE AND COLLECTION METHODS	ANALYSIS PROCEDURE/ TRIANGULATION	DATA AVAILABILITY/ RELIABILITY
Effectiveness	How effective is/was the intervention in providing the population with access to a safe and reliable transport systems?	To what degree did the project create a modal shift from passenger vehicles to BRT?	**IRMF core indicator 2:** Direct and indirect beneficiaries reached (number of individuals, disaggregated by gender, age). (Quantitative indicator)	Statistics on people using bike sharing; data on numbers of travellers provided by BRS administrators (e.g., ticket sales statistics); demographic data on areas reached by cycle and pedestrian lanes.	Adding numbers of people using BRT and bike sharing services (accounting for a certain percentage of overlap) to determine direct beneficiaries; counting inhabitants of areas reached by intervention as indirect beneficiaries.	Fair – number of direct beneficiaries may be higher than those buying tickets since travelling without a ticket is common. Indirect beneficiaries include those of secondary benefits such as cleaner air which is difficult to quantify as this may affect inhabitants not directly connected to the transport network, too.
			Number of beneficiaries who perceive the BRT as safe and reliable.[36] (Quantitative indicator)	Survey of a representative sample of users of the BRS.	Survey analysis. Triangulation with KIIs with female beneficiaries (see gender equity criterion). Focus group discussions with a group of users to understand the factors that contributed to the shift in behaviour.	Strong.
	How effective is the intervention in helping city x meets its GHG emission reduction targets?	How much tCO_2eq were avoided due to the BRT operationalization?	**IRMF core indicator 1:** GHG emissions reduced, avoided or removed/ sequestered (tCO_2eq). (Quantitative indicator)	Bus emission data.	Calculations based on average emission data of old buses x the buses that were replaced with new buses.	Strong.
		How many tCO_2eq were avoided over the lifespan of the intervention? (Contribution)	**IRMF core indicator 1:** GHG emissions reduced, avoided or removed/ sequestered (tCO_2eq). (Quantitative indicator)	CO_2 emission data for city x.	Following MRV methodology.	Strong.

[36] *This overlaps with the indicator on gender equity 'Women's perception of safety on the BRT'. Data for these two indicators will be used for triangulation.*

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE AND COLLECTION METHODS	ANALYSIS PROCEDURE/ TRIANGULATION	DATA AVAILABILITY/ RELIABILITY
Efficiency	How efficient is the intervention in providing the population of city x with access to a safe and reliable transport system?	To what extent did the project deliver the intended outputs within budget and schedule?	Cost per unit (per beneficiary) in comparison to similar interventions.[37] (Quantitative indicator)	Calculations based on administrative data and project management data.	Efficiency is assessed against benchmarks from other comparable interventions and/or over time within the project.	Fair – limited number of similar interventions that can be used as benchmark. Comparison over time to be used as supplementary data source.
			Efficiency of co-operations, governance, and learning. (Qualitative indicator)	Document reviews and key informant interviews or workshops conducted during the evaluation.	Qualitative analysis.	Strong.
	How efficient is the intervention in reducing GHG emissions?	In comparison to similar interventions, how efficient is the intervention in reducing GHG emissions?	Cost per unit (per tCO_2eq avoided) in comparison to similar interventions. (Quantitative indicator) Fuel consumption data (Quantitative indicator)	Calculations based on CO_2 emission data supplied by BRT administrator.	Efficiency is assessed against benchmarks from other comparable interventions.	Fair – limited number of similar interventions that can be used as benchmark.
Impact	To what extent did the project/ programme contribute to a paradigm shift in city x towards a low emission and sustainable development pathway?		Overall narrative from the paradigm shift scorecard findings.			Strong.
Sustainability	How likely is it that the net benefits of the project/ programme will continue beyond implementation?	Has the programme generated changes in behaviour? Has the programme created an enabling environment for climate resilient practices? Has the programme set up the structures necessary to sustain benefits over time?	Paradigm shift score on Sustainability.	Surveys with beneficiaries. Interviews with government officials.	Survey and interview analysis, scorecard approach suggested by the GCF to assess enabling environment.	Strong.

[37] *Where external benchmarks cannot be found, comparison over time can show whether the programme is becoming more efficient.*

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE AND COLLECTION METHODS	ANALYSIS PROCEDURE/ TRIANGULATION	DATA AVAILABILITY/ RELIABILITY
Coherence in climate finance delivery with other multilateral entities	To what extent is the intervention coherent with the climate finance delivery of other multilateral entities? And climate funds?	To what extent is the programme compatible with and/or complements (and not duplicates) other climate finance funding? Does the programme avoid creating parallel systems through creating synergies in funding mechanisms and interventions?	Compatibility of intervention with other climate interventions (score from 1 to 5).	Government documents on climate projects in the target area; interviews with Ministry of Environment officials.	Desk-based document analysis; interview analysis	Strong.
Gender equity	To what extent has the intervention lead to more gender equity?	To what extent has the intervention lead to more gender equity in access to safe and reliable public transport?	Percentage of BRT passengers who are women. (Quantitative indicator)	Data on numbers of travellers provided by BRS administrators.	Quantitative analysis, disaggregation of beneficiaries by gender.	Fair – BRT does not capture traveller data disaggregated by gender. A methodology for doing so will have to be revised.
			Women's perception of safety on the BRT. (Qualitative indicator)	Key informant interviews with female passengers of the BRT.	Interview analysis.	Strong.
Country ownership	To what degree is the project consistent with the capacity to deliver implementing entities?	To what extent does the intervention build and use local capacity?	Percentage of national/local project staff. (Quantitative indicator)	Project management data.	Quantitative analysis	Strong
			Capacity building efforts integrated into the project implementation. (Qualitative indicator)	Key informant interviews with project staff.	Interview analysis.	Strong.
Innovativeness in result areas	To what extent does the project create opportunities for targeting innovative solutions, new market segments, developing or adopting new technologies, business models, modal shifts and/or processes that bring about a paradigm shift?	Has the intervention made use of innovative technology?	Degree to which the intervention has made use of innovative technology. (Qualitative indicator)	BRT data.	Comparison of BRT technology with other available technology.	Strong.

EVALUATION CRITERIA	EVALUATION QUESTION	SUB-QUESTIONS	INDICATOR	DATA SOURCE AND COLLECTION METHODS	ANALYSIS PROCEDURE/ TRIANGULATION	DATA AVAILABILITY/ RELIABILITY
Replication	To what degree can activities be replicated in other countries or other locations within the same country?	Have lessons from this intervention been generated and disseminated to inform replication in other cities/countries/ sectors?	Number of generated knowledge products/reports/ presentations about this intervention. (Quantitative indicator)	Project management data.	Desk-based document review.	Strong.
		Have lessons from this intervention been perceived as useful and taken up by other actors?	Perception of usefulness and uptake of knowledge products /reports/ presentations by other actors. (Qualitative indicator)	Key informant interviews with other actors.		
Scalability	Does the programme have the potential for expanding its scale and impact without increasing the total costs of implementation?		Paradigm shift score on Scalability.	Economic, demographic, price data; interviews; secondary publications	Economic analysis, narrative analysis.	
Unexpected results, both positive and negative	To what degree has the project generated (or is expected to generate) significant positive or negative, intended, or unintended, higher-level effects? Does the intervention integrate M&E as a key element of project design and implementation to identify and respond to emerging or unintended problems or risks and/or capitalize on positive unintended results?	Are there any positive or negative unintended mitigation and/ or adaptation results that were generated as an outcome of the project? Has the intervention been able to identify and respond to unexpected consequences?		Document reviews, interviews.	Thematic comparison.	Strong.
			Responsiveness of the programme to unexpected results and ability to course-correct. (Qualitative indicator) Mechanisms for beneficiary feedback.	M&E Framework, Logframe, learning strategy M&E Framework, feedback collection processes	Assessment of the M&E framework and its integration into the project design; assessment of the learning strategy. Assessment of feedback collection processes.	Strong.

ANNEX 2: GCF SECRETARIAT EVALUATION QUALITY ASSURANCE AND REVIEW FRAMEWORK FOR AE-LED EVALUATIONS

PROJECT NAME	
EVALUATION TYPE	
ACCREDITED ENTITY	
REVIEWERS	

OVERALL ASSESSMENT

ASSESSMENT CONSIDERATIONS	GCF SECRETARIAT COMMENTS	GCF SECRETARIAT RECOMMENDATIONS
1. IMPARTIALITY, OBJECTIVITY AND BIAS MITIGATION		
Independent roles and responsibilities • Are the roles and responsibilities of the stakeholders involved (e.g., commissioner, evaluator, NDA, EE, etc.) in the evaluation and how this delineation of roles facilitated/ impeded the evaluation's independence clearly described? • Is it clear who undertook the evaluation and how it was undertaken?		
QA processes • Is there an independent quality assurance process put in place either by the independent evaluator or the commissioning AE?		
Mitigation of bias • Did the evaluation report confirm how it ensured that the various aspects of its evaluation, such as design, framework, data collection, analysis, findings, conclusions, and recommendations are free from external influence and bias?		

ASSESSMENT CONSIDERATIONS	GCF SECRETARIAT COMMENTS	GCF SECRETARIAT RECOMMENDATIONS
Ethical standards • Did the evaluation report explain the ethical considerations within the evaluation, and how the evaluators adhered to ethical standards during the design and implementation of the evaluation? • Was it clear how the evaluation upheld the GCF's policies and procedures, i.e., the Gender Policy, Indigenous Peoples Policy, Policy on the Prevention and Protection from Sexual Exploitation, Sexual Abuse, and Sexual Harassment? i.e., • Did the evaluators ensure participation and representation of different groups of beneficiaries such as women, children, indigenous people, etc.? • Was there free, prior and informed consent amongst data collection partners? • Did the evaluation ensure anonymity and non-attributability of statements in the report? • Did the evaluators consider potential adverse impacts of the evaluation activities on partners (e.g., burden on time from data collection activities, safety of women and children, etc.)?		
2. RELEVANCE, USE AND PARTICIPATION		
Relevance • Did the report provide adequate description of the intervention to be evaluated? • Did the product adequately describe and analyse the logical framework and theory of change?		
Use and participation • Was there a well-defined dissemination and knowledge management plan, which identifies target users and specific ways to reach them? • Does the timing of the evaluation maximize benefits for stakeholders and encourage participation of relevant stakeholders? • Have issues of gender equity and participation at all levels been considered in the selection of stakeholders? • Did the evaluation justify where the data collection did not involve a wide range of stakeholders that could provide different perspectives on the evaluation question?		
3. CREDIBILITY AND ROBUSTNESS		
A. Evaluation questions • Do the evaluation questions address GCF's evaluation criteria? • Are the questions sufficiently specific, and do they address the evaluation objectives?		
B. Evaluation protocol/approach/matrix Was the evaluation based on a clear protocol, approach, or evaluation framework/matrix that provides the line of sight to commissioners from the evaluation design to analysis and findings?		

ASSESSMENT CONSIDERATIONS	GCF SECRETARIAT COMMENTS	GCF SECRETARIAT RECOMMENDATIONS
C. Evaluation methodology • Is the proposed methodology adequately explained and justified? • Was the proposed methodology applied as described to address the evaluation questions satisfactorily? • Does the methodology allow for the complementation of quantitative and qualitative methods and triangulation of data and sources? If not, was a clear explanation and justification for this provided? • Did the report explain the methodological limitations and their impact on the quality of the evaluation? How were these limitations and their impacts mitigated?		
D. Analysis and recommendations • Did the evaluation use high-quality independent and relevant data and independent analyses? • Are the recommendations fit-for-purpose and useful for the primary users of the evaluation?		
4. STRUCTURE AND CLARITY OF THE EVALUATION		
• Is the report structured logically and written in an accessible manner? • Does the structure include relevant elements of the report (i.e., executive summary, relevant annexes, sources etc.)? • Is the executive summary of reasonable length and does it communicate the essential and key messages of the evaluation?		

ANNEX 3: TECHNICAL METHODOLOGICAL GUIDANCE

3.A EVALUATING REDUCED GHG EMISSIONS/INCREASED CARBON SEQUESTRATION

Due to the specific mandate of the GCF, and in line with the newly adopted Integrated Results Management Framework (IRMF), Evaluation Policy, as well as the guidance by the GCF Board on the steps for enhancing the climate rationale (B.33/13), the GCF encourages the incorporation of elements linked to evaluations (or verification) of the GHG emissions reductions as part of the interim and final evaluations conducted by AEs in line with the Evaluation Policy and the MAF. Specified at the outcome level, the reduced GHG emissions/increased sequestration assessment will typically fall under the Effectiveness criterion and therefore associated questions shall be specified within the evaluation matrix.

OVERALL APPROACH

The interim and final evaluations of any mitigation or cross-cutting project funded by the GCF shall include an evaluation of the actual GHG emission reductions achieved by the project, as well as an assessment of the potential of the project to achieve the lifetime emission targets that were originally established in the funding proposal or any agreements with the Fund.

The Fund is not recommending the use of any verification standards as part of this assessment. However, it is strongly recommended that a standard linked to the methodology used for the estimation of GHG emission reductions and their reporting is used. Alternatively, references should be made to international good practices for GHG emission verification, including those under the Clean Development Mechanisms or Joint Implementation of the Kyoto Protocol, Article 6 of the Paris Agreement, voluntary carbon markets and other, while taking into consideration project-specific conditions.

Each interim evaluation and final evaluation are encouraged to cover the following aspects:

- Examine whether the mid-term and final targets, as well as lifetime targets provided in the log-frame, are based on a robust methodology, and whether this methodology is being followed by the project/programme M&E.
- Verify the data and their sources used for reporting GHG emissions in APRs. This may include desk reviews of relevant documents, GPS data and others.
- Where the project/programme involves multiple data sources, evaluators may employ sampling approaches. Examples of cases where sampling is appropriate include transport projects, cookstove projects, rooftop solar panel projects and others.
- Consider/use actual activity data of the project/programme (e.g., amount of electricity generated, actual reforested area, etc.). Approaches where the emission reductions are linked only to the amount of funding disbursed are not considered good practice.
- Cross-check primary data with other data sources and identify any potential disparities.
- Undertake site visits, particularly for multicounty projects or projects involving many data sources, users, etc.
- Assess project/programme performance against mid-term target, final target or lifetime target and provide an analysis of whether the project/programme is on track or not, and why or why not.

With the above recommended verification activities, evaluation reports will contain the following elements:

- a description of the overall GHG verification methodology
- a summary of the evaluation approach
- a qualified opinion on the veracity of the GHG emission reduction reported in APRs, as well as the projected emission reductions until the end of the project lifetime as defined in the funding proposal or legal agreements with the Fund
- an annex with a recalculation of the GHG emission reduction results as well as a spreadsheet with the recalculation of the annual GHG emission projections. All data sources and assumptions shall be clearly stated.

EVALUATION TEAM

It is strongly recommended that each evaluation team for mitigation or cross-cutting projects includes at least one member who has experience in GHG accounting and verification of GHG emission reductions.

3.B EVALUATING ADAPTATION BENEFICIARIES

Mandated by the Integrated Results Management Framework (IRMF), measurement of adaptation beneficiaries is a key requirement for adaptation and cross-cutting projects funded by the GCF. Project/programme evaluations have a role to play in ensuring that the reported adaptation beneficiaries results are robust. Specified at the outcome level, the adaptation beneficiaries assessment will typically fall under the Effectiveness criterion, and therefore associated questions shall be specified within the evaluation matrix.

As the interim and final evaluations refer to different stages of a project cycle, it is important to devise an analytical framework to distinguish them for the evaluation to assess project/programme performance in view of adaptation impacts.

It is the responsibility of the evaluators to:

- Validate the reported adaptation beneficiaries reached at mid-term or end of the project, and any lifetime projects where relevant.
- Collect additional evaluative evidence and data and provide illustrative examples of adaptation beneficiaries.

The GCF encourages evaluators to use this guidance when interrogating the adaptation beneficiaries results reported by the project/programme in an interim evaluation and a final evaluation.

EVALUATING ADAPTATION BENEFICIARIES IN AN INTERIM EVALUATION

At mid-term, the logical framework of adaptation and cross-cutting projects/programmes will typically include a target for Core Indicator 2: Number of beneficiaries reached. As the reported figure of this indicator comes from the project/programme monitoring, the role of the evaluation at this stage of implementation is to check whether the methodology for calculating and defining adaptation beneficiaries across the selected GCF results areas is being applied appropriately and whether the required disaggregation (e.g., direct vs. indirect, by sex, etc.) exists. To strengthen the credibility of the reported results, it is also ideal for evaluators to collect additional data and information as a form of results verification.

BOX 8. CHECKLIST FOR EVALUATING ADAPTATION BENEFICIARIES AT MID-TERM

- Examine whether the mid-term and final targets provided in the project's log-frame are based on a robust methodology, and whether this methodology is being followed by the project/programme M&E.
- Assess whether a proper vulnerability assessment was carried out with clear outline of the adaptation and resilience needs in the project area as part of the funding proposal development.
- Assess whether the adaptation benefits expected to accrue are identified and defined, and if they were appropriate.
- Assess the relevance of the materiality threshold applied for defining an adaptation benefit.
- Assess whether direct and indirect adaptation beneficiaries are clearly defined.
- Assess whether the direct and indirect adaptation beneficiaries are mapped against those adaptation benefits expected to accrue during the project.
- Assess whether the adaptation beneficiaries are disaggregated across adaptation result areas of the IRMF (in case the project covers more than one adaptation result area).
- Assess whether the adaptation beneficiaries are disaggregated by sex.
- For multi-country projects – assess whether the steps outlined above are followed for **each** country.
- Evaluate the project/programme performance on reaching the number of adaptation beneficiaries against the mid-term target and provide an analysis of whether the project/programme is on track or not, and why.

EVALUATING ADAPTATION BENEFICIARIES IN THE FINAL EVALUATION

At the end of the implementation, the focus of the final evaluation, insofar as adaptation beneficiaries results are concerned, is on the final achievements of a project/programme. It is then critical for the evaluators to identify any departure from the methodological measurements and definitions, and how they were applied during implementation, and in case of departures, identify how it might have affected the

reported results. At this stage, it is also useful for evaluators to verify if there were any changes to the adaptation needs in the project area as identified in the vulnerability assessment undertaken during the funding proposal development stage.

BOX 9. CHECKLIST FOR EVALUATING ADAPTATION BENEFICIARIES AT THE END OF PROJECT/PROGRAMME IMPLEMENTATION

- Assess whether there were changes to the adaptation needs identified in the vulnerability assessment carried out as part of the funding proposal development.
- Examine whether there was a consistent application of the identification and definition of direct and indirect adaptation beneficiaries during the implementation.
- Examine how the project's progress in reaching adaptation beneficiaries were assessed.
- Examine any changes in approach, methodology or framework with respect to adaptation benefits and beneficiaries made during project implementation and assess whether they are pertinent to the project's objectives.
- Assess if any recommendations made during project implementation (e.g., in the interim evaluation or following APR submissions) in terms of improving the measurement of adaptation beneficiaries and their appropriateness were adequately implemented.
- Validate the reported results through sampling and where possible, provide illustrative examples and case studies based on a clear data collection and analytical framework.
- Assess how the adaptation beneficiaries aligned across the adaptation result areas of the IRMF.
- Assess the quality of sex disaggregated data for adaptation beneficiaries.
- For multi-country projects – assess whether the steps outlined above were followed for **each** country.
- Evaluate the effectiveness of the project in terms of adaptation beneficiaries reached.

3.C EVALUATING THE PARADIGM SHIFT

The evaluation of paradigm shift is *required* during the interim evaluation and final evaluation.

The GCF's **Investment Framework** and IRMF require Accredited Entities to provide a short description of a project/programme's paradigm shift potential within the funding proposal. This headline description is based on the analysis and logic behind the project/programme's theory of change. An initial assessment of the paradigm shift potential along with the baseline is also provided in the logframe. These assessments are based on three dimensions – scale, replicability, and sustainability. Specified at the impact level of the IRMF, paradigm shift assessments will typically fall under the Impact criterion and therefore associated questions must be specified in the Evaluation Matrix.

The IRMF's measurement of a paradigm shift is based on three key principles:

- **Contribution, not attribution:**[38] The GCF's Governing Instrument states that 'the purpose of the Fund is to make a significant and ambitious contribution to the global efforts towards attaining the goals set by the international community to combat climate change.' Accordingly, the IRMF is designed to identify contributions to a paradigm shift, not attributions.
- **Learning, not accountability:** The primary purpose of assessing a project/programme's contribution to a paradigm shift is for learning rather than accountability. A paradigm shift will take place over and above GCF-supported

[38] *Attribution is the idea that a change is solely due to interventions which a project/programme is undertaking while contribution is the idea that the influence of an intervention is just one of many factors which contribute to a change.*

activities, is most likely to occur over a long timeframe (typically beyond a GCF project/programme period) and will depend on multiple actors and externalities. Consequently, GCF project/programmes alone cannot be held accountable for whether a paradigm shift takes place. At the same time, the focus on learning should identify positive and negative lessons and should not just report on success.

- **A paradigm shift is context-specific:** A paradigm shift is context-specific, so it will not always be possible to directly compare metrics across different projects/programmes. For example, the absolute number of beneficiaries cannot be used to compare the scale of a paradigm shift across countries of different sizes. The measurement of a paradigm shift therefore needs to be firmly based on the context in which the project/programme operates.

OVERALL APPROACH TO EVALUATING A PARADIGM SHIFT

The GCF recommends a **scorecard-based approach** to the assessment of a paradigm shift, whereby progress is reviewed towards each paradigm shift dimension against a series of pre-defined statements.

Considering the paradigm shift baselines established within the funding proposal, evaluators will assess progress across each of the three paradigm shift dimensions (scale, replicability, sustainability).

Assessments will be scorecard-based, and – against a three-point scale – will identify the extent of change that is evident. All 'scores' should be supported by a qualitative, narrative assessment. Evaluators should draw on a range of evidence sources when making their assessment, including project/programme documentation (such as APRs), stakeholder interviews, and secondary data that can illustrate broader changes, such as national statistics, media reports and reports from other organizations.

While the baselines and anticipated contributions identified in a project/programme's funding proposal will provide the basis for the assessment, evaluators will also need to review evidence of **unanticipated changes and unanticipated contributions across all dimensions**. Given the complex, dynamic nature of a paradigm shift, it is highly possible that unexpected progress may occur due to unforeseen circumstances or new opportunities arising.

The focus when assessing a paradigm shift is learning, so assessments should not only record evidence of the paradigm shift but should also explain **why** any shift has or has not yet occurred, and how the project/programme is contributing or has contributed to the shift.

TABLE 9. RECOMMENDED SCORECARD FOR ASSESSING A PARADIGM SHIFT

DIMENSION	LOW	MEDIUM	HIGH
SCALE	Limited or no evidence of a pathway towards quantifiable impact in mitigation and adaptation measures beyond existing GCF intervention targets.	Clear evidence of a pathway towards increased quantifiable impact in mitigation and adaptation measures is emerging beyond project or programme targets. Evidence might include increased commitment/interest from existing project holders or new interested parties; the development of strategies covering larger target areas/ populations; or signs of better-than-expected results from GCF funded or influenced interventions.	Clear evidence of a pathway to a significant increase in quantifiable results. This evidence might include the evidence in the 'medium' score plus a significant expansion of GCF funded or influenced programmes based on increased resources allocated from new or existing sources and/or actual significant increase in measurable, quantifiable results within and beyond the scope of the project/programme by a range of similar interventions/ actors.
REPLICABILITY	Limited or no evidence of examples of intervention models funded/supported by the GCF that reduce emissions and/or increase resilience being considered in different geographical or sectoral settings or by new organizations.	Examples of intervention models which are similar or influenced by GCF funded/supported interventions are being planned and/or piloted in different contexts by one or more different organizations.	Clear evidence as in the 'medium' score plus the evidence that there are multiple examples of models similar to, or drawing from GCF funded interventions, being extensively funded and implemented, including appropriate adaptation to meet local context.
SUSTAINABILITY	No or limited evidence that the institutional structures and behavioural norms required to sustain the climate mitigation and adaptation benefits are sufficiently robust to exist without external funding and support.	Clear examples of where good practice norms and institutional structures have become embedded across a range of stakeholders and where intended outcomes are maintained without being reliant on external funding and support.	Clear evidence as in the 'medium' score plus the evidence that institutional structures and a range of stakeholder groups are able to lead, facilitate and support interventions that expand and further improve climate mitigation and adaptation benefits and the associated good practice norms.

INTERIM EVALUATION

Given the usually long-term nature of paradigm shifts, interim evaluations may realistically focus on the following elements:

- Assess if the qualitative baselines contained in the funding proposals are complete and appropriate.
- Examine how realistic (or not) the target scenarios are as specified in the funding proposal.
- Identify and analyse the conditions and contextual factors that will allow the project/programme to contribute towards the defined paradigm shift.

- Test the relevant assumptions in the theory of change to check which ones are holding (or not) along the causal pathways that may strengthen or weaken the project/programme contributions to the intended paradigm shift target scenarios.
- Assess the prospects for a paradigm shift based on the above, state the reasons for the conclusion.
- Fill up the scorecard and provide qualitative justifications for the scores.

FINAL EVALUATION

The final evaluation will be carried out for a more definitive assessment of the paradigm shift. The assessment will be guided by the same principles and dimensions but will be enriched by more information and evidence.

The evaluators need to assess how the project has contributed to achieving the paradigm shifts in a country, sector, and other spheres. While doing so, the following questions need to be answered.

- SCALE: To which degree there has been a significant increase in quantifiable results within and beyond the scope of the project/programme.
- REPLICABILITY: To which degree key structural elements of a project/programme are exported elsewhere within the same sector and/or to other sectors, regions, or countries.
- SUSTAINABILITY: To which degree the results of a project/programme are sustained beyond completion, through the creation of a structural and/or financial base, as well as through climate resilient practices.

At the end of the project/programme implementation, an evaluation may be in a better position to provide evidence of contributions or clearer prospects for contributions. As such, the focus of a final evaluation could include the following elements:

- Identify and analyse the conditions and contextual factors that will likely continue to allow or hinder the project/programme's contribution towards the paradigm shift.
- Re-test the relevant assumptions in the theory of change to check which ones are holding (or not) along the causal pathways that may strengthen or weaken the project/programme contributions to the intended paradigm shift target scenarios.
- Assess the project/programme's actual or likely contributions to the paradigm shift based on the above, state the reasons for the conclusion. If the assessment is based on likelihood, expound on the timeframe required for the materialization of contributions.
- Fill up the scorecard and provide qualitative justifications for the scores.
- Generate lessons learned for future programming.

3.D EVALUATING THE ENABLING ENVIRONMENT

The measurement of enabling environment indicators is *required* during the interim evaluation and final evaluation.

As part of these evaluations, AEs will provide substantive assessments of their chosen indicators (as selected during the funding proposal development stage). In both instances, the assessment is carried out by the evaluator(s) and will ideally closely involve the project/programme's key stakeholders.

CORE INDICATOR 5	Degree to which GCF projects/programmes contribute to strengthening institutional and regulatory frameworks for low-emission climate-resilient development pathways in a country-driven manner.
CORE INDICATOR 6	Degree to which GCF projects/programmes contribute to technology deployment, dissemination, development or transfer and innovation.
CORE INDICATOR 7	Degree to which GCF projects/programmes contribute to market development/transformation at the sectoral, local or national level.
CORE INDICATOR 8	Degree to which GCF projects/programmes contribute to effective knowledge generation and learning processes, and use of good practices, methodologies and standards.

Considering the baselines established within the funding proposal, evaluators will assess progress against each of the project/programme's selected enabling environment indicators. Each indicator has a corresponding scorecard based on a series of elements that break down the indicator, allowing for a more granular definition of the enabling environment being measured.

Evaluators will assess progress against each of these scorecard elements with a three-point scale, with all 'scores' to be supported by a qualitative, narrative assessment. Once all elements have been assessed and scored, an overall indicator score is calculated.

Evaluators should draw on a range of evidence sources when making their assessment, including project/programme documentation (such as APRs), stakeholder interviews, and secondary data that can illustrate broader changes, such as national statistics, media reports and reports from other organizations. While the baselines and anticipated contributions identified in a project/programme's funding proposal will provide the basis for the assessment, evaluators will also need to review evidence of unanticipated changes and unanticipated contributions. It is possible that unexpected changes to the enabling environment may have arisen due to unforeseen circumstances or new opportunities arising.

OVERALL APPROACH TO EVALUATING THE ENABLING ENVIRONMENT

The scorecards below shall be used by evaluators during the interim and final evaluations. All assessments – baseline development, interim evaluation, final evaluation – first require the **scoring of all the individual elements** that underpin an indicator. The scorecard statements represent an illustrative description of what each numerical score means. For example, in core indicator 5 the first element focuses on the degree to which an effective and socially inclusive regulatory/policy framework is developed for low emission climate resilient pathways. The statements for scores 1-3 show a progression towards meeting that outcome. It is important to note that the

statements are *indicative* and there will always need to be some judgment involved in making an assessment.

Once each element has been scored, **final indicator-level scores** can be calculated (**low, medium or high**). To calculate these indicator-level scores, all the element-level scores should be totalled up. The final indicator-level scores are then allocated according to the total of the element scores, as follows:

	LOW	MEDIUM	HIGH
CORE 5 (INSTITUTIONS)	Total element score: <=10	11-14	=>15
CORE 6 (TECHNOLOGY)	<=8	9-11	=>12
CORE 7 (MARKETS)	<=5	6-9	=>10
CORE 8 (KNOWLEDGE)	<=8	9-11	=>12

CORE INDICATOR 5:

Degree to which GCF projects/programmes contribute to strengthening institutional and regulatory frameworks for low-emission climate-resilient development pathways in a country-driven manner.

ELEMENT	SCORE 1	SCORE 2	SCORE 3
1	Lack of or limited legal/ regulatory/policy frameworks in place for low emission climate resilient pathways.	Clear evidence and examples of improved legal/regulatory/policy frameworks being developed and put into place which show an appreciation of low emission climate resilient pathways.	Effective socially inclusive legal/regulatory/ policy frameworks developed and implemented at a local and/or national level and clear evidence of enforcement of a regulation.
2	Limited or no financial and/or human resources allocated to support the development and implementation of institutional and regulatory frameworks.	Clear budgets and resources allocated to supporting the development of institutional and regulatory frameworks with some evidence of progress being made.	Significant and regular financial resources and organizational units focused on the development, implementation, and enhancement of institutional and regulatory frameworks.
3	Public sector actors do not have an organizational structure/system or trained staff to respond to climate change challenges.	Clear efforts being made to identify skills and capacity gaps to addressing the climate change crisis at both an organizational and individual level, with evidence of training and learning being underway.	Public sector actors have an organizational structure/system or are fully staffed with trained and knowledgeable individuals to address climate change challenges.
4	No horizontal or vertical cross government coordination in the response to climate change.	Evidence that government departments/ ministries and/or national and local governments are aware of the need for coordination and have initiated the development of coordination mechanisms to respond to climate change challenges.	Clear functioning coordination mechanisms at both horizontal and vertical levels effectively coordinating climate change responses.
5	Private sector players unaware of their contribution to climate change and do not have structures or skills to respond in a timely manner.	Clear examples of private sector companies developing and funding initiatives and strategies that directly identify and respond to climate change challenges.	Private sector players fully understand their role in addressing climate change and possess business models/strategies/ expertise to proactively address appropriate climate change challenges.
6	Civil society organizations have insufficient knowledge and skills to address relevant climate change challenges or to hold the public and private sector to account.	Clear evidence and examples that some civil society organizations understand climate change challenges and are developing strategies, interventions, and capabilities to ensure they are addressing those that are relevant to them.	Civil society organizations understand the contribution they can make and are collectively focused on ensuring their interventions address climate change challenges, and that they hold other stakeholders/duty bearers to account.

CORE INDICATOR 6: Degree to which GCF projects/programmes contribute to technology deployment, dissemination, development or transfer and innovation.			
ELEMENT	**SCORE 1**	**SCORE 2**	**SCORE 3**
1	Limited evidence that new technologies are being considered to address climate change challenges.	Clear examples of organizations assessing the possible use of new technologies to address climate change challenges including initial trialling or piloting.	Evidence of successful deployment and uptake of new technologies as part of regular/routine ways of working.
2	Limited financial resources being made available to fund innovation or to try new technologies or processes.	Clear evidence/examples of financial commitment and fund flows to improving innovation and/or utilizing new or transferring existing technologies to address climate change challenges.	Regular and routine allocation of funds made available for innovation, technology development and transfer.
3	Key organizations (public, private or civil society) have limited numbers or no staff with either the skills or time allocated to work on innovation or knowledge transfer.	There are some examples of organizations where they have staff with the skills to innovate, develop new or apply existing technologies in new ways and a structure which provides them with the opportunity.	There are a large number of organizations with models in place where skilled staff can be deployed in a timely way to develop and transfer new technologies and innovations.
4	No incentivization – such as tax relief, access to funding, grants, or tax breaks – provided at sectoral, local or national levels to support innovations.	Some evidence of incentivization processes/ products – such as access to funding, grants, or tax breaks – being developed at local or national levels and being taken up by firms or private sector bodies.	Clear evidence of incentives for developing and testing innovations, including acknowledgement and acceptance of possible failure leading to increased levels of innovation and technological deployment at sectoral, local, or national levels.
5	No sectoral, local or national level capacity within the government, amongst business associations or civil society groupings to promote, disseminate or transfer innovations.	Some evidence of mechanisms and structures to support and facilitate the dissemination of new innovations and technology transfer are being developed and appropriately resourced.	Mechanisms and platforms available and widely used to promote and transfer innovations to a range of audiences at a sectoral, local, national and possibly international level.

CORE INDICATOR 7: Degree to which GCF projects/programmes contribute to market development/transformation at the sectoral, local or national level.			
ELEMENT	**SCORE 1**	**SCORE 2**	**SCORE 3**
1	Limited or no market assessments being undertaken to analyse future market opportunities for low-emission, climate-resilient solutions.	Examples of progress towards effective market development and transformation are emerging based on assessments and analysis of potential opportunities.	A culture of looking for low-emission, climate-resilient market transformation opportunities has been created in a sector or region.
2	No evidence of projects/programmes contributing towards market development and transformation.	Clear examples of projects and programmes that are contributing to market development – for example through new sources of finance, market consolidation, improved value chains job creation (particularly for women and other excluded groups), economies of scale.	Strong evidence that there has been a systematic shift in market dynamics and a transformation in a climate change-affected sector due to projects/ programmes.

3	No evidence of projects/programmes incentivizing market participants by reducing costs or risks, or through eliminating barriers to the deployment of low-emission, climate-resilient solutions.	Clear examples of where projects and programmes are on a positive trajectory towards improving market conditions through cost or risk reductions or by addressing clear market barriers.	Strong evidence that projects/programmes have effectively and sustainably reduced the costs and risks of deploying effective low-emission and climate-resilient market solutions.
4	Little or no demand exists for the targeted market.	Clear evidence of increased demand and higher levels of interest from possible new market players.	Extensive consumer/institutional demand has been created supporting a vibrant competitive market and attracting new entrants.

CORE INDICATOR 8: Degree to which GCF projects/programmes contribute to effective knowledge generation and learning processes, and use of good practices, methodologies and standards			
ELEMENT	**SCORE 1**	**SCORE 2**	**SCORE 3**
1	No routine capturing or sharing of lessons learned by projects/programmes.	Examples of rigorous and credible lesson learning exercises being undertaken and shared at a regional or national level which highlight good practice examples and provide evidence for future action.	Routine and systemized, rigorous documented reflection of what has and has not worked at the project/programme level being shared at a national level and influencing future intervention design.
2	No effective project/programme-level monitoring, evaluation, action and learning systems.	Evidence that monitoring and evaluation is understood by a significant number of project/programme stakeholders within a region/sector including the development of plans and the allocation of a realistic level of resources to develop and implement a MEL system.	Effective and resourced monitoring, evaluation, action and learning system in place at a sectoral or national level, able to report on project and programme progress and to influence future project/programme designs.
3	No mechanism for sharing relevant knowledge of good practice and methodologies between and among projects/programmes.	Clear understanding of the need and commitment to the need to develop some form of shared learning platform, including allocated resources, a critical mass of stakeholders and clear leadership at a local or regional level.	Credible learning hub/mechanism in place which facilitates effective peer-to-peer knowledge exchange between and among projects/programmes at sectoral, regional or national levels.
4	No evidence that learning and knowledge generated at a project/programme level is being used to inform the development of improved methodologies or new standards.	Clear example(s) of how learning or knowledge has informed standards and/or improved methodologies at a sectoral, regional or national level.	Evidence that the use of knowledge to inform good practice and to revise expected standards has become routine and the norm at a sectoral, regional or national level.
5	No evidence of changes in direction, based on learning and knowledge generated at project/programme level.	Examples of organizations showing they are able to take on board lessons learned and have the flexibility and capability to change what they are doing based on those lessons.	Clear evidence of routine adaptive management across organizations in a region or sector based on learning generated through good practice M&E or structured reflective practices.

INTERIM EVALUATION

The interim evaluation will assess how the project/programme is making progress in bringing about the intended changes in the enabling environment as identified in the funding proposal.

- Assess if the qualitative baselines contained in the funding proposals are complete and appropriate.
- Examine how realistic (or not) the target scenarios are as specified in the funding proposal.
- Test the relevant assumptions in the theory of change to check which ones are following (or not) the causal pathways that may (or may not) allow the project/programme to achieve its enabling environment results.
- Assess the project/programme's progress.
- Fill up the scorecard and provide with qualitative justifications for the scores.

The evaluation will cover each indicator supported by a detailed analysis of granularity and quantitatively backed score cards. Evaluators will assess and assign a score against performance on each of these indicators along with a qualitative assessment.

Evaluators should draw on a range of evidence sources when making their assessment, including project/programme documentation stakeholder interviews and secondary data that can illustrate broader changes, such as national statistics, media reports and reports from other organizations. Again, a participatory process could be designed to collect the evidence and provide illustrative examples.

FINAL EVALUATION

The final evaluation will be carried out for a more definitive assessment of achievement in terms of creating an enabling environment. The assessments will be undertaken for all indicators but will be enriched by more information, detailed analysis and evidence.

- Assess the project/programme's achievements for each indicator, derive an indicator-level score.
- Provide a detailed narrative assessment of the performance along with illustrative examples.
- Review evidence of unanticipated changes due to unforeseen circumstances or new opportunities.

ANNEX 4: MANAGEMENT RESPONSE AND ACTION PLAN TEMPLATE

Accredited Entities are required to submit a Management Response and Action Plan to the GCF Secretariat for the interim evaluation and final evaluation. The GCF Secretariat recommends the use of the template below for this purpose.

PROJECT/PROGRAMME				
EVALUATION TITLE/YEAR				
ACCREDITED ENTITY	*Indicate who completed the management response and identified the follow-up actions.*			
OVERALL RESPONSE	*Summarize the Accredited Entity's overall view on the recommendations and how they will be implemented.*			
RESPONSE TO SECRETARIAT ASSESSMENT	*Summarize the Accredited Entity's response to the Secretariat assessment. This summary may be based on the evaluator's detailed response to the Secretariat's feedback based on the quality assurance framework in Annex 2 of these guidelines.*			
EVALUATION REPORT RECOMMENDATION NO. 1	*Detail the recommendation from the evaluation report that has warranted the following response.*			
PROJECT/PROGRAMME ACTIVITY/OUTPUT/ OUTCOME TO WHICH THE RECOMMENDATION IS RELEVANT	*State which activity, output or outcome (as stated in the approved FP) is relevant to the recommendation.*			
RESPONSIBILITY FOR IMPLEMENTATION	*The stakeholders who are expected to implement the assigned key actions.*			
MANAGEMENT RESPONSE (DELETE AS APPROPRIATE)	Agree	Partially agree	Disagree	Rejected
EXPLANATION FOR MANAGEMENT RESPONSE	*Provide justification for the chosen response.*			
KEY ACTIONS	*The planned actions to achieve the recommendations.*			
INTENDED RESULTS FOR KEY ACTIONS	*Set out what the actions aim to achieve.*			
PRIORITY LEVEL OF IMPLEMENTATION	**1** – urgent/to be addressed first	**2** – non-urgent, but important	**3** – less urgent and somewhat important	

TIMEFRAME FOR IMPLEMENTATION	*The timeframe provides information about the planned period during which the key actions are expected to be implemented, which is also indicative of the urgency of a recommendation and its applicability as indicated above in 'Priority level of implementation'.*			
STATUS (DELETE AS APPROPRIATE)	**Completed:** the action has been implemented.	**Initiated:** implementation of the action has begun but is not completed.	**Not yet initiated:** the action is yet to be implemented.	**Key action no longer applicable:** the action has been deemed no longer relevant.
IF ACTION IS NO LONGER APPLICABLE, PLEASE PROVIDE JUSTIFICATION				
TIMETABLE FOR PERIODIC REVIEW	*Include a timetable for addressing the recommendations and a frequency for tracking the various follow-up steps and actions agreed.*			
EVALUATION REPORT RECOMMENDATION NO. 2				
RESPONSIBILITY FOR IMPLEMENTATION				
MANAGEMENT RESPONSE (DELETE AS APPROPRIATE)	Agree	Partially agree	Disagree	Rejected
EXPLANATION FOR MANAGEMENT RESPONSE				
KEY ACTIONS				
PRIORITY LEVEL OF IMPLEMENTATION	**1** – urgent/to be addressed first	**2** – non-urgent, but important	**3** – less urgent and somewhat important	
INTENDED RESULTS FOR KEY ACTIONS				
TIMEFRAME FOR IMPLEMENTATION				

STATUS (DELETE AS APPROPRIATE)	**Completed:** the action has been implemented.	**Initiated:** implementation of the action has begun but is not completed.	**Not yet initiated:** the action is yet to be implemented.	**Key action no longer applicable:** the action has been deemed no longer relevant.
IF ACTION IS NO LONGER APPLICABLE, PLEASE PROVIDE JUSTIFICATION				
TIMETABLE FOR PERIODIC REVIEW				

Add additional rows following this template outline as needed for further recommendations.

ANNEX 5: INDICATIVE TERMS OF REFERENCE FOR INTERIM AND FINAL EVALUATIONS

This annex aims to provide Accredited Entities (AEs) with an indicative terms of reference (TOR) template for commissioning evaluation services that has, as far as possible, already been tailored to the specific requirements of AE-led evaluations. The template should be adapted to the specific needs and priorities of the project/programme by deleting/revising and adding content as appropriate.

The TOR can be used for commissioning an interim evaluation, a final evaluation, or a package of evaluation-related deliverables comprised of the interim evaluation, final evaluation, and related independent evaluative data collection activities (i.e., baseline, annual, mid-line, end-line).

[TITLE] INTERIM/FINAL/INDEPENDENT EVALUATION OF [PROJECT/PROGRAMME]

I. PROJECT/PROGRAMME BACKGROUND

Short description of the project/programme, including aims and objectives, components, intended results (impact, outcomes, outputs), implementation timescales, budget, and implementation arrangements. Where relevant, include co-financing information.

II. OBJECTIVES AND PURPOSE OF THE EVALUATION

The project is seeking to procure the services of an independent evaluator/evaluation team to design, plan and conduct an interim/final evaluation of the [project] over the next [X] years, which is funded through the GCF. The evaluator will provide an independent and rigorous evaluation function, designing and implementing a framework which will assess the project along the GCF evaluation criteria and consistent with the GCF Evaluation Policy. It will report the findings and lessons learned and provide recommendations.

The findings from the evaluation will primarily be used [Select all that apply and add more if needed]:

- by the project management team, delivery partners and other stakeholders to inform improvements in project implementation.
- by the project management team to support the on-going development and implementation of the project's sustainability.
- to demonstrate accountability for the funding received from the GCF.
- by partners, stakeholders, the government to learn lessons from the project for the purpose of replicating what works elsewhere and/or taking up approaches and activities that have been proven to work to scale up the project.

III. SCOPE AND FOCUS OF THE EVALUATION

a. Scope of work

The interim/final evaluation will assess project/programme implementation, its progress, overall management, credibility of results/reporting and achievement of results and/or contributions towards expected results, including behavioural changes necessary to achieve the expected results. The evaluation will apply the GCF Evaluation Criteria set out in the GCF Evaluation Policy and further defined in the Evaluation Guidelines.

b. Evaluation questions

The following questions are intended to guide evaluators to deliver credible and trusted evaluations. AEs will **select and prioritize** the questions that will be most relevant to the project's learning and accountability needs.

RELEVANCE

- Were the context, problem, needs and priorities well analysed and reviewed during project initiation?
- Are the planned project objectives and outcomes relevant and realistic to the situation on the ground?
- Is the project theory of change (TOC) and intervention logic coherent and realistic? Does the TOC and intervention logic hold, or does it need to be adjusted?
- Do outputs link to intended outcomes which link to the broader paradigm shift objectives of the project?
- Are the identified, planned inputs and strategies realistic, appropriate and adequate to achieve the results? Were they sequenced sufficiently to efficiently deliver the expected results?
- How realistic are the risks and assumptions of the project?

EFFICIENCY

- To what extent did project deal with issues and risks in implementation in an efficient manner?
- Have project resources been utilized in the most economical, effective and equitable ways possible (considering value for money, absorption rate, commitments versus disbursements and projected commitments, co-financing, etc.)?
- Are the project's governance mechanisms functioning efficiently?
- To what extent did the design of the project help or hinder achieving its own goals?
- To what extent was the M&E tools such as the TOC and logframe used in performance management and progress reporting?
- To what extent did the project's M&E data and mechanism(s) contribute to achieving project results?
- Were there clear baseline indicators and/or benchmarks for performance measurements? How were these used in project management?
- To what extent and how did the project apply adaptive management?

EFFECTIVENESS

- Are the outputs being achieved in a timely manner? Is this achievement supportive of the TOC and pathways identified?

- What and how much progress has been made towards achieving the overall outcomes such as adaptation beneficiaries and/or reduced GHG emissions/ increased carbon sequestration of the project (including contributing factors and constraints)? How strong is the evidence base for the achievements of outcomes, and to what extent are they based on the application of a well-defined methodology?
- To what extent has the project contributed to an enabling environment? What is the strength of evidence for this finding based on the scorecard assessment?
- To what extent is the project able to demonstrate changes against the baseline (assessment in approved funding proposal) for the GCF investment criteria (including contributing factors and constraints)?
- What, if any, alternative strategies would have been more effective in achieving the project objectives?

COHERENCE IN CLIMATE FINANCE DELIVERY WITH OTHER MULTILATERAL ENTITIES

- Who are the partners of the project and how strategic are they in terms of capacities and commitment?
- Is there coherence and complementarity by the project with other actors for other local climate change interventions?
- To what extent has the project complimented other on-going local-level initiatives (by stakeholders, donors, governments) on climate change adaptation or mitigation efforts?
- How has the project contributed to achieving a stronger and more coherent integration of the shift to low-emission sustainable development pathways and/or increased climate-resilient sustainable development (GCF RMF/PMF paradigm shift objectives)? Please provide concrete examples and make specific suggestions on how to enhance these roles going forward.

GENDER EQUITY

- To what extent has the project relied on and goes beyond sex-disaggregated data per population statistics?
- Are financial resources/project activities explicitly allocated to enable women, youth, people with disability, indigenous people and other marginalised groups to benefit from project interventions?
- Does the project account in activities and planning for local power dynamics and how project interventions affect different marginalised groups as beneficiaries?
- Do all groups of beneficiaries know their rights and/or benefits from project activities/interventions?
- How do the results for women compare to those for men?
- Is the decision-making process transparent and inclusive of all relevant marginalised groups?
- To what extent are the beneficiaries satisfied with the project's results?
- Did the project sufficiently address cross-cutting issues, including gender?

COUNTRY OWNERSHIP OF PROJECTS AND PROGRAMMES

- To what extent is the project aligned with national development plans, national plans of action on climate change, or sub-national policy as well as projects and priorities of national partners?

- How well is country ownership reflected in the project governance, coordination and consultation mechanisms or other consultations?
- To what extent are country-level systems for project management or M&E utilized in the project?
- Is the project as implemented responsive to local challenges and relevant/appropriate/strategic in relation to SDG indicators, national indicators, GCF RMF/PMF indicators, AE indicators, or other goals?
- Were the modes of deliveries of the outputs appropriate to build essential/necessary capacities, promote national ownership and ensure sustainability of the results achieved?

INNOVATION IN RESULTS AREAS

- Which role has the project played in the provision of 'thought leadership', 'innovation', or 'unlocked additional climate finance' for climate change adaptation/mitigation in the project and country context? Please provide concrete examples and make specific suggestions on how to enhance these roles going forward.

REPLICATION AND SCALABILITY

- What are the project's lessons learned, failures/lost opportunities to date? What might have been done better or differently?
- How effective were the exit strategies and approaches to phase out assistance provided by the project, including contributing factors and constraints?
- Which factors of the project achievements are contingent on a specific local context or enabling environment factors?
- Are the actions and results from project interventions likely to be sustained, ideally through ownership by the local partners and stakeholders?
- What are the key factors that will require attention to improve prospects of sustainability, scalability or replication of project outcomes/outputs/results?

IMPACT

- To which extent has the project contributed or will be contributing to the desired paradigm shift? What is the strength of evidence for this finding based on the scorecard assessment?

UNEXPECTED RESULTS

- How has the project's ability to adapt and evolve been based on continuous lessons learned and the changing development landscape? Please account for factors both within the AE/EE and external.
- Can any unintended or unexpected positive or negative effects be observed as a consequence of the project's interventions?
- What factors have contributed to the unintended outcomes, outputs, activities, results?

IV. METHODOLOGY

The interim/final evaluation should be rigorous allowing for verification and triangulation of findings. It will seek, to the extent possible, to be inclusive and participatory, involving principal stakeholders and beneficiaries in the analysis. As per the policy, a mixed methods approach is recommended, which at a minimum will include:

1. Desk review of relevant documents including baseline studies, progress reports and any records of surveys conducted during the project, stakeholder maps, etc.
2. Qualitative methods such as interviews and focus group discussions to explore and explain observed changes. The type of methods used may vary depending on the cultural constraints or opportunities presented by the project's context.
3. Quantitative methods such as survey or secondary data analysis to assess and validate reported results.

Other data collection should be undertaken as needed (government data/records, field observation visits, CDM verifications, public expenditure reporting, GIS data, etc.) to validate evidence of results and assessments (including but not limited to assessment of the TOC, activities delivery and results/changes occurred).

V. QUALITY ASSURANCE AND RISK MANAGEMENT

Bidders are required to submit a quality assurance plan that sets out the systems and processes for ensuring that the evaluation deliverables and all evaluation activities are of a high quality and meet ethical standards from inception to the end of the contract. In addition, bidders will demonstrate understanding of the risks associated to the delivery of the evaluation with a contingency plan to mitigate the impacts of the risks on the evaluation outputs.

VI. EXPECTED OUTPUTS AND DELIVERABLES

The consultant is expected to deliver the following outputs:

1. An **inception report.**
2. A **draft evaluation report and presentation**, to be presented at a debriefing meeting with the AE and EE project teams.
3. A **final report**, including a two to three-page executive summary, a set of limited and strategic recommendations (not to exceed 10 recommendations in total), and a response addressing issues raised during the presentation of the draft.
4. Inputs to the Management Response and Action Plan.

VII. EVALUATION GOVERNANCE ARRANGEMENTS

During the implementation of the contract, the evaluator will report to the [AE focal point], who will provide guidance and ensure satisfactory completion of interim/final evaluation deliverables. There will be coordination with the project team who will assist in connecting the evaluator with senior management, government and development partners, beneficiaries, and other relevant key stakeholders. In addition, the project team will provide key project documentation prior to fieldwork and assist in developing a detailed programme to facilitate consultations as necessary.

Where applicable, describe the constitution of an Evaluation Steering Group, which could advice the AE on the definition of the TOR and quality assurance of reports. An Evaluation Steering Group could be comprised of a Chair, the Evaluation Manager, independent sector, and evaluation specialists (subject specialists), academics, and representatives from delivery partners and beneficiaries. Ideally, the group is constituted before commencing an evaluation.

VIII. QUALIFICATIONS REQUIRED

Bidders are required to clearly identify and provide CVs for all those proposed for the evaluation team, clearly stating their roles and responsibilities for this evaluation.

The proposed evaluators should include the technical expertise and practical experience required to deliver the scope of work and evaluation outputs.

Bidders should provide evidence of previous project experience for the provision of similar evaluation services and the design and implementation of similar evaluation activities as required by this TOR.

IX. CONTRACT DURATION

Number of working days: X working days in Y calendar days/months
Contract start date: date, month, year
Contract end date: date, month, year

> Where AEs decide to commission different evaluators for each evaluation, contract duration will be shorter and contingent on the timeframe for delivering the evaluation reports to the GCF.
>
> Where AEs decide to package the evaluations along with other evaluative activities such as independent baseline, mid-line, or end-line (in some cases annual) data collection, then the contract duration may run in parallel to the project/programme.

X. BUDGET

The estimated budget for this work is [specify]. This budget is inclusive of all costs covering team member costs, travel, communications, quality assurance and any other costs associated with the completion of the work. Bidders are required to organize and fund their own duty of care arrangements as required.

XI. ANNEXES

Existing information sources

- GCF funding proposal, including the logical framework and theory of change
- APRs
- Monitoring data and information
- GCF Evaluation Policy
- GCF Evaluation Standards
- GCF Evaluation Operational Procedures and Guidelines
- GCF Integrated Results Management Framework Handbook

INCEPTION REPORT SAMPLE OUTLINE

I. Project/programme description, including a brief description of the requirements of the TOR and evaluation audience

II. Evaluation questions

III. Evaluation approach, including overall design, data collection methods and analytical procedures

IV. Ethical considerations

V. Stakeholder engagement and dissemination plan

VI. Quality assurance, risk management plan

VII. Roles and responsibilities

VIII. Detailed evaluation work plan indicating the activities at each phase, timing of delivery, key deliverables, and milestones

Annex: Evaluation matrix, draft data collection tools

EVALUATION REPORT SAMPLE OUTLINE

I. Executive summary – no more than three pages containing a summary of the key findings and recommendations.

II. Introduction – including but not limited to context, description of project/programme under evaluation, evaluation scope, evaluation methodology, methodological limitations and mitigation measures, audience, and dissemination plan.

III. Findings – can be structured by evaluation criterion. Ensure that the findings are based on multiple sources and be clear on the strength of evidence supporting/refuting the findings.

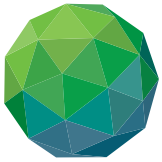
IV. Lessons learned – should directly link to the key findings, which will then shape the recommendations.

V. Recommendations –

Annex: Evaluation matrix, mission reports, list of interviewees, list of documents reviewed, data sources used, detailed calculations and supporting evidence for mitigation and adaptation results, detailed scorecards and others.

CONTACT INFORMATION

For queries regarding the GCF Evaluation Policy or these guidelines, please contact at evaluation_policy@gcfund.org.



GREEN
CLIMATE
FUND