



GREEN
CLIMATE
FUND

Independent
Evaluation
Unit



GCF Impact Evaluation Standards
August 2026

Green Climate Fund Impact Evaluation Standards

Table of Contents

I.	Context and purpose of the Green Climate Fund Impact Evaluation Standards.....	1
II.	Definition of impact and impact evaluation at the GCF	2
III.	Purpose of impact evaluation at the GCF.....	2
	3.1 Accountability	2
	3.2 Learning.....	2
	3.3 Evidence-informed decision-making.....	2
IV.	Consideration for conducting impact evaluation	3
	4.1 Strategic learning and evidence priorities	3
	4.2 Relevance and potential use of evidence.....	3
	4.3 Evaluability and feasibility	3
V.	GCF Impact Evaluation Standards	3
Annex I.	Glossary of terms.....	8
Annex II.	Suggested checklist for application of GCF Impact Evaluation Standards	10
Annex III.	References	16

I. Context and purpose of the Green Climate Fund Impact Evaluation Standards

1. The *Evaluation Policy for the GCF*¹ establishes an evaluation function that “helps the Fund credibly and objectively assess and measure its performance, results, effectiveness and efficiency in delivering its mandate, including its contribution to promoting a paradigm shift towards low-emission and climate-resilient development pathways”.² The Policy provides that “the IEU [Independent Evaluation Unit] shall develop standards, in collaboration with the Secretariat, and the Secretariat will develop guidelines to implement the Policy, in collaboration with the IEU, that ensure the Fund is able to inform its overall results, success and unintended consequences in a credible and measurable manner”.³ In line with this provision, the *Green Climate Fund Evaluation Standards* were produced in 2022.⁴

2. The Policy further outlines “the long-term aim that approximately 30 per cent of the Fund’s projects and programmes approved annually by the Board will include real-time impact assessments as part of their evaluation plans”.⁵ Such evaluations are expected to generate credible evidence on the causal effects of Green Climate Fund (GCF) supported interventions and contribute to learning, accountability, adaptive management and strategic decision-making across the GCF portfolio.

3. Although the GCF’s Evaluation Standards provide overarching principles applicable to all evaluations, impact evaluations raise additional methodological, operational and ethical considerations. In particular, impact evaluations require careful attention to causal inference, counterfactual analysis, risks related to bias and implementation realities. Climate change mitigation and adaptation interventions also pose specific challenges for impact evaluation, including long and often non-linear causal pathways, complex and multi-component intervention designs, heterogeneous impacts across populations and geographies, and the interactions between climatic, environmental, institutional and socioeconomic factors.

4. The present GCF Impact Evaluation Standards therefore complement the broader GCF Evaluation Standards by defining principles and expectations specifically for the design, implementation, assessment and use of high-quality impact evaluations in the context of climate change interventions. The Standards are informed by international good practice and draw on lessons and approaches from impact evaluation guidance developed across the international development and United Nations systems.

5. The target audience for these Standards includes those who commission, design, conduct, manage and use impact evaluations, such as the IEU, the GCF Secretariat, accredited entities, implementing partners, independent evaluators and research institutions. Applying and implementing the Standards is a shared responsibility. These Standards are principles-based rather than prescriptive and should be applied in a proportionate manner, taking into account the nature, scale, evaluability and operational context of the intervention.

¹ The Evaluation Policy of the GCF was adopted by the Board following decision B.BM-2021/07.

² Green Climate Fund, *Evaluation Policy for the GCF*, para. 10.

³ *Ibid.*, para. 50.

⁴ Following decision B.BM-2021/07, the IEU was requested to develop the GCF Evaluation Standards. The Standards were produced in May 2022, developed in collaboration with the Secretariat.

⁵ Green Climate Fund, *Evaluation Policy for the GCF*, para. 58(e).

II. Definition of impact and impact evaluation at the GCF

6. Within the GCF context, impact refers to medium- to long-term changes resulting from GCF-supported interventions, including intended and unintended and positive and negative effects. The GCF Integrated Results Management Framework (IRMF) establishes a results architecture for assessing how GCF investments deliver climate results and contribute to paradigm shift toward low-emission and climate-resilient development pathways.
7. Consistent with the definition proposed by the United Nations Evaluation Group, impact evaluations aim to assess intended changes for a population that are attributable to a specific intervention, using counterfactual analysis. In the context of the GCF, this includes assessing changes in climate mitigation and adaptation outcomes attributable to GCF-supported interventions.
8. Accordingly, not all evaluations of impact-level outcomes are impact evaluations. Results monitoring, impact assessments, performance evaluations, and other evaluation approaches may assess progress towards or achievement of impact-level results but do not necessarily establish causal attribution. Under these Impact Evaluation Standards, an impact evaluation refers specifically to evaluations that apply appropriate methodological approaches to establish a credible counterfactual and identify the causal effects of GCF-supported interventions on climate mitigation and adaptation outcomes.

III. Purpose of impact evaluation at the GCF

3.1 Accountability

9. Impact evaluations should provide robust evidence on whether and how much observed changes in climate mitigation and adaptation outcomes are attributable to GCF-supported interventions. By establishing causal relationships between interventions and observed outcomes, impact evaluations strengthen accountability for results and support the Fund's mandate to assess its performance, effectiveness and efficiency. This evidence contributes to demonstrating how GCF investments advance climate-resilient and low-emission development pathways.

3.2 Learning

10. Impact evaluations can generate insights into what works, why it works, and under what conditions climate-related interventions produce results. Beyond measuring impacts, these evaluations help identify causal mechanisms, implementation dynamics and contextual factors that influence outcomes. This learning supports the continuous improvement of project design, implementation strategies and programmatic approaches within the GCF portfolio.

3.3 Evidence-informed decision-making

11. Impact evaluations should generate evidence that informs strategic and operational decisions within the GCF. By robustly estimating the changes in the observed outcomes that are attributable to the intervention, including intended and unintended effects, both negative and positive ones, as well as by analysing observed changes not attributable to the intervention, impact evaluations should reliably guide decisions on replication, scaling-up and resource allocation, helping the Fund prioritize the investments that are most likely to maximize positive impact on climate outcomes.

IV. Consideration for conducting impact evaluation

12. Impact evaluations should be conducted selectively and strategically where they can generate the greatest value for learning, accountability and evidence-informed decision-making. In line with the GCF Evaluation Policy, which envisages that approximately 30 per cent of GCF-funded projects will include real-time impact assessments, the selection of projects for impact evaluation should consider a combination of complementary factors. No single factor is determinative; rather projects may be prioritized when multiple considerations align and where rigorous evaluation is feasible and likely to generate meaningful evidence.

4.1 Strategic learning and evidence priorities

13. Impact evaluations should be prioritized when interventions address important evidence gaps related to climate mitigation, adaptation or climate finance, particularly where causal evidence remains limited or where innovative approaches, technologies or implementation models are being tested.

4.2 Relevance and potential use of evidence

14. Impact evaluations are particularly valuable when findings can inform future programming, replication, scaling or policy adoption within the GCF portfolio or beyond. Priority may therefore be given to interventions with high potential to generate actionable and/or generalizable evidence with significant expected climate mitigation or adaptation outcomes, transformative potential, or lessons that can inform replication, scaling or broader strategic learning.

4.3 Evaluability and feasibility

15. Priority should also be given when a credible evaluation design is feasible. Considerations include clarity of the theory of change, availability of suitable comparison groups or counterfactuals, implementation arrangements that support rigorous evaluation, and the practicality of collecting high-quality data over an appropriate time frame.

V. GCF Impact Evaluation Standards

16. The GCF Impact Evaluation Standards set out the core principles that should guide the design, implementation, assessment and use of impact evaluations in the context of the GCF. Building on the overarching GCF Evaluation Standards, the Impact Evaluation Standards adapt their application to the specific methodological and ethical considerations of impact evaluation, including causal attribution, counterfactual analysis, rigorous evidence generation, and responsible conduct of evaluation. The five standards are Relevance, Credibility, Utility, Transparency and Ethics. They are mutually reinforcing and should be applied together to ensure that impact evaluations generate credible, useful and responsible evidence for learning, accountability and evidence-informed decision-making.

Standard 1. Relevance: What should be evaluated?

17. Relevance in evaluation concerns whether the evaluation addresses questions and outcomes that matter in relation to the intervention objectives, context and intended users. For impact evaluation at the GCF, relevance requires a clear focus on outcomes that are directly linked to the intervention being evaluated and contribute to climate mitigation and/or adaptation objectives. These outcomes should align with the Fund's objectives of promoting a paradigm shift towards low-emission and climate-resilient development pathways.
18. Relevance is demonstrated when an impact evaluation clearly articulates causal questions, outcomes and indicators that assess whether, how and to what extent a GCF-supported intervention contributes to climate-related change. Evaluations should therefore focus on outcomes that are directly linked to the intervention's climate objectives and theory of change, including relevant changes in resilience, emissions, adaptive capacity, ecosystem conditions or institutional systems, where appropriate. Where relevant, evaluations may also consider broader development co-benefits that arise alongside climate outcomes.
19. Relevance also requires that evaluation questions, indicators and units of analysis are appropriate to the nature of the intervention and the pathways through which change is expected to occur. Given the wide range of GCF-supported interventions, their objectives, delivery models and expected outcomes, impact evaluations should be tailored to assess climate-relevant change at the most appropriate level, whether at the level of individuals, households, communities, institutions, markets or ecosystems.
20. Given the often long-term and systemic nature of climate interventions, relevant impact evaluations may incorporate the analysis of intermediate outcomes and causal mechanisms along the project's theory of change that are expected to contribute to longer-term mitigation or adaptation impacts. Such outcomes may include behavioural changes, institutional arrangements, or adoption of climate-relevant practices that plausibly lead to longer-term climate resilience or emissions reductions.
21. Finally, relevant impact evaluations should employ climate-appropriate indicators, data sources and measurement approaches capable of capturing meaningful climate-related outcomes and assessing whether observed changes can reasonably be attributed to the intervention being evaluated. This includes selecting indicators and data sources that reflect the scale, timing and pathways of expected change, while remaining grounded in the realities of implementation and the information needs of intended users.

Standard 2. Credibility: Are the causal findings trustworthy?

22. Credibility requires that impact evaluations use rigorous and objective methods and high-quality evidence to ensure that causal findings are trustworthy. Consistent with GCF Evaluation Standard 5, on "Credibility and Rigour",⁶ impact evaluations should apply methodological approaches that are sufficiently robust and transparently documented, such that findings can be independently assessed and scrutinized.
23. Impact evaluations must establish a counterfactual. The choice of impact evaluation design should be appropriate to the context, implementation modality and evaluation questions, and should support credible causal inference. Where feasible, evaluations should employ experimental designs such as randomized controlled trials (RCTs). Where randomization is not feasible, evaluations should use quasi-experimental approaches that construct a credible comparison group, such as difference-in-differences (DID), regression discontinuity (RD),

⁶ Independent Evaluation Unit, *Green Climate Fund Evaluation Standards*, 8.

matching or instrumental variables (IV). The evaluation should clearly explain how the chosen design aligns with the project's implementation realities, available data and evaluation questions, and how the design supports causal attribution rather than simple correlation.

24. Credibility is strengthened when the evaluation design is specified ex-ante through a clear impact evaluation strategy. This strategy should set out the evaluation questions, theory of change, key indicators, identification strategy, sampling approach, statistical power considerations where relevant, data-collection protocols, and planned analytical methods – before outcome data are analysed. Ex-ante specification improves methodological transparency, reduces the risk of ad hoc analytical choices, and helps ensure that evaluation design is aligned with implementation realities and learning objectives.

25. Credibility also requires careful attention to internal validity and potential sources of bias. Evaluations should identify and mitigate risks such as selection bias, attrition, spillovers between treatment and comparison groups, measurement error, and deviations from the intended intervention design. Where such risks arise, evaluators should apply appropriate design or analytical strategies and transparently report assumptions, limitations and robustness checks.

26. Credibility can also be strengthened through mixed-methods approaches and the use of fit-for-purpose data sources, enabling findings to be triangulated using both quantitative and qualitative evidence. Impact evaluations should integrate quantitative causal estimates with qualitative evidence to examine causal mechanisms, implementation dynamics and contextual factors, and to interpret heterogeneous effects.

27. Data used for analysis should meet appropriate quality standards and be suitable for causal analysis. Relevant data sources should be selected based on the intervention context, evaluation design and causal questions under study. Such sources may include, among others, project monitoring data, administrative records, household or community surveys, geospatial or remote-sensing data, climate and environmental monitoring data, and other context-appropriate quantitative or qualitative sources.

Standard 3. Utility: Will the evidence inform decisions?

28. Impact evaluations should be designed and implemented to inform learning, accountability and evidence-informed decision-making. Drawing on the GCF's Evaluation Policy and Evaluation Standard 3 on "Utility and Value Added",⁷ an impact evaluation achieves utility when its findings, conclusions and recommendations are relevant to the needs of intended users and are available in a form and at a time that enables action. Utility also depends on the timely availability of evaluation findings so that evidence can inform ongoing implementation, adaptive management, strategic planning and future investment decisions.

29. Beyond informing project-level and institutional decision-making, impact evaluations supported by the GCF should also generate actionable learning. Useful impact evaluations should go beyond reporting statistical results and help explain why impacts occur, through which mechanisms, for whom and under what conditions. This includes interpreting heterogeneous effects, identifying implementation factors that influence outcomes, and drawing lessons that can inform adaptation, replication and future programme design.

30. Effective utilization of evaluation findings also requires active planning for dissemination, learning and uptake throughout the evaluation process. Impact evaluations should include a dissemination and learning plan that specifies how findings will be

⁷ Independent Evaluation Unit, *Green Climate Fund Evaluation Standards*, 7.

communicated to different audiences and how they will inform decision-making, learning, and knowledge-sharing among relevant stakeholders, including within the GCF where appropriate.

Standard 4. Transparency: Clear and transparent documentation of evaluation methods, data and findings

31. Transparency requires clear, complete and systematic documentation of impact evaluation processes, methods and findings so that users can understand what was done, why it was done and how conclusions were reached. In line with the GCF Information Disclosure Policy,⁸ impact evaluations supported by the Fund should ensure the transparent disclosure of relevant information generated through the evaluation process, including the timely publication of evaluation reports and supporting materials, where appropriate. Transparent disclosure should, however, be balanced with appropriate safeguards for confidentiality, privacy and the protection of sensitive information.

32. In addition to general transparency requirements, such as communicating the purpose of the evaluation, the criteria applied, the methods used and the intended use of findings, impact evaluations should clearly document the causal logic and analytical procedures underlying the evaluation. This includes explaining how the counterfactual was constructed, how data were collected and analysed, and what assumptions and methodological choices influence the interpretation of results. Where substantive changes are made to the evaluation design, analytical plan or data-collection protocols, such deviations should be clearly documented and justified.

33. Transparency also requires balanced reporting of evaluation findings. Impact evaluations should report positive and negative results, as well as unintended consequences. Reports should clearly distinguish evidence from interpretation and communicate limitations, uncertainty and methodological constraints so that decision makers do not overinterpret findings and future evaluations can learn from both successes and challenges identified through the evaluation.

Standard 5. Ethics: Was the evaluation conducted responsibly?

34. Ethics in impact evaluation requires that evaluations be conducted in a manner consistent with the principles of respect for persons, beneficence and justice, while upholding integrity, honesty and responsibility throughout the evaluation process. As established in GCF Evaluation Standard 11 on Ethics,⁹ ethical conduct prioritizes the rights and interests of all those involved and requires appropriate engagement, sensitivity to context, protection of human dignity and attention to power imbalances, particularly when working with vulnerable or marginalized groups.

35. In practical terms, impact evaluations should include robust procedures to ensure respect for participants, including free, prior and informed consent and the protection of privacy and personal data. Participation must be voluntary and non-coercive, and participants should be clearly informed about the purpose of the evaluation, how information will be used, and the limits of confidentiality. Personal and sensitive information should be protected through appropriate confidentiality safeguards and data management protocols. Where impact evaluations involve household surveys, key informant interviews, focus group discussions or

⁸ Green Climate Fund, *Information Disclosure Policy*. The Information Disclosure Policy of the GCF was adopted by decision B.12/35. The document sets out the policy regarding the information that the Fund makes available to the public either as a routine matter or upon request.

⁹ Independent Evaluation Unit, *Green Climate Fund Evaluation Standards*, 11.

other forms of primary data collection, appropriate safeguards should be applied for the collection, storage, use, sharing and retention of personal and household-level data, including anonymization, secure data management and controlled access to sensitive information. Where relevant, impact evaluations should comply with applicable ethical review procedures, including an institutional review board (IRB) or equivalent ethical approval mechanisms.

36. Ethical practice in impact evaluation also requires adherence to principles of research integrity. Evaluation questions, methods, data collection, analysis and reporting should be designed and conducted honestly, objectively and impartially, with due consideration to fairness, scientific integrity and the ethical implications of methodological choices. Where experimental or quasi-experimental designs are used, evaluators should carefully consider the ethical implications of treatment assignment, exclusion, incentives, and any potential unequal distribution of risks or benefits among participants. Potential conflicts of interest should be identified and appropriately managed to preserve the independence and integrity of the evaluation process.

37. Ethical practice also requires adherence to the principle of “do no harm”. Evaluations should be designed and implemented in ways that minimize potential risks to participants and communities. Where evaluations uncover potential wrongdoing, abuse or serious harm, those responsible for conducting the evaluation should follow applicable legal, institutional, ethical and safeguarding protocols while maintaining confidentiality and protecting those who provided information.

Annex I. Glossary of terms

TERM	DEFINITION WITHIN THE GCF CONTEXT
Causal attribution	“The ascription of a causal link between observed (or expected to be observed) changes and a specific intervention. Note: Attribution refers to that which is to be credited for the observed changes or results achieved. It represents the extent to which observed development effects can be attributed to a specific intervention or to the performance of one or more partners, taking account of other interventions, (anticipated or unanticipated) confounding factors or external shocks” (Organisation for Economic Co-operation and Development, 2023, p. 20).
Comparison group	“Also known as a <i>control group</i>. A valid comparison group will have the same characteristics on average as the group of beneficiaries of the program (<i>treatment group</i>), except for the fact that the units in the comparison group do not benefit from the program. Comparison groups are used to estimate the <i>counterfactual</i>” (World Bank, 2016, p. 326).
Counterfactual	“The situation or condition which hypothetically may prevail for individuals, organizations, or groups were there no development intervention” (Organisation for Economic Co-operation and Development, 2023, p. 25).
Difference-in-differences (DID)	“Also known as <i>double difference</i> or <i>DD</i>. Difference-in differences compares the changes in <i>outcomes</i> over time between the <i>treatment group</i> and the <i>comparison group</i>. This eliminates any differences between these groups that are constant over time” (World Bank, 2016, p. 327).
Institutional Review Board (IRB)	“A committee that has been designated to review, approve, and monitor research involving human subjects. Also known as an <i>independent ethics committee</i> (IEC) or <i>ethical review board</i> (ERB)” (World Bank, 2016, p. 328).
Instrumental variable (IV)	“Also known as <i>instrument</i>. The instrumental variable method relies on some external source of variation or IV to determine treatment status. The IV influences the likelihood of participating in a program, but it is outside of the participant’s control and is unrelated to the participant’s characteristics” (World Bank, 2016, p. 328).
Internal validity	“An <i>evaluation</i> is internally valid if it provides an accurate estimate of the <i>counterfactual</i> through a valid <i>comparison group</i>” (World Bank, 2016, p. 329).
Matching	“A nonexperimental <i>impact evaluation</i> method that uses large data sets and statistical techniques to construct the best possible <i>comparison group</i> for a given <i>treatment group</i> based on observed characteristics” (World Bank, 2016, p. 329).
Mixed methods	“A mixed-method evaluation systematically integrates two or more evaluation methods, drawing on both quantitative and qualitative data, incorporating, for example, both randomized control trial experiments and case studies and different types of data collection techniques (e.g. structured observations, key informant interviews, household surveys, and

TERM	DEFINITION WITHIN THE GCF CONTEXT
	reviews of existing secondary data)” (Green Climate Fund, 2021, p. 17).
Randomized controlled trials (RCTs)	Impact evaluation method whereby every eligible <i>unit</i> (for example, an individual, household, business, school, hospital, or community) has a probability of being selected for treatment by a program. With a sufficiently large number of <i>units</i> , the process of randomized assignment ensures equivalence in both observed and unobserved characteristics between the <i>treatment group</i> and the <i>comparison group</i> , thereby ruling out <i>selection bias</i> . Randomized assignment is considered the most robust method for estimating <i>counterfactuals</i> and is often referred to as the gold standard of impact evaluation (World Bank, 2016, p. 332).
Regression Discontinuity Design (RDD)	“A <i>quasi-experimental</i> impact evaluation method that can be used for programs that rely on a continuous index to rank potential participants and that have a cutoff point along the index that determines whether potential participants are eligible to receive the program or not. The cutoff threshold for program eligibility provides a dividing point between the <i>treatment group</i> and the <i>comparison group</i> . Outcomes for participants on one side of the cutoff are compared with outcomes for nonparticipants on the other side of the cutoff. When all units comply with the assignment that corresponds to them on the basis of their eligibility index, the RDD is said to be ‘sharp.’ If there is noncompliance on either side of the cutoff, the RDD is said to be ‘fuzzy’” (World Bank, 2016, p. 332).
Theory of change (TOC)	“A theory of change is a method that explains how a given intervention, or set of interventions, is expected to lead to specific ... change [in this case, evidence-based decision-making], drawing on a causal analysis based on available evidence. ... A theory of change also helps to identify the underlying assumptions and risks that will be vital to understand and revisit throughout the process to ensure the approach will contribute to the desired change” (United Nations Development Group, 2017, p. 4).

Sources: Paul J. Gertler et al., *Impact Evaluation in Practice*, 2nd ed. (World Bank, 2016); Green Climate Fund, *Evaluation Policy for the GCF* (Green Climate Fund, 2021); Organisation for Economic Co-operation and Development, *Glossary of Key Terms in Evaluation and Results-Based Management for Sustainable Development*, 2nd ed. (OECD Publishing, 2023); United Nations Development Group, *United Nations Development Assistance Framework Guidance* (United Nations Development Group, 2017).

Annex II. Suggested checklist for application of GCF Impact Evaluation Standards

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
1. Commissioning, planning and design			
Is the impact evaluation justified by clear GCF decision-making, accountability or learning needs?			Relevance
Has the suitability, feasibility and appropriateness of conducting an impact evaluation been assessed?			Relevance
Is the intervention sufficiently stable, well-defined and measurable to support credible impact evaluation?			Credibility
Was the impact evaluation considered sufficiently early in the programme cycle to support baseline design, counterfactual construction and data-collection planning?			Credibility
Is the timing of the evaluation appropriate given the expected results pathway and maturity of outcomes?			Relevance
Have relevant stakeholders been appropriately engaged in defining the evaluation scope and design?			Relevance
Are the intended users and uses of the impact evaluation clearly identified?			Utility
Is the evaluation aligned with identified GCF policy, programming or strategic evidence needs?			Utility
Is the evaluation designed to inform implementation, scaling, replication or future investment decisions, where relevant?			Utility
Are evaluation questions, outputs and timelines aligned with the intended objectives and uses of the evaluation?			Utility
Is a credible and appropriate identification strategy specified and justified?			Credibility

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
Is the counterfactual approach clearly defined, justified and feasible in the implementation context?			Credibility
Is there adequate opportunity to establish a credible counterfactual and collect appropriate baseline or comparison data?			Credibility
Is the selected evaluation method appropriate to the intervention, context, ethical considerations and available data?			Credibility
Where appropriate, does the design incorporate mixed methods to strengthen interpretation and triangulation?			Credibility
Has the unit of analysis (e.g. household, community, institution, ecosystem) been clearly specified and justified?			Credibility
Is the theory of change explicitly linked to outcomes, assumptions and hypotheses to be tested?			Credibility
Are key outcome indicators measurable, relevant and aligned with the theory of change and climate objectives?			Relevance/Credibility
Have confounding variables and external influencing factors been identified and considered in the design?			Credibility
Are sampling, statistical power and key risks to validity identified and adequately addressed?			Credibility
Are anticipated implementation risks and operational constraints adequately considered in the evaluation design?			Credibility
Is a comprehensive data collection and management plan established for both treatment and comparison groups?			Credibility
Are baseline, follow-up and longitudinal data requirements clearly identified?			Credibility

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
Are quality assurance and technical oversight arrangements clearly defined?			Credibility
Do the terms of reference and inception report clearly document methods, assumptions, analytical choices and limitations?			Transparency
Are procedures for documenting deviations from the original evaluation design clearly specified?			Transparency
Is there a clear plan for documenting instruments, data sets, codebooks and analytical decisions?			Transparency
Are evaluation management, oversight and implementation responsibilities clearly defined?			Transparency
Has an ethics risk assessment been conducted, including risks specific to impact evaluation designs?			Ethics
Are informed consent, confidentiality and data-protection procedures clearly defined?			Ethics
Does the evaluation design adhere to the do-no-harm principle and consider equity and inclusion, particularly for vulnerable groups?			Ethics
Are independence and conflict-of-interest considerations appropriately addressed?			Ethics/Transparency
2. Implementation and data collection			
Is the impact evaluation implemented in accordance with the approved design and workplan?			Credibility
Is treatment assignment or comparison group selection implemented with fidelity?			Credibility
Are attrition, spillovers, non-compliance and contextual changes monitored and documented?			Credibility
Are major implementation deviations, delays or operational disruptions systematically documented?			Transparency
Are operational realities appropriately considered when interpreting emerging findings?			Credibility

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
Are data quality assurance procedures clearly defined and implemented?			Credibility
Are data validation protocols and quality control mechanisms implemented throughout data collection?			Credibility
Are data management systems sufficiently secure, documented and maintained to support analysis and reproducibility?			Transparency
Is coordination maintained between evaluators, implementing entities and stakeholders while preserving evaluator independence?			Credibility/Ethics
Is informed consent obtained and respected in practice?			Ethics
Are safeguarding risks monitored and addressed during implementation and fieldwork?			Ethics
Are data privacy and secure data management protocols followed?			Ethics
Are data-collection instruments, protocols and codebooks clearly documented?			Transparency
Are implementation decisions, adaptations and deviations systematically recorded?			Transparency
3. Analysis and reporting			
Are findings clearly linked to the stated evaluation objectives, questions and theory of change?			Credibility
Are robustness checks, sensitivity analyses, assumptions and key limitations adequately reported and discussed?			Credibility
When appropriate, is qualitative or mixed-methods evidence used to support and interpret quantitative findings?			Credibility

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
Are implementation realities and contextual factors adequately reflected in the interpretation of findings?			Credibility
Are threats to internal validity, including selection bias, spillovers and measurement bias, adequately discussed?			Credibility
Are limitations regarding external validity and transferability of findings appropriately discussed?			Credibility
Are assumptions, uncertainties and contextual limitations transparently documented?			Transparency
Are methods described in sufficient detail to understand how findings were produced?			Transparency
Are deviations from the original design transparently disclosed and explained?			Transparency
Are positive, negative, null and unintended findings reported without selective emphasis?			Transparency
Are analytical methods, assumptions, data sets and limitations sufficiently documented to support verification or replication, where appropriate?			Transparency
Are findings translated into actionable implications for GCF policy, programming, operations or future climate investments?			Utility
Are conclusions, lessons and recommendations clearly tailored to intended users?			Utility
Are findings documented in a manner that supports future synthesis, cross-project learning or evidence aggregation?			Utility
Does reporting protect confidentiality and avoid identification of individuals or communities?			Ethics
Are findings presented in a manner that minimizes risks of harm, stigmatization or misinterpretation?			Ethics

CHECKLIST	YES / PARTIALLY / NO	COMMENTS	RELEVANT STANDARDS
4. Dissemination, learning and use			
Is engagement with intended users planned to support interpretation and uptake of findings?			Utility
Are dissemination products tailored appropriately to different intended audiences?			Utility
Is there a plan to integrate evaluation findings into future project design, implementation or strategic learning processes?			Utility
Are dissemination approaches designed to support learning among accredited entities, implementing partners and relevant stakeholders?			Utility
Are findings communicated in formats accessible to both technical and non-technical audiences?			Utility
Is it clear what information will be disclosed, to whom and for what purpose?			Transparency
Are dissemination approaches consistent with applicable disclosure, confidentiality and data-protection requirements?			Transparency
Does dissemination adequately mitigate risks of harm, stigmatization or unintended disclosure?			Ethics
Are data-sharing practices consistent with informed consent, confidentiality and do-no-harm principles?			Ethics

Annex III. References

- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. *Impact Evaluation in Practice*. 2nd ed. World Bank, 2016. <https://hdl.handle.net/10986/25030>.
- Green Climate Fund. *Evaluation Policy for the GCF*. Green Climate Fund, 2021. <https://www.greenclimate.fund/document/evaluation-policy-gcf>.
- Green Climate Fund. *Information Disclosure Policy*. Green Climate Fund, 2016. <https://www.greenclimate.fund/document/information-disclosure-policy>.
- Independent Evaluation Unit. *Green Climate Fund Evaluation Standards*. Green Climate Fund, 2022. <https://ieu.greenclimate.fund/document/green-climate-fund-evaluation-standards>.
- Organisation for Economic Co-operation and Development. *Glossary of Key Terms in Evaluation and Results-Based Management for Sustainable Development*. 2nd ed. OECD Publishing, 2023. <https://doi.org/10.1787/632da462-en-fr-es>.
- United Nations Development Group. *United Nations Development Assistance Framework Guidance*. United Nations Development Group. 2017. <https://unsdg.un.org/resources/united-nations-development-assistance-framework-guidance>.
-