The Third Evaluability assessment of the Green Climate Fund's Funding Proposals

Alejandro González-Caro, Peter Mwandri, Andreas Reumann, Galyna Uvarova, Susumu Yoshida

05/2025

175, Art center-daero
Yeonsu-gu, Incheon 22004
Republic of Korea
Tel. (+82) 10-3458-6476
Email: ieu@gcfund.org
https://ieu.greenclimate.fund

First Edition

This evaluation is a product of the Independent Evaluation Unit at the Green Climate Fund (GCF/IEU). It is part of a larger IEU effort to provide open access to its research and work and to make a contribution to climate change discussions around the world.

While the IEU has undertaken every effort to ensure the data in this report is accurate, it is the reader's responsibility to determine if any and all information provided by the IEU is correct and verified. Neither the author(s) of this document nor anyone connected with the IEU or the GCF can be held responsible for how the information herein is used.

Citation
The citation details for this evaluation are as follows:
Independent Evaluation Unit (2025). *The Third Evaluability Assessment of the Green Climate Fund's Funding Proposals* (May). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund.

Credits
Head of the GCF Independent Evaluation Unit: Andreas Reumann
Task manager: Susumu Yoshida, Evaluation Specialist - Implementation Science, Independent Evaluation Unit
Editing: Toby Pearce
Layout: Ewnetu Kefale; Josephine Ngala
Cover design: Therese Gonzaga

A FREE PUBLICATION

Printed on eco-friendly paper

Acknowledgements

The authors express their sincere gratitude to SISTME for their meticulous work on data extraction and management, which provided the essential foundation for this evaluability assessment. Our appreciation extends to the IEU evaluation team members who contributed to various stages of this study.

Special thanks to Daisuke Horikoshi and Genta Konci from the IEU for their insightful review and constructive feedback, which significantly enhanced the analytical depth and practical applications of this assessment.

This report represents a collaborative effort to advance the GCF's learning agenda at a critical juncture in climate finance. By mapping the evolving landscape of project evaluability, we aim to contribute to the larger conversation about measuring what matters in climate action, ensuring that each investment not only addresses immediate climate needs but also builds our collective knowledge about effective interventions. The authors are grateful to all who have contributed to this pioneering work in strengthening the evidence base for climate finance decisions. All remaining errors are the sole responsibility of the authors.

List of authors

The authors of this paper are listed in alphabetical order by last name.

Full name	Affiliation
Alejandro González-Caro	Independent Evaluation Unit, Green Climate Fund
Peter Mwandri	Independent Evaluation Unit, Green Climate Fund
Andreas Reumann	Independent Evaluation Unit, Green Climate Fund
Galyna Uvarova	Independent Evaluation Unit, Green Climate Fund
Susumu Yoshida	Independent Evaluation Unit, Green Climate Fund

Table of Contents

List of Tables

List of Figures

List of Boxes

Abbreviations

AE	Accredited entity
DAE	Direct access entity
EARF	Energy Access Relief Facility
FAA	Funded activity agreement
FP	Funding proposal
GCF	Green Climate Fund
IAE	International accredited entity
IE	Impact evaluation
IEU	Independent Evaluation Unit
LDC	Least developed country
M&E	Monitoring and evaluation
SAP	Simplified Approval Process
SIDS	Small island developing State
ToC	Theory of change

I. Introduction

1. The Green Climate Fund (GCF) is a multilateral fund created to make significant and ambitious contributions to global efforts to combat climate change. The GCF contributes to achieving the objectives of the United Nations Framework Convention on Climate Change and the Paris Agreement. It aims to promote a paradigm shift towards low-emission and climate-resilient development pathways by helping developing countries reduce their greenhouse gas emissions while supporting countries' specific needs in adapting to and combating climate change's adverse effects. For developing countries, the GCF provides support through various financial modalities, including grants, loans, and market instruments such as bonds and equity.

2. Following the first assessment conducted in 2019, this document assesses the quality of proposals approved for financing by the GCF (also called funding proposals (FPs)). It asks the following question: To what extent are GCF-supported programmes and projects capable of credibly reporting their impacts, efficiency, and effectiveness in an evidence-based and robust way?

3. We ask this question for two reasons. First, the GCF's overall goal is to support a paradigm shift towards low-carbon, high-resilience pathways. Therefore, it is critical to understand if a paradigm shift is occurring and how much of this shift is attributable to the GCF. The GCF's contribution to the shift requires GCF investments to credibly commit to and measure the results to which they statedly aim. Second, measurement in the climate change space is difficult. Climate change action requires that large numbers of people act simultaneously to individually effect change that together must represent a large enough and critical change to make a difference. Results from individual actions on overall global climate change will only be apparent after hundreds of years, if not longer. However, we can assess the extent to which current investments are likely to yield these results. It is important the GCF examines projects for the likelihood of these results. This is to understand the probability of success and the credibility of results reporting (should it occur) and enable the GCF to reliably report its overall contribution to this climate action effort. It is even more important to assess, test and establish the credibility of these results.

4. The GCF invests its resources using several criteria. Among these are the investment criteria, which require that projects show proof of impact potential, sustainability, paradigm shift potential, country ownership, climate relevance, effectiveness, and efficiency. These are also among the criteria the GCF's Independent Evaluation Unit (IEU) uses to assess the quality of the GCF portfolio's performance, activities, and results.

5. This study presents the results of an IEU desk assessment of the GCF portfolio. The study builds on the findings of the first assessment conducted in 2019, as well as the second one conducted in 2022. [1, 2] The study has two main aims. Firstly, to assess the quality of the proposals for the FPs the GCF has approved and is currently supporting. Project managers can learn from these and produce stronger proposals in the future that have a higher likelihood of reporting measured results and a greater likelihood of achieving success. Secondly, the study aims to inform the GCF investment criteria, introduce evidence-based learning opportunities into GCF projects and processes, and inform the implementation and overall impact of GCF resources.

[1] Nathan Fiala, Jyotsna Puri and Peter Mwandri (2019). Becoming bigger, better, smarter: A summary of the evaluability of Green Climate Fund proposals. Working Paper No. 1. Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund. Available at https://ieu.greenclimate.fund/sites/default/files/document/working-paper-becoming-bigger-better-smarter-summary-evaluability-gcf-proposals.pdf.
[2] Independent Evaluation Unit (2022). Evaluability assessment of the Green Climate Fund funding proposals. IEU Learning Paper (December). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund. Available at https://ieu.greenclimate.fund/sites/default/files/document/230214-evaluability-study-top_2.pdf.

6. These two aims help us meet three purposes. First, to help inform, where possible, risks that may arise in currently supported projects and to alert project managers. Second, to improve the quality of proposals overall. Third, to help projects measure better and discuss methods FPs may use for this purpose. Hopefully, this discussion and the use of robust methods will enable the GCF to report its overall impact measurably and credibly.

II. Summary of GCF's funded projects as of 31 December 2023

2.1 Overview of GCF's portfolio

7. Projects represent the primary mechanism through which the GCF invests in low-emission, high-resilience development pathways. All GCF-supported activities must demonstrate climate rationale to receive funding. The GCF aims to drive paradigm shifts in both climate mitigation and climate adaptation efforts. Approved projects are classified into three categories: mitigation, adaptation, and cross-cutting:

(i) Mitigation projects that help developing countries reduce their greenhouse gas emissions

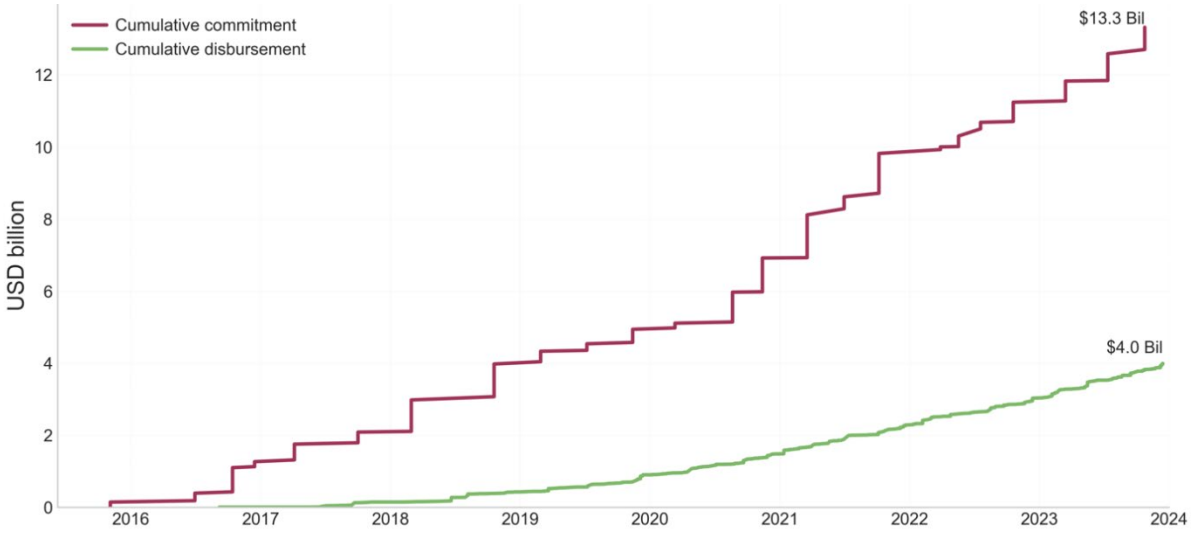
(ii) Adaptation projects that enhance countries' abilities to withstand climate and weather shocks while increasing community resilience

(iii) Cross-cutting projects that simultaneously address both mitigation and adaptation objectives.

8. This analysis examines the GCF's approved project portfolio, consisting of 241 projects (including both FPs and Simplified Approval Process (SAP) projects), approved up to the thirty-seventh meeting of the Board (B.37) in 2023. This total reflects all approved projects maintained in the portfolio excluding: FP031 (never submitted); FP032, FP079, FP088 (now FP110), and FP123 (withdrawn by accredited entities (AEs)); FP029, FP030, FP006, FP038, FP054, FP065, and FP104 (approvals lapsed).

9. Of the 241 projects in the approved portfolio, 206 (approximately 85 per cent) have effective funded activity agreements (FAAs) and have entered the implementation phase, while the remainder were still in post-approval stages. The portfolio has achieved a 30 per cent disbursement rate, with USD 4 billion disbursed for funded activities as of 31 December 2023 (see **Figure 1**).

Figure 1. Cumulative GCF commitment and disbursement



Source: GCF Project API data as of 31 December 2023 (accessed 28 February 2024), including all projects approved through B.37.

2.2 Portfolio by theme – adaptation vs. mitigation funding

10. In this analysis, adaptation funding encompasses both standalone adaptation projects and the adaptation component of crosscutting projects. Similarly, mitigation funding includes standalone mitigation projects and the mitigation component of crosscutting projects. This approach provides a comprehensive view of how resources are allocated across climate action themes. Mitigation activities make up 56 per cent of the total funding in the GCF approved portfolio, while adaptation accounts for 44 per cent of the funding, as per the cutoff date of 31 December 2023. This distribution underscores several key trends in climate finance allocation:

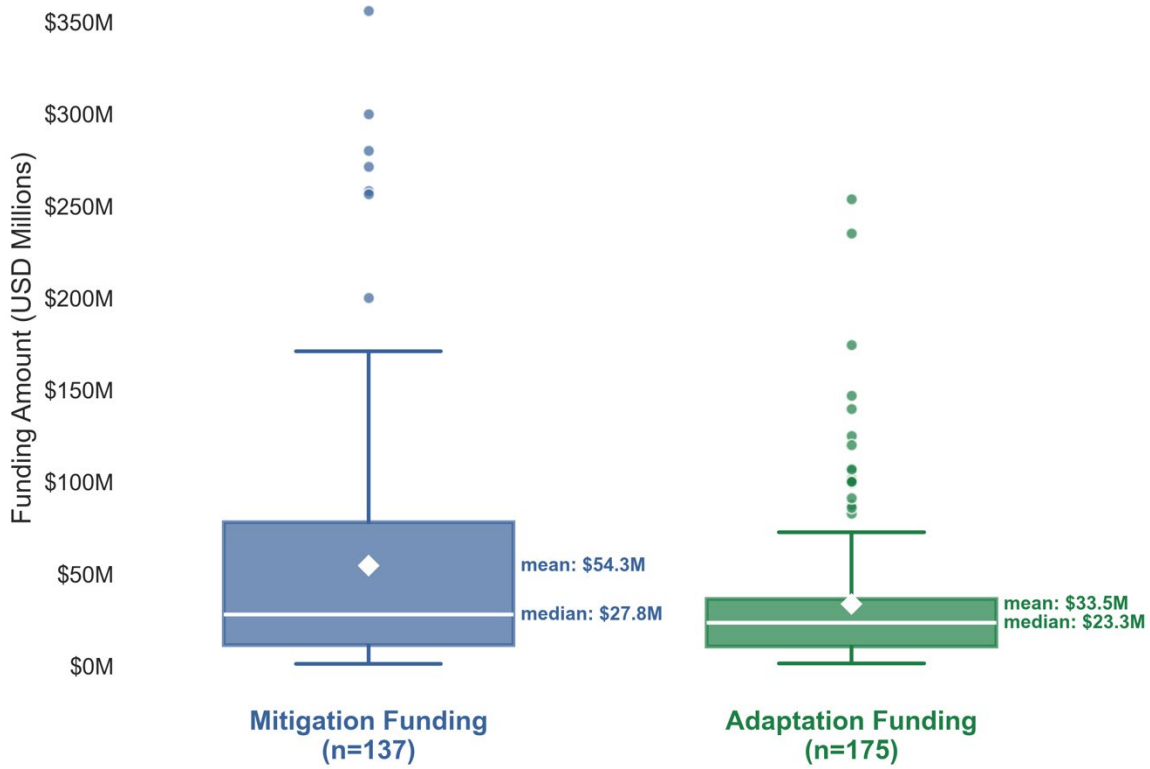
(a) **Project range**: Mitigation funding ranges from USD 0.9 million to USD 356 million, while adaptation funding ranges from USD 1.2 million to USD 253.8 million.

(b) **Average funding**: The average funding for mitigation projects is substantially higher (USD 54.3 million) compared to adaptation projects (USD 33.5 million), indicating strategic emphasis on larger-scale mitigation investments.

(c) **Median values**: The median funding for adaptation projects is USD 23.3 million, while that of mitigation projects is USD 27.7 million. This relatively moderate difference in medians indicates that the higher average funding for mitigation is driven by some very large projects.

(d) **Total GCF approved portfolio**: The analysis encompasses a total of 241 projects in the approved portfolio, with 175 projects having adaptation components (totalling USD 5,860.2 million) and 137 projects with mitigation components (amounting to USD 7,440.5 million). It is important to note that these figures include crosscutting projects that address both adaptation and mitigation objectives.

Figure 2. GCF funding distribution and comparison of adaptation and mitigation projects



Source: GCF Project API data as of 31 December 2023 (accessed 28 February 2024), including all projects approved through B.37.

Note: The sum of mitigation and adaptation funding does not equal the total number of approved projects up to B.37 (241), as some approved projects are cross-cutting, incorporating both mitigation and adaptation components.

11. The boxplot visualization shows the greater spread and higher outliers in mitigation funding compared to adaptation funding, while also showing the difference in median values. The white diamonds represent the mean values, illustrating how these are pulled upward by large outliers, particularly in the mitigation portfolio.

2.3 Funding by GCF region

12. There is a significant disparity between the number of approved FPs and the volume of finance allocated per GCF region. Although Africa and the Asia-Pacific region together claim the most proposals, their per-project finance is not necessarily the highest. For instance, Latin America and the Caribbean has fewer proposals (62), yet its total allocation of USD 3.2 billion translates to a level of per-project funding on par with Africa – both are around USD 51 million per proposal – while the Asia-Pacific region averages closer to USD 45 million per proposal. These differences suggest that factors beyond the number of proposals – such as project size, complexity, and region-specific funding priorities – can significantly influence how finance is ultimately allocated.

Table 1. Distribution of the GCF portfolio by region

GCF REGION	NUMBER OF FPS	FINANCE VOLUME (USD MIL)	SHARE OF TOTAL (%)
Africa	101 (42%)	5,184	39%
Asia-Pacific	101 (42%)	4,513	34%
Eastern Europe	13 (5%)	436	3%
Latin America and the Caribbean	62 (26%)	3,185	24%
Total	241*	13,317	100%

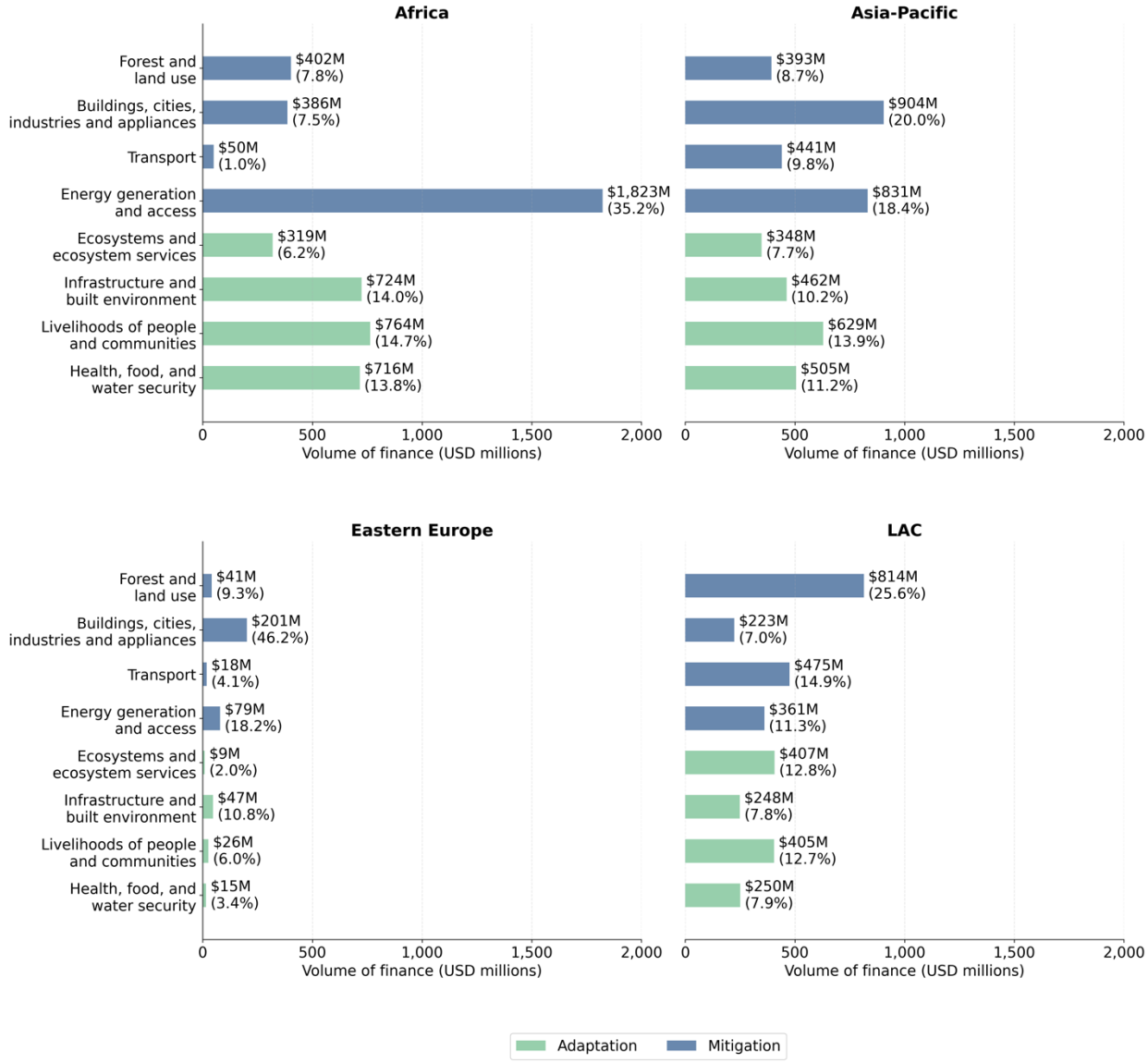
Source: GCF Project API data as of 31 December 2023 (accessed 28 February 2024), including all projects approved through B.37.

Note: * The number of FPs up to B.37 does not aggregate into a total due to some approved proposals covering several regions

13. The breakdown of the GCF's investments across its mitigation and adaptation result areas offers valuable insights into its allocation strategies and priorities. Among adaptation interventions, the largest investment – USD 1.7 billion – is directed towards enhancing the livelihoods of people and communities. This allocation reflects the GCF's commitment to strengthening community resilience as a key strategy for mitigating climate change impacts. Additionally, significant investments support infrastructure and the built environment, with USD 1.4 billion allocated, and USD 1.3 billion directed toward health, food, and water security. These allocations highlight the Fund's focus on strengthening essential services and resilience mechanisms in vulnerable regions.

14. Regarding mitigation interventions, GCF investments are predominantly directed towards energy generation and access, with a substantial sum exceeding USD 3 billion. This significant investment reflects the Fund's commitment to advancing renewable energy sources and improving access to clean energy technologies, crucial steps in reducing greenhouse gas emissions.

Figure 3. Result areas finance by region



Source: GCF Project API data as of 31 December 2023 (accessed 28 February 2024), including all projects approved through B.37.

2.4 Breakdown of funding by energy vs. non-energy projects

15. The analysis in **Table 2** presents a classification of projects based on their nature and categorization, encompassing both energy-related and non-energy projects.[3] Under the energy category, projects are subdivided into energy access and generation, energy efficiency, and mixed energy. Overall, energy-related projects comprise 43 per cent of the GCF portfolio, and non-energy projects comprise 57 per cent of the GCF portfolio.

[3] Independent Evaluation Unit (2024). Independent Evaluation of the Green Climate Fund's Energy Sector Portfolio and Approach. Evaluation report No. 17 (February). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund. Available at https://ieu.greenclimate.fund/evaluation/ES2023.

Table 2. GCF portfolio by energy vs. non-energy categorization

NATURE OF THE PROJECT	CLASSIFICATION	NUMBER OF PROJECTS
Energy	Energy access and generation	50
	Energy efficiency	25
	Mixed energy	24
Subtotal		**99 (43%)**
Non-energy		129 (57%)
Total		**228**

Source: GCF Tableau server as of 15 July 2023. Categorization by the Energy evaluation team.

III. Methods overview

16. This section discusses the methods used to assess proposals submitted and approved for funding by the GCF. We ask, "To what extent are approved projects likely to be able to inform results that they claim in a credible and measured manner?"

17. Board-approved FP proposals were assessed along several dimensions. These dimensions are grouped into several common topics.[4] An important caveat here is that the comments in this study are not indicative of the capacity or ability of proposal submitters, mainly because – to the extent we know – the attributes we analysed were neither GCF requirements nor funding conditions when these proposals were prepared. Inferences made in this paper are made from data and information in the proposals. They indicate whether proposals and the information within them are fit for purpose in fulfilling the GCF's objectives. While noting this is a desk review, our guiding principle is only to examine what has been submitted in proposals. We also provide constructive comments on how these proposals may be improved and comment mainly on the potential "internal validity" of these projects – that is, the feasibility and ability of the projects to inform the results/changes they aspire to, as stated in their proposals.

3.1 Lenses to assess whether proposals are fit for purpose

18. We use four lenses to assess the potential for the internal validity of FPs. The first is the **theory of change** (ToC). We assess whether proposals include an explicit discussion of the project's overall ToC. A clear ToC is critical for understanding if the proposed activities will lead to the investment's intended outcomes, including their size. In many FPs, ToCs are not laid out coherently in a single place. In all cases, we analysed the full proposal and piece together an "implicit" ToC if an explicit one is omitted.

19. Second, to understand if the programme activities can achieve the impacts claimed in the proposal, we examine proposals for their potential to **measure** and report **causal change and report results using impact measurement**. Many proposals make bold claims about what

[4] The individual assessment documents present our detailed comments on each of the proposals and are available on request.

their investments can accomplish. This causality and impact lens enables us to objectively estimate if the project investment will cause the claimed impact(s) or whether some proportion of the anticipated effect would still have occurred in the absence of programme activities/GCF investment. Observing this counterfactual scenario is impossible: we cannot observe what happens to a beneficiary if it receives a project intervention and what happens if it does not. But there is now a sizeable discipline showing this can be done using experimental construction or observation of valid comparison groups to generate accurate estimates of causal impacts.[5, 6]

20. Third, we assess the project's ability to inform the GCF investment criteria credibly. We assess the extent to which the proposal's credibly responds to fulfilling the GCF's required investment criteria, including whether the proposal demonstrates a strong potential for delivering impact, paradigm shift, sustainable development, recipient needs, country ownership, effectiveness, and efficiency. Specifically, we recognize that **targeting** is a primary concern for many projects. We assess each project based on its targeting criteria. For instance, if a programme plans to reach 50 per cent of women or vulnerable groups, we examine if the proposal has articulated its targeting criteria clearly, and the extent to which the programme is likely to achieve this goal, based on the programme model and ToC.

21. Fourth, we examine FPs for how well they have set up systems to help report on their progress and their **fidelity to implementation** plans. We also examine their stated monitoring and evaluation (M&E) systems to see if they are sufficient in their current state to assess the projects' capabilities in this area.

3.2 Building a stoplight

22. To illustrate results, we built a stoplight for each FP proposal that summarizes risks and other issues related to results measurement and the information presented in each FP. Four criteria inform the stoplight. We use a likelihood or risk framework for each criterion to assess the quality with which the proposal meets each requirement (see **Box 1**). The following decision rule is used:

(a) If the FP has done well on a criterion, and it is highly likely that the criterion will be achieved, the proposal is marked as "low risk" for that criterion.

(b) If, based on the information provided in the FP, there appears to be a moderate probability the proposed programme or project will perform well relative to the stoplight criterion, then the proposal is marked as "medium risk" for that criterion.

23. A proposal is marked as "high risk" for a given criterion if there appears to be a high probability that the proposed programme or project will not perform well relative to the criterion.

[5] Paul J. Gertler and others (2016). *Impact Evaluation in Practice*, 2nd ed. (Washington, D.C.: World Bank, 2016). Available at https://openknowledge.worldbank.org/bitstream/handle/10986/25030/9781464807794.pdf?sequence=2&isAllowed=y;

[6] Emmanuel Jimenez and Jyotsna Puri (2017). The Wicked Cases of Education and Climate Change: The Promise and Challenge of Theory-Based Impact Evaluations. In: Rob D. van den Berg, Indran Naidoo, and Susan D. Tamondong (eds.), Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability. Exeter, UK: International Development Evaluation Association (IDEAS). Available at https://sdgs.un.org/sites/default/files/publications/2455IDEASwebREV_08Dec.pdf.

Box 1. Why use a risk framework?

The stoplight assessments associated with each GCF-funded proposal are constructed based on the information provided within the proposals. The proposals submitted to the GCF do not include every minute detail about the proposed project or programme. The GCF recognizes that the information in the proposals may be further adjusted based on feedback from the GCF, resulting from the evolving needs of target recipients or ongoing M&E efforts during implementation. Because the proposals are used as input for evaluating proposed projects or programmes, the project's quality vis-à-vis each stoplight criterion cannot be evaluated with absolute certainty before implementation. However, projects and programmes can be evaluated in terms of the likelihood they will meet each stoplight criterion based on the information in the proposal. Because the assessments gauge probabilities of success rather than the observed performance against the stoplight criteria, a risk framework provides a useful assessment tool. As described above, a project is rated as "high risk" for a given stoplight criterion when there is a high probability that the project described in the proposal will not adequately perform relative to that criterion. Alternatively, a "low risk" rating corresponds to a low probability of poor performance against a given criterion. This framework recognizes that our assessments are not based on observed progress but on the projected success of the proposed projects and programmes.

3.3 Theory of change and discussion of causal pathways

24. We use the following questions and rating rules to assess the quality of the ToCs and causal pathways discussed in the FPs.

(a) **What is the quality of the (implicit or explicit) theories of change and programme logic? (See the annex for a ToC checklist.)**

- (i) Low risk. ToC is well articulated.
- (ii) Medium risk. Logic framework or ToC is present but needs some clarification. (Missing information is specified.)
- (iii) High risk. Logic framework or ToC either does not exist, or it exists but relies on unverified assumptions or is missing critical details about implementation and/or causal pathways. (Missing information is specified.)
- (iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating the ToC.

(b) **Are unintended consequences referred to and identified robustly in the programme ToC and/or in the surrounding literature reviews?**

- (i) Low risk. Unintended consequences are well articulated. (These are drawn from discussion of the ToC.)
- (ii) Medium risk. Unintended consequences are discussed but need some clarification. (Missing information is specified.)
- (iii) High risk. Unintended consequences are not discussed and are potentially very large, given the programme design. (Missing information is specified.)
- (iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating how it addresses unintended consequences.

(c) **Are causal pathways clearly identified and discussed? (This is discussed in the context of the ToC and the credibility and feasibility of the pathways.)**

(i) Low risk. Causal pathways are well articulated and supported with credible evidence.

(ii) Medium risk. Causal pathways are described or implied, but the proposed links need some clarification about the assumptions on which they rely. (Missing information is specified.)

(iii) High risk. The causal pathways implied in the proposal do not have a clear description and/or are based on unfounded assumptions.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents evaluating the proposed causal pathways adequately.

(d) **How robust are the causal linkages (implicit or explicit) and are they well informed by high quality evidence?**

(i) Low risk. Causal linkages are well articulated and are well informed by high quality evidence.

(ii) Medium risk. Causal linkages are discussed but need clarification and/or support by additional high-quality evidence. (Missing information is specified.)

(iii) High risk. Causal linkages are either not discussed or implied but lack any foundation in credible evidence. (Missing information is specified.)

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents evaluating the proposed causal pathways.

(e) **Is good quality evidence cited to discuss the efficacy of causal linkages?**

(i) Low risk. Evidence is of good quality and well-articulated.

(ii) Medium risk. Evidence is used but needs some clarification. (Missing information is specified.)

(iii) High risk. Evidence is not discussed, or the quality of the evidence cited is inferior. (Missing information is specified.)

(iv) Unclear. The quality of the evidence cited to discuss the efficacy of causal linkages is unclear.

3.4 Potential for measurement of casual change and evaluability

25. We ask the following questions to determine if causal change can be attributed to the GCF programme/GCF investment through impact evaluation (IE).

(a) **Does the proposal design allow for credible reporting of causal change?**

(i) Low risk. The proposal design allows for credible evaluation methods to be used to report casual change.

(ii) Medium risk. More details are needed to determine what could be a relevant comparison group or if there are feasible options to create comparison groups.

(iii) High risk. There does not appear to be a way to create a comparison group.

(iv) Unclear. There is not enough information to determine whether a credible measurement of causal change is possible.

(b) **To what extent are included requirements for M&E adequate and able to cover the costs of undertaking high quality IEs?**

(i) Low risk. Requirements for M&E are likely adequate to cover the costs of a high-quality evaluation.

(ii) Medium risk. Requirements for M&E are specified but are likely insufficient to support a high-quality IE.

(iii) High risk. Requirements for M&E are not specified or cannot be determined from the information provided.

(iv) Unclear. Information about the M&E requirements is ambiguous, making assessing this information impossible.

(c) **Activities included in the proposal focus on "economic analyses" and "overall M&E" – are these sufficient for high quality credible evaluations?**

(i) Low risk. Both are specified and are of high quality.

(ii) Medium risk. Both are specified but are of low quality. (Missing information is specified.)

(iii) High risk. Only one is specified or neither is specified. (Missing information is specified.)

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequate evaluation of the quality of proposed economic analyses and M&E activities.

(d) **Are methods for measuring attributable causal changes (outcomes, impact, or other) discussed?**

(i) Low risk. Measurement of attribution is well articulated.

(ii) Medium risk. Measurement of attribution is discussed and/or the need for causal impact measurement is acknowledged, but strategies for doing so are not well articulated. (Missing information is specified.)

(iii) High risk. Measurement of causal impact attribution is not discussed and/or the need for causal impact measurement is not acknowledged.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating any proposed methods for measuring attributable causal changes.

(e) **Are there potential areas of bias that are likely to creep in?**

(i) Low risk. There is a low risk of bias considering the proposed method of evaluating causal impact.

(ii) Medium risk. There is a medium risk of bias considering the proposed method of evaluating causal impact. We specify what could lead to biases.

(iii) High risk. There is a high risk of bias. The proposal either does not discuss a strategy for causal IE, or the strategy that is discussed has a high risk of producing unbiased impact estimates.

(iv) Unclear. Cannot judge the likelihood of bias due to insufficient information.

3.5 Implementation of fidelity and performance against investment criteria

26. We ask the following questions to determine if implementation and performance are likely to fit with the investment criteria.

(a) **Are eligibility and targeting criteria well-articulated in submitted documents?**

- (i) Low risk. Eligibility and targeting criteria are well articulated.
- (ii) Medium risk. Eligibility and targeting criteria are discussed but need some clarification. We specify the missing information.
- (iii) High risk. Eligibility and targeting criteria are either not discussed or are discussed but do not appear feasible, given the programme design. (Missing information is specified.)
- (iv) Unclear. Insufficient or ambiguous information in the proposal prevents evaluating eligibility and targeting criteria adequately.

(b) **Is there adequate and reliable information included in the proposal regarding implementation fidelity?**

- (i) Low risk. Implementation fidelity appears to be strong.
- (ii) Medium risk. There is a medium level of risk related to implementation fidelity. Some risks to implementation fidelity need to be addressed. (Missing information is specified.)
- (iii) High risk. There is a high level of risk related to implementation fidelity. Substantial risks need to be addressed. We specify the missing information.
- (iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating the information regarding implementation fidelity.

(c) **To what extent is impact potential identifiable and measurable in the proposal?**

- (i) Low risk. Impact potential is well articulated in the proposal and appears to be measurable using high quality methods.
- (ii) Medium risk. Impact potential is specified but needs some clarification. We specify the missing information. Impact potential is measurable, but high-quality methods may not be feasible given the programme design.
- (iii) High risk. Impact potential is specified, but it relies on significant assumptions that are not verified, and/or impact indicators are vaguely described. Measurement and evaluation potential appears to be low.
- (iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating the impact potential description and the feasibility of high-quality impact measurement.

(d) **To what extent is paradigm shift potential identifiable and measurable in the proposal?**

- (i) Low risk. Paradigm shift potential is well articulated in the proposal and appears to be measurable using high quality methods.
- (ii) Medium risk. Paradigm shift potential is specified but needs some clarification. (Missing information is specified.) Paradigm shift potential is measurable, but high-quality methods may not be feasible given the programme design.

(iii) High risk. Paradigm shift potential is specified, but it relies on significant assumptions that are not verified and/or paradigm shift indicators are vaguely described. Measurement and evaluation potential appears to be low.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating the description of the potential paradigm shift and the feasibility of high-quality measurement.

(e) **How well are other GCF investment criteria informed, and are these measurable and verifiable with high credibility and quality?**

(i) Low risk. Other investment criteria are likely to be credible.

(ii) Medium risk. Other investment criteria have some limitations. (Missing information is specified.)

(iii) High risk. Other investment criteria are not likely sufficient. We specify the missing information.

(iv) Unclear. The credibility of other investment criteria cannot be determined from the information provided.

3.6 Data collection and reporting credibility

27. We ask the following questions to determine if data collection and reporting will likely be of good quality.

(a) **Are current reporting requirements sufficient for regular M&E?**

(i) Low risk. Reporting for M&E is well articulated.

(ii) Medium risk. Reporting for M&E is discussed but needs some clarification. We specify the missing information.

(iii) High risk. Reporting for M&E is discussed, but it is insufficient for credible and useful M&E. We specify the missing information.

(iv) Unclear. The quality of reporting plans for M&E cannot be determined from the information provided.

(b) **How feasible is it to measure and report credibly on the progress of investment criteria, given M&E plans, budget and indicators for investment criteria?**

(i) Low risk. M&E and reporting plans have a high potential to measure progress on investment criteria.

(ii) Medium risk. M&E and reporting plans are discussed but are likely not of high enough quality or backed by sufficient resources to measure progress against investment criteria adequately.

(iii) High risk. M&E and reporting plans related to progress on investment criteria are not well articulated and/or clearly lack the resources needed to measure progress.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating the potential for the project to monitor and report on progress associated with investment criteria credibly.

(c) **To what extent did the proposal provide additional impact indicators beyond those proposed by the GCF? Can the proposal's indicators measure the magnitude of causal change?**

(i) Low risk. Indicators and measurements are well defined and can be used to measure impact.

(ii) Medium risk. Indicators and measurements lack specificity and measuring impact using the specified indicators may be challenging.

(iii) High risk. Indicators and measurements are vague and/or unclear. More detailed indicators are needed to measure impacts credibly.

(iv) Unclear. Insufficient information in the proposal to deduce the quality of indicators and measurements.

(d) **Have baseline data been collected and/or is there a requirement for this?**

(i) Low risk. Project will use baseline data, and the methods for collecting are well articulated.

(ii) Medium risk. Baseline data are discussed but need some clarification. Missing information to be specified.

(iii) High risk. Plans for collecting baseline data are not discussed despite a need to collect baseline data to inform an IE.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents adequately evaluating plans for collecting baseline data.

(e) **What is the potential quality of data and are they suitable for IEs?**

(i) Low risk. Data to be collected will be of high quality.

(ii) Medium risk. Data are likely to be of good quality.

(iii) High risk. Data are likely low quality, or data collection plans are not specified/unclear.

(iv) Unclear. Insufficient or ambiguous information in the proposal prevents evaluating the potential quality of data adequately.

28. We use these questions to assign each FP to high risk, medium risk, low risk or unclear categories for each of the questions. Assigning FPs to different risk categories should be made in conjunction with the individual project assessments (available on request). This will help the reader better understand the impact and feasibility of the proposals. Each approved project is assessed using these four lenses. The table summarizes the basic information of the proposal (number, implementer, period of funding, countries, and funding amount) and the results of the assessments, focusing on the ToC's quality, causal linkages, targeting strategy, and whether the proposal can rigorously inform the GCF investment criteria. We then aggregate these rankings. The overall results are discussed in the next section.

IV. Stoplight analysis – results

29. In this section, we present the key findings from the stoplight analysis of the GCF project portfolio as of 31 December 2023. We first analysed overall changes across the four assessment areas. Then, we show the overall trends, followed by comparing the portfolio between the Initial Resource Mobilization (IRM) period (2015–2019) and the GCF-1 period (2020–2023). Finally,

we provide a sector-specific analysis. In this report, we compare the energy and non-energy portfolios.[7]

30. As of B.37, the GCF has approved a total of 241 funded projects (both FP and SAP) after accounting for withdrawn, and lapsed projects. The following assessment examines these 241 funded projects, comprising 120 projects approved during the IRM period and 121 approved during the GCF-1 period.

4.1 Overall changes across the four assessment areas

31. The stoplight data are used to assess the quality and evaluability of the proposals at entry across four main areas as described earlier in the methods section. These areas are:

(i) Theory of change

(ii) Potential for measurement of causal change and evaluability

(iii) Implementation fidelity and performance against investment criteria

(iv) Data collection and reporting credibility

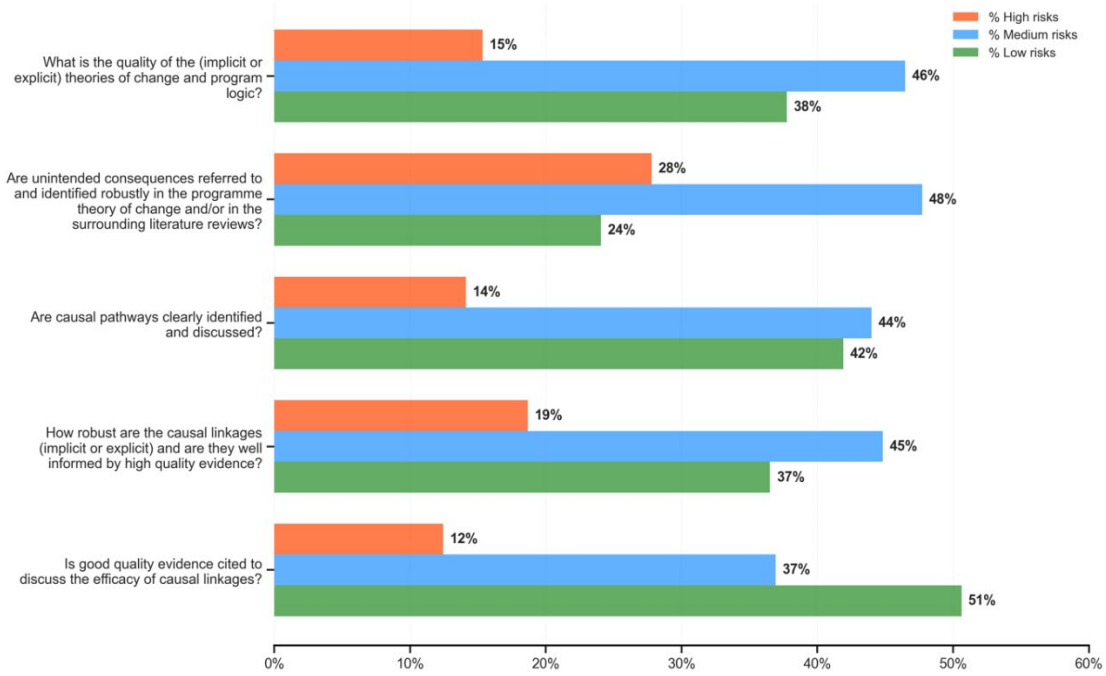
32. In category A, we assess the extent to which the pathways to impact are outlined using a ToC or logic model. The assessment seeks to determine if the ToC is explicit or implicit, to assess if the proposal identifies and/or is cognizant of the potential externalities of its requested financing, and to ascertain the robustness of the evidence cited and used to build their programme (see **Figure 4**).

33. The best-performing criterion (most low-risk evaluations) concerns whether good quality evidence is cited when discussing causal linkages. Areas concerning the identification of causal pathways and the quality of the ToC exhibit a more balanced distribution of risk levels, indicating variability in how well these aspects are typically handled within the GCF portfolio. Referral to unintended consequences in the ToC continues to be a challenge for many projects (28 per cent at high risk).

34. Overall, there is still a consistent high-risk rating in all the categories, suggesting areas where programme design or evaluation methodologies may need improvement to reduce uncertainties or vulnerabilities in their theoretical frameworks.

[7] Data last verified on 18 February 2025.

Figure 4. Stoplight assessment of the theory of change

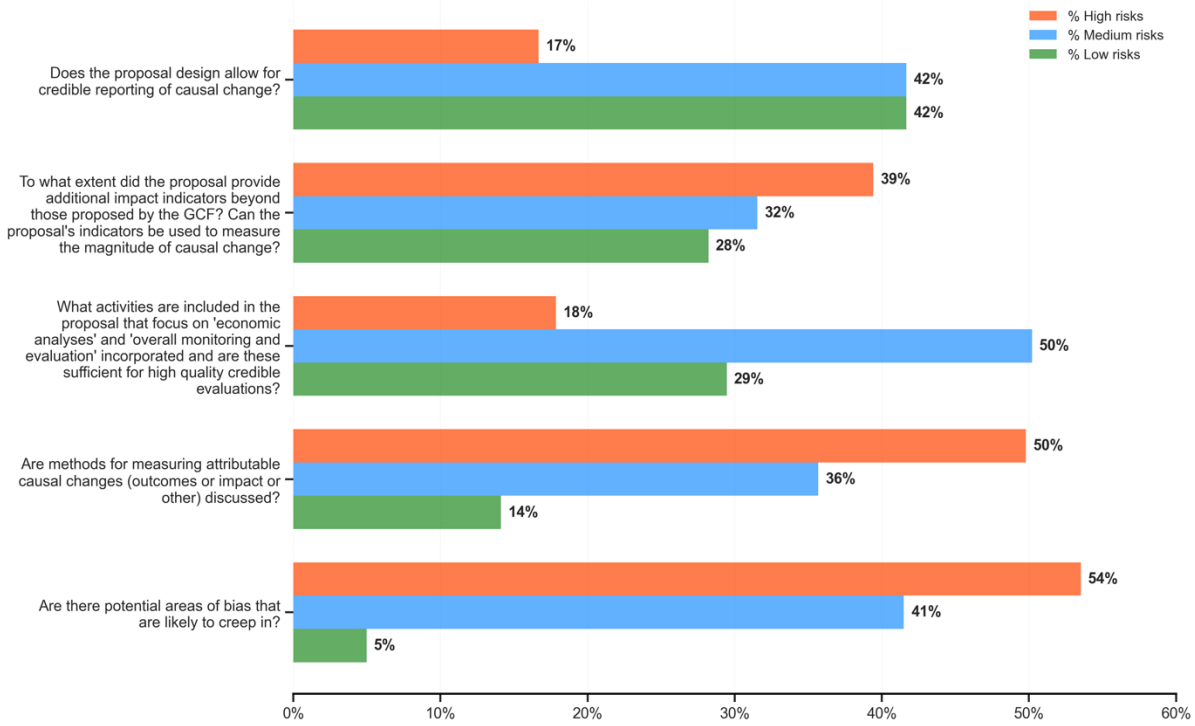


Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis performed by the authors.

35. In category B, we assess the ability to accurately measure, report and evaluate the economic impact and other changes due to the proposed activities. In other words, we seek to determine if the (claimed) causal effects of the proposed activities can be credibly measured by examining the FP's plans for M&E (see **Figure 5**).

36. The proposals generally indicate a high level of risk across most criteria in these categories, with particularly acute concerns regarding potential biases, where approximately 54 per cent are assessed as high risk. The areas involving economic analyses, monitoring and evaluation practices, and reporting of causal changes also display predominantly high risk, reflecting significant concerns about the adequacy and effectiveness of these proposals. Methods for measuring attributable causal changes also show a high risk in the majority, highlighting potential weaknesses in how outcomes or impacts are evaluated. Overall, these results suggest that there are considerable risks and areas for improvement in the measurement of causal changes and evaluability assessed in these proposals

Figure 5. Stoplight assessment of the potential for measurement of causal change and evaluability

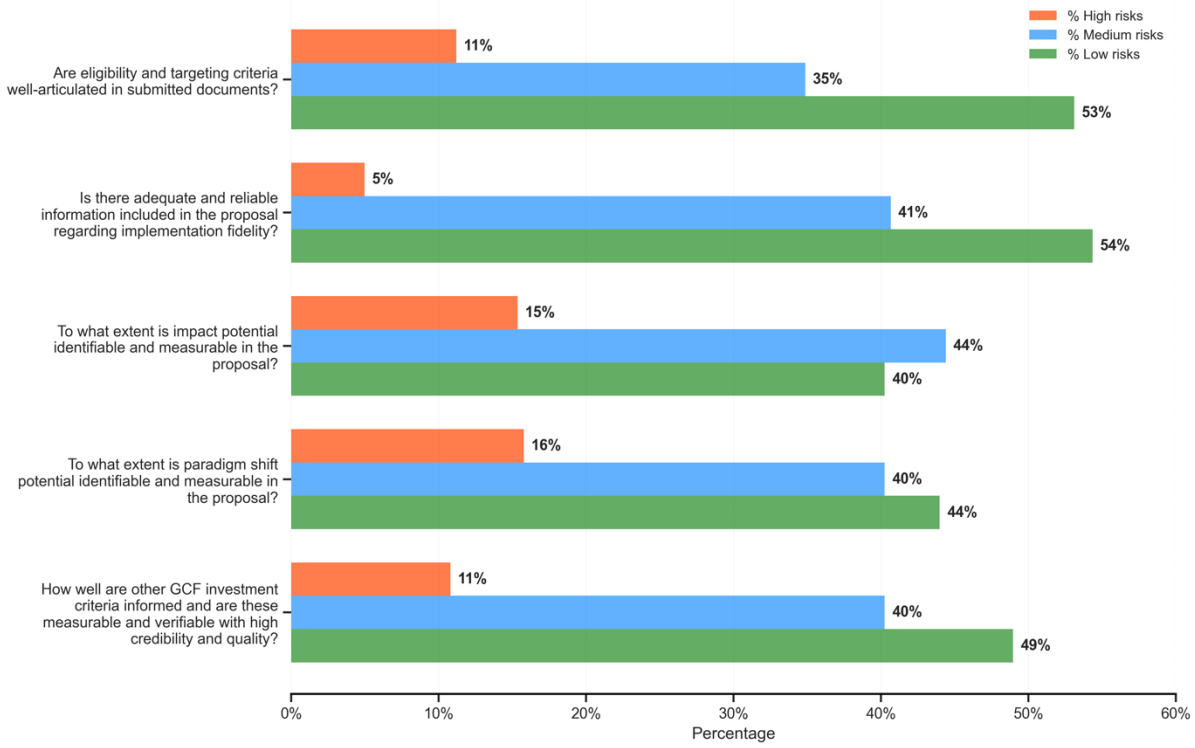


Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis performed by the authors.

37. In category C, we seek to assess if the project activities are well-targeted and to determine the programme's performance against GCF's investment criteria, the feasibility of the implementation plans for the overall programme, and if the proposal identifies relevant barriers to implementation and includes plans for recourse in the event of such constraints (see **Figure 6**).

38. The proportion of high risk is consistently lower across all the evaluation criteria, which reflects well on the proposal preparation and suggests there is overall good quality and thoroughness in the proposals against GCF investment criteria. The proposals generally perform well in articulating eligibility and targeting criteria and providing information on implementation fidelity, both showing a significant proportion classified as low risk. The evaluation of impact potential and paradigm shift potential also shows a more favourable low-risk distribution, indicating that these aspects are generally well-addressed in the proposals.

Figure 6. Stoplight assessment of implementation fidelity and performance against investment criteria

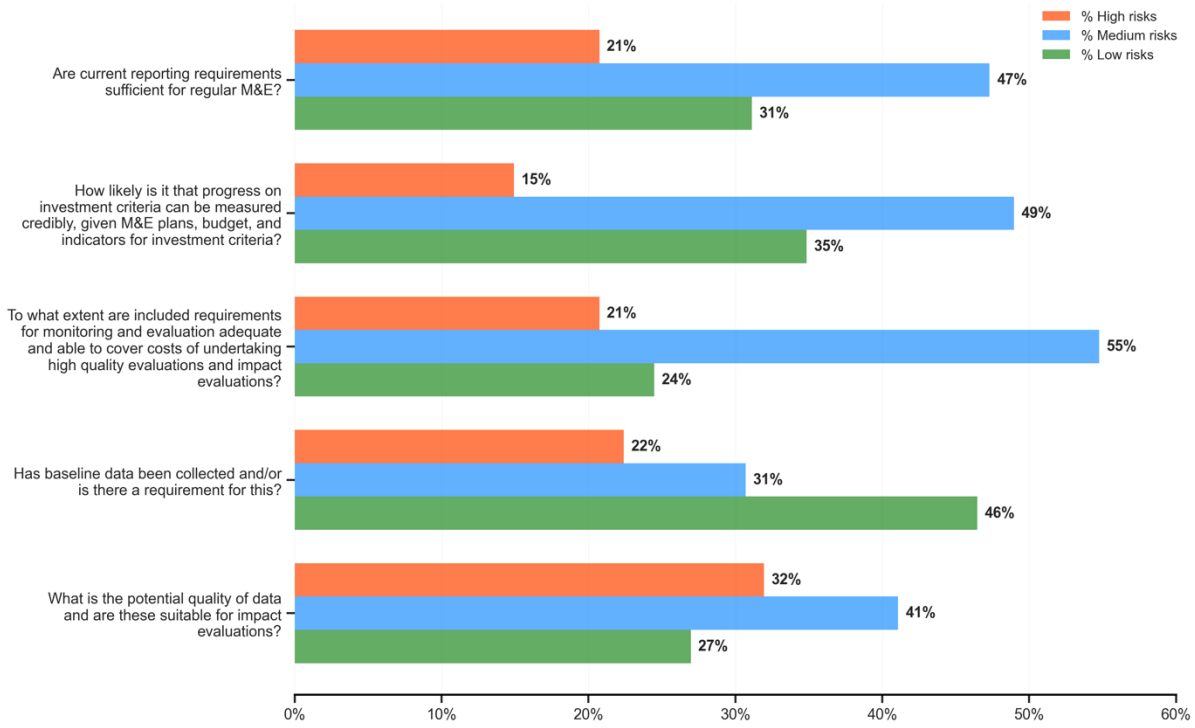


Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis performed by the authors.

39. Lastly, in category D, we assess whether the data collection and reporting processes outlined in the proposals are rigorous enough to help identify the causal effects of the GCF investment (see **Figure 7**).

40. The criterion on data quality for impact evaluations shows significant concerns with a large portion evaluated as high risk, indicating potential issues with data quality and suitability. Availability of baseline data is more favourably assessed with approximately 46 per cent at low risk, suggesting better practices or clarity in baseline data requirements. In terms of collecting or planning to collect additional impact indicators, the majority falls into medium risk, indicating that while there are efforts to extend impact indicators beyond GCF requirements, there is room for improvement. Credibility in reporting on investment criteria shows a substantial amount at medium risk, reflecting ongoing challenges but a reasonable foundation in M&E planning and budget alignment. Sufficiency of reporting requirements was predominantly assessed at medium risk, suggesting that while existing mechanisms are in place, they might not fully meet the M&E needs.

Figure 7. Stoplight assessment of data collection and reporting credibility



Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis performed by the authors.

4.2 Trajectories over time

41. **Figure 8** visualizes the distribution of criteria ratings by risk level (high, low and medium) over time, illustrating how the risk level of approved projects has evolved across Board meetings.[8] The data are derived by adding each project's risk scores (based on the stoplight rubric criteria), calculating the percentage of these scores relative to the total for each Board meeting or set of approved projects, and then dividing these percentages into the different risk categories.

42. At the initial data point (November 2015), more than 40 per cent of the criteria in project proposals were deemed high risk, whereas only about 20 per cent were rated as low risk. Over time, the proportion of criteria rated high risk has trended downward, evidenced by a slight decrease in the share of these ratings.

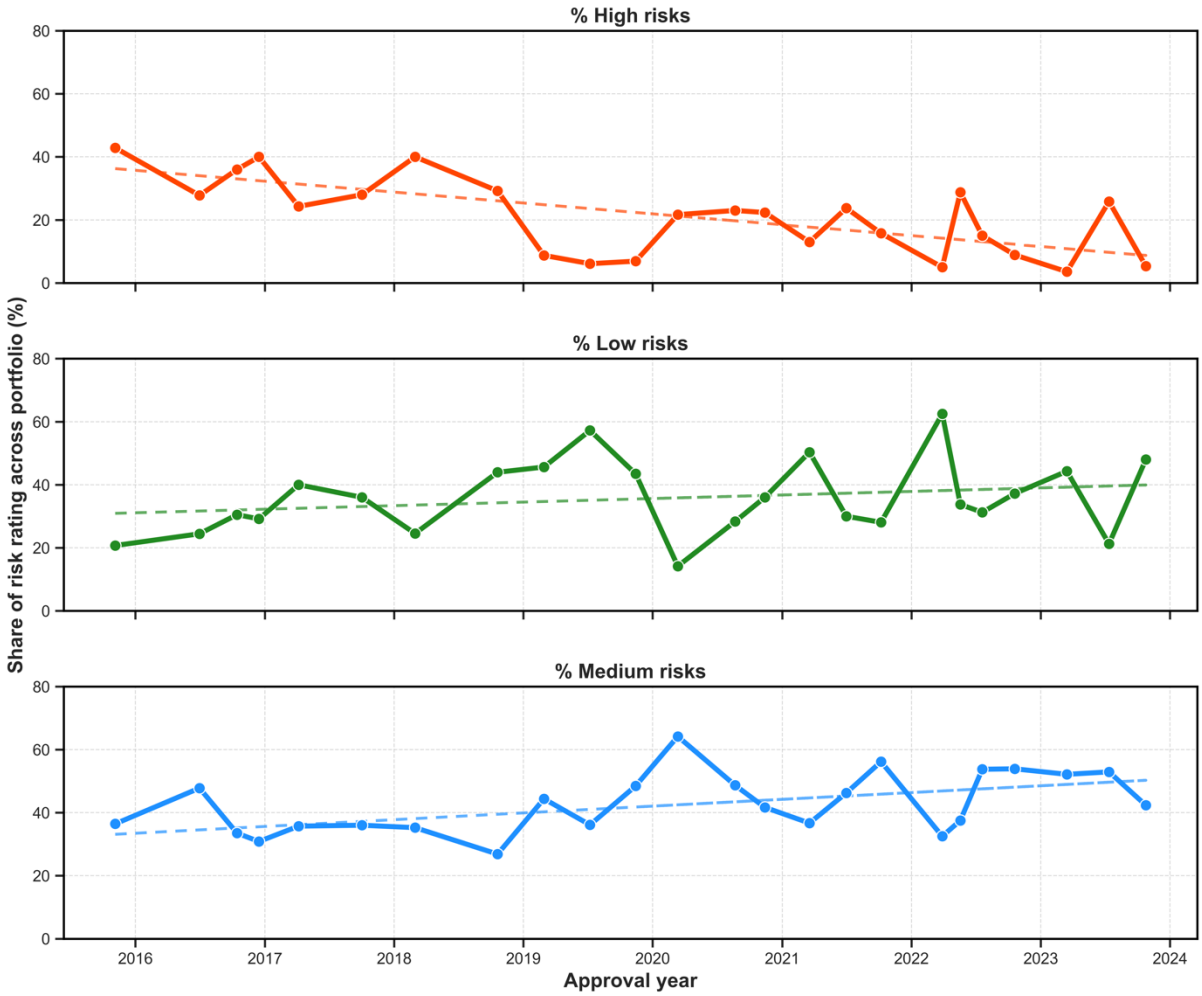
43. The solid lines in the graph represent actual data, while the dotted lines indicate linear regression trends – these trend lines provide a simplified view of how risk ratings have changed overall.[9] A rising dotted line suggests increasing risk exposure, whereas a declining line reflects a decrease in risk exposure. The negative slope for high-risk ratings corroborates a notable decline over the observed period.

[8] Panels are arranged with high risk (top), low risk (middle), and medium risk (bottom) to enable clear comparison between risk categories while maintaining visual clarity.

[9] Note on dotted lines (linear regression trends):

- The dotted lines represent linear regression trend lines for each of the three categories (high, low and medium risk).
- These trend lines provide a simplified view of the overall directions of change, smoothing out short-term fluctuations.
- A rising dotted line indicates an increase in the respective risk exposure, whereas a falling line points to a decrease in that risk level over time.

Figure 8. Trends in portfolio risk level over time [10]



Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis by authors.

4.3 Stoplight assessment (risk index) by IRM vs. GCF-1

44. In the previous section, we showed the risk trend identified by the stoplight analysis over time. The share of the "high" risk rating decreases over time compared to that of the "low" or "medium." In this section, we compare the risk ratings of the four main assessment areas between the IRM period (2015–2019) and the GCF-1 period (2020–2023). To compare the aggregated ratings of the two periods, we constructed a risk index. It is a share of the risk ratings (low, medium, high) of the corresponding portfolio in the reference period. The index takes 1 for a low, 3 for a medium, and 5 for a high-risk rating. It means the lower the risk index, the better when comparing the index between two assessment periods.

[10] Dotted lines show linear regression trends. These trend lines provide a simplified visualization of the overall direction of change over the time period. A rising line indicates an increasing trend in risk exposure, while a declining line indicates a decreasing trend.

Table 3. A Comparison of the stoplight assessment (risk index) between IRM (2015–2019) and GCF-1 (2020–2023)

ASSESSMENT CRITERIA	IRM	GCF-1	CHANGE
A. Theory of change	2.61	2.57	-0.04
B. Measurement of causal change	3.4	3.07	-0.36
C. Implementation fidelity	2.17	2.36	+0.19
D. Data collection and reporting	3.12	2.45	-0.67

Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis by authors.

45. Analysis of the changes between periods reveals:

(i) **Theory of change**: A minimal decrease (-0.04) in risk score, suggesting slight improvement in project logical frameworks.

(ii) **Measurement of causal change**: A notable decrease (-0.36) in risk score, indicating moderate improvement in how projects establish and measure cause-effect relationships.

(iii) **Implementation fidelity**: An increase (+0.19) in risk score, signalling a minor deterioration in how closely projects adhere to their implementation plans.

(iv) **Data collection and reporting**: A significant decrease (-0.67) in risk score, demonstrating substantial improvement in how projects gather, analysed, and report data.

46. When interpreting these changes, shifts greater than 0.5 (as seen in data collection and reporting) represent moderate improvements, while changes approaching or exceeding 1.0 would indicate major shifts in project risk profiles. The overall trend suggests meaningful progress in risk reduction across most assessment criteria during the GCF-1 period.

4.4 Stoplight assessment by energy vs. non-energy projects

47. As part of the evaluability study's stoplight assessment, the team also assessed the quality at entry of the energy-related projects in comparison to non-energy projects. This disaggregated analysis was part of the Independent Evaluation of the GCF's Approach to the Energy Sector and contributed as an input to the evidence collected for portfolio-level monitoring and evaluation of the energy projects.

48. The team used the four lenses of the evaluability study (the theory of change; potential to measure and report causal change; implementation fidelity and performance against investment criteria; and data collection and reporting credibility) to assess the potential for the internal validity of FPs. Based on the stoplight assessment framework, a key finding was that energy projects present weaker requirements for monitoring and evaluation compared to the rest of the portfolio. More details are provided below.

49. Reflecting on the ToC, the study found that 82 per cent of energy-related projects have a well-articulated ToC or that a logic framework/TOC is present but needs some clarification (rated low risk or medium risk), compared to a similar proportion (82 per cent) of the non-

energy related projects. On the other hand, 37 per cent of energy-related projects are rated as high risk, as their unintended consequences are neither identified nor discussed anywhere in the TOC, although they are potentially substantial given the project design.

50. The team also assessed the potential ability of the energy-related projects to credibly measure and report causal change. For 56 per cent of the energy-related projects, the requirements for monitoring and evaluation are not specified or cannot be determined from the information provided in the proposals; this figure was 36 per cent for the non-energy-related projects. Overall, for more than half of the energy-related projects, the GCF will not be able to determine whether causal change can be attributed to its investment in a credible manner.

V. Discussion

5.1 Summary discussion, comparison of IRM and GCF-1 periods

51. Through the stoplight analysis, we reviewed the FPs and summarized risks, and other issues related to results measurement and the information presented in each FP. The distribution of risk ratings across four main assessment areas and respective assessment criteria can be found in appendix 1. In this section, we highlight assessment criteria which showed improvement or require attention between the IRM and GCF-1 periods, focusing specifically on "high risk" ratings.

52. Several criteria showed notable improvement. The percentage of high-risk proposals regarding "quality of theories of change and programme logic" decreased from 20 per cent to 11 per cent (a 9-percentage point improvement). Similarly, proposals with high-risk ratings for "GCF investment criteria being measurable and verifiable with high credibility" decreased from 16 per cent to 6 per cent (a 10-percentage point improvement). The most significant progress was observed in baseline data collection, where high-risk designations dropped dramatically from 38 per cent to just 7 per cent (a 31-percentage point improvement). It is apparent that clear guidance and alterations to the templates and requirements from the Secretariat contributed to clarifying GCF requirements and helping AEs address aspects related to the ToC, investment criteria, and baseline data in the FPs.

53. On the other hand, some areas showed deterioration. The percentage of high-risk proposals regarding "credible reporting of causal change" increased from 13 per cent to 20 per cent (a 7-percentage point worsening). Similarly, proposals with high-risk ratings for "adequate and reliable information regarding implementation fidelity" increased from 3 per cent to 7 per cent (a 4-percentage point worsening). Since credible reporting of causal change is a fundamental requirement for assessing the impact of GCF investments, it is important for AEs to clearly articulate how they will document and report the causal change of projects in their FPs.

Table 4. A Comparison of high-risk assessment criteria: IRM (2015–2019) vs. GCF-1 (2020–2023)

Observations	Assessment criteria	IRM (N=120)	GCF-1(N=121)	Change(pp)[11]
		% High risks	% High risks	
Improved areas	A: What is the quality of the (implicit or explicit) theories of change and programme logic?	20%	11%	**-9 pp ↓**
	C: How well are other GCF investment criteria informed and are these measurable and verifiable with high credibility and quality?	16%	6%	**-10 pp ↓**
	D: Have baseline data been collected and/or is there a requirement for this?	38%	7%	**-31 pp ↓**
Deteriorated areas	B: Does the proposal design allow for credible reporting of causal change?	13%	20%	**+7 pp ↑**
	C: Is there adequate and reliable information included in the proposal regarding implementation fidelity?	3%	7%	**+4 pp ↑**

Source: Assessments of funding proposals as of 31 December 2023. Assessment and analysis by authors.

Note: pp = percentage point.

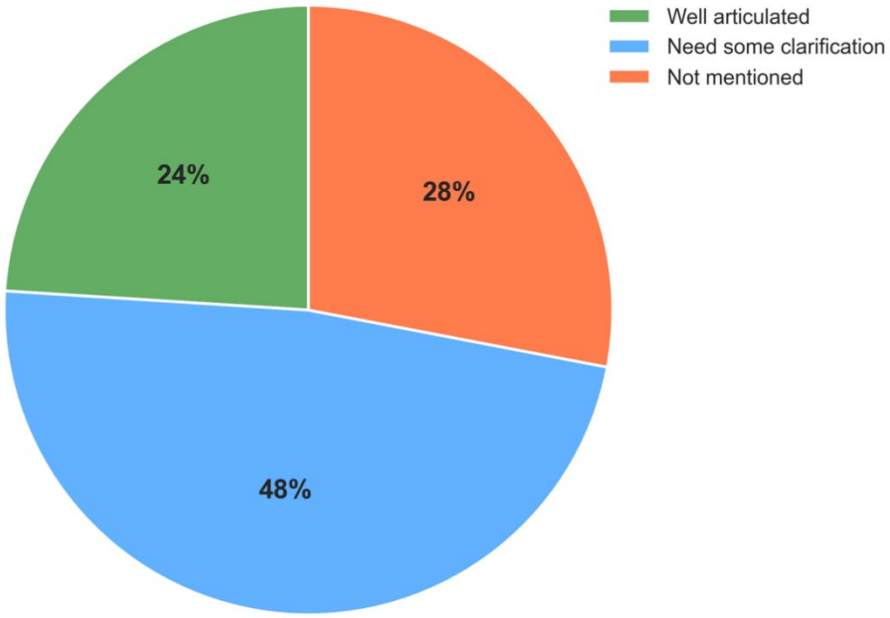
5.2 Discussion of risk for unaccounted-for unintended consequences and their implication for the GCF

54. Another point is around unintended consequences of the funded activities. Addressing unintended consequences in the FP continues to be a challenge at large. The analysis of this chart highlights a significant area for improvement in how unintended consequences are addressed in programme theories of change. Specifically, there is a need to enhance the articulation and inclusion of unintended consequences to ensure that they are adequately considered and mitigated in programme planning and execution. Only 24 per cent of the FPs well-articulated unintended consequences at the project proposal stage while another 28 per cent have no mention of unintended consequences. Due to its nature and the imposed word limit of the new FP template, it might be difficult to fully identify unintended consequences ex-

[11] Negative values indicate improvement (reduction in high risk), while positive values indicate worsening (increase in high risk). The percentage point change shows the magnitude of difference between the GCF-1 and IRM assessment periods.

ante. However, since it can adversely affect the impact of GCF investments, it would be important to include such considerations when developing and reviewing the FPs.

Figure 9. Are unintended consequences referred to and identified robustly in the programme ToC and/or in the surrounding literature reviews?



Source: Assessments of FPs as of 31 December 2023. Assessment and analysis by authors.

5.3 Conclusion and the way forward

55. Evaluability is defined as the extent to which an activity or project can be evaluated in a reliable and credible fashion among international development communities.[12] The study of evaluability is not undertaken to assess an intervention itself nor what has been achieved but is done to provide indication of the likelihood of successfully evaluating interventions in the future. Depending on its use it can be a powerful tool for management, accountability, and learning.

56. This study reviewed 241 FPs approved up to B.37 in 2023. We assessed and compared the quality at entry of these proposals based on a risk assessment described above. Our findings indicate that overall evaluability has improved over time, reflected in a decline in high-risk ratings and a corresponding increase in low and medium risk ratings. We also analysed the risk ratings across the four main assessment areas by comparing the IRM period (2015–2019) and the GCF-1 period (2020–2023). To do so, we constructed a risk index (with ratings of 1–5), where 1 represents low risk and 5 represents high risk, with a lower index indicating better outcomes. Our analysis shows improvements in three out of the four assessment criteria during the GCF-1 period compared to the IRM period. However, for the criterion C *Implementation Fidelity*, we observed a slight increase of 0.19 in the risk index, and further investigation is required into the underlying causes of this trend.

[12] Organisation for Economic Co-operation and Development (2010). Quality Standards for Development Evaluation. DAC Guidelines and Reference Series, p.21. Paris. Available at https://www.oecd.org/en/publications/dac-quality-standards-for-development-evaluation_9789264083905-en.html.

57. We observe that some of the efforts and improvements in streamlining the information provided in the FPs' capacity building efforts, by both the Secretariat and the IEU over time, have likely contributed to improvements in the quality of FPs, their data quality, and reporting.

58. However, ensuring the continued improvement, effectiveness, and ultimate impact of GCF-funded projects calls for several further actions:

(i) **Establish project development based clear guidelines and templates**: Consider developing guidelines and templates for proposal submissions that emphasize the importance of robust theories of change, clear causal pathways, and the identification of potential unintended consequences. Updates of these guidelines should be informed by feedback and lessons learned from previous assessments to ensure they remain relevant, timely and effective. Too frequent changes may put additional burden on implementing entities.

(ii) **Support the strengthening of M&E systems**: Further support the capacity of AEs in designing and implementing high-quality M&E systems. This includes training on best practices for project and impact evaluation, data collection, and reporting. Encourage the use of innovative M&E tools and methodologies that can provide more accurate and timely insights into project performance and impact. The use of established national monitoring systems may become important.

(iii) **Improve data quality and reporting mechanisms**: Consider more standardized and yet flexible approaches to data collection and reporting requirements to ensure consistency and reliability across all projects. This includes establishing clear criteria for baseline, midline and endline data collection, beyond the integrated results management framework (IRMF). A centralized data management system could also facilitate real-time tracking and analysis of project performance and impact.

(iv) **Promote a culture of learning and continuous improvement**: Actively encourage a learning-oriented approach within the GCF ecosystem by regularly sharing evaluation findings, lessons learned, and best practices. Create platforms for knowledge exchange and peer learning among AEs, project implementers, and other stakeholders to facilitate the dissemination of effective strategies and innovations.

(v) **Strengthen stakeholder engagement and collaboration**: Strengthen collaboration with local communities, governments, and other stakeholders to ensure that projects are contextually relevant and have strong local ownership. Encourage multi-stakeholder partnerships to leverage additional resources, expertise, and support for GCF-funded projects.

Appendix. Summary Tables

Table A - 1. Stoplight assessment of the theory of change (by replenishment period)

PERIOD	IRM			GCF-1		
A. Theory of change	**% Low risks**	**% Medium risks**	**% High risks**	**% Low risks**	**% Medium risks**	**% High risks**
What is the quality of the (implicit or explicit) theories of change and programme logic?	38%	42%	20%	38%	51%	11%
Are unintended consequences referred to and identified robustly in the programme ToC and/or in the surrounding literature reviews?	38%	43%	19%	11%	53%	36%
Are causal pathways clearly identified and discussed? (This is discussed in the context of the ToC and the credibility and feasibility of the pathways.)	37%	43%	20%	47%	45%	8%
How robust are the causal linkages (implicit or explicit) and are they well informed by high quality evidence?	33%	42%	24%	40%	47%	13%
Is good quality evidence cited to discuss the efficacy of causal linkages?	50%	37%	13%	51%	37%	12%

Table A - 2. Stoplight assessment of the potential for measurement of causal change and evaluability (by replenishment period)

PERIOD	IRM			GCF-1		
B. Potential for measurement of causal change and evaluability	**% Low risks**	**% Medium risks**	**% High risks**	**% Low risks**	**% Medium risks**	**% High risks**
Does the proposal design allow for credible reporting of causal change?	53%	34%	13%	31%	50%	20%
To what extent are included requirements for M&E adequate and able to cover costs of undertaking high-quality impact evaluations?	24%	50%	26%	25%	60%	16%
Are activities included in the proposal that focus on "economic analyses" and "overall monitoring and evaluation," and are these sufficient for high-quality credible evaluations?	32%	48%	21%	29%	55%	16%
Are methods for measuring attributable causal changes (outcomes or impact or other) discussed?	15%	28%	57%	13%	43%	43%
Are there potential areas of bias that are likely to creep in?	2%	28%	70%	7%	55%	37%

Table A - 3. Stoplight assessment of implementation fidelity and performance against investment criteria (by replenishment period)

PERIOD	IRM			GCF-1		
C. Implementation fidelity and performance against investment criteria	**% Low risks**	**% Medium risks**	**% High risks**	**% Low risks**	**% Medium risks**	**% High risks**
Are eligibility and targeting criteria well-articulated in submitted documents?	59%	26%	15%	48%	44%	8%
Is there adequate and reliable information included in the proposal regarding implementation fidelity?	79%	18%	3%	30%	64%	7%
To what extent is impact potential identifiable and measurable in the proposal?	49%	37%	14%	31%	52%	17%
To what extent is paradigm shift potential identifiable and measurable in the proposal?	47%	31%	22%	41%	50%	9%
How well are other GCF investment criteria informed and are these measurable and verifiable with high credibility and quality?	44%	40%	16%	54%	40%	6%

Table A - 4. Stoplight assessment of data collection and reporting credibility (by replenishment period)

PERIOD	IRM			GCF-1		
D. Data collection and reporting credibility	**% Low risks**	**% Medium risks**	**% High risks**	**% Low risks**	**% Medium risks**	**% High risks**
Are current reporting requirements sufficient for regular M&E?	28%	46%	26%	34%	50%	16%
How feasible is it to measure and report credibly on the progress of investment criteria, given M&E plans, budget and indicators for investment criteria?	25%	59%	16%	46%	40%	14%
To what extent did the proposal provide additional impact indicators beyond those proposed by the GCF? Can the proposal's indicators be used to measure the magnitude of causal change?	19%	21%	60%	38%	43%	19%
Have baseline data been collected and/or is there a requirement for this?	28%	34%	38%	66%	28%	7%
What is the potential quality of data, and are these suitable for impact evaluations?	20%	32%	48%	34%	50%	16%

References

Fiala, N., and others (2019). Becoming bigger, better, smarter: A summary of the evaluability of Green Climate Fund proposals. IEU Working Paper No. 1. Songdo, South Korea: Green Climate Fund Independent Evaluation Unit. Available at https://ieu.greenclimate.fund/sites/default/files/document/working-paper-becoming-bigger-better-smarter-summary-evaluability-gcf-proposals.pdf.

Gertler, Paul J. and others (2016). *Impact Evaluation in Practice*, 2nd ed. (Washington, D.C.: World Bank, 2016). Available at https://openknowledge.worldbank.org/bitstream/handle/10986/25030/9781464807794.pdf?sequence=2&isAllowed=y.

Green Climate Fund (2023). Assessments of funding proposals as of 31 December 2023. Assessment and analysis by authors. Internal data.

_____ (2023). GCF Project API data as of 31 December 2023, including all projects approved through Board Meeting B.37. Available at https://api.gcfund.org/v1/projects. Accessed on 28 February 2024.

Independent Evaluation Unit (2022). Evaluability assessment of the Green Climate Fund funding proposals. IEU learning paper (December). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund. Available at https://ieu.greenclimate.fund/sites/default/files/document/230214-evaluability-study-top_2.pdf.

_____ (2024). Independent Evaluation of the Green Climate Fund's Energy Sector Portfolio and Approach. Evaluation report No. 17 (February). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund. Available at https://ieu.greenclimate.fund/evaluation/ES2023.

Jimenez, E., and Puri, J. (2017). The Wicked Cases of Education and Climate Change: The Promise and Challenge of Theory-Based Impact Evaluations. In: Rob D. van den Berg, Indran Naidoo, and Susan D. Tamondong (eds.), *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability*. Exeter, UK: International Development Evaluation Association (IDEAS). Available at https://sdgs.un.org/sites/default/files/publications/2455IDEASwebREV_08Dec.pdf.

Organisation for Economic Co-operation and Development (2010). Quality Standards for Development Evaluation. DAC Guidelines and Reference Series. Paris. Available at https://www.oecd.org/en/publications/dac-quality-standards-for-development-evaluation_9789264083905-en.html.