
From Ongoing Streams of Evidence to Final Synthesis

Journal of MultiDisciplinary Evaluation
Volume 21, Issue 50, 2025

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
<http://www.jmde.com>

Juha I. Uitto

Environmental Law Institute & International Evaluation Academy

Background: Protecting the global environment requires complex program and project designs intended to produce multiple benefits in the various environmental domains (biodiversity, land degradation, water, climate change, and pollution) while benefiting the people, especially local communities living in the areas of intervention. In this context, learning is a key requirement to which independent evaluation must contribute. This necessitates bringing together streams of evidence from multiple perspectives and sources.

Purpose: The article brings out lessons learned in designing and conducting comprehensive evaluations of large-scale programs and projects funded by the Global Environment Facility (GEF) using streams of evidence. It also describes the utility and utilization of such evaluations for making decisions about future funding.

Setting: GEF is the oldest public financier for the global environment and serves as a financial mechanism to five multilateral environmental agreements in the areas of biodiversity, climate change, land degradation, and pollutants.

Intervention: Comprehensive evaluations of the GEF

Research Design: Not applicable.

Data Collection and Analysis: Not applicable.

Findings: The article draws upon the sixth and seventh comprehensive evaluations of the GEF, which were key inputs to the Fund's replenishment negotiations. The article outlines the transition from stand-alone studies to continuous evaluation streams, which provided real-time insights to improve decision-making in the GEF. These evaluations assessed the GEF's continued relevance and ability to effectively support the recipient countries in protecting the global environment. The evaluations covered project performance, governance, integration in programming, and transformational impact, as well as methodologies and addressing emerging challenges like financial constraints and the pandemic's impact. The findings of these evaluations directly influenced GEF's strategic direction and policy, leading to program refinements. The comprehensive evaluation experience demonstrates the power of bringing together evidence from multiple studies into a stream that can provide both timely information on specific aspects and a comprehensive diagnostic of an organization's performance, strengths, and weaknesses.

Keywords: *Global Environment Facility; comprehensive evaluation; integrated approaches; evidence streams; evaluation influence; environmental finance.*

In 2014, when I started my term as the director of the Independent Evaluation Office (IEO) of the Global Environment Facility (GEF), the fifth Overall Performance Study of the GEF (OPS5) was finished (see contribution by Rob D. van den Berg, this issue of *JMDE*), and my immediate task was to start gearing up toward the next one, OPS6. Although I would only present the approach paper to the GEF Council in its June 2016 meeting, a lot of thinking, analysis, and preparation went into it. First of all, following a recommendation from OPS5, the proposal to change the name of the evaluation to “Comprehensive Evaluation” was accepted because these evaluations had come to deal with much more than just the performance of GEF programs and projects (however, the well-known acronym “OPS” stuck; consequently, the sixth comprehensive evaluation of the GEF would still be known as OPS6). The evaluation would focus on two related themes: (1) institutional, governance, strategy, and programming issues; and (2) the performance and impact of the GEF.

The overarching theme for the evaluation was chosen as “the GEF in the changing landscape of environmental finance,” reflecting on the growing importance of other funds, especially in the climate finance area (notably the Green Climate Fund and the Climate Investment Funds). IEO stated its intention to analyze the continued relevance of the GEF to the program countries, to the global environmental conventions the GEF serves as funding mechanism, and to the environmental issues it seeks to address.

The concept and conduct of OPSs in the GEF is quite unique within the international development architecture. The GEF’s governing body, the GEF Council, pays close attention to the evaluative findings as it considers the programmatic directions and policy agenda of the GEF partnership, especially at times of replenishment of the fund. While it was important to bring forward the strategic evaluative synthesis to the replenishment process, it was equally important to feed evaluative evidence on critical aspects to the system on an ongoing basis. Irrespective of the final synthesis, the ongoing evaluation streams would provide information on GEF’s performance—both programmatic and institutional—and allow for improvement and course correction along the way. My decision at the time was that there would be no studies conducted specifically for OPS6; instead, we would have a stream of evaluations that reflected

the ongoing decision-making needs of the GEF Council, the Secretariat, and the broader partnership.

Streams of Evaluations for Ongoing Improvement and Synthesis

OPS6 would again build on multiple streams of evidence that would in the end of the replenishment cycle be brought together for a comprehensive picture of the GEF and its relevance and ability to support countries in protecting the global environment. At the same time, each of the evaluations and studies would be fed to the GEF partnership on an ongoing basis. They would be discussed at the council meetings and would provide a cumulative evidence base on the issues facing the partnership and potential solutions.

The project-level terminal evaluations conducted by the implementing agencies continued as important building blocks for evidence of performance at the field level. By now, the GEF portfolio had risen to 4,047 projects covering \$15.47 billion in GEF grants. Of these, 581 projects had been completed since OPS5 and had terminal evaluations that IEO could utilize.¹

The stream on program performance and impact would cover project success ratings, progress toward impact, the catalytic role of the GEF, country ownership and the GEF’s role in contributing to policy and regulatory reforms in countries, the longer-term impact of the GEF, and other issues related to project quality and implementation.

These building blocks were supplemented by evaluations of higher order at thematic, country, or regional levels conducted by the IEO. One innovation was to introduce what was called strategic country cluster evaluations (SCCEs), which we conducted of important categories of countries, including least developed countries (LDCs), small island developing states (SIDS), and GEF programming in the African belt of the Sahel and Sudan-Guinea savanna biomes. In addition, we conducted a number of thematic and institutional evaluations.

We would also look at many institutional issues, such as the resource allocation and results-based management systems, crosscutting policies pertaining to gender and stakeholder participation, the involvement of civil society organizations, Indigenous peoples and the private sector, and the

¹ Although the project terminal evaluations were either commissioned by the agencies or conducted by their independent evaluation offices and each had their own

evaluation system, the minimum standards and validation process established by the GEF IEO guaranteed their reliability.

health of the expanded partnership, which now consisted of 18 implementing agencies.

Toward Integration in Programming

A key development that had started in the GEF programming was the move toward integrated approaches. The GEF's sixth replenishment period contained three integrated approach pilots, which were somewhat contested, as they presented a departure from the traditional focal area programming. Countries and conventions wanted to ascertain that the GEF continued to serve them to solve the global environmental problems the fund had been established to address. GEF's scientific and technical advisory panel (STAP) conducted a study on the scientific underpinnings of integration, which it presented to the council (Bierbaum et al., 2018).

IEO's work reflected this trend toward integration and the need for evidence that it would work. We designed and conducted evaluations that looked back to assess the extent to which the GEF's programs and projects had generated multiple benefits (defined as benefits accruing to two or more focal areas, including any local environmental and social benefits).² We also conducted an analysis of the value for money of GEF land degradation projects in terms of quantifying the benefits to other focal areas, such as increased forest cover, reduced deforestation, and enhanced carbon sequestration.³ These evaluations generated much discussion in the council when it considered how far and fast the integration agenda should spread.

The GEF Council wanted timely updates, which the ongoing changes to the GEF strategy made even more important. There was pressure on IEO to conduct more ongoing and formative evaluations. We thus embarked on a thorough formative review of the three integrated approach pilots, which focused on the coherence of their design, their alignment with convention guidance, and their capacity to deliver synergies between focal area strategies while accounting for country needs and ownership.⁴

All evaluations were presented to the council as they were completed throughout the replenishment period. Each of them contained specific recommendations and management responses thereto, which the council could consider and

adopt. As these recommendations pertained to the specific evaluations focused on their own area, not all of them would make it to OPS6, which distilled the higher-level, strategic recommendations from multiple evaluations. These ongoing component evaluation recommendations were very important; they were recorded and their implementation tracked on an annual basis.

Influential Evaluations

As in the past, some of these individual evaluations were quite influential. One such was the evaluation of the GEF's contributions to transformational change.⁵ The evaluation was based on a purposeful sample of projects that were initially identified and suggested by the agencies as having led to transformational change. The IEO screened the proposed projects against specific criteria and terminal evaluation reports to focus on a small group of completed projects where there was evaluative evidence that they had indeed been transformative. The IEO then analyzed these projects in depth to come up with a list of common factors, *sine qua non*, for transformation to happen. These included factors, such as ambition of design, that targeted transformation from the get-go, establishing a clear mechanism for transformation, including mainstreaming the project objectives in national policy and harnessing market forces. The beauty of the study was that its findings could equally be turned into criteria for designing and screening new project concepts.

In order to meet the council's demands for timeliness, the first GEF-7 replenishment meeting started with a half-day session in which I was able to present the findings of all completed component evaluations to the replenishment participants. Most of these had already been presented to the council in the preceding couple of years, although some new evidence was brought directly to the replenishment meeting. At this stage, we had not yet attempted to conduct a meta-analysis to synthesize the findings at a higher level—this would come about half a year later—but presenting all of the evaluative evidence in one package provided a solid basis for the discussions on the GEF-7 programming directions.

² <https://www.gefio.org/evaluations/evaluation-multiple-benefits-gef-support-through-its-multifocal-area-portfolio-map-2016>

³ <https://www.gefio.org/evaluations/value-money-analysis-land-degradation-projects-gef-2016>

⁴ <https://www.gefio.org/evaluations/formative-review-integrated-approach-pilot-iap-programs-2017>

⁵ <https://www.gefio.org/evaluations/evaluation-gef-support-transformational-change-2017>; see also Uitto et al., 2019.

The full OPS6 was based on a synthesis of 29 separate evaluations and studies.⁶ All of its recommendations were incorporated into the next GEF phase's programming directions and policy agenda.

Learning from Experience

The process for the seventh comprehensive evaluation of the GEF (OPS7) followed a similar pattern. This time around, we were even more deliberate in designing the IEO's work program so that each of the evaluations would feed into the comprehensive evaluation in a systematic manner, while providing critical information in a timely manner. It was like adding pieces into a larger jigsaw puzzle so that the overall picture became clearer one part at a time. We designed the evaluation reflecting upon what had worked in OPS6 and what could be improved.

The negotiations for the GEF's eighth replenishment would take place in a difficult financial situation in which the global economy was growing slowly and the recent pandemic had clearly demonstrated societal vulnerability—and how closely related human and environmental security were. We chose an overarching theme for the evaluation: “working for a greener global recovery.”

We would again focus on GEF's performance, longer-term impacts, and sustainability, as well as its additionality, based on an evaluative approach to assessing the latter that we had recently developed. This was a theory-based approach to track how GEF's programs and projects would add value in several domains, including environmental, legal and regulatory, institutional and governance, financial, socioeconomic, and innovation.

A notable development in the GEF programming was the further expansion of integrated approaches. While the formative evaluation of the pilots had given inputs into the design and operation of the programs, we had not yet been able to evaluate their results. In the meantime, the GEF expanded the approach to five impact programs.⁷ This time, IEO would have to conduct an evaluation that looked at what had happened to the three earlier pilots, as well as the five new programs. Given the time lag between program implementation and impact, this would still be a formative evaluation.⁸

Evaluations of innovation and scaling up, and support to countries, as well as private sector engagement, were seen by the council as critical to leverage GEF's catalytic role. There was a sense of urgency to provide evidence on what the experiences told us about these.

In the traditional focal areas, instead of comprehensive studies, we decided to focus on selected important topics from each. The portfolio had grown to 4,786 projects. Since OPS6, 608 new project-level terminal evaluations had been submitted to IEO. In the spirit of timeliness, we included in the analysis the 504 new projects that had been approved during the 3 first years of the GEF-7 period. The travel restrictions due to the pandemic in 2020 and 2021 had affected our work. Consequently, many of the evaluations relied on innovative online data gathering efforts and use of geospatial analysis (Anand et al., 2025). We also mined data from earlier field missions, repurposing what was possible for the task at hand.

We again presented each completed evaluation to the council as they became ready, and the full set of completed evaluations to the first GEF-8 replenishment meeting, this time attempting to highlight emerging preliminary findings. The presentation led to a two-hour discussion in which 18 donor nations, all recipient country groupings, and CSOs took turns asking questions.

The final OPS-7 report,⁹ based on the findings of 34 evaluations and studies, contributed significantly to the design and approval of the GEF-8 programming directions and policy agenda, which led to the record replenishment of \$5.33 billion. Apart from the importance of the overall synthesis, the ongoing streams of evaluative evidence throughout the replenishment period had contributed significantly to adjustments in policies and practices along the way.

OPS7 presented a suite of high-level recommendations, all of which were adopted. A notable recommendation was that the GEF should continue pursuing integration in programming but should clearly demonstrate the additionality of this approach in terms of environmental benefits, socioeconomic co-benefits, policy influence, and inclusion. The evaluation further recommended that the GEF should develop and implement a more strategic and coherent approach to country-level engagement; it should increase its support to priority country groupings, LDCs, and SIDS; and it

⁶ <https://www.gefio.org/evaluations/gef-changing-environmental-finance-landscape-ops6>

⁷ In the following phase, GEF-8, there would be 11 integrated programs.

⁸⁸ <https://www.gefio.org/evaluations/formative-review-gef-integrated-approach-tackle-drivers-environmental-degradation-2020>

⁹ <https://www.gefio.org/evaluations/seventh-comprehensive-evaluation-gef-ops7>

should strengthen private sector engagement with targeted support. The recommendation that the GEF should pursue innovative projects to advance transformational change led to the establishment of a specific innovation window in the next phase.

Conclusion

These observations indicate why the move from studies to streams was possible and welcomed in the case of the GEF. In a thorough postmortem of OPS6, it was concluded that there were a number of factors that made the evaluation influential (Batra et al., 2022). These included the ongoing and timely delivery of quality evaluations, recognizing the need for balancing updated information with early delivery. Time was of the essence, and it was not acceptable to wait for a synthesis at the end of a 4-year replenishment period.

In order to provide useful evaluative evidence for decision-making, evaluators must understand the internal priorities and the political economy of the organization. Credibility of evaluation depends on a number of factors, including the independence of the evaluation, the rigor of the methodologies utilized, transparency, and stakeholder engagement (pp. 131–135).

The GEF experience with the OPSs, as described here, is testimony to the power of bringing together evidence from multiple studies into a stream that can provide both timely information on specific aspects and a comprehensive diagnostic of an organization's performance, strengths, and weaknesses.

References

- Anand, A., Batra, G., & Uitto, J.I. (2025). Harnessing geospatial approaches to strengthen evaluative evidence. In S.B. Nielsen, F.M. Rinaldi, & G.J. Petersson (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation* (pp. 196-218). Routledge.
- Batra, G., Uitto, J. I., & Feinstein, O. (2022). *Environmental evaluation and global development institutions: A case study of the Global Environment Facility*. Routledge.
- Bierbaum, R., Cowie, A., et.al. (2018). *Integration: To solve complex environmental problems*. Scientific and Technical Advisory Panel to the Global Environment Facility.
- Uitto, J. I., Puri, J., Williams, A., Dickman, J., et.al. (2019). Evaluating transformational change—lessons from international environmental funds. In R. D. Van den Berg, C. Magro, & S. Salinas Mulder (Eds.), *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals* (pp. 105–130). International Development Evaluation Association (IDEAS).