



Investing in rural people

## Evaluation Committee

133<sup>rd</sup> Session

Rome, 18 June 2026

---

# Report of the third External Peer Review of IFAD Evaluation Function

---

Document: EC 2026/133/W.P.5

Agenda: 6

Date: 9 June 2026

Distribution: Public

Original: English

### **FOR: REVIEW**

**Useful references:** Revised IFAD Evaluation Policy ([EB 2021/132/R.5/Rev.1](#)). Multi-year Evaluation Strategy ([EC 2021/114/W.P.8](#)). Evaluation Manual ([EB 2022/135/R.29](#))

**Action:** The Evaluation Committee is invited to review the report of the third External Peer Review of IFAD Evaluation Function.

---

---

### **Technical questions:**

#### **Veronique Salze-Lozac'h**

Former Chief Evaluator

Independent Evaluation Department

European Bank for Reconstruction and

Development

Email [vsalzelozach@gmail.com](mailto:vsalzelozach@gmail.com)

---

# **2026 External peer review of IFAD's evaluation function**

**Final report**

**May 2026**

## Table of contents

<b>Acronyms</b> .....	ii
<b>Foreword</b> .....	iii
<b>2026 EPR key findings and messages</b> .....	iv
<b>1 Introduction and background</b> .....	1
<b>1.1 Introduction</b> .....	1
<b>1.2 Background – implementation of the 2019 EPR recommendations</b> 2	
<b>2 Main findings</b> .....	4
<b>2.1 Overall assessment: A strong function that will need to become more strategic to accompany IFAD in the future</b> .....	4
<b>2.2 Credibility: strong and trusted, with the need for a more strategic approach</b> .....	5
<b>2.3 Independence and governance: A well-established and respected function with a need to ensure that independence is maintained</b> .....	10
<b>2.4 Utility: Recognized value but limited uptake</b> .....	11
<b>2.5 Complementarity of independent and self-evaluation: Improvements with remaining challenges</b> .....	14
<b>2.6 Management and staffing: Need for a broader skills-mix</b> .....	15
<b>2.7 Enabling environment for learning and development: A great potential for further improvement</b> .....	16
<b>3 Recommendations</b> .....	21
<b>4 Appendices</b> .....	23
Appendix 1: EPR Evaluation Framework (LINK) .....	24
Appendix 2: EPR approach and methodology .....	24
Appendix 3: List of interviewees.....	26
Appendix 4: Surveys of staff and members of the Executive Board .....	29
Appendix 5: Quality review of IOE evaluation reports .....	35

## Acronyms

AI	artificial intelligence
ARIE	Annual Report on the Independent Evaluation of IFAD
CCR	country strategy completion report
CLE	corporate-level evaluation
COSOP	country strategic opportunities programme
DAC	Development Assistance Committee (OECD-DAC)
EAP	Evaluation Advisory Panel (IOE)
ECG	Evaluation Cooperation Group (of multilateral development banks)
EPR	external peer review
EvalNet	Evaluation Network (of OECD/DAC)
IOE	Independent Office of Evaluation of IFAD
M&E	monitoring and evaluation
ODE	Office of Development Effectiveness (IFAD)
OECD	Organisation for Economic Co-operation and Development
PCR	project completion report
PRISMA	President's Report on the Implementation Status of Evaluation Recommendations and Management Actions
RIDE	Report on IFAD's Development Effectiveness
UNEG	United Nations Evaluation Group

## Foreword

The external peer review (EPR) presented in this report was initiated by the IFAD Executive Board consistent with the Fund's evaluation policy. It was carried out in line with international standards by an external panel of heads of evaluation.

The review assesses the quality of the evaluation function – both Management-led self-evaluation and independent evaluation managed by the Independent Office of Evaluation. It finds the evaluation function presently to be credible, and well-regarded, with the 2019 EPR recommendations implemented.

However, to sustain and enhance its impact in the future, the review finds that a more strategic, selective and learning-oriented approach is needed. This includes a stronger emphasis on strategic topics, timelier outputs, more robust follow-up, better alignment between independent and self-evaluation, and increased focus on the opportunities and challenges arising from greater use of artificial intelligence (AI).

This report presents the main findings of the EPR related to the six core criteria: credibility, independence, utility, complementarity of self- and independent evaluation, management and staffing and an enabling environment for learning and development, followed by recommendations for further improvement of the evaluation function.

The panel members are thankful for broad engagement in the review from all parts of IFAD and the excellent support provided during the process. A peer review is always a mutual learning experience, where panel members also gain insights and learn from the good practices displayed in the reviewed organization. We appreciate being asked to undertake this review.

Saint-Lunaire/Washington DC/Bonn  
May 2026

Véronique Salze-Lozac'h (chair)  
Geeta Batra  
Jörg Faust

-----  
Per Øyvind Bastøe (consultant)

## 2026 EPR key findings and messages

*"Solid evaluation function, with self-evaluation in the process of being strengthened. Not broken but needs to evolve to be fit for purpose for the future"*

### Credibility

#### Finding

IFAD's evaluation function – particularly the Independent Office of Evaluation (IOE) – is widely regarded as credible, professional and authoritative, with strong trust from Management, the Executive Board and external stakeholders. Country-level evaluations (CSPEs) are seen as especially credible and influential.

#### Key message

Credibility is a core strength of IFAD's evaluation function and should be preserved, even as changes are made to focus, selectivity and delivery.

Increased use of artificial intelligence (AI) has the potential to increase efficiency in evaluation processes and accessibility of evaluation evidence.

### Governance and Independence

#### Finding

The general enabling environment of the evaluation function in IFAD is satisfactory. The existing evaluation policy of IFAD is aligned with international standards. Some details in the annexes of the Policy should be clarified (like roles and responsibilities of ODE, and human resources processes and practices for the selection of the IOE Director) but a revision of the policy itself is not required.

Independence of IOE is well-established, well-understood and not contested. Stakeholders consistently view independence as essential to IOE's value, enabling it to raise sensitive or uncomfortable issues. Recent improvements in engagement have strengthened the credibility of the reports without undermining this independence.

#### Key message

The Independent Office of Evaluation operates from a position of strong independence. The challenge is to ensure that the current evolution of self-evaluation ensures quality and timeliness of self-evaluations as well as clear complementarity with independent evaluation.

Evolving practices of interaction and complementarity between the two evaluation functions should contribute meaningfully to the relevance, learning value and impact of evaluation findings.

### Utility (relevance and use)

#### Finding

Evaluations from IOE are mostly perceived as useful at Board and senior management level, particularly for accountability and strategic reflection at the country level.

Self-evaluations are more uneven in their quality and usefulness, and the transformation of this function should focus on quality and utility of self-evaluation processes and reports.

However, use by operational staff is uneven, and learning uptake is sometimes constrained by the length, timing and volume of evaluation outputs and the absorption capacity of Management and Board members.

**Key message**

Utility and use remain uneven across the organization, strong at Board and Management level, it is more limited at operational level, limiting the usefulness and influence of evaluation for the institution.

Improving the strategic focus, timeliness, prioritization and accessibility is critical to strengthening uptake and decision-making. The use of AI can help address this issue, common to most IFIs, by making it possible to have rapid access to a wealth of evaluation insights by various stakeholders, on demand.

**Complementarity of independent and self-evaluation****Finding**

The recalibration and creation of the Office of Development Effectiveness (ODE) have reduced tension and improved collaboration between self-evaluation and IOE. Collaboration has also improved drastically thanks to a constructive approach to collaboration, and processes that favor consultation and exchanges, without undermining the independence of IOE.

Nonetheless, methodological inconsistencies between self and independent evaluation, uneven self-evaluation quality and data quality across regions, and weak articulation between functions persist, limiting coherence and learning.

**Key message**

Complementarity has improved but remains partial; clearer roles, stronger methodological alignment and more systematic interaction are needed.

**Product mix and strategic focus****Finding**

The evaluation product mix has expanded and become increasingly dominated by country evaluations, largely driven by Board expectations and Management demand. While CSPEs are valued, stakeholders express concern that the current mix overstretch resources and leaves insufficient space for strategic thematic and corporate learning.

**Key message**

The core issue is not quality but selectivity: a more focused product mix, with fewer and more strategically chosen evaluations, on cross-cutting topics that are of priority to the institution would significantly increase relevance and impact. IOE could also use its evaluative findings to deliver cross-countries synthesis and just on time notes before key high level discussions.

**IOE recommendations and follow-up****Finding**

IOE evaluation recommendations are generally considered useful and relevant, but are often fragmented into sub-recommendations, or insufficiently prioritized, or too prescriptive, making implementation and follow-up on the implementation difficult given IFAD's resource constraints and absorptive capacity.

**Key message**

Fewer, sharper, more strategic and more prioritized recommendations would strengthen implementation and learning.

## **Learning environment and absorptive capacity**

### **Finding**

IFAD has a strong evaluation culture, and learning is valued in principle. In practice, however, learning absorptive capacity is limited due to IFAD's staff time and evaluative capacity. Delayed evaluation products and access to evaluation insights that require time and efforts reduce opportunities for timely learning and course correction.

### **Key message**

Strengthening learning depends less on producing more evaluations than on delivering strategic, timely, concise and targeted outputs that align with decision cycles and a more tailored evaluation product mix, enabling lessons to be taken up and acted upon. Easy access to reports is also key to creating a demand for evaluation insights. New AI-supported tools can be a gamechanger in providing easy and meaningful access to evaluation findings and lessons, on the supply side. However, the demand side also needs to be strengthened to ensure learning is valued and incentivized by the institution.

### **Overall EPR message**

- IFAD's evaluation function is credible, independent and highly valued.
- For IOE, the main challenge is strategic focus and use, not quality or independence.
- Greater selectivity, improved complementarity with self-evaluation, and more actionable recommendations are critical to enhance impact.
- For self-evaluation, the main challenge is the consistency of the quality as well as the relevance and usefulness for learning purposes.
- Addressing evaluation capacity (at the country and institutional level) and organizational absorptive capacity of evaluation lessons are central to ensuring evaluation continues to support IFAD's development effectiveness and institutional efficiency.
- An AI strategy would help identify opportunities for increased efficiency and effectiveness and provide a framework for a sustainable and ethical use of AI in evaluation. This should be done, like for capacity development, in close collaboration and mutual reliance with other IFIs and partners.

# 1 Introduction and background

## 1.1 Introduction

1. The external peer review (EPR) is initiated in accordance with the requirements of IFAD's Evaluation Policy and aims to provide an external assessment of IFAD's evaluation function, including both Management-led self-evaluation and independent evaluation by the Independent Office of Evaluation (IOE).
2. The main purpose of the EPR is to assess IFAD's evaluation function in order to enhance its contribution to the Fund's development effectiveness and institutional efficiency.
3. The review comes at a critical juncture, shaped by the current operating environment IFAD is facing. A decade of institutional reforms and evolving expectations for development effectiveness is transforming the organization to be better equipped to meet evolving needs for food security and agricultural production in the poorest regions of countries in a global context of environmental, financial and political uncertainties. IFAD has established a new financial architecture with increased focus on private sector operations and access to alternative funding sources with the purpose of increasing operational effectiveness and field impact. An important contextual factor is the upcoming IFAD replenishment (IFAD14 for 2028-2031). The trend of shrinking official development assistance is also a source of rising concern in IFAD.
4. Of particular relevance for the evaluation function is the recalibration of the Office of Development Effectiveness (ODE) as the institutional hub for evidence and effectiveness, with functions related to quality assurance, results, resources, knowledge management, innovation and impact assessment. In 2026 a corporate strategy function was added to ODE to strengthen and better connect strategic efforts across the institution, with the purpose of supporting and operationalizing the strategic vision set by the President. The intention is to ensure strategic coherence across institutional priorities, financial and operational directions, evidence, institutional positioning and external engagement priorities.
5. While the review in its nature is retrospective as it assesses the performance of the evaluation function since the 2019 EPR and adherence with standards agreed by the multilateral development banks' Evaluation Cooperation Group (ECG), the United Nations Evaluation Group (UNEG) and the Evaluation Network (EvalNet) of the Organisation for Economic Development Co-operation's Development Assistance Committee (OECD/DAC), it is intentionally formative and forward-looking, also anticipating what is required to strengthen the evaluation function in the near future.
6. The panel has taken an impartial view, based on knowledge of best practice, and with respect to ethical norms and confidential information. The review approach and methodology are consistent with those agreed by ECG, UNEG and EvalNet. As outlined in EPR's evaluation framework (appendix 1) responding to the EPR approach paper, the assessment builds on multiple sources of information including systematic document reviews, interviews, focus group sessions and a survey of all staff and members of the Executive Board (see appendix 2 for details about the methodology).
7. In line with best practice of external reviews of evaluation functions, management (ODE) and IOE have done self-assessments of self-evaluation and independent evaluation respectively. These self-assessment reports provide valuable information to the EPR. While the self-assessments are not validated by the EPR panel, they provide valuable information for readers of this report. The self-assessment reports are available on request by contacting the concerned departments.

8. Unlike the two previous EPRs (2009 and 2019), this EPR covers both Management-led self-evaluation and independent evaluation by IOE and assesses how the two parts of the evaluation function are aligned and complementary.

## 1.2 Background – implementation of the 2019 EPR recommendations

9. The recommendations of the 2019 EPR form the background for this EPR. The eight recommendations were made in the context of IFAD at that time. The findings of this EPR show that all recommendations have been implemented.
10. According to IOE's and Management's self-assessments, the implementation has been extensive including evaluation policy revisions, strengthened methodological guidance, improved supervision and completion processes, development effectiveness framework roll-out, and the creation of ODE. According to Management's self-assessment, the next phase requires embedding new standards into everyday practice through incentives, capacity-building, and clarified workflows.
11. Here are some of the details of the implementation:
12. **Revision of the Evaluation Policy.** The revised IFAD Evaluation Policy was published in 2021. For the first time, it covered both independent and self-evaluation, broadening the scope and inclusivity of evaluation practices within IFAD.
13. **Simplification of IOE internal processes and procedures.** Recent reviews have shown improvements in CSPE efficiency. The timeline for CSPEs has been reduced by about 10 per cent, with evaluations now often completed ahead of schedule, reflecting better planning and execution. IOE has met the increased coverage of CSPEs with a less than 30 per cent increase in budget and staff resources. The average planned and actual budgets for CSPEs have decreased (by around 9 per cent and 22 per cent respectively), and the use of virtual workshops has further reduced costs.
14. **Revision of IOE's product mix.** IOE has diversified its outputs, including thematic and corporate evaluations, review series, learning notes, research papers, books and peer-reviewed articles. New products such as infographics, video series, podcasts, blogs and an independent magazine have been introduced to enhance outreach and engagement.
15. **Preparation of a multi-year strategy for IOE to implement the evaluation policy.** The multi-year evaluation strategy (2022-2027) outlines four main objectives, including fostering a transparent, learning-oriented and accountable corporate culture, and engaging with Management, Member States and external partners. According to IOE,<sup>1</sup> the strategy provides the framework to increased coverage of country evaluations over the past two years with a 333 per cent rise in country coverage and a 1,578 per cent increase in the dollar value of the assessed portfolio, all managed with less than a 30 per cent increase in budget and staff.
16. **Preparation of a new Evaluation Manual.** In 2022, the 3<sup>rd</sup> edition of the IFAD Evaluation Manual was co-signed with Management – a first for any United Nations agency or international financial institution. This manual covers both independent and self-evaluation processes, reflecting a more integrated approach to evaluation within IFAD.
17. **Simplification and harmonization of IOE's budget with IFAD's overall administrative budget.** Despite a significant expansion in evaluation coverage and activities the last two years, IOE has kept budget increases below 30 per cent. The average cost per CSPE has decreased, and overall expenditures have been reduced through process streamlining and the adoption of digital tools.

---

<sup>1</sup> IOE's self assessment.

**Key points**

The 2026 EPR is done in accordance with the IFAD Evaluation Policy and in line with international standards. The review builds on multiple data sources and extensive consultation. It is intentionally high-level and forward-looking.

The 2019 EPR triggered important reforms (policy revision, rating alignment, creation of ODE). The current review finds that the direction is right, but implementation depth, prioritization and agility as well as system-wide coherence now matter more than further expansion.

## 2 Main findings

### 2.1 Overall assessment: A strong function that will need to become more strategic to accompany IFAD in the future

18. The EPR finds that the evaluation function (including both independent and self-evaluation) in the current context is performing well without any major flaws. As described above, the recommendations from the 2019 EPR have been implemented. The function in its present form is widely appreciated and respected in IFAD.
19. The evaluation policy and evaluation manual are consistent with common practice in other international financial institutions (IFIs) and United Nations organizations. The EPR panel finds the policy sufficiently clear and flexible as a framework for further improvements of the evaluation function. Some elements in the annexes should be clarified (like roles and responsibilities of ODE, and clarity of human resources processes and practices) but a revision of the policy itself is not required.
20. Independent evaluation managed by IOE has matured <sup>2</sup> over the last few years with the efficient delivery of a mix of evaluation products and broad dissemination and communication of findings and recommendations. The core criteria for a reliable independent evaluation function related to credibility, independence and utility are largely met.
21. Self-evaluation is, despite its long history in IFAD, at an earlier stage of maturity compared to similar functions in other IFIs. However, it has credibly identified its weaknesses and has established a well-crafted structure and convincing plans for application (ref. the IOE self-assessment). The recalibration of ODE and the appointment of a managing director of development effectiveness and corporate strategy create positive expectations across the institution.
22. Despite this overall positive assessment, it is uncertain how ready the evaluation function is to meet the anticipated future. As IFAD is purposely reforming to effectively meet an increasingly challenging global context, it is questionable if the evaluation function is sufficiently aligned with these reforms to be able to play its intended role moving forward. ODE has a way to go before the self-assessment structure is fully implemented and the quality of self-evaluation is acceptable throughout the organization. IOE's product mix and capacity is currently not fully adjusted to meet the future as outlined in IFAD's reform agenda.
23. In a fast-changing environment, the evaluation function (both independent and self-evaluation) cannot fulfil its role by primarily providing evidence about the performance of completed operations. More attention must be given to planned and ongoing operations and less to compliance and past operations. Focus needs to shift from ex-post to ex-ante. Quality at entry matters for achieving results in the future. The evaluation function (in particular IOE) has the potential to provide even more strategically relevant evidence<sup>3</sup> for senior management and the Executive Board. This requires selectivity, focus, timeliness and efficiency. IOE's evaluation plans must be closely aligned with replenishment processes and other key corporate strategic decision points.
24. Increased use of AI is another key ingredient in a proactive evaluation function. AI offers opportunities to improve evaluation coverage, timeliness and use of evidence. But it also brings challenges around ethics, bias, accountability, skills and governance that need deliberate management. Some of the opportunities are to get faster evidence synthesis and meta-evaluation, access broader data sources and evaluative knowledge and gain efficiency in evaluation processes. At the same time, acknowledging the challenges related to ethics, human rights, and fairness, bias and

---

<sup>2</sup> Compared to the situation described by the 2019 EPR.

<sup>3</sup> Being strategic implies being proactive by providing timely evidence on current topics and giving attention to issues that matter for the future

representativeness, accountability and transparency is critical. It is worth noting the OECD AI principles<sup>4</sup> that call for traceability of datasets, processes and decisions. This implies documenting how AI tools influenced findings and ensuring that evaluative judgments remain clearly attributable to human evaluators.

25. As presented in the EPR evaluation framework (appendix 1), the review focuses on six core criteria: credibility, independence, utility, complementarity of self- and independent evaluation, management and staffing, and creating an enabling environment for learning and development. The main findings related to these criteria are presented below.<sup>5</sup>

### Key points

IFAD's evaluation function is credible, independent and widely valued, and compares favorably with other IFIs and United Nations agencies.

However, to remain relevant and impactful in a rapidly changing institutional context, particularly in light of IFAD's new financial architecture, the upcoming IFAD14 replenishment, and the growing role of artificial intelligence, the function must become significantly more strategic, selective, and learning-oriented, with more relevant and timely products, stronger follow-up, and enhanced alignment between independent and self-evaluation.

Attention to the opportunities and challenges in increased use of AI is essential.

## 2.2 Credibility: strong and trusted, with the need for a more strategic approach

26. The credibility criteria assess to what extent the evaluation function operates in line with accepted practices when it comes to the quality of work, including the methodology for collecting and analyzing data, the timeliness, transparency and involvement of stakeholders in the evaluation processes, the application of ethical standards in the work processes and the relevance of evaluation issues.
27. **The evaluation guidance documents follow common standards.** The main guidance documents for evaluation (the 2021 Evaluation Policy and the 2022 Evaluation Manual) are aligned with ECG, UNEG and EvalNet standards, and in line with similar documents in other multilateral organizations. The exception is that the recalibration of ODE is not yet fully reflected in the documents.
28. **The accountability role is stronger than the learning role.** While respondents in interviews and in the survey find that the evaluation function plays both an accountability role and a learning role, the accountability role is currently dominating (figure 1)<sup>6</sup>,

<sup>4</sup> OECD. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449.

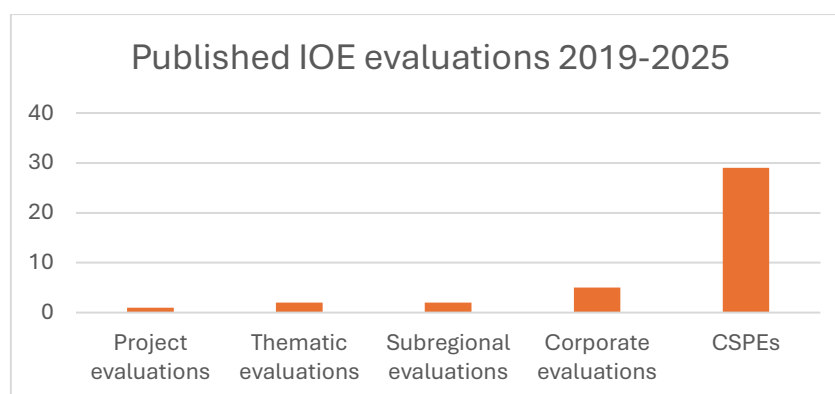
<sup>5</sup> A detailed overview of responses to the evaluation questions listed in the EPR evaluation framework can be provided upon request.

<sup>6</sup> The focus on accountability is commented by respondents, including the following: "IOE seems to have more a compliance role than to support learning with teams." "The lesson learnt from PCR can be disseminated better to inform the new generation of projects in the country."



34. **IOE and other respondents acknowledge limited country-level monitoring and evaluation (M&E) capacity.** This is one of the limiting factors for credible and reliable project data. Rather than doing their own country-level evaluation capacity-building, IOE has decided to cooperate with the Global Evaluation Initiative, including the regional Centers for Learning on Evaluation and Results. The EPR panel supports this approach.
35. **The product mix in the three-year IOE work programme differs from work programmes of other IFIs.** The main difference is the large proportion of CSPEs among the 40 published evaluations in the review period (2019-2025). IOE has completed 29 CSPEs, 5 corporate evaluations, 2 thematic evaluations, 1 project cluster evaluation and 2 subregional evaluations (figure 2).

Figure 2: Published IOE evaluations 2019-2025



36. The broad coverage of CSPEs is a result of increasing demands from operations and the Executive Board and fills the previous gaps in reliable assessments of implementation of IFAD's country strategic opportunities programmes (COSOPs). Among survey respondents, 39 per cent rated CSPEs as the most useful IOE product, while 36 per cent rated thematic evaluations as most useful and 34 per cent felt corporate evaluations were most useful.
37. **The quality of IOE evaluations is generally good but can be further improved.** Most of the survey respondents expressed agreement (including 29 per cent that strongly agreed) with the statement that IOE's approaches and methodologies are rigorous.
38. As part of the EPR process, a review of a sample of 20 IOE evaluations was done to provide an overview of the methodologies and overall quality of evaluation products (see appendix 5). The sample includes all corporate-level, thematic, project cluster and subregional evaluations published during the review period, as well as nine CSPEs, randomly selected. Overall, the evaluations reviewed demonstrate high quality and reflect a strong appreciation of methods and evidence, even in the context of ambitious evaluation scopes and complex evaluands.
39. **At the same time, several areas for improvement were identified.** These include greater transparency regarding data quality and triangulation, more systematic attention to the environment in which IFAD operates and enhanced reflection on the external validity<sup>9</sup> of findings. Most evaluations do not provide sufficient clarity on how evidence from different sources was triangulated and used to substantiate findings relative to the initial evaluation questions. The review also found limited emphasis on contextual understanding which may constrain deeper insight into the environments in which IFAD-supported interventions operate. Strengthening the use of the theory of change, particularly its articulation in relation to empirical analysis, would further improve analytical clarity. More broadly, adopting

<sup>9</sup> External validity (or generalizability) refers to how well the findings from an evaluation can be expected to apply in other settings.

a more rational and formal approach to evaluation design, moving beyond a collection of methods and tools toward a coherent, structured framework, would support a deeper understanding of evaluation quality. The integration of innovative approaches, such as contribution analysis, could also enhance assessments of coherence and causal inference.

40. Expanding attention to these dimensions would improve the validity, reliability and overall utility of future evaluations, while maintaining the high standards of methodological rigour already evident across the sample analyzed.
41. **More strategic prioritization in IOE's work programming will strengthen coherence and impact.** Documents presented to IFAD's Governing Council and inputs from IOE show that IOE's work programme follows common good practice among IFIs when it comes to consultation with Management and IFAD's governing bodies. The consultations give Management and the Board possibilities for raising priority areas, flag COSOPs and graduation issues and discuss areas of strategic or cross-cutting importance at corporate level. Higher plane evaluations are guided by Management inputs, evaluation coverage, strategic relevance, critical issues raised from previous evaluations and timing of strategy and/or policy implementation and revisions. IOE develops initial scoping notes which further highlight key areas and timing considerations. For example, IOE decided in 2026 to upgrade a rapid review of human resource issues to do a full corporate-level evaluation (CLE), noting the lack of evaluation coverage and strategic issues raised in the decentralization and IFAD 11/12 evaluations. Emerging areas relevant for determining future directions for IFAD were also prioritized, including supplementary funding and private sector engagement.
42. In particular corporate evaluations have brought insights on important strategic topics. One example is the evaluation of decentralization, another is the IFAD11&12 evaluation. One of the survey respondents commented: "*The evaluation of IFAD11 and IFAD12 has been particularly useful because it candidly raised many issues that remain relevant for the entire organization.*" *Several of the interviewees made similar comments.*
43. Despite the broad consultations and internal considerations, doubts arise regarding the strategic relevance of some parts of the work programme, in particular the continuation of doing project evaluations<sup>10</sup> and initiating CSPEs where the IFAD portfolio is very small or a COSOP is not planned. IOE has the capacity to go in depth. The demand-driven nature of the work programme — driven by Management and Board requests for country evaluations — may be pulling IOE away from an even more impactful role. The IFAD11 and IFAD12 evaluation<sup>11</sup> was cited by multiple Board members as exemplary models of strategic, high-value evaluations. These evaluations are described as 'valuable and much appreciated' precisely because they raised difficult institutional issues in a candid and useful manner.
44. The argument is that this capacity should be concentrated on the most important issues where new knowledge could be brought to the table or on topics that are high on the agenda of the Board and senior management. This implies that it is less useful for IOE to do project evaluations and full CSPEs in some cases. One example of the latter is the Peru CSPEs where apparently only one project has been completed since the previous CSPE. The EPR panel underlines that there should always be a close linkage between a CSPE and a new COSOP. Another concern is that CSPEs use a standard approach. Sometimes the scope is too wide (for instance by covering ten years in a country that has undergone huge changes) or not sufficiently addressing current challenges (for instance by not discussing graduation in a country where this is on the agenda). Increased focus and selectivity in the scope could also help ensure

<sup>10</sup> IOE notes that the number of PPEs conducted has been significantly reduced, from 10-12, to 4-5 maximum when projects are still ongoing

<sup>11</sup> Corporate-level evaluation of the institutional and operational performance under IFAD11 and IFAD12  
Document: EC 2026/132/W.P.7

that the IOE capacity is used in areas that are most useful. This will help release the potential for IOE to play an even more important proactive and strategic role in the future. Conscious and well-managed use of AI is important in this context.

45. IOE should rebalance its work-programme to prioritize:
  - Strategic corporate-level evaluations aligned with replenishment cycles and major institutional reform processes;
  - Thematic and cross-cutting evaluations on topics of direct institutional relevance (food security, private sector engagement, sustainability of impact, AI);
  - Meta-evaluations and/or synthesis studies drawing on existing CSPE findings, rather than initiating full CSPEs on limited portfolios or without a subsequent COSOP.
46. As guidance for determining the strategic relevance of an evaluation, the advice from the EAP on what constitutes influential evaluations is worth noting: (1) strategic orientation and policy relevance; (2) rigorous methodology and credibility; (3) stakeholder engagement; (4) clarity of findings; (5) actionability of recommendations; (6) time horizon of evaluative knowledge; (7) timeliness; and (8) dissemination and communication.
47. **The credibility of self-evaluation is developing.** The Development Effectiveness Framework serves as the backbone of IFAD's performance and results architecture. The 2021 updated Development Effectiveness Framework 2.0 integrates different streams of evidence – project completions reports (PCRs), COSOP reviews, impact assessments and core outcome indicators into a coherent system that is used to assess progress toward IFAD's strategic objectives and commitments under IFAD12 and IFAD13. The focus of impact assessments is shifting from being primarily tools for corporate reporting to being tools for strategic learning and adaptive management.
48. Management's self-assessment concludes that IFAD's self-evaluation function is significantly more mature now than at the time of the 2019 EPR: *The revised Evaluation Policy, the Development Effectiveness Framework (DEF 2.0), the strengthening of ODE, improvements in supervision reporting, enhanced quality assurance standards for PCRs and CCRs<sup>12</sup>, and the growing integration of impact assessments have all contributed to greater methodological coherence and a clearer articulation of roles and responsibilities*" The EPR panel share this conclusion.
49. IFAD's corporate reporting tools – particularly the annual Report on IFAD's Development Effectiveness (RIDE) and the midterm review of replenishment commitments – has according to the self-assessment, strengthened the incentive for self-evaluation to provide data that is not only accurate but also strategically relevant.
50. The RIDE presents the Fund's performance in a given year and focuses on indicators and targets from the Results Management Framework. The RIDE reports in a cross-cutting manner on a range of indicators from human resources to institutional efficiency, country programme performance, project-level results and portfolio management. As with similar reports in other IFIs, RIDE focuses on IFAD's positive performance and achievements, with limited discussion of weaknesses and challenges. A more balanced presentation would increase the credibility of the report.
51. **Findings from the interviews and the survey show concerns with the variable quality of data for self-evaluation.** According to the ODE self-assessment, alignment with international standards has improved substantially. UNEG principles of evaluability, transparency and utility are reflected in standardized reporting formats and strengthened guidance. OECD-DAC criteria are operationalized across supervision, PCRs and country strategy completion reviews, while ECG principles inform methodological coherence. However, uneven implementation across regions –

---

<sup>12</sup> Country Strategy Completion Reviews

driven by varied partner and division capacities, staffing and operational realities – is still to be addressed for full institutionalization.

52. Management admits that its future effectiveness will depend on IFAD's ability to focus on consolidation, quality, learning and use. Rather than introducing further reforms, the emphasis should be on consistent application, quality assurance and managerial follow-through. The EPR panel fully agrees with this approach.
53. Key enabling conditions include sustained leadership by ODE, clear accountability arrangements with operational divisions and alignment between corporate guidance and operational realities on the ground. Furthermore, the self-assessment highlights the need to institutionalize learning cycles that systematically translate self-evaluation evidence into decision-making. It also points to opportunities to expand the strategic use of digital tools, data analytics and emerging technologies to strengthen evidence generation, synthesis and use. The EPR panel supports these considerations, including the need to take advantage of AI to improve efficiency and effectiveness.

### Key points

IOE evaluations are widely viewed as credible and methodologically sound, with improving efficiency. However, stronger strategic prioritization, stakeholder engagement, data quality, learning orientation and more focused, influential reporting are needed to maximize relevance, use and institutional impact.

The credibility of self-evaluation is better structured and more mature now than in 2019 and has promising future developments.

## 2.3 Independence and governance: A well-established and respected function with a need to ensure that independence is maintained

54. The main question under the independence criteria is whether IOE operates according to the definition of independence (organizational, behavioral, protection from outside interference, avoidance of conflicts of interest). This has implications for budgeting, work planning and staffing. This does not mean isolation but rather a credible level of interaction and relationship with stakeholders.
55. **IOE is independent.** The EPR findings do not raise any major concerns with IOE's independence along all dimensions of independence. On the contrary, independence is widely considered to be one of IOE's main assets. One of the survey respondents commented: *"IOE under its current leadership has shown full independency and rigour."*
56. The evaluation policy has a clear description of the basic structural and behavioural factors needed to ensure independence. IOE reports directly to the Executive Board, with no reporting line to Management. IOE has a separate budget and workplan, which is reviewed by the Evaluation Committee and approved by the Executive Board. IOE has final decision authority on its work programme and on the content of all evaluation products.
57. **The EPR panel finds the 2021 Revised Evaluation Policy to be in line with international standards and common practice** among similar organizations. The policy focuses on strategic and higher-level principles and does not, as stated in paragraph 17 in the policy, provide details of products, processes and methods of evaluation that may need frequent updating. It does not either provide details on the type of evaluations (independent and self) or on the underlying methodologies and approaches that should be left to the appreciation of the unit managers and be particularly respectful of the independent and informed choices to the head of the

IOE, after consultations. As mentioned in chapter 2.1 the EPR panel finds this sufficiently flexible as a framework for further improvements of the evaluation function. Apart from making some clarifications in the annexes, the policy itself does not require a revision.

58. The EPR panel strongly recommends that IFAD not only reaffirm its commitment to the 2021 Evaluation Policy but also take concrete steps to safeguard independence in practice.

#### Key point

IOE is widely regarded as independent, with strong structural safeguards, direct Board reporting, separate budget and full authority over its work and evaluation outputs.

Safeguarding independence of the evaluation function is essential for IFAD's credibility.

The 2021 Evaluation Policy is adequate and does not require fundamental revision. Targeted clarifications in the annexes are needed.

## 2.4 Utility: Recognized value but limited uptake

59. The utility criteria bring attention to the relevance and utility of the evaluation function (both independent and self-evaluation) with respect to improving IFAD's institutional performance, for meeting Management's need for strategic inputs and for learning for IFAD's end beneficiaries. It also focuses on dissemination and communication related to evaluation.
60. **Utility for IOE is generally good, with some exceptions.** Evaluation influence is contingent on multiple factors, including relationships with Management relating to recommendations' uptake, correctness, details and timeliness. The findings show that IFAD's Management and operations in general consider IOE's work to be relevant and useful.
61. IOE receives repeatedly positive feedback from government stakeholders on the utility of evaluations, in the form of informal written and verbal communications and feedback within stakeholder workshops. The direct use of CSPE findings and recommendations in the development of new COSOPs has increased as a result of a closer linkage between CSPE and COSOP. In 2025/26, 69 per cent of COSOPs will be preceded by a CSPE – up from 20 per cent in 2021. One of the survey respondents commented: *"I have used IOE's assessments for my daily basis job, either for portfolio review as part of a regional division's FO (former position), or for better assessing the quality of a COSOP being presented for approval (current job in ODE)."*
62. An interesting view expressed in some interviews is that being a subject for an IOE evaluation for some is considered positive, as it may help to get attention from senior Management and the Executive Board. Other interviewees raise concerns with the capacity of operations to respond to the large number of requests from IOE in connection with the evaluation processes and thus would appreciate clearer prioritization of the importance of IOE products. While individual evaluations may be useful, the large number of evaluations may put a burden on staff that undermines their usefulness. One of the survey respondents made the following comment: *"I have seen that the way IOE engages IFAD staff in evaluation is very diverse, depending on whether they are working on a thematic evaluation or a country strategy evaluation."*
63. The survey showed that almost half of the respondents (47 per cent) have read IOE products and 40 per cent declared they have used them to inform their work. About one third of the survey respondents are using IOE's work regularly (16 per cent) or very often (15 per cent), yet half (49 per cent) are using them at least few times.

- Only 16 per cent of respondents declared they have never used IOE's work in the past 12 months.
64. Four of the seven members of the Executive Board who provided complete responses to the survey expressed absolute familiarity with the work of IFAD's evaluation function and three are somewhat familiar with it. Six represents list A countries and one is from a list B country. There were no responses from list C countries. One board member noted: *"I mainly read evaluations. I found them to be very good, balanced and realistic: I like that they aim to facilitate learning and thus further improve IFAD's work."*
  65. **Potential for increased external collaboration and mutual reliance to strengthen utility.** Collaboration with other independent evaluation functions through evaluation networks is one possibility for increasing the quality, relevance and usefulness of evaluations. IOE is a member of both ECG and UNEG. IOE is currently prioritizing its involvement with ECG. ECG is dedicated to harmonizing evaluation work among multilateral development banks by working to strengthen the use of evaluation; providing a forum to share lessons; developing harmonized performance indicators and evaluation methodologies and approaches; enhancing the professionalism of evaluation; and helping build evaluation capacity.
  66. The Mutual Reliance Framework, one of the collaboration initiatives between the development banks, is intended to reduce duplication and increase collaboration to generate efficiencies and more seamless coordination, including when it comes to evaluation and development and use of AI tools. This offers possibilities for enhanced utility for IOE.
  67. **Although considered useful, criticism has been raised regarding IOE's recommendations.** In line with its guidance documents, IOE usually presents four to six recommendations. However, the recommendations are often multi-part and comprehensive, and there is a continuous debate on the level of flexibility versus prescription required. Interviews confirm this view. While most of the survey respondents reported that the recommendations were useful (39 per cent considered them to be very useful for their work and 43 per cent consider them to be somewhat useful), some also had critical comments.
  68. One of the criticisms is regarding the vagueness of recommendations. Survey respondents made the following comments: *"In a recent evaluation conducted by IOE of the country strategy, the findings and recommendations were very general and did not provide much support for the development of the next country strategy. It would be useful for the recommendations to be more specific."*
  69. *"Some of the IOE's recommendations seem theoretical and non-practical in certain context."*
  70. *"Sometimes recommendations are quite general and not really applicable to the project being assessed."*
  71. **Another criticism is related to the number of recommendations and sub-recommendations.** The quality review of IOE reports (appendix 5) found that although recommendations were generally evidence-based, there is a tendency to accumulate what could be described as "bags of recommendations." Evaluations typically included four to seven overarching recommendations, but approximately half of the evaluations also specified up to 22 sub-recommendations. While balancing the specificity and breadth of recommendations is inherently challenging, the "bag of recommendations" approach risks diluting the key messages of the evaluation and potentially creating confusion about priorities and focus. The EPR has heard many critical comments regarding this. One view is that the many sub-recommendations make the recommendations difficult to untangle. One of the survey respondents noted: *"At the time of the last EPR, IOE had started to reduce the number of recommendations to highlight the key recommendations. However, in recent years,*

*IOE has been increasing the number of recommendations through sub-recommendations. Five main recommendations become 25 sub-recommendations which are challenging to implement. Many of the thematic evaluations seem to present cut and paste recommendations. There is often (the recommendation) to increase the budget and staffing for nutrition, gender, youth, etc.”*

72. A specific and actionable proposal would be for IOE to conduct a technical meeting, to be held between IOE and Management before recommendations are finalized. This would serve as a space for mutual understanding, and agreement on the 'what' — not a compromise on evaluative independence, but a mechanism to ensure recommendations are clear, mutually understood and actionable.
73. The President’s Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA) tracks recommendation follow-up but lacks credible verification other than agreement from IOE and qualitative assessment of implementation.<sup>13</sup> Recommendations are tracked annually through PRISMA, prepared by IFAD Management and presented to the Evaluation Committee and the Executive Board. As responsibility for recommendations tracking falls under IFAD Management, there is no mechanism for independent verification or qualitative or systematic assessment of the results of implementation. There are also challenges in interpreting what the reporting of implementation implies.
74. **The influence of the Annual Report on the Independent Evaluation of IFAD (ARIE) may be further heightened.** The IOE annual flagship report has changed during the EPR review period both content-wise and in length (from 111 pages plus annexes in 2019 to 39 pages plus annexes in 2025). The previous Annual Report on Results and Impact of IFAD Operations was re-pitched and renamed as the Annual Report on the Independent Evaluation of IFAD (ARIE) in 2022 to provide the Executive Board and Management with a more comprehensive account of the evaluation activities undertaken by IOE, their findings and lessons to be learned. While continuing the approach of comparing findings across evaluations and of presenting and analyzing time series of ratings to identify performance trends, the new approach focuses on substantive findings. The 2024 ARIE examined two key themes: the relationship between cofinancing and project performance and inclusive rural finance interventions. The 2025 ARIE examines the sustainability of benefits from IFAD-funded projects. Even though the influence and use of the report have improved from 2019 to 2025, there is a potential to make the ARIE even more influential by sharpening the main findings and lessons further.
75. A more succinctly and focused report could be presented to the Executive Board as a basis for an annual session on evaluation results and recommendations.
76. **Variable views on the usefulness of PCRs.** Most survey respondents consider PCRs to be useful in assessing project achievements, facilitating sharing of experiences and insights among project stakeholders and contributing to the improvement of future project design and implementation. One of the critical comments is illustrated by the following quote: *“I think PCR reports are very good, but the timing is sometimes not useful. I do not see the value in completing a PCR report for projects around completion date when most the activities are incomplete or the project is seeking extensions. Some projects get extensions of about 5 months before closure, allowing (them) to execute sometimes up to 30%-40% of activities. This does not allow evaluators to receive complete information on the project.”*
77. **According to the 2025 RIDE, Management has introduced stronger quality assurance mechanisms for PCRs.** In addition, regional divisions have introduced stricter quality control mechanisms with enhanced peer review arrangements. In contrast to the former provisions, IFAD Management approves PCRs and can request that ratings be revised in cases where the justification is not robust enough.

---

<sup>13</sup> This could be commissioned to an external partner every two or three years.

### Key points

IOE evaluations are widely valued and increasingly used, but utility is constrained by recommendations with numerous sub-recommendations, staff burden, weak independent follow-up verification, mixed PCR timing, and untapped potential for collaboration to enhance relevance and efficiency.

A succinct and focused ARIE could be presented to the Executive Board as a basis for an annual session on evaluation results and recommendations.

Management's self-evaluations are generally seen as relevant and useful for learning and performance improvement, but their effectiveness is weakened by staff burden and variable engagement.

## 2.5 Complementarity of independent and self-evaluation: Improvements with remaining challenges

78. In theory Management's self-evaluation and IOE's independent evaluations should be complementary and avoid overlap and repetitions. This implies a clear division of labor in monitoring and evaluation. It also implies that in-depth independent evaluations can rely on data collected through self-evaluation to be credible. The prerequisites for complementarity include shared guidelines and interpretation of rating criteria and a level of trust between the two parts of the evaluation function. Jointly the two parts should fill a major part of IFAD's need for accountability and learning.
79. **Improved collaboration between IOE and ODE.** The findings indicate a more productive and aligned relationship between the two functions compared to 2019. One significant change since the 2019 EPR has been greater collaboration between IOE and Management, specifically ODE, in developing the IFAD Evaluation Policy and the Evaluation Manual. The revised policy is the first to include self-evaluation, ensuring a holistic view and approach of the IFAD evaluation function, and enabling greater harmonization and collaboration. ODE has been assigned a role as the Management focal point for IOE, and coordinates Management responses for corporate and thematic evaluations. ODE is invited to all consultations for the IOE work programme and budget and is copied on all key evaluation communications. IOE in turn provides comments on ODE-developed annual products such as the RIDE and the PRISMA.
80. The ODE self-assessment emphasizes the strategic importance of reinforcing complementarity between self-evaluation and independent evaluation. Greater methodological alignment, routine technical exchange and coherent quality assurance can enhance credibility, reduce duplication and strengthen learning. This depends on sustained collaboration between ODE and IOE, shared standards where appropriate, and governance arrangements that promote mutual reinforcement while safeguarding independence.
81. The EPR panel finds that the clearer delineation of roles and the introduction of shared conceptual foundations have created an opportunity to build a genuinely complementary system. Capitalizing on this joint foundation requires deep alignment in analytical expectations, strong cross-functional dialogue during key stages of the project and country cycles, and a systematic approach to synthesizing evaluative insights to support organizational learning. The findings show that the evaluation function is moving in this direction but has still a way to go.
82. **Differences in rating of project performance.** Management and IOE<sup>14</sup> rate the same evaluation criteria, with the exception of impact, which is not always rated by

---

<sup>14</sup> IOE validates PCRs

Management (10 out of 23 PCRs in 2024 did not include impact ratings). The 2025 ARIE identifies statistically significant differences between the IOE and PCR ratings across all evaluation criteria. IOE ratings are more critical on average than Management ratings. The greatest disconnects are in relevance, (-0.46), followed by scaling up (-0.43) and gender equality and women's empowerment (-0.39). The disconnect has improved over time. According to the 2025 RIDE the gradual application of the revised evaluation manual is yielding results in bridging this gap. Management does not consider the disconnect to be a cause for concern; on the contrary, it generally enriches the debate on results and development effectiveness measurement. One of the explanations of the disconnect expressed in the interviews is that Management and IOE follow slightly different rating manuals and have differences in the interpretation of the criteria.

83. The view of the EPR panel is that, if disconnects in rating are natural and "healthy" given Management's and IOE's different perspectives, differences in guidelines are problematic and should easily be adjusted. The differences in interpretation of the guidelines should furthermore be resolved through dialogue between ODE and IOE. The current disconnect in IFAD is not any bigger than in other IFIs.

### Key points

Complementarity and collaboration between Management self-evaluation and independent evaluation have been improved during the review period. Rating disconnects still exist, some of it is due to differing guidelines.

There is a need for further alignment, dialogue and shared standards to strengthen accountability and learning.

## 2.6 Management and staffing: Need for a broader skills-mix

84. These criteria deal with the capacity, effectiveness and efficiency of IOE and Management's self-evaluation function, and identifies what is needed for the evaluation function to play its intended role now and in the foreseeable future.
85. IOE has sufficient resources but needs a broader skills mix to accompany the evolution of IFAD towards providing more private sector and financial support. IOE has, according to its self-assessment, generally had sufficient resources to carry out its work programme. However, the increase in coverage in 2026, and the growth in the number of evaluations has outpaced the growth in budget, which has required an increased workload of staff as well as greater efficiencies. The addition of new staff has helped to meet some of the demand. Capacity is supplemented by the use of consultants, analysts and interns.
86. To match the implications of IFAD's new financial architecture, the introduction of new investment instruments, increase in non-sovereign operations and greater use of artificial intelligence in the work process, IOE needs to adjust its capacity to meet these needs and develop expertise in areas such as:
- Non-sovereign operations and blended finance instruments;
  - Private sector development and value chain finance;
  - Impact measurement methodologies adapted to private sector contexts.
87. Findings indicate that IOE has taken some steps to move in the right direction but is not yet where it needs to be to creditably deal with the new challenges IFAD is facing.
88. In order to play a strategic role, the evaluation function needs to develop further its proactive orientation. This may also require different skills to the current range. Increased training and capacity development is important to make sure IOE staff are motivated, technically strong and professionally up to date.
89. **Credible self-evaluation requires more than updated guidelines.** While the system and guidelines for self-evaluation have been updated to be in line with

common practice among IFIs and strengthen the methodological clarity, the adjustment in the application has not yet fully happened. Management's own self-assessment states that the system now requires a shift from structural reform toward consistent, organization-wide application. This places emphasis on improving the reliability and integration of evidence across the project, country and corporate cycles, rather than on further procedural redesign.

90. **Varying levels of M&E capacity across project management units and country offices is an obstacle.** Management's view is that this variability is not a reflection of lack of commitment, but stems from structural constraints – competing operational demands, turnover or the need for more sustained hands-on support in applying methodological guidance. The teams operate under different resource conditions, staffing models and external environments, which naturally affect the regularity and quality of data collection. Management finds in its self-assessment evidence of persistent gaps in workload balance, staff time, technical skills and incentives for evaluative thinking.
91. Many staff highlighted the importance of continuing to invest in practical training, targeted accompaniment, and clearer guidance that translates evaluative principles into feasible, context-appropriate practices.
92. Several interviews point out the weak M&E capacity in project management units and partner governments as major obstacles for getting sufficient project data quality. It is obviously essential to enhance efforts to address this problem.
93. Addressing these constraints is essential not only to improve the quality of self-evaluation products, but also to embed evaluative reasoning into operational decision-making. Strengthening capacities at the country office level, where evaluability challenges are often most acute, is especially important. Feedback the EPR panel has received from country representatives emphasizes the need for systematic training in basic skills related to procurement, financial management, M&E and human resources issues.
94. According to the 2025 RIDE, IFAD's knowledge and learning team (part of ODE) in early 2025 assumed management of IFAD's Operations Academy, a key institutional mechanism designed to strengthen the operational capacities of country teams, technical experts and partners. The Operations Academy is composed of three main instruments: e-learning courses (level 1), interactive webinars and in-person training events (level 2), and a structured mentorship programme. It is recommended to closely monitor and assess after two to three years to validate whether the efforts have been successful.

### Key points

IOE has resources but needs broader skills and proactive capacity to meet IFAD's evolving demands, especially in non-sovereign operations.

Varying level of M&E capacity across regions and country offices is a concern. Stronger, consistently applied self-evaluation requires practical training, country-level capacity-building, and better integration of evidence into decision-making.

## 2.7 Enabling environment for learning and development: A great potential for further improvement

95. Included in these criteria is an assessment of IFAD's organizational culture, and which factors support and restrict use and learning from evaluations.
96. **IFAD is to some degree considered to be a learning organization.** In its self-assessment IOE staff rate the learning and evaluation culture in IFAD highly. Furthermore, IOE sees the high demand for evaluation from IFAD Management as a

reflection of a positive organizational environment for evaluation, as well as an understanding of the value brought by independent evaluation to corporate- and country-level strategies and programmes. The survey supports these views. One of the comments were: *"Evaluations help staff to think out of the day-to-day routine in order to reflect on the need and efficiency what the organization is presently doing. Lessons learnt should stimulate staff and senior management to explore ways for improvement."*

97. Most of the survey respondents (90 per cent) report that IFAD is a learning organization. This result is consistent across the tenure period but tends to increase slightly for respondents with more than six years of work experience at IFAD. However, administrative staff are slightly less likely to consider IFAD as a learning organization with 15 per cent reporting that it is not at all the case.
98. Findings from interviews offer a more nuanced picture. Evidence is not always used when planning new projects. For instance, when evidence from evaluations shows that a type of operation is not working, this is not always taken into account when planning new operations.
99. **Timeliness of the publication of IOE reports influences broader learning.** Findings show that delays in publishing IOE reports prevent broad use of their findings and recommendations. One survey respondent commented: "IOE products are often published with significant delays. This affects their timeliness and reduces their potential influence, as only a small group of internal staff may see draft versions, while the wider stakeholder community – including governments, project teams and development partners – gains access much later, when opportunities for learning and course correction are more limited."
100. **Enabling and restricting factors for learning.** The most cited enabling factors for learning by the respondents are collaboration and teamwork across projects and departments (71 per cent) and access to training, resources and development programmes (58 per cent). Feedback mechanisms for learning are reported by 41 per cent of the respondents.
101. Respondents also shared a variety of factors that may restrict learning at IFAD, like political priorities and organizational culture. Discussing knowledge management (KM) one stated: *"There is no clear structure for KM in place anymore, so nobody feels responsible for leading KM. By making it everyone's responsibility, it became nobody's responsibility. Also the KM relations between HQ and the field are unclear."*
102. Management's self-assessment notes that staff highlight that the practical uptake of tools and guidance depends heavily on leadership engagement, resource availability and the perceived usefulness of evaluative evidence for operational decision-making.
103. The self-assessment suggests that IFAD is moving toward a more cohesive evaluative culture. It points at encouraging examples such as thematic clinics, structured debriefs and peer-exchange formats, but finds that these practices remain uneven and are not yet anchored in a systematic cycle that routinely consolidates evidence across supervision, completion reporting, surveys and impact assessments. Disparities persist between regions, project types, and levels of M&E capacity. While reporting systems have improved, learning processes remain uneven and are often dependent on individual initiative rather than institutionalized practice.
104. The EPR panel agrees with Management that strengthening the enabling environment – through targeted capacity-building, sequencing of evidence generation and deliberate institutional learning cycles – is critical to ensure that evaluative evidence consistently informs project design, supervision and strategic positioning.
105. Management also finds that aligning the timing of self-evaluation products with key operational decision points – particularly project design, supervision cycles and COSOP reviews – would ensure that evaluative evidence is available when decisions are made.

106. The use of artificial intelligence to process large volumes of evaluative data – such as PCRs, supervision records and M&E datasets – would allow IFAD to systematize lessons learned more efficiently, classify them by thematic area, project type or geography and transform dispersed information into actionable knowledge.
107. The benefits of closer alignment between Management’s self-evaluation and IOE’s independent evaluation include increased credibility of evaluative evidence, a more predictable evaluative environment, and a stronger culture of accountability and learning.
108. **IOE applies a neuroscience-based approach to communication.** IOE has utilized a neuroscience-based approach in its engagement with IFAD stakeholders, understanding that communication styles can have a significant impact on receptivity to findings, particularly when they are critical or negative. The approach is centered around using compassion and empathy to create psychologically safe spaces for expression and sharing unbiased views, managing complexity, understanding intrinsic motivations and using future-focused framing.
109. The EPR panel members’ own experience indicates that a holistic approach to communication is preferred by looking at content, relationships, context and channels together, so that everything works in a coherent way. Timeliness of dissemination and communication is particularly important in a complex organizational setting like IFAD.
110. **Extraordinary variety in IOE communication.** The EPR panel is impressed by the large variety of dissemination and communication products coming from IOE, which exceeds what most other IFI evaluation departments are doing. The Board members responding to the survey consider IOE to be effective or very effective in disseminating and communicating evaluation results, as well as in facilitating the sharing of experiences among project stakeholders and promoting lesson-learning. Board members are less familiar with IOE’s outreach to partner countries, as three members were unable to assess these aspects.
111. The IOE website has over 400,000 views. An exercise using AI for social media sentiment analysis of IOE was conducted from January – July 2025. It analyzed engagement on IOE’s social media platforms (X and LinkedIn) which have approximately 3,600 and 8,000 followers respectively.
112. **Variable views on IFAD’s learning absorptive capacity.** The absorptive capacity is the organization’s ability to identify, assimilate, transform and use the findings and lessons from the evaluation function and other providers of evidence. The absorptive capacity is dependent on openness to new knowledge, but also the capacity to respond and the timeliness of being exposed to the findings and lessons and the incentive for doing so. For an evaluation function it is a challenge to find the right balance between too little and too much when it comes to evaluations and evaluation processes. IOE staff consider IFAD’s absorptive capacity to be adequate and mention that one to two corporate or thematic evaluations per year should be within IFAD’s ability to respond to and absorb. Several of the interviewees question if this is the case and find that Management’s absorptive capacity is a challenge. The argument is that IOE is better resourced than other parts of the organization. The large number of evaluations challenge the capacity to respond to requests during the evaluation processes and to deal with the lessons and recommendations emerging from the evaluations.
113. **Responsible and strategic use of artificial intelligence in independent evaluation.** Artificial intelligence (AI) is likely to play an increasingly important role in the future of independent evaluation, particularly in the analysis of large and complex bodies of qualitative and quantitative evidence, but also for improved access to evaluation insights by all in IFAD to feed project design in a systematic and “compulsory” way.

114. While these tools can significantly support evaluators, their use also raises important questions related to bias, transparency, explainability, data protection and overreliance on automated outputs. Addressing these risks is essential to safeguard the independence, credibility and trustworthiness of evaluation. To this end, there is merit in developing clear guidance on the ethical and appropriate use of AI in evaluation, building on and contributing to ongoing work with peer institutions, including other IFIs, ECG, UNEG and related networks. Such coordination would help promote coherence with emerging international good practice while avoiding fragmented or inconsistent approaches.
115. The EPR panel finds the “Strategy for Integrating Artificial Intelligence in IOE 2024-2026” to be a good start for providing a structured framework for prioritizing use cases, managing risks, clarifying governance arrangements and ensuring alignment with evaluation standards and values. The use of AI should be encouraged by providing complementary investments in staff training and AI literacy. This is essential to ensure that evaluators remain informed and critical users of these tools, with AI serving to augment – rather than replace – professional judgment and human oversight.
116. AI is a strategic opportunity for the evaluation function – to improve coverage, timeliness, accessibility of evidence, and efficiency. But it requires deliberate governance: clear ethical guidelines, staff capacity development, and coordination with peer institutions. AI should augment professional evaluative judgment, not replace it. IOE's current strategy is a good foundation; implementation and staff literacy are the next priority steps.
117. **Roles, Responsibilities, and the Imperative of Independence in Evaluation.** The 2021 Evaluation Policy of IFAD establishes a robust framework for independent evaluation, clearly defining the roles and responsibilities of the Executive Board, the Evaluation Committee, IFAD Management, and the Independent Office of Evaluation (IOE). The Executive Board and the Evaluation Committee are tasked with monitoring the policy's implementation, ensuring that independent evaluation is sufficiently resourced, recognized, and systematically followed up. The Board also holds the critical responsibility of appointing the IOE Director through a detailed and transparent process, as outlined in the policy annex. IFAD Management, in turn, is accountable for safeguarding IOE's independence, engaging in regular consultation with IOE, and providing full and unrestricted access to information and support necessary for evaluations. IOE, reporting directly to the Executive Board, is responsible for adhering to international norms and standards set by ECG, UNEG, and OECD-DAC EvalNet.
118. However, a well-designed policy alone is not sufficient to guarantee independence of IOE. The credibility and effectiveness depend on the practical preservation of independence in all processes, including Work-Program development, evaluation implementation and publication, and Budget and human resources (HR) management and staff hiring. Independence must be actively upheld in day-to-day operations, decision-making, and institutional culture.
119. The process of hiring the Director of IOE is a cornerstone of IOE's independence. This process must be fully transparent, merit-based, and free from external influence. To avoid conflicts of interest and ensure objectivity, recruitment and management of IOE staff must be fully autonomous from IFAD Management within the IFAD HR framework. Beyond HR and hiring, IOE's operational autonomy must be preserved in all aspects of its work, including unrestricted access to Information and freedom to publish and disseminate.
120. Independence of IOE is not merely a procedural requirement; it is the foundation of trust in IFAD's independent evaluation function. When independence is compromised—whether through interference in HR processes, politicized hiring, or restrictions on operational autonomy—the credibility of evaluations is eroded. This,

in turn, weakens IFAD's ability to learn, adapt, and demonstrate accountability to its stakeholders and the populations it serves.

121. Management led self-evaluation is by nature different from IOE when it comes to structural independence. Even though self-evaluation is conducted by those in charge of the project or process to be evaluated, also these evaluations should be sound and free of biases. As the policy states, the six principles for evaluation – Usefulness, Impartiality and credibility, Transparency, Partnership, consultation and collaboration, Evaluability and Value for money/cost-effectiveness – applies to both self- and independent evaluation. The two parts of the evaluation function are to be complementary to ensure reliable and timely evidence for learning, accountability and further improvements of IFAD's operations.

**Key points**

IFAD values evaluation and learning, but evidence of its use is uneven. Delays, capacity gaps, and weak knowledge management limit uptake.

Better timing, leadership, systematic learning cycles and AI-enabled synthesis could strengthen impact.

Given the growing use of artificial intelligence, the Independent Office of Evaluation needs a clear strategy and explicit guidance to ensure AI is used consistently, responsibly and in line with evaluation principles, independence and accountability.

### 3 Recommendations

122. The overall objective of the recommendations is to increase the alignment between the evaluation function and IFAD's strategy and reform processes. The evaluation function has the potential to provide even more evidence and knowledge for strategically relevant processes like replenishment and reform initiatives. Alignment means close interaction between IOE and ODE/Management and the Executive Board. For the evaluation function it has implications for work planning, for selection of evaluation products, and for timeliness of dissemination and communication. For Management it has implications for sharing knowledge needs, willingness to seek and base decisions on evidence, reflect on it, improve and adapt. Close alignment has the potential of enhancing development effectiveness and ultimately the achievement of IFAD's operational objectives. Increased and conscious use of AI and other technologies can play a pivotal role.
123. **Recommendation 1:** *IOE should rebalance its product mix by improving selectivity, so that resources are focused on strategically chosen evaluations that enhance the usefulness of evaluation for timely and informed decision-making.*
124. This implies clearer prioritization of the strategically most relevant evaluations while taking into account Management's absorptive capacity for involvement in the evaluation processes and implementation of the recommendations. While IOE in its work programme should listen to demand from the Board and Management, it should be equally cognizant of current and future needs for evaluative evidence, including taking into account the IFAD14 replenishment timeline. Instead of continuing to do project evaluations and CSPes on limited project portfolios or when it is not followed by a COSOP, IOE should do meta-evaluations and synthesis studies or information notes on cross-cutting and timely issues, taking advantage of existing evaluation findings and lessons to inform the priorities and strategic direction of IFAD, including for non-sovereign operations. This has implications for work planning, staff skills-mix, methodology for involvement of country-level stakeholders and the use of technologies.
125. **Recommendation 2:** *Management should increase efforts to improve self-evaluation and together with IOE refine complementarity with independent evaluations, enabling stronger learning loops and more consistent use of evaluative insights in decision-making.*
126. As self-evaluation is maturing through implementation of the current planned development by ODE as expressed in the IOE self-assessment, the division of labor between ODE and IOE should be reviewed and revised to avoid overlap and strengthen the comparative advantage of the two parts of the evaluation function. More coherence across methodologies and criteria should also be ensured through appropriate measures. This has implications for independent and self-evaluation of projects and country programmes. A follow-up assessment of the self-evaluation function should be undertaken in two years' time when the reform is more advanced to assess the relevance, effectiveness and overall coherence of the new system, practices, quality and results of the revised self-evaluation function.
127. **Recommendation 3:** *IFAD should strengthen its learning culture by fostering closer collaboration between Management and the evaluation function, with the objective of improving the uptake and use of evaluation lessons in strategic and operational decisions.*
128. This includes improving learning loops throughout the Fund, introducing areas for systematic reflection, and revising the methods for reporting on and validating the implementation of IOE recommendations. It also means ensuring that key evaluation lessons from evaluation are identified and used at the early stages of a project and COSOP design. This implies proactive use of learning methods in all parts of IFAD, and deeper involvement of the Executive Board including sessions devoted to

discussion of an improved version of IOE's annual flagship report ARIE and other key evaluation products. Enhanced learning depends more on behavior, incentives and well-organized processes than sophisticated and costly learning measures.

129. **Recommendation 4:** *IOE should increase the focus and clarity of its recommendations to strengthen their actionability and decision-making value.*
130. Closer and better engagement with Management in formulating and prioritizing recommendations will ensure that recommendations are mutually understood, clear, useful, realistic and actionable. In order to safeguard the principle that recommendations are well-substantiated by findings, IOE should impose greater transparency regarding data quality and triangulation, more systematic attention to the environment and enhanced reflection on the external validity of findings. Sub-recommendations should be avoided, and monitoring of recommendations should be verified by independent evaluations. Timeliness is essential to strengthen possibility for uptake of lessons and recommendations.
131. A technical meeting on recommendations could be held between IOE and Management before recommendations are finalized, not to discuss the substance of the recommendations based on evaluative independence, but as a mechanism to ensure recommendations are mutually understood, realistic, and actionable.
132. An independent regular verification mechanism for PRISMA by an external partner could be considered to assess uptake and results of the recommendations and their implementation.
133. **Recommendation 5:** *ODE and IOE should strengthen collaboration with relevant global initiatives<sup>15</sup> on M&E capacity development to clarify roles, strengthen capabilities, and improve the decision-making value of monitoring and evaluation both internally and with country partners.*
134. Adequate M&E capacity among IFAD's operational staff at all levels and partner countries' representatives and project management units is essential for credibility of data and capacity to implement projects. IOE and ODE should continue to strengthen their own M&E capacity-building efforts and build stronger collaboration with relevant global initiatives.
135. **Recommendation 6:** *The Executive Board should reaffirm its commitment to the 2021 Evaluation Policy and safeguard the independence of IOE and the role of ODE by clarifying the annexes in the policy.*
136. The EPR finds that the policy itself is in line with best practice of comparable organizations and sufficiently flexible as a framework for further improvements of the evaluation function. The annexes should explain the roles and responsibilities of ODE, clarify that the Board is in full charge of the selection and dismissal of the evaluation director without interference, and that sufficient time is allocated in the Evaluation Committee for each evaluation presented<sup>16</sup>.

---

<sup>15</sup> Like the Global Evaluation Initiative (GEI), including the Center for Learning on Evaluation and Results (CLEAR)

<sup>16</sup> The annexes should clarify: (i) The role and responsibilities of ODE, (ii) That sufficient time be allocated to each evaluation product presented to the Evaluation Committee, (iii) That the Board is exclusively responsible for selecting and dismissing the IOE Director, and (iv) That the recruitment process should be merit-based, fully transparent, and free from external interference, including from IFAD Management, the President and the outgoing IOE Director

## **4 Appendices**

Appendix 1: EPR Evaluation Framework ([LINK](#))

Appendix 2: EPR approach and methodology

Appendix 3: List of interviewees

Appendix 4: Surveys of staff and members of the Executive Board

Appendix 5: Assessment of IOE evaluation reports

## Appendix 1: EPR Evaluation Framework ([LINK](#))

### Appendix 2: EPR approach and methodology

The Evaluation Framework (submitted by the EPR in November 25 and approved by IOE and Management) was developed in response to the approach paper (9 October 2025) and recognizes guidelines for such reviews (ECG, UNEG and OECD/DAC) including conversations with selected key stakeholders and an initial identification of issues related to the evaluation function in the current and emerging internal and external context. It was also based on the panel members' broad expertise and experience and the time and resources available to complete the EPR. The findings and recommendations of the 2019 EPR formed the basis for this EPR.

The EPR has also assessed how the evaluation function complies with the requirements of the 2021 revised Evaluation Policy. This policy presents, for the first time, a comprehensive framework through which self-evaluation (conducted under the aegis of Management) and independent evaluation will be planned, conducted and used. The revised policy seeks to promote complementarity and synergy between the two. The principles underlying the policy include usefulness, impartiality and credibility, transparency, partnership, consultation and collaboration, evaluability, and value for money/cost-effectiveness. The policy also describes the procedures for selection, appointment, dismissal and performance assessment of the IOE Director.

The review has been formative and forward-looking, anticipating how the evaluation function can be strengthened in the future. The panel has taken an impartial view, based on knowledge of best practice, with respect to ethical norms and confidential information.

The methodology included the following elements:

- Review of the 2019 EPR and the implementation of the recommendations.
- Review of IOE's self-assessment report (December 2025) which consists of a summary of IOE's assessment of its functions as per the six key clusters guided by the evaluation questions. At the end of each cluster is a SWOT analysis. The document also contains an update on the implementation of the 2019 EPR recommendations. It focuses specifically on independent evaluation in IFAD and will accompany a separate self-assessment developed by IFAD Management on self-evaluation.
- Review of the report on self-assessment of IFAD's self-evaluation function (draft 16 January 2026). Covering the period 2019–2025, it examines how the self-evaluation system has evolved since the 2019 EPR, how it has been shaped by successive institutional reforms, and how it now supports accountability, learning and performance improvement across the full operational cycle, from design through supervision and completion.
- Review of relevant documents related to the evaluation function such as the evaluation policy, the evaluation manual, guidelines for project completion reports, strategic plans, work programmes and annual reports.
- Systematic assessment of a sample of IOE corporate and country evaluation reports published in the review period. The objective was to provide an overview of the methodologies and overall quality of a sample of IOE evaluation products. A sample of 20 evaluations was selected which includes all corporate-level, thematic, project cluster and subregional evaluations, as well as a purposive selection of nine CSPEs, randomly selected with the support of Copilot AI. Evaluations were assessed against a set of indicators grouped into seven quality dimensions: scope, reliability, construct validity, internal validity, external validity, data analysis validity, consistency.

- Assessment of a sample of self-evaluation reports produced in the review period.
- Review of relevant corporate IFAD documents.
- Survey of IFAD staff and members of the Executive Board to assess familiarity with, assessment of and use of the knowledge from the evaluation function. The survey was prepared in November 2025. To ensure that the survey was well adapted and intelligible from a beneficiary's point of view, the survey was then tested and reviewed by IFAD. It was open for two and a half weeks in November/December 2025 and was closed with 512 views. In total, the survey received 400 individual responses: 280 complete answers as well as 125 incomplete. Among the respondents, eight were members of the IFAD Executive Board, and the rest were from IFAD staff. In sum, the response rate for board members was 22.2 per cent (8 respondents out of 36), and the response rate for IFAD staff was 43.5 per cent (392 respondents out of 900 IFAD staff).
- Interviews/focus groups in-person and virtual with stakeholders (IOE, the Executive Board, the President and other members of senior Management, country directors, country partners, selected staff, and other key informants).

## Appendix 3: List of interviewees

### IOE staff

Hannah Den Boer, Evaluation Analyst

Mona Fetouh, Deputy Director

Massiel Jimenez, Evaluation Officer

Steven Jonckheere, Senior Evaluation Officer

Hansdeep Khaira, Senior Evaluation Officer

Diane Abi Khalil, Senior Evaluation Officer

Mikal Khan, Evaluation Officer

Kouessi Maximin Kodjo, Lead Evaluation Officer and Chief of Section

Monica Lomena-Gelis, Senior Evaluation Officer

Anne Sophie Lenoir, Evaluation Assistant

Sarah Mahdar, Temporary Administrative Assistant

Indran Naidoo, Director IOE

Oanh Nguyen, Lead Evaluation Officer and Chief of section Paolo Silveri, Lead Evaluation Officer and Chief of section

Shaun Ryan, Administrative Associate

Laure Vidaud, Administrative Associate

Alexander Voccia, Senior Evaluation Communication & Knowledge Management Specialist  
| Coordinator, Evaluation Communication Unit

### IFAD staff

Carola Alvarez, Managing Director, Office of Development Effectiveness

Oscar Anaadumba, MEL specialist

Stella Atela, Director, People and Culture Division

Madina Bazarova, Director, Office of Audit and Oversight

Thomas Bousios, Director, Information, Communication and Technology Division

Nigel Brett, Director, Sustainable Production, Markets and Institutions Division

Donal Brown, Associate Vice President, Department for Country Operations

Han Ulac Demirag, Chief, ODE (Quality Assurance and SECAP Unit)

Chitra Deshpande, Results specialist, ODE

Federica Diamanti, Associate Vice-President, External Relations Department

Roger Fischer, Chief Risk Officer, Office of Enterprise Risk Management

Lenyara Fundukova, Senior Knowledge Management Specialist, ODE (Evidence for Development Impact and Innovation)

Ron Thomas Hartman, Director, Global Engagement, Partnerships and Resource Mobilization

Bernard Hien, Director, West and Central Africa Division

Carlos Manuel Icaza Lara, Specialist Policy and Results, ODE (Results, Resources and Systems Unit)

Alvaro Lario, IFAD President

Sara Mbago-Bhunu, Director, East and Southern Africa Division

Rocio Medina Bolivar, Director, Latin America and the Caribbean Division

Vibhuti Mendiratta, Senior Economist, ODE (Evidence for Development Impact and Innovation)

Juan Carlos Mendoza, Director, Environment, Climate, Gender and Social Inclusion Division

Diane Menville, Associate Vice-President, Financial Operations Department

Gerardine Mukeshimana, Vice-President

Robert Palmer, delegated by Jenny Scharrer, Director, Private Sector Operations Division

Tisorn Songsermsawas, Senior Economist, ODE (Evidence for Development Impact and Innovation)

Reehana Rifat Raza, Director, Asia and the Pacific Division

Naoufel Telahigue, Director, Near East, North Africa, Europe and Central Asia Division

Claudia Ten Have, Secretary of IFAD

Leon Williams, Senior Specialist (M&E), ODE (Results, Resources and Systems Unit)

Guoqi Wu, former Associate Vice President, Corporate Services Department

### **Executive Board Representatives**

Ms Lorena Alvarqado Quezada, Alternate Permanent Representative, Permanent Mission of Mexico to the United Nations Organizations for Food and Agriculture in Rome

Mr Jan Bade, Chargé d'Affaires a.i., Permanent Representation of the Kingdom of the Netherlands to the United Nations Organizations for Food and Agriculture in Rome

Mr Nicolas Cattaneo, Collaborateur diplomatique, Embassy of the Swiss Confederation

Mr James Catto, Director, Office of International Development Policy, Department of the Treasury of the United States of America

Mr Edoardo Ciammariconi, Communication Officer and Assistant to the Ambassador, Swiss Mission to the United Nations Agencies in Rome

Mr Miguel Garcia Winder, Former Mexico Executive Board Representative

Ms Marlene Heeb, Deputy Permanent Representative, Counsellor, Swiss Mission to the United Nations Agencies in Rome

Ms Aliisa Ketonen, Advisor, Embassy of the Republic of Finland

Ms Léa Kvaternik-Borderon, Alternate Permanent Representative, Permanent Representation of France to the United Nations Agencies in Rome

Mr Ronald Meyer, Former Germany Executive Board Representative

Ms Susann Nilsson, Executive Board Representative of Sweden to WFP and IFAD

Dr Yaya Olaniran, Minister, Permanent Representative of the Federal Republic of Nigeria to the United Nations Food and Agriculture Agencies in Rome and Evaluation Committee Chairperson

Mr Mohamed Sawy, Executive Board Representative of the Arab Republic of Egypt to the United Nations Food and Agriculture Organizations in Rome

Ms Divyadharshini SHANMUGAM, Director (OMI), Department of Economic Affairs, Ministry of Finance of the Republic of India

Ms Sui Li, Deputy Permanent Representative of People's Republic of China to the United Nations Agencies for Food and Agriculture in Rome

Ms Anja Wagner, Executive Board Representative of the Federal Republic of Germany to the International Organizations in Rome

### **Country-level Stakeholders**

Mr Daniel Chakwasha, International Cooperation, Ministry of Finance, Economic Development and Investment Promotion. Harare, Zimbabwe

Ms. Tu Duong, Administrative Director, RYNAN SMART AGRICULTURE JSC, Viet Nam

Dr Mohamed Yammai, Director General, Resource Mobilization, Resource Mobilization Department, Ministry of Finance and Planning, Male', Maldives

### **Other**

Ms Anne-Claire Luzot, Director, Office of Evaluation, World Food Programme

Mr Fabrizio Felloni, Deputy Director, Independent Evaluation Office, Global Environment Facility

Dr Srini Pillay, Chief Medical Officer and co-founder of Reulay

Dr Vinod Thomas, Senior Adviser for IOE

Rob D. van den Berg, member for IOE's External Advisory Panel 2022-2024

## Appendix 4: Surveys of staff and members of the Executive Board

The survey was launched on 26 November 2025 and disseminated by email to IFAD staff by the IOE team. It remained open for two and a half weeks and closed on 12 December 2025, which aligns with good practice for internal organizational surveys. The survey was designed to be short, with an average completion time of 10 minutes, balancing sufficient participation time with respondent engagement. However, the timing of the survey, towards the end of the calendar year, may have affected participation rates, particularly among senior staff and Executive Board members.

The survey was closed with 512 views. In total, the survey received 400 individual responses: 280 complete answers, as well as 125 incomplete. Among the respondents, eight respondents were members of the IFAD Executive Board, and the rest were regular IFAD staff. In sum, the **response rate for Board members was 22.2 per cent** (8 respondents out of 36), and **the response rate for IFAD staff was 43.5 per cent** (392 respondents out of 900 IFAD staff).<sup>17</sup> About 20 per cent of the respondents from IFAD staff were not familiar with the evaluation work conducted in the organization and thus were taken out of the survey. The **final sample for the rest of the survey is 272 IFAD staff and 8 members of the Executive Board.**

The survey successfully reached a diverse set of IFAD staff:

- Most **respondents are professional staff (54 per cent), followed by consultants (25 per cent)** and administrative staff (15 per cent). Only 3 per cent hold positions at the level of Deputy Director or above. Notably, three respondents identified as interns and two as United Nations volunteers.
- Most of the **respondents (54 per cent) work in the headquarters of IFAD in Rome**, and over a quarter (27 per cent) work in country offices. The next largest group at 12 per cent work in regional offices and, more marginally, 2 per cent (6 respondents) worked remotely, 2 per cent as consultants (5), 1 per cent in a multi-country office (4) and in a liaison office (3). Over a third (34 per cent) of respondents work in the Country Operations Division, followed by 13 per cent each in the Corporate Services Division and the Office of Technical Delivery. The Independent Office of Evaluation accounts for 9.7 per cent of the sample.
- Nearly **half of respondents (47 per cent) have worked at IFAD for one to five years**, while a quarter (25 per cent) has been with the organization for over a decade. Thirteen per cent have less than one year of service and 15 per cent have worked at IFAD for six to ten years.
- The survey sample is **48 per cent female and 46 per cent male**, with 5 per cent of respondents choosing not to disclose their gender.

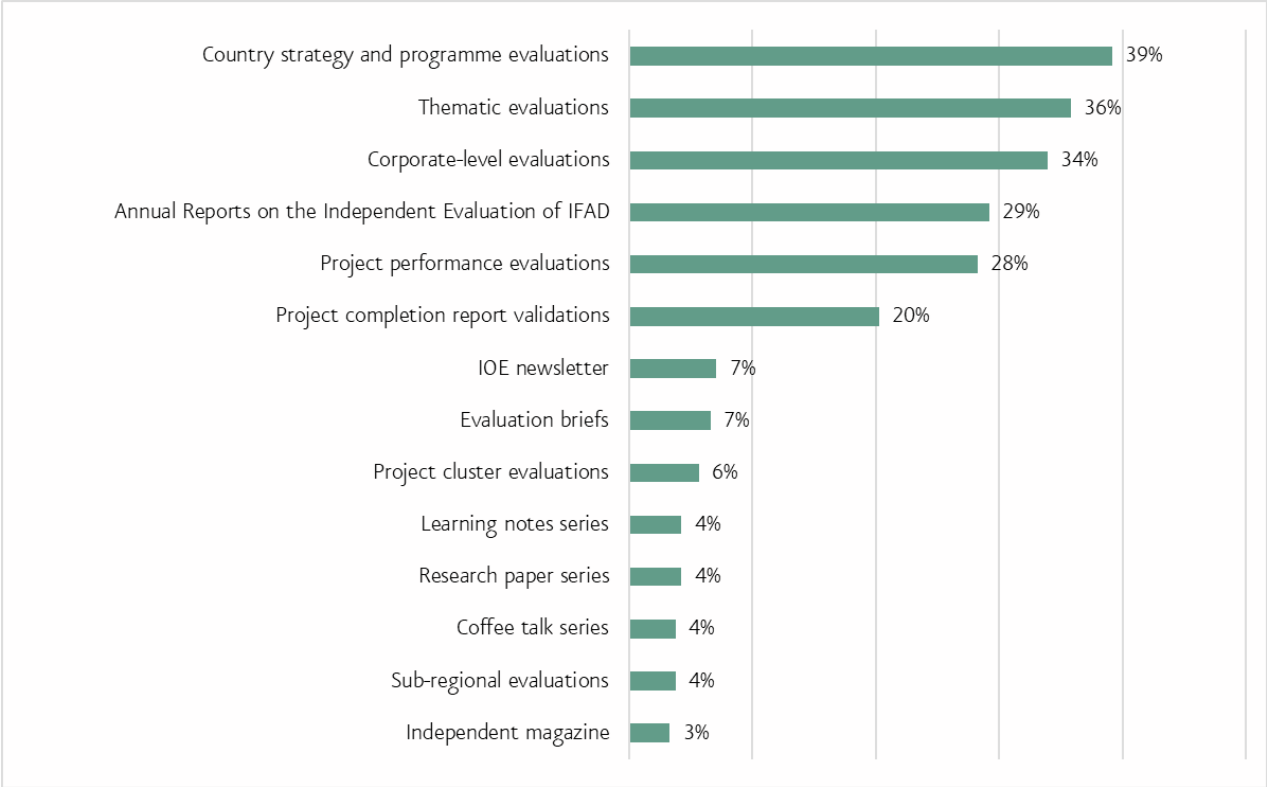
While the survey reached a diverse range of staff in terms of division, tenure and occupation, limited background data on the overall staff population prevented a full assessment of sample representativeness. Although the survey targeted permanent IFAD staff, responses to screening questions suggest that some non-permanent staff also participated. As a result, the findings should not be considered representative of, or generalized to, the broader population of consultants working with IFAD.

---

<sup>17</sup> The IFAD staff population figures were supplied by IFAD Human Resources and correspond to data as of October 2025. From our experience, these response rates fall within acceptable ranges (30-50 per cent) for voluntary organizational surveys conducted in multilateral institutions.



**Figure 2 : Which of the IOE products do you find the most useful for your work? (n = 212, IFAD staff)**



Overall, the **survey collected evidence on a very positive assessment of the IOE:**

- The independence of IOE is the most strongly endorsed quality of the evaluation function, with 51 per cent of respondents strongly agreeing. Assessments are more mixed with regard to the rigour of IOE approaches and methodologies, as well as IOE’s engagement with stakeholders. These areas received the highest levels of negative feedback (between 3 to 8 per cent), although they also recorded the largest share of “don’t know” responses.
- The survey also collected **highly positive feedback on the usefulness of IOE evaluation recommendations**. Most respondents (82 per cent) reported that the recommendations were useful, including 39 per cent who rated them as very useful. This is **consistent across occupation status and tenure period** at IFAD among all the staff, but it is particularly the case for administrative staff and the deputy directors and above.
- On the effectiveness of IOE’s results and communication, the results are somewhat positive with respectively 62 per cent and 55 per cent of respondents stating that it is effective or very effective. Less than a third of respondents said it was only a little effective.
- The assessment of the impact of IOE’s work is also positive, with between 65 per cent and 74 per cent of respondents providing positive ratings. However, it should be noted that a relatively high share of respondents, up to 20 per cent, were unable to assess these elements, while negative responses accounted for between 10 per cent and 15 per cent.

The negative feedback regarding IOE mainly relates to the significant workload required from staff during evaluations, the timing of evaluations not always aligning with operational needs and a perceived distance between operational teams and IOE’s independent role, as well as concerns that some recommendations are too general or theoretical and not sufficiently adapted to specific operational contexts.

*"In my previous role as CPM, after CSPE I tried to involve IOE colleagues at the early stage of a programme design but the participation level or the notion was IOE is "independent" and (we) therefore didn't get much value or input at early stage of the design but rather at a later stage to check on compliance which was late in my view."*

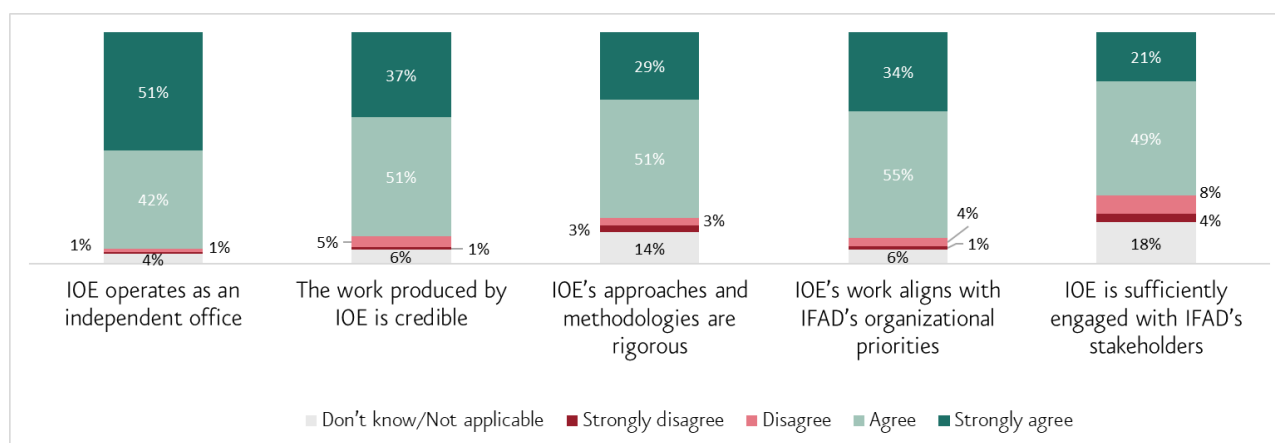
The recommendations call for broadening the scope of evaluations beyond IFAD operations, improving communication through shorter and more accessible outputs and easier access to report findings, strengthening dialogue between evaluation functions and Management, and further decentralizing evaluation functions to bring them closer to country and regional contexts. A selection of survey respondents' comments is illustrative.

*"IOE products are somewhat hard to find. They should have more presence in the intranet and more diagrams for easy digest of information."*

*"In my opinion, staff members and other stakeholders require assessments to be presented in two ways. There are times when in-depth rigorous research is required, of course, but I would also contend that there is a time when 'lighter' briefs, learning notes, and shorter reports are needed, especially when staff are busy. These give readers an understanding of a particular issue, keeping them updated and informed; if they need more information, they can always turn to more in-depth work."*

*"Given the organizational structure in terms of the independence of IOE, its role in facilitating lessons learning or uptake of evaluations particularly at the regional and country level is somewhat limited. This would work better if the organization had a decentralized evaluation function which has a dotted line management to IOE. This perhaps requires either partial decentralization of IOE in line with what other UN agencies have done or strengthening the self-evaluation function to enhance uptake."*

**Figure 3 : Based on your experience, how much do you agree with the following statements about these aspects of the IOE? (n = 213, IFAD staff)**

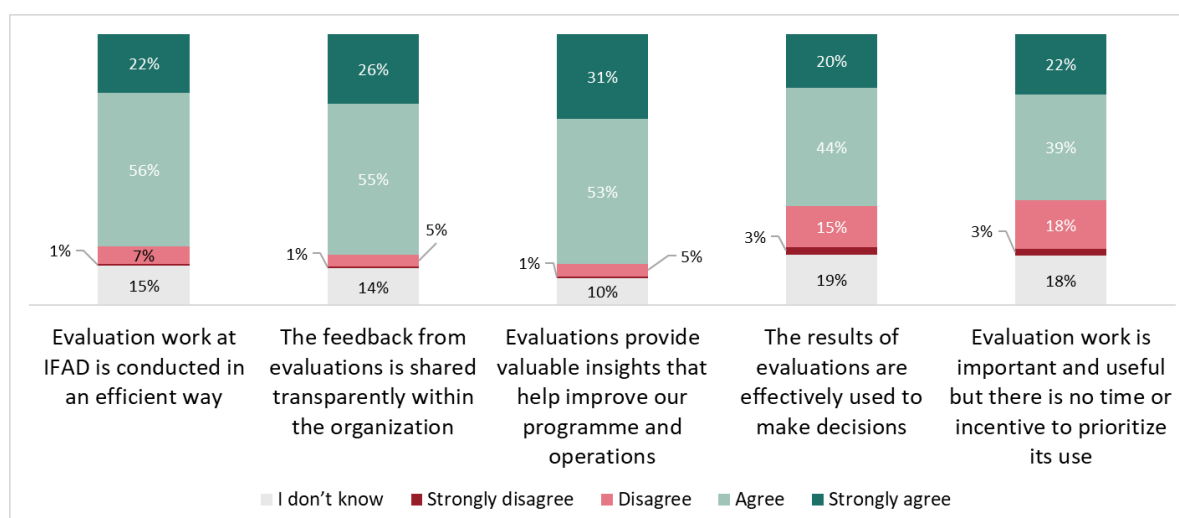


Concerning **Management-led self-evaluation**, less than one third of respondents (31 per cent) reported having been involved in a project completion report (PCR), mostly consultants and professional staff. However, for those concerned with PCRs, their reported use was high, with 39 per cent indicating they use them very often and 51 per cent sometimes; only 5 per cent reported never using them. PCRs are very positively assessed by respondents, with overwhelming majorities (88–97 per cent) agreeing that they support the assessment of achievements, outcomes and impacts, facilitate experience-sharing and lesson-learning, and contribute to improvements in future project design and implementation.

The last section of the survey was dedicated to a **broader assessment of the evaluation conducted** at IFAD:

- Overall, a large majority of respondents express positive views on evaluation at IFAD (with 75–84 per cent positive agreement across statements), particularly regarding the usefulness of evaluations for improving programmes and informing decisions, although a notable minority highlights constraints related to time and incentives for prioritizing their use.
- The effective use of evaluation results in decision-making is slightly less positively assessed with 18 per cent of the respondents disagreeing with this statement.
- The lack of time and incentive to prioritize the use of evaluation also shows a higher level of disagreement among respondents (21 per cent).
- Most respondents (90 per cent) consider IFAD to be a learning organization, including 37 per cent who strongly affirm this view. This perception is generally consistent across tenure and slightly more positive among staff with over six years of service, although administrative staff are somewhat less positive, with 15 per cent stating that this is not at all the case.

**Figure 4 : Based on your experience, how much do you agree with the following statements about evaluation work conducted at IFAD (including self-evaluation and independent evaluation)? (n = 197, IFAD staff)**



- The most frequently cited enablers of learning are collaboration and teamwork across projects and departments (71 per cent) and access to training, resources and development programmes (58 per cent), followed by feedback mechanisms (41 per cent), with additional factors including professional interest, internal communication and learning events, peer learning and IOE events and newsletters.
- The main barriers to learning identified by respondents are limited resources (71 per cent) and competing priorities (60 per cent), followed by ineffective knowledge management (42 per cent) and hierarchical rigidity (40 per cent), while fewer respondents cite unclear learning objectives (25 per cent), and additional constraints include challenges in headquarters and country office relations regarding decentralized knowledge management and potential tensions with political priorities.

## II. Main results from the Executive Board

As described above, the survey collected the complete responses of seven members of IFAD's Executive Board. Six of them are from List A countries, one of them is from List B.

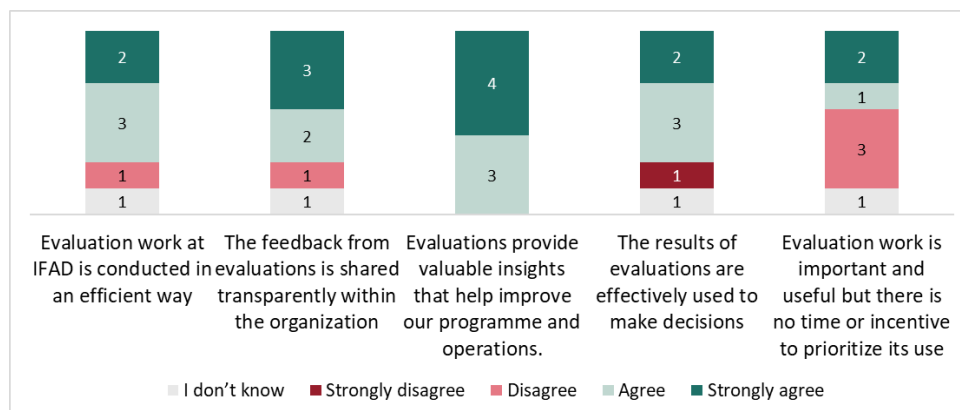
We had no answers from list C countries. This distribution suggests that IFAD’s evaluation function may face challenges in engaging Executive Board representatives from List B and List C countries. Almost all of the board member (86 per cent) has read some IOE products, but only about half of them (43 per cent) have used them to be informed about their work.

*“I mainly read evaluations. I found them to be very good, balanced and realistic: I like that they aim to facilitate learning and thus further improve IFAD's work.”*

There is broad agreement that IOE effectively contributes to improving the design and implementation of future projects, with most Executive Board members viewing IFAD as either fully or partially a learning organization. Executive Board members generally find IOE’s recommendations useful (57 per cent agree they are very useful) and recognize its independence, credibility, methodological rigour and alignment with IFAD’s priorities. Most also view IOE as effective in communicating results, sharing experiences and promoting lesson learning, though some note limited stakeholder engagement and familiarity with outreach to partner countries. One board member expressed some concern:

*“IOE talk a lot about themselves but less about the outcome or results of their work.”*

**Figure 5 : Board member assessment of IOE (n = 7, Executive Board)**



Leadership support, a learning-friendly culture, collaboration and access to training are seen as the main enablers of learning at IFAD, though feedback mechanisms and the knowledge management system are less recognized, with some caution about problem cover-ups and external pressures. Key barriers include competing priorities, limited time and resources, ineffective knowledge management, hierarchical rigidity, unclear learning goals, and, for some, pride or protection of financial resources.

*“There seems to be more cover-up if problems arise than interest in learning - this is a huge worry for us.”*

## Appendix 5: Quality review of IOE evaluation reports

This meta-evaluation is based on a review of IOE evaluations conducted between 2019 and 2025. The quality review was prepared in January 2026, and the report was drafted in early February 2025. The analysis and report were produced by Matthieu Sarnin, Consultant at Quadrant Conseil.

The objective of this review was to provide an overview of the methodologies and overall quality of a sample of IOE evaluation products, in order to inform the independent external review of IOE. While an in-depth review of individual evaluations was beyond the scope of this exercise, the review sought to identify recurring strengths and weaknesses across evaluations conducted by IFAD.

The review universe comprises 40 evaluation reports, including five corporate-level evaluations (CLEs), two thematic evaluations (TEs), two project cluster evaluations (PCEs), two subregional evaluations (SREs), and 29 country strategy and programme evaluations (CSPEs). From this universe, a sample of 20 evaluations was selected. The sample includes all corporate-level, thematic, project cluster, and subregional evaluations, as well as a purposive selection of nine CSPEs, randomly selected with the support of Copilot AI. The complete list of reviewed reports is presented in annex I.

The systematic quality review of IOE evaluations applied a simplified version of the meta-evaluation framework developed by the Independent Evaluation Group (IEG).<sup>18</sup> Evaluations were assessed against a set of indicators grouped into seven quality dimensions: scope, reliability, construct validity, internal validity, external validity, data analysis validity, consistency. For each of these attributes, we assess a series of indicators based on a grading scale (1) adequate, (2) partial, (3) inadequate, or inapplicable (NA). The key questions of the framework can be found in annex II. Based on the findings, the review highlights several key quality concerns and identifies areas requiring further attention.

### I. Limitations

Finally, several limitations affecting the quality review should be noted. First, contrary to the IEG meta-evaluation framework, the analysis relies exclusively on final evaluation reports published by IFAD<sup>19</sup> and comparison with methodology presented in approach papers was not possible. In addition, the IEG framework includes a second phase consisting of an in-depth quality review of a limited subsample of evaluations, complemented by interviews with evaluation staff to capture contextual information on evaluation design and implementation. This level of qualitative depth was beyond the scope of the present study. Given time constraints and the larger number of evaluations reviewed, a more systematic but less in-depth approach was applied to collect information on evaluation. Secondly, the scope and nature of the evaluations included in the review vary considerably. However, evaluations were treated as equivalent types of products during the data collection phase, meaning that the same items were reviewed for each evaluation, the diversity of evaluation types and their respective idiosyncrasies were considered during the analysis phase and should also be taken into account when interpreting the results presented in this report.

<sup>18</sup> For brevity, this report omits detailed definitions and rationales for each attribute in the framework. The full methodology is available in: World Bank. 2021. A Meta-Evaluation of Independent Evaluation Group Evaluations (Fiscal Years 2015–19). Independent Evaluation Group. Washington, DC: World Bank.

<sup>19</sup> The exercise did not cover dimensions of credibility and quality that could not be assessed based on the reports alone, such as consultations between evaluators and counterparts, the expertise and composition of evaluation teams, integrity in data collection, the nature of quality assurance processes, and the quality of peer review, where applicable.

Considering these limitations, this report should not be seen as a standalone product, but rather as a complementary input intended to inform the external evaluation review.

## **II. Results**

The assessment finds that overall, the quality of the IOE reports is very good and that in most cases, the evaluation reports clearly describe and report the key elements under review, as presented in the score summary in Appendix 5. Accordingly, the challenges highlighted below mainly reflect areas where further improvements could be made from an already strong baseline.

### **Assessment of the overall evaluation design**

The scope of evaluations varies substantially depending on the type of evaluation product. While some evaluations focus on a limited number of programmes (e.g. PCEs) or a single country (e.g. CSPEs), corporate and thematic evaluations tend to adopt a much broader and more ambitious scope. In addition, the timespan covered by the evaluations is often extensive, with half of the sample covering periods of more than ten years. As a result, the evaluands described in the sample are frequently highly complex, reflecting the multisite, multilevel and multi-actor nature of IFAD-supported interventions. Most evaluations in the sample explicitly stated a learning objective with only three which did not refer to learning or lesson drawing as a stated purpose. In addition, the evaluations can broadly be grouped according to how their objectives are framed: one group adopts a more causal perspective, while the other emphasizes performance assessment.<sup>20</sup> Finally, regarding evaluation questions, all evaluations except four formulated between two and five overarching questions to guide the evaluation process. None subsequently introduced an excessive number of sub-questions that would resemble a “bag-of-questions” approach to scoping. Most overarching questions were evaluative in nature. To a lesser extent, approximately half of the evaluations included exploratory questions, nine out of fourteen incorporated design-oriented questions, and only two raised more descriptive questions aimed at understanding programme contexts. Given the ambitious scope of many evaluations, this limited emphasis on contextual understanding is somewhat surprising, as it constrains deeper insight into the environments in which IFAD-supported interventions operate.

Despite this, most recommendations were logically grounded in the evidence presented throughout the evaluation reports. One minor caveat is that, although recommendations were generally evidence-based, there is a tendency to accumulate what could be described as “bags of recommendations.” Evaluations typically included four to seven overarching recommendations, but approximately half of the evaluations also specified up to 22 sub-recommendations. While balancing the specificity and breadth of recommendations is inherently challenging, the “bag of recommendations” approach risks diluting the key messages of the evaluation and potentially creating confusion about priorities and focus.

### **Reliability and quality of evidence issue**

All evaluation relied on a well-developed evaluation design matrix and a theory of change, which provides a clear rationale for the methods used to address specific evaluation questions. Findings were generally well described and logically linked to the underlying data and methods. However, one area of concern is that most evaluations do not provide sufficient clarity on how evidence from different sources was triangulated and used to substantiate findings relative to the initial evaluation questions. Several important cautions should be noted regarding the transparency and clarity of the analysis of evidence.

---

<sup>20</sup> This distinction appears largely driven by the type of evaluation, with more cross-sectoral evaluations (e.g. CLEs and TEs) focusing on IFAD's contribution to broader social outcomes, and more narrowly-scoped evaluations (e.g. PCEs and CSPEs) concentrating on project-level results in relation to predefined objectives, often linked to country strategies.

This relates to four elements reviewed: the description of data analysis methods, the explanation of how evaluative evidence was triangulated and synthesized, the articulation of quality assurance principles applied during coding and synthesis, and the depth of the discussion of limitations. Only two evaluations explicitly described how data were analyzed and how evidence was triangulated. Similarly, only two evaluations included elements related to analysis and triangulation within the evaluation design matrix itself. The evaluation titled Joint evaluation of collaboration among the United Nations Rome-Based Agencies (ID3CLE)<sup>21</sup> offers a good example of a clearly-formulated data analysis approach: “Data analysis was structured around the questions, sub-questions and indicators in the evaluation matrix. The evaluation team developed a standardized findings matrix template, structured by evaluation questions and sub-questions. Team members entered data from each of the country, deep dive and thematic studies into one of these matrices, and all matrices were then combined for easy analysis. Structuring the data in this way facilitated triangulation. Analysis of documentation and review of interview notes supplemented the combined findings matrices. The evaluation team integrated its understanding and analysis from all these sources through a series of workshop discussions. [...]” (ID3CLE). Conversely, the evaluation Knowledge Management Practices in IFAD Corporate-Level Evaluation (ID1CLE) provides a useful example of the “black box” approach to analysis and triangulation often found in the evaluations: “During the synthesis phase, the evaluation **team assessed the extent to which the evidence was adequate and could be triangulated through the various sources and methods**. Constraints in the availability and quality of data were, in particular, noted for evidence blocks 1 and 2 (see limitations).” (ID1CLE, own emphasis)

Second, a closely-related issue identified in the systematic review concerns the quality assurance principles applied during coding and synthesis. All evaluations but one included in the sample underwent an internal peer review process within IOE. Eight evaluations reported the involvement of external independent consultants in quality assurance and enhancement activities, and half explicitly referred to compliance with IOE evaluation standards or internal guidance. However, only a limited number of evaluations provided detailed information on the quality assurance principles applied throughout the evaluation process, the coding of data and the analysis.

Finally, a well-developed discussion of limitations is an important factor affecting the reliability of an evaluation. In this respect, most evaluations performed somewhat well, as all included a dedicated section or paragraph outlining key limitations and, in most cases, the mitigation approaches adopted. However, while limitations on data sources and methods are generally well documented, only three evaluations addressed limitations in the evaluation design and reflected on the evaluation process. Finally, the extent to which mitigation measures were discussed in response to identified limitations varies considerably across the evaluations analyzed. Some evaluations clearly explained how specific methods were used to address evidence gaps. This is the case, for example, in the country strategy and programme evaluation (CSPE) of the People’s Republic of China: “To ensure adequate coverage of evidence, qualitative interviews, an extensive literature review and an online stakeholder survey complemented this CSPE to the extent possible and allowed triangulation of quantitative and qualitative information.” (ID13CSPE). By contrast, other evaluations remained vague regarding the mitigation measures undertaken. The CSPE of the Republic of Türkiye illustrates this limitation: “In light of these limitations, the team triangulated using diverse sources of data and information, where accessible, before concluding.” (ID12CSPE).

---

<sup>21</sup> It should be noted that this evaluation is the only one in the sample conducted by an external private company, rather than internal evaluators. While the emphasis on transparency and clarity in source analysis is understandable in this context, internally-conducted evaluations should also adhere to similar standards.

A final concern regarding the reliability of evaluation results relates to the extent to which findings can be expected to apply in other settings. Overall, temporal validity was generally well addressed in the evaluation reports, with evaluations reflecting on the maturity of projects and the timing of data collection. In most cases, external validity was discussed in relation to the selection of primary data sources, such as portfolio analyses, country case studies and field visits, and was also considered when assessing secondary data. For example, in the Republic of Türkiye country strategy and programme evaluation (ID12CSPE): “In the AKADP, the samples used in the impact assessment survey had no links to those in the baseline survey. In the MRWDP, the Government of Türkiye impact assessment did not use the baseline data. The validity of its comparison of results between beneficiaries and control groups is questionable without an understanding of these two groups before the project”. However, discussions of external validity specifically regarding the generalizability of evaluation findings were relatively scarce and, in some cases, omitted entirely. Population validity was particularly weak in most reports, leading to partial ratings for several evaluations. Some evaluations, however, acknowledged limitations related to external validity in a clear and meaningful way. For example, the Sub-Regional Evaluation of Countries with Fragile Situations in IFAD-WCA, which examined IFAD’s engagement in the G5 Sahel countries and northern Nigeria, explicitly addressed generalization issues arising from limited data sources.

Overall, while the evaluations reviewed in the sample were generally robust with respect to data analysis validity, concerns related to data quality and the strategies used to mitigate potential biases arising from weaker data emerged as key areas for improvement under this attribute. The main issue raised by these findings is that, although evaluations collect large volumes of evidence from a wide range of data sources and methods, it is often difficult to assess the relative strengths and weaknesses of that evidence, which tends to be reported in a largely linear and non-hierarchical manner that can lack analytical clarity. A more explicitly articulated link between evidence, methods and the scope of inference would improve the interpretability and robustness of the findings. This could be done through a tighter articulation of the theory of change in relation to the empirical components of the study and could enhance transparency in how evidence is synthesized and weighted, thereby improving the overall rigour of the evaluation.

### **IFAD-centred approach**

All the evaluation reports included in the analysis used a theory-based approach, reconstructing a theory of change, sometimes situated within a broader conceptual framework and they often provided a compelling section to operationalize key concepts within the formulation of the theory or the conceptual framework used. However, only half of the evaluations mobilized and incorporated external research following structured review principles. Consequently, most evaluations tend to adopt an IFAD-centred perspective in defining and operationalizing key concepts and fall short in describing the broader policy context or the roles of external actors engaged alongside IFAD interventions. A notable exception is the corporate-level evaluation of Knowledge Management Practices in IFAD (ID1CLE), which illustrates strong use of external literature. In this case, the conceptual framework and nested theory of change draw on an extensive review of the six generations of knowledge management functions to explore “a set of hypotheses that cover causal linkages between KM strategies, tools, practices and key enabling factors, and how they interact to deliver KM results at corporate, regional division and country programme levels” (ID1CLE). The assessment of construct validity thus highlights that including external knowledge more systematically and applying structured literature review methods could broaden the evaluation perspective beyond IFAD-specific interventions.

### III. Conclusion

Overall, the evaluations reviewed demonstrate high quality and reflect a strong appreciation of methods and evidence, even in the context of ambitious evaluation scopes and complex evaluands. At the same time, several areas for improvement were identified. These include greater transparency regarding data quality and triangulation, more systematic attention to the environment in which IFAD operates, and enhanced reflection on the external validity of findings. Strengthening the use of the theory of change, particularly its articulation in relation to empirical analysis, would further improve analytical clarity. More broadly, adopting a more rational and formalized approach to evaluation design, moving beyond a collection of methods and tools toward a coherent, structured framework, would support a deeper understanding of evaluation quality. The integration of innovative approaches, such as contribution analysis, could also enhance assessments of coherence and causal inference. Expanding attention to these dimensions would improve the validity, reliability and overall utility of future evaluations, while maintaining the high standards of methodological rigour already evident across the sample analyzed.

#### Annex I : List of evaluations in the review sample

ID	Title
ID1CLE	Corporate-level evaluation of knowledge management practices in IFAD. December 2024
ID2CLE	Corporate-level evaluation of IFAD's decentralization experience 2023. November 2023
ID3CLE	Joint evaluation on the collaboration among the United Nations Rome-based agencies. October 2021
ID4CLE	Corporate-level evaluation on IFAD's support to innovations for inclusive and sustainable smallholder agriculture. September 2020
ID5CLE	Corporate-level evaluation on IFAD's engagement in pro-poor value chain development. September 2019
ID6TE	Thematic evaluation of IFAD's support for smallholder farmers' adaptation to climate change. May 2023
ID7TE	Thematic evaluation of IFAD's support to gender equality and empowerment. March 2025
ID8PCE	Project cluster evaluation on rural finance in the East and Southern Africa Region. December 2023
ID9PCE	Project cluster evaluation on rural enterprise development. August 2023
ID10SRE	Sub-regional evaluation of IFAD's experience in the Dry Corridor of Central America. August 2024
ID11SRE	Sub-regional evaluation of countries with fragile situations in IFAD-WCA. Learning from experiences of IFAD's engagement in the G5 Sahel countries and northern Nigeria. March 2023
ID12CSPE	Türkiye country strategy and programme evaluation. 23 April 2025
ID13CSPE	China country strategy and programme evaluation. 21 February 2024
ID14CSPE	Colombia country strategy and programme evaluation. 7 February 2024
ID15CSPE	Malawi country strategy and programme evaluation. 24 January 2023
ID16CSPE	Islamic Republic of Pakistan country strategy and programme evaluation. 4 March 2022
ID17CSPE	Uganda country strategy and programme evaluation. 7 July 2021
ID18CSPE	Ecuador country strategy and programme evaluation. 3 March 2021
ID19CSPE	Madagascar country strategy and programme evaluation. 1 December 2020
ID20CSPE	Tunisia country strategy and programme evaluation. 8 March 2019

**Annex II : Quality review assessment**

<b>Attribute</b>	<b>Review</b>
<b>Scope</b>	<ul style="list-style-type: none"> <li>- Are the context and rationale of the evaluation adequately described? (Graded)</li> <li>- Are the evaluation goals adequately formulated? (Graded)</li> <li>- Are the evaluation questions adequately formulated (also in relation to each other)? (Graded)</li> <li>- Are the evaluation questions adequately linked to the evaluation goals? (Graded)</li> <li>- Has the scope and delimitation of the evaluation been adequately described? (Graded)</li> </ul>
<b>Reliability</b>	<ul style="list-style-type: none"> <li>- Is the methodology of the evaluation adequately described? (Graded)</li> <li>- Is there a theory of change or theory of action, or conceptual framework? (Y/N question)</li> <li>- Is there identification and analysis of an operations portfolio? (Y/N question)</li> <li>- Are the quality assurance principles in coding and synthesis described? (Y/N question)</li> <li>- Are the sampling and selection considerations described? (Y/N question)</li> <li>- Is there a description of data collection methods and sources of data? (Y/N question)</li> <li>- Is there a description of data analysis methods? (Y/N question)</li> <li>- Is there a triangulation and synthesis of findings, including how (different) findings from different methods or designs have been integrated to reach (general) conclusions? (Y/N question)</li> <li>- Are the limitations of the evaluation adequately described (resulting from limitations in scope, methods or data and validity of findings)? (Y/N question)</li> </ul>
<b>Construct validity</b>	<ul style="list-style-type: none"> <li>- Has the evaluation clearly defined and operationalized key concepts into measurable attributes, and are these concepts adequately interrelated within a coherent theoretical framework (theories of action, change and conceptual models)? (Graded)</li> <li>- Has the evaluation effectively utilized and integrated external literature following structured review principles, and if applicable, has it successfully implemented a theory-driven approach with a clear articulation of mechanisms, contexts and outcomes? (Graded)</li> </ul>
<b>Internal validity</b>	<ul style="list-style-type: none"> <li>- How does the evaluation address causality and attribution? (If there are no causality questions, this does not apply. If there are causality questions, more information is needed to grade the dimension) (Graded)</li> <li>- In the presence of causal questions: <ul style="list-style-type: none"> <li>- Is the identification strategy based on theory or counterfactual analysis, or is there an evident triangulation of different sources? (in this particular question, you must extract the identification strategy defined in the evaluation).</li> </ul> </li> </ul>
<b>External validity</b>	<ul style="list-style-type: none"> <li>- Has the report paid adequate attention to population validity or generalizability of findings (the ability to generalize the study results to individuals or target groups, organizations or regions not included in the study)? (Graded)</li> </ul>
<b>Data analysis validity</b>	<ul style="list-style-type: none"> <li>- Has the evaluation addressed risks of bias resulting from data weaknesses or demonstrated any signs of incorrect methodological application in the qualitative or quantitative analyses? (Graded)</li> </ul>

<b>Consistency</b>	<ul style="list-style-type: none"> <li>- Are the methods and data sources logically linked to the evaluation questions? (Graded)</li> <li>- Have the methods that are reported as being applied indeed been applied? (Graded)</li> <li>- Do the findings logically relate to the underlying data and methods used? (Graded)</li> <li>- Do the findings respond to the original evaluation questions? (Graded)</li> <li>- Do the recommendations logically flow from the findings? (Graded)</li> </ul>
--------------------	---

code name	Scope	Reliability	Construct validity	Internal validity	External validity	Data analysis validity	Consistency
ID1CLE	Adequate	Partial	Adequate	Adequate	Adequate	Adequate	Adequate
ID2CLE	Adequate	Partial	Partial	Partial	Adequate	Partial	Adequate
ID3CLE	Adequate	Partial	Partial	Adequate	Adequate	Adequate	Adequate
ID4CLE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID5CLE	Adequate	Adequate	Adequate	Partial	Partial	Adequate	Adequate
ID6TE	Adequate	Adequate	Adequate	Adequate	Partial	Partial	Adequate
ID7TE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID8PCE	Partial	Adequate	Partial	Adequate	Adequate	Adequate	Adequate
ID9PCE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID10SRE	Adequate	Adequate	Partial	Adequate	Partial	Adequate	Adequate
ID11SRE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID12CSPE	Adequate	Partial	Adequate	Adequate	Partial	Adequate	Adequate
ID13CSPE	Adequate	Adequate	Partial	Partial	Adequate	Partial	Partial
ID14CSPE	Adequate	Partial	Partial	Adequate	Partial	Adequate	Adequate
ID15CSPE	Adequate	Adequate	Adequate	Partial	Partial	Adequate	Adequate
ID16CSPE	Partial	Adequate	Adequate	Partial	Adequate	Adequate	Adequate
ID17CSPE	Partial	Partial	Partial	Adequate	Partial	Adequate	Partial
ID18CSPE	Adequate	Adequate	Partial	Partial	Adequate	Adequate	Adequate
ID19CSPE	Partial	Adequate	Partial	Adequate	Adequate	Partial	Adequate
ID20CSPE	Partial	Adequate	Partial	Adequate	Adequate	Adequate	Adequate