

Thirteenth Steering Committee
18 February 2026

Terms of Reference for SOFF Independent External Evaluation

Decision 13.4

Systematic Observations
Financing Facility

**Weather
and climate
data for
resilience**



Decision 13.4: Terms of Reference for SOFF Independent External Evaluation

The SOFF Steering Committee:

Recalls

- [Decision 6.6](#), which included the evaluation in the SOFF work programme and budget for the first implementation period (July 2022–June 2025);
- [Decision 7.2](#), which endorsed the outcomes of the independent external review and management response prepared by SOFF Secretariat.
- [Decision 11.3](#), which further updated the evaluation timeline to the 16th Steering Committee (February 2027) by when the final evaluation results are to be presented.

Welcomes the approach to ensure an inclusive and transparent evaluation process through the establishment of a multi-stakeholder “Reference Group”, which will guide the work of the evaluator(s) and review key evaluation deliverables.

Endorses the Terms of Reference for the fund-level SOFF Independent External Evaluation.

Requests

- The SOFF Secretariat
 - To convene and coordinate the “Reference Group” and to facilitate the evaluation process in accordance with the evaluation Terms of Reference.
 - To include in the “Reference Group” an organization representing the civil society, such as the Global Network of Civil Society Organisations for Disaster Reduction (GNDR), member of the SOFF Advisory Board.
 - To manage the administration of the Independent External Evaluation. This includes facilitating the procurement process conducted by WMO in consultation with the “Reference Group”, as well as provisioning of logistical support to the evaluation team, while safeguarding the independence of the evaluation.
 - To facilitate the presentation of the final findings at the 16th Steering Committee meeting.
- The UNMPTF Office to disburse USD 100,000 to WMO to contract the evaluator.

Purpose of the document

With [Decision 11.3](#), the SOFF Steering Committee decided to conduct an independent external evaluation with the following timeline: (i) February 2026 (13SC): Steering Committee considers and approves the scope, approach, Terms of Reference, and budget; (ii) February 2027 (16SC): Final evaluation results are presented for consideration.

The present document responds to this decision. It provides the background and rationale for the evaluation, outlines evaluation purpose, approach, scope, and management of the evaluation including budget.

Table of contents

1. Background and Rationale	5
2. Evaluation Purpose and Approach.....	6
3. Evaluation Scope	7
3.1. Evidence Period	7
3.2. Dimensions, Criteria and Methods	8
3.3. Scope Boundaries and Exclusions	12
3.4. Data Collection.....	12
4. Management of the Evaluation.....	13
4.1. Evaluation Reference Group	13
4.2. Key Deliverables and Timeline.....	14
4.3. Budget	15
Annex	16

Terms of Reference for SOFF Independent External Evaluation

1. Background and Rationale

The Systematic Observations Financing Facility (SOFF) was established in 2021 at the request of the 193 countries and territories of the World Meteorological Congress to address critical gaps in weather and climate observations and to support sustained compliance with the Global Basic Observing Network (GBON). SOFF operates as a United Nations Multi-Partner Trust Fund (UN MPTF), that pools contributions from SOFF funding partners, with the UN MPTF Office serving as a SOFF Trustee. With over an initial ten-year horizon, SOFF is structured across a Start-up Period (January–June 2022), a First Implementation Period (July 2022–June 2025, subsequently extended to June 2027), and an Expansion and Sustaining Period (until December 2031). Its business model combines long-term, grant-only financing with peer-to-peer technical assistance delivered through interlinked Readiness, Investment, and Compliance phases. Following Decision 6.6, which included the evaluation in the SOFF work programme and budget for the first implementation period (July 2022–June 2025), [Decision 11.3](#) further updated the evaluation timeline with final evaluation results to be presented for consideration of the 16th Steering Committee meeting. SOFF initially focuses its Readiness, Investment and Compliance support on Small Island Developing States (SIDS) and Least Developed Countries (LDCs). However, WMO governing bodies and partners have progressively recognized the need to consider a wider possible scope for SOFF in order to ensure the effectiveness, sustainability and global integrity of the Global Basic Observing Network.

In 2021, the World Meteorological Congress, through Resolution 2 (Cg-Ext (2021)), established the Global Basic Observing Network (GBON), including surface marine observations as part of the required global network, and through Resolution 3 requested WMO participation in SOFF as Technical Authority to support GBON implementation. Building on this mandate, in 2024 the WMO Commission for Observation, Infrastructure and Information Systems (INFCOM) approved guidance for surface marine GBON stations in Exclusive Economic Zones and urged Members to consider contributing to SOFF, while requesting that SOFF consider expanding its support to this domain. This direction was elevated by the WMO Executive Council through Resolution 14 (EC-78, 2024), which urged increased financial contributions to SOFF and invited consideration of expanding its scope, subject to resource availability.

In 2023, the World Meteorological Congress (Resolution 21, Cg-19) requested the SOFF Steering Committee to explore opportunities to provide SOFF financial and technical support to Middle-Income Countries in need, while continuing to prioritize LDCs and SIDS. This guidance has been reinforced by multiple WMO Regional Association decisions, notably from Africa, the Pacific, the Caribbean and the Americas, which have called for regional SOFF programmes, expanded support to Middle-Income Countries, and stepped-up resource

mobilization to close persistent GBON gaps (see also [Decision 9.3](#) for more details). The direction was also echoed by conclusions of the [UNFCCC Subsidiary Body for Scientific and Technological Advice \(SBSTA\) in 2025](#) inviting the Facility to consider extending its provision of support for systematic observation to more developing countries.

SOFF's design and early trajectory have been underpinned by independent analytical work. [An Independent External Review](#) undertaken in the second half of 2023 assessed SOFF's design and early implementation, concluding that the facility represents a viable and scalable approach to sustainably upgrading, operating, and maintaining observation systems, with strong early performance during the Readiness Phase. In December 2023, [a study by the Scaling Community of Practice](#) found that SOFF has embedded scaling into its institutional design and aligns closely with established scaling principles. The European Centre for Medium Range Weather Forecasts (ECMWF) in collaboration with the World Meteorological Organization (WMO) undertook SOFF Impact Studies, including quantification of how additional Global Basic Observing Network (GBON) observations reduce uncertainty in short-range weather forecasts. The Steering Committee with [Decision 11.2](#) welcomed the results that scientifically demonstrated that targeted investments in GBON infrastructure in under-observed regions dramatically improve forecast accuracy, both locally and globally.

The Independent External Review applied selected [DAC evaluation criteria](#) to SOFF's design and early implementation but was necessarily limited to early evidence, while the subsequent Scaling Community of Practice study was design-focused and non-evaluative. The management response to the Independent External Review was taken into account and endorsed by the 7th SOFF Steering Committee ([Decision 7.2](#)).

As SOFF prepares for decisions on scaling its activities during the Expansion and Sustaining Period, building on learning from the First Implementation Period and requiring increased resource mobilization, an independent evaluation is essential.

2. Evaluation Purpose and Approach

The purpose of the evaluation is to assess SOFF's performance to date and inform decision-making on whether, how, and under what conditions SOFF should scale its operations during the Expansion and Sustaining Period.

It will also aim to ensure accountability, support learning and provide recommendations for further development of the program work.

The assessment will examine three key dimensions: SOFF's value proposition, business model, and operating model, as defined in Table 1.

For the purpose of this evaluation, **scaling** is defined along two dimensions, reflecting recent intergovernmental recommendations, as follows:

- **Increasing the number of SOFF-supported countries**, including the potential expansion of financial support to middle-income countries, as recommended by WMO Congress Resolution 21 (21 (Cg-19, 2023)) and echoed at UNFCCC COP30 by the [conclusions of the Subsidiary Body for Scientific and Technological Advice](#) (SBSTA 63, November 2025).
- **Expanding the scope of SOFF support** to include the remaining elements of the Global Basic Observing Network (GBON), namely marine observations, as recommended by WMO Congress Resolution 2 (Cg-Ext 2021) and further by WMO Executive Council (EC-78, 2024, [Resolution 4.1.2\(2\)/1](#)) and Regional Association IV¹ in 2025 (Decision 5.2(1)/1 (RA IV-19(I)).

This is a exclusively a fund-level evaluation, carried out in accordance with [OECD-DAC criteria](#), [United Nations Evaluation Group \(UNEG\) Norms and Standards](#) and UN MPTF evaluation guidance. The SOFF Gender Action Plan (Decision 3.3) should be used as a reference framework for any gender sensitivity component or question addressed through the evaluation.

It will be conducted by an external evaluation team of individual contractors or a company, selected through an open and competitive process to ensure independence, transparency, and credibility. Proposals will be assessed against technical and financial criteria, including the evaluator's relevant experience, independence, team qualifications, familiarity with climate funds, and cost-effectiveness.

3. Evaluation Scope

3.1. Evidence Period

As part of the SOFF business model, the Readiness and Investment phases will be assessed based on observed performance and other evidence available from the inception of the UN fund in November 2021 up to 30 June 2026. This period allows for a retrospective assessment of progress to date.

The planned Compliance Phase should be considered using a theory-based approach. This means that, given the longer-term nature of the Compliance Phase, the evaluation will examine whether the underlying assumptions, institutional arrangements, and implementation pathways are credible and likely to deliver the intended results based on

¹ RA IV (North America, Central America and the Caribbean)

the Draft Compliance Phase Framework, which is expected to be finalized at the 14th Steering Committee in May 2026.

3.2. Dimensions, Criteria and Methods

Table 1 below presents a summary of the evaluation dimensions, criteria applied, and illustrative assessment methods applied.

Table 1: Evaluation Dimensions, Applied Criteria, and Illustrative Assessment Methods

Evaluation purpose			
Assess SOFF's performance to date and inform decision-making on whether, how, and under what conditions SOFF should scale its operations during the Expansion and Sustaining Period.			
Evaluation Dimensions	What This Dimension Entails	OECD DAC Criteria Applied ²	Examples of potential Assessment Methods
1. Value Proposition	Examines why SOFF exists and whether it provides a clear and distinctive value in addressing gaps in the Global Basic Observing Network (GBON), including its comparative advantage in addressing the global data gap as a foundational global public good and strategic fit within the broader climate and early warning financing landscape. It assesses SOFF's performance to date in delivering value for money and examines whether its value proposition is likely to remain compelling and relevant as SOFF scales, including in relation to future demand for SOFF services	Relevance Coherence Efficiency Effectiveness Impact Sustainability	Past performance analysis; benchmarking against comparable funds and financing mechanisms; counterfactual and comparative analysis; stakeholder value-perception analysis; assessment of future relevance under different scaling and policy scenarios
2. Business Model	Examines whether SOFF's Readiness–Investment–Compliance phased model, and the intended		Past performance analysis; pattern and variance

² See OECD-DAC Criteria definitions: <https://www.oecd.org/en/topics/sub-issues/development-co-operation-evaluation-and-effectiveness/evaluation-criteria.html>

	<p>roles of countries, Implementing Entities, WMO Technical Authority, and peer advisors within that logic, have functioned as intended to date.</p> <p>It also assesses whether this business model is likely to remain appropriate, efficient, and effective as the number and diversity of participating countries increase, taking into account operational complexity and progression through phases.</p>	<p>analysis across country cohorts; stress-testing of the phased model (“what must be true” for effective scaling); scenario analysis of increased participation and operational complexity; comparative analysis with similar phased financing models</p>
3. Operating Model	<p>Examines whether SOFF’s Secretariat capacity and governance interfaces—including interactions with the Steering Committee, Advisory Board, co-founders, and the UN MPTF Administrative Agent—have enabled efficient delivery to date and are fit for purpose as SOFF scales. This includes specific attention to resource mobilization capacity; the predictability and sustainability of funding; the comparative efficiency gains of pooled funds versus alternative non-pooled mechanisms, and value added of UN MPTF administrative and disbursement processes; cost structures and potential economies of scale; and the Secretariat’s ability to adapt operational and coordination arrangements under increased scale and complexity.</p>	<p>Capacity and process stress-testing under scaling scenarios; analysis of resource mobilization pathways and funding sustainability; assessment of administrative efficiency and transaction costs; cost and scale analysis; learning and adaptation analysis; benchmarking</p>

To operationalize the application of the evaluation dimensions and DAC/UNEG criteria, Table 2 presents indicative evaluation questions. These questions are intended to guide analysis and evidence collection in a backward- and forward-looking manner. The questions are illustrative and do not constitute a prescriptive reporting structure.

Table 2: Matrix of Indicative Evaluation Questions by Evaluation Dimensions and Analytical Criteria

DAC Criteria Applied	Evaluation Dimensions		
	Value Proposition	Business Model	Operating Model
Relevance	To what extent does SOFF address identified GBON gaps, now and looking ahead?	To what extent is the Readiness–Investment–Compliance phased model appropriate for achieving sustained GBON compliance across different country contexts?	To what extent are Secretariat capacity and operational systems appropriate to current needs and anticipated scaling requirements?
Coherence	How well does SOFF's value proposition align with and complement other climate and early warning financing initiatives, including as stated in the Collaboration framework signed with other climate funds?	How coherent and complementary are the roles of countries, Implementing Entities, WMO Technical Authority and Peer Advisors within the three SOFF phases?	How well aligned and coordinated are the Secretariat, WMO Technical Authority, implementing entities, and Peer Advisors in supporting delivery at scale?
Efficiency	To what extent does SOFF deliver value for money ³ relative to plausible alternatives, and how might this be	To what extent does the phased model enable timely and efficient progression across readiness,	To what extent do existing processes, systems, and workflows as structured for a pooled funding

³ In the context of this evaluation, and within the Value Proposition dimension, value for money refers to the extent to which SOFF offers a compelling and efficient approach to delivering sustained, GBON-compliant systematic observations as a foundational global public good at the lowest feasible total system cost over time, while maximizing national and global forecasting and early warning benefits, compared to credible alternative financing and delivery approaches. These alternatives include project-based funding from other climate and development funds, satellite-centric observation strategies, and increased downstream or analytical investments, including modeling and AI-enabled forecasting approaches.

	impacted by an increased scale?	investment, and compliance as operations expand?	mechanism (including multi-stakeholder governance, Administrative Agent and pass-through functions, and fund allocation and disbursement modalities), enable timely, predictable, and cost-effective implementation as operations expand?
Effectiveness	To what extent is SOFF strengthening systematic observations, and how likely is it to continue doing so at scale?	To what extent has the business model functioned as intended and how likely is it to do so at greater scale?	To what extent has the operating model enabled effective implementation to date, and how likely is it to do so at scale?
Impact	What evidence shows that SOFF contributions will be sustained at national level and contribute to global benefits, and how might these impacts evolve with expansion?	How does the business model contribute to longer-term institutional strengthening of national Met offices capacity for sustained data collection and sharing, and how might these impacts vary with scale?	How does the operating model support learning, adaptation, and performance management in ways that influence results over time?
Sustainability	To what extent is SOFF's value proposition likely to remain valid and compelling over time, including under scenarios of increased scale?	To what extent is the business model resilient to increased scale and diversity, and what adaptations may be required to sustain results over time?	To what extent is the operating model institutionally and operationally sustainable, including required Secretariat and resource mobilization capacity?

3.3. Scope Boundaries and Exclusions

To ensure focus, proportionality, and stage of SOFF maturity, the evaluation will not cover the following areas:

- **Project-level or country-by-country performance assessments**, including detailed verification of outputs or attribution of results to individual country operations, recognizing that this assessment is conducted at the fund level rather than at the individual project level.
- **Attribution of long-term or downstream impacts**, including full assessment of socio-economic, climate resilience, or development impacts that extend beyond the current evidence window and SOFF's system-enabling role.
- **Re-assessment or validation of SOFF's Theory of Change**, including its fundamental assumptions or results logic. The evaluation will use the Theory of Change as a reference framework for assessing performance and readiness to scale. It may examine how it is operationalized in practice.
- **Re-assessment of SOFF's fundamental mandate or purpose**, which has been established by 193 countries and territories of the World Meteorological Congress and endorsed by the SOFF Steering Committee.
- **Standalone performance assessments, ratings, rankings, or audits of Implementing Entities, Peer Advisors, or technical partners**, except insofar as their roles and incentives affect SOFF's operating model and performance.
- **Financial, fiduciary, or compliance audits**, which are addressed through separate UN MPTF and implementing partner assurance mechanisms.
- **Broader assessments** of national meteorological or climate systems not directly linked to SOFF-supported interventions.
- **Technical verification of individual GBON stations or datasets**, except where analytically necessary to test business-model assumptions at a system level.

3.4. Data Collection

Data collection techniques are the tools used to gather evidence. For the SOFF evaluation, these will include desk (document) review; key informant interviews; surveys; validation workshops or focus groups; analysis of internal portfolio, administrative, and financial data (including budgets, expenditures, and cost structures); and comparative document review to support benchmarking.

Three to six country case studies will be used selectively, while ensuring geographic balance to test key assumptions of the business and operating models.

Field visits are not foreseen as part of the evaluation, in light of the fund-level focus of the evaluation, the nature of the evaluation questions, the opportunity to conduct interviews remotely, and the need to ensure proportional and efficient use of evaluation resources, in line with practice in comparable fund-level evaluations.

4. Management of the Evaluation

4.1. Evaluation Reference Group

Following the positive experience of other UNMPTF funds such as the [Central African Forest Initiative \(CAFI\)](#), and [Joint SDG fund](#), and to ensure an inclusive and transparent evaluation process, a multi-stakeholder “Reference Group” will be established to guide the work of the evaluator(s).

The “Reference Group” will be composed of representatives from each of the following constituencies: SOFF funders, Implementing Entities, SOFF co-founders, peer advisors, beneficiary countries.

The SOFF Secretariat will facilitate the discussions of the evaluator(s) with a multi-stakeholder representative “Reference group”.

This Reference group will:

- Be consulted during the recruitment process of the evaluator(s)
- Discuss drafts and endorse the inception report
- Discuss drafts and endorse the preliminary findings report
- Be consulted in the preparation of the management response
- Endorse the final evaluation report before it is submitted to the SOFF Steering Committee for adoption.

The proposed composition of the Reference Group⁴ is outlined in Table 3.

⁴ The participating organizations still to be confirmed.

Table 3: Proposed composition of the SOFF Evaluation Reference Group

Role	Organization/Country
Funder representatives	e.g. current and previous SOFF Co-chairs
Co-Founders representative	e.g. WMO
Implementing Entity representative	e.g. WFP
Peer Advisor representative	TBC (elected)
LDC representative	e.g. Rwanda (most advanced SOFF operations)
SIDS representative	e.g. AOSIS
Independent financial expert	TBC
Facilitator	SOFF Secretariat

4.2. Key Deliverables and Timeline

Deliverables will include:

- **Inception Report – May 2026**
Evaluation framework, methodology, workplan, and stakeholder mapping.
- **Data collection and analysis – June - October 2026**
Desk review, interviews, portfolio and financial analysis.
- **Preliminary Findings Brief / Presentation – October 2026**
Synthesis of emerging findings discussed with the Reference Group to sense-check interpretations.
- **Draft Evaluation Report – November 2026**
Full analytical report incorporating feedback from the preliminary findings' discussion.
- **Final Evaluation Report – January 2027**
Revised report reflecting written comments.
- **Executive Summary and decision/action-oriented synthesis – January 2027**
Finalized alongside the report for Steering Committee use.
- **Presentation to the 16th Steering Committee – February 2027**

The Executive Summary will present a coherent narrative highlighting SOFF's value, achievements, strengths, critical conditions for success, and future potential, with a clear focus on implications for scaling in the Expansion and Sustaining Period.

Recommendations will focus on enablers and conditions for success, rather than as corrective actions alone, and will be grouped, as appropriate, into:

- Priorities requiring immediate adjustment
- Focus areas to be addressed prior to scaling
- Areas to monitor and optimize over time

The evaluation will be subject to appropriate quality assurance processes consistent with [OECD-DAC and UNEG standards.](#)

4.3. Budget

The total budget approved by the 11th Steering Committee for the external evaluation amounts to USD 100,000 ([Decision 11.3](#), Table 2). The contract shall be awarded on a fixed-price basis and shall be directly linked to the achievement and acceptance of the deliverables specified in Section 4.2. The approved budget shall cover all costs associated with the assignment, including professional fees and any other applicable charges necessary for the successful completion of the services.

Annex

The annex below provides an overview of common practices applied across selected multi-partner trust funds and financing facilities in conducting external evaluations. While each fund operates within its own institutional context and governance structure, several shared approaches emerge.

Table 4: Comparative Practices of External Evaluations in Selected Funds

	<u>Climate Risk and Early Warning Systems (CREWS) Initiative</u>	<u>Central African Forest Initiative (CAFI)</u>	<u>UN-IFI Partnership Facility / UN Peacebuilding Fund</u>	<u>Joint SDG Fund</u>
Fund type and size⁵	World Bank-administered Financial Intermediary Fund ~USD 134 million ⁶	UN MPTF fund ~USD 1 billion ⁷	UN MPTF fund ~USD 1.5 billion ⁸	UN MPTF fund ~USD 491 million ⁹
Primary purpose of the evaluation	CREWS external evaluation to inform CREWS Vision 2025 and future project and portfolio planning, while the external review will inform the 2030 Strategy.	Assess CAFI's performance to inform future strategic decisions	Shape the next strategic phase and governance adjustments	Provide actionable, evidence-based recommendations to maximize the Fund's catalytic impact on the SDGs and UNDS reform

⁵ Cumulative commitments since inception

⁶ Source: <https://crews-initiative.org/>

⁷ Source : UNMPTF web site: CAFI (<https://mptf.undp.org/fund/afi00>),

⁸ Source : UNMPTF web site: UN-IFI (<https://mptf.undp.org/fund/pb000>)

⁹ Source : UNMPTF web site <https://mptf.undp.org/fund/ips00>

Evaluation Approach	Use OECD DAC and mixed methods like desk reviews and interviews.	Use OECD DAC and mixed methods like document reviews, interviews, and governance scoring.	Evaluations incorporate case studies, interviews, and contribution analysis.	Use of OECD/DAC Evaluation Criteria for Global and country level methodology
Evaluator Type	Lead external evaluator and additional experts – individual contractors.	External consultants (individual contractors) hired via UNDP.	External consultants (individual contractors) commissioned by the Steering Committee.	Evaluation team consists of two consultants (individual contractors) - one team leader and one technical expert).
Hiring Process	Secretariat-led selection Recruitment followed WMO procedures; SC informed.	Secretariat-led (UNDP-hosted) with donor Reference Group oversight.	UN Peacebuilding Support Office (PBSO) with UNOPS support.	UNDP procurement
Field missions	No	Limited use	Limited use	Limited use