

Evaluation Synthesis Report

2018-2019

Evaluation Office

UNITED NATIONS ENVIRONMENT PROGRAMME

**Credits:**

This report is a product of the Evaluation Office of the UN Environment Programme (UNEP). The findings and conclusions expressed herein do not necessarily reflect the views of the UNEP Senior Management.

For further information on this report, please contact:

Evaluation Office of UNEP
P. O. Box 30552-00100 GPO
Nairobi Kenya
Tel: +254 20 762 3389
Email: Michael.Spilsbury@un.org

March 2020

All rights reserved.

© (2020) Evaluation Office of UNEP

Content list

Executive Summary	vi
1 Introduction	1
1.1 Scope and objectives of the synthesis report.....	1
1.2 Evaluation Office mandate and mission	1
2 UNEP's results definitions, evaluation criteria and coverage	3
2.1 UNEP's Results definitions.....	3
2.2 Evaluation criteria and rating scales	6
2.3 Evaluation coverage	9
2.4 Sample sizes	14
2.5 Tools supporting the evaluation process.....	14
2.6 Learning from evaluations - recommendations tracking and coding evaluation reports	15
2.7 Limitations	15
3 Analysis of project performance	17
3.1 Performance in the 2018-19 biennium.....	17
3.2 Strategic relevance	22
3.3 Quality of project design	23
3.3.1 <i>Relationship between Quality of Project Design and Effectiveness Ratings</i>	24
3.4 Nature of external content	24
3.5 Effectiveness	25
3.5.1 <i>Availability of outputs</i>	26
3.5.2 <i>Achievement of project outcomes</i>	27
3.5.3 <i>Likelihood of impact</i>	29
3.6 Financial management.....	29
3.7 Efficiency.....	31
3.8 Monitoring and reporting	32
3.9 Sustainability.....	35
3.10 Factors affecting project performance	36
3.10.1 <i>Preparation and readiness</i>	36
3.10.2 <i>Quality of project management and supervision</i>	37
3.10.3 <i>Stakeholder's participation and cooperation</i>	38
3.10.4 <i>Responsiveness to human rights and gender equity</i>	38
3.10.5 <i>Country ownership and driven-ness</i>	40
3.10.6 <i>Communication and public awareness</i>	40

4	Key findings from strategic evaluations	42
4.1	Resource Efficiency Sub-programme.....	42
4.2	Clean Seas Campaign	45
4.3	Montevideo Programme on Environmental Law	49
4.4	Biochar for Sustainable Soils.....	53
4.5	La Plata Basin Project Evaluation.....	56
4.6	Poverty-Environment Initiative (a Joint UNEP / UNDP Evaluation)	58
5	Compliance with evaluation recommendations	63
5.1	Overview of Compliance with Evaluation Recommendations	63
6	Performance of the Evaluation Office and evaluation quality.....	69
6.1	The OIOS evaluation dashboard study of UNEP Evaluation Office	69
6.2	GEF Independent Evaluation Office assessment of the quality of UNEP evaluation reports	69
6.3	UN System-Wide Action Plan on gender equality and the empowerment of women (UN-SWAP)	70
6.4	United Nations Evaluation Group Annual General Meeting 2019	71
	Annex 1. List of Evaluations in the 2018-19 Biennial Report.....	73
	Annex 2. List of evaluations in 2018-19 rated in the satisfactory range for responsiveness to human rights and gender equity.....	78

List of figures

Figure 1. Distribution by sub-programme of projects evaluated in 2018-19.....	vii
Figure 2. Summary of projects attaining ‘Satisfactory’ or better performance against the main evaluation criteria in 2018-19.....	viii
Figure 3. Summary of projects attaining ‘Satisfactory’ or better performance in the factors affecting performance in 2018-19.....	viii
Figure 4. Performance against all evaluation criteria in projects evaluated during the 2018-19 biennium ..	x
Figure 5. Comparison of overall project performance by biennium	xi
Figure 6. Completion of evaluation management response (recommendation implementation plans) by Division	xii
Figure 7. Accepted evaluation recommendations with a completed compliance process 2018-2019	xiii
Figure 8. Comparison of distribution of evaluations by sub-programme in 2010-19 and the 2018-19 biennium.....	10
Figure 9. Comparison of evaluations conducted for GEF and non-GEF projects in 2018-19.....	11
Figure 10. Distribution of projects evaluated in 2018-19 by UNEP Division	11
Figure 11. Distribution of projects evaluated in 2018-19 by global region.....	12

Figure 12. Alignment of projects evaluated in 2018-19 with SDG Targets	13
Figure 13. Overview of project performance by criteria in the 2018-19 biennium.....	17
Figure 14. Performance against all evaluation criteria in projects evaluated during the 2018-19 biennium	18
Figure 15. Comparison of overall project performance by biennium	22
Figure 16. Comparison of overall project performance by sub-programme in 2018-19	22
Figure 17. Comparison of strategic relevance (overall) by biennium.....	23
Figure 18. Comparison of quality of project design in 2018 and 2019	24
Figure 19. Comparison of nature of external context in 2018 and 2019.....	25
Figure 20. Comparison of effectiveness (overall) by biennium	26
Figure 21. Comparison of availability of outputs by biennium	27
Figure 22. Comparison of achievement of outcomes by biennium.....	28
Figure 23. Comparison of likelihood of impact by biennium	29
Figure 24. Comparison of financial management (overall) by biennium	30
Figure 25. Financial management by sub-criteria in 2018-19.....	31
Figure 26. Comparison of efficiency by biennium	32
Figure 27. Comparison of project monitoring by biennium	33
Figure 28. Monitoring by sub-criteria in 2018-19.....	33
Figure 29. Comparison of sustainability (overall) by biennium	35
Figure 30. Comparison of sustainability sub-criteria in 2018-19.....	36
Figure 31. Comparison of preparation and readiness by biennium	37
Figure 32. Quality of project management and supervision in 2018-19	37
Figure 33. Comparison of stakeholder participation and cooperation by biennium	38
Figure 34. Responsiveness to human rights and gender equity in 2018-19	39
Figure 35. Comparison of country ownership and driven-ness by biennium.....	40
Figure 36. Comparison of communication and public awareness in 2016-17 and 2018-19	41
Figure 37. Completion of evaluation management response (recommendation implementation plans) by Division	65
Figure 38. Acceptance of evaluation recommendations 2018-19.....	65
Figure 39. Implementation status of project evaluation recommendations issued, shown by Division (2018 – 19).....	66
Figure 40. Accepted evaluation recommendations with a completed compliance process 2018-19	66
Figure 41. Accepted evaluation recommendations with a completed compliance process 2016-17	67

List of tables

Table 1. UNEP results definitions relevant to evaluations	4
Table 2. UNEP performance evaluation criteria and adjustments over time	6
Table 3. Performance rating scale	9
Table 4. Number of completed evaluations with performance ratings by sub-programme	14
Table 5. Comparison of S/HS performance between 2016-17 and 2018-19 biennia.....	19
Table 6: Rate of Management Responses	64
Table 7. Compliance status for Corporate\ High Level Recommendations issued 2018-19	68

Boxes

Box 1. Summary of 2018-19 text analysis findings: Completeness of financial information	30
Box 2. Summary of 2018-19 text analysis findings: Communication between project and financial management staff	31
Box 3. Summary of 2018-19 text analysis findings: Monitoring design and budgeting	34
Box 4. Summary of 2018-19 text analysis findings: Monitoring of project implementation.....	34
Box 5. Summary of 2018-19 text analysis findings: Project reporting	34
Box 6. Summary of 2018-19 text analysis findings: Gender responsiveness	39
Box 7. Institutional History of the water resource management in the La Plata Basin.....	58

Executive Summary

1. This Evaluation Synthesis Report summarises all independent evaluations undertaken by the Evaluation Office of the UN Environment Programme (UNEP) in the 2018-19 biennium. During this period 61 evaluations of UNEP interventions that, together, encompass a resource envelope in excess of USD 265 million¹ of expenditure, were completed. Highlights of the aggregated performance trends across these evaluations for the 2018-19 biennium are presented. The report also provides brief highlights drawn from a selection of evaluations of high strategic importance including: the Resource Efficiency Sub-programme Evaluation, the joint UNEP / UNDP Poverty and Environment Initiative, the Clean Seas Campaign and an assessment of the Montevideo IV Programme on International Law Panel to name but a few.

2. All evaluations should improve organizational learning and help stakeholders to hold UNEP accountable for contributing to development results at different levels. In short, the work of the Evaluation Office enhances accountability, transparency and learning. Evaluations generate evidence to identify 'what works' and 'what doesn't' and provide feedback for the improvement of planning and management processes.

Coverage and performance of projects completed during the 2018–19 biennium

3. The evaluations undertaken by the Evaluation Office are strongly influenced by project timelines. Evaluations are scheduled and initiated as projects reach, or approach, their operational completion.

4. Looking at the distribution of project evaluations, in Figure 1 below, the largest proportion of evaluations were for projects in the following sub-programmes: Climate Change (CC), Chemicals, Waste and Air Quality (CW), Environmental Governance (EG), and Healthy and Productive Ecosystems (HPE). The Evaluation Office also notes that the number of evaluations for EG is considerably increased by the numerous GEF Biosafety projects that have been housed under this sub-programme.

5. Project evaluations for the Resilience to Disasters and Conflict (D&C) and Environment Under Review (EUR) sub-programmes were both few in number and it is noted that these sub-programmes both have a relatively small number of projects in their portfolio compared to others. Resource Efficiency (RE) is also less well represented than normal this biennium and this may be due to the recent completion of the RE Sub-programme evaluation, which was preceded by a period where RE project evaluations were deliberately prioritized.

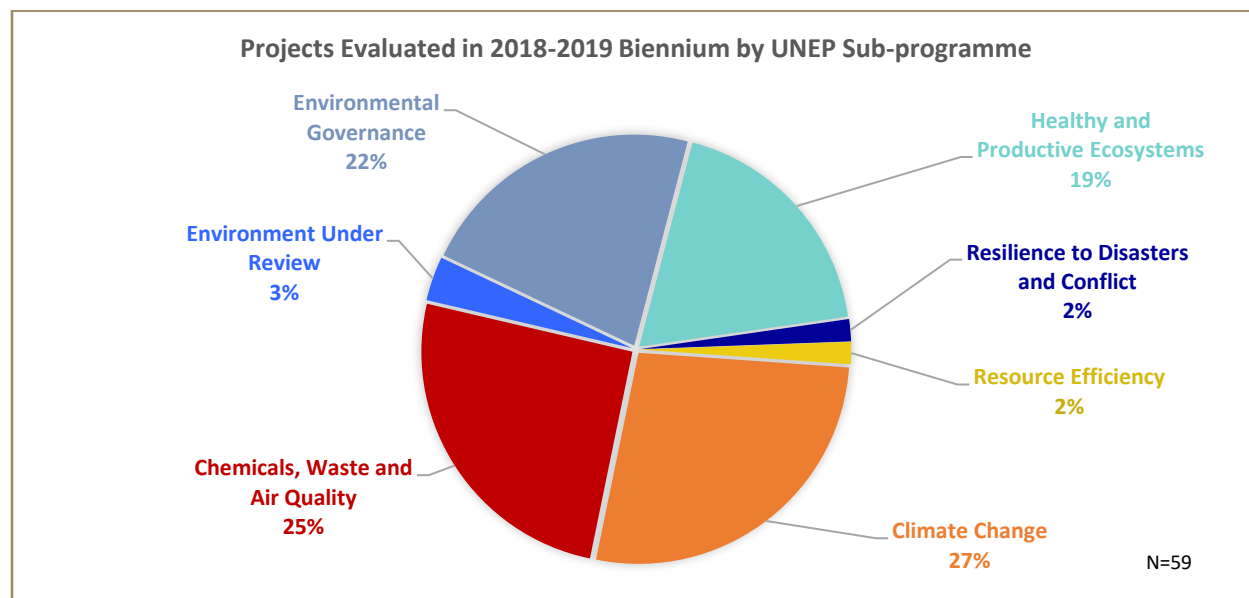
6. The main report also breaks down evaluation coverage by funding modality, Division, geographic region and by the links between the project intervention evaluated and the relevant SDGs (Figure 9, Figure 10, Figure 11 and Figure 12).

7. Consistent with UNEP's Evaluation Policy, some performance assessment needs are met through a 'Review' modality, which is led by the project manager and supported by the Evaluation Office. Reviews are carried out to meet the terminal performance needs of projects under a threshold of USD 1million and to the majority of mid-point assessments. Data on the number of

¹ This is the approximate total of the reported expenditures declared in each project evaluation report and includes GEF grants, Extra Budgetary funding and cash co-finance.

reviews has not been collected in a systematic manner in the past but, for indicative purposes, the Evaluation Office supported a total of 16 project reviews during 2019 (7 Terminal Reviews; 9 Mid-Term Reviews).

Figure 1. Distribution by sub-programme of projects evaluated in 2018-19



8. All projects are evaluated against a standard set of evaluation criteria that are consistent with international good practice. Performance against all criteria is rated on a six-point scale from 'Highly Unsatisfactory' through to 'Highly Satisfactory'². The benchmark for good performance is set by the UNEP Evaluation Office at ratings of 'Satisfactory' or 'Highly Satisfactory' (S/HS). The set of performance evaluation criteria have evolved over time in order to:

- Be responsive to UNEP's policies and areas of priority
- Make necessary areas of learning visible
- Create a comprehensive and transparent framework against which all performance aspects of UNEP's work can be evaluated

9. As has been reported in previous biennia, the Evaluation Office develops tools and procedures that are intended to help evaluate projects with greater rigor, consistency and objectivity. In doing so, the Office liaises with other parts of the house to ensure that the way in which projects are evaluated is consistent with the guidance UNEP provides during project design, development and implementation. Such liaison between departments also contributes to an ongoing process of institutional sharing and learning.

10. A significant number of evaluation tools were developed during the 2016-17 biennium and these have been tested and subsequently revised during 2018-19. In particular, the Evaluation Criteria Ratings Matrix, that provides descriptions of the performance levels and evidence

² The six-point rating scale is as follows: Highly Unsatisfactory (HU); Unsatisfactory (U); Moderately Unsatisfactory (MU); Moderately Satisfactory (MS); Satisfactory (S) and Highly Satisfactory (HS). For the criteria Sustainability and Likelihood of Impact, the same scale is used and the word 'Satisfactory' is replaced with the word 'Likely' and for the criterion Nature of External Context the word 'Favourable' is used. The abbreviations are altered as appropriate.

requirements, (e.g. 'Highly Satisfactory' or 'Moderately Unsatisfactory') was applied in most of the evaluations during 2018-19 and has recently been revised to incorporate suggestions for improvement from both evaluation consultants and project teams.

11. The proportion of projects that received a S/HS ratings against the evaluation criteria is summarised in Figure 2 and Figure 3 below:

Figure 2. Summary of projects attaining 'Satisfactory' or better performance against the main evaluation criteria in 2018-19³

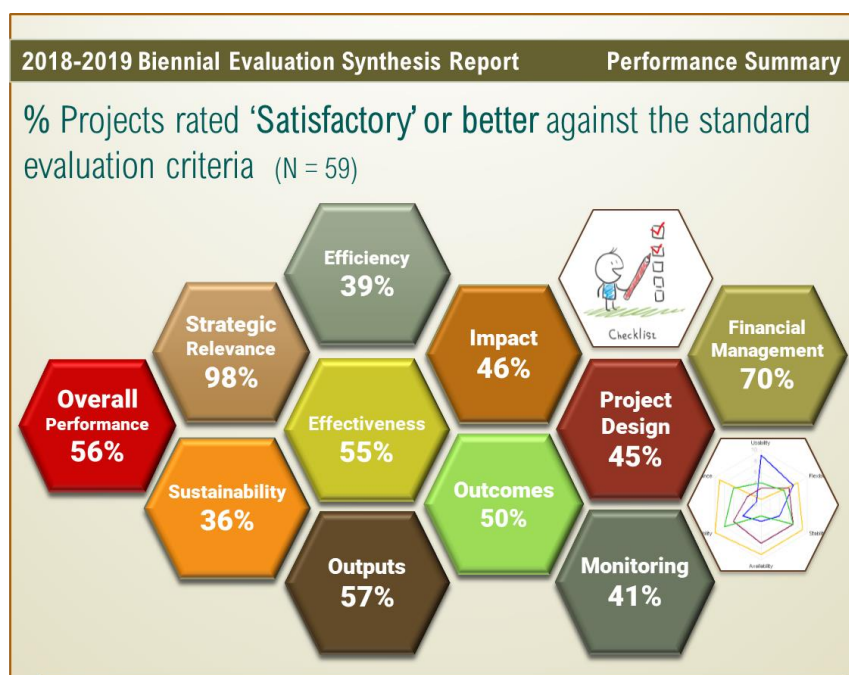
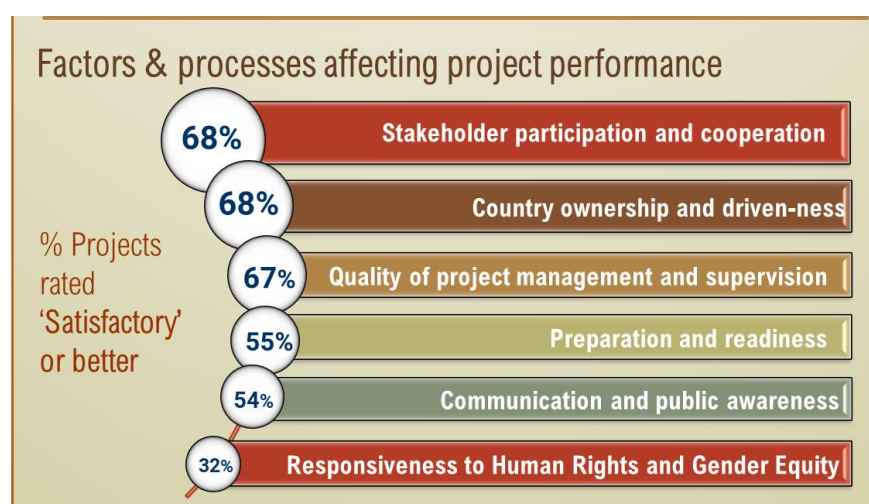


Figure 3. Summary of projects attaining 'Satisfactory' or better performance in the factors affecting performance in 2018-19



³ Although more than 60 evaluations were conducted, some were of strategic initiatives or arrangements that were not rated against standard evaluation criteria, and could therefore not be included in the comparative data

12. Several overall observations are evident from the figures above:

Positive performance

- Strategic Relevance (98% S/HS ratings) remains the highest-performing criterion and is closely linked to the project design and approval process, during which alignment with strategic priorities is checked.
- Financial Management (70% S/HS ratings) has moved from a lower-performing criterion to the higher-performing group, in large part due to a re-definition of the criterion, which has seen strong performance ratings for communication between project and financial management staff members.
- Effectiveness performance ratings (55% S/HS ratings) show the familiar pattern of stronger performance at output level (57% at S/HS levels), decreasing to lower ratings for likelihood of impact (46%). From a results perspective, performance at Project Outcome level is a critical indicator; currently 50% of all projects achieve performance ratings at S/HS levels.
- Overall Project Performance ratings (56% S/HS) are heavily influenced by the performance at outcome level and their sustainability. Currently, just over half of those projects evaluated achieve a S/HS rating for overall project performance.
- Among the Factors Affecting Performance the highest performing factors are critical for change to happen: stakeholder participation and cooperation (68% S/HS); country ownership and driven-ness (68% S/HS) and quality of project management and supervision (67% S/HS).

Areas for attention

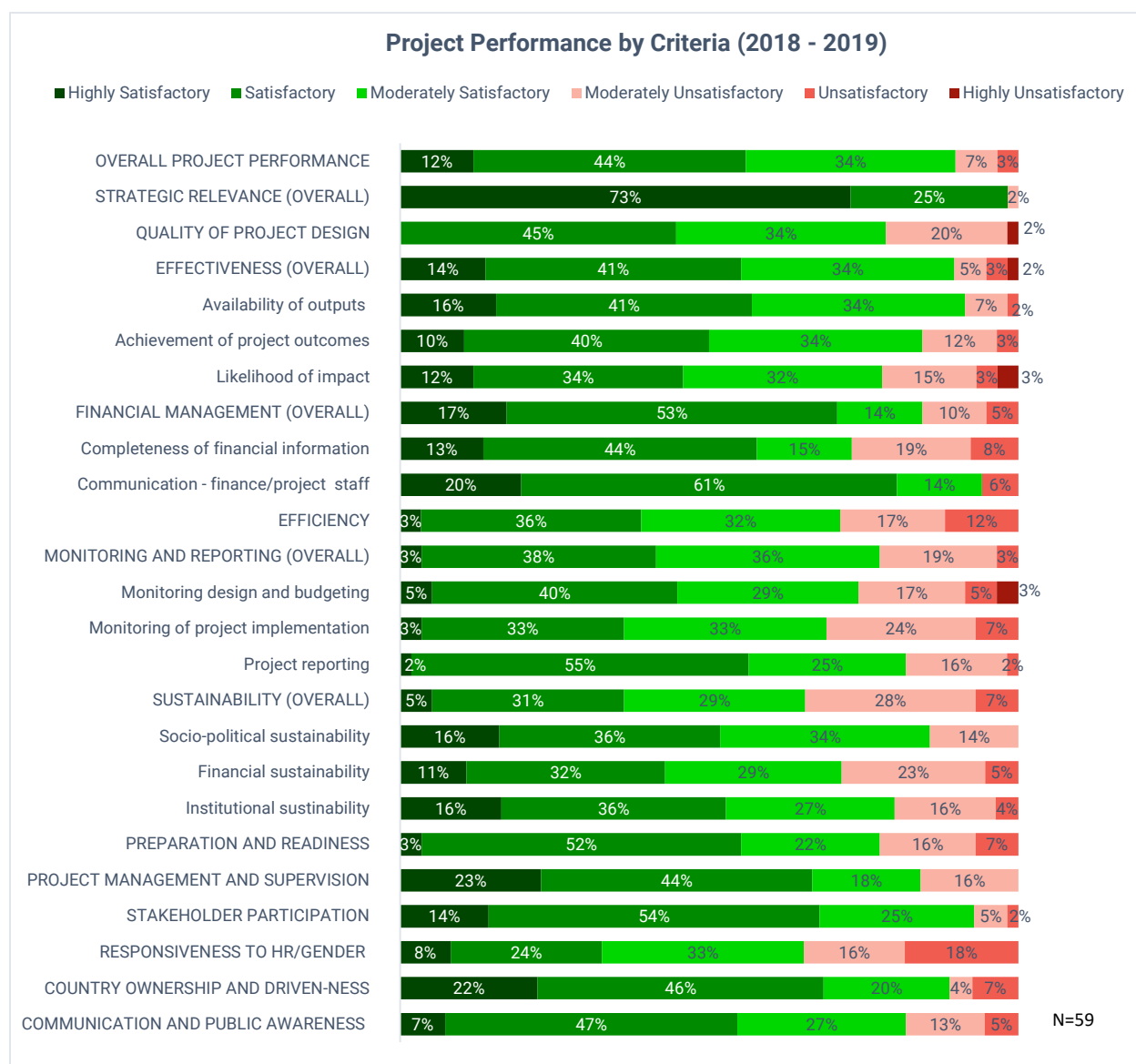
- Among the Factors Affecting Performance the lower performing factors that would benefit from attention are: preparation and readiness (55% S/HS), which is defined as the period between project approval and first disbursement, and communications and public awareness (54% S/HS).
- Quality of project design (45% S/HS) is on the low side and should improve substantially in the coming years as evaluations in this biennium are currently assessing projects approved between 2011 and 2014.
- Monitoring (41% S/HS) still shows low levels of performance in the monitoring of project implementation (as opposed to the design of, and budgeting for a monitoring plan and project reporting). Monitoring project implementation is at the heart of the potential for effective monitoring to improve project performance through adaptive management and course correction.
- Efficiency (39% S/HS) remains a challenging criterion to assess as financial information is frequently not available at the disaggregated levels necessary to consider cost-effectiveness. Evaluations view no-cost extensions as having a negative effect on efficiency.
- Sustainability (36% S/HS) shows a lower than desired level. This is, in part, influenced by performance at the level of outcomes because, under the 'sustainability' criterion,

evaluations assess the likelihood of the benefits derived from the achievement of outcomes being of an enduring nature.

- The criterion that shows the lowest performance levels is Responsiveness to Human Rights and Gender (32% S/HS). This may show some improvement in the coming years as project proposal templates began incorporating this element from 2010 and institutional capacity development and take-up should have followed thereafter. However, this is an area where stories of positive achievement in UNEP projects need to be shared to promote learning and improvement.

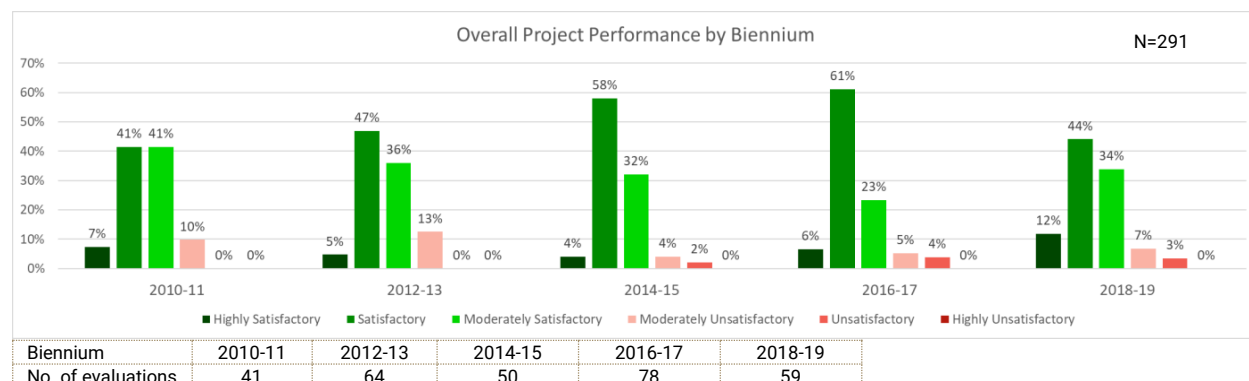
13. Figure 4 below provides an aggregated view of project performance across all six rating categories by criteria, and their sub-categories, based on data collated from evaluation reports completed in the 2018-19 biennium.

Figure 4. Performance against all evaluation criteria in projects evaluated during the 2018-19 biennium



14. General trends over a longer period, back to the 2010-11 biennium, have also been reflected upon, see Figure 5 below. Over the past nine years, UNEP's projects have maintained steady performance within the 'Satisfactory' range (i.e. 'Moderately Satisfactory', 'Satisfactory' and 'Highly Satisfactory'). There is a strong pattern of the majority (40 – 60%) of projects in each biennium reaching the fifth point on the six-point scale, i.e. the 'Satisfactory' level, which represents the performance threshold applied by the UNEP Evaluation Office. The proportion of projects being assessed at the 'Highly Satisfactory' level has noticeably increased in the current biennium, although this is not a statistically significant change.

Figure 5. Comparison of overall project performance by biennium



15. In addition to performance ratings against set criteria, evaluation reports also contain considerable insights into different aspects of project design and implementation.

16. During 2019 the UNEP Evaluation Office began the process of coding the texts of each evaluation report, using dedicated software, so that searches can be performed across evaluations on topics of interest and value to the house. Initial findings on **Monitoring and Reporting** (Box 3, Box 4, Box 5), **Financial Management** (Box 1, Box 2) and **Responsiveness to Gender** (Box 6) are included in the report based on this emerging approach. Together with tracking of the implementation of evaluation recommendations, the coding of evaluation report findings is intended to strengthen the capacity of the Evaluation Office to systematically share learning with targeted audiences, especially in-house.

17. The criterion '**Quality of Project Design**' was incorporated into the evaluation ratings table in 2017. The Evaluation Office undertook a study of data to-date to examine whether there is a link between the ratings for the 'Quality of Project Design' and later project performance against effectiveness criteria (Section 3.3.1). Whilst there is no clear statistical correlation, this study did highlight the importance of country ownership and political will in achieving the intended project results.

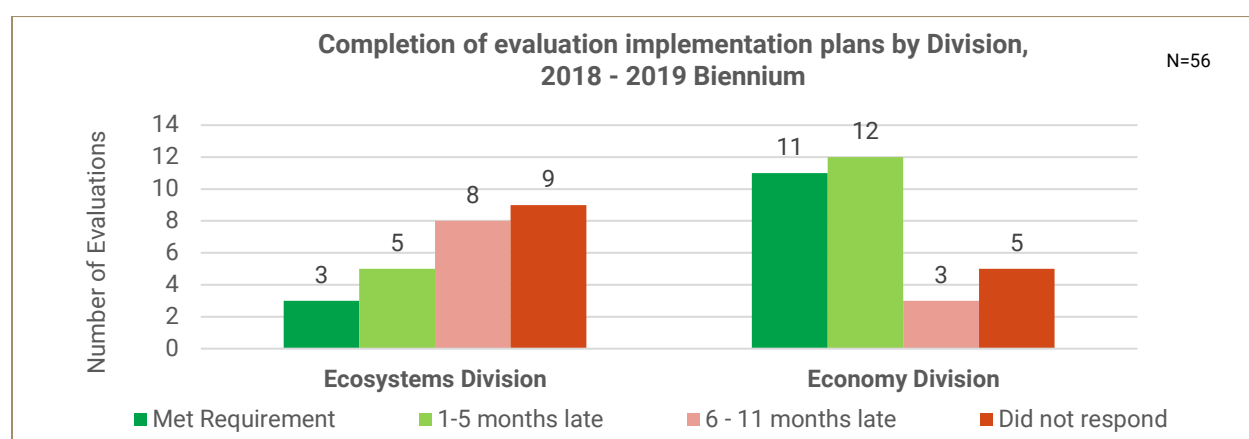
Evaluation recommendation compliance

18. The Evaluation Office undertakes robust independent evaluations that are followed by a recommendation compliance process. In the 2018-19 biennium, 61 evaluations were completed with 56 of them issuing recommendations. The evaluations that required a management response were exclusively from projects residing within Ecosystems and Economy Divisions.

19. The preparation of a management response in the form of a recommendation implementation plan is a **mandatory** requirement for all evaluations.

20. Fifty-six evaluation recommendation implementation plans, comprising a total of 296 recommendations, were sent to UNEP substantive offices. Overall, **75 %** (42) of project/programme officers prepared an implementation plan that provided responses to the evaluation recommendations. Only **25%** (14) of the evaluation management responses received were submitted within one month of the issuance of the evaluation report as per the requirements set out in the UNEP Programme Manual⁴. Therefore, for **25%** (14) of evaluations that were completed and issued recommendations, project / programme managers **failed to prepare a management response**.

Figure 6. Completion of evaluation management response (recommendation implementation plans) by Division



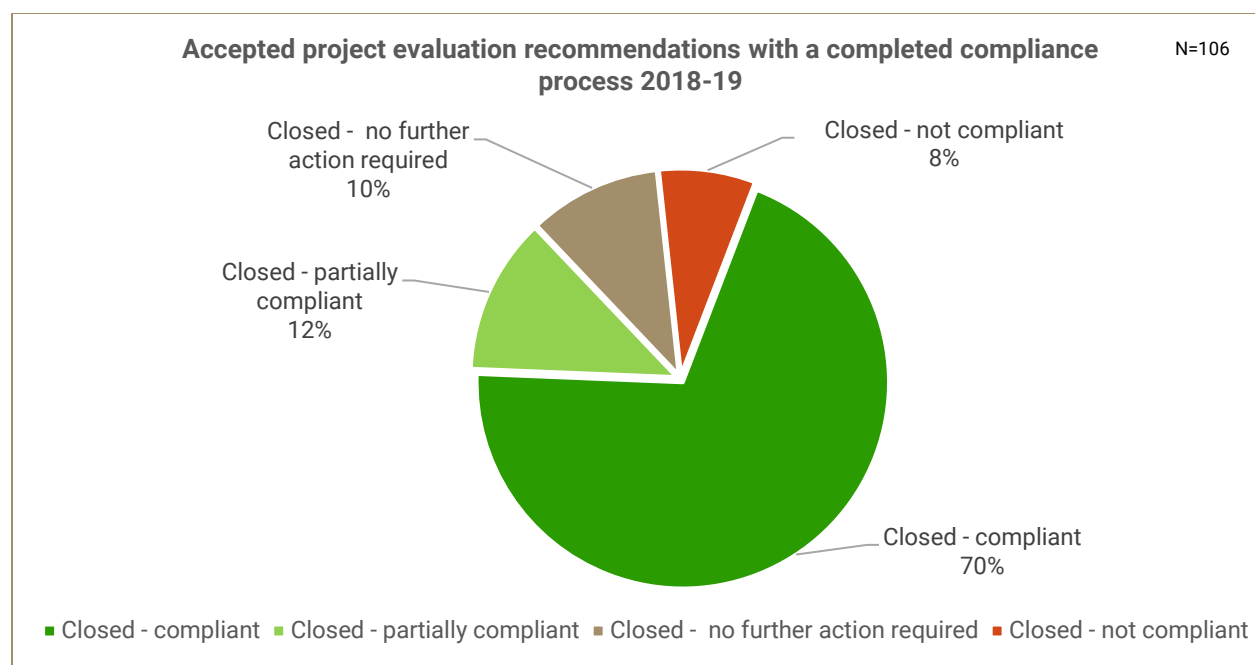
21. Compliance rates for completion of management responses to evaluation **should be 100%** and the timelines for these standard requirements should be respected. This is a performance dimension needing strong management attention.

22. Figure 7 shows the compliance performance for project / programme evaluation recommendations that were accepted and where a recommendation implementation plan was completed. This shows a strong performance with only 8% of recommendations being closed as not compliant for the biennium.

23. The performance in recommendation compliance in the 2018-19 biennium compared with that for the 2016-17 biennium shows broadly similar patterns but with a slight reduction in compliance performance as shown by the increase from 3% to 8% of recommendations closed as 'not compliant'. This needs to be monitored closely during the 2020-21 biennium, and it is anticipated that evaluation recommendation compliance will feature in internal corporate monitoring 'dashboards' that are currently under development.

⁴ <https://wecollaborate.unep.org/pages/viewpage.action?pageId=9374201> (internal intranet reference)

Figure 7. Accepted evaluation recommendations with a completed compliance process 2018-2019



Performance of the UNEP Evaluation Office

24. The UNEP Evaluation Office was externally assessed during the 2018-19 biennium by the Office of Internal Oversight Services of the United Nations Secretariat (OIOS) as part of a study of evaluation capacity and practice in every entity in the Secretariat⁵. The 'Evaluation Dashboard' includes an independent assessment based on 19 indicators of evaluation practice. The OIOS 'Dashboard' report noted that the UNEP *"had robust evaluation systems in place across all relevant Evaluation Dashboard indicators. Its function was organized into a stand-alone evaluation unit, and its most senior evaluation professional was at the D-1 level. A strong evaluation policy, plan, and procedures were in place."*

25. The Independent Evaluation Office of the GEF also publishes an annual comparative assessment of the performance of GEF projects, including an assessment of the quality of evaluation reports received from GEF Agencies. In the most recent Annual Performance Report (APR), published by the GEF in October 2019, the assessment of the performance of the UNEP Evaluation Office continues to be very strong. The data presented in the most recent APR also indicates that **100% of UNEP evaluation reports** of GEF projects it recently assessed were rated 'Moderately Satisfactory' or better for quality, this compares very favourably against other GEF agencies including UNDP and the World Bank.

⁵ United Nations Evaluation Dashboard 2016-2017. OIOS Assignment No: IED-19-002, 30 April 2019.

1 Introduction

1.1 Scope and objectives of the synthesis report

26. This report presents **an analysis of evaluations conducted in the 2018-19 biennium**. It utilizes information drawn from in-depth evaluations including: a review of the Resource Efficiency Sub-programme Evaluation, review of the United Nations System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women, an in-depth assessment of the Montevideo Programme on Environmental Law, GEO-6, The Economics of Ecosystems and Biodiversity (TEEB), and **over fifty other in-depth project evaluations** spanning the UNEP Programme of Work and Global Environment Facility project portfolios. Summaries of a selection of project evaluations are included in Chapter 4. The evaluation synthesis report also contains a review of the implementation status of evaluation recommendations and presents external assessments of the performance of the Evaluation Office.

27. The report is prepared as an inter-sessional document of the UN Environment Assembly of UNEP and serves as part of the input of UNEP to the Secretary-General's report on evaluation to the General Assembly. The report provides stakeholders such as Governments, UNEP senior management and partners with an evaluative assessment of UNEP's programme and project performance in the 2018-19 biennium. The main objective of the report is to help UNEP reflect on, and learn from, its programme performance through evaluative evidence and lessons from programme and project design and implementation.

1.2 Evaluation Office mandate and mission

28. The mandate for conducting, coordinating and overseeing evaluation in UNEP is vested in the Evaluation Office and has been stated in various General Assembly resolutions and UNEP Governing Council / Environment Assembly decisions. The Governing Council recognized the importance of evaluation as an integral part of the programme planning cycle, while retaining its independence, and requested the Executive Director to continue to refine evaluation methodologies in collaboration with Governments and partners within the United Nations system (Governing Council decisions 75 IV, 6/13, 13/1 and 14/1).

29. The mandate for evaluations in UNEP covers all programmes and projects of the Environment Fund, related trust funds, earmarked contributions, partnership agreements and projects implemented by UNEP under the Global Environment Facility (GEF) and the Green Climate Fund but does not extend to cover work conducted by UNEP MEA Secretariats.

30. The Evaluation Office comprises five Professional positions and three General positions. The team is supplemented on a regular basis by three or four consultants as Individual Contractors who work on a part-time basis.

31. The evaluation team undertakes a variety of evaluations and management studies, in accordance with the requirements of the United Nations General Assembly (UNGA), the United Nations Environment Assembly, and in conformity with the United Nations Evaluation Group Norms and Standards for Evaluation. The activities of the Evaluation Office can include high-level strategic evaluations of UNEP's thematic sub-programmes, evaluations of the UNEP strategic

planning processes, in-depth project evaluations, portfolio evaluations, cross-cutting thematic evaluations and management studies.

32. The Evaluation Office closely follows-up on the implementation of all accepted evaluation recommendations and, when requested, provides technical backstopping to project and programme managers undertaking project reviews by advising on consultant selection, providing a suite of 19 tools, templates and guidelines to support the review process, providing technical guidance on a needs-basis and completing a quality assessment of the final review report. Guidelines, formal requirements and practical advice in planning for evaluations have been specified in detail in the UNEP Programme Manual and are further elaborated on the UNEP Evaluation website at: www.unenvironment.org/about-un-environment/evaluation

33. This evaluation synthesis report has been prepared as part of the mission of the Evaluation Office to promote a results focus in UNEP that reflects an organizational culture of learning, informed decision-making and accountability. According to the Secretary General's bulletin on programme planning, monitoring and implementation (ST/SGB/2000/8), which consolidates the General Assembly decisions on the evaluation function, *"The objective of evaluation is: (a) To determine as systematically and objectively as possible the relevance, efficiency, effectiveness and impact of the Organization's activities in relation to their objectives; (b) To enable the Secretariat and Member States to engage in systematic reflection, with a view to increasing the effectiveness of the main programmes of the Organization by altering their content and, if necessary, reviewing their objectives."*

2 UNEP's results definitions, evaluation criteria and coverage

2.1 UNEP's Results definitions

34. In the 2018-19 biennium, the UNEP Evaluation Office collaborated with the Programme Coherence/Assurance Unit (PCA) of the Policy and Programme Division, and colleagues from other departments in a process to generate a glossary of UNEP results definitions⁶. A summary of those terms most closely associated with evaluation processes is shown below in Table 1.

35. During this work, four distinctions were usefully explored and made clear:

Activity - Deliverable - Output

36. From a results perspective the attention is placed on those who are affected by projects rather than those who are responsible for delivering them. This work on results' definitions provided an opportunity to clarify that **activities** (work performed) and **deliverables** (products or services derived from a completed activity) differ from **outputs** (immediate gains experienced from the work performed and products or services delivered) by the inclusion of the intended beneficiary or user. Importantly, this refinement in the definition of an output only shifts the perspective from that of the supplier (what has been delivered) to that of the beneficiary (what has been gained); it does not change the ambition underpinning the output, nor does it change the timeframe within which outputs are made available. For example, in this way the 'provision of technical assistance' (activity) builds the knowledge and understanding of a target audience (output), which may lead to a change of behaviour, based on that new knowledge and understanding, by that target audience (outcome).

Direct Outcome - Project Outcome

37. Evaluations aim to assess what has been achieved against what was planned (and committed to). In order to clearly represent the ambition of, and commitment made by, a project the term **Project Outcome** will be used by the house to refer to those outcomes a project is expected to achieve by the time it is operationally complete. However, it is recognised that some outcomes (i.e. the use or uptake of outputs) occur before the end of the project and these will be referred to, in all parts of the house, as **Direct Outcomes**. Project outcomes are the highest level of result to which UNEP evaluations hold projects to account.

Attribution - Contribution - Credible Association

38. The gold standard of proving that a project has achieved its results is proving attribution between the demonstrated results and the efforts of the project. However, this is not always possible, and it is helpful to consider a wider range of ways of establishing claims between project efforts and observed results.

39. In order to *attribute* any outcomes and impacts to a project intervention, one needs to consider the difference between what has happened with, and what would have happened without, the project (i.e.

⁶ The UNEP results definitions build on those of OECD-DAC, UNDA, UNDP, UNEG, and the World Bank.

take account of changes over time and between contexts in order to isolate the effects of a project). This requires appropriate baseline data and the identification of a relevant counterfactual, both of which are frequently not available for evaluations. Establishing the *contribution* made by a project in a complex change process relies heavily on prior intentionality (e.g. approved project design documentation, logical framework) and the articulation of causality (e.g. narrative and/or illustration of the Theory of Change). Robust evidence that a project was delivered as designed and that the expected causal pathways developed can be used to support claims of contribution and this is strengthened where an alternative theory of change can be excluded (e.g. that the 'business as usual' behaviour pattern is no longer evident or prevalent). Finally, a *credible association* between the implementation of a project and observed positive effects can be made where a strong causal narrative, although not explicitly articulated, can be inferred by the chronological sequence of events, active involvement of key actors and engagement in critical processes.

Programme - Portfolio

40. There is a tendency for these two terms to be used either interchangeably or at least, loosely. There is, however, an important distinction between the two. Whereas the projects within both a portfolio and a programme will share common characteristics with each other, only those projects in a programme will be working towards common outcomes and be managed in a coordinated way that generates additional benefits. Evaluations of portfolios can provide insights on key issues based on multiple experiences of working on a common theme etc. but only evaluations of programmes can provide evidence of cumulative effects across a group of projects.

Table 1. UNEP results definitions relevant to evaluations

UNEP Results Definitions - relevant to evaluations	
TERM	DEFINITION
Activity	An action taken, or work performed, through which inputs are utilized to realize specific results.
Assumption	An assumption is a significant external factor or condition that needs to be present for the realization of the intended results but is beyond the influence of the project and its partners. Assumptions are often positively formulated risks. (see also Driver)
Attribution	Attribution can be claimed when comprehensive evidence proving the cause and effect relationship between the project and the observed results is presented.
Contribution	Contribution can be claimed when compelling evidence is presented that supports a cause and effect relationship through which intended collective results are achieved by the combined efforts of more than one project.
Credible Association	Credible association can be claimed when compelling evidence supporting a cause and effect relationship between the project and the observed results.
Driver	A driver is a significant external factor that, if present, is expected to contribute to the realization of the intended results of a project. Drivers can be influenced by the project and its partners. (See also Assumption)
Direct Outcome(s)	A direct outcome is an outcome that is intended to be achieved from the uptake of outputs and occurring prior to the achievement of Project Outcome(s).
Deliverable	A deliverable is a specified product or service derived from completed activity(ies). For example, a draft report, a revised manual, X number of workshop facilitation days are all

UNEP Results Definitions - relevant to evaluations	
TERM	DEFINITION
	frequent deliverables in a UNEP project context. (Deliverables are viewed from the perspective of the supplier/provider of the product or service)
Effect	An effect is a change which is a consequence of an action or other social, economic, political or environmental cause. These changes can be intended, unintended, positive or negative.
Evaluand	The evaluand is the entity being evaluated.
Evaluand parameters	An evaluand's parameters are the features of an evaluand that define its scope (e.g. timeframe, funding envelope, results framework, geographic dimensions).
Impact	Impacts are long-lasting results arising, directly or indirectly from a project. Impacts are intended and positive changes and must relate to UNEP's mandate.
Impact Pathway / Causal Pathway / Results Chain	Impact or Causal Pathways and Results Chains all describe cause and effect relationships between outputs, outcomes and impacts and are the basis of a programme/project's "Theory of Change".
Intermediate states	Intermediate states are changes (i.e. changes at the outcome level) beyond the Project Outcome(s) that are required to contribute towards the achievement of the intended impact of a project.
Milestone	A milestone is a scheduled event or achievement that represents a major stage in the progress of the project towards expected indicator targets and which is verifiable.
Monitoring	Monitoring is a continuing function that uses the systematic collection of data on project / programme implementation (e.g. completion of activities, rate of expenditure, emergence of risks, milestone delivery, inclusive participation of intended stakeholders, etc.) to provide management with indications of the extent of progress against plans and targets.
Outcome(s)	An outcome is the use (i.e., uptake, adoption, application) of an output by intended beneficiaries, observed as a change in institutions or behaviours, attitudes or conditions.
Outputs	An output is the availability (for intended beneficiaries/users) of new products and services and/or gains in knowledge, abilities and awareness of individuals or within institutions. For example, access by the intended user to a report; new knowledge held by a workshop participant at the end of a training event; heightened awareness of a serious risk among targeted decision-makers. (Outputs are viewed from the perspective of the intended beneficiary or user of the output rather than the provider).
Portfolio	A group of projects and/or programmes that share a common characteristic relevant to the organization's strategic objectives (e.g. funded by the same donor, operating in the same thematic area etc).
Programme	A programme is a group of synergistic projects contributing to a common outcome(s) and managed in a coordinated way to obtain benefits not available from managing the projects individually.
Project	A project is a time-bound intervention with a specific funding envelope that addresses a defined set of results within an identified implementation context or geographic area. The main components of the project must be interlinked/interdependent to achieve the project outcome(s).

UNEP Results Definitions - relevant to evaluations	
TERM	DEFINITION
Project Outcome(s)	Project Outcome(s) are those outcomes that are intended to be achieved by the end of project timeframe/funding envelope.
Results	Results are intended changes in a state or condition that derive from a cause-and-effect relationship. Such changes must be describable and measurable/discernible. A results statement, the associated targets and their indicators should conform collectively to the SMART ⁷ and/or CREAM ⁸ principles. Outputs, outcomes and impact are considered 'results' (as opposed to inputs and activities).
Theory of Change	A Theory of Change is a method used for planning a project, describing the participation that will be needed by different actors and for evaluating the project's performance. It articulates long lasting intended impact and then maps backward to identify the preconditions necessary to achieve this impact(s). It is a comprehensive description and illustration of how and why a desired change is expected to happen in a context. A Theory of Change also allows for unintended positive and/or negative effects to be depicted.

2.2 Evaluation criteria and rating scales

41. The UNEP Evaluation Office assesses performance against a total of eight evaluation criteria, six of which have sub-categories (see Table 2 below). While these criteria reflect standard internationally accepted⁹ performance criteria (Relevance, Efficiency, Effectiveness, Impacts and Sustainability), they are also adapted to be relevant to those areas of performance of most interest to UNEP. For example, UNEP's evaluation criteria allow for assessments against Financial management and Monitoring and reporting but also cover a range of other factors affecting performance and cross-cutting issues.

42. This set of performance evaluation criteria have evolved over time in order to: a) be responsive to UNEP's policies and areas of priority; b) make necessary areas of learning visible and c) create a comprehensive and transparent framework against which all performance aspects of UNEP's work can be evaluated. In Table 2 below the major adjustments made during 2018-19 are noted.

Table 2. UNEP performance evaluation criteria and adjustments over time

UNEP Evaluation Criteria	
A. Strategic Relevance	Prior to 2018 this criterion included a sub-category related to gender. At that time, this placement reflected the place of gender in UNEP's policies and strategic priorities. In 2018-19 'gender' was moved to its own sub-category under cross-cutting issues to better reflect its central importance, along with human rights, during project implementation.
<i>1. Alignment to MTS and POW</i>	
<i>2. Alignment to Donor/GEF strategic priorities</i>	
<i>3. Relevance to regional, sub-regional and national environmental priorities</i>	
<i>4. Complementarity with existing interventions</i>	

⁷ SMART refers to targets that are Specific, Measurable, Achievable, Realistic and Time-Bound.

⁸ CREAM refers to indicators that are Clear; Relevant, Economic, Adequate, Monitorable.

⁹ The evaluation criteria described by the OECD/DAC are widely accepted as an international standard. UNEP Evaluation Office also works in accordance with the descriptions reflected in the UN Evaluation Group Norms and Standards.

UNEP Evaluation Criteria	
B. Quality of Project Design	Assessment against this dimension has been longstanding – with formal ratings for the criterion continuous since 2017. The assessment template used in evaluations is consistent with that used during the project design approval process.
C. Nature of External Context	Introduced in 2018, this is not a performance criterion and its rating is not used in the calculation of overall project performance. However, when a project is found to have faced some specific unfavourable external conditions that could not have been anticipated, the ratings on Effectiveness and Sustainability may be adjusted to take unexpected disruptions into account (e.g. conflict, natural disasters and political upheaval).
D. Effectiveness	
<i>1. Availability of outputs</i>	Since 2019 the emphasis has been placed on the availability of outputs to intended beneficiaries rather than the less specific term ‘delivery’ of outputs. This is in keeping with UNEP’s revised results definitions, 2019.
<i>2. Achievement of project outcomes</i>	During 2018-19 emphasis has been placed on those outcomes (project outcomes) that a project is expected to achieve within its project timeframe and with its secured funding.
<i>3. Likelihood of impact</i>	Since 2018 an excel-based tool that supports a combined assessment of the extent to which outcomes have been achieved and the contributing conditions needed to make long-lasting change happen (i.e. assumptions and drivers) have been seen to hold, has been tested and strengthened.
E. Financial Management	
<i>1. Adherence to UNEP’s policies and procedures</i>	Since late 2019 this sub-category has been introduced to capture insight that has been evident but less easy to articulate in evaluation reports.
<i>2. Completeness of project financial information</i>	Assessment against this sub-category has been continuous since 2018.
<i>3. Communication between finance and project management staff</i>	Assessment against this sub-category has been continuous since 2018.
F. Efficiency	Assessment against this criterion has been continuous as a stand-alone criterion since 2014.
G. Monitoring and Reporting	This criterion was changed from ‘Monitoring and Evaluation’ in 2016 and new sub-categories were defined during 2018-19.
<i>1. Monitoring design and budgeting</i>	The design of monitoring is, since 2018, combined with its budgeting.
<i>2. Monitoring of project implementation</i>	Since 2018, emphasis has been placed on the adequacy and effectiveness of monitoring project implementation rather than the delivery of a monitoring plan. This allows evaluations to give credit to projects that develop a range of ways of monitoring that were not part of the monitoring design.
<i>3. Project reporting</i>	Assessment against this sub-category has been continuous since 2018. It allows evaluations to give full credit to those projects that meet reporting requirements outside of an effective monitoring approach.

UNEP Evaluation Criteria	
H. Sustainability	Assessment against these sub-categories has been continuous since 2010, although the way in which the terms are understood and how likelihood should be assessed has been developed and refined over time.
1. Socio-political sustainability	
2. Financial sustainability	
3. Institutional sustainability	
I. Factors Affecting Performance and Cross-Cutting Issues	
1. Preparation and readiness	Assessment against this sub-category has been continuous since 2010 although the category included an assessment of the quality of project design up until 2015.
2. Quality of project management and supervision	Project management and supervision were brought together as one sub-category from 2018 as the boundaries between the two terms was found to be unhelpful in terms of learning and accountability.
3. Stakeholder's participation and cooperation	Assessment against this sub-category has been continuous since 2018.
4. Responsiveness to human rights and gender equity	Introduced as a stand-alone sub-category since 2018. Previously assessed as part of UNEP's policies and strategic priorities.
5. Environmental, social and economic safeguards	Introduced as a stand-alone sub-category in late 2019. Previously assessed as a negative, unintended effect under Effectiveness.
6. Country ownership and driven-ness	Assessment against this sub-category has been continuous since 2010.
7. Communication and public awareness	Assessment against this sub-category has been continuous since 2016.

43. In December 2019, the OECD DAC Network on Development Evaluation (EvalNet) published revised definitions of the OECD DAC evaluation criteria for relevance, effectiveness, efficiency, impact and sustainability, and included one new criterion, coherence. The EvalNet stressed that the criteria should be used thoughtfully and adjusted to the context of the intervention and the intended users' needs. The UNEP criteria remain consistent with the revised DAC framework although several more specific criteria are used in order to be able to explore areas of interest to the organisation (e.g. monitoring; financial management; safeguards etc.).

44. Across all UNEP Evaluation Office project evaluations, a six-point scale is used to rate performance against each evaluation criterion, as presented in Table 3 below. The rating system and evaluation quality control processes used by UNEP are consistent with those applied by the GEF, United Nations Development Programme (UNDP) and the World Bank. UNEP's evaluation ratings are regularly benchmarked against those of other relevant organisations.

45. UNEP Evaluation Office sets the desired performance level at 'Satisfactory'. On a six-point scale, Satisfactory is the fifth point and, therefore, represents a percentage score of 83%. Throughout this report, therefore, commentary is provided on the percentage of project evaluations with performance ratings in the 'Satisfactory' or better' category. Effectively, this means a project has been assessed with either a 'Satisfactory' or 'Highly Satisfactory' rating. The abbreviation for 'Satisfactory or better' is S/HS and will be used throughout this report.

Table 3. Performance rating scale

Category	Rating	Abbrev.	Abbrev.*		
Strategic relevance; Quality of project design; Effectiveness; Availability of outputs; Achievement of project outcomes; Likelihood of impact; Financial management; Efficiency; Monitoring and reporting; Sustainability; and Factors and processes affecting project performance	Highly Satisfactory	HS	HL		 'Satisfactory' range 
	Satisfactory	S	L		
	Moderately Satisfactory	MS	ML		
	Moderately Unsatisfactory	MU	MU		 'Unsatisfactory' range 
	Unsatisfactory	U	U		
	Highly Unsatisfactory	HU	HU		

*Sustainability and Impact are rated against a 6-point 'likelihood' scale, ranging from 'Highly Likely' to 'Highly Unlikely'

46. In order to arrive at an overall project performance rating the UNEP Evaluation Office applies a weighted scale to the individual ratings. In this weighted scale 30% of the overall performance score is determined under Effectiveness by the rating for Achievement of project outcomes and a further 20% is determined by the rating for Sustainability. What this means is that all UNEP evaluations place great importance on the achievement of results and the likelihood that the benefits derived from the project will be enduring.

2.3 Evaluation coverage

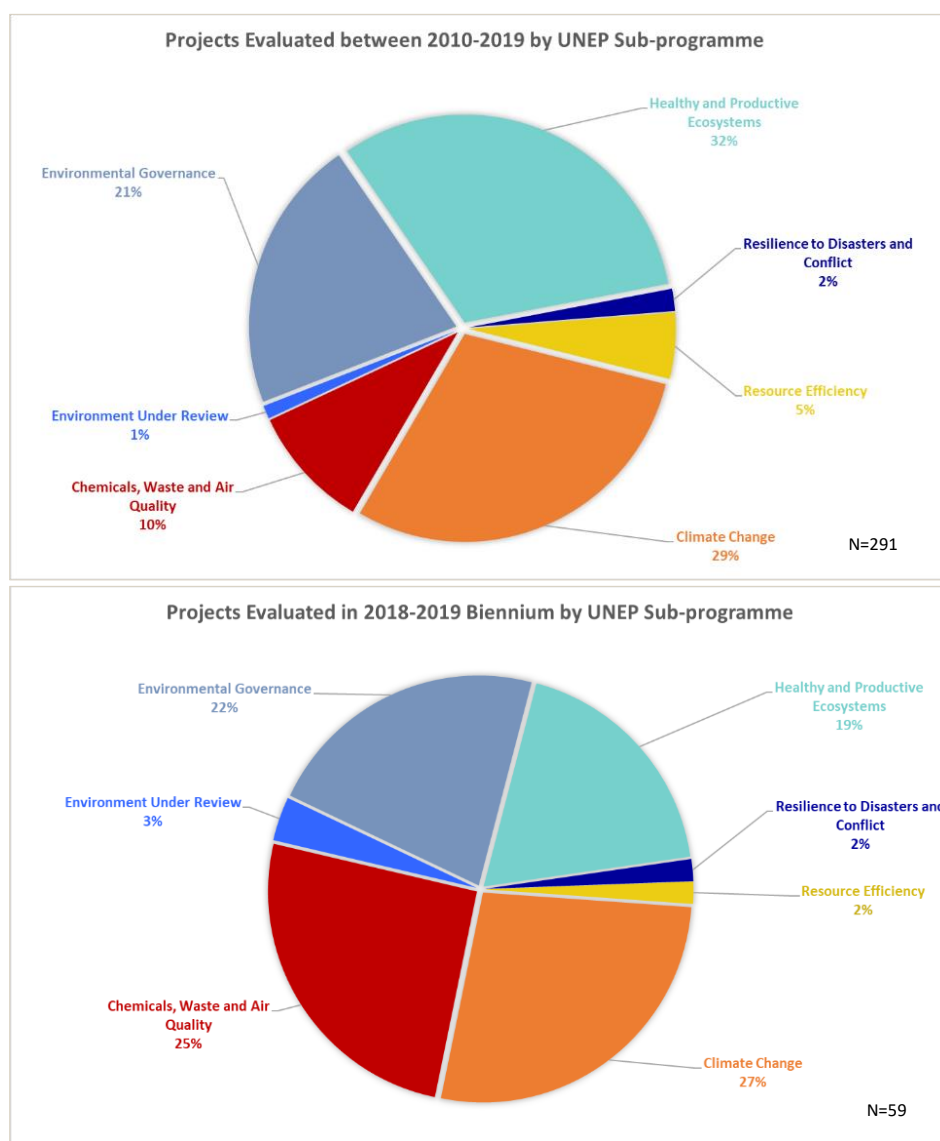
47. The mandate of the UNEP Evaluation Office is to evaluate projects implemented under the Programme of Work. At present, the sample of projects that is taken to represent the Programme of Work is determined by the operational completion date of projects from across all seven Sub-programmes. Over time, and as the number of projects has increased to the current total of 292 project evaluations carried out since 2010, one would expect such a sample to accurately reflect the characteristics of the whole of UNEP's portfolio. However, in order to check for unknown biases in this sample, the Evaluation Office has begun to examine its portfolio of project evaluations across a number of dimensions relevant to the house: by sub-programme; funding modality; division; region and by Sustainable Development Goals (SDGs). As information on all of these dimensions has not been readily available in the past, these data are drawn only from project evaluations completed during 2018-19. As mentioned above, this is still a small sample size for many of these dimensions and within each sample there is considerable diversity in types of project that have been evaluated (e.g. single country, global, field-based, normative etc.). Although conclusions cannot, therefore, be drawn from these data, the approach forms a foundation for further monitoring of UNEP's evaluation coverage in the future.

48. In addition, each of the dimensions listed above has some element of complexity: some projects contribute to more than one sub-programme or are managed by more than one division; the data on funding modality has only been captured on a few named donors; it is assumed that a common definition of a 'global' project is being applied (i.e. any project operating in more than one region or if there is no involvement of specific countries) and alignment to the SDGs has been done retrospectively as projects evaluated in 2018-19 were mostly approved prior to the SDGs being defined. Looking at distribution of project evaluations by sub-programme the cumulative figures for 2010-19 give reasonable sample sizes for Climate Change (CC), Chemicals, Waste and Air Quality (CW), Environmental Governance (EG), Healthy and Productive Ecosystems (HPE) and, marginally, Resource Efficiency (RE). However, project

evaluations for Resilience to Disasters and Conflict (D&C) and Environment Under Review (EUR) are both few in number (it is noted that the first evaluation of the EUR sub-programme is currently underway). The Evaluation Office also notes that the sample size for EG is considerably increased by the numerous GEF Biosafety projects that have been housed under this sub-programme since 2014.

49. Apart from the nature of each sub-programme, with some sub-programmes having a small number of projects in their portfolio and others having many, it is noted that the evaluation policy of key donors also plays a part in the overall number of project evaluations, notably, the HPE sub-programme receives the majority share of GEF grants and the GEF evaluation policy requires a mandatory evaluation, preferably within one year of operational completion. However, in this biennium RE is less well represented than normal and this may be due to the recent completion of the RE Sub-programme evaluation, which was preceded by a period where RE project evaluations were deliberately prioritized. CW is better represented than in the longer time period and HPE accounts for a lower proportion of the project evaluations than over the longer time period.

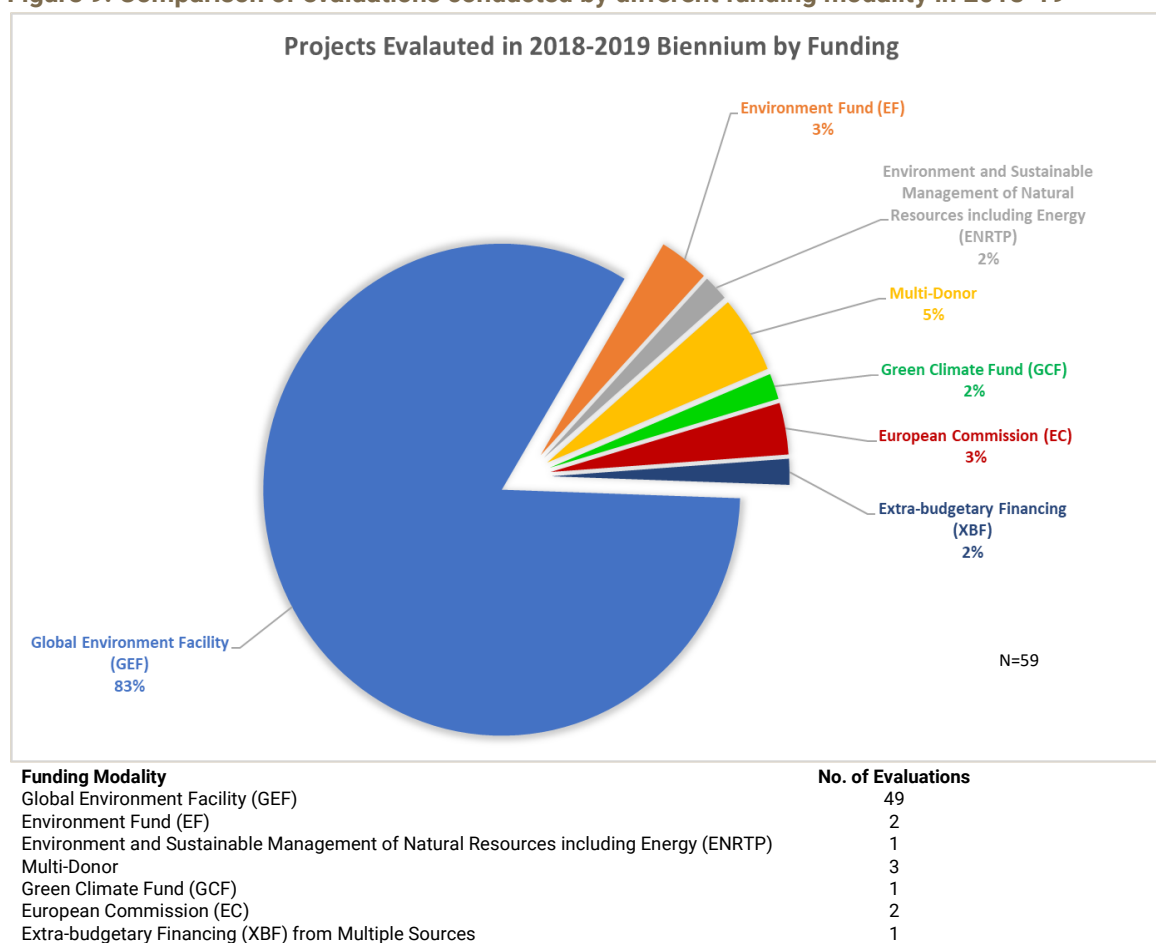
Figure 8. Comparison of distribution of evaluations by sub-programme in 2010-19 and the 2018-19 biennium



Sub-programme	2010-2019	2018-2019
Climate Change	86	16
Chemicals, Waste and Air Quality	28	15
Environment Under Review	3	2
Environmental Governance	62	13
Healthy and Productive Ecosystems	92	11
Resilience to Disasters and Conflict	5	1
Resource Efficiency	15	1

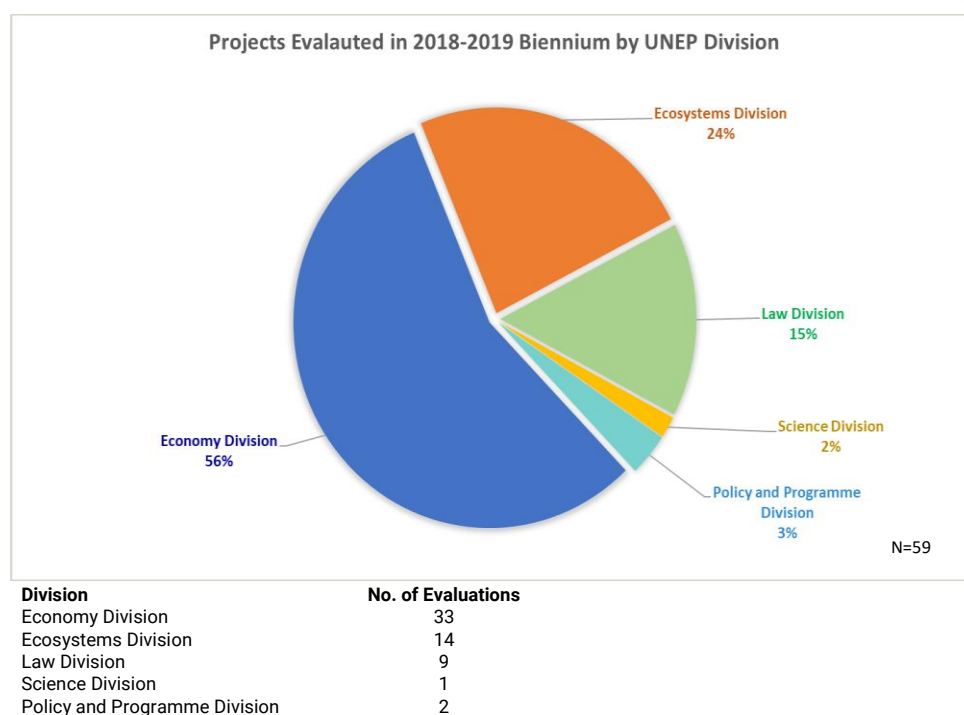
50. A comparison of evaluations conducted by their funding modality puts the effect of the GEF mandatory evaluation policy into stark relief as it accounts for 83% of the project evaluations carried out in the current biennium.

Figure 9. Comparison of evaluations conducted by different funding modality in 2018-19



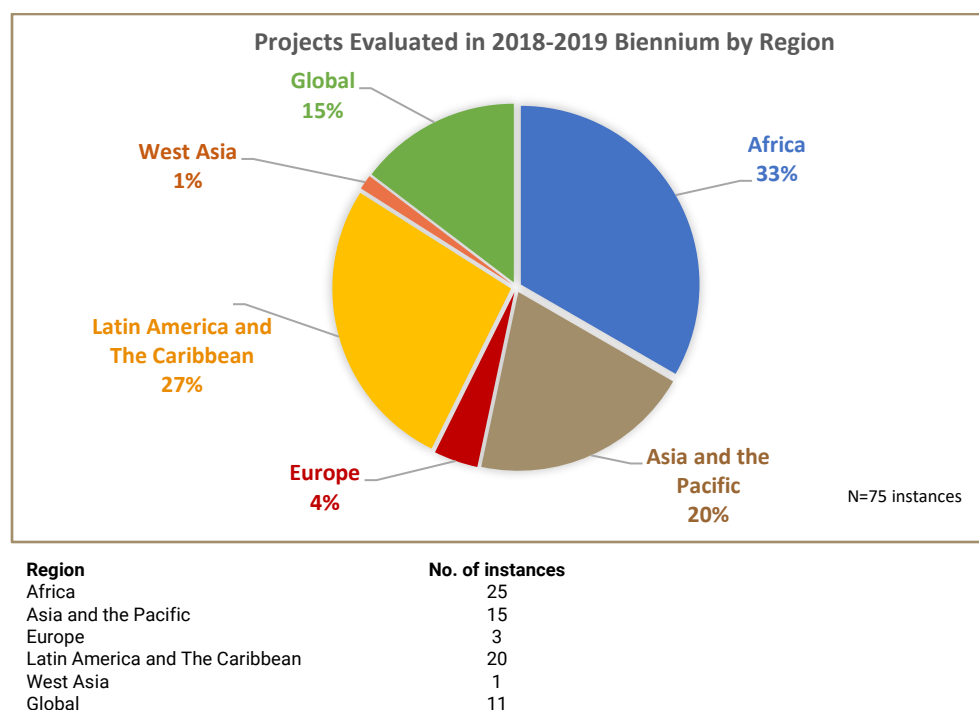
51. A comparison of evaluations conducted by their host Division is presented below (Figure 10). The number of project evaluations in any given time period is a factor of project completion dates but also reflects the nature and composition of the Divisions themselves (e.g. some Divisions have few, large projects while others have multiple Branches and many projects within them).

Figure 10. Distribution of projects evaluated in 2018-19 by UNEP Division



52. As mentioned above, a 'global' project is one that is operating in more than one region. This geographic disaggregation illustrates UNEP's focus on Africa (33%), Latin America and the Caribbean (27%) and Asia and the Pacific (20%).

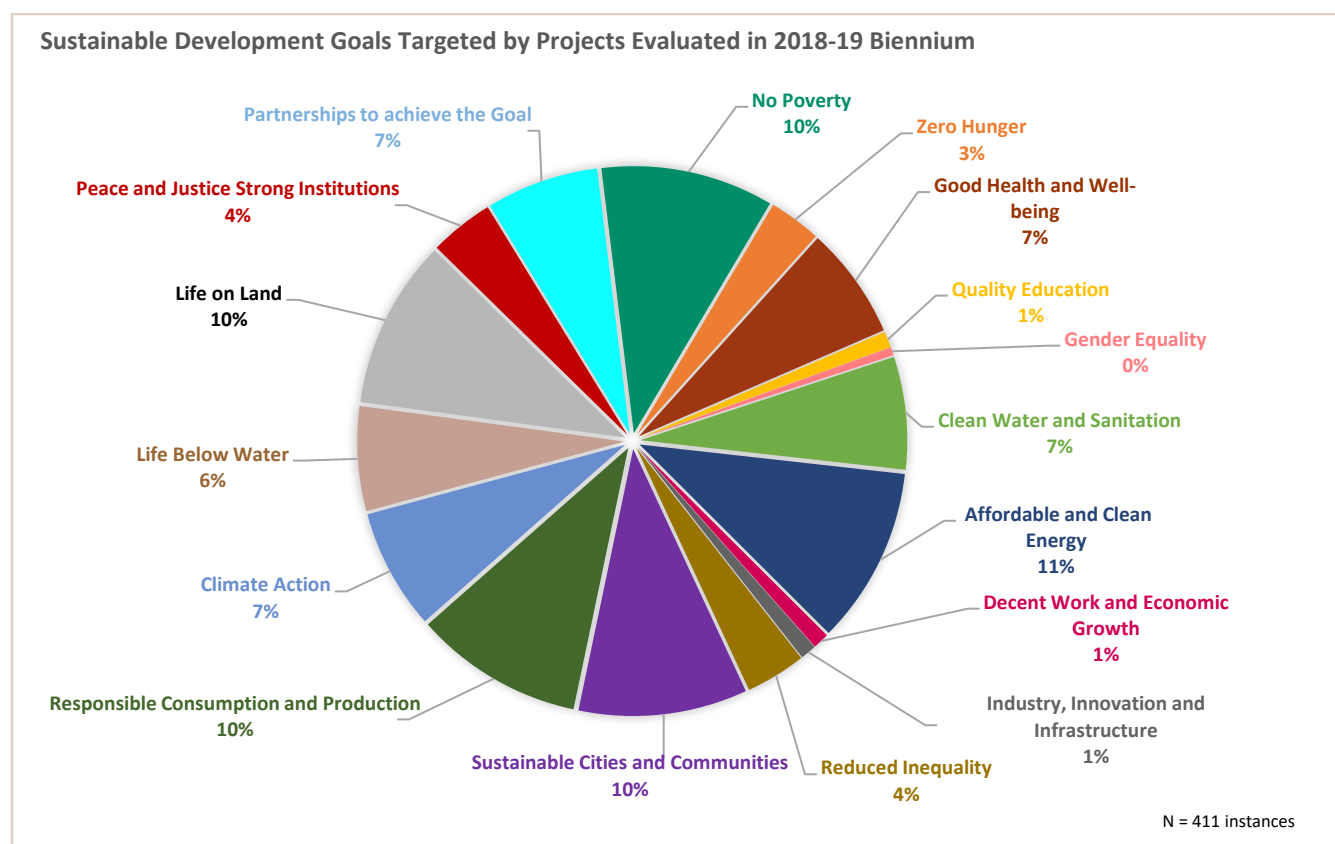
Figure 11. Distribution of projects evaluated in 2018-19 by global region



53. With regard to the link between a project intervention and relevant SDGs, it is emphasized that this alignment has been created retrospectively and that the majority of projects evaluated in this biennium were designed and approved before the SDGs came into play. In addition, where a project has been considered as aligned to more than one SDG, all have been counted. This underpins the high number of instances (411).

54. The house may wish to reflect on whether, and how, it counts alignment with SDGs in the future. In this analysis the SDGs to which project interventions most frequently link are: SDG 7 (affordable and clean energy), SDG 1 (no poverty), SDG 11 (sustainable cities and communities), SDG 12 (responsible consumption and production) and SDG 15 (life on land).

Figure 12. Alignment of projects evaluated in 2018-19 with SDG Targets



Sustainable Development Goal	No. of instances
SDG 1 No Poverty	43
SDG 2 Zero Hunger	13
SDG 3 Good Health and Well-being	28
SDG 4 Quality Education	4
SDG 5 Gender Equality	2
SDG 6 Clean Water and Sanitation	28
SDG 7 Affordable and Clean Energy	44
SDG 8 Decent Work and Economic Growth	4
SDG 9 Industry, Innovation and Infrastructure	4
SDG 10 Reduced Inequality	15
SDG 11 Sustainable Cities and Communities	42
SDG 12 Responsible Consumption and Production	42
SDG 13 Climate Action	30
SDG 14 Life Below Water	26
SDG 15 Life on Land	42
SDG 16 Peace and Justice Strong Institutions	16
SDG 17 Partnerships to achieve the Goal	28

2.4 Sample sizes

55. In the period spanning from 2010 to 2019, UNEP's Evaluation Office completed over¹⁰ two hundred and ninety-two (292) project and programme evaluations. With the current staffing complement the Evaluation Office has the capacity to complete somewhere between 30-35 evaluations a year and this number fluctuates depending on the duration of actual evaluation processes.

56. In the 2018-19 biennium 61 project and programme evaluations were completed, and among these were two strategic evaluations¹¹. The data used for the analysis in this chapter covers those projects for which performance ratings by criteria were included in the evaluation report (i.e. 59 instances of evaluation ratings)¹². At the end of 2019 there were 37 ongoing evaluations carried over into the 2020-21 biennium. Table 4 below shows a breakdown by sub-programme of the number of completed evaluations within the time periods listed.

Table 4. Number of completed evaluations with performance ratings by sub-programme

	Climate Change	Resilience to Disasters & Conflicts	Healthy and Productive Ecosystems	Environmental Governance	Chemicals, Waste & Air Quality	Resource Efficiency	Environment Under Review	Total
POW 2018- 2019	16	1	11	13	15	1	2	59
MTS 2014-2017	30	4	41	35	7	9	1	127
MTS 2010-2013	40	0	40	14	6	5	0	105
2010-2019	86	5	92	62	28	15	3	291¹³

57. The number of projects evaluated in any time period is dependent on the number of projects reaching operational completion within that period. Findings derived from small sample sizes, (less than 20 units in a sample is generally held to be 'too small') cannot be taken to represent the reality of the larger group with any reliability. In some instances, specifically Environment Under Review, Resilience to Disasters and Conflict and Resource Efficiency, the 2018-19 sample includes less than five evaluations per sub-programme. This means that any performance data disaggregated by sub-programme must be treated with extreme caution as it cannot always be generalised to the performance of the sub-programme itself.

2.5 Tools supporting the evaluation process

58. As has been reported in previous biennia (i.e. 2012-13, 2014-15 and 2016-17), the Evaluation Office develops tools and procedures that are intended to help evaluate projects with greater rigor, consistency and objectivity. In doing so, the Office liaises with other parts of the house, e.g. PCA within the Policy and Programme Division, the GEF Coordination Unit etc.) to ensure that the way in which projects are evaluated is consistent with the guidance UNEP provides during the project design and

¹⁰ In some cases, evaluations were completed that did not include an evaluation ratings table and are therefore not included in this total.

¹¹ These are an evaluation of the Resource Efficiency Sub-Programme (2018) and the UN SWAP Review for 2018.

¹² In some instances, programme evaluations and special studies are not conferred any criteria-based performance ratings, so references to the number of completed evaluations may vary within this report.

¹³ The N value for rated projects in the period 2010-19 is actually 292 but by sub-programme we only record 291 because one project in 2016 (TUNZA) was cross-cutting.

development process. Such liaison between departments also contributes to an ongoing process of institutional sharing and learning.

59. A significant number of evaluation tools were developed during the 2016-17 biennium and these have been tested and subsequently revised during 2018-19. In particular, the Evaluation Criteria Ratings Matrix that provides descriptions of the performance levels and evidence requirements (e.g. 'Highly Satisfactory' or 'Moderately Unsatisfactory') was applied in most of the evaluations during 2018-19 and has recently been revised to incorporate suggestions for improvement from both evaluation consultants and project teams.

2.6 Learning from evaluations - recommendations tracking and coding evaluation reports

60. The dual purpose of evaluations is to meet accountability requirements and also, importantly, to contribute to learning. Evidence that learning has been taken up is provided by an assessment of compliance with the recommendations that are derived from evaluations. During 2019 the Evaluation Office has developed a revised process to a) support the development of recommendations based on the findings from project evaluations and b) to track the uptake of associated actions at both project and institutional levels. This tool takes full advantage of UNEP's adoption of SharePoint as an institution-wide platform and can be used to support more in-depth discussions around desired actions and the kind of evidence that will meet compliance reporting needs. This approach has been introduced in January 2020 and will be applied to all projects evaluated during the 2020-21 biennium.

61. In addition to the lessons learned and recommendations developed during an evaluation, the reports themselves contain considerable insights into different aspects of performance. During 2019 the UNEP Evaluation Office began the process of coding the texts of evaluation reports so that searches can be performed on topics of interest and value to the house. In this report, initial findings on Monitoring and Reporting (Box 3, Box 4, Box 5), Financial Management (Box 1, Box 2) and Responsiveness to Gender (Box 6) are included based on this emerging approach. Together with the recommendations' development and tracking tool mentioned above, the coding of evaluation report findings is intended to strengthen the capacity of the Evaluation Office to systematically share learning with targeted audiences.

2.7 Limitations

62. In order to fulfil its mandate of evaluating projects forming the Programme of Work the UNEP Evaluation Office is reliant on information regarding the operational completion of projects from all organizational units implementing them, and the provision of an adequate evaluation budget. As neither of these features are within the direct control of the Evaluation Office there is a limit as to how far the Office can apply a more strategic sampling approach to interrogate the Programme of Work as a results-delivering framework.

63. Given the manual nature by which the list of projects eligible for evaluation is compiled, there is a possibility of a positive bias in the projects the Evaluation Office evaluates. This is because one can logically assume that being able to complete a project on time and making sure an adequate evaluation budget is available by the time it is needed are both characteristics of good project management. One could expect that strong project management skills are associated with stronger project performance.

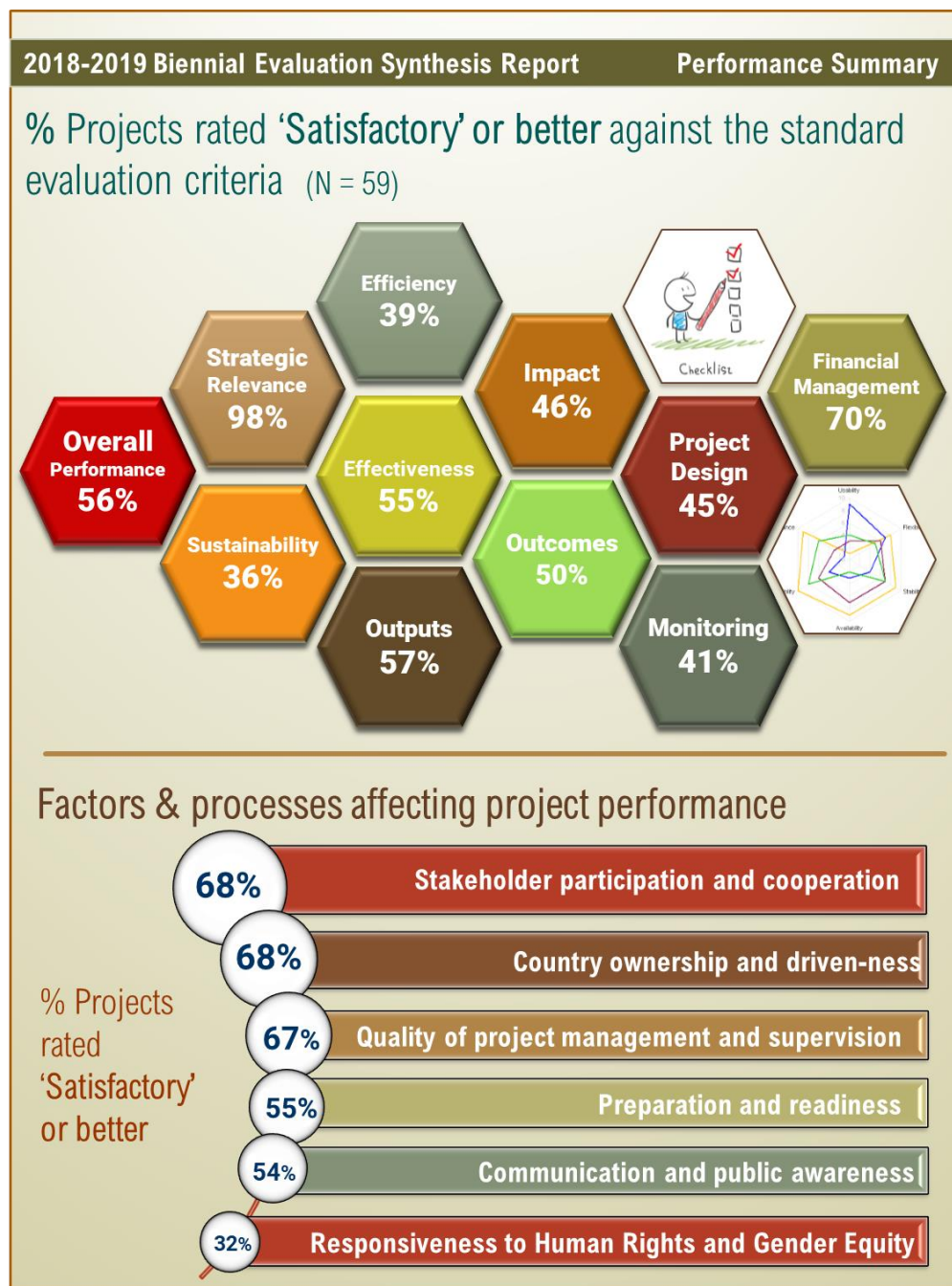
64. While the overall sample size (61 project evaluations in the current biennium) forms a robust basis for learning lessons, when this sample is disaggregated by dimensions of interest to the house (e.g. sub-programmes, or projects in Africa etc.), its reliability for providing insights is low.

3 Analysis of project performance

3.1 Performance in the 2018-19 biennium

65. Figure 13 below provides a general view of performance by projects evaluated in the 2018-19 biennium, with a focus on the percentage of projects rated 'Satisfactory' or better (S/HS) in selected evaluation criteria.

Figure 13. Overview of project performance by criteria in the 2018-19 biennium

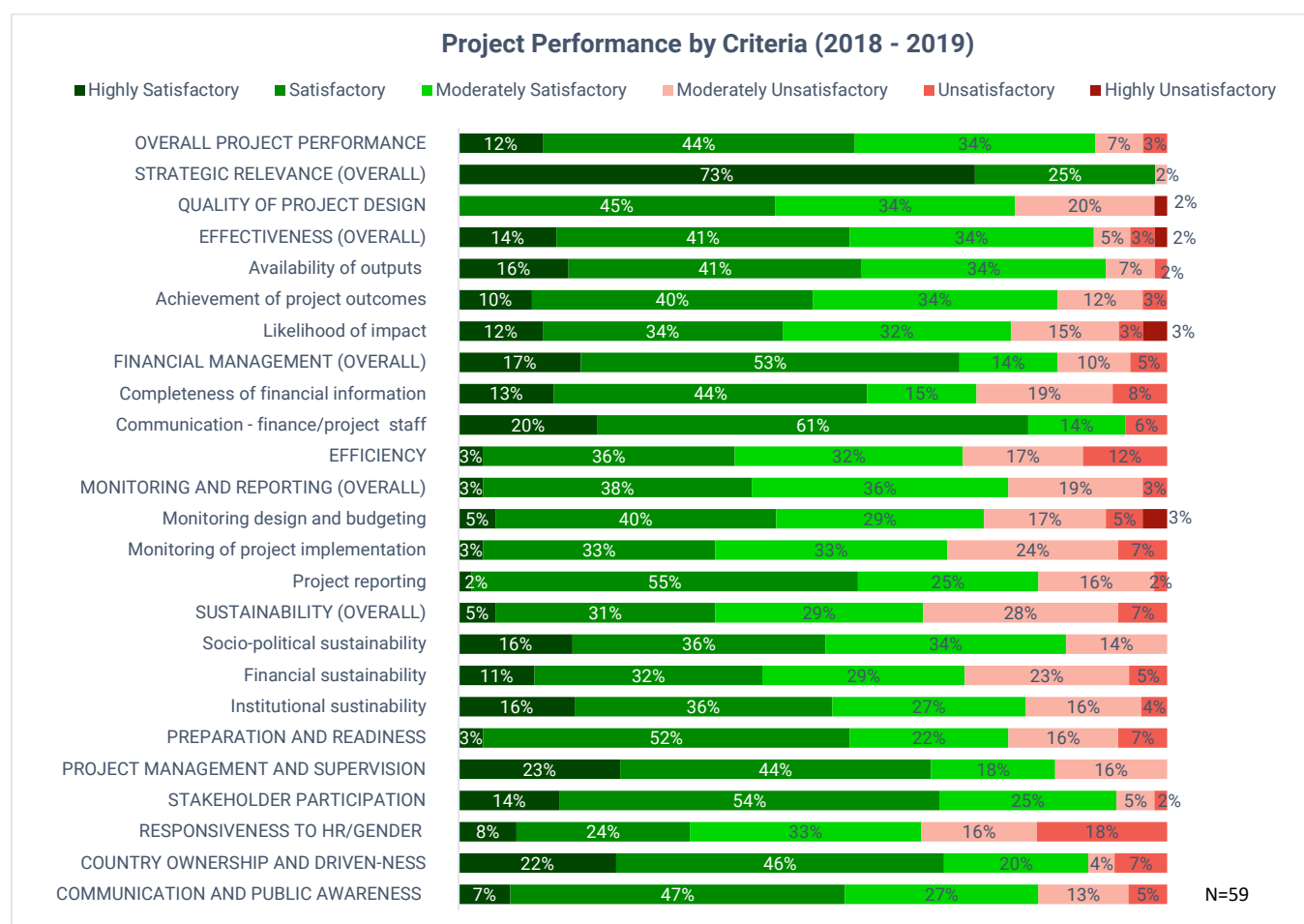


66. As mentioned above, para. 45, the UNEP Evaluation Office sets the desired performance level at 'Satisfactory' (i.e. 83 % on a six-point ratings scale) and focuses its attention on the percentage of cases where ratings have been awarded at the levels of 'Satisfactory or better' (i.e. either 'Satisfactory' or 'Highly Satisfactory'). This is abbreviated throughout this report as S/HS.

67. Although the Evaluation Office has made improvements to the descriptions of evaluation criteria over time made adjustments to improve the visibility of key areas of performance, introduced concepts that are of importance to UNEP and improved the rigour of performance assessments, summarized in Table 2 above, it remains relevant to compare performance against an individual criterion (or their sub-categories) from one biennium to another. However, the number of changes makes a visual representation of 2018-19 figures side by side with those from 2016-17 inappropriate and potentially misleading. Comparisons over time are reported in this chapter with some caution and with consideration of whether other substantive knowledge of general evaluation findings supports the interpretation that the percentage figures suggest. Where the percentage change figures are not supported by wider evaluation findings, this inconsistency is noted, or points to bear in mind are given when making comparisons over time.

68. Figure 14 provides an aggregated view of project performance across all six rating categories by criteria, based on data collated from evaluation reports completed in the 2018-19 biennium.

Figure 14. Performance against all evaluation criteria in projects evaluated during the 2018-19 biennium



69. Data (specifically, ratings for S/HS performance) for the evaluation criteria performance scores from 2016-17 was compared to the scores from 2018-19 in order to: a) to gain some insight on whether there were differences in the proportion of projects that reached a satisfactory level on evaluation criteria when comparing those evaluated in different biennia, and b) to reflect on what could be causing any significant changes.

70. The null hypothesis is: There is no change in the percentage of projects with satisfactory or highly satisfactory scores from 2016-17 to 2018-19.

71. The level of satisfactory achievement on each of 10 evaluation criteria and 6 factors that might affect outcomes was measured on 6-point scales. The two highest levels (5 and 6) were combined to indicate satisfactory or highly satisfactory. The remainder of the levels (1 through 4) were combined to indicate less than satisfactory. The percentage of projects achieving the satisfactory or higher rating in each time period was calculated for each criterion. The statistical significance of differences in these percentages was calculated using the chi-square statistic. Conclusions based on the comparisons between time periods must take into consideration two factors: (1) the projects evaluated in the two periods are not equivalent and may differ, for example, in the likelihood of achieving a satisfactory level; (2) clarity of criteria for rating the levels of satisfaction were improved for the 2018-19 evaluations and may have resulted in different views of what scale values meant.

72. In most cases the changes in % scores are not statistically significant. This suggests that although comparable numbers of UNEP projects are reaching satisfactory levels of performance for these two time periods, there is also no major improvement. Where there are signs of changed performance (i.e. availability of outputs, financial management and preparation and readiness), the cause of the change is not yet clear because of improvements that the Evaluation Office has been making during these two biennia in the list of stand-alone criteria; clarity of criteria descriptions and the boundaries and characteristics of each category of the ratings scale. However, having established this methodology in this biennium, in subsequent years and biennia it will be possible to provide comparisons that can be more directly related to changes in performance.

73. Table 5 below summarises the findings of the statistical analysis:

Table 5. Comparison of S/HS performance between 2016-17 and 2018-19 biennia

Evaluation Criteria/sub-category	S/HS performance 2016-17 and 2018-19	Chi-square (p-value)	t-test (p-value)	Difference over time
Overall Project Performance	Performance: 67% in 2016-17; 56% in 2018-19 Some of this change is likely to be due to the cumulative effects in changes in the definition of evaluation criteria. In 2020-21 a better comparison over time will be possible.	0.181	Not calculated	Change not significant
Strategic Relevance	Remains the highest performing criterion, currently at 98% compared to 96% in 2016-17 The sustained high performance (over 90%) here does raise the	0.459	0.439	Change not significant

Evaluation Criteria/sub-category	S/HS performance 2016-17 and 2018-19	Chi-square (p-value)	t-test (p-value)	Difference over time
	question of the utility of this evaluation criterion in terms of institutional learning.			
Quality of Project Design	Performance: not measured in 2016-17; 45% in 2018-19 Newly introduced within the ratings table, no previous biennium data			Not calculated
Effectiveness	Performance: 52% in 2016-17; 55% in 2018-19	0.662	0.665	Change not significant
Availability of outputs	Performance: 79% in 2016-17 to 57% in 2018-19 Some of this decrease is likely to be due to the introduction of the Ratings Matrix that describes the characteristics underpinning each of the points on the six-point rating scale.	0.005	0.006	Significant change (decrease)
Achievement of outcomes	Performance: 53% in 2016-17; 50% in 2018-19	0.758	0.760	Change not significant
Likelihood of impact	Performance: 41% in 2016-17; 46% in 2018-19	0.563	0.567	Change not significant
Financial Management	Performance: 47% in 2016-17; 70% in 2018-19 Some of this increase may be due to creating a stand-alone criterion on financial management, articulating sub-categories and providing clear guidance on ratings in the Ratings Matrix	0.005	0.005	Significant change (increase)
Efficiency	Performance: 49% in 2016-17; 39% in 2018-19	0.228	0.230	Change not significant
Monitoring and Reporting	Performance: 31% in 2016-17; 41% in 2018-19 The introduction of 3 sub-categories under this criterion allowed for good performance in project reporting to be captured as well as recording the lower performance in monitoring project implementation.	0.220	0.227	Change not significant
Sustainability	Performance: 39% in 2016-17; 36% in 2018-19	0.726	0.728	Change not significant
Factors Affecting Performance/Cross Cutting Issues				

Evaluation Criteria/sub-category	S/HS performance 2016-17 and 2018-19	Chi-square (p-value)	t-test (p-value)	Difference over time
Preparation and readiness	Performance: 37% in 2016-17; 55% in 2018-19	0.037	0.038	Significant change (increase)
Quality of project management/supervision	Performance: not measured in 2016-17; 67% in 2018-19 The criterion was spread over two criteria in previous biennia and the present criterion ratings only began in 2018.		-	Not calculated
Stakeholder participation and cooperation	Performance: 72% in 2016-17; 68% in 2018-19 Given the broad definition of stakeholders used by UNEP, there can be a wide-ranging interpretation of the actual nature of participation under this sub-category.	0.608	0.613	Change not significant
Responsiveness to human rights and gender equality	Performance: not measured in 2016-17; 32% in 2018-19 Newly introduced within the ratings table, no previous biennium data			Not calculated
Country ownership and driven-ness	Performance: 61% in 2016-17; 68% in 2018-19	0.401	0.401	Change not significant
Communication and public awareness	Performance: 60% in 2016-17; 54% in 2018-19	0.611	0.615	Change not significant

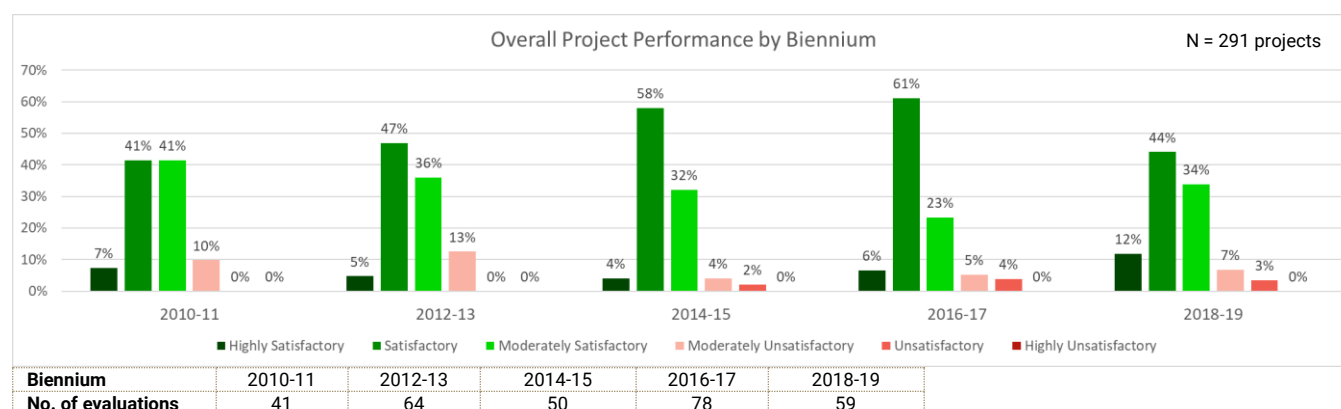
74. General trends over a longer period, back to the 2010-11 biennium, can now be reflected upon, see Figure 15 below.

75. One can summarize to say that, over the past nine years, UNEP's projects have maintained steady performance within the 'Satisfactory' range (i.e. Moderately Satisfactory, Satisfactory and Highly Satisfactory). There is a strong pattern of the majority (40 – 60%) of projects reaching the fifth point on the six-point scale, i.e. the 'Satisfactory' level, which represents the performance threshold applied by the UNEP Evaluation Office. The proportion of projects being assessed at the 'Highly Satisfactory' level have noticeably increased¹⁴ in the current biennium although this is not statistically significant¹⁵.

¹⁴ There is nothing in the way that evaluation criteria have been adjusted over time, nor the way in which points on the ratings scale have been characterised in the Evaluation Criteria Ratings Matrix that would obviously bring about this increase. The only factor that might be at play here, is greater consistency among evaluation consultants and within the Evaluation Office team in interpreting what 'Highly Satisfactory' "looks like".

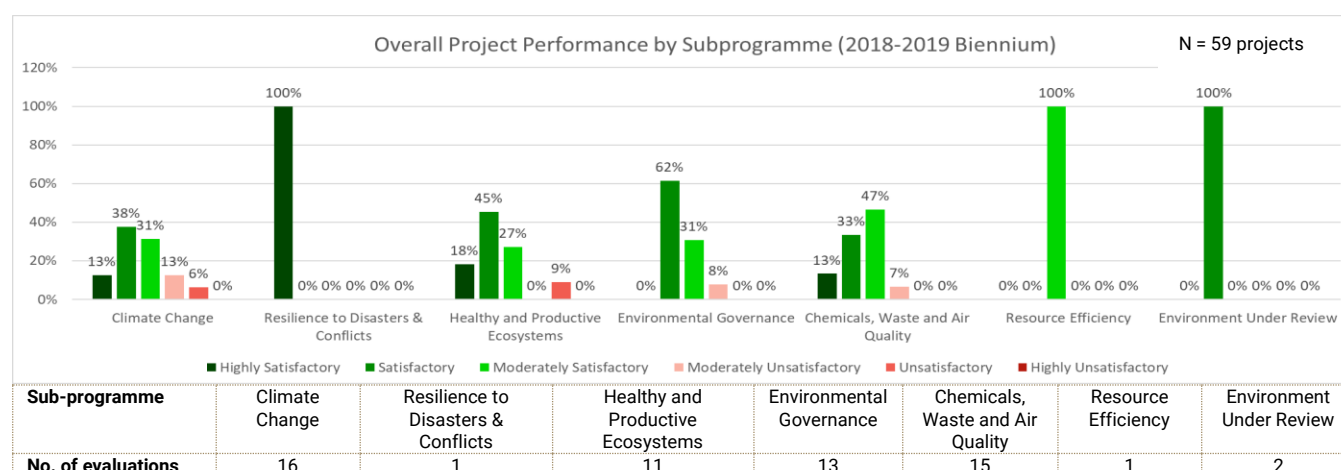
¹⁵ Using the Mann-Whitney U test with a null hypothesis of no difference in ratings between biennia.

Figure 15. Comparison of overall project performance by biennium



76. While these data are presented by sub-programme for the purpose of transparency and deeper learning, the distribution of the sample, with fewer than five projects having been evaluated with the Resilience to Disasters and Conflict (D&C), Resource Efficiency (RE) and Environment Under Review (EUR) sub-programmes during this biennium, limits its utility.

Figure 16. Comparison of overall project performance by sub-programme in 2018-19



3.2 Strategic relevance

77. Strategic relevance remains the evaluation criterion with the highest proportion of S/HS ratings. Even in the earlier biennia of 2010-11 and 2012-13 the percentage of projects with S/HS ratings was above 90% and since the 2014-15 biennium this percentage has steadily increased (88% in 2014-15; 96% 2016-17 and 98% in 2018-19).

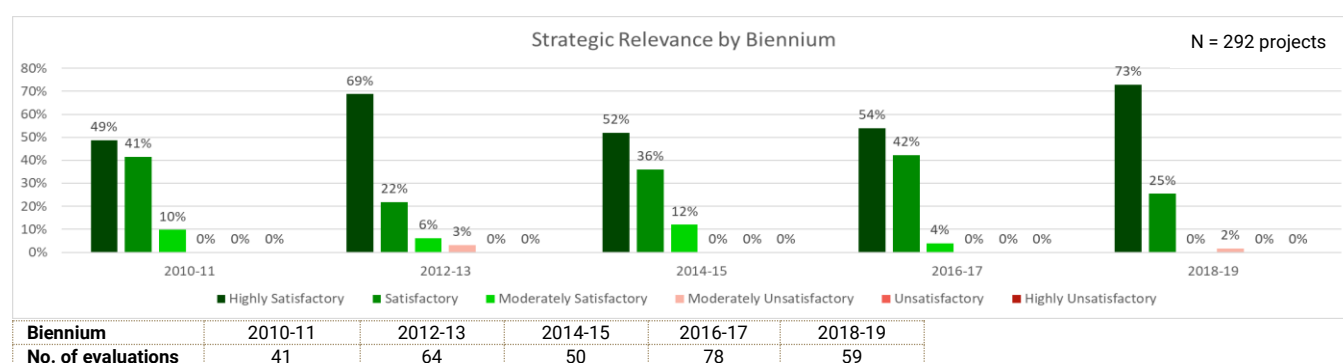
78. However, a number of issues need to be considered. Firstly, this criterion assesses 'alignment' with strategic frameworks and priorities at the levels of UNEP, donors, regions and countries as well as the complementarity between UNEP's project and other existing/planned interventions. Consistent alignment is, therefore, largely a function of various approval processes. If the criterion were to look at

the ‘substantive contribution’ being made by projects under evaluation to the strategic frameworks and priorities to which they are aligned, the performance would look considerably different.

79. Perhaps more importantly, this is a ‘soft’ indicator of performance and the continued high performance levels are currently offering very little learning to the house. Going forwards, the criterion may prove to be more useful when UNEP has articulated a sharper definition of priorities and strategic focus.

80. One project in the 2018-19, notably, had a rating of ‘Moderately Unsatisfactory’ against this criterion. A closer examination of the evaluation report shows that the project did not work in conjunction with, or liaise closely with, other relevant UNEP projects, nor with other similar projects in the region (e.g. DFID funded projects and a large number of related GEF funded projects).

Figure 17. Comparison of strategic relevance (overall) by biennium



3.3 Quality of project design

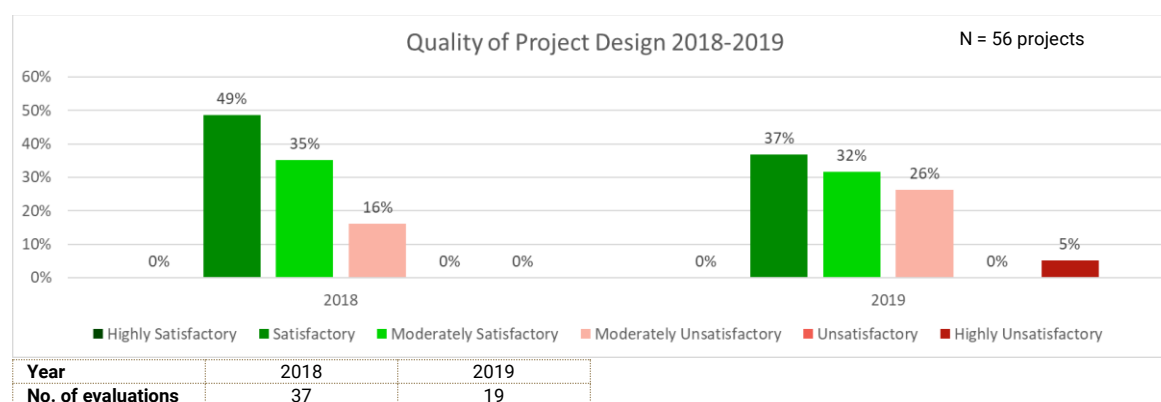
81. The criterion Quality of Project Design was incorporated into the evaluation ratings table in 2017 and therefore performance against the criterion cannot be compared over biennia. The quality of project designs has been assessed by the Evaluation Office since 2011 and documented in evaluation Inception Reports so there is more knowledge within the Evaluation Office about the strengths and weaknesses of project designs, than these data suggest. Common weaknesses in project design across projects were also highlighted through a lessons learned framework first published in 2007 and subsequently updated in 2017¹⁶.

82. Looking at the data for each of the two years within the biennium the pattern of distribution is similar with the spread of ratings concentrating around the ‘Satisfactory’ and ‘Moderately Satisfactory’ levels. A further exploration of the relationship between quality of project design and effectiveness is provided in para 84, below.

83. In 2019 one project was evaluated which was rated as having a ‘Highly Unsatisfactory’ project design (i.e. representing 5% of the 2019 sample). This project was designed in 2010 as a merging of two existing projects, but no guiding project design document that articulated the proposed outputs and outcomes was produced. The scope of work for the newly designed project was limited to brief lists of key activities in the Memoranda of Understanding between UNEP and the funding partner.

¹⁶ <http://hdl.handle.net/20.500.11822/184>

Figure 18. Comparison of quality of project design in 2018 and 2019



3.3.1 Relationship between Quality of Project Design and Effectiveness Ratings

84. UNEP's interest lies in strengthening its results. From an evaluation perspective this means, in the first instance, improving performance ratings against Effectiveness criteria, especially in the achievement of outcomes. A member of the evaluation team undertook a study into whether the quality of project design (as reflected in the ratings awarded by evaluation consultants during the evaluation process) could be said to influence Effectiveness ratings (both at an overall level and also at the levels of outputs, outcomes and the likelihood of impact).

85. The study found, based on a sample of 49 project evaluations carried out in 2018 and 2019, that there was no statistical correlation between quality of design and any aspect of effectiveness. What this suggests is that there are several factors influencing effectiveness and that the quality of project design is not, on its own, a strong enough factor to determine the level of effectiveness.

86. Although no statistical correlation could be found, frequency patterns show that the predictive power of quality of project design ratings is stronger in relation to outputs and gradually weakens in relation to the likelihood of impact (i.e. in 86% of cases ratings at output level are the same or better than the quality of project design rating; in 82% of cases the ratings at outcome level are the same or better than the quality of project design rating and for the likelihood of impact, 71%).

87. In order to gain further insight into other influential factors the study looked at exceptional cases, particularly those few where the rating for the quality of project design was low but ratings for effectiveness were relatively high and vice versa. It found that political will and country ownership are critical factors. In one case where the project design lacked a complete logical framework, had no theory of change, no stakeholder analysis etc. the government was a strong driving force throughout the project and strong results were achieved. On the other hand, a project with a solid design failed to realise the expected level of effectiveness largely because the project's ambitions were a poor match for the country's capacity and there was instability among, and a lack of coherent support from, government parties.

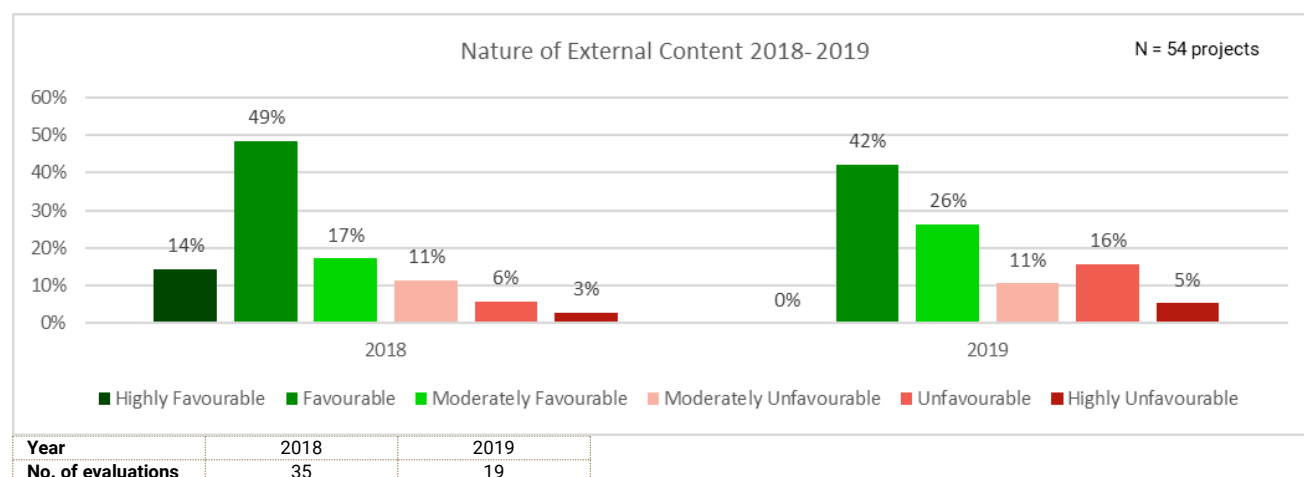
3.4 Nature of external content

88. This criterion was introduced in 2017 and differs from other criteria. This criterion is used, at evaluation inception stage, to establish the stability of the project's external operating context, considering the prevalence of conflict, natural disasters and political upheaval during the project's

implementation period. We note that these are unexpected and substantive factors that may disrupt a project's delivery. So, for example, 'political upheaval' does not include regular national election cycles, which should be factored into the project's design and timeline but would capture any unanticipated unrest or prolonged disruption surrounding a national election process. Where a project has been rated as having encountered either an 'Unfavourable'¹⁷ or 'Highly Unfavourable' external operating context, and/or a negative external event has occurred during project implementation, the ratings for Effectiveness, Efficiency and/or Sustainability may be increased at the discretion of the Evaluation Consultant and Evaluation Manager together. This allows the evaluation to still have utility, and to be able to give a project 'credit' for what it has achieved, even if the planned project design or implementation plan could not be followed as expected.

89. The only data available to-date are for 2018 and 2019, see below. The proportion of projects encountering such disruptive external factors (20% in the Unfavourable range in 2018 and 32% in that range in 2019) might provide the house with an alternative perspective on some elements or risk. In the future, as more projects that were subject to UNEP's safeguarding assessment at the time of project approval reach their operational completion, the Evaluation Office will be able to make comparisons with Environmental, Social and Economic safeguard assessments.

Figure 19. Comparison of nature of external context in 2018 and 2019



3.5 Effectiveness

90. Assessments of the effectiveness of UNEP's work is at the core of any analysis of performance. In the assessment of overall project performance this criterion, made up of three sub-categories (availability of outputs; achievement of outcomes and likelihood of impact), account for 45% of the overall score (5%, 30% and 10% respectively). Here we consider the overall ratings for Effectiveness, followed by discussions of each of the fundamental sub-categories.

91. The proportion of evaluated projects with S/HS ratings is surprisingly constant over the biennia (59% in 2010-11; 55% in 2012-13; 55% in 2014-15; 52% in 2016-17 and 55% in 2018-19). In the 2018-19

¹⁷ For this criterion a similar six-point rating scale is used by the labels are changed to: Highly Unfavourable; Moderately Unfavourable; Unfavourable; Moderately Favourable; Favourable and Highly Favourable. The two lowest rating points in the Unfavourable range trigger may lead to an adjustment in the assessment of other performance criteria.

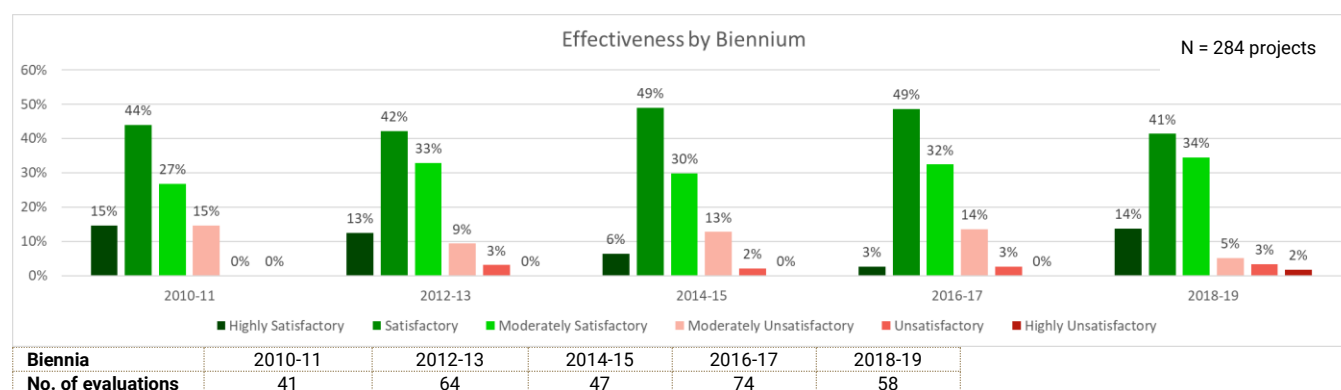
biennium we also see the split in the proportion between 'Satisfactory' and 'Highly Satisfactory' performance move in favour of higher performance.

92. Underlying these data may be more substantial changes in this area, especially in the 2018-19 biennium, due to adjustments made to evaluation criteria, the description of characteristics of different ratings levels and the use of a weighted table to aggregate sub-categories as well as to determine the overall project performance score.

93. All of these developments have brought more rigour and consistency to the evaluation ratings process. It is noted that in 2018-19 the distribution of ratings across the Satisfactory range shows a higher proportion of projects with 'Highly Satisfactory' ratings. Without wanting to detract from good project performance, some of this increase may be due to the development and use of the Evaluation Criteria Ratings Matrix, although the introduction of clearly specified standards of evidence might, a priori, have been expected to drive the distribution of ratings in the opposite direction.

94. While the constancy of this performance on overall effectiveness may, on the one hand, be encouraging, one could also argue that this is the area where UNEP would wish to see indications of faster improvement.

Figure 20. Comparison of effectiveness (overall) by biennium



3.5.1 Availability of outputs

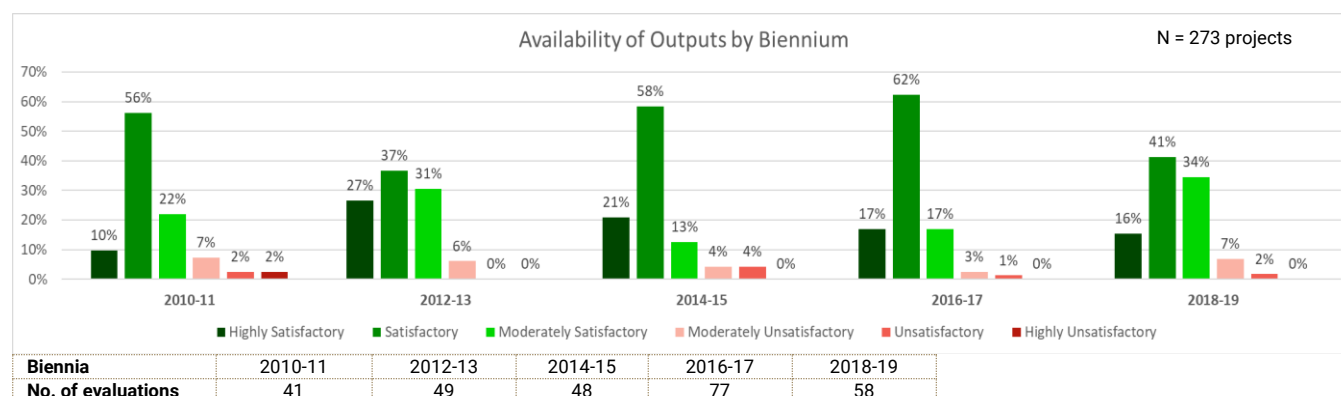
95. Outputs are typically the level at which evaluations start a project's Theory of Change (TOC). This is because it is at the output level that a project's intended beneficiaries or users come into view and a potential change process can begin to be discerned. Up until the output level the focus during project implementation is very much on development and completion of activities, all of which are under the strong control of implementing and executing parties.

96. The importance of project outputs actually reaching intended beneficiaries or users is reflected in considerations of access to, or the availability of the products and services delivered by a project. For example, if a project has worked to support a partner to review and revise a technical manual but the final copy of the manual is only available on the laptop of the Project Manager then it is not yet possible to consider any potential change process. Once that revised technical manual is in the hands of those who are intended to use it, even if they have not yet used it, it becomes possible to identify the people who are expected to take part in a change process or contribute to the progression of a causal pathway.

97. Typically, and as evidenced by the high proportion of S/HS ratings for this sub-category of Effectiveness, UNEP projects perform well at output level. Over the biennia the percentage of S/HS ratings has varied from 66% in 2010-11 to 89% in 2016-17. Underlying the level in 2018-19, (57%), may be a more demanding understanding of outputs from a results perspective. Prior to 2018 evaluations were more likely to give greater credit to the completion of activities (e.g. 'manual revised').

98. In addition, through the work of developing an Evaluation Criteria Ratings Matrix, evaluations since 2018 have more clearly and consistently considered the following elements in the provision of outputs: proportion of planned outputs made available; timeliness and suitability for purpose of those outputs most important for the achievement of outcomes; quality of the outputs and levels of ownership of the outputs by the intended beneficiaries or users. This more holistic perspective may have been more difficult to achieve in this biennium but it is hoped that such feedback will help to 'raise the bar' for UNEP's project-level performance.

Figure 21. Comparison of availability of outputs by biennium



3.5.2 Achievement of project outcomes

99. At the outcome level evaluations assess: levels of achievement (e.g. partial or full) of planned outcomes; whether achievements have been evidenced in those outcome areas most important for intermediate states to emerge and whether the conditions that contribute to the intended change process (i.e. 'assumptions' and 'drivers') have been seen to hold.

100. The proportion of S/HS ratings for project outcomes is more variable than that at output level and typically noticeably lower in each biennium: S/HS ratings under this sub-category were 55% in 2010-11; 61% in 2012-13; 49% in 2014-15; 53% in 2016-17 and 50% in 2018-19. A coarse calculation of the average S/HS ratings over time for both sub-categories shows that outputs perform at an average of 73% on S/HS ratings and outcomes at 54%.

101. The proportion of projects being assessed at the 'Highly Satisfactory' level in this biennium is at 10%. Although this does not mean, necessarily, that 'more' projects are being rated as Highly Satisfactory for achievement of outcomes in 2018-19 than in 2016-17, it does mean that the distribution during this one biennium is more towards this high level of achievement than in the previous biennium.

102. However, project outcomes are defined as those outcomes that a project is expected to achieve at the end of its timeframe and within its secured budget. It is the highest level of result to which

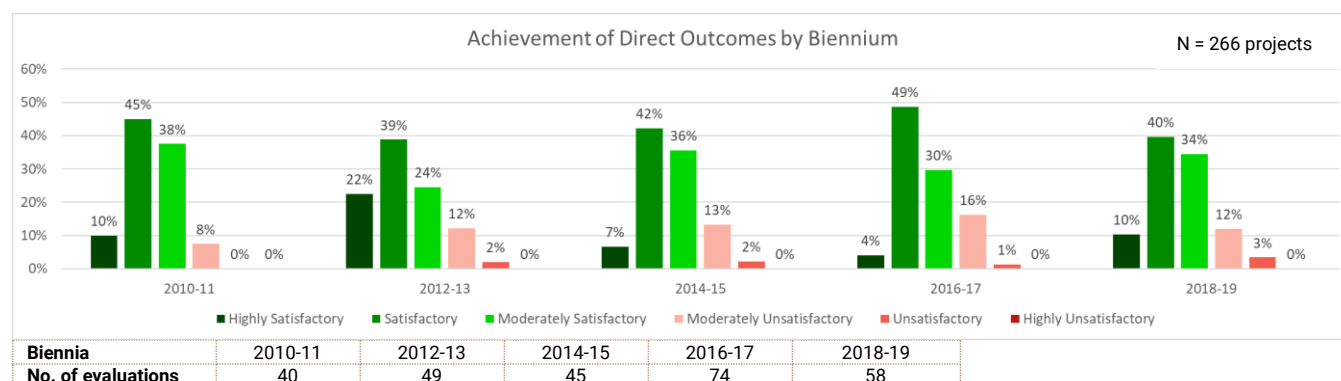
evaluations hold projects to account. Performance at outcome level is, therefore, the clearest indication of the direct and short-term effect of UNEP's work.

103. Outcomes are also the points in the change process at which agency¹⁸ is clearly transferred from those who are responsible for delivering the project to those who are expected to benefit from, or use and take forwards, a project's results. For a more visual analogy, between outputs and outcomes the relay baton should be passed from one team member to another. Only when those who are not dependent on a project's inputs have the capacity to move forwards on the intended change trajectory, can long-lasting effects be realized. The drop of in S/HS ratings between the availability of outputs and achievement of outcomes is, therefore, an area that requires attention.

104. Some of the challenges noted through evaluations in this area are:

- outcome level results statements are poorly formulated, often without the inclusion of a verb that denotes use, application or uptake;
- the ambitions implicit in outcome level results are often too high, or unrealistic, for the timeframe, budget and scope of work of the project, but may have seemed impressive during approval processes;
- measures to show changes in behaviour (i.e. use of outputs) from before and after project implementation are either not envisaged in the project design or not tracked during implementation;
- no, or few, activities are designed and carried out between output and outcome level. Typically project designs, logical frameworks, budgets and even TOCs are formulated with activities only feeding into the output level but the activities needed to ensure the uptake or use of those outputs are either not considered or are left at an implicit level; and
- Theories of Change are still designed with more regard to how a project will be operationalized (e.g. components, work packages, areas of activity etc.) than as a set of interlinked and interdependent causal pathways.

Figure 22. Comparison of achievement of outcomes by biennium



¹⁸ 'agency' - definition "the capacity, condition, or state of acting or of exerting power". Merriam-Webster.

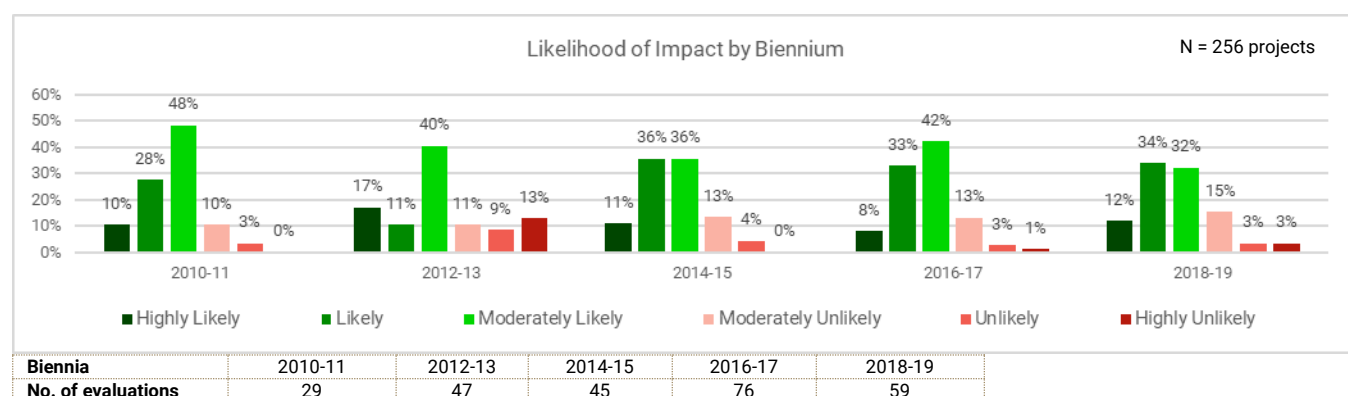
3.5.3 Likelihood of impact

105. It is recognized that impacts, by definition long-lasting and therefore stable changes of state, are not under the control, or even sole influence, of a project. As evaluations are undertaken as soon after operational completion as is possible, it is often too early to gather clear evidence of impact. Evaluations, therefore, assess the *likelihood* of impact. The approach taken is based on the TOC and considers the assessment of the achievement of outcomes combined with assessments of whether contributing conditions have held and whether there are any nascent signs of long-lasting (i.e. embedded) changes developing, or having developed.

106. The proportion of higher performance levels against this sub-category of Effectiveness has been relatively constant since 2014-15: L/HL¹⁹ ratings were at 47% in 2014-15; 41% in 2016-17 and 46% in 2018-19.

107. Similar to achievement of outcomes, the proportion of projects reaching high performance levels against this criterion is higher (12%) in 2018-19 than in 2016-17 (8%). There is no known change in evaluation methods or definitions etc. that could be responsible for changes in performance against this criterion, other than increased consistency in the way the rating is established.

Figure 23. Comparison of likelihood of impact by biennium



3.6 Financial management

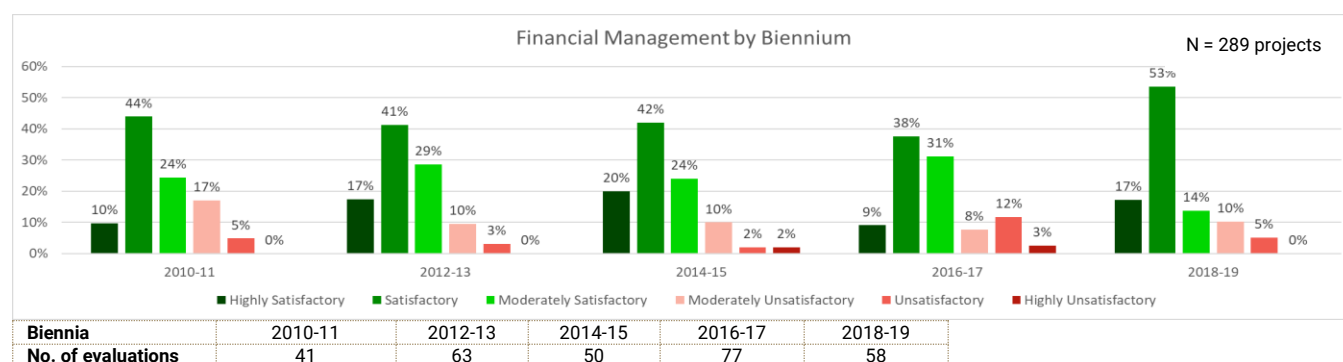
108. Up until the 2016-17 biennium, financial management was assessed as a sub-category of Factors Affecting Performance. Apart from having noticeably lower S/HS ratings than other criteria the amount and depth of evaluative insights were limited. This may have been partially due to a lack of clarity and guidance given to external evaluation consultants on how to assess performance in this area. In 2016 financial management was established as a stand-alone evaluation criterion and two sub-categories were defined: completeness of financial information and communication between project and financial project team members. At first a third category on 'compliance' was proposed but the boundaries between evaluation and audit on financial matters became blurred and the sub-category was not continued. In late 2019 this area was re-visited, with the benefit of further experience of examining this

¹⁹ Likelihood of Impact and Sustainability are both rated against a six-point rating scale, labelled as: Highly Unlikely; Unlikely; Moderately Unlikely; Moderately Likely; Likely; Highly Likely. Abbreviations are therefore: HU, U, MU, ML, L and HL. Likely (L) and Highly Likely (HL) equate to Satisfactory (S) and Highly Satisfactory (HS).

area of performance, and a third sub-category has been introduced to assess ‘adherence to UNEP’s financial policies and systems’. This will be included in the 2020-21 biennial synthesis report.

109. The proportion of S/HS ratings under this criterion have fluctuated over time: 54% in 2010-11; 58% in 2012-13; 62% in 2014-15; 47% in 2016-17 and reaching an all-time high of 70% in 2018-19. Although this criterion does have a comparatively high proportion of ratings at ‘Highly Satisfactory’ in the past, the current level of 17% is worthy of note as an area of strong performance. It is likely that the fluctuation in ratings on financial management is linked to the introduction of Umoja. Although projects are designed several years before evaluation, the daily financial management of projects is sensitive to the real-time financial systems operating within the house. As Umoja was introduced in 2015, it is likely that the low S/HS ratings in 2016-17 reflects some of the effect of the disruption to normal services and the learning curve needed to adapt to the new system.

Figure 24. Comparison of financial management (overall) by biennium



110. For this Biennial Synthesis Report, brief studies were carried out on three criteria that have previously shown weak performance, one of which is financial management. Text-coding software was used to explore, categorize and analyse the findings on this topic from twenty-four (24) evaluation reports completed during 2018-19.

Box 1. Summary of 2018-19 text analysis findings: Completeness of financial information

- Financial data often have to be combined, at evaluation, from different sources (e.g. implementing parties, partners etc.) to gain a picture of the overall status of the project’s expenditure
- Financial information is often provided late in the evaluation process (which takes place at, or after, project operational completion and before financial closure) and/or is incomplete. This is particularly with regard to information disaggregated by component or results category
- Not all financial information is recorded consistently between UNEP and its implementing partners
- There are occasions where an audit report raises an issue but there is no management response from UNEP
- The institutionalization of financial information is threatened by weak back-up systems that do not protect the organization against malfunctions on individual computers

Box 2. Summary of 2018-19 text analysis findings: Communication between project and financial management staff

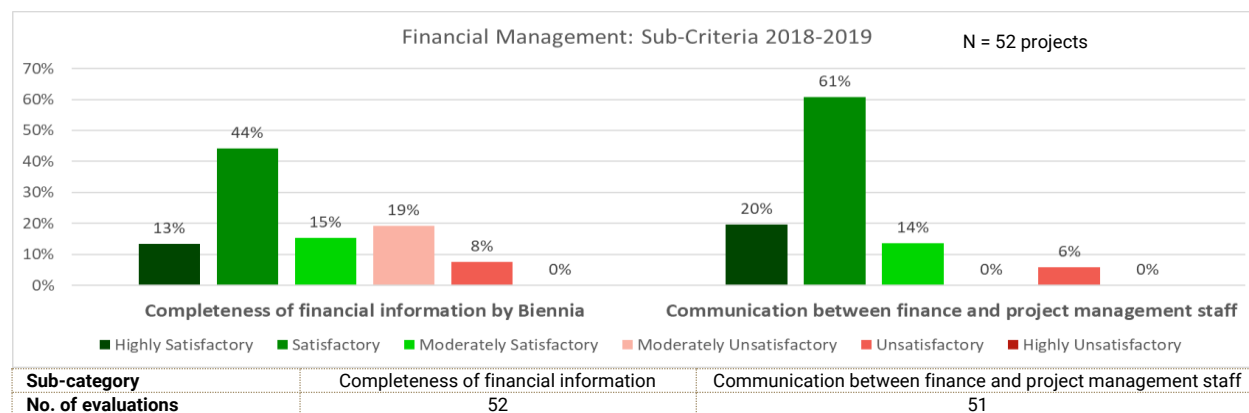
- There is frequent evidence of positive communication between project and financial management staff
- There is also evidence of a lack of the effective transfer of information between managing units
- Training on donor procedures/requirements is either absent or an area of weakness
- There is a lack of adequate back-up systems for information being stored on individual computers, (PIMS alone is not found to be an adequate back-up system)

111. There are instances where financial management has achieved excellent performance ratings despite changes in staff; the project Development of Tools to Incorporate Impacts of Climate Variability and Change, in Particular Floods and Droughts, into Basin Planning Processes, for example, achieved 'Highly Satisfactory' ratings on both completeness of financial information and financial communication despite having three Task Managers during its implementation period.

112. An issue that frequently arises in evaluations is the lack of institutional guidance in the area of defining, estimating reporting and verifying co-finance (cash and in-kind contributions).

113. Figure 25 speaks for itself in showing the breakdown in S/HS ratings for the two sub-categories under this criterion. The data confirm that this good performance (70% S/HS in the biennium) is largely due to higher performance in the area of communication between finance and project management staff (81% of projects rated as S/HS in this sub-category), while completeness of financial information has S/HS ratings of 57% of the projects evaluated in 2018-19.

Figure 25. Financial management by sub-criteria in 2018-19



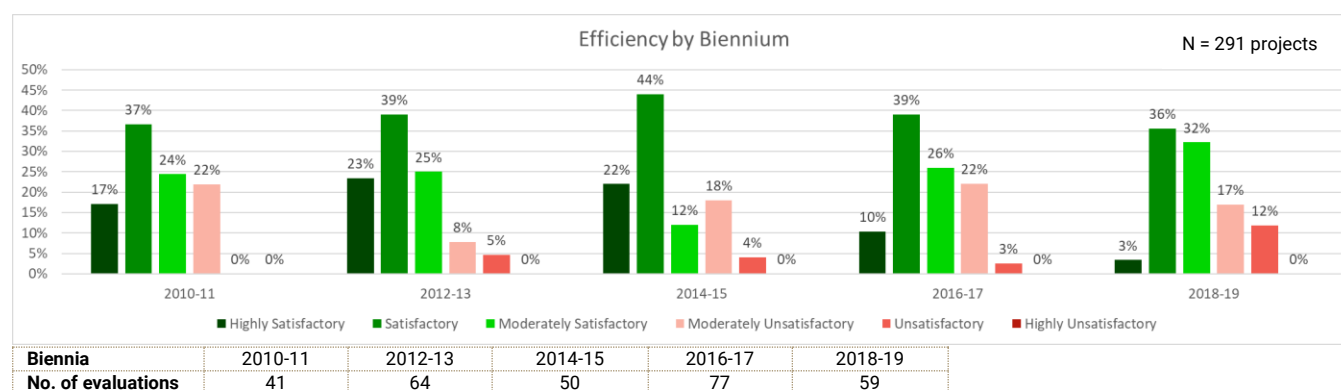
3.7 Efficiency

114. The proportion of S/HS ratings under this evaluation criterion show some fluctuation and recent signs of continued decline: S/HS ratings are 54% in 2010-11; 62% in 2012-13; 66% in 2014-15; 49% in 2016-17 and 39% in 2018-19. The proportion of ratings at the 'Highly Satisfactory' level, in particular, have dropped over time and are now at 3%.

115. This evaluation criterion focuses on two main elements: timeliness and cost-effectiveness. Given the lack of disaggregated financial information that is suited to analyses of the costs associated with specific areas of result, the interpretation of performance against this criterion is constrained. Evaluations assess projects in terms of: the timeliness of their project delivery; the efficient sequencing and arrangement of activities and the level to which projects have built on existing structures and institutions.

116. Under timeliness, evaluations consider the reasons underpinning any 'no-cost' extensions and the number of such extensions projects have had. No cost extensions are considered under efficiency because any extension to the end date of a project is usually not accompanied by an increase in project support costs. Any extension to the original project timeframe brings additional overhead costs to both implementing and executing agencies.

Figure 26. Comparison of efficiency by biennium



3.8 Monitoring and reporting

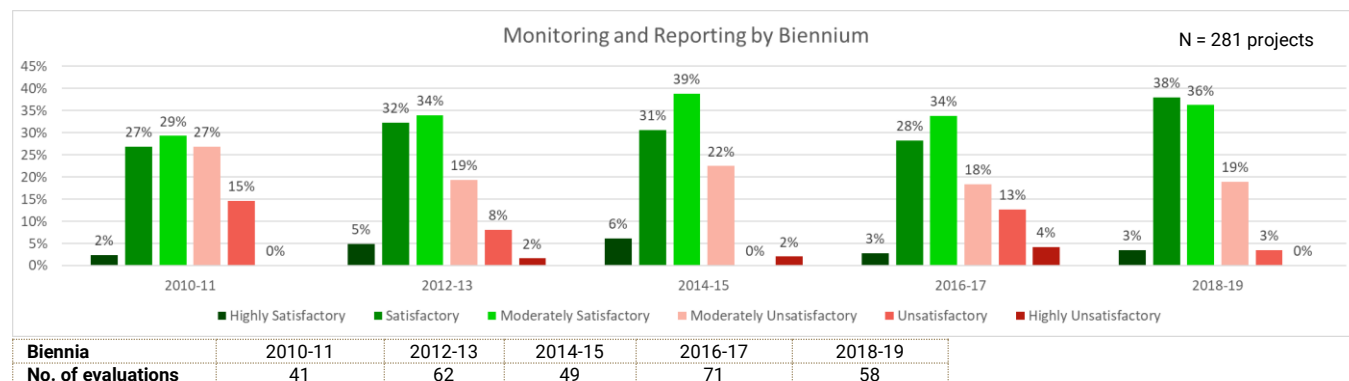
117. Monitoring is the second aspect of project performance that was, up until the 2016-17, assessed as a sub-category of Factors Affecting Performance. Not only was the topic established as a stand-alone evaluation criterion, but it was also reframed from 'Monitoring and Evaluation' to 'Monitoring and Reporting'. This is because the responsibilities for monitoring and evaluation lie in different places and only the monitoring aspect is properly evaluable during a project-level evaluation. Reporting, like monitoring, is a responsibility that also lies with the project team and is more appropriately combined with monitoring.

118. One of the key areas of weakness that comes through in evaluation reports and during the evaluation process is the lack of awareness of the key distinction between the monitoring of project implementation and reporting on project implementation. The critical role that monitoring can play in informing management of the need to make corrective action when progress against an agreed workplan (with due regard for timing and funding levels) and results framework is tracked is frequently overlooked. The role can also be falsely confused with the role of reporting in terms of describing 'what the project has done'.

119. Whilst 'evaluation' no longer features as a performance criterion in project level evaluations, it is noted that UNEP's evaluation function is regularly assessed for performance by a number of external bodies including OIOS, MOPAN, and GEF. During this biennium OIOS conducted an external assessment and this and other assessments of the Evaluation Office are described in Chapter 6.

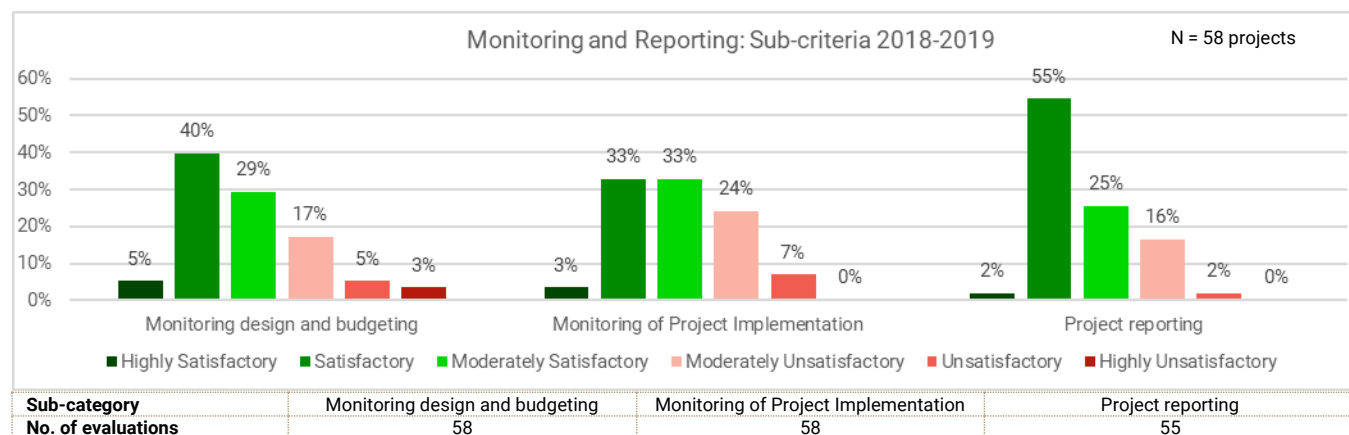
120. The proportion of S/HS ratings for overall monitoring performance has shown a gradual, and slightly unsteady, increase over time: S/HS ratings were 29% in 2010-11; 37% in 2012-13; 37% in 2014-15; 31% in 2016-17 and 41% in 2018-19.

Figure 27. Comparison of project monitoring by biennium



121. At a sub-category level we see that, during this biennium, project reporting performs best and monitoring of project implementation has the weakest performance (S/HS ratings for monitoring design and budgeting is 45%; 36% for monitoring of project implementation and 57% for project reporting).

Figure 28. Monitoring by sub-criteria in 2018-19



122. Text coding software was again used to explore this theme in 24 project evaluations completed during 2018-19. This analysis supports the data presented above: in most cases, plans for monitoring are described in the project designs, some level of budget provision is made and the reporting requirements for both UNEP and donors, especially the GEF, are followed. However, the actual operationalization of the monitoring function during project implementation is still either infrequent or weak. This means that despite planning and providing for monitoring, the substantive contributions that project monitoring can bring to the achievement of results, are not being realized.

123. Project reporting is carried out in accordance with UNEP or donor guidelines, but frequently fails to cover the full scope of the evaluand, reflects inconsistencies in different versions of results frameworks (e.g. UNEP approved logical frameworks and agreements made with donors).

124. Learning from the text coding exercise across evaluation reports is summarized below:

Box 3. Summary of 2018-19 text analysis findings: Monitoring design and budgeting

- The majority of projects lack a robust monitoring plan. This is despite evidence of Monitoring and Evaluation plans and budgets being included in project design templates from 2010 onwards
- Costs associated with monitoring are often limited to budgets for a mid-term review and terminal evaluation. These budgets are, in turn, often under-estimated. Typically, there are either no, or insufficient, resources allocated to monitoring
- Frequently there is no identified staff member responsible for monitoring, which undermines the likelihood of the role being played in full and has a negative effect on the recording of monitoring data at an institutional level. Gaps in monitoring data are then worsened when there is significant staff turnover

Box 4. Summary of 2018-19 text analysis findings: Monitoring of project implementation

- In the majority of cases, the project plan is optimistic and the time allocated for project implementation is underestimated. This is associated with a lack of up-to-date/revised (or implemented) workplans and the tracking of implementation progress through monitoring. This weakness may be linked to the number of 'no-cost' project extensions requested
- In the majority of cases, results statements and the associated indicators of project performance are either: a) not SMART or b) not appropriate for the results' level (e.g. a count of number of participants to indicate the achievement of an outcome)
- Where a proper and systematic monitoring plan is not followed there is a negative impact on: the quality of project implementation (timing and sequencing of activities, adjustments to reach targets and realise results etc.); efficiency and the scope of evaluation and potential learning

Box 5. Summary of 2018-19 text analysis findings: Project reporting

- Not all project reporting is systematic and reporting guidelines are not always followed. For example, different versions of the same report (different template/same year) can be found and in other cases the reports provided to the evaluation by the project team are not those officially submitted
- There are cases where project teams and implementing parties over-rate the project's achievements in reports and downplay the risks the project was facing
- Common findings on UNEP's Project Information Management System (PIMS) are that: the system cannot accommodate umbrella projects with numerous sub-projects or work covering a large geographic area; it does not act as a substitute for project progress documentation, adjustment and planning; it should not be taken as an alternative to substantive financial reporting and the system requires a more robust back up-system (evaluations have encountered instances where key project documentation has been lost because of failure of an individual computer and the documents have not been loaded onto PIMS)
- There are instances where high rates of staff turnover have not been supported with systematic and institutionalized handover systems and project implementation records have suffered as a result
- Where training on PIMS reporting was given, the consistency of project reporting improved

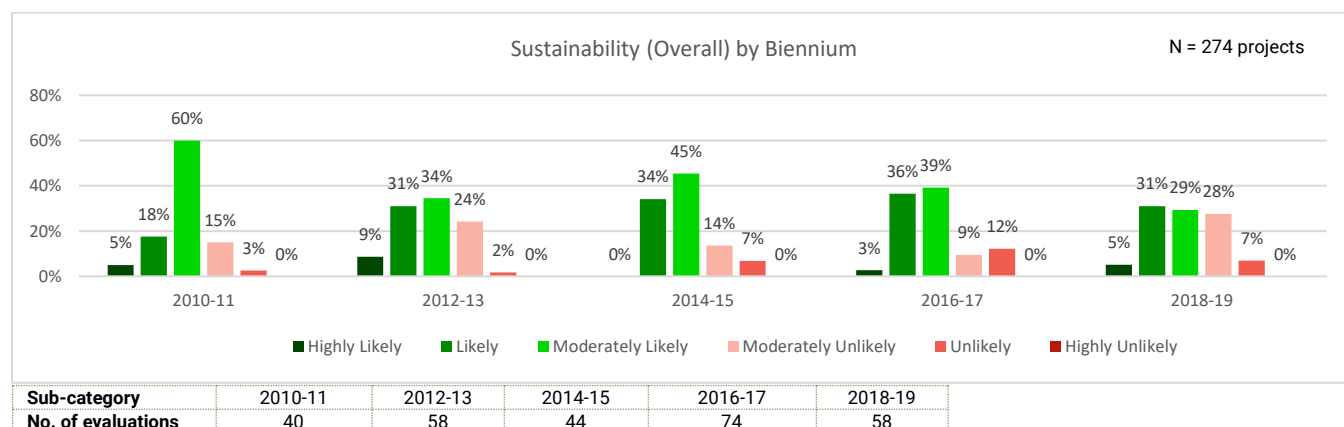
125. There are also some positive cases within the projects evaluated during 2018-19 which, although not reaching the highest performance ratings, show effort has been made in this area: for example, the project *Promoting Peace Over Natural Resources in Darfur and Kordofan* produced a document which included an introduction to the Results Based Management (RBM) approach, proposed an RBM matrix for the project and guidelines for developing community level M&E frameworks for project interventions and in the *Biochar for Sustainable Soils (B4SS)* project the project team received on-the-job training in monitoring and this improved the monitoring of project accomplishments.

3.9 Sustainability

126. UNEP Evaluation Office applies a 20% weighting to Sustainability across the three sub-categories of socio-political, institutional and financial sustainability. Interpreted as the likelihood that the benefits achieved at outcome level will be enduring, under this criterion evaluators assess: a) whether a project's outcome level achievements are sensitive to socio-political, financial and institutional factors within the implementing context and b) whether the project has put measures in place to mitigate these sensitivities. Taken together, these two dimensions give a likelihood rating.

127. Data from across the five biennia show some fluctuation in the proportion of projects that were assessed to have reached 'Likely' (L) or 'Highly Likely' (HL) levels of sustainability: L/HL ratings were 23% in 2010-11; 40% in 2012-13; 34% in 2014-15; 39% in 2016-17 and 37% in 2018-19.

Figure 29. Comparison of sustainability (overall) by biennium



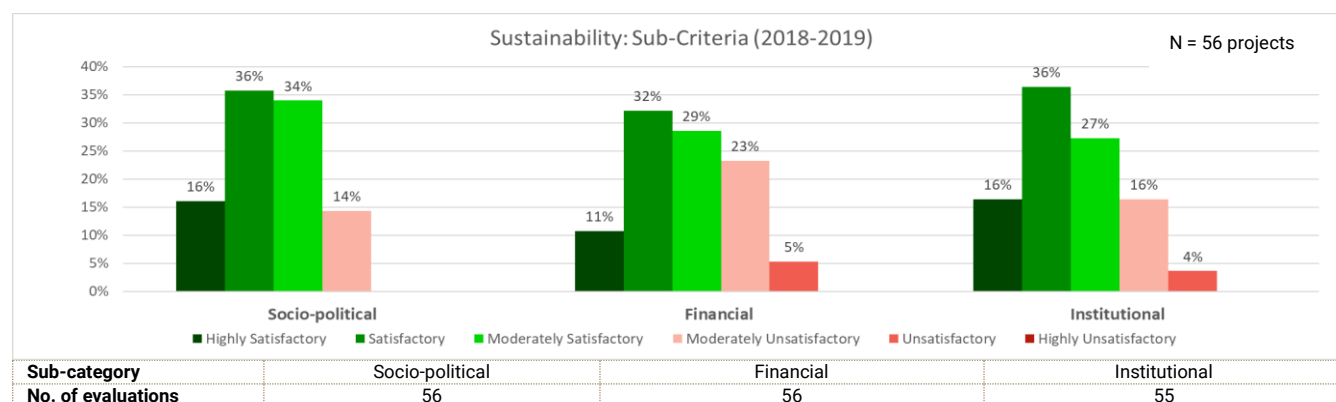
128. At the level of the sub-categories UNEP shows stronger performance in socio-political and institutional sustainability (both have L/HL ratings for 52% of the projects evaluated in 2018-19) and weaker performance in financial sustainability (L/HL ratings for 43% of evaluated projects). As the concepts of sustainability and continued effects are not widely or commonly understood across the house, a discussion of each sub-category follows.

129. Under the first sub-category, evaluations assess the extent to which **social or political factors** are required to support the continuation of project outcomes, including the level of ownership, interest and commitment among government and other stakeholders to take the project achievements forwards, and to what extent the project has responded to the factors that affect this. One would expect good performance in this area to be accompanied with strong levels of country ownership and driven-ness.

130. Establishing a common understanding of **financial** sustainability poses the most challenges as many people interpret this sub-category as having secured further funding. Evaluations assess the extent to which the continuation of project results and the eventual impact of the project are dependent on financial resources. This requires more consideration of the design of the project and the ways in which it has been delivered. With regard to future funding, evaluations also assess the likelihood that adequate financial resources will be, or will become, available to use capacities built by the project, and whether there are any financial risks that may jeopardize continuance of project results and onward progress towards impact.

131. The **institutional framework** component of sustainability considers the extent to which institutional structures (especially those supporting policies and legislation) and developed capacity are required to continue delivering the benefits associated with the outcomes beyond the completion of project activities and how robust the relevant institutional features are. One would expect this good performance here to be accompanied with positive findings at outcome level in projects that have capacity development components for ministry or other institutional staff.

Figure 30. Comparison of sustainability sub-criteria in 2018-19



3.10 Factors affecting project performance

132. This section addresses the following factors that affect project performance: Preparation and readiness; Quality of project management and supervision; Stakeholders participation and cooperation; Responsiveness to human rights and gender equity; Country ownership and driven-ness; and Communication and public awareness.

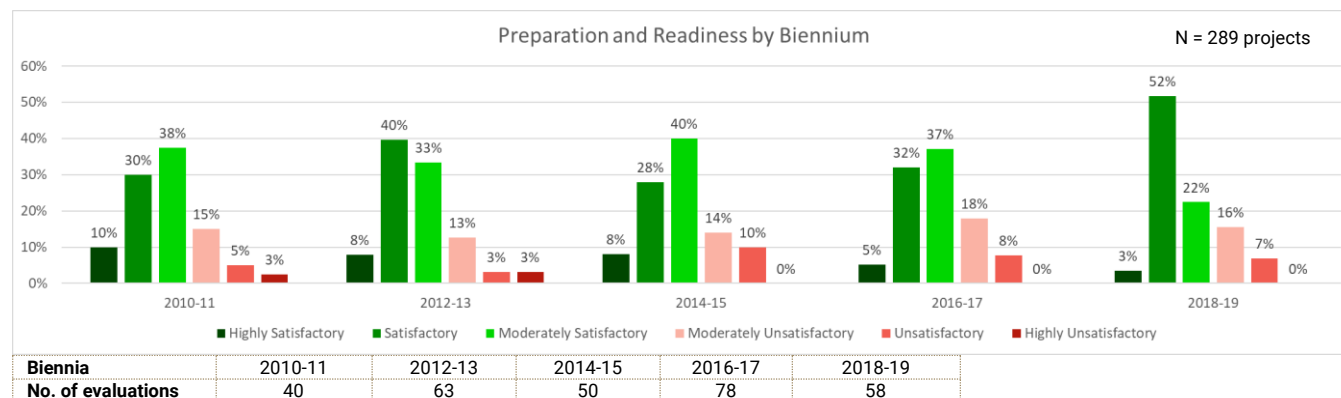
3.10.1 Preparation and readiness

133. This criterion focuses on the inception or mobilisation stage of the project, which the Evaluation Office defines as the time between project approval and first disbursement. The evaluation assesses whether appropriate measures were taken to either address weaknesses in the project design and/or respond to changes that took place between project approval, the securing of funds and project mobilisation.

134. In particular, the evaluation considers the nature and quality of engagement with stakeholder groups by the project team, the confirmation of partner capacity and development of partnership agreements as well as initial staffing and financing arrangements.

135. The proportion of higher level performance against this sub-category shows some fluctuation, no particular pattern: S/HS rates are 40% in 201-11; 48% in 2012-13; 36% in 2014-15; 37 % in 2016-17 and at an increased proportion of 55% in 2018-19.

Figure 31. Comparison of preparation and readiness by biennium



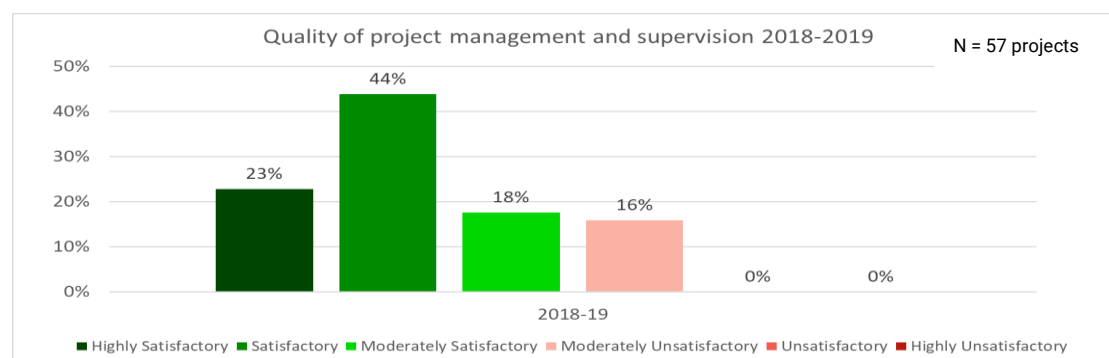
3.10.2 Quality of project management and supervision

136. Evaluations assess the effectiveness of project management with regard to: providing leadership towards achieving the planned outcomes; managing team structures; maintaining productive partner relationships (including Steering Groups etc.); communication and collaboration with UNEP colleagues; risk management; use of problem-solving; project adaptation and overall project execution. Where there is evidence of strong adaptive management, this is also assessed under this sub-category.

137. In some cases 'project management and supervision' refers to the supervision and guidance provided by UNEP to implementing partners and national governments while in others, specifically for GEF funded projects, it refers to the project management performance of the executing agency and the technical backstopping and supervision provided by UNEP (the GEF Implementing Agency). In the data presented here this distinction is not captured.

138. As the sub-categories under Factors Affecting Performance were adjusted in 2017, data only exist from 2018 onwards. The data show S/HS ratings at a level of 67% in this biennium. The closest comparator to this sub-category is the category 'Supervision, guidance and technical backstopping' which had a similar proportion (64%) of S/HS ratings in 2016-17.

Figure 32. Quality of project management and supervision in 2018-19



3.10.3 Stakeholder's participation and cooperation

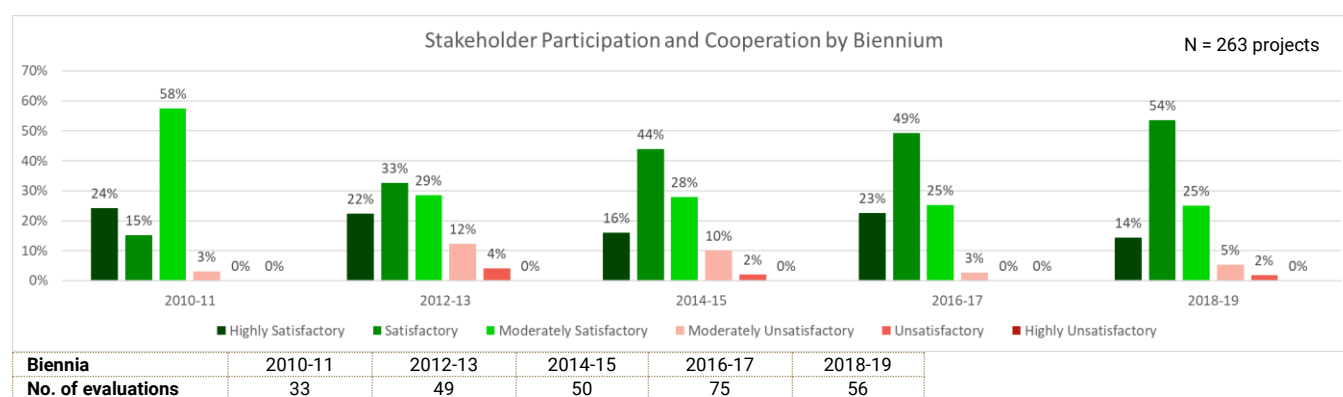
139. In an RBM context, and from the perspective of evaluation, UNEP's broad definition of stakeholders as 'those who are affected by, or who could affect (positively or negatively) the project's results' fails to provide the disaggregated categorization or vocabulary to support nuanced analysis. At project design a stakeholder analysis is required and this looks at the role played by different groups within the project context. However, this tends to result in either specific individuals or organizations being identified, perhaps those partnering UNEP etc., or a broad generalization such as 'civil society' or 'the general public'.

140. The Evaluation Office advises evaluators to identify key groups at a more disaggregated level, such as: implementing partners; government officials and duty bearers (e.g. national focal points, coordinators); civil society leaders (e.g. associations and networks) and beneficiaries (e.g. households, tradespeople, disadvantaged groups, members of civil society etc.). The Evaluation Office notes that in order to carry out any kind of analysis of the effects of project's on human rights or gender, a more nuanced understanding of stakeholders is needed.

141. Under this sub-category evaluations consider the quality and effectiveness of all forms of communication and consultation with stakeholders throughout the project life and the support given to maximise collaboration and coherence between various stakeholders, including sharing plans, pooling resources and exchanging learning and expertise. The inclusion and participation of all differentiated groups, including gender groups should be considered.

142. Performance against this sub-category shows some fluctuation and no particular pattern: S/HS ratings were at 39% in 2010-11; 55% in 2012-13; 60% in 2014-15; 72% in 2016-17 and 68% in 2018-19. However, strong performance under this sub-category is shown in consistently high percentages of project evaluations awarding 'Highly Satisfactory' ratings within each of the biennia.

Figure 33. Comparison of stakeholder participation and cooperation by biennium



3.10.4 Responsiveness to human rights and gender equity

143. In the past, the Evaluation Office has assessed gender equality as part of the criterion, Strategic Relevance, as it was a topic being introduced at policy and strategy levels by various organisations. In 2018 the Evaluation Office introduced 'Responsiveness to Human Rights and Gender Equality' as sub-category under Factors Affecting Performance and Cross-Cutting Issues to better reflect the status of human rights and gender as a topic that should be mainstreamed.

144. Having isolated the performance criterion, the weak performance in this area is extremely evident. In the forty-nine evaluated projects that had ratings for this sub-category, none received a 'Satisfactory' rating and only 8% (i.e. 4 projects) received a 'Highly Satisfactory' rating. Most projects were assessed as performing at the 'Moderately Unsatisfactory' level (33%) followed by the 'Moderately Satisfactory' level (24%).

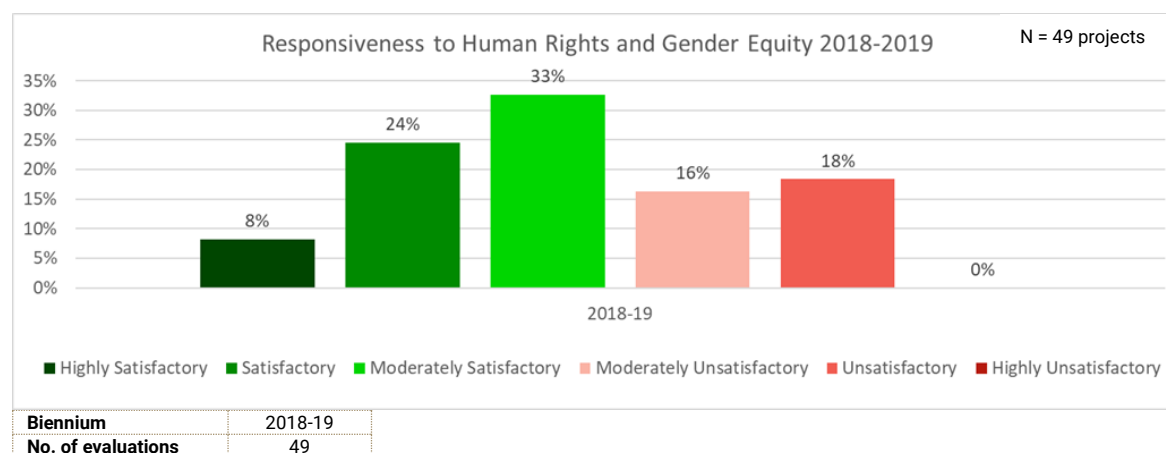
145. UNEP takes part in the UN Sector Wide Approach on Gender Equality and the Empowerment of Women (UN SWAP) and, as part of this process, an external consultant is commissioned by the Evaluation Office to review all the project evaluations completed in the year. Although this exercise primarily focuses on how well the evaluation function is addressing gender equality in its evaluation processes and approaches, the exercise also provides some insight into UNEP's performance at project level (see para 269). In addition, this sub-category was the third to be selected for text analysis. The following insights into those projects evaluated in 2018-19, which were largely designed and approved between 2011-14, were gathered:

Box 6. Summary of 2018-19 text analysis findings: Gender responsiveness

- In several cases, the project proposal intended to include women, but there was no evidence of this in the project implementation
- In the majority of the projects evaluated, gender indicators or disaggregated targets/differentiated strategies were not included in the project design
- It was noted that in a few projects, a gender and social assessment/analysis was carried out during project implementation and, in the case of one project, led to the development of a rapid assessment tool to address gender and social dimensions
- There are isolated cases of good performance, often where there is strong community involvement

146. Despite the overall weak performance in this area, 3-4 projects were rated as 'Highly Satisfactory' against this criterion and serve as positive examples. Amongst the 11-12 projects scoring 'Moderately Satisfactory' in this area there are also promising examples. Projects scoring in the satisfactory range are listed in Annex 2.

Figure 34. Responsiveness to human rights and gender equity in 2018-19

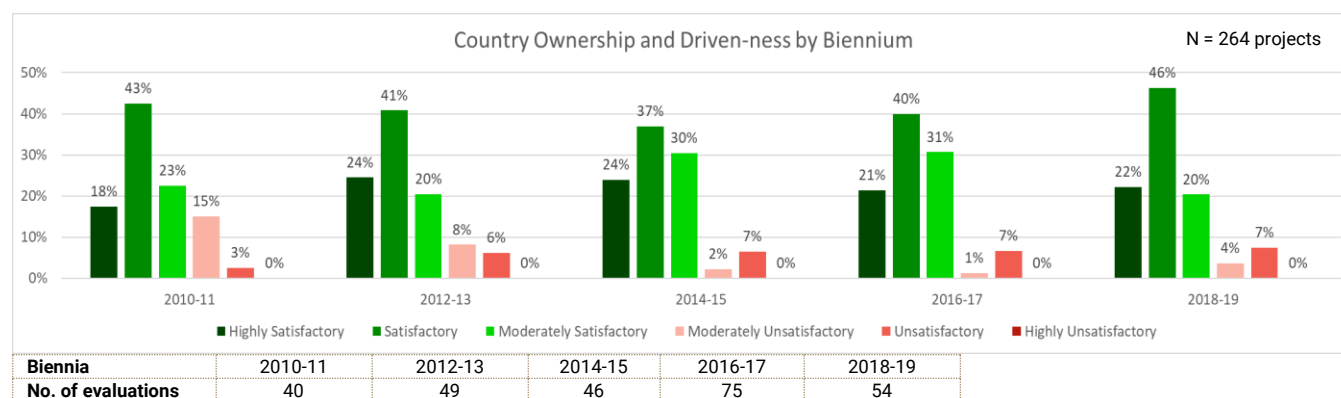


3.10.5 Country ownership and driven-ness

147. Similar to performance under the Stakeholder Participation and Cooperation, project performance under this sub-category has the high proportions of 'Highly Satisfactory' ratings within each of the biennia. In addition, the S/HS ratings are consistently high: S/HS ratings at 61% in 2010-11; 65% in 2012-13; 61% in 2014-15; 61% in 2016-17 and rising to 68% in 2018-19.

148. Evaluations assess the quality and degree of engagement of government / public sector agencies in the project. They consider the involvement not only of those directly involved in project execution and those participating in technical or leadership groups, but also those official representatives whose cooperation is needed for change to be embedded in their respective institutions and offices (e.g. representatives from multiple sectors or relevant ministries beyond Ministry of Environment). This factor is concerned with the level of ownership generated by the project over outputs and outcomes and that is necessary for long term impact to be realised.

Figure 35. Comparison of country ownership and driven-ness by biennium



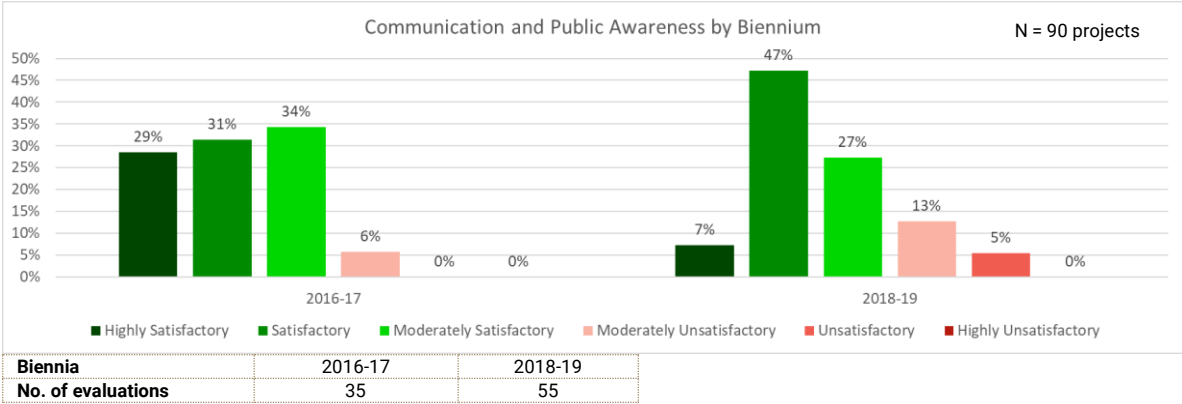
3.10.6 Communication and public awareness

149. Under this sub-category evaluations assess the effectiveness of: a) communication of learning and experience sharing between project partners and interested groups arising from the project during its life and b) public awareness activities that were undertaken during the implementation of the project to influence attitudes or shape behaviour among wider communities and civil society at large. Evaluations are expected to assess whether existing communication channels and networks were used effectively, including meeting the differentiated needs of gendered or marginalised groups, and whether any feedback channels were established.

150. When projects establish knowledge sharing platforms it is particularly important that the sustainability of the communication channel is assessed and this can be done under this sub-category or under either socio-political, institutional or financial sustainability, depending on where the knowledge platform plays a role in the TOC.

151. Introduced as a sub-category in 2016, data across two biennia are presented here. The data show similar levels of performance at the combined S/HS levels (60% of projects evaluated at S/HS level in 2016-17 and 54% in 2018-19), although a smaller proportion in 2018-19 reached the Highly Satisfactory level.

Figure 36. Comparison of communication and public awareness in 2016-17 and 2018-19



4 Key findings from strategic evaluations

4.1 Resource Efficiency Sub-programme

152. The UNEP Evaluation Office aims to carry out an evaluation of each of the seven sub-programmes over a period of two Medium-Term Strategies, this translates into approximately one sub-programme evaluation per year. The evaluation of the Resource Efficiency Sub-programme was completed in late 2018 and findings were disseminated during 2019. It assessed performance during 2014-17 and provided recommendations for the latest Medium-Term Strategy period, 2018-21. Sub-programme evaluations tend to raise institution-wide and strategic insights as well as learning specific to the sub-programme itself.

153. In this evaluation an effort was made to respond to the nature of a sub-programme as a thematically-bound entity that provides the framework for the articulation of UNEP's higher-level results. This is quite different from any Divisions that operationalise the sub-programme, (in this case largely the Economy Division) or Regional Offices that play geographically-defined roles in the sub-programme's delivery.

154. Resource Efficiency is a well-established (created in 2010) and sizeable sub-programme, with a portfolio of over fifty projects. In 2016 the sub-programme reported expenditure of USD 43m with 73% of those funds coming from extra-budgetary sources. The evaluation was designed to develop findings and insights organised under the following main topics.

Synthesis of findings from 25 project evaluations

155. With a base of twenty-five project evaluations it possible to derive an aggregation of the performance of the projects within its portfolio. The performance findings were consistent with the findings already established in the 2016-17 Biennial Synthesis Report. Specifically, the Strategic Relevance of this sub-programme is high and it is well positioned to make substantial contributions to SDGs 8 and 12. Project outputs are delivered effectively and as planned across the portfolio and some projects, such as the Green Growth Knowledge Platform and International Resource Panel, show a high level of stakeholder engagement. In other projects, stakeholder participation and cooperation was assessed as relatively poor, but the evaluation notes that assessments of the participation of, and effects on, different groups of people is constrained by UNEP's lack of a description of different stakeholder groups.

156. Some areas of the portfolio that would benefit from further attention are: the translation of outputs into outcomes and long-lasting impact; the sustainability of benefits achieved at outcome level (especially with a high dependency on external sources of funding); active application of UN approaches to human rights and the rights of indigenous people and gender equality and in the area of monitoring, which is over-reliant on reporting rather than being a responsive process whereby progress against a plan is tracked and real-time data is used to inform corrective action where needed. As with other areas of UNEP's work, any assessment of efficiency is limited by a lack of detailed financial information relevant to a results focus. Given the close link between this sub-programme and SDGs 8 and 12 the sub-programme needs to align itself quickly and effectively to these goals and embed global results at a strategic level.

Reporting on higher-level results

157. During 2010-17, the sub-programme was seen to have made considerable progress in developing a robust and transparent results reporting system. Based on a review of this system, the related backing documents and interviews with country representatives, the results (i.e. Expected Accomplishments) that have been reported were found to be evidence-based and consistent with external sources of information. However, reported results offer only a partial and fragmented picture of the change processes that are taking place in countries, business sectors, cities and among people, among others for a number of reasons: the deliverables that are counted are close to project outputs; there is a narrow definition of monitoring that does not support the development of in-depth understanding; the time necessary for change is outside reporting periods and the Project Information Management System is inadequate for gathering the required high level data. As a result of PIMS limitations, higher-level reporting carries high transaction costs for UNEP.

158. In a sub-programme context, the 'attribution' of results to the work of the sub-programme is highly complex and difficult to demonstrate for various reasons and primarily because UNEP is driving an integrated change process with a variety of actors, countries and topics. The sub-programme therefore relies on project-level results, aggregated against common indicators, to report progress against high-level results. Claims of 'contribution' or establishing a case for the 'credible association' between sub-programme effort and observed results, would be strengthened by a better understanding/articulation of causality and a stronger narrative across the sub-programme.

159. Despite a transparent and well-structured data collection and reporting process being in place there are fundamental limitations to how well the aggregation of project results (or in some cases single-case results) can represent the systemic changes that UNEP aims to achieve. This is an institution-wide issue.

160. The long-term and cumulative effects (multiple UNEP projects; UNEP contributions to country-level change) of UN Environment's work at country and regional levels are difficult to capture within the existing results framework.

Regional aspects of the sub-programme

161. This evaluation offers broad-based reflections on regional aspects of the sub-programme. It confirms that Regional offices, currently managing the implementation of several large projects, can support the integration of project activities across the sub-programme and recognizes that the development of local partnerships and stakeholder engagement is a key activity that can be performed at a regional or sub-regional level. It is important that such relationships are valued and understood within the context of a change process and UNEP's business model.

162. The evaluation also explored the evidence symbiotic relationship within which substantive divisions and regional offices operate. Within this relationship early inclusive planning, two-way communication and recognition of the equal value of technical and contextual knowledge are essential. The evaluation presents a diagrammatic representation of this two-way exchange.

163. Further insights into the potential for regional offices to strengthen the achievement of results at a sub-programme are constrained by the lack of clarity around the operationalization of the Strengthened Regional Strategic Presence Policy.

Identification of institutional narratives to establish global influence

164. New social and economic pathways, like the sub-programme is advocating, face the challenge of competing with existing mainstream systems. During this sub-programme evaluation, a workshop was held with the senior managers to explore the potential for Transition Management as an approach to help achieve the changes the sub-programme aims to bring about. Transition Management advocates a disruptive strategy for change followed by an 'evolutionary stabilization' and impact strategy, which is in line with the UNEP Business model to catalyse a steep learning curve until a take-off stage handing over further implementation to partners.

165. In addition, the sub-programme evaluation explored the potential for 'institutional narratives' to articulate how key projects within the sub-programme effect or influence change in order to support claims of key areas of work having an influence at the global level. This involved tracking the history of global processes over time where this work was present (e.g. examining UNEP's engagement with opinion leaders; contributions to key events; pivotal moments; chronological links between presence and key decisions etc). To do this, the evaluation team looked at 'Snapshots' of influential bodies of work such as the Green Economy Initiative; Finance Initiative and the Inquiry; Resource Efficiency (through the International Resource Panel) and Sustainable Consumption and Production. Strong examples of two distinct change models were identified, a 'disruptive' model that operates by interrupting the prevailing narrative and aims to redirect its future direction and an 'evolutionary' model which focuses stewardship of major bodies of work until it can be handed over to other partners who are already engaged in that sector. Green Economy and Inquiry are good representations of a more disruptive model of action, influencing the processes of G7 and G20. Resource Efficiency and Sustainable Consumption and Production represent a more evolutionary model of action. It was confirmed that the sub-programme has had a substantial influence on initiating and accelerating novel concepts in global processes as well as on stewarding implementation.

166. The SP would benefit from becoming more conscious of these connected models and take them into account in the design of SP and Projects

Recommendations

167. The evaluation makes a number of recommendations at both the sub-programme and institutional level:

Recommendations specific to the sub-programme:

- Strengthen the sub-programme Theory of Change to better articulate the causal pathways that are intended to drive the intended change (reflect science-policy process; align with SDGs etc)
- Focus attention on long-lasting impact and results reporting (e.g. global agendas, work with other units in UNEP)

- Strengthen portfolio of projects delivering on the sub-programme (strong narrative, linkages and synergies, conscious testing of models)
- Strengthen project designs in areas of translation of outputs to outcomes; sustainability; human rights and gender
- Proactive approach to donor contributions and needs (common reporting formats, pooled funding)
- Work more closely with Regional Offices (continued connectivity)

Recommendations for the institution:

- Improve the institutional narrative of UNEP
- Allocate internal budgets strategically
- Strengthen project design capacity and support mechanisms
- Maximise the potential of Umoja 2.0
- Share regional workplans, which are organized on a sub-programme basis, more widely

4.2 Clean Seas Campaign

168. The Clean Seas Campaign - implemented across three one-year phases from 2017 to 2022 - addresses the pressing global issue of marine plastic litter and contributes to the broader UNEP project, Global Partnership on Marine Litter (GPML 2015 - 2018). The use of plastics has increased exponentially in recent decades and much of it ends up in the oceans harming wildlife and contaminating the food chain. Marine litter is increasingly becoming a significant economic, environmental, human health and aesthetic problem, posing complex and multi-dimensional challenges. UNEP advocates for change, proposing that the ultimate solution is a systemic reduction of the consumption and production of plastic. The Mid-term Evaluation of the Clean Seas Campaign was designed to assess campaign performance to mid-2019 (in terms of relevance, effectiveness and efficiency), and determine the likelihood of the Campaign achieving its intended outcomes and impacts, including their sustainability.

169. The Campaign strategy is based on the premise that compelling communications will help increase public and consumer concern, shift attitudes towards the need for immediate widespread action and help drive tangible changes in public behaviour. In turn, increased signs of public concern and changes in behaviour are expected to motivate decision-makers in government and industry to take further action. The Campaign aims to achieve attitudinal and behavioural shifts through a broad-based, global, public-facing set of interventions, using a mix of online and offline channels and vehicles, events and consumption and production partnerships. This project focused on work with governments and industry to promote and support action at national and private sector levels.

Achievements:

Stimulating and Sharing Informed Dialogue

170. The evaluation found that, to-date the Campaign has achieved much of what it was designed to do in terms of amplifying public concern and getting people to engage on the issue: it has helped increase the salience of the problem of marine litter and potential solutions to ending single use plastics on social media platforms. Tapping into public attention and concern about marine plastic pollution, the project has augmented discussions with technical input and examples of good practice. Campaign-led events such as beach clean-up operations, the Clean Seas Virtual Reality experience and the Flipflop Dhow (a 9m, multi-coloured dhow made of re-used plastic and flip flops, launched from Lamu and sailed in the region) attracted high participation rates and received a great deal of positive feedback. Partner-organised events significantly extended the reach of Campaign initiatives, although partnerships themselves could have been more strategically targeted. Social media has been a key vehicle for Campaign communications and there are examples of substantial spikes in sharing on Campaign-related threads and topics after planned posts. There is also anecdotal evidence of some wider shifts in behaviour in certain geographies (e.g. Brazil, China, India, Kenya). For example, immediately after the Chinese micro-blogging platform was launched there was a spike of over 1.7 million posts and onward sharing of posts centred around a range of topics deemed relevant to the broader Campaign. Project outputs such as this attracted more than 600 million readers and generated 32 million discussions.

Shifts in Public Concern, Attitudes and Opinion

171. The Campaign's contribution to positive and rapid action is part of a wider movement for change. The Campaign's communications work forms part of a longer-term narrative, which has also included many marine experts and scientists, civil society and grassroots groups and media outlets. The first two years of the Campaign witnessed an increase in the salience of the issue as well as public concerns, with the Campaign playing a central role within a network of public communications. Data from social media analytics shows that the spike in posts discussing the Campaign and related topics after the launch, culminated in an increase of close to 2 million posts, resulting in an estimated 18 billion potential impressions. The dialogue has demonstrated a stronger focus on finding solutions and taking personal and collective action. Specific commitments to act have increased over the same period.

Shifts in Consumer Behaviour and Calls for Action

172. Shifts in consumer behaviour and the effects of calls for action are difficult to measure and even harder to attribute to a single cause. However, the social media data reveals that the online conversation was qualitatively different before and after the launch of the Campaign, not only in terms of increasing awareness, but eventually moving towards articles and media content about the wider issue of plastic pollution and marine litter, rather than mentioning the Campaign specifically. For example, phrases such as 'turn the tide on plastic' have developed into stand-alone phrases, often used without reference to the Clean Seas Campaign, particularly by the media. The evaluation confirmed that posts featuring statistics on the devastating effects of plastics and stories about the effects on wildlife were retweeted the most and that adopting a re-usable water bottle and giving up plastic straws were the most popular pledges among consumers.

173. In terms of raising public awareness, the Campaign shows that actions in one place have had a reverberating effect. Social media analytics show that when governments have acted, public interest in the same topic has increased with substantially more online activity and visibility. In the follow up period a ripple effect - the topic and discussion being shared among a wider and fundamentally global audience - is evident. The most highly shared media content, particularly during the second phase of the Campaign, centred around successful initiatives by countries taking a stance against plastic pollution.

Increased Political Will and Commitments to Act

174. A significant number of governments have made promises to act, which, along with UNEA resolutions calling on countries to significantly reduce single use plastics, are specific markers of progress in the political arena. Fifty-one countries had joined the Campaign by June 2018 – some have already taken action and other have made substantive commitments to act. For example, India made an ambitious pledge in June 2018 to eliminate all single use plastics by 2022 and Kenya banned the use of plastic bags in 2017. In both cases the close involvement of the two countries with the GPML project and this campaign, indicates a positive influence of UNEP's work.

175. Although UNEP's influence is clearer in some contexts than others, it was found to have likely influenced and certainly supported Governments to take action in a number of cases during the Campaign. The promises made by the Modi government, for example, have clear links to UNEP's work. In Kenya too, there are strong grounds for acknowledging that UNEP helped promulgate a tough plastics ban. A key added value of UNEP, throughout the Campaign, has been its leverage and convening power to get decision-makers to the table in ways that NGOs cannot do so easily. UNEA resolutions calling on countries to significantly reduce single-use plastics are specific markers of progress in the political arena and are indicative of contributing towards catalytic change in the future.

Industry Policy and Practices

176. Work with industry has produced some significant results, although it has generally proved to be an area of either slower take up, or less robust promises and subsequent action. As a positive example, Volvo made two ambitious commitments: aiming to remove single-use plastic from all their offices, restaurants and events by the end of 2019, and to ensure 25% of plastics in new cars are made from recycled sources by 2025. The follow through on these commitments can only be seen in the future. Industry partners noted that events organised by the Campaign worked to engage their employees in the issues around plastic pollution.

Key Factors Affecting Performance and Impact

177. The Campaign's aim of reducing marine litter is central to UNEP's priorities, underpinned as they are by UNEA resolutions. Particularly significant in this case was the fact that the Campaign was contextually relevant to the wider public and supported by burgeoning public and political interest in environmental issues. Some respondents highlighted that the Campaign made use of innovative and globally relevant communications' techniques to become part of the wider movement for change. Significantly there has been an increasing amount of support for the Campaign shown in the Global South.

178. While the Campaign has not always been entirely well coordinated with the Global Partnership on Marine Litter, it has managed to contribute to the achievement of the Project Outcomes through the development of Campaign and awareness materials. However, the design and use of Key Performance Indicators (KPIs), baseline data and the Campaign's monitoring system to support the evaluation across all three of its intended phases has been sub-optimal and the campaign strategy was not closely used to guide the implementation of the Campaign. This is largely the result of inefficiencies at the strategic level stemming from a lack of clarity over decision-making, faltering governance procedures within the framework of the project, and the dynamic relating to how funding is decided.

179. Challenges in the relationship between those responsible for the Campaign and those responsible for managing finances were cited in reporting, with concerns around delays in funding, and short-term deadlines. This made forward-planning difficult. At the same time, levels of resourcing have been described as inadequate for meeting both the global challenge of marine litter and the opportunity represented by the growing number of government pledges. As well as funding, the forward planning necessary to capitalise on government pledges was also lacking. Although the Campaign was found to be reasonably well designed, based on a detailed analysis of the context and the issue, it may have benefited from developing a more detailed mechanism for ensuring that countries that signed up to the Campaign actually took further action.

180. Despite these issues, the problem of marine litter is widely recognised as an important challenge in global agenda and declarations. Policy decisions and national actions to reduce the negative impact of marine litter on the environment have also increased significantly since the launch of the Campaign.

Contributions to Impact: Reduction of Plastics in the Oceans

181. The Campaign's long-term impact goal is a closed-loop plastic material chain whereby plastic's management is improved, non-recoverable plastics are phased out, single-use plastics are drastically reduced, and all plastics consumed re-enter the product life cycle, with no waste generated. In cases like this, where widespread change is dependent on many actors playing significant roles and when changes take time to both become evident and to take effect, it is difficult to identify the specific effects of a single cause. Currently, records suggest that plastic production and marine litter are continuing to grow. However, looking at the nature of engagement by consumers, suppliers and policy/regulatory authorities in the Campaign's central issues as well as the positive context provided by the Global Partnership on Marine Litter and the UNEA resolutions there is reason for cautious optimism that the changes seen in attitudes towards plastic pollution, reports of changes in user habits, calls for action and commitments towards action at national levels will be long-lasting and lead to a reduction in marine litter. At the national level links between the Campaign and GPML project and government/private sector action can be more confidently made (e.g. Brazil, India, Kenya etc).

182. This optimism is founded on the recognition that knowledge and awareness is increasingly circulating in dynamic ways that promote participation; changes in public opinion are emerging, as is a willingness to engage in collective activity and national decision-makers have shown signs of being prepared to initiate new regulations. Private sector actors may at present still be weighing up the costs of changing production practices but with evidence of a shift in consumer attitudes and increasing salience of the problem of marine litter by both the public and

governments, future changes to more sustainable practices within the private sector are plausible.

4.3 Montevideo Programme on Environmental Law

183. From 1982 to date, UNEP's environmental law activities have been organized and coordinated through a series of 10-year programmes for the development and periodic review of environmental law (known as the Montevideo Programme), adopted by UNEP's Governing Council. This Programme has set the strategy for UNEP's engagement in the field of environmental law, and the latest version, Montevideo IV, was also intended to provide a broad strategy for the international community at large.

184. The Montevideo Programme IV was designed to be a broad strategy for formulating environmental law activities. As such, it provides a comprehensive framework of activities that address the whole spectrum of environmental law challenges and the specific functional challenges that countries face in the development, implementation and enforcement of environmental law.

185. However, because the Montevideo Programme IV was not designed as a programme in the conventional sense, it lacked many of the attributes that would allow a formal evaluation. The Programme lacked a clear results framework that specifies intended outcomes, indicators of achievement, quantitative target indicators and timelines. The Programme did not specify a discreet set of activities to be implemented and outputs to be delivered or have a defined resource envelope. Baselines against which progress might be measured were absent. The lack of these key attributes meant a formal evaluation could not be undertaken and an in-depth performance assessment was conducted instead.

Accomplishments with the Montevideo IV Programme

186. The Montevideo Programme IV was instrumental in steering the efforts of the international community to develop environmental law and in supporting countries to develop, implement and enforce environmental law. A high-level overview of the key achievements under each of Montevideo IV's four pillars and some highlights of initiatives undertaken under each programme area follows.

Pillar 1- Effectiveness of environmental law

187. Concerning support for *implementation and enforcement*, UNEP's work on the environmental rule of law has been pivotal. As a result of the UNEP Governing Council Decision 27/9, the development of the term 'environmental rule of law' has become increasingly embedded in national legislative frameworks and judicial rulings. Consequently, this has enabled many countries to open up new frontiers of adjudication on environmental rights, and to integrate environmental crime in national penal codes. Increased opportunities for the exchange of best practices, including the ever-important regional ministerial platforms, have increasingly contributed to a higher standard of practice as well as more coherence within environmental decision-making at all levels.

188. Through *capacity-building* efforts, UNEP has equipped law and policymakers with information to support the implementation of MEAs in national legislation, and the improvement

of environmental law jurisprudence. UNEP's efforts to support the further development of international environmental law have also been noteworthy. A number of new environmental law and global policy frameworks have been adopted during Montevideo IV. These range from the Nagoya-Kuala Lumpur Protocol on Liability and Redress Protocol (2010) through to the 2015 Sendai Framework on Disaster Risk Reduction.

189. Moreover, UNEP has also played a key role in the strengthening of transboundary freshwater governance regimes, thereby enhancing the capacity of basin organisations to address legal challenges inherent in the management of transboundary basins.

190. Additionally, UNEP's work on promoting *harmonised approaches* in environmental law-making and implementation has helped countries to develop national strategic frameworks to deliver on the environmental dimension on the SDGs and MEAs in a coherent coordinated manner.

191. Furthermore, UNEP's work on promoting *public participation* has contributed to adoption of the Regional Agreement on Access to Information, Public Participation and Justice in Environmental Matters in Latin America and the Caribbean, the world's second regional agreement on access, participation and justice and heralded as one of the most important human rights agreements of the last 20 years.

192. And finally, UNEP's consultative process on *International Environmental Governance (IEG)* has enabled governments to articulate important reform options, which in turn were adopted at Rio+20. The upgrading of UNEP and the creation of the United Nations Environment Assembly (UNEA) represent a significant milestone in the strengthening of IEG. Moreover, UNEP's support to regional and sub-regional ministerial fora has helped to strengthen environmental cooperation at these levels

Pillar 2 - Conservation, management and sustainable use of natural resources

193. UNEP's efforts to enhance *freshwater, coastal and marine and ecosystems* have involved a range of activities for the development and implementation of national, regional and global instruments. As a result of this work, the capacity of key actors to strengthen freshwater laws, governance frameworks and institutions, as well as SDG 6, is being enhanced. Another important milestone in this area is the recent UNEA resolution on marine pollution that was adopted by Member States in 2017 (UNEP/EA.3/L.20). In this resolution, Member States have committed to step up actions including policies and measures by 2025 to address marine pollution. Another important law-related dimension of the resolution is the agreement by Member States to cooperate to develop harmonised standards and methodologies for measuring and monitoring marine litter and microplastics.

194. As regards the *sustainable management and use of soils*, the principal milestone has been the adoption of a draft resolution "managing soil pollution to achieve sustainable development." This resolution was adopted at the third session of the UNEP Environment Assembly. The resolution addressed the legal dimensions of managing soil pollution and provides a new framework to guide country-level law-making in this area.

195. Regarding the need to *conserve and to manage all types of forests*, most of UNEP's work in this area has been conducted through the unique inter-agency partnership—the UN REDD Programme (which comprises UNEP, UNDP and FAO). The UN REDD Programme supports forest

countries in developing strategies, legal and regulatory frameworks as well as institutional arrangements to reduce emissions from deforestation and degradation. As a result of the UN REDD Programme, countries are increasingly strengthening laws, policy and regulations to integrate environmental and social safeguards in REDD+ activities.

196. As regards biodiversity, UNEP has supported countries in developing implementing legislation for CITES as well as the Nagoya and Kuala Lumpur Protocols. UNEP's work in this area has helped to deepen understanding on the part of law and policy makers of the significance of these international instruments and their implementation challenges. Another important milestone in the area of biodiversity is UNEP's work in supporting the coherent and integrated implementation and synergies of biodiversity Multi-Lateral Environmental Agreements (MEAs). UNEP's work in promoting synergies at the international level is helping countries to more effectively meet their legal obligations under the rapidly expanding sphere of biodiversity related MEAs.

Pillar 3 - Challenges to environmental law

197. On *climate change*, UNEP's support to countries has helped decision-makers and legal drafters to incorporate measures for adapting to the adverse impacts of climate change into their national climate strategies. UNEP also played an important role in the negotiations of the 2015 Paris Agreement (ratified by 179 countries) and works actively with countries to help them implement their Paris commitments into national legislation. The *Guidebook on National Legislation for Adaptation to Climate Change* has helped lawmakers to incorporate adaptation measures into national sustainable development policies, plans and programmes. Additionally, the *2017 Status of Climate Change Litigation: A Global Review*, prepared by UNEP and Columbia University provides judges, advocates, researchers with a survey of global climate change litigation, an overview of litigation trends, and descriptions of key issues that courts must resolve in the course of climate change cases

198. On *drinking water*, a central milestone has been UNEP's work with countries to assist them in greening national water law. To tackle ecosystem conservation and protection, UNEP has worked with more than 100 countries to implement policies and strategies on biodiversity, desertification, biosafety, agriculture, forestry and access and benefit-sharing.

Pillar 4 - Relationship with other fields

199. On human rights and the environment, UNEP's ongoing collaboration with the Special Rapporteur on Human Rights and the Environment and the UN Office of the High Commissioner on Human Rights (UNHCHR) has enhanced the capacity of states and citizens to strengthen and operationalise the linkages between human rights and the environment. One of the notable achievements is UNEP's joint global work with the Special Rapporteur to raise awareness of the judiciary on constitutional environmental rights. As well, in 2016, UNEP contributed towards the Human Rights Council Resolution on Human Rights and the Environment, which encouraged countries to adopt a normative framework for environmental human rights. Another important area is UNEP's work to protect environment defenders, which has contributed to the historic 2018 Human Rights Council resolution on Environmental Human Rights Defenders.

200. Equally important is the recent Human Rights Council resolution, whereby the Council recognised the ongoing need to clarify key aspects of human rights and the environment. In

response to this request, the Special Rapporteur submitted a report on *framework principles on human rights and the environment* to the 37th session of the Council. And finally, UNEP supported Latin American and Caribbean countries and stakeholders in the negotiation and the 2018 adoption of the Regional Agreement on Access to Information, Public Participation and Justice in Environmental Matters in Escazú, Costa Rica, heralded as one of the most important environmental human rights agreement adopted in the past 20 years.

201. UNEP's work on *Trade and the Environment* has had an important legal dimension in its work tackling the illegal trade in environmentally sensitive commodities. As well, an important development in this field is the adoption of the UNEA-1 resolution addressing illegal trade in wildlife and wildlife products. Under the Green Customs Initiative (GCI) and Regional Enforcement Network for Chemicals and Waste (REN) umbrellas, the capacity of over 360 national enforcement officers, representing customs, police and environmental authorities has been strengthened, to detect, prevent and control the operation of illegal trade streams in environmentally sensitive commodities.

202. Concerning linkages between *security and the environment*, UNEP has continued to promote awareness of environmental crime as one of the top global law enforcement priorities. UNEP has been working closely with INTERPOL to combat transnational organised crime in relation to environmental crime. In November 2016, the UN General Assembly adopted a resolution which recognises for the first time ever that environmental crime is part of other transnational organised crimes. This represents a milestone in raising awareness about the increasing threat that environmental crime poses for peace and security, sustainable development and environmental rule of law.

Shortcomings with the Montevideo IV Programme

203. **Broad strategy nature of the Montevideo Programme** - The Montevideo IV Programme, far from being designed as an outcome-oriented and measurable programme, was designed as a broad strategy on environmental law. As such, it is formulated as a compendium of priorities, which are already embedded throughout UNEP's Law Division and its seven thematic sub programmes. The Montevideo IV Programme does not have any of the classic characteristics of UN programmes, notably: robust institutional arrangements with clearly defined roles and responsibilities; governance arrangements; programme areas that are defined by clear objectives, and which are measurable in accordance with performance metrics.

204. **Low profile** - It is clear that despite the longevity of the Montevideo Programme, its profile was rather low within the UN system, among national governments and the international environmental law community.

205. **Overly broad scope** - As noted above, because Montevideo IV was designed as a broad strategy for environmental law for UNEP and the rest of the international community, its scope is extremely broad. The intention at the time of the adoption of Montevideo IV was to secure a mandate for UNEP to deal with existing as well as new and emerging issues. Montevideo IV is defined by a set of four overarching pillars, across which 27 programme areas are distributed. There is an average of 10 to 15 action areas under each of the 27 programme areas. This represents almost 250 action areas that have been mandated by the Programme, which are difficult to monitor because of the lack of monitoring mechanisms and performance metrics.

206. **Lack of results-based management approach** - Each of the four Montevideo programmes share a common design flaw, namely the lack of a results-based management (RBM) framework.

207. **Lack of Theory of Change methodology** - Equally important is the fact that Montevideo IV was not underpinned by a Theory of Change (TOC) or Logical Framework analysis. The TOC is an essential part of programme design, articulating the pathways of change and ideas and assumptions of how change is expected to happen. It also provides a basis for tracking progress towards outcomes and final intended impacts and provides an essential framework for expressing the intervention logic of a programme. A robust TOC will facilitate the early identification of the inevitable problems within a programme and move forward with sound adaptive management.

208. **Lack of implementation structure** - The establishment of national focal points is a good start but a more robust system needs to be put in place both to promote action and to monitor the implementation of the programme.

4.4 Biochar for Sustainable Soils

209. The Biochar for Sustainable Soils (B4SS) project was implemented between April 2015 and December 2018. It evaluated the diverse formulations and application rates of biochar (as one type of Sustainable Land Management) for different scenarios of soil types, climates and agricultural systems, with a focus on six diverse countries, namely China, Ethiopia, Indonesia, Kenya, Peru and Vietnam. It also aimed to connect those groups having strong scientific expertise with those who have strong capacities for rural agricultural extension and worked with farmers to test biochar applications in different contexts. The B4SS project is an example of a well-executed research-action project that integrates scientific rigour, technical innovation and social spread. It was awarded a performance rating of 'Highly Satisfactory'.

210. The project was a GEF-funded (USD 1,826,484), UNEP project and was executed by Starfish Enterprises. The Norwegian Geotechnical Institute (Indonesia), Thai Nguyen University of Sciences (Vietnam), Nanjing Agricultural University (China), Jimma University (Ethiopia) World Agroforestry Centre – ICFRAF (Kenya), and APRODES (Peru) implemented the project at country level. The project was advised and technically supported by a Scientific Advisory Panel, which consisted of the key international experts in biochar.

Relevance

211. Given the rates of land degradation globally and the interconnectedness between agricultural transformation and achieving the Land Degradation Neutrality (LDN) indicators for sustainable food security, it has become imperative to socially disperse Sustainable Land Management (SLM) innovations as quickly as possible. The new GEF-7 programming has strengthened its efforts in this regard, and a more strategic focus to harness and leverage SLM innovations is of utmost importance if we are to halt and reverse the loss of healthy soils, and halt regime shifts in vital ecosystems. The B4SS project plays an important part in this bigger system as it is one of the innovations that improves soil fertility and health, and also supports sustainable farming practices. The project was consistent with the global programming efforts of GEF and UNEP, as well as the Global Agenda for Transformation 2030 and its Sustainable Development Goals. It was also based on country-demand, seeking to answer each of the six

countries' soil challenges, which included dry soils, acidic soils, heavy metals, erosion, over-use of fertilizers, among others.

Performance

212. The B4SS project was an exemplary project with many successes and lessons that can be learned for application elsewhere. In fact, it has been used as a flagship project in both UNEP and GEF platforms with regard to innovations in SLM and climate change mitigation. It is a good example of integration of scientific rigour, SLM innovation and social spread in agricultural communities. The six countries were well-chosen and provided a diverse testing ground to biochar application in different contexts. The capacity strengthening element of the country partners was a highlight, both in terms of building a bigger scientific and practitioner community working on biochar both in each country, as well as internationally, but also to enhance country ownership and sustainability, and further the spread of biochar applications in the countries (and beyond). The project was designed as a research-action project. The project overachieved on all its outputs, and in most of its outputs, delivered more than 10-fold what was envisaged at project design. This is a testament to the project partners' commitment to the project, and shows the importance of engagement, ownership and commitment by partners beyond project life-spans. The various materials, including the good practice guides, videos, many peer-reviewed scientific publications, are accessible on the project website (<https://biochar.international/>). These are valuable resources not only for the implementation countries, but also for other countries, and the global community at large who work on soil health and reversing land degradation. As mentioned by way of dedication of project partners, the level of championship was a strong contributor to project success, and the project was also able to strengthen and empower women farmers and women scientists.

Key Lessons Learned

213. The following lessons are meant to be useful for future project design and implementation (UNEP/GEF, in the three main areas: Land Degradation, Biodiversity, Climate Change), as well as useful for project partners in their continued work in biochar application.

Lesson 1: Championship is key to project results attainment and sustainability, catalytic role and replication

214. This lesson can be taken in two contexts, namely (1) champion farmers, or lead farmers, who allow for social spread more rapidly as they have trust and respect in their communities, and (2) champion project partners, whose commitment goes beyond tick-box project implementation, and often results in a much higher and more impactful project. In relation to champion farmers, the project was clear – and Ethiopia is one of the strongest examples here – that farmers follow other farmers (they see success, and they become willing to replicate). The lead farmer approach has been used several times in many SLM-related projects. It is important, in any project, to create an enabling environment for champion farmers to develop and lead, and then investigate how best to support, empower, and measure the impacts of their innovation and dissemination abilities to ensure the highest level of sustainability and uptake of any innovation. In Vietnam, championship through women farmers, and most notably, women farmer groups made a strong impact due to the strong organisational structure and process of women's groups in the implementation areas (and this spread to more homesteads because of regular meetings and sharing among the women's groups). This also speaks to strong organisational and social

structures within champion farmers as important contributors to the social spread of technical innovations. Despite the importance of social bonds and deeper connections, as well as passion, to project results attainment, these are rarely focused on in project design. It helps if there are foundations already laid in countries (as was the case in this project). Generally, those who are committed with greater purpose to their area of expertise, place much more emphasis on achieving the outcomes.

Lesson 2: Capacity development throughout the project through exposure can create strengthened ownership and abilities to sustain results

215. This is true particularly in the case of strengthened capacities of project partners through the implementation of the project (in some countries more than others). It was an unintended outcome that project partners would have their capacities (in scientific design, reporting, implementing projects) strengthened. In some cases, country-level scientists were greatly empowered through their exposure and forged relationships with the top scientists in the field. Another strong lesson is the involvement of students in a project like this. Many in-country students were able to rapidly increase their own capacities through exposure to the project (through PhDs, Masters, smaller undergraduate programmes, co-publishing, training, on farm support, etc). This was a direct result of the project including universities as implementers. This has ensured that in most countries, there is a growing and upcoming network of biochar-related scientists.

Lesson 3: Being as participatory as possible and co-designing with farmers has the potential to improve design and increase uptake

216. The ParTriDes approach by ICRAF in Kenya, i.e. involving the farmers and other key stakeholders, which influence directly or indirectly SLM decisions was a critical success factor in the country and contributes to the sustaining of results. Having farmers actively involved in the design of the project strengthens ownership. Farmers are the ones who understand their challenges the most and having them own the solutions through support (rather than the other way around) is a much more powerful approach to project implementation. Training and implementation of B4SS in Kenya was directly related to farmer needs and priorities. Ethiopia had a similar approach, and where farmers were given the freedom to lead experimentation, there was a stronger drive for championship.

Lesson 4: Diversity of countries in the context of this project was a powerful force for mutual learning and sharing

217. All project respondents highlighted that the diversity of countries, and the platform created to actually meet in person and share project experiences were of utmost importance to improve and strengthen capacity among individuals and partners who could take their initiatives back to their own countries. As an example, the training opportunity in China for the project coordinator in Vietnam strengthened capacity but also provided a relationship that extended to collaboration (e.g. soil samples were sent to China to be tested). Another example, where Ethiopian colleagues were exposed, in Indonesia, to the Kon Tiki kilns, allowed for them to copy the technology through their engineering students building similar kilns in Ethiopia. Testing biochar in different country contexts also allowed the project to identify which formulations and application rates are most appropriate where, and thus allowed the project to then be able to share results with wider audiences.

Lesson 5: Strong professional relationships lay a foundation for life-long commitments to achieving sustainability results

218. Many project respondents felt a very strong need to highlight life-long commitments and collaborations made in biochar and SLM through their professional relationships that were strengthened or forged through the project implementation. It is something worth thinking about, the role of partnerships and professional (often also tied to personal bonding and mentoring) relationships, and their role in the sustainability transformation.

4.5 La Plata Basin Project Evaluation

219. The GEF-UNEP project, “Sustainable Management of the Water Resources of the La Plata Basin with Respect to the Effects of Climate Variability and Change” (LPB project) was implemented between March 2011 and December 2017. The project was funded by the GEF (\$10,730,000), implemented by UNEP and co-executed by the General Secretariat of the Organisation of American States (GS/OAS) and the Intergovernmental Coordinating Commission (CIC). This project is an example of a) the length of time it takes to bring about substantive change in complex and multi-country situations and b) how country ownership and driven-ness can complement, and/or compensate for, limitations in project designs.

220. The objective of the La Plata Basin project was to strengthen transboundary cooperation among the riparian country governments of Argentina, Bolivia, Brazil, Paraguay and Uruguay to ensure management of shared water resources of the La Plata Basin in an integrated sustainable manner, within the context of climate variability and change, while capitalising on development opportunities.

221. The overall evaluation of the project was rated as ‘Moderately Satisfactory’: Strategic Relevance, Effectiveness and Financial Management were rated in the ‘Satisfactory’ range, while the project had lower ratings for Project Design Quality, Efficiency, Monitoring and Reporting, Sustainability and Factors Affecting Performance.

Context

222. Covering an area of approximately 3.1 million square km, La Plata is one of the world’s largest river basins. The basin has 3 major river systems (Parana, Paraguay and Uruguay rivers) which drain approximately one fifth of the South American continent rivers (with an output of 25,000 m³/ sec) into the Atlantic Ocean. The basin’s extensive mineral resources, forests and soil fertility are conducive to economic development, contributing to 70% of the GDP of the 5 countries. The basin’s population in 2008 was an estimated 100 million inhabitants and includes all five capital cities: Buenos Aires, Sucre, Brasilia, Asuncion and Montevideo.

223. The preparation of an integrated water resources management programme in relation to climate change was agreed upon by the five riparian countries and the preparation of an Action Plan began in 2001, during the IV Inter-American Water Management Dialogue. The Action Plan based on the common vision of the five countries and defining strategies to guide the development in the LPB during short- (5 years), medium- (10 years) and long-term (more than 15 years) timeframes, was approved by the CIC in 2003. Between 2005-07, the Framework Programme was designed and this UNEP-GEF project was flagged as Phase 1 of the LPB Framework Program, with a start date of March 2011.

Accomplishments

224. The project's strategic relevance stands out as a particular strength due to the fact that it was designed and implemented based on the main national and regional priorities of water management of the five riparian countries, with good complementarity with existing interventions. In addition, the project promoted an unprecedented generation of knowledge of hydro-climatic resources of the La Plata Basin and the creation of a network between national governments, technicians and specialists within the whole region, strengthening relationships and collaborative work.

225. Country ownership and driven-ness was a strong feature of this project. Each country took a national leadership role on strategic guidance of project delivery, endorsing project results, provision of in-kind resources and co-finance (~US\$ 60 million), and to some extent advocating for changes to achieve higher-level results. The 17 thematic groups and 5 inter-ministerial working groups were also positive features of the project, promoting integration across sectors among the five countries. The results of the evaluation findings, and especially lessons learned, and proposed recommendations have been readily taken up by the CIC in the follow-on phase of the project.

Shortcomings

226. The project design had weaknesses: the logical framework was weak, with absent casual pathways and a weak stakeholder analysis. The project lacked a robust monitoring system with no one in charge to monitor the progress using indicators and no data collection methods.

227. The project was large and complex with 25 outputs and 12 outcomes, which contained 4 priority activities, 4 pilot projects, and several studies to cover a wide range of issues. The project was made up of three main components (strengthening basin-wide co-operation by developing a harmonised legal framework; integrated water resource management; and development of hydro-climatic modelling and scenarios for adaptation), that would lead to the fourth component for the development of a Strategic Action Plan amongst the five countries and put a decision system support in place.

228. Despite being implemented in 2011, the project missed an opportunity to consider possible gender inequalities in the access and control of water and natural resources; the vulnerabilities of women, children and indigenous communities to environmental degradation; and/or the role of women and indigenous people in mitigating or adapting to climate change, and their participation in the protection and rehabilitation of water resources.

Evaluation

229. A considerable amount of time was spent by the evaluation team in reconstructing the Theory of Change (TOC) from the project documents, interviews with key stakeholders and best practices on the subject matter. This was to ensure that the project's complexity could be unpacked, key drivers could be identified and assumptions stated, as well as to complete the stakeholder analysis. A positive by-product of this lengthy activity was that the project team was closely involved in the development of the TOC and gained first-hand experience in the usefulness of the TOC approach in project management. A clear and comprehensive TOC was particularly

important in a project with such complexity and the need to identify key drivers and map out stakeholders for each aspect of the project to successfully effect change was critical.

230. Part of the success of the project's achievements can be attributed to the strong country ownership and driven-ness, especially with regard to the inter-ministerial relationships and decision making. Each country had a political and technical representative, making decisions on its behalf. Integrated into the project decision making were 5 inter-ministerial working groups, as well as 17 thematic working groups.

Box 7. Institutional History of the water resource management in the La Plata Basin

1967: First Meeting of the Ministers of Foreign Affairs of the LPB – governments of Argentina, Bolivia, Brazil, Paraguay and Uruguay established the Intergovernmental Coordinating Commission (CIC) as the basin wide coordinating body.

1969: Riparian countries sign the La Plata Basin Treaty, the primary legal instrument for the La Plata Basin. CIC was *"in charge of the promotion, co-ordination, and follow up of multinational integrating development programmes in the La Plata Basin, with the technical and financing assistance of international agencies and to execute the decisions approved by the Ministries of Foreign Affairs."* FONPLATA was the financing agency, the Intergovernmental Committee for the Paraguay-Parana Waterway (CIH) was the inter-modal transportation agency and the CIC has technical and political representation.

2001: During the IV Inter-American Water Management Dialogue, the preparation of an integrated water resources management programme in relation to climate change was agreed upon. In December 2001, the preparation of the Action Plan commenced.

2003: Action Plan approved by the CIC in preparation of the Framework Strategic Action Programme, based upon the common vision of the five countries and defining strategies to guide the development in the LPB during the short (5 years), medium (10 years) and long term (more than 15 years). Within this context, CIC requested the GEF PDF block A support to identify the priority transboundary programme elements and GEF PDF block B support to prepare the Framework Programme.

2005-07: the Framework Programme was designed and this project was flagged as Phase 1 of the LPB Framework Program.

2011-17: La Plata Basin project implementation (GEF-UNEP). Executed by the General Secretariat of the Organisation of American States GS/OAS.

2018-20: 'Preparing the Ground for the Implementation of La Plata Basin Strategic Action Plan' (GEF-Latin American Development Bank - CAF). Also executed by GS/OAS.

4.6 Poverty-Environment Initiative (a Joint UNEP / UNDP Evaluation)

231. The Poverty Environment Initiative (PEI) is a joint Global Programme between the United Nations Development Programme (UNDP) and UNEP, that supports country-driven efforts to mainstream poverty-environment linkages into national, sectoral and sub-national development plans and budgets. The project 2013-18 worked in 25 countries, four regions (Africa, Asia, Europe and the Commonwealth of Independent States (ECIS), and Latin America and the Caribbean (LAC)) and at the global level.

232. UNEP and UNDP and jointly commissioned the final evaluation of PEI to assess the effectiveness, efficiency and sustainability of the PEI's interventions. The evaluation aimed to provide lessons that can inform the PEI's successor project - the Poverty Environment Action (PEA) 2018-2022, as well as initiatives by others to mainstream poverty and environment into policies, plans and budgets in support of the delivery of the Sustainable Development Goals (SDGs).

233. The PEI started in 2005 and was in its 13th year of operation at the time of the evaluation. The PEI has evolved through a number of phases starting with its Africa Pilot phase 2005-07. Each phase has built on the preceding one, consolidating the successes achieved, pushing into more difficult P-E mainstreaming areas and honing approaches based on lessons learned. Looking forward, a new four-year joint programme - Poverty-Environment Action (PEA) (2018–2022), aims to build on the work of the 2013-18 phase. The focus of Poverty-Environment Action is to align finance and investment with poverty, environment and climate objectives to accelerate Sustainable Development Goal implementation.

234. The intended **Global Programme Outcome** of PEI (2013-18) was: “Enhanced implementation of development policies, plans and budgets that combine environmental sustainability and poverty reduction to contribute to inclusive and sustainable development goals.”

235. The three global **outcomes** of the PEI were:

- P-E approaches and tools for integrated development policies, plans and coordination mechanisms applied;
- Cross-sectoral budget and expenditure processes, and environment-economic accounting systems institutionalised; and
- P-E approach and experience documented and shared to inform country, regional and global development programming by the UN and Member States.

236. The PEI, as a joint UNDP and UNEP effort, operated through: (1) joint teams at the regional and global levels; (2) a single Managing Agent - UNDP; and, (3) a pooled fund (i.e. a single account).

237. The PEI organizational structure consisted of: a global programme (managed by the Poverty Environment Facility (PEF); a joint institutional board; a Technical Advisory Group (TAG) and Donor Steering Group (DSG)); and 4 regional projects and a number of country projects. The **Poverty-Environment Facility (PEF)**, based in Nairobi, provided the overall programme, operations and financial management coordination. The PEF was answerable to, and under the instruction of, the **Joint PEI Management Board (JMB)**, composed of senior UNDP and UNEP staff. The PEF and the JMB were supported by the **Technical Advisory Group (TAG)**. The country projects were implemented by joint government-UN PEI country teams (CT), with support from, and in coordination with, joint UNDP/UNEP regional teams (RTs).

238. The total planned donor and in-kind contributions for 2013-2018 totalled US\$ 60 million. PEI 2013-2018 was funded by the Governments of Germany, Norway, Spain, Sweden, the United Kingdom and the European Union. In addition, Switzerland contributed around US\$ 4 million for Lao PDR at the country level.

Evaluation Findings

239. PEI is a well-lauded Programme, which has achieved ground-breaking work with a relatively small budget. It has developed specialized knowledge, a comprehensive P-E mainstreaming tool kit and a number of strong case studies that can inspire others. The demand for PEI's services remains strong, especially given the recognized support PEI can offer on SDG implementation. The PEI UNDP-UNEP collaboration is a leading example of joint working at the forefront of the UN reform process towards a One UN.

240. The project's achievements in countries reflect the maturity of the services provided by PEI, which evolved over 13 years. The project saw the deepening of PEI's efforts in mainstreaming into sectors and budgets (e.g. the development of Public Environmental Expenditure Reviews in a number of countries) and increasing activity at the subnational level. More attention was placed on the political economic aspects of poverty-environment mainstreaming, including equity and the social inclusion of marginalized groups (especially women). The project facilitated Government's efforts to localize the SDGs, a role highly valued by countries. In a discreet number of cases it is possible to link PEI's policy work, often initiated in the previous stage of PEI, through to improvements in the lives of the poor.

241. PEI's impact on institutions, policies and investments derived from a very wide range of interventions including: capacity building for decision-makers in sustainability and climate change adaptation; economic research and analysis; tracking public spending on climate change; and improving enforcement of environmental regulations. PEI's extensive toolkit is an asset for the whole development community. The quality and importance of PEI's toolkit has been acknowledged through OECD's peer review of best practice on environment and development integration. However, there is a view that the tools are not widely enough understood and used by the development community.

242. The project had considerable success at the country level and is held in high esteem in several countries. The PEI has strategically worked to break down the marginalization of the environment. It has done this by not solely working with stand-alone and generally less well-resourced environment departments, but by convincing more influential Government departments such as the Ministry of Finance / Planning of the importance of Poverty-Environment-Climate-Gender (P-E-C-G) mainstreaming. The PEI built strong relationships with Ministries of Planning / Finance in the countries in which it worked – something that many other environmental initiatives have been unable to do. The PEI achieved this by putting in place strong technical advisors, who were able to build trust through their long-term presence and clear understanding of Government priorities, and through the development of tools that helped these Ministries reach their objectives. Engagement with these ministries greatly facilitated the mainstreaming of P-E-C-G into plans and budgets and is now resulting in higher country-level investments in and budget allocations for poverty-environment objectives and climate change adaptation. A considerable amount of core funds has been provided by UNDP at the country level, supporting country level operations, and demonstrating commitment to the joint project.

243. PEI 2013-2018 faced a period of management change (from 2015 to mid-2018), which caused division and adversely affected morale particularly among the Global team (PEF) and the Africa Regional team who are co-located in Nairobi. The main changes were: (i) Changes to PEI rules and procedures to better align with UNDP procedures and new requirements on EU PAGODA funding; (ii) Staff changes and losses within the Global and Regional Teams; and, (iii) the

uncertainty associated with the transition to PEA. **Delays in cash disbursements** to regional and country teams affected delivery and caused inefficiencies and frustration. **Strategic leadership** was lacking in a number of areas including: (i) development of the project document for PEA; (ii) knowledge management and communications; and, (iii) timely action to identify and address management issues affecting the project.

244. The success and sustainability of PEI was to a large extent dependent on its ability to leverage its work. The project has operated in the context of millions of dollars as opposed to the billions of dollars channelled through the Global Environment Facility (GEF) and the Green Climate Fund (GCF). The aim of PEI was to influence catalytic interventions that could have a major impact in the quality of the trillions of investments made by the private sector (DSG, May 2017). As a small program therefore, the PEI was designed to act as a catalyst and to leverage the small resource allocation it had through partnerships and spreading the word on PEI achievements at the country level, so that they could be taken-up by others. It was generally felt that the many avenues for leveraging PEI's work were underexploited such as: (i) integrating PEI into other UNEP / UNDP programs and projects; (ii) developing partnerships; (iii) knowledge management, including South-South learning; and, (iv) resource mobilization.

245. Overall the project rating is evaluated as '**Satisfactory**'. While virtually all the project's targets were reached, the project is rated as 'Moderately Unsatisfactory' in terms of management (implementation and financial). The project faced prolonged management difficulties, which severely affected morale and efficiency and it is probable that the project's achievements could have been even better were it not for an extended period of operational challenges which blighted the efficient running of the programme. The successful delivery at the country level (against which the majority of the targets depend) under PEI occurred despite the management difficulties, but it cannot be assumed that this will be the case for PEA.

Lessons learned

246. PEI has accrued a wealth of lessons through its 13 years of implementation that can inform the mainstreaming work of others and the strengthening of the UNDP-UN Environment joint-work modality. Nine key lessons are:

- **PEI as a small project needed to be strategic and catalytic.** PEI with its small budget could not do everything, therefore PEI needed to be clear on the most strategic entry points and catalyse support from strategic partners to ensure its sustainability. However, the identification of synergies with on-going and planned activities across different sectors at the local, national and regional levels can be time-consuming and needs to be assigned to capable people and resourced.
- **P-E mainstreaming requires long term support.** P-E mainstreaming is complex and becomes more demanding as progress is made along its key stages, which can be broadly characterized as: (i) integrating P-E objectives into a national development plan; (ii) integrating P-E objectives into a range of sector plans, policies and strategies, while concurrently engaging in influencing national the sector budget and monitoring processes; and, (iii) integrating P-E objectives and implementation at the sub-national level. It is a long-term process of institutional change across Government and capacity building. Sustainable shifts in the approach of country governments therefore require long term funding.

- **Integrated approaches should target existing processes rather than creating parallel systems.** The most effective way to promote integrated approaches will usually be by targeting existing planning, budgeting and institutional coordination mechanisms and tools and enabling them to better respond to the three dimensions of sustainable development.
- **Without finance, plans cannot be implemented.** Working with the Ministry of Finance and tracking expenditures are both key to increasing budget allocations to P-E. The financial challenge is more acute at the sub-national level, where the links between planning and budgets are weak and there is limited capacity / understanding of P-E mainstreaming. It is, therefore, important to support Governments to ensure that the delegation of powers to the sub-national level is accompanied by relevant budgets—including target transfers from national to local budgets (AR, 2017).
- The inclusion of P-E objectives in national development plans does not automatically lead to their integration in sector and sub-national plans. If the national poverty-environment objective is not transformed into concrete actions through sector and district plans, change is not realized. Engaging in a small number of pilot districts and sectors and seeking to integrate P-E objectives more broadly through the inclusion of P-E elements in central Government guidelines to all districts, provinces and sectors has proved to be the most realistic option for achieving this in Africa, given PEI's resources.
- **Influencing policy is very much relationship-based.** A small contribution can have a big impact if good relations with the Government are built. Effective technical assistance support requires patience, perseverance and presence. A long-term presence allows a technical advisor to be perceived as more of part of the Government team, rather than an outsider, and better able to understand the sensitivities around policy changes and how systems operate.
- **Pilots can be powerful tools for shifting policy.** Buy in at the National Government level is important and pilots can be used to test ideas and build a business case that can be used to influence policy. For example, in Tanzania, PEI pilots on fish farms led to their inclusion in its Fisheries policy on account of the demonstrated environment, social and economic benefits.
- **Capacity gaps are substantial** in the Lower Middle-Income Countries (LMICs) especially at the local level, and on-going training and capacity building is required to ensure skills and expertise are broad and deep enough to sustain P-E-C-G mainstreaming.
- **Data gaps are a key challenge** and efforts are needed to build up data needed for policy design and monitoring purposes, especially environmental data.

5 Compliance with evaluation recommendations

5.1 Overview of Compliance with Evaluation Recommendations

247. This section reports on the compliance levels for the biennium 2018-19 both for the submission of management responses to evaluations and regarding the implementation of evaluation recommendations.

Management responses to evaluation - overview of policy and process

248. As per the UNEP Evaluation Policy (2016). "The findings from evaluations will be discussed with relevant programme staff and senior management where specific evaluation recommendations will be reviewed. Subsequently, programme staff / senior management will prepare a formal management response to the evaluation that specifies which recommendations were accepted, what action will be taken and a timeline. They will provide a justification for any rejected recommendations. Every six months, the Evaluation Office of UNEP (EOU) prepares compliance assessments of the implementation of accepted recommendations based on submitted compliance reports prepared by the organizational unit responsible for the management response. The EOU ensures that implementation actions meet required standards."

249. A one-month deadline for submission of the recommendation implementation plan is imposed from the date the Evaluation Office sends out the final evaluation report. The recommendation implementation plan should specify:

- Whether a recommendation has been accepted
- How the recommendation will be implemented
- Who is responsible for its implementation
- The date by which the implementation of the recommendation is to be completed
- What actions have already been taken (if any)

250. After the implementation plan has been completed, the Evaluation Office follows-up with the substantive office on the status of implementation of recommendations at six-month intervals and reports on the levels of compliance. During this biennium this was done in September and in March every year.

251. At each assessment point, the progress in implementing agreed recommendations is assessed by the Evaluation Office. The compliance level is based on implementation progress updates and supporting evidence provided by the responsible programme staff. Recommendations are deemed to be:

- Fully implemented (compliant)
- Partially implemented (partially compliant)
- Not implemented (not compliant)
- No further action required (if events overtake what is planned)

252. After the first assessment point when a recommendation has been fully implemented it is recorded as 'compliant' and its status changes from 'open' to 'closed', and no further follow-up is required. All other recommendations remain 'open'. When recommendations reach the second assessment point (i.e. 12 months after initial completion of the recommendation implementation plan) their implementation status is updated again and then all recommendations are recorded as 'closed'. This means that staff responsible for implementing the evaluation recommendation have 12 months from the date the implementation plan is received by the Evaluation Office to implement them fully.

Management Response compliance 2018 - 2019

253. In the 2018-19 biennium, 61 evaluations were completed and 56 of them issued recommendations. The evaluations that required a management response were exclusively from projects residing within Ecosystems and Economy Divisions.

254. The preparation of a management response in the form of a **mandatory** recommendation implementation plan and an optional narrative is a requirement for **all** evaluations.

255. Fifty-six evaluation recommendation implementation plans were sent to UNEP substantive offices. Overall, **75 %** (42) of project/programme officers prepared an implementation plan that provided responses to the evaluation recommendations. **Only 25%** (14) of the evaluation management responses received were submitted within one month of the issuance of the evaluation report as per the requirements set out in the UNEP Programme Manual.²⁰ For **25%** (14) that were completed, project / programme managers **failed to prepare a management response**.

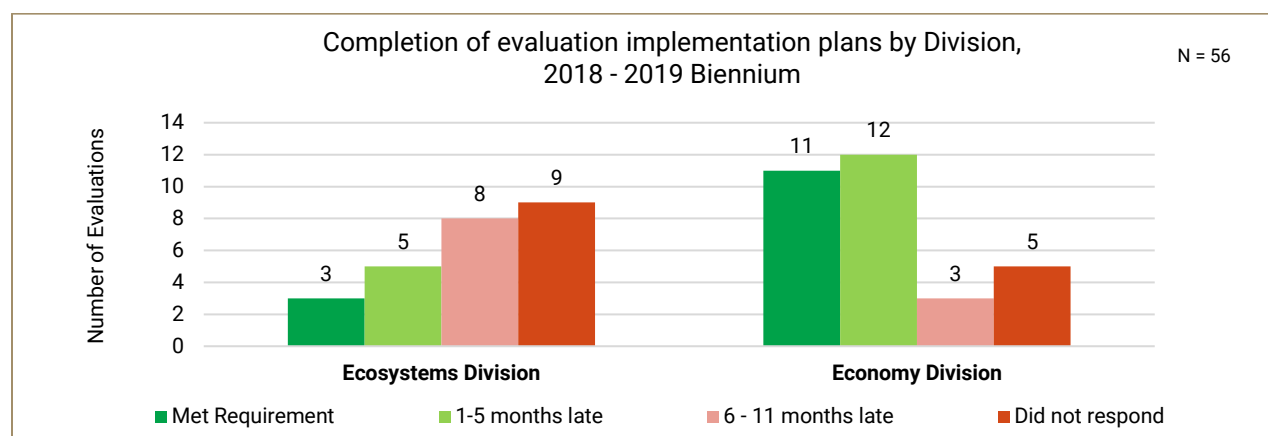
Table 6: Rate of Management Responses

Action	Number of instances	Notes
Completed evaluations	61	
Recommendations plans sent to substantive office	56	5 project evaluations issued no recommendations.
Recommendation management responses (within one month)	14	For the issued recommendations plans, 25% had management responses provided within the required timeframe.
Recommendations management responses late (i.e. more than one month after recommendations plan issued)	42	For the issued recommendations plans, 75% had management responses provided late.
No management response received	14	For the issued recommendations plans, 25% provided no management response.

²⁰ <https://wecollaborate.unep.org/pages/viewpage.action?pageId=9374201> (internal intranet reference)

256. Compliance rates for completion of management responses to evaluation **should be 100%** and the timelines for these standard requirements should be respected. **This is a performance dimension needing strong management attention.**

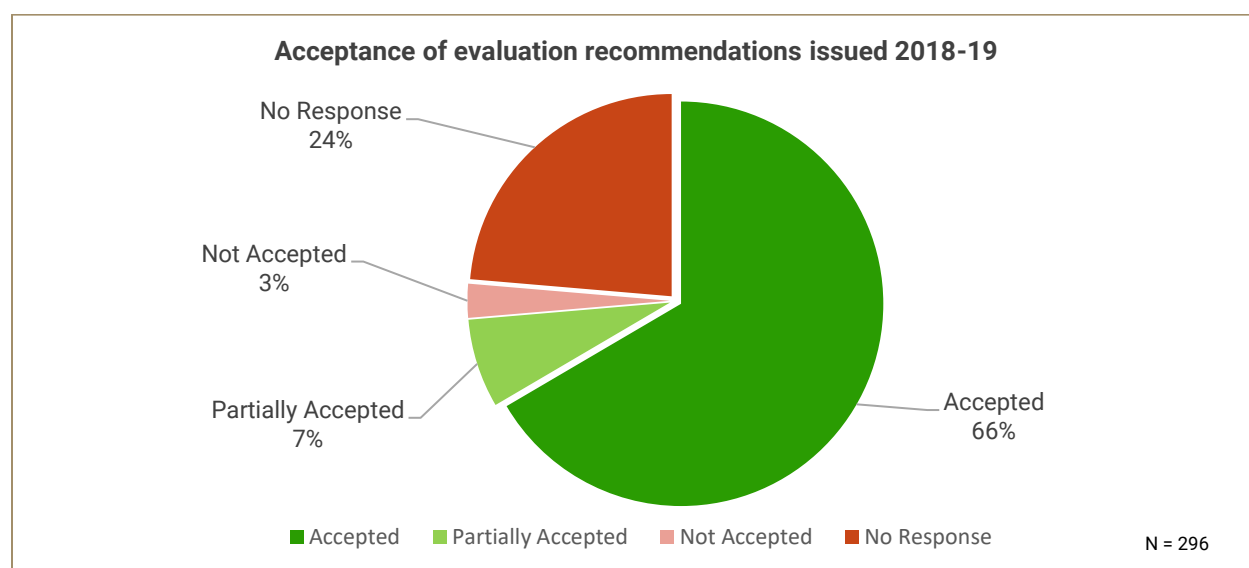
Figure 37. Completion of evaluation management response (recommendation implementation plans) by Division



Acceptance of Recommendations (2018-19 Biennium)

257. Two hundred and ninety-six (296) recommendations were issued in the 2018–19 biennium comprising of 261 programme\project related recommendations and 35 recommendations from higher level evaluations. Figure 38 shows the level of acceptance by project / programme staff of the evaluation recommendations issued. Responses were **not** received for 24 % of the recommendations because, as noted above, the corresponding implementation plans were not submitted by the responsible project officers.

Figure 38. Acceptance of evaluation recommendations 2018-19.

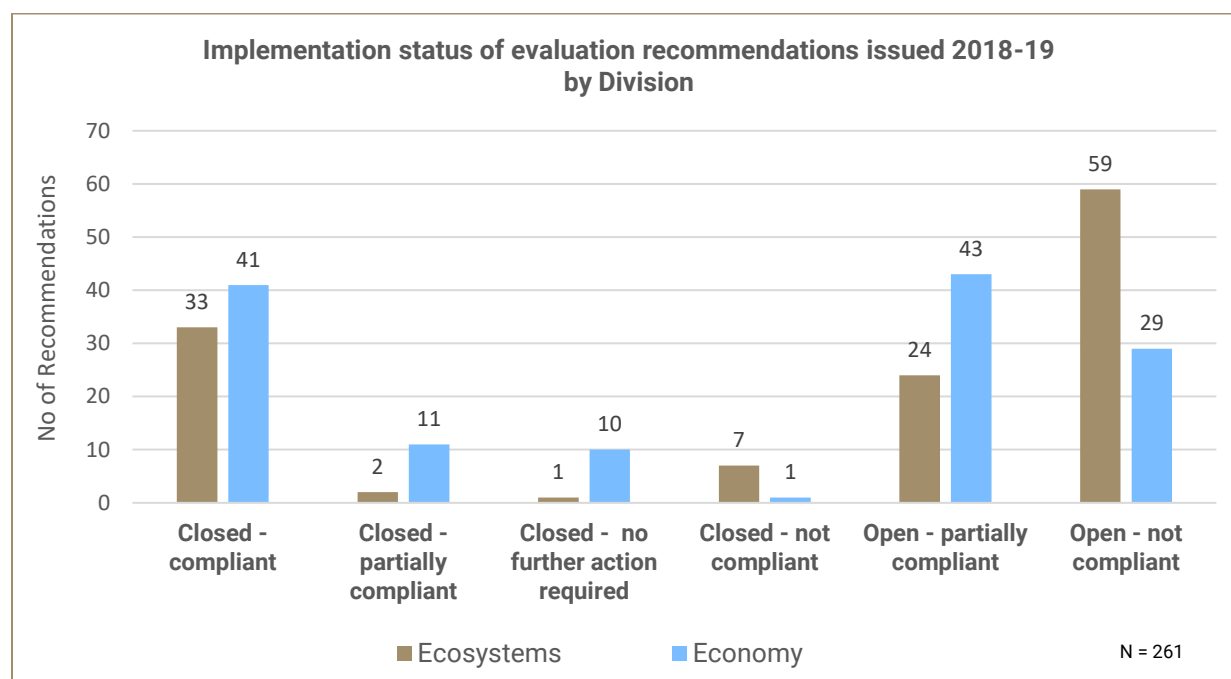


258. Reasons for non-acceptance or partial acceptance of recommendations included: lack of funds for implementation, project closure / follow on projects not envisaged and recommendations being beyond the purview of UNEP or the executing agency. Rejected recommendations were reviewed by the Evaluation Office and determination made as to whether they should be 'closed' with 'no further action required' or whether alternative ways to address the recommendation should be considered.

Implementation of accepted evaluation recommendations (2018 -19 Biennium)

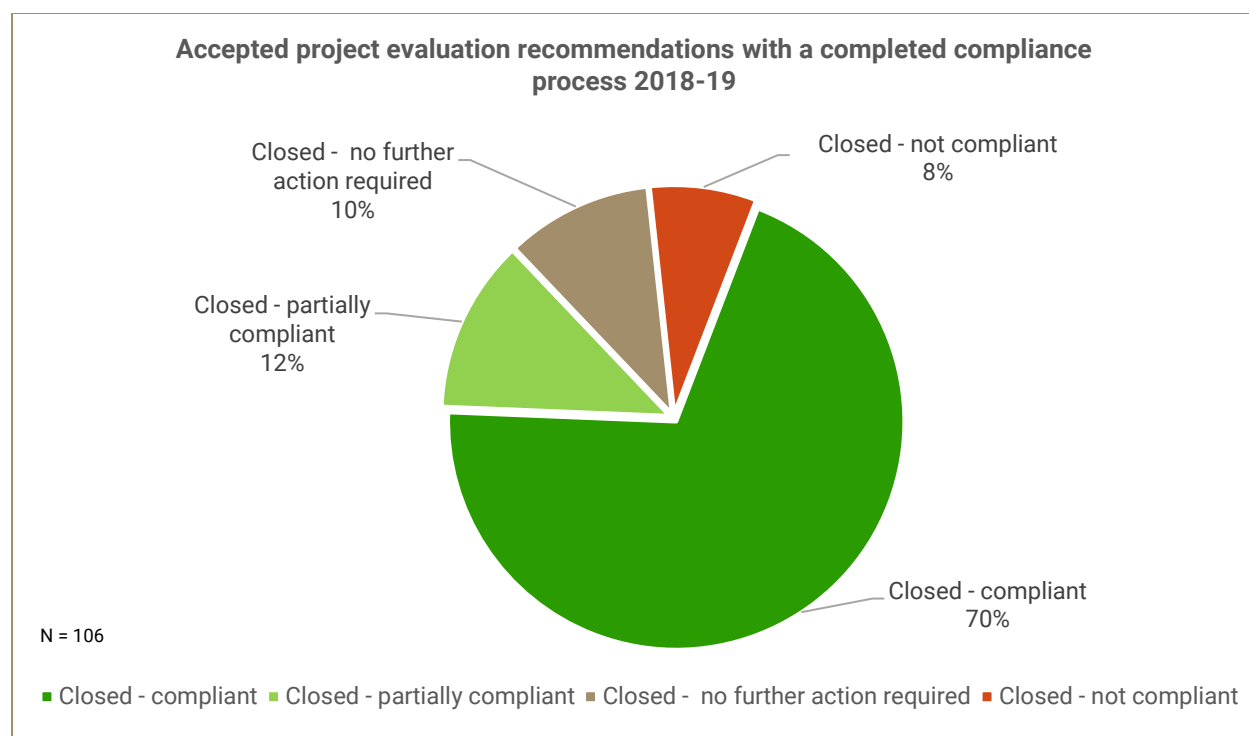
259. Figure 39 below, shows the status of implementation across the 261 project / programme level recommendations issued. 'Open' recommendations are those that were still within their designated implementation timeframe at the end of the biennium. By contrast 'closed' recommendations had either passed their implementation deadline or had already been deemed as being fully implemented (compliant) during the biennium.

Figure 39. Implementation status of evaluation recommendations issued, shown by Division (2018 – 19)



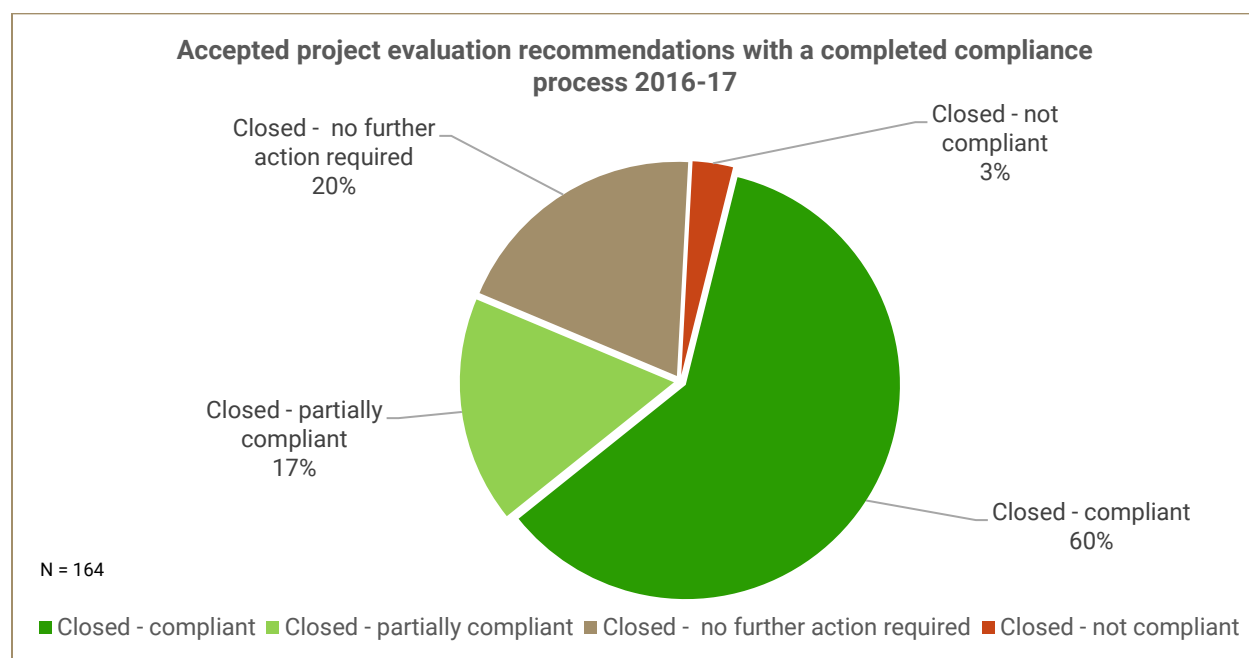
260. Figure 40 shows the compliance performance for project / programme evaluation recommendations that were accepted and where a recommendation implementation plan was completed. This shows that when a management response is provided it is often followed by a strong compliance performance with only 8% of recommendations being closed as not compliant for the biennium.

Figure 40. Accepted evaluation recommendations with a completed compliance process 2018-19



261. The performance in recommendation compliance in the 2018-19 biennium compared with that of the 2016-17 biennium shows broadly similar patterns but with a slight reduction in compliance performance as shown by the increase from 3% to 8% of recommendations closed as 'not compliant'.

Figure 41. Accepted evaluation recommendations with a completed compliance process 2016-17



Corporate and High-Level Recommendations (2018-19)

262. This category of recommendations was introduced midway through the biennium to capture recommendations where the required management actions need to be taken by several actors and at levels beyond the management authority of the project or programme manager that received the implementation plan (for GEF projects this may also include recommendations directed at the GEF Secretariat). The nature of such recommendations means they often take more time to implement. Table 7 shows the current implementation status of these recommendations and it can be seen that the majority of them (82%) have yet to reach their implementation deadlines.

Table 7. Compliance status for Corporate\ High Level Recommendations issued 2018-19

Compliance category	No. of Recommendations
Closed - compliant	2 (6%)
Closed - partially compliant	0
Closed - no further action required	1 (3%)
Closed - not compliant	3 (9%)
Open - partially compliant	11 (31%)
Open - not compliant	18 (51%)

6 Performance of the Evaluation Office and evaluation quality

6.1 The OIOS evaluation dashboard study of UNEP Evaluation Office

263. The Office of Internal Oversight Services of the United Nations Secretariat (OIOS) conducted an assessment of evaluation capacity and practice for every entity in the Secretariat for the 2016-2017 biennium, this was published in April 2019²¹. The 'Evaluation Dashboard' includes an independent assessment conducted by the Inspection and Evaluation Division in OIOS based on 19 indicators of evaluation practice. The indicators used are based on the United Nations norms and standards for evaluation developed by the United Nations Evaluation Group (UNEG) and supported by Member States in General Assembly resolution 67/226.

264. The OIOS Dashboard report noted that the UNEP *"had robust evaluation systems in place across all relevant Evaluation Dashboard indicators. Its function was organized into a stand-alone evaluation unit, and its most senior evaluation professional was at D-1 level. A strong evaluation policy, plan, and procedures were in place."*

265. OIOS also stated in their report to the UN General Assembly on evaluation across the UN Secretariat *"The number of evaluation reports meeting the OIOS screening criteria increased to 335 evaluation reports, up by about 23 per cent from 273 in the previous biennium. This increase was driven primarily by a few high -producing entities, namely the United Nations Entity for Gender Equality and the Empowerment of Women (UN-Women) and the United Nations Environment Programme (UNEP), which each produced 75 evaluations."*²²

6.2 GEF Independent Evaluation Office assessment of the quality of UNEP evaluation reports

266. The Independent Evaluation Office of the GEF publishes an annual comparative assessment of the performance of GEF projects including an assessment of the quality of evaluation reports received from GEF Agencies. In the most recent Annual Performance Report (APR), published in October 2019, UNEP project performance compared favourably against other GEF agencies (primarily UNDP, UNIDO, World Bank and others²³) and assessment of the performance of the UNEP Evaluation Office continues to be very strong. The data presented in the most recent APR also indicates that **100% of UNEP evaluation reports** of GEF projects it recently assessed were rated 'Moderately Satisfactory' or better for quality.

²¹ United Nations Evaluation Dashboard 2016-2017. OIOS Assignment No: IED-19-002, 30 April 2019.

²² Report on the activities of the Office of Internal Oversight Services to the General Assembly. Strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives. A/74/67, 20 March 2019

²³ GEF notes 'other' agencies as including: Asian Development Bank, Food and Agriculture Organization of the United Nations, Inter-American Development Bank, and International Fund for Agricultural Development.

6.3 UN System-Wide Action Plan on gender equality and the empowerment of women (UN-SWAP)

267. Standards for Gender Equality and Empowerment of Women (GEEW) were introduced across the UN network in 2014. The performance against these standards is assessed through 15 GEEW indicators. Fourteen of these indicators assess the progress made within the institution in mainstreaming gender-responsive strategies and systems, the evaluand being the institution itself, including its systems and the actions it has taken to promote gender equality and the empowerment of women. The scores are based on a self-assessment guided by a common approach developed by UN Women, in conjunction with the UN Evaluation Group. The one exception within the set of 15 indicators is the Evaluation Performance Indicator, which is assessed by an external consultant who reviews the coverage of GEEW in all the evaluation reports completed in the preceding year. The UN-SWAP method for measuring performance against the indicators was revised with more specific sub-criteria by which to examine the evaluations. A new set of indicators was applied from 2018, referred to as UN System-Wide Action Plan 2.0, which needs to be considered when comparing the results prior, and since, 2018.

268. In 2018, the review of the Evaluation Performance Indicator included 37, and in 2019, 32 evaluation reports²⁴. The evaluation reports were assessed against three performance criteria: 1) GEEW is integrated in the evaluation scope of analysis and indicators are designed in a way that ensures GEEW related data will be collected; 2) a gender-responsive methodology, methods and tools, and data analysis techniques are selected; and 3) the evaluation findings, conclusions and recommendation reflect a gender analysis, including one linked to meeting the gender-related UN Evaluation Group norms and standards. The performance indicators are rated against a five-point scale ('not applicable'; 'missing'; 'approaches requirements'; 'meets requirements'; and 'exceeds requirements').

269. In 2018, the overall score of the UNEP Evaluation Performance indicator was 5.16 points – “approaches requirements” whereas in 2019, the overall score was 6.53 points – “meets requirements”. This indicates a significant improvement from 2018²⁵. According to the review consultant, the improvement was attributed to UNEP Evaluation Office guidance template, which includes clear gender criteria in all terms of references for evaluations; a greater number of evaluations were found to be responsive to the Evaluation Office guidance template than in the previous year.

270. The review found that challenges remain when the subject matter is highly technical and does not have elements of field impact, or when the project did not collect relevant data during implementation. In addition, there is generally still very little within evaluation reports on the notion of gender empowerment or equality as concepts and evaluations are generally more able to provide quantitative, rather than qualitative, data on gender. With a well-structured evaluation template, the Evaluation Office will now have to find ways to deepen some of the analysis in the evaluations, so that the reporting on gender captures some of the sociological impacts, unanticipated effects, and how project activities affect differentiated groups.

²⁴ The total number of project evaluations assessed for each of these two annual reporting periods (69) exceeds the number of project evaluations summarized in this Biennial Synthesis Report (61) because the annual cut off dates in the UN-SWAP process differ from the cut off dates that define the 'biennium'.

²⁵ The overall score of the Evaluation Performance Indicator has been steadily improving; in 2016 UNEP scored 3.7 on the Evaluation Performance Indicator and improved this to 5.3 in 2017 both scores translating into a rating of “approaching requirements”.

271. A number of key recommendations have been provided in the 2018 and 2019 UN-SWAP reviews of the UNEP Evaluation Performance Indicator as summarised below:

- In cases where projects do not include gender indicators, evaluators should highlight gaps in the reconstructed Theory of Change
- Evaluators of highly technical projects without local level activities need to identify the gender implications of this work
- Evaluators should make gender-relevant actionable recommendations or lessons learned. Project evaluations that have provided recommendations on how to enhance results frameworks and gender indicators should feed into the project design process
- Evaluators should explicitly state their gender methodology and data sources to represent women's voices
- Evaluation Office should require ethics statements in all evaluations to confirm that confidentiality/anonymity commitments have been honoured
- Evaluation reports should present a more holistic background/contextual piece to assist in identifying gender risks and consequences. In cases where there are security concerns in reaching vulnerable communities a more holistic background/contextual piece can assist in capturing some of the challenges faced by marginalized groups

6.4 United Nations Evaluation Group activities

UNEG Evaluation Week 2019, Nairobi Kenya

272. UNEP Evaluation Office, in close collaboration with the UN-Habitat Evaluation Unit, organised and hosted the 2019 'UNEG Evaluation Week' between the 15th and 19th May 2019 at the Nairobi UN compound. The event brings together UN evaluation professionals from across the system including evaluation staff from agencies' headquarters, regions and country offices. UNEG Evaluation Week was very well attended with more than 150 participants representing the evaluation functions of more than 30 UN entities, and those of the GEF, Green Climate Fund, OECD-DAC Evalnet, ADB, AfDB and the World Bank. The agenda included:

- A series of one-day professional development seminars delivered by international experts on the following topics;
 - *Big data for evaluation: How to walk the walk*
 - *Evaluation Theories and Approaches: Applications in evaluation practice*
 - *Approaches to impact evaluation*
 - *Use of Qualitative Analysis Software for Evaluations e.g. NVIVO*
 - *Revisiting the Foundations of Humanitarian Evaluation: Reflections on current practice and future directions*
- A two-day series of peer-learning Seminars – 'The Evaluation Practice Exchange'
 - *This consisted of a series of 16 interactive sessions on contemporary topics, using different (non-PowerPoint) modalities and led by different UNEG member agencies across three themes: 1. Managing Evaluations, 2. Evaluation Standards, Norms and Principles and 3. Evaluation Methods.*

- The final segment of the week on the 18th and 19th May was the Annual General Meeting of the UN Evaluation Group. The meeting was opened by the Deputy Executive Director of UNEP, Joyce Msuya.

Peer Review of UN Evaluation Functions

273. UNEP is an active member of the United Nations Evaluation Group and participates in many of its substantive activities. For example, the UNEP Director of Evaluation Chairs the UNEG Peer Review Working Group and has undertaken reviews of the evaluation functions of UNODC, UNICEF and, during 2018, UNFPA.

Annex 1. List of Evaluations in the 2018-19 Biennial Report

Project evaluations

1. Terminal Evaluation of the UNEP project “The Economics of Ecosystems and Biodiversity - TEEB Phase III, including the project “TEEB National Implementation: Reflecting the Value of Ecosystems and Biodiversity in Policymaking” (funded by the EU Thematic Programme for Environment and Sustainable Management of Natural Resources including Energy (ENRTP), By Sarah Humphrey and Justine Braby (Consultants), January 2018
2. Terminal Evaluation of the UN Environment/GEF Project “Low Carbon-Energy Islands: Accelerating the Use of Energy Efficient and Renewable Energy”, by Mr. Alfredo J Caprile and Ms. Sirikul Prasitpianchai (Consultants), January 2018
3. Terminal Evaluation of the UN Environment/GEF project “Expanding Forest Stewardship Council (FSC) Certification at Landscape Level Through Incorporating Additional Ecosystem Services”, Mr. Trond Norheim (Consultant), January 2018
4. Terminal Evaluation of the UN Environment/Global Environment Facility project “Pilot Project on the Development of Mercury Inventory in the Russian Federation” By Wouter Pronk (Consultant), February 2018
5. Mid Term Evaluation of Energy for Sustainable Development in the Caribbean, by Mr. David Simmons (Consultant), March 2018
6. Follow-up to finalise the Terminal Evaluation of the UN Environment/GEF project: “Support for Implementation of the National Biosafety Framework for Egypt – GFL/2328-2716-4954, By Camillo Risoli (Consultant), March 2018
7. Terminal Evaluation of the UN Environment/GEF project: “Implementation of National Biosafety Framework for Nigeria” GEF ID 3655 By Camillo Risoli (Consultant), April 2018
8. Terminal Evaluation of the UN Environment/GEF project: “Implementation of National Biosafety Framework for Ghana”, GEF ID 3045 By Camillo Risoli (Consultant) May 2018
9. Terminal Evaluation of the UN Environment/GEF project “A Transboundary Waters Assessment Programme: Aquifers, Lake/Reservoir Basins, River Basins, Large Marine Ecosystems, and Open Ocean to Catalyze Sound Environmental Management (TWAP), Mr. Peter Whalley (Consultant), April 2018
10. Terminal Evaluation of the UN/GEF project Reducing Greenhouse Gas Emissions with Bus Rapid Transit and Non-Motorized Transport Gef Id 1917 GFL-2328-2723-4825, By Roland Wong (Consultant), April 2018
11. Terminal Evaluation of the UN Environment Project “Green Climate Fund (GCF) Readiness”, By Mr. Arthur D. Long (Consultant) and Mr. Albert Stocker, May 2018
12. Terminal Evaluation of the UN Environment/GEF projects: Implementation of National Biosafety Framework for Turkey GEF Id 4067, by Emilia Venetsanou (Consultant), May 2018.

13. Terminal Evaluation of the UN Environment/Project “Capacity Building for Implementation of the Cartagena Protocol on Biosafety in India” – Phase II – GEF Id 3751, By Ms. Emilia Venetsanou (Consultant), April 2018
14. Terminal Evaluation of the UN Environment/Project “Reducing Global and Local Environmental Risks from Primary Mercury Mining in Khaidarkan, the Kyrgyz Republic”, By Petr Sharov (Consultant), May 2018
15. Terminal Evaluation of the UN Environment/GEF project: “Implementation of National Biosafety Framework for Liberia”, GEF ID 3040, By Camillo Risoli (Consultant), May 2018
16. Terminal Evaluation of the UNEP project “Improvement of Early Warning System to Reduce Impacts of Climate Change and Capacity Building to Integrate Climate Change into Development Plans” (NAPA Lesotho Evaluation), By Ms. Veronica Nyawira Muthui (Consultant), June 2018
17. Terminal Evaluation of the UN Environment/GEF project Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the COMESA Sub Region (GEF Id 3968), By Mr. Richard Thompson and Mr. Cristobal Vignal (Consultants), June 2018 (On website – one main report and some Exec Summaries in different languages)
18. Terminal Evaluation of the UN Environment/GEF Project Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the SADC Sub-Region (GEF Id 3942), By Mr. Richard Thompson and Mr. Cristobal Vignal (Consultants), June 2018 (On website, above)
19. Terminal Evaluation of the UN Environment/GEF Project Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the ECOWAS Sub-Region (GEF Id 3969), By Mr. Richard Thompson and Mr. Cristobal Vignal (Consultants), June 2018 (On website, above)
20. Mid Term Evaluation of the GEO 6, Mr. David Annandale (Consultant), June 2018
21. Terminal Evaluation of the UN Environment/GEF project “Mitigation Options of Greenhouse Gas (GHG) Emission In Key Sectors in Brazil”, By Ms. Pamela Ransom, Mr. Goncalo Nuno Pera Cavalheiro and Mr. Gustavo de Melo Ribeiro (Consultants), June 2018
22. Terminal Evaluation of the UN Environment/GEF projects: A: Institutional Capacity Building Towards the Implementation of the Biosafety Act 2006 and related obligations to the Cartagena Protocol on Biosafety – Namibia” GEF Id no 3644, by Mr. Segbedzi Norgbey and Mr. Charles Gbedemah (Consultants), August 2018
23. Terminal Evaluation of the UN Environment/GEF Projects: B: Capacity Building for the Implementation of the National Biosafety Framework of Swaziland” GEF Id no 4077. By Mr. Segbedzi Norgbey and Mr. Charles Gbedemah (Consultants), July 2018
24. Terminal Evaluation of the UN Environment/GEF Project” Development and Institution of a National and Control System (Framework) for Living Modified Organisms (LMOs) and Invasive Alien Species (IAS) – Cameroon” – GEF Id 3651, By Mr. Segbedzi Norgbey and Mr. Charles Gbedemah (Consultants), August 2018

25. Terminal Evaluation of the UN Environment/GEF projects: Building Capacity for the Detection and Monitoring of LMOs in Cambodia – GEF Id 3636, by Mr. John Retnon Mauremootoo (Consultant), August 2018
26. Terminal Evaluation: Secretariat Support to the Intergovernmental Negotiating Committee for the Minamata Convention on Mercury, Ms. Linda Ghanime (Consultant), July 2018
27. Terminal Evaluation of the UNEP/GEF project “Phasing Out Incandescent Lamps Through Lighting Market Transformation in Vietnam, By Mr. Brahmanand Mohanty and Mr. Phuc Vinh Tiet (Consultants), September 2018
28. Terminal Evaluation of the UNEP Project on Strategic Approach to International Chemicals Management Secretariat (SAICM), by Mr. Robert Nurick (Consultant), December 2018 (On website.)
29. Terminal Evaluation of the UNEP/GEF project “Demonstrating and Scaling up Sustainable Alternatives to DDT for the control of Vector Borne Diseases in Southern Caucasus and Central Asia (Georgia, Kyrgyzstan, Tajikistan)”, By Ecaterina Melnicenco (Lead Consultant) and Petr Sharov (Supporting Consultant), August 2018
30. Terminal Evaluation of the UN Environment/GEF Project “Demonstration of a regional approach to environmentally Sound Management of PCB Liquid Wastes and Transformers And Capacitors containing PCBs”, By Mr. Cristobal Vignal (Consultant), August 2018
31. Terminal Evaluation of the UN Environment/GEF project “Seed Capital Assistance Facility, Phase I GEF Id 1609 (SCAF I), By Ms. Christine Woerlen (Consultant), August 2018
32. Terminal Evaluation of UN environment/GEF project “Greening the Cocoa Industry – GEF Id 3077, Ms. Maryline Raymonde Jeanne Guiramand-Gois and Mr. Giorgio Brandolini (Consultants), November 2018
33. Terminal Evaluation of the UN Environment/GEF Project On “Lead Paint Elimination Project in Africa”, By Mr. Nee Sun Choong Kwet Yive (Consultant), November 2018
34. Terminal Evaluation of the UN Environment/GEF Project On Establishing the Tools and Methods to Include the Nine New POPs into Global Monitoring Plan, By Mr. Nee Sun Choong Kwet Yive (Consultant), November 2018
35. Terminal Evaluation of the UN Environment/GEF Project On A Regional Focus on Sustainable Timber Management In the Congo Basin, By Mr. Denis Ruysschaert (Consultant), June 2019
36. Terminal Evaluation of the UN Environment/GEF project “Scalingup the SE4ALL Building Efficiency Accelerator (BEA), Ms. Katherine Conway (Consultant), October 2018
37. Terminal Evaluation of the UN Environment/GEF Project “Establishing the Foundations of a Partnership to Accelerate the Global Market Transformation for Efficient Appliances and Equipment (SE4ALL Global Project)”, Mr. Roland Wong (Consultant), November 2018
38. Terminal Evaluation of the UN Environment/GEF project “Market Lighting Transformation in Peru”, Mr. Hugo Navajas and Mr. Douglas Ingaroca (Consultants), October 2018

39. Terminal Evaluation of the UN Environment project “World Heritage Inscription Process as a Tool to Enhance Natural and Cultural Resources Management of the Iraqi Marshlands” Pims Id no. 000547, By Mr. Walid Abed Rabboh (Consultant), December 2018
40. Mid-Term Evaluation of the UN Environment project “Seed Capital Assistance Facility, Phase II, PIMS id 1657 (SCAF II), By Christine Woerlen (Consultant), August 2018
41. Terminal Evaluation of the UNEP Project “Promoting Sustainable Transport Solutions in East African Cities”, Mr. Roland Wong (Consultant), January 2019
42. Terminal Evaluation of the UN Environment/GEF project on “Promoting Energy Efficiency in Buildings in East Africa”, GEF Id no 3788, By Ms. Mari-Louise Van der Walt (Consultant), December 2018
43. Terminal Evaluation of the UN Environment/GEF Project on “Mainstreaming Sustainable Management of Tea Production Landscapes” GEF Id no 5750. By Mr. Trond Norheim (Consultant), December 2018
44. Terminal Evaluation of the Un Environment/GEF project “Integrated Responses to Short Lived Climate Forcers Promoting Clean Energy and Energy Efficiency” GEF Id 4999. By Ms. Pamela Ransom (Lead Consultant); Mr. Goncalo Cavalheiro and Mr. Ramiro Barrios (Consultants), March 2019
45. Terminal Evaluation of the UN Environment project “Delivering Sustainable Development and Enabling the Transition to Greener Economies Through Sustainable Public Procurement PIMs id no. 01609, By Mr. Bjorn Bauer (Consultant), December 2018
46. Terminal Evaluation of the UN Environment/GEF Project “Contributing to the Safe Use of Biotechnology in El Salvador” (also referred to as Implementation of National Biosafety Frameworks for El Salvador), GEF Id no 3332, By Ms. Tea Garcia-Huidobro (Consultant), May 2019
47. Terminal Evaluation of the UN Environment/GEF project “Regional Project for Implementing National Biosafety Frameworks in the Caribbean Sub-Region” by Mr. Camillo Risoli (Consultant), May 2019
48. Terminal Evaluation of the UN Environment/GEF project: Access to and Benefit Sharing and Protection of Traditional Knowledge to Promote Biodiversity Conservation and Sustainable Use” (ABS Guatemala”), By Mr. Camillo Risoli (Consultant), April 2019.
49. Terminal Evaluation of the UN Environment/GEF project “Implementing Integrated Water Resources and Wastewater Management in Atlantic and Indian Ocean SIDS”. By Mr. Gaetan Quesne and Mr. James Lenoci (Consultants), June 2019
50. Terminal Evaluation of the UN Environment/GEF project “Development of a National Implementation Plan in Brazil as a first step to implement the Stockholm Convention on Persistent Organic Pollutants (POPs)”, by Mr. Cristobal Vignal (Consultant), January 2019
51. Mid-Term Evaluation – UN Environment Project Mediterranean Investment Facility (MIF), by Ms. Christine Woerlen and Ms. Martina Greib (Consultants), January 2019 (a.k.a “Creating Enabling Conditions for Renewable Energy and Energy Efficiency Investment” PIMS ID No 1715)

52. Terminal Evaluation of the UN Environment project “Promoting Peace Over Natural Resources in Darfur and Kordofan, By Mr. Amr Abdalla (Consultant), June 2019
53. Terminal Evaluation of the UN Environment/GEF project on “Development of Mercury Risk Management Approaches in Latin America” GEF Id No 5494, By Mr. Jorge Leiva (Consultant), August 2019
54. Terminal Evaluation of the UN Environment/GEF Project on “Global Project on the Updating of National Implementation Plans for POPs” (covering GEF Ids no 5307 & 5525 & related Country Level Enabling Activities. By Mr. Cristobal Vignal, Ms. Suman Lederer, Mr. Katim Touray and Mr. Petr Sharov (Consultants), June 2019
55. Terminal Evaluation of the UN Environment/GEF Project on “Best Practices for PCB Management in the Mining Sector of South America” GEF Id no 3814, By Mr. Jorge Leiva (Consultant), August 2019
56. Terminal Evaluation of UN Environment/GEF project “Multiplying Environmental and Carbon Benefits in High Andean Ecosystems” GEF Id 4750 & TOR 2 – to provide information and Related materials on the selected topics for PROVIA (Science Div), By Robert Hofstede (Consultant), September 2019
57. Mid-Term Evaluation of the UN Environment project “Clean Seas Campaign” contributing to the Global Partnership on Marine Litter project (GPML) PIMs Id no 01882, By Mr. Stephen Tibbet and Mr. Jeremy Smith (Consultants), October 2019
58. Terminal Evaluation of the UN Environment/GEF project “Development of Tools to Incorporate Impacts of Climatic Variability and Change, in Particular Floods and Droughts, into Basin Planning Processes”, By Mr. Peter Whalley and Mr. Christopher Gordon (Consultants), August 2019
59. Terminal Evaluation of the UN Environment/GEF project “Sharing Knowledge on the Use of Biochar for Sustainable Land Management, or Biochar for Sustainable Soils (B4SS)”, By Ms. Justine Braby (Consultant), October 2019

Strategic/programme evaluations

60. Sub-Programme Evaluation on Resource Efficiency and Sustainable Consumption and Production, Mr. Marcel Crul, Mr. Dick Van Beers and Mr. Derek Eaton (Consultants), September 2018
61. UN SWAP Review for 2018, by Ms. Erum Hasan (Consultant), December 2018.

Annex 2. List of evaluations in 2018-19 rated in the satisfactory range for responsiveness to human rights and gender equity

Projects that scored a 'Highly Satisfactory' Rating in 2018-19

1. Terminal Evaluation of the UN Environment / Global Environment Facility Project "Expanding Forest Stewardship Council (FSC) Certification at landscape level through incorporating additional ecosystem services" (GEF ID 3951)
2. Terminal Evaluation of the UN Environment Project Promoting Peace over Natural Resources in Darfur and Kordofan
3. Terminal Evaluation of the UN Environment Project "ABS Guatemala: Access to and Benefit Sharing and Protection of Traditional Knowledge to Promote Biodiversity Conservation and Sustainable Use"
4. Terminal Evaluation of UN Environment/GEF Project "Multiplying Environmental and Carbon Benefits in High Andean Ecosystems" EF 4750

Projects that scored a 'Satisfactory' Rating in 2018-19

1. Terminal Evaluation of the Global Environment Facility/UN Environment Project: "Low Carbon-Energy Islands: Accelerating the Use of Energy Efficient and Renewable Energy Technologies in Tuvalu, Niue and Nauru"
2. Terminal Evaluation of the UN Environment Project: "Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the COMESA, Sub-regions"
3. Terminal Evaluation of the UN Environment Project: "Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the SADC Sub-regions"
4. Terminal Evaluation of the UN Environment Project: "Capacity Strengthening and Technical Assistance for the Implementation of Stockholm Convention National Implementation Plans (NIPs) in African Least Developed Countries (LCDs) of the ECOWAS Sub-regions"
5. Terminal Evaluation of the UN Environment/Global Environment Facility Project "Capacity Building for the Implementation of the National Biosafety Framework of Swaziland"
6. "Terminal Evaluation of the UN Environment / GEF project: Lighting Market Transformation in Peru"
7. Mid-Term Evaluation of the UN Environment Project: Global and Regional Integrated Environmental Assessments ("GEO-6")
8. Terminal Evaluation of 'The Economics of Ecosystems and Biodiversity (TEEB)' Phase III and ENRTP-funded sub-component: National Implementation: Reflecting the Value of Ecosystems and Biodiversity in Policymaking

9. Terminal Evaluation of the UN Environment-GEF Project “Demonstration of a Regional Approach to Environmentally Sound Management of PCB Liquid Wastes and Transformers and Capacitors Containing PCBs”
10. Terminal Evaluation of the GEF-UN Environment Project “Establishing the Tools and Methods to Include the Nine New POPs into Global Monitoring Plan”
11. "Terminal Evaluation of the UN Environment / GEF Project: Integrated Responses to Short-Lived Climate Forcers Promoting Clean Energy and Energy Efficiency"
12. Terminal Evaluation of the UNEP/GEF Project “BIOCHAR FOR SUSTAINABLE SOILS B4SS” (GEF ID: 5824)

Projects that scored a ‘Moderately Satisfactory’ Rating in 2018-19

1. Terminal Evaluation of the UN Environment/GEF Project: “Scaling up the Sustainable Energy for All Building Efficiency Accelerator” (2016–2017), GEF ID 9329”
2. "Terminal Evaluation of the Global Environment Facility/UN Environment Project: “Transboundary Waters Assessment Programme (TWAP)”"
3. "Terminal Evaluation of the UN Environment project: “Pilot project on the Development of Mercury Inventory in the Russian Federation” "
4. Terminal Evaluation of the UN Environment Project: “Support for Implementation of the National Biosafety Framework for Egypt”
5. Terminal Evaluation of the UN Environment/Global Environment Facility Project: “Development and Institution of a National Monitoring and Control System (Framework) For Living Modified Organisms (LMOs) and Invasive Alien Species (IAS) - Cameroon”
6. Terminal Evaluation of the UN Environment Project: “UN Environment/UNDP/WRI Green Climate Fund (GCF) Readiness Programme”
7. Terminal Evaluation of the UN Environment Project: Secretariat Support to the Intergovernmental Negotiating Committee for the Minamata Convention on Mercury
8. Terminal Evaluation: UN Environment/Global Environment Facility Project “Seed Capital Assistance Facility, Phase I (Renewable Energy Enterprise Development)”
9. "Terminal Evaluation of the UN Environment Project: “Greening the Cocoa Industry” "
10. Terminal Evaluation of the GEF-UN Environment Project “Lead Paint Elimination Project in Africa”
11. Mid-Term Evaluation of the UN Environment Project “Seed Capital Assistance Facility, Phase II”
12. Terminal Evaluation of the UN Environment–Global Environment Facility Project “Reducing global and local environmental risks from primary mercury mining in Khaidarkan, the Kyrgyz Republic”
13. "Terminal Evaluation of the UN Environment Project: “Global Project on the Updating of National Implementation Plans for POPs” Vol 1"

14. "Terminal Evaluation of the GEF/ UN Environment / UNDP Project: "Implementing Integrated Water Resources and Wastewater Management in Atlantic and Indian Ocean SIDS""
15. Terminal Evaluation of the UN Environment - GEF Project "Contributing to the Safe Use of Biotechnology in El Salvador"
16. Terminal Evaluation of the Global Environment Facility/UN Environment Project: Development of Tools to Incorporate Impacts of Climate Variability and Change, in Particular Floods and Droughts, into Basin Planning Processes