



GEF Evaluation Office

Annual Performance Report 2012

Approach Paper

Contacts:

Neeraj Kumar Negi

Senior Evaluation Officer GEF EO. Nnegi1@thegef.org

GEF Evaluation Office
1818 H Street NW
Washington DC 20433
Tel: 202 473-2259
Fax: 202 522-1691
www.gefeo.org

A. Issues covered by Annual Performance Report

1. The *Annual Performance Report* (APR) is a regular feature of the GEF Evaluation Office's (GEF EO's) performance evaluation stream of work. The report presents a detailed account of the performance of the Global Environment Facility (GEF) portfolio in terms of results; processes that may affect results; and, monitoring and evaluation (M&E) arrangements.
2. The issues that will be addressed in the APR 2012 include assessment of project outcomes, risks to sustainability of outcomes, quality of M&E arrangements in completed projects, quality of terminal evaluation reports, and management action records. The report would also provide information on the progress on the ongoing work of the performance evaluation stream. The APR 2012 would also be prepared as an input to the Fifth Overall Performance Study (OPS5) – APR 2012, including its intermediary products, will feed into both the first and the second report of OPS5.
3. In the APR reporting on issues such as outcome achievements and factors that affect outcome achievements, risks to sustainability of outcomes, and quality of M&E arrangements, is primarily based on the evidence provided in the terminal evaluations of the completed GEF projects. The GEF Agencies are required to prepare these reports at operational completion of all GEF projects and programs following the terminal evaluation guidelines. The APR 2012 report would provide a detailed account of the performance of the cohort of completed projects for which terminal evaluations and, where applicable, their independent reviews were submitted during the past year. To assess trends, the performance of the APR 2012 cohort would be compared with that of the preceding cohorts.
4. The management action records (MARs) keep track of the level of adoption of GEF Council decisions, which were based on the findings, conclusions and recommendations presented in the Council working documents prepared by the GEF EO, by relevant partners within the GEF partnership. APR 2012 will present an updated account of the adoption of decisions that have not yet graduated from the MARs and would also cover the relevant Council decisions made during fiscal year 2012.
5. During the fiscal year 2013 a review on quality of supervision for GEF projects that are under implementation would be initiated. This would be the third quality of supervision review that the Office would undertake. Findings of the earlier reviews were presented in APR 2006 and APR 2009. The pilot review (reported on in APR 2006) covered a representative sample of GEF projects implemented by the World Bank, UNDP and UNEP. The follow up review (APR 2009) tracked changes in the quality of supervision. These reviews also contributed to better understanding of the supervision practices in these agencies. The third iteration of the quality of supervision review (FY 2013-14) will cover the agencies that have so far not been covered by the quality of supervision review (i.e. implementing agencies other than the World Bank, UNDP and UNEP). Earlier, the other agencies had not been covered because their GEF portfolio was quite small. This constraint has been mitigated to a large extent as in the interim the portfolios of these agencies have increased significantly. While reporting on performance of some of these agencies might still be difficult due to small size of the portfolio, it would be possible to report on performance of some of the agencies and their performance as a group. This would help the Office plug an important gap as the supervision practices of other agencies have not been reviewed by the Office so far. Another reason to focus primarily on the agencies that have not yet been covered was to keep the costs of undertaking the review low. The quality of supervision review would be completed in July 2013; the progress on the review would be reported on in APR2012.

6. Two mid-term evaluations were started during the fiscal year 2013. The Mid Term Evaluation of the System for Transparent Allocation of Resources (STAR) will assess STAR's design and its performance in meeting its objectives. The STAR provides country allocations for biodiversity, climate change and land degradation, and has now been under implementation for more than two years. STAR's performance would also be compared with that of the Resource Allocation Framework (RAF), which was implemented in GEF-4. The other midterm evaluation covers the National Portfolio Formulation Exercises (NPFES). The NPFE initiative was started in 2010 and has been under implementation for more than two years. As part of this initiative GEF provides a support of US \$ 30,000 to countries that agree to undertake a NPFE to identify activities that are consistent with the national and GEF priorities to utilize their indicative country allocations. The GEF support covers costs involved in holding of necessary meetings, broader consultations, and bringing on board requisite expertise to conduct NPFES. The Mid Term Evaluation of NPFES will assess the implementation progress, uptake of NPFES and the extent these exercises are serving their intended purpose. The mid-term evaluations on STAR and NPFE would be completed in August 2013. APR 2012 will report on the progress on these evaluations. The draft approach papers for these evaluations may be accessed at the GEF EO website (<http://www.thegef.org/gef/OngoingPerformance>).

B. Methodological Considerations for APR 2012

Reporting on completed projects

Defining the APR 2012 Cohort

7. The assessment on performance of completed projects presented in the APR is primarily based on the evidence provided in the terminal evaluation reports. In addition to the completed projects that have been reported on in preceding APRs, APR 2012 would report on 78 additional completed projects for which terminal evaluations have been submitted to the Evaluation Office up to 30th September 2012. Some recently completed projects for which terminal evaluations were submitted before 30th September 2012 were dropped from the APR 2012 cohort despite submission of the terminal evaluations because the respective evaluation offices of the agencies were undertaking independent reviews of the terminal evaluations for these but had not completed the process by January 31st 2013¹. Annex 1 provides the list of 78 projects that are included in the APR 2012 cohort.

8. Of the 78 projects for which terminal evaluations are included in the APR2012 cohort, 51 were implemented by the United Nations Development Programme (UNDP), 12 by the United Nations Environment Programme (UNEP), six by the World Bank Group, three by the Inter-American Development Bank (IADB) and three by United Nations Industrial Development Organization (UNIDO). Three projects were jointly implemented by the agencies. Table 1 provides summary statistics on the completed projects included in the APR2012 cohort.

¹ This was especially true for the terminal evaluations submitted by the World Bank Group. Several of the implementation completion reports from the Bank were dropped from the cohort because the IEG was in the process of preparing independent reviews for these but had not completed the process by January 31st 2013. As the reviews for these projects become available, these will be included in the future APR cohorts.

Table 1: Terminal Evaluations for Completed Projects Submitted

Agency	Full Size Projects	Medium Size Projects	Total TE for APR 2012
IADB	3	0	3
UNDP	26	25	51
UNEP	7	5	12
UNIDO	1	2	3
World Bank Group	6	0	6
Joint Projects	3	0	3
Total	46	32	78

Rating Scales

APR 2012 will report on outcomes achievements, risks to sustainability of outcome achievements, quality of M&E arrangements, and quality of implementation and execution of completed projects. These risks to sustainability of outcome achievements would be rated on a four point scale and others would be rated on a six point scale. The APR would also assess quality of terminal evaluation reports. The use of balanced scale would preclude middle values. Annex 2 provides more information on the rating scales.

Terminal Evaluation Review Process

9. The Evaluation Office prepares reviews for some of the terminal evaluation reports. The Office intends to prepare reviews for about 30 terminal evaluations. For the terminal evaluations that would not be reviewed, the GEF EO would adopt the ratings provided by the independent evaluation offices of the implementing agencies.

10. The terminal evaluation reviews verify the ratings on outcome achievements, risks to sustainability, implementation and execution, and M&E, provided in the terminal evaluations. The reviews also assess the quality of the terminal evaluations in terms of the extent these evaluations address the project objectives and outcomes specified in the project document, its comprehensiveness, internal consistency, and information provided on topics such as financing, M&E and lessons.

11. The terminal evaluation reports are reviewed by Evaluation Office staff using a set of detailed guidelines to ensure the application of uniform criteria. When deemed appropriate, a reviewer may propose to upgrade or downgrade project ratings provided in the terminal evaluation report.

12. The reviews are examined by a peer reviewer, who provides feedback on the draft reviews. The feedback by the peer reviewer is incorporated by the primary reviewer in subsequent versions of the review. When the ratings of projects are downgraded below moderately satisfactory (for outcomes) or below moderately likely (for sustainability), a senior evaluation officer in the Evaluation Office also examines the review to ensure that the new ratings are justified. The reviews are then shared with the Implementing Agencies, and, after considering their feedback, the reviews are finalized.

13. If the terminal evaluation reports provide insufficient information to make an assessment of or verify Agency ratings on outcomes, sustainability, or quality of project M&E systems, the Evaluation Office classifies the projects as “unable to assess” and excludes them from any further analysis on the respective dimension.

14. Among the ratings that are presented in the APR, the rating on outcome achievements of completed projects is probably the most important. While a complete picture of final outcome ratings for the APR 2012 cohort of completed projects would be available only after the terminal evaluation reviews, including reviews by evaluation offices of agencies that have been adopted by the GEF EO, become available, ratings provided in the terminal evaluations does provide some indication of how the overall picture of the portfolio might look like. Table 2 provides a summary of the ratings for completed projects of different APR cohorts.

Table 2: Projects with Outcome ratings in the satisfactory range (number of rated projects in parentheses)

Cohorts	Number of completed projects in the cohort	PIR rating	Terminal Evaluation Rating	IA evaluation office rating	GEF EO rating or adopted IA evaluation office rating
APR 2005	41	100% (38)	97% (30)	100% (12)	82% (39)
APR 2006	66	98% (48)	94% (35)	90% (20)	84% (64)
APR 2007	41	100% (36)	84% (34)	83% (12)	73% (40)
APR 2008	62	95% (61)	93% (54)	83% (23)	84% (62)
OPS4 cohort	210	98% (183)	92% (153)	88% (67)	80% (205)
APR 2009	55	98% (52)	100% (51)	91% (46)	91% (55)
APR 2010	46	98% (44)	93% (42)	89% (28)	91% (46)
APR 2011	102	96% (98)	85% (88)	80% (87)	81% (102)
APR 2012	78	97% (77)	93% (67)	—	—
OPS5 cohort	281	97% (271)	91% (238)	—	—

Adoption of Agency Evaluation Office Ratings

15. The GEF EO encourages independent review of the findings and ratings, and assessment of quality of terminal evaluations by independent evaluation offices of the agencies. Several GEF Agencies have established a process wherein their independent evaluation office reviews and validates terminal evaluations and assesses the quality of terminal evaluation reports. In instances, where there is sufficient track record to indicate convergence in the ratings provided by the independent evaluation office of an agency and the GEF EO, the GEF EO accepts the ratings provided by the evaluation office of such agencies. Adoption of the ratings provided by the independent evaluation offices of the implementing agencies reduces duplication of work involved in the reviews.

16. Where a GEF Agency does not have an independent evaluation unit or lacks an independent review process, the GEF EO reviews the terminal evaluation reports. The GEF EO also reviews those terminal evaluations where despite an agency having an independent evaluation unit and an independent review process some terminal evaluations that pertain to joint projects or to MSPs may not have been covered through the independent review process or where for some of the reviews the outcome ratings provided by the evaluation unit are not compatible with the rating scales used by the

Office.² The GEF EO may also conduct additional reviews to assess whether the ratings provided by the evaluation offices of the agencies continue to be consistent with those provided by the GEF EO.

17. Table 3 provides a comparison of the outcome ratings given by different providers. The table shows that the outcome ratings provided in the PIRs tend to be more optimistic than by other providers. It also shows that the difference in percentage of projects rated in the satisfactory range in terms of their outcome ratings by the evaluation offices of the implementing agency and those provided by the GEF Evaluation Office is quite small (3.6 percent).

Table 3: Comparison of ratings by different providers/provider groups (row vs column)

Rating comparison	PIR Outcome Rating*	TE Outcome Rating	IA EO Outcome Rating	GEF EO Outcome Rating
PIR Outcome Rating*	—	5.5 % (309)	10.3% (213)	13.0% (261)
TE Outcome Rating	– 5.5% (309)	—	4.8% (209)	11.3% (231)
IA EO Outcome Rating	– 10.3% (213)	– 4.8% (209)	—	3.6% (111)
GEF EO Outcome Rating	– 13.0% (261)	– 11.3% (231)	– 3.6% (111)	—

*The rating on progress towards objectives provided in the last PIR for a completed project is used as a proxy for outcome ratings. This rating is provided by the project team of the implementing agency in its annual reporting to the GEF.

18. Table 4 compares the difference in the ratings provided by the GEF Evaluation Office and those provided by the evaluation offices of the agencies. It shows there is little difference across the agency evaluation offices in terms how closely the ratings provided by the GEF EO tracks the ratings by them. Nonetheless, in terms of observations there is more data on the World Bank projects than on projects by other agencies. The Evaluation Office would continue to track ratings provided by the evaluation offices of agencies. This would allow it to correct for the small bias that the data available so far indicates.

Table 4: Comparison of Outcome rating by GEF EO and agency evaluation office rating

Implementing Agency	Total number of observations	Percentage rated MS or above by agency evaluation offices	Percentage rated MS or above by the GEF EO	Difference
ADB	1	100.0% (1)	100.0% (1)	0.0% (0)
UNDP	20	85.0% (17)	80.0% (16)	5.0% (1)
UNEP	28	92.9% (26)	89.3% (25)	3.6% (1)
World Bank	62	87.1% (54)	83.8% (52)	3.2% (2)
Total	111	88.2% (98)	84.6% (94)	3.6% (4)

*Percentages may not add up due to rounding off.

² As is the case with outcome ratings, the GEF EO adopts the ratings provided by the independent evaluation offices of the agencies on other parameters only when these have been provided using an approach that is congruent with the GEF EO approach. However, absence of congruence in the rating approach for ratings other than the outcome rating may not lead to a re-review by the GEF EO.

19. Based on the information presented in table 2, 3 and 4, it is likely that the overall figure – in terms of percentage of completed projects of the OPS5 cohort whose outcomes are rated in the satisfactory range by the GEF Evaluation Office – would be quite close to the figures for OPS4 cohort.

Management Action Records (MARs)

20. Management action records (MARs) facilitate review and follow-up on the implementation status of evaluation recommendations that have been accepted by management (that is, the GEF Secretariat and/or the GEF Agencies) and/or the GEF Council. For each MAR, the Evaluation Office completes the columns pertaining to recommendations, management responses, and Council decisions. The management is then invited to provide a self-rating of the level of adoption of Council decisions and add any comments as necessary. After the management's response is included in a MAR, the Evaluation Office verifies actual adoption and provides its own ratings, with comments, in time for presentation to the Council. The GEF Evaluation Office rates adoption of the relevant Council decision using a four point scale: high adoption; substantial adoption; medium adoption; and negligible adoption. The Office may also choose not to rate the level of adoption of a Council decision. When it so chooses, the Office provides a reason for it. These reasons could be: it's not possible to verify the level of adoption; the given Council decision is no longer relevant; assessment of adoption of the decision is not applicable in this case.

21. For APR 2012 MARs would be prepared for all the Council decisions based on GEF EO evaluations taken during or before the November 2012 Council meeting but not yet graduated from the MAR process. In APR 2011, 12 Council Decisions had been tracked. Of these two had been graduated due to full adoption. For the APR 2012, in addition to the un-graduated decisions for the APR 2011 relevant Council decisions taken during June 2012 and November 2012 Council meetings would also be tracked.

Performance Matrix

22. The performance matrix provides a summary of the performance of GEF Implementing and Executing Agencies and of the GEF Secretariat on relevant parameters. Presently, the matrix tracks performance on 10 parameters. For most of these parameters—including *project outcomes*, *implementation completion delays*, *materialization of co-financing*, *quality of M&E during project implementation*, and *quality of project terminal evaluations*—the Office already assesses performance on an annual basis. On these parameters the ratings for agencies would be updated. For quality of supervision and most of the parameters related to quality of M&E, the ratings would not be updated as new data on these parameters would not be available.

Quality of Supervision

23. Supervision is understood as *the identification and tracking of, and response to, risks and other issues affecting project implementation and achievement of project objectives*. Quality of supervision review will assess how GEF projects being implemented by the agencies are being supervised. The approach followed to assess quality of supervisions would be the same as was followed for the pilot review reported on in APR 2006 and the follow review undertaken in 2009. The three key dimensions that are considered in the assessment are: focus on results; supervision inputs and processes; and, candor and quality of project performance reporting. Performance on these parameters will be rated on a six point scale ranging from highly satisfactory to highly unsatisfactory.

24. The review would cover projects that were under implementation during the FY 2011-12. Earlier quality of supervision reviews (pilot review in 2006 and follow up review in 2009) have focused entirely on the three major implementing agencies – World Bank, UNDP and UNEP. To economize on the costs involved and to facilitate better understanding of the supervision processes in other agencies, the focus for this review would be on the full size projects being implemented by other agencies. According to Annual Monitoring Report 2012(GEF/C.43/05/Rev.01) during the FY2012, 747 projects – including 551 full size and 196 medium size projects – were under implementation. Of these 134 projects including 104 full size projects were being implemented by agencies other than World Bank, UNDP and UNEP. A representative sample of about 30 projects would be drawn for the review.

25. The quality of supervision review is expected to be completed in FY2014, in time for its findings to be reflected in OPS5. APR 2012 will report on the progress made in the evaluation. Its full results would be presented in APR 2013. Annex 3 provides more details on the approach that would be followed.

c. Contents of the APR 2012

26. The APR 2012 will include the following chapters.

Chapter 1. Conclusions and Recommendations

27. This chapter will present the background, main conclusions, and recommendations of the study and issues for the future.

Chapter 2. Scope and Methodology

28. This chapter will include a detailed discussion of the scope of the APR and the methodology used in its assessments. The project terminal evaluation reports submitted by the GEF Agencies to the Evaluation Office will continue to be the core source of information for much of the APR, particularly for topics that are reported on annually. This chapter will also briefly describe the methodologies adopted for terminal evaluation reviews and for assessment of quality of supervision.

Chapter 3. Results

29. This will be the fourth year that the APR will report on the results of completed GEF projects. This chapter will provide an overview of the extent to which GEF projects are achieving their stated objectives. It will also report the aggregated verified ratings of terminal evaluations of completed projects, which will consist of the Evaluation Office's assessment of the extent to which the completed projects for which GEF Agencies submitted terminal evaluation reports for APR 2012 achieved their expected outcomes and the likely sustainability of the achieved outcomes.

Chapter 4. Factors Affecting Results

30. Past portfolio performance reviews (the predecessor of the APR) and other evaluations conducted by the Evaluation Office have identified specific issues that seem to affect project attainment of outcomes. In this chapter, the Office will report on factors that lead to poor performance and delays in project completion based on findings from the terminal evaluation report reviews. Information on co-financing will be presented in a separate chapter.

Chapter 5. Quality of terminal evaluation

31. This chapter will report on the quality of the terminal evaluation reports based on the year these reports were completed.

Chapter 6. Management Action Record

32. The APR 2012 will provide an overview of the management action record of Council decisions regarding Evaluation Office reports and evaluations. This chapter will provide an overview of the progress made and will identify issues that require Council attention.

Chapter 7. Ongoing work of the performance stream

33. This chapter would report on the ongoing work of the performance evaluation stream. This includes reporting on the progress in the review on quality of supervision, and mid-term evaluations on STAR and NPFs. It would also report on the progress made on the guidance for agencies for preparation of terminal evaluation and PMIS related work.

E. Dissemination of Findings

34. The APR is largely intended for the GEF Council and a GEF corporate audience, including the GEF Secretariat, STAP and the GEF Agencies. The report will be delivered in time for the June 2013 Council meeting. At that point, it will have been circulated among and discussed with GEF Secretariat and GEF Agency staff. Thus, by the time the report is presented during the Council meeting and discussed by the Council, the most relevant audiences would have already read it, commented on it, and identified ways to address its findings. The report will be published in hard copy and distributed via email among the GEF Council members, GEF country focal points, GEF Secretariat, Agency staff with GEF responsibilities, and members of other interested organizations. The report will also be posted on the Evaluation Office Web site.

35. The Office may develop knowledge products and services targeted to specific audiences. Specific audiences for knowledge products that may be developed from this APR will be identified. The APR team will work with the knowledge sharing team of the Office to develop and disseminate such knowledge products and services.

F. Calendar of Activities

36. Table 5 presents the calendar of activities for the APR 2012.

Table 5: Calendar of Activities

Process milestone	Schedule
Approach paper	Jan 2013
Terminal evaluation reviews	Jan to Feb 2013
Draft Report of Preliminary Findings	April 15, 2013
Presentation of draft APR at Inter-Agency Meeting	Mid- April 2013
Council working/information document uploaded	May 15 th 2013
Presentation of report at June 2013 Council	June 2013
Publication and knowledge sharing	October 2013

Annex 1: List of Projects included in APR 2012 cohort

GEF_ID	Name	Agency	Type	Focal Area
963	Environmental Protection and Maritime Transport Pollution Control in the Gulf of Honduras	IADB	FP	IW
1515	Consolidation of Ecosystem Management and Biodiversity Conservation of the Bay Islands	IADB	FP	BD
2686	Integrated Management of the Montecristo Trinational Protected Area	IADB	FP	BD
503	Paraguayan Wildlands Protection Initiative	UNDP	FP	BD
668	Coastal and Wetland Biodiversity Management at Cox's Bazar and Hakakuki Haor	UNDP	FP	BD
776	Conservation and Sustainable Use of Medicinal Plants in Arid and Semi-arid Ecosystems	UNDP	FP	BD
834	Promoting Biodiversity Conservation and Sustainable Use in the Frontier Forests of Northwestern Mato Grosso	UNDP	FP	BD
843	Removal of Barriers to Rural Electrification with Renewable Energy	UNDP	FP	CC
1029	Renewable Energy Technology Development and Application Project (RETDAP)	UNDP	MSP	CC
1036	Conservation of "Tugai Forest" and Strengthening Protected Areas System in the Amu Darya Delta of Karakalpakstan	UNDP	MSP	BD
1043	Establishing Conservation Areas Landscape Management (CALM) in the Northern Plains	UNDP	FP	BD
1100	Community-based Conservation of Biological Diversity in the Mountain Landscapes of Mongolia's Altai Sayan Ecoregion	UNDP	FP	BD
1104	Conservation of the Montane Forest Protected Area System in Rwanda	UNDP	FP	BD
1128	Biodiversity Management in the Coastal Area of China's South Sea	UNDP	FP	BD
1137	Promoting the Use of Renewable Energy Resources for Local Energy Supply	UNDP	FP	CC
1148	In-Situ Conservation of Kazakhstan's Mountain Agrobiodiversity	UNDP	FP	BD
1177	Biodiversity Conservation in the Russian Portion of the Altai-Sayan Ecoregion	UNDP	FP	BD
1246	Partnerships for Marine Protected Areas in Mauritius	UNDP	MSP	BD
1308	Strategic Planning and Design for the Environmental Protection and Sustainable Development of Mexico	UNDP	MSP	MFA

GEF_ID	Name	Agency	Type	Focal Area
1338	South Africa Wind Energy Programme (SAWEP), Phase I	UNDP	FP	CC
1343	Demonstrations of Integrated Ecosystem and Watershed Management in the Caatinga, Phase I	UNDP	FP	MFA
1399	Capacity Building for Implementation of Malaysia's National Biosafety Framework	UNDP	FP	BD
1557	Removing Barriers to the Reconstruction of Public Lighting (PL) Systems in Slovakia	UNDP	MSP	CC
1612	Second National Communication of Brazil to the UNFCCC	UNDP	FP	CC
1713	Improved Management and Conservation Practices for the Cocos Island Marine Conservation Area	UNDP	MSP	BD
1725	Biodiversity Conservation in Altos de Cantillana	UNDP	MSP	BD
1854	Biodiversity Conservation and Sustainable Development in the Gissar Mountains of Tajikistan	UNDP	MSP	BD
1899	Regional Programme on Electrical Energy Efficiency in Industrial and Commercial Service Sectors in Central America	UNDP	FP	CC
2068	Integrating Protected Area and Landscape Management in the Golden Stream Watershed	UNDP	MSP	BD
2104	Catalyzing Sustainability of the Wetland Protected Areas System in Belarusian Polesie through Increased Management Efficiency and Realigned Land Use Practices	UNDP	FP	BD
2107	Removing Barriers to Energy Efficiency Improvements in the State Sector in Belarus	UNDP	FP	CC
2193	Enabling Sustainable Dryland Management Through Mobile Pastoral Custodianship	UNDP	MSP	LD
2257	Demonstration of Fuel Cell Bus Commercialization in China, Phase 2	UNDP	FP	CC
2440	Sustainable Land Management in Drought Prone Areas of Nicaragua	UNDP	FP	LD
2492	Strengthening the Protected Area Network (SPAN)	UNDP	FP	BD
2509	Sustainable Land Management for Combating Desertification (Phase I)	UNDP	FP	LD
2589	Institutionalizing Payments for Ecosystem Services	UNDP	FP	BD
2730	Conservation of Globally Important Biodiversity in High Nature Value Semi-natural Grasslands through Support for the Traditional Local Economy	UNDP	MSP	BD
2800	Developing Institutional and Legal Capacity to Optimize Information and Monitoring System for Global Environmental Management in Armenia	UNDP	MSP	MFA

GEF_ID	Name	Agency	Type	Focal Area
2836	Conservation and Sustainable use of Biodiversity in the Kazakhstani Sector of the Altai-Sayan Mountain Ecoregion	UNDP	FP	BD
2848	Improved Conservation and Governance for Kenya Coastal Forest Protected Area System	UNDP	MSP	BD
2863	Ensuring Impacts from SLM - Development of a Global Indicator System	UNDP	MSP	LD
2915	CPP Namibia: Adapting to Climate Change through the Improvement of Traditional Crops and Livestock Farming (SPA)	UNDP	MSP	CC
3062	Strengthening Institutional Capacities for Coordinating Multi-Sectoral Environmental Policies and Programmes	UNDP	MSP	MFA
3068	Mainstreaming the Multilateral Environmental Agreements into the Country's Environmental Legislation	UNDP	MSP	MFA
3069	Strengthening Capacity to Integrate Environment and Natural Resource Management for Global Environmental Benefits	UNDP	MSP	MFA
3163	Strengthening Capacity to Implement the Global Environmental Conventions in Namibia	UNDP	MSP	MFA
3235	CACILM Rangeland Ecosystem Management-under CACILM Partnership Framework, Phase 1	UNDP	MSP	LD
3237	Demonstrating Local Responses to Combating Land Degradation and Improving Sustainable Land Management in SW Tajikistan-under CACILM Partnership Framework, Phase 1	UNDP	MSP	LD
3310	Environmental Learning and Stakeholder Involvement as Tools for Global Environmental Benefits and Poverty Reduction	UNDP	MSP	MFA
3355	CPP Namibia: Enhancing Institutional and Human Resource Capacity Through Local Level Coordination of Integrated Rangeland Management and Support (CALLC)	UNDP	MSP	LD
3557	Catalyzing Financial Sustainability of Georgia's Protected Area System	UNDP	MSP	BD
3620	The Caspian Sea: Restoring Depleted Fisheries and Consolidation of a Permanent Regional Environmental Governance Framework	UNDP	FP	IW
3706	CBPF: Emergency Biodiversity Conservation Measures for the Recovery and Reconstruction of Wenchuan Earthquake Hit Regions in Sichuan Province	UNDP	MSP	BD

GEF_ID	Name	Agency	Type	Focal Area
886	Implementation of Strategic Action Program for the Bermejo River Binational Basin: Phase II	UNEP	FP	IW
1022	Integrated Ecosystem Management of Transboundary Areas between Niger and Nigeria Phase I: Strengthening of Legal and Institutional Frameworks for Collaboration and Pilot Demonstrations of IEM	UNEP	FP	MFA
1097	Development of a Wetland Site and Flyway Network for Conservation of the Siberian Crane and Other Migratory Waterbirds in Asia	UNEP	FP	BD
1281	Solar and Wind Energy Resource Assessment	UNEP	FP	CC
1353	Nature Conservation and Flood Control in the Yangtze River Basin	UNEP	FP	MFA
1776	Strengthening the Network of Training Centers for Protected Area Management through Demonstration of a Tested Approach	UNEP	MSP	BD
2178	Promoting Sustainable Transport in Latin America (NESTLAC)	UNEP	MSP	CC
2538	Assessment of Risk Management Instruments for Financing Renewable Energy	UNEP	MSP	CC
2796	Building the Partnership to Track Progress at the Global Level in Achieving the 2010 Biodiversity Target (Phase I)	UNEP	FP	BD
3037	Conservation and Use of Crop Genetic Diversity to Control Pests and Diseases in Support of Sustainable Agriculture (Phase 1)	UNEP	FP	BD
3309	Participatory Planning and Implementation in the Management of Shantou Intertidal Wetland	UNEP	MSP	IW
3811	International Commission on Land Use Change and Ecosystems	UNEP	MSP	BD
1254	Integrating Watershed and Coastal Area Management (IWCAM) in the Small Island Developing States of the Caribbean	UNEP/UNDP	FP	IW
1520	Development of a National Implementation Plan in India as a First Step to Implement the Stockholm Convention on Persistent Organic Pollutants (POPs).	UNIDO	FP	POPs
2715	Disposal of PCB Wastes in Romania	UNIDO	MSP	POPs
3011	Introduction of BAT and BEP methodology to demonstrate reduction or elimination of unintentionally produced POPs releases from the industry in Vietnam	UNIDO	MSP	POPs
87	Protected Areas Management Project	World Bank	FP	BD

GEF_ID	Name	Agency	Type	Focal Area
1081	Lima Urban Transport	World Bank	FP	CC
1221	Coastal and Biodiversity Management Project	World Bank	FP	BD
1531	Coral Reef Targeted Research and Capacity Building for Management	World Bank	FP	IW
2654	Consolidation of the Protected Area System (SINAP II) - Third Tranche	World Bank	FP	BD
1092	Integrated Ecosystem Management in Indigenous Communities	World Bank/IADB	FP	BD
112	Photovoltaic Market Transformation Initiative (IFC)	World Bank / IFC	FP	CC
1093	Reversing Land and Water Degradation Trends in the Niger River Basin	World Bank / UNDP	FP	IW

Annex 2. Terminal Evaluation Report Review Guidelines and Rating Scales

The assessments in the terminal evaluation reviews will be based largely on the information presented in the terminal evaluation report. If insufficient information is presented in a terminal evaluation report to assess a specific issue such as, for example, quality of the project's monitoring and evaluation system or a specific aspect of sustainability, then the preparer of the terminal evaluation reviews will briefly indicate so in that section and elaborate more if appropriate in the section of the review that addresses quality of report. If the review's preparer possesses other first-hand information such as, for example, from a field visit to the project, and this information is relevant to the terminal evaluation reviews, then it should be included in the reviews only under the heading "Additional independent information available to the reviewer." The preparer of the terminal evaluation review will take into account all the independent relevant information when verifying ratings.

1. Criteria for Outcome Ratings

Based on the information provided in the terminal evaluation report, the terminal evaluation review will make an assessment of the extent to which the project's major relevant objectives were achieved or are expected to be achieved, relevance of the project results, and the project's cost-effectiveness³. The ratings on the outcomes of the project will be based on performance on the following criteria⁴:

- a. Relevance. Were project outcomes consistent with the focal area/operational program strategies and country priorities? Explain.
- b. Effectiveness. Are project outcomes commensurate with the expected outcomes (as described in the project document) and the problems the project was intended to address (the original or modified project objectives)?
- c. Efficiency. Include an assessment of outcomes and impacts in relation to inputs, costs, and implementation times based on the following questions: Was the project cost-effective? How does the project's cost-time versus outcomes equation compare to that of similar projects? Was the project implementation delayed due to any bureaucratic, administrative, or political problems and did that affect cost-effectiveness?

An overall outcome rating will be provided according to the achievement and shortcomings in the three criteria ranging from highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory,

³ *Objectives* are the intended physical, financial, institutional, social, environmental, or other development results to which a project or program is expected to contribute (OECD DAC 2002).

⁴ *Outcomes* are the likely or achieved short-term and medium-term effects of an intervention's outputs. Outputs are the products, capital goods, and services that result from a development intervention; these may also include changes resulting from the intervention that are relevant to the achievement of outcomes (OECD DAC 2002). For the GEF, environmental outcomes are the main focus.

unsatisfactory, highly unsatisfactory, and unable to assess.

The reviewer of the terminal evaluation will provide a rating under each of the three criteria (relevance, effectiveness, and efficiency). Relevance of outcomes will be rated on a binary scale: a 'satisfactory' or an 'unsatisfactory' rating will be provided. If an 'unsatisfactory' rating has been provided on this criterion, the overall outcome achievement rating may not be higher than "unsatisfactory". Effectiveness and Efficiency will be rated as following:

- Highly satisfactory. The project had no shortcomings.
- Satisfactory. The project had minor shortcomings.
- Moderately satisfactory. The project had moderate shortcomings.
- Moderately unsatisfactory. The project had significant shortcomings.
- Unsatisfactory. The project had major shortcomings.
- Highly unsatisfactory. The project had severe shortcomings.
- Unable to assess. The reviewer was unable to assess outcomes on this dimension.

The calculation of the overall outcomes score of projects will consider all three criteria, of which relevance criterion will be applied first - the overall outcome achievement rating may not be higher than "unsatisfactory". The second constraint that is applied is that the overall outcome achievement rating may not be higher than the "effectiveness" rating. The third constraint that is applied is that the overall rating may not be higher than the average score of effectiveness and efficiency criteria calculated using the following formula:

$$\text{Outcomes} = (b + c) \div 2$$

In case the average score is lower than the score obtained after application of the first two constraints, then the average score will be the overall score. The score will then be converted into an overall rating with mid values being rounded upwards.

2. Criteria for Risks to Sustainability Ratings

Sustainability will be understood as the likelihood of continuation of project benefits after completion of project implementation (GEF 2000). To assess sustainability, the terminal evaluation reviewer will identify and assess the key risks that could undermine continuation of benefits at the time of the evaluation. Some of these risks might be related to inadequate financial resources, legal and regulatory barriers, lack of commitment from key stakeholders, political instability, and economic context.

The reviewer will provide a rating as follows:

- Likely. There are no risks to sustainability of outcomes.
- Moderately likely. There are moderate risks to sustainability of outcomes.
- Moderately unlikely. There are significant risks to sustainability of outcomes.
- Unlikely. There are severe risks to sustainability of outcomes.
- Unable to assess. Unable to assess risks on this dimension.
- Not applicable. Risks on this dimension are not applicable to the project.

A number rating 1–4 will be provided in each category according to the achievement and shortcomings with likely = 4, moderately likely = 3, moderately unlikely = 2, unlikely = 1, and not applicable = NA. A rating of unable to assess will be used if the reviewer is unable to assess any aspect of sustainability. In such instances, it may not be possible to assess the overall sustainability. The rating would be based on an overall assessment of whether the overall risk would jeopardize continuation of benefits.

3. Criteria for Assessment of Quality of Project M&E Systems

GEF projects are required to develop M&E plans by the time of work program inclusion, to appropriately budget M&E plans, and to fully carry out the M&E plan during implementation. Project managers are also expected to use the information generated by the M&E system during project implementation to improve and adapt the project to changing situations. Given the long-term nature of many GEF projects, projects are also encouraged to include long-term monitoring plans that measure results (such as environmental results) after project completion. Terminal evaluation reviews will include an assessment of the achievement and shortcomings of M&E systems.

M&E design. Project should have a sound M&E plan to monitor results and track progress in achieving project objectives. An M&E plan should include a baseline (including data, methodology, and so on), appropriate indicators and data analysis systems, and evaluation studies at specific times to assess results. The time frame for various M&E activities and standards for outputs should have been specified. The questions to guide this assessment include: In retrospect, was the M&E plan at entry practicable and sufficient (sufficient and practical indicators identified; timely baseline; targets created; effective use of data collection; analysis systems including studies and reports; practical organization and logistics in terms of what, who, and when for M&E activities)?

M&E plan implementation. The M&E system was in place and allowed the timely tracking of results and progress toward project objectives throughout the project. Annual project reports were complete, accurate, and with well-justified ratings. The information provided by the M&E system was used to improve and adapt project performance. An M&E system should be in place with proper training for parties responsible for M&E activities to ensure that data will continue to be collected and used after project closure. The questions to

guide this assessment include: Did the project M&E system operate throughout the project? How was M&E information used during the project? Did it allow for tracking of progress toward project objectives? Did the project provide proper training for parties responsible for M&E activities to ensure data will continue to be collected and used after project closure?

A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = UA. The reviewer of the terminal evaluation will provide a rating under each of the criterion (M&E design, and M&E plan implementation) as follows:

- Highly satisfactory. There were no shortcomings in that criterion of the project M&E system.
- Satisfactory. There were minor shortcomings in that criterion of the project M&E system.
- Moderately satisfactory. There were moderate shortcomings in that criterion of the project M&E system.
- Moderately unsatisfactory. There were significant shortcomings in that criterion of the project M&E system.
- Unsatisfactory. There were major shortcomings in that criterion of the project M&E system.
- Highly unsatisfactory. There was no project M&E system.

The rating for M&E during implementation will be the overall rating of the M&E system:

Rating on the Quality of the Project Monitoring and Evaluation System = b

4 Criteria for Assessment of Quality of Terminal Evaluation Reports

The ratings on quality of terminal evaluation reports will be assessed using the following criteria:

- The report presents an assessment of all relevant outcomes and achievement of project objectives in the context of the focal area program indicators if applicable.
- The report was consistent, the evidence presented was complete and convincing, and ratings were well substantiated.
- The report presented a sound assessment of sustainability of outcomes.
- The lessons and recommendations are supported by the evidence presented and are relevant to the portfolio and future projects.
- The report included the actual project costs (totals, per activity and per source) and actual cofinancing used.
- The report included an assessment of the quality of the M&E plan at entry, the M&E system used during implementation, and whether the information generated by the M&E system was used for

project management.

A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = UA. Each criterion to assess the quality of the terminal evaluation will be rated as follows:

- Highly satisfactory. There were no shortcomings in the terminal evaluation on this criterion.
- Satisfactory. There were minor shortcomings in the terminal evaluation on this criterion.
- Moderately satisfactory. There were moderate shortcomings in the terminal evaluation on this criterion.
- Moderately unsatisfactory. There were significant shortcomings in the terminal evaluation on this criterion.
- Unsatisfactory. There were major shortcomings in the terminal evaluation on this criterion.
- Highly unsatisfactory. There were severe shortcomings in the terminal evaluation on this criterion.

The first two criteria (of all relevant outcomes and achievement of project objectives and report consistency and substantiation of claims with proper evidence) are more important and have therefore been assigned a greater weight. The quality of the terminal evaluation reports will be calculated by the following formula:

Quality of the Terminal Evaluation Report = $0.3 \times (a + b) + 0.1 \times (c + d + e + f)$

The total number will be rounded and converted to the scale of highly satisfactory to highly unsatisfactory.

Annex 3. Quality of Supervision Assessment

The GEF Evaluation Office would follow the same approach for the quality of supervision 2012 review as followed for the pilot assessment in 2006 and follow up assessment in 2009. Supervision is still understood as *the identification and tracking of, and response to, risks and other issues affecting project implementation and achievement of project objectives*. The Office would examine the Agency supervision systems for the agencies covered through the review and a representative sample of projects. The quality of project supervision during implementation, as was the case with preceding reviews, would be assessed using the following three criteria:

- Focus on results
- Supervision inputs and processes
- Candor and quality of project performance reporting⁵

Each of these criteria would have sub-criteria on which performance would be assessed. The performance on the sub-criteria would form a basis for the overall rating on a criterion. Similarly, performance on the three criteria would form a basis for the overall quality of supervision rating. The instrument that would be used would be the same as that was used for the 2009 review.

A representative sample would be drawn from the projects that were under implementation during FY2011 and 2012, and were being implemented by agencies other than the World Bank, UNDP and UNEP. In all about 30 projects would be covered. Agencies that have only a few projects may be oversampled. The effect of oversampling of projects would corrected by using probability weights.

The reviews would be conducted by panels convened by the GEF Evaluation Office. Each panel would consist of two or more reviewers. For each sampled project, the project appraisal documents and PIRs would be accessed through the GEF PMIS. In addition, the respective Agencies would be asked to provide documentation on supervision activities undertaken for the sampled project. Such documentation would include audit reports, emails, management letters, field mission reports, and—where applicable—midterm review and terminal evaluation reports. The panelists would review these materials to gather information on supervision-related aspects and to identify issues for which more information is required. Conference calls would be made to the project management team to gather more information on issues identified through the desk review, to seek further clarification on such issues, and to gather new information. The information thus gathered would be recorded in an instrument (the instrument used in the 2009 review is listed later in this annex). In making modifications in the instrument, precautions would be taken to ensure comparability in ratings with preceding assessments.

⁵ The quality of supervision assessment would look into issues related to fiduciary standards in the sampled projects. However, since the review panels for the assessments would not include fiduciary specialists, these issues would not be examined in depth. However, when appropriate, performance on fiduciary responsibilities related issues would be considered in providing an overall supervision rating for a project.

Table C.1

Instrument Used for 2009 Quality of Supervision Assessment

Overall Assessment	
Comments on quality of project design at entry:	
Comments on quality of project design at the time of assessment:	
Comments on quality of project execution:	
Comments on likelihood of project achieving its environmental objectives?	
Overall Comments	
Use the following rating scale for assessments:	
Highly satisfactory (HS = 6) Satisfactory (S = 5) Moderately satisfactory; (MS = 4) Moderately unsatisfactory (MU = 3) Unsatisfactory (U = 2) Highly unsatisfactory (HU = 1) Unable to assess (UA)	
Supervision Quality Parameter (only ratings need to be included)	
Part A. Context (this section assesses project quality at entry ^b)	
	Parameter and performance explanation
(a)	Assessment of project design at entry
(b)	Assessment of quality of logframe or results framework to achieve the direct objectives, including realism of causal links between the project's inputs, activities, outputs, and intended outcomes
(c)	Quality of arrangements for M&E (including clarity and precision of the performance indicators to assess progress and outcomes and availability of baseline data)
(d)	Adequacy of measures incorporated in the project design to address policy constraints
(e)	Host government commitment to the project
(f)	Extent of integration and quality of fiduciary and safeguards aspects in project design
(g)	Quality of institutional framework for the project (consider the adequacy of the institutional assessment and the appropriateness and realism of the institutional capacity building measures and TA arrangements).
(h)	Quality of risk assessment and management

(i)	Major issues affecting project implementation
(j)	Residual risks affecting project implementation at the time of the review
(k)	Changes made in the project design prior to the assessment period
(l)	Changes made in the project design during the assessment period and the reasons why it was required
Part B. Supervision Ratings	
1.	Focus on Results
1.1	Identification and assessment of problems (<i>assess performance on this parameter based on timely identification and assessment of implementation problems and potential threats to global environmental objectives</i>)
1.2	Focus on sustainability (<i>extent to which supervision paid attention to this aspect through emphasis on host country and stakeholder ownership; and technical assistance, training, and capacity building</i>)
1.3	Actions taken and follow-up during the assessment period (<i>this assessment should take into account subparameters such as appropriateness of advice and proposed solutions; appropriateness and speed of Implementing Agency follow-up action; impact and effectiveness of Implementing Agency actions; quality and timeliness of midterm review [if any]; and quality and timeliness of restructuring plan [if any]</i>)
	a) Quality and timeliness of midterm review (only rating).
1.4	Performance Monitoring (<i>this assessment should take into account the extent to which the task team made use of the global development objective and intermediate outcome indicators to assess the project's implementation, as reported in the PIR; and extent to which the indicators [both quantitative and qualitative] have been used to identify and address potential obstacles to the achievement of the global development objective [attention to long-term objective]</i>)
2.	Fiduciary/safeguards aspects (<i>this assessment will take into account quality of oversight of procurement; financial management; governance, anticorruption, and legal aspects; environmental aspects; social aspects</i>)
	(<i>For tracking of compliance with the GEF Council decision on incorporation of social science expertise in supervision of GEF projects when appropriate; responses for these questions need to be recorded. They will not be rated.</i>)
	a. Given the nature of the project was there a need to include a social scientist expert for providing supervision inputs?
	b. Did a social scientist expert provide supervision inputs—in what form?
	c. Did s/he participate in the supervision missions—if yes in how many out of ...?
3.	Adequacy of supervision inputs and processes

3.1	Staffing <i>(for this assessment, take into account staff continuity and quality of supervision skill mix)</i>
3.2	Supervision activities <i>(for this assessment, take into account the quality of mission preparation and effectiveness of time spent in the field; and if there was joint supervision with cofinancers, quality of coordination and other supervision activities)</i>
3.3	Quality of supervision documentation
3.4	Effectiveness of relationships with government, donors, and other stakeholders
3.5	Management inputs <i>(this assessment will take into account adequacy and speed of management attention and actions; adequacy of supervision budget; and effectiveness of budget use)</i>
	a) Adequacy and speed of management attention and actions (only rating)
	b) Adequacy of supervision budget (only rating)
Part C. Reporting	
4.	Candor and quality of project performance reporting
4.1	Extent to which ratings reflect actual conditions
4.2	Adequate explanation of ratings, and of any change in ratings
4.3	Accuracy of ratings of project components and risk assessment
4.5	Quality and timeliness of data (including intermediate outcome indicators) to support the key performance indicators

a. Quality at entry ratings are not aggregated with the Part B ratings.