GEF Evaluation Office

GEF IMPACT EVALUATION

Approach Paper

Impact Evaluation Information Document No. 1

Prepared by David Todd and Josh Brann
GEF Evaluation Office

September 2007

This paper was commissioned by the GEF Evaluation Office (GEF EO) as an input into its program of **Impact Evaluation**.

A first annual report on this program will be presented to the GEF Council at its November 2007 meeting. The findings, interpretations, and conclusions expressed herein are those of the authors and do not necessarily represent the views of GEF Evaluation Office, the GEF Council, or the Governments they represent. The authors of this document would welcome any comments or suggestions on its contents.

The papers in the Impact Evaluation information document series, as of September 2007, are:

1. Approach Paper to GEF Impact Evaluation – *Brann and Todd*
2. Final Report on proposed Approach to GEF Impact Evaluation - *Foundations of Success*
3. GEF Biodiversity Policy Review - *Foundations of Success*
4. Methodological Challenges in Impact Evaluation: The Case of the Global Environment Facility – Todd and Vaessen
5. Priorities and indicators for Global Environment Benefits from Biodiversity: The current international architecture – *Nair*
6. Case Study Methodology – *Conservation Development Centre*
7. Case Study: Bwindi Impenetrable National Park and Mgahinga Gorilla National Park Conservation Project - *Conservation Development Centre*
8. Case Study: Lewa Wildlife Conservancy – *Conservation Development Centre*
9. Case Study: Reducing Biodiversity Loss at Cross-Border Sites in East Africa *Conservation Development Centre*
10. Impacts of Creation and Implementation of National Parks and of Support to Batwa on their Livelihoods, Well-Being and Use of Forest Products – *Namara*
11. Protected Areas and Avoided Deforestation: A Statistical Evaluation – *Andam, Ferraro, Pfaff and Sanchez-Azofeifa*

Global Environment Facility
Director of the GEF Evaluation Office: Robert D. van den Berg

Impact Evaluation Team
Task Manager: David Todd, Senior Evaluation Officer
Evaluation Analyst: Divya Nair, Junior Evaluation Professional
Co-reader: Lee A. Risby, Evaluation Officer

Contents:

GEF IMPACT EVALUATIONS

INITIATION AND PILOT PHASE – FY06

1: BACKGROUND AND RATIONALE

The GEF portfolio has now matured to a stage at which consideration can be given to the long-term impacts of its projects. This possibility conforms well to the increasing emphasis of the GEF Council on results, exemplified by the development of a Results-based Allocation Framework and the establishment of an Independent Evaluation Office.

GEF Council document GEF/ME/C.25/3 of May 6, 2005, "Four Year Work Program and Budget of the Office of Monitoring and Evaluation – FY06-09 and Results in FY05", states that the "Council has requested on several occasions that GEF Office of Monitoring and Evaluation conduct Country Portfolios Reviews and Impacts Evaluations, in particular post project completion evaluations. These evaluation modalities are essential elements for an independent Office, as presented in its TORs. GEF Office of Monitoring and Evaluation proposes to develop a methodology for both types of reviews during FY06, as well as implementing at least one pilot for each. A full implementation of these two modalities will then be undertaken in the subsequent three years, Furthermore, both types of evaluations will be carried out as much as possible and appropriate in collaboration with evaluation departments of GEF partners and taking into full consideration their own impact assessments and country evaluations." (Para.20, P5).

2: OBJECTIVES

The same document (GEF/ME/C.25/3) specifies that: "In the case of Impact Evaluations, an approach paper will be developed in FY06 and at least one pilot evaluation completed. The objective of this modality will be to evaluate the long-term results of GEF interventions, a few years after GEF support is concluded and to assess the sustainability and replication of the support as well as to extract lessons learned. An interesting theme to explore will be the GEF long term contribution to the countries' institutional capacities (i.e., policies, organizations, knowledge, etc.) and the countries' willingness to address global environmental concerns" (P6). Furthermore, the impact evaluation will seek to place impacts at the project level within the global context of the GEF's overall goal of financing the incremental costs of global environmental benefits.

According to the Council document, "there are several approaches that could be selected: (1) a cluster of projects with similar objectives, geographic area, focal area, or even a combination of these could be evaluated together, (2) a particular completed project in a country chosen for country portfolio review; (3) use of a location-specific approach which looks at trends and developments in a certain area (for example, allocations to "hot spots" or a particular water basin) and establish the role of GEF interventions in that area." (Para. 23, P6).

Keeping in mind the fact that the GEF Evaluation Office will only be able to undertake a limited number of impact evaluations, it will increase the value of the studies by focusing on projects that have used approaches with common goals or objectives, or which deal with issues common or important within the full portfolio. In addition, it will be useful to focus on evaluating the impact of projects, which are expected to provide insights for other studies being undertaken or to be undertaken by the Office. Examples could include projects that involve the Protected Area Management Effectiveness Tracking Tool, support to the private sector, indigenous peoples, the GEF's catalytic role and capacity building. Linkages to the country portfolio evaluations could also be valuable.

3: CHALLENGES

The study of impacts of the GEF portfolio will face substantial challenges, arising from the low starting point of such activities. The OPS3 study team, which was required to report on impacts, described the situation facing it as follows: "there would be problems relating to reporting at the level of long-term quantifiable results or impacts (global environmental benefits). …… such as the following:

- Most projects do not generate information at the level of long-term quantifiable impacts and, more important, many projects still do not have clear and agreed baselines, indicators of impacts, or methodologies to calculate them.
- Environmental change may take decades to be perceived or measured, while GEF projects on average span four or five years.
- The GEF does not systematically conduct post-completion studies to look at long-term results.
- The GEF, as an institution, does not have an overall results measurement framework or methodology to aggregate from project-level impacts to program-level or GEF-wide impacts. There is no unified framework in place for systematically defining, measuring, and aggregating results of GEF activities, particularly in terms of global environmental benefits for each of the GEF focal areas.

OPS3 observed that while mechanisms appeared to be in place to guide development of goals and results during project design, implementation, and reporting… and individual projects have been assessed against their implementation performance as part of various…reporting, there remains a large gap in the effectiveness of such project-level mechanisms in capturing results at the impact level. Apart from this constraint, there were no mechanisms in place to support the roll-up of impacts should they be identified,"
(OPS3, P22).

Clearly, the GEF Evaluation Office impact evaluations will need to develop a strategy, which can produce the most useful results, with a limited budget and a low starting point. Key elements of this strategy will include; maximizing use of existing evaluation data, collaboration with the evaluation offices of IAs and adopting a sound theoretical basis for deriving an understanding of environmental impacts and the possibilities to measure these.

4: EVALUATION MODELS

What type of impact evaluation is appropriate for the GEF EO studies?

The OECD's Development Assistance Committee (DAC) defines impacts as "Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended."[1] The Operations Evaluation Department of the World Bank defines impact evaluation as the "systematic identification" of these effects.[2] The majority of impact evaluation work within the World Bank has focused on the poverty impacts of activities such as social funds, conditional cash transfer programs and health, water, urban and transport sector initiatives. Environmental activities have not featured to any major extent in the impact evaluation work of OED and this area presents an opportunity for the GEF EO to develop a "product" in relatively uncharted territory, namely the evaluation of the impacts of environmentally-targeted interventions. Some pioneering efforts have also been made by the GEF Office of the World Bank in this direction (see Section 5 and Annex A).

Impact evaluations face a set of conceptual and practical difficulties. These revolve around the necessity to be able to demonstrate that changes, which are observed to have occurred, can be attributed in part or in full to the intervention being evaluated. This issue has been addressed through radically different evaluation approaches. For many years, evaluators relied on the concept of a "counterfactual" as the baseline against which project impacts could be measured. As the conceptual and practical difficulties of this approach gained acceptance, there was a trend towards a streamlined version, which attempted to measure impacts as best possible through less intensive and expensive means. At the same time, alternative approaches developed which questioned the utility of a hypothetically-determined baseline and placed a theoretical and empirical assessment of the cause and effect chain of an intervention as the most accurate means of attributing change.

4a: The Counterfactual-Based Approach and Attribution

According to this approach, from a conceptual point of view, attribution requires an explicit counterfactual, which establishes what would have happened if the intervention had not been made. The impact of a project can then be assessed as the difference between the observed outcome and the counterfactual. Since the counterfactual did not actually occur, it is at best a hypothetical construction and it is therefore never possible to say precisely what difference an intervention has made. Discussions in evaluation circles have focused on what degree of precision is acceptable or reasonable in making such hypotheses and to what extent changes can be attributed to projects or programs evaluated. Evaluation practitioners adopting positions based on economics or statistics have maintained that it should be possible to specify statistically (with a reasonable degree of confidence) the difference between the recorded results and the counterfactual.

[1] DAC, 2002. "Glossary of Key Terms in Evaluation and Results Based Management."

[2] www.worldbank.org/oed/ei

In development cooperation evaluation several methods have been used to attempt to deliver such results, including:

- Randomization or Experimental Design
- Propensity Score Matching
- Pipeline Comparison
- Simulated Counterfactual
- Difference in Means (Single Difference)
- Difference-in-Difference or Double Difference
- Instrumental Variables.

These statistically-based approaches have been categorized by their advocates as "rigorous impact evaluation methodologies."[3] They are said to have the advantage of providing a reasonable means of establishing causality and for estimating the scale of impact attributable to the intervention. However, they are also costly ($200,000 to $900,000 per project) and time consuming, taking up to two (or sometimes more) years to complete. The World Bank's OED has conducted 23 "rigorous" impact evaluations since 1980, at an average cost of around $500,000. In view of its budget limitation of around $23 million per annum, OED has only been able to conduct about one such evaluation per annum. The modest budget of the GEF Evaluation Office restricts the possibilities of pursuing such an approach, even if it were considered appropriate.

Regarding practical difficulties for impact evaluation, data describing the initial situation, or baseline, where an intervention is targeted, are necessary to be able to identify change over time. This is especially true when attempting to compare the impacts of the intervention under review to the counter-factual. Impact evaluations are often faced with the difficulty that such baselines are not available in a complete and comprehensive data set. This is a common challenge in the environmental field, where appropriate indicators in some fields, notably biodiversity conservation, may not yet have been identified or broadly applied. Even where clear and appropriate indicators have been identified, projects often fail to carry out the data collection necessary to establish a complete baseline, which includes both environmental and socio-economic data sets.

4b: Contribution, Narrative/Historical Evaluation, and Shoestring Approaches

Another school of thought in development cooperation evaluation suggests that it is more appropriate to use impact evaluations to identify an intervention's *contribution* to observed changes, rather than to attempt to establish *causality*.[4] Because of the conceptual and practical difficulties discussed above, the real world is proving far too complex and dynamic to be able to clearly identify causal linkages between an intervention and subsequent changes. Identifying such linkages can pose particular difficulties in relation to GEF projects and programs; because

[3] See, for example, "OED and Impact Evaluation – A Discussion Note"

[4] See: van den Berg, R. D., 2005. "Results Evaluation and Impact Assessment in Development Cooperation," *Evaluation.* Vol. 11(1): 27-36.

they target long term and global changes, to which any individual project will make a limited, but definable contribution. In most cases the GEF supported intervention may be a necessary, but not sufficient, condition for the desired changes to occur.

In addition to the "rigorous" methods of evaluation discussed previously, another often-employed methodology is the narrative/historical approach to evaluation. This involves a full description of the external intervention or interventions, coupled with an assessment of changes in development or environmental status which appear to have been associated with it. Many different techniques such as stakeholder surveys, interviews, desk studies, field visits and economic research, have been used by narrative/historical evaluations to show the extent to which the changes observed can be attributed to the interventions under evaluation.

The narrative/historical approach has been commonly used in evaluations to assess the effectiveness of aid development programs over long periods of time. One example is the "Evaluation of the Netherlands Development Programme with Bangladesh, 1972-1996"[5]. This evaluation reviews activities within a single country over a period of nearly 25 years. In view of the broad scope adopted, narrative/historical evaluations usually require more time and money than will be available for the impact evaluations. However, many of the individual methods of such studies may be included in the approach adopted.

Another approach to overcome the limitations imposed by the cost, time and data requirements of "rigorous" and narrative/historical impact evaluations has been the development of a variety of "shoestring methodologies"[6] to best approximate the results sought from the more elaborate approaches. "They are based on a five-step[7] approach that includes the following:

1. Definition of the evaluation design and prioritization of the clients' information needs (which also implies participatory identification of what is possible or is less likely in terms of analytical results);
2. Identification of options for reducing costs and time of data collection;
3. Reconstruction of baseline conditions and reconstruction of control groups based on a preliminary survey;
4. Identification of possible threats to validity and assessment of the adequacy of the evaluation during the elaboration of the data collection strategy;
5. Limiting some of the threats to validity and adequacy in the evaluation design by recommending practical measures"[8].

[5] Netherlands Ministry of Foreign Affairs, 1998. "Bangladesh: Evaluation of the Netherlands Development Programme with Bangladesh, 1972-1996." The Hague: Ministry of Foreign Affairs, Policy and Operations Evaluation Department.

[6] See: Bamberger, M. and N. Fujita. 2003. Impact Evaluation of Development Assistance. A Practical Handbook for Designing Methodologically Sound Impact Evaluations under Budget, Time, and Data Constraints, FASID Evaluators Handbook. Tokyo: FASID; and Bamberger, M., J. Rugh, M. Church, et L. Fort. 2004. Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints. American Journal of Evaluation 25, no. 1:5-37.

[7] This is a simplified presentation as the model can be broken down into more steps that separate the different components (time, budget, data) of the evaluation design.

[8] See, "Impact Assessment: biodiversity protection project in Ecuador, Proposed Evaluation Methodology, preliminary draft," P5, Le Groupe-conseil baastel Ltee, Canada.

An examination of this approach suggests that it is not substantially different from good practice in conventional end of project evaluations. The major area of difference from such studies is the attempt to define control groups, which have been found to be less convincing in field-based studies than in scientific or quasi-scientific studies.

In the context of impact evaluations of GEF projects, it is clear that the "rigorous impact evaluation" model is neither appropriate nor affordable, so that studies will be nearer to the "shoestring" model. Furthermore, one must keep in mind that the target of GEF interventions – the global environment - necessitates additional modifications of conventional impact evaluation methodologies used in development cooperation, even of the "shoestring" variety, since these have been primarily developed to investigate social and poverty impacts, rather than change in environmental status.

Another potentially useful method of investigating impact of interventions is being piloted by GHK Consulting, who recently discussed this methodology with the GEF Evaluation Office. The method is currently known as Systematic Evaluation of Environmental and Economic Results (SEEER). The method has been designed to assess the environmental and economic effects of resource and environmental decisions. It involves the creation of an index of effects through a process of triangulation of expert and stakeholder inputs and feedback, which is validated through a facilitated meeting with a panel of scientific experts. The method has been found accurate in a number of US and Canadian environmental interventions and appears promising for use as part of the impact evaluation of GEF projects The GEF Evaluation Office will therefore further explore the potential of this method in the impact evaluation pilot phase.

4c: Theory Based Approaches

Theory based evaluation adopts a logical model of cause-effect linkages, though which an intervention intends to achieve its objectives and explores empirically the extent to which events followed the anticipated sequence. The evaluation assesses the validity of the theory of change adopted (implicitly or explicitly) by the intervention, as well as the interaction between the specific local circumstances and the general principles expected to generate the desired change. The approach enables detailed examination of the nature of linkages between a complex set of causes and a set of effects, which vary in the extent to which they are susceptible to quantifiable measurement. The theoretical model can clarify the assumptions upon which the intervention was based and also incorporate opposing theories, which may be applied to the same situation.

A theory-based evaluation begins with an analysis of the assumptions made at the project design stage[9]. These may be relatively accessible, where a project has a detailed Logical Framework as part of its funding proposal. Where there is no Logical Framework, or a poorly-defined one, the evaluation will begin by reconstructing the intended change logic on the basis of project documents, interviews with staff involved in project preparation and approval, comparison with

[9] See, for example, "Theory-Based Evaluation: The Case of Social Funds," Soniya Carvalho and Howard White, American Journal of Evaluation, Vol. 25, No.2, Pages 141-160, 2003.

similar interventions elsewhere and broader literature review. During this process, the identification of any counter-theories is an important step, which considers the extent to which interventions can be expected to generate responses, which run against those anticipated in the change theory on which the intervention was based.

Once the underlying assumptions and counter-assumptions have been fully explored, the evaluation proceeds to examine progress reports, project evaluations and other documentation, to develop an understanding of what happened during the implementation process. This material generates a set of key questions, which can be examined through two complementary routes: firstly, comparison with projects with similar theoretical models and secondly, field work to enable discussion with people engaged with and affected by the intervention, as well as to collate evidence of changes (for the GEF, mainly in environmental status), which may have been gathered by project staff or other actors on the spot.

Primary and secondary data concerning the project achievements are then analyzed to see to what extent the anticipated chain of events led to the attainment of project objectives and to examine those local factors which inter-acted with project design and implementation to produce specific local outcomes and impacts. This analysis is placed in a broader framework of evaluation of similar interventions to see to what extent impacts have been of the nature and magnitude, which could be anticipated from such an intervention and what were the factors promoting, modifying or hindering achievement.

Theory-based evaluations do not necessarily have to take place years after a project has been completed. There are various moments through project implementation and completion at which the "theory of change" can be evaluated. During project implementation it can be evaluated whether or not the mechanism employed in the intervention, based on the identified assumptions, is in fact resulting in the expected outcomes, and whether the assumptions are holding.

At the completion of the project, it is then possible to evaluate whether or not the mechanism employed under the theory of change as resulted in the expected outputs and outcomes. Through the expected outputs and outcomes the anticipated path towards impact should be visible, and it may be possible at this stage to identify some initial impact resulting from the intervention. Following the end of the intervention, the full impact would be expected to be visible within five to ten years.

With regard to impact evaluations of GEF activities, a theory based approach offers the opportunity to explore the achievements of the portfolio, on the basis of empirically evaluated projects, set in the broader context of theories of environmental change. The approach is well adapted to the particular circumstances of the GEF, which has a relatively small amount of funds, targeting a high (global) level of change, since it can show the direct impacts, which can be attributed to an intervention, as well as any broader contribution towards the strengthened understanding and development of theories of change in the global environment.

Theory-based evaluation has been refined by advocates of "realistic evaluation." This modification of the approach also seeks to determine attribution of change on the basis of the theory underlying an intervention. However, it also emphasizes the critical nature of the specific

context in which an intervention is implemented. Its analytical approach uses a model based on Context, Mechanism and Outcomes (CMOs). For outcomes to be achieved, the contextual conditions must be appropriate to allow the intended mechanism to work. Any particular intervention is likely to use a variety of mechanisms, each of which will respond to its context to produce anticipated or unanticipated outcomes. "Through measurement of a series of CMOs it should be possible to deduce the features of contexts that allow different mechanisms to work to achieve particular outcomes. Thus 'transferable lessons' may be learned"[10]. The approach, while a valuable tool, needs to be used circumspectly since an "infinite number of CMO configurations can be drawn for the simplest of programmes.[11]" It will be valuable to the GEF impact evaluation program as a means of testing patterns, which emerge from documentary analysis, to explore how particular contexts interact with broad based mechanisms to contribute to or detract from their intended impacts. For example, what lessons can be drawn from varying experiences of eco-tourism as a generator of community income to design future interventions with a higher probability of success?

The emphasis on the importance of the interaction of a specific context with a general mechanism provides a strong challenge to the theoretical basis of "conventional" counterfactual-based approaches. Realistic evaluation proposes that each interaction is unique, so that the construction of a theoretical, "context-free", counterfactual provides only an illusion of comparability, since the effectiveness of any intervention depends, amongst other factors, on the degree of fit between the mechanism and the context in which it is delivered. "Whereas the question which was asked in traditional experimentation was, "Does this work?" or "What works?' the question asked by us in realistic evaluation is "What works for whom in what circumstances?" Thus, we begin by expecting measures to vary in their impact depending on the conditions in which they are introduced. The key problem for evaluation research is to find out how and under what conditions a given measure will produce its impacts.[12]" The concern of realistic evaluation is then to understand under what conditions underlying causal mechanisms become activated to produce specific outcomes.

[10] "The Development of Realistic Evaluation Theory through the Evaluation of National Crime Prevention Programmes", P2, Martin Gill and Angela Spriggs, University of Leicester.

[11] Gill and Spriggs, P3.

[12] "Realistic Evaluation: An Overview", P4, Nick Tilley, Nottingham Trent University, 2000.

5: THE WORLD BANK GEF OFFICE IMPACT EVALUATIONS

The World Bank's GEF Office has already commissioned several impact evaluation studies of GEF projects[13]. A cluster of four energy efficiency projects was evaluated, as well as one biodiversity project in Ecuador. Although commissioned by the World Bank GEF Office, these studies were conducted by independent consultants and do not show any major biases. The quality and approach of the studies is sound and they therefore provide valuable insights for the proposed EO impact evaluations. Some key lessons to be incorporated into planning the EO studies are:

- Impact evaluations are likely to begin with projects, which were developed early in the GEF experience. Many such projects had no logical framework and poor monitoring information, which makes impact evaluation extremely difficult. Counterfactuals developed by the Bank studies to overcome this difficulty varied considerably in the extent to which they were grounded by any factual data or were "scenario construction", making comparison between project achievements and counterfactuals highly speculative.
- Studies encountered reluctance to participate on behalf of Bank staff and national counterparts, who did not appreciate the value of the study and were reluctant to devote time to a project of no current interest.
- Consideration needs to be given to what level of conclusions and lessons are expected from impact evaluations. Many of the lessons from the Bank-GEF studies do not differ substantially from "good practice" examples generated by terminal evaluations. Is it appropriate to use impact evaluations to generate such lessons, or should this level of study target a higher level of lesson? What level would that be?
- Lessons from a set of related impact evaluations are more compelling than those from an isolated study of an individual project. The latter cannot be placed in a suitable framework to triangulate findings and separate those that are unique to the project from those which have broader significance. The set of Energy Efficiency studies provide interesting lessons, whereas it is difficult to interpret the one biodiversity study completed. It would be better for the EO study to avoid "one off" evaluations, except of very large or otherwise important interventions.
- How will impact evaluations relate to changes in the global environment, which are the long term objective of GEF activities? It will rarely be possible to show that one project contributed to this level of change. It might be possible to see such changes at the level of a set of related activities in one country or region. Major IW projects might be expected to make a discernible contribution at this level, but it seems that progress is slow in moving from the preparatory TDA stage to the implementation of a SAP, at which point physical impacts might be generated. It may be still too early to assess impacts from IW activities.
- Could impact evaluations be used to focus on specific macro-level changes intended to be monitored by global indicators (e.g., protected areas brought under effective management) to give substance to these indicators?

[13] These are discussed in Annex A

- Could impact evaluations be brought to bear on the assumptions of the RAF in any way? The RAF is partly based on the performance of country level delivery of global environment gains. Can these be substantiated by impact evaluations of major projects or sets of activities? This could be tied in to country portfolio reviews in future.
- Could impact evaluations be used to assess the long-term results of "higher-level" GEF outcomes, such as changes in policies or capacity at institutional or national level?

6: PROPOSED MODEL FOR IMPACT EVALUATION

In the light of the above discussion, it is proposed to produce an approach to impact evaluation, tailored to the GEF context, which builds on theory-based approaches and takes into account the experience of the World Bank GEF Office impact studies, the possible use of the SEEER method (discussed on Page 8) and the range of "shoestring" adaptations available to evaluate complex interventions on a limited budget. A major challenge for the impact evaluations will be to relate impacts identified at project level to the intentions of the GEF to promote global environmental change. The GEF portfolio is far too large for the Evaluation Office to begin to evaluate the impact of a statistically representative sample of projects. However, the adoption of a theory-based approach will enable empirical results to be used to assess the robustness of the theories of change underlying the key strands of the portfolio. Furthermore, the adoption of a "realistic" perspective will enable the studies to focus on the extent to which the theories have promoted effective mechanisms, which in turn have been appropriately adapted to specific project contexts to generate the desired impacts, whilst avoiding or minimizing any unexpected and undesirable impacts.

In parallel with the adoption of a theory-based approach, the study team will develop a systematic method of tracing linkages between project impacts and global environmental status. The elements of this method are summarized below and discussed in detail in Annex B.

The first step toward understanding the GEF's contribution is to know the current global environmental status within the GEF focal areas. At a basic level, for example, this means knowing the amount and rate of change of Green House Gas (GHG) emissions, or the rate of global biodiversity loss. However, these important global trends cannot be captured in a single number and there are many challenges to producing meaningful summary indicators. Because of the complexity of global environmental trends, the large geographic (and atmospheric) scale involved, and the limited amount of data available in many areas, the ability to report accurately and comprehensively on the state of the global environment is extremely limited.

There have been some important endeavors in this direction. For example, the UNEP produced Global Environment Outlook (2002), draws together data on a wide range of environmental issues. In addition, the continuing expansion of information technology has allowed widely accessible global databases on specific environmental issues to be developed. Examples in the biodiversity focal area include the World Bird database, the World Database on Protected Areas, and Reefbase. Other global level assessments of environmental resources are also being conducted, such as the Millennium Ecosystem Assessment, the Global Forest Resources Assessment, and the Global Amphibian Assessment.

The intention of GEF impact evaluations will be to draw linkages between field level results and these global level information resources whenever possible and appropriate. This will be done through the identification of indicators, based on the project's Global Environmental Objectives, as defined by the project document, that are applicable at the project level, but also relevant within the global context. These objectives will be based on the theory of change underlying specific portfolios and projects, which will form the theoretical base of the impact evaluation. It will therefore be possible to incorporate available environmental status indicators into the project based evaluations. In the pilot study, using the indicators chosen for the impact evaluation, the project's results will be analyzed using the Pressure-State-Response (PSR) framework[14] to determine the impact at the project level, which will then be extended to the global level based on relevant macro level data. The PSR framework provides a relatively simple basis for the development of change theories, which can form the basis of a theory-based approach. The efficacy of combining this approach with that of PSR frameworks and relating the impacts observed through these methods to global environment databases will be assessed by the pilot study and fed into the design of the main phase of impact evaluations.

The biodiversity focal area has a variety of resources of global level data. Some of these databases and assessments, as well as those from other focal areas, are further described in Annex B. Other focal areas will require fine-tuned approaches based on the objectives and strategies applied in these focal areas. Some initial discussion on each of the other focal areas has been included in Annex B.

Since the biodiversity focal area has a broader range of developed datasets, at this pilot stage it will be easier to apply the model in this focal area than in other GEF focal areas. For example, forest cover within a particular ecosystem type may be identified as an appropriate indicator for a GEF project. It is not known whether forest cover is a good of an indicator of habitat quality as it is of habitat quantity, but it is also important to recognize that it may be an imperfect indicator. At the same time, for various indicators it is necessary to make use of the data available. An impact evaluation would use the PSR framework to review the project's contribution toward maintaining or increasing forest cover within the project area. These results could then be analyzed in the context of the Global Forest Resources Assessment to understand the relationship between project level results and global environmental status of this particular resource.

Within the biodiversity focal area, impact evaluations can also build on the operational model of the Resource Allocation Framework (RAF). The RAF develops its Global Benefits Indicator for biodiversity (GBI_{Bio}) by drawing on a range of global databases developed for various components of biodiversity. Although the Convention on Biodiversity (CBD) defines biodiversity as having ecosystem, species, and genetic components, for practical purposes the RAF considers genetic diversity to be represented through diversity at the ecosystem and species levels. The Millennium Ecosystem Assessment posits that global genetic diversity is decreasing (see references in Annex B), but actually measuring the level of and trends in global genetic diversity has so far proven beyond the capacity of the scientific community at a time when it is estimated that only 10% of the planet's species have been identified.

[14] See Annex B.

7: KEY QUESTIONS

Even within focal areas there is an extremely wide range of project approaches to various environmental challenges. Thus, the impact evaluation approach will have to be flexible in its application of the proposed model and the key questions identified below. The interaction between project level experiences and the theories of change on which they are built will provide an important mechanism to progressively refine the approach to the assessment of impacts. Individual project impact evaluations will seek to answer the following questions:

- What was the intended series of cause and effect linkages (mechanisms), which were expected to generate impacts?
- What were the key features of the project context, which interacted with these linkages to determine results achieved (outputs, outcomes, and impacts)
- How did the project respond to its specific context to generate results?
- How do impacts at the project level relate to global environmental status and the overall objectives of the GEF?
- Are additional results from the project anticipated in the future, and if so, to what extent?
- Are project results likely to be sustained?
- What does the project tell us about the underlying theory of change on which the intervention was based?
- What does the project tell us about the interaction between the intended change mechanisms and project context, which could form the basis of lessons for future interventions?
- What lessons does the evaluation present concerning possible improvements to impact evaluation methodology?

8: KEY AUDIENCE AND PARTNERS INVOLVED

The initial audience targeted is the GEF Council, which has a specific interest in the issue of impacts. However, it appears likely that there are differing expectations and motivations among Council members in pressing for impact evaluations. The pilot study (together with a summary of the WB GEF impact evaluations) should therefore be used to draw out the key perspectives of Council members, to allow the main phase of the studies to deliver products, which are likely to satisfy the major expressed needs of Council.

GEFSEC and the IAs also have a major interest in establishing the impacts of the GEF's work. The initiative of the World Bank GEF Office in this respect provides a valuable body of work, which will be drawn upon by the EO study. Discussions will be held during the pilot phase to establish the extent of interest of the IA's Evaluation Offices in collaborating with the GEF EO pilot and main studies.

The broader environmental community, including the major international NGOs, is also likely to have a strong interest in the work to be undertaken and one possibility would be to host or co-host, as an integral part of the pilot study, a workshop to share the experience of impact evaluation of environmental interventions within this broader community. With regard to

biodiversity, for example, this could be linked to the process to develop indicators under the CBD. This has not been budgeted.

Country level partners, particularly in countries receiving large GEF inputs, should also have considerable interest in assessing the impacts of the GEF portfolio, particularly in view of the results-based intentions of the RAF.

The impact results would be an important input into OPS4, enabling this to say something about long term results of the GEF's work for the first time.

9: THE PILOT STUDY

Since the Climate Change portfolio has figured heavily in the WB impact studies and the IW portfolio is likely to have a long lead time, before an impact evaluation approach can be applied to its long-term TDA-SAP method, it is appropriate to focus on the biodiversity focal area for the GEF EO pilot. The pilot study therefore will develop an impact evaluation approach for the biodiversity focal area. This approach will be theory-based and can be adapted for use in other focal areas. The adoption of a theory-based approach will mean that a mix of projects can be included in the impact evaluation program, including those under implementation; since it will be possible to review the degree of fit between the change theory on which they are based, the local context, which will be an important element influencing their achievement of results and the implementation approach adopted.

The pilot study will undertake the following activities:

- Discussions with biodiversity focal area team of the GEF Secretariat on the Approach Paper and refinement of details
- Discussions with Implementing Agency Evaluation Offices to determine their potential collaboration in the pilot and main phase of the impact evaluations
- Desk review of biodiversity policy, strategy and other planning documents, including guidance from the Convention and COPs
- Detailed desk review of systems of measurement of environmental change, further developing the approach of Annex B of this Approach Paper
- Exploration of the SEEER methodology for assessing environmental effects of resource and environmental decisions, as developed by GHK Consulting
- Establishment of major theories of change underlying the focal area as a whole and specific approaches within it (e.g. protected areas, production landscapes, capacity building)
- Theory-based analysis of sample of projects in the biodiversity portfolio, which represent the major underlying theories of change in the focal area activities, focusing on the results reported in available terminal evaluations, program and cross-cutting studies
- Preparation of overview of reported results against those intended and summary of previous analyses of performance, including reported deficiencies in data on impacts
- Generation of key questions concerning impacts
- Development of detailed evaluation strategy, in order to provide a meaningful exploration of impact, based on aggregations of interventions based on common approaches to

environmental change (e.g. policy strengthening, capacity building, protected areas, production landscapes, Income Generating Activities, eco-tourism)

- Workshop with EO, biodiversity team and IA evaluators to review and revise research strategy
- Selection of cluster of projects to be evaluated during the first year of the impact evaluation program
- Development of detailed and budgeted Terms of Reference for first year of impact evaluation, including field-based evaluation of several projects within one or more common approaches
- Preparation and submission of pilot study report to Council, including TOR and budget for the first year's evaluation work.

It should be noted that the development of this Approach Paper has shown that the initial intention, outlined in Document GEF/ME/C.25/3, to undertake a field-based pilot impact evaluation in FY06, was premature. The approach adopted will be to develop a sound and affordable approach to impact evaluation in FY06, to test and refine this on the basis of desk reviews and to obtain the collaboration of Implementing Agency Evaluation Offices and Focal Area Teams, before commencing to the fieldwork stage in FY07.

10: PROCESS AND WORK PLAN

Key dates:

- February 10th 2006 - Final Draft Approach Paper Approved
- February 15th 2006 - Finalization of TOR for Pilot Phase and for consultant inputs
- February 28th 2006 - Consultant contracted to undertake desk reviews and assist in development of research strategy
- April 5th, 2006 – Report on research strategy for next phase circulated
- April 14th 2006, Workshop with EO, biodiversity team and IA Evaluators
- May 5th, 2006 – Pilot impact evaluation report finalized, providing details of evaluation strategy, TORs and budget for first year of main study.

11: RESOURCE REQUIREMENTS

GEF EVALUATION OFFICE STAFF INPUTS:	
Director: 3 weeks	
Senior Staff: 10 weeks	
Junior Staff: 10 weeks	
Travel and miscellaneous expenses	$2,000
Workshop costs	$1,000
Contingency (10% of pre-contingency budget)	$6,000
Total GEF EO Cost	**$9,000**
CONSULTANT	
60 days at $500	**$30,000**
Total Budget	**$39,000**

12: COMMUNICATION AND DISSEMINATION OF RESULTS

A Final Report on the Pilot Phase will be provided to Council and placed on the GEF web site. Emerging findings and developments of the impact assessment methodology will be disseminated through working papers, presentation at appropriate fora and publication in evaluation journals. The study team will collaborate with the knowledge management group in EO to determine the most appropriate dissemination strategy.

ANNEX A: IMPACT EVALUATIONS OF GEF PROJECTS (World Bank GEF Office)

Lessons and Questions for the Evaluation Office provided by the World Bank GEF Office Impact Evaluations

The World Bank's GEF Office has already commissioned several impact evaluation studies of GEF projects[15]. A cluster of four energy efficiency projects was evaluated, as well as one biodiversity project in Ecuador. Although commissioned by the World Bank GEF Office, these studies were conducted by independent consultants and do not show any major biases. The quality and approach of the studies is sound and they therefore provide valuable insights for the proposed EO impact evaluations. Some key lessons to be incorporated into planning the EO studies are:

- Impact evaluations are likely to begin with projects, which were developed early in the GEF experience. Many such projects had no logical framework and poor monitoring information, which makes impact evaluation extremely difficult. Counterfactuals developed by the Bank studies to overcome this difficulty varied considerably in the extent to which they were grounded by any factual data or were "scenario construction", making comparison between project achievements and counterfactuals highly speculative.
- Studies encountered reluctance to participate on behalf of Bank staff and national counterparts, who did not appreciate the value of the study and were reluctant to devote time to a project of no current interest.
- Consideration needs to be given to what level of conclusions and lessons are expected from impact evaluations. Many of the lessons from the Bank-GEF studies do not differ substantially from "good practice" examples generated by terminal evaluations. Is it appropriate to use impact evaluations to generate such lessons, or should this level of study target a higher level of lesson? What level would that be?
- Lessons from a set of related impact evaluations are more compelling than those from an isolated study of an individual project. The latter cannot be placed in a suitable framework to triangulate findings and separate those that are unique to the project from those which have broader significance. The set of Energy Efficiency studies provide interesting lessons, whereas it is difficult to interpret the one biodiversity study completed. It would be better for the EO study to avoid "one off" evaluations, except of very large or otherwise important interventions.
- How will impact evaluations relate to changes in the global environment, which are the long term objective of GEF activities? It will rarely be possible to show that one project contributed to this level of change. It might be possible to see such changes at the level of a set of related activities in one country or region. Major IW projects might be expected to make a discernible contribution at this level, but it seems that progress is slow in moving from the preparatory TDA stage to the implementation of a SAP, at which point physical impacts might be generated. It may be still too early to assess impacts from IW activities.

[15] These are discussed in Annex A

- Could impact evaluations be used to focus on specific macro-level changes intended to be monitored by global indicators (e.g., protected areas brought under effective management) to give substance to these indicators?
- Could impact evaluations be brought to bear on the assumptions of the RAF in any way? The RAF is partly based on the performance of country level delivery of global environment gains. Can these be substantiated by impact evaluations of major projects or sets of activities? This could be tied in to country portfolio reviews in future.

Could impact evaluations be used to assess the long-term results of "higher-level" GEF outcomes, such as changes in policies or capacity at institutional or national level?

I. Four Energy Efficiency Projects

A set of four Energy Efficiency projects was evaluated, in Poland, Mexico, Thailand and Jamaica. Key issues which the impact evaluations addressed were:

- Contribution of outcomes to the achievement of expected impacts;
- The project's impacts on global environmental benefits;
- The project's impacts on institutional development;
- The project's impacts on beneficiaries (e.g., savings, knowledge);
- The project's impacts on market development in the energy efficiency sector;
- The project's impacts on country organizations including capacity development;
- The project's impacts on mainstreaming global environment concerns into national development and sector policies;
- Replicability of project outcomes achieved and catalytic effect of the project;
- Lessons learned for the sustainability of project impacts and market transformation;
- Lessons learned regarding achievement and measurement of impacts; and
- Lessons learned for improving the design and management of future activities.

The studies also attempted to separate project-specific impacts from those due to other sources, including any follow-up projects. The methodology employed was developed by two independent consultancy groups, on the basis of TOR produced by the GEF Office of the World Bank. Basic elements were:

- Development of impact maps (Log Frames, which did not exist since these were early projects)
- Identification of key indicators
- Development of an approach to attribution
- Development of an approach to information gathering.

All four projects had objectives of developing DSM (Demand Side Management) capacity, reducing electricity consumption and GHG emissions, and developing technical and financial program models that could be replicated. Findings from the four studies were synthesized and key results identified were:

- "Major market transformation in the residential sector, primarily with respect to lighting, but also including refrigerators and air conditioners (in Thailand)

- No significant transformation in the institutional, commercial or industrial sectors, despite having targeted those sectors in Thailand and Jamaica
- Significant and sustainable energy savings and GHG emission reductions associated with the transformation of the residential markets
- Significant program replication and extension, both in the countries themselves and in surrounding countries
- Some development of capacity for DSM and energy efficiency within government institutions but with moderate to significant gaps remaining
- Significant benefits for consumers in terms of cost savings and improved product quality
- Significantly enhanced opportunities for distributors and retailers of energy efficient equipment
- Significantly improved competitiveness of manufacturers (in Thailand only)
- Small contribution to the integration of energy efficiency objectives into energy policies
- Minimal to modest contribution to the mainstreaming of global environmental issues into energy policies
- Modest contribution to the development of procedures and tools needed for global flexibility mechanisms such as CDM/JI".

Main **impacts** were reported as follows:

- "Energy savings and corresponding GHG reductions have been sustained and expanded upon since the end of the projects.
- Annual reductions throughout the analysis period are in the range of 0.5% of electricity sector emissions (except Thailand = 3.5%).
- There are corresponding reductions in SO_2 and NO_X emissions, of local environmental interest.
- Program costs are in the range of US$1 to $5 per tonne of ultimate reduction, except in the case of Jamaica at $40/tonne.
- Benefit-Cost ratio is in the range of 2 to 5, except for Jamaica at 0.4".

One of the main outputs of the studies for the World Bank was a series of lessons learned, grouped under project cycle stages, namely; project identification, project preparation, public education and outreach, capacity building and project implementation.

The synthesis report of the four DM projects also derived **lessons with regard to the conduct of impact evaluations** (confusingly referred to as "impact assessments"). These were broken down into three stages; planning the studies, data collection and analysis. Key lessons in this area were as follows:

<u>Planning an impact evaluation:</u>

- "Impact assessments are greatly facilitated when information is collected throughout the project life-cycle and an appropriate evaluation framework is established from the beginning.
- There is substantial confusion (particularly in recipient countries) between the various forms of audits, completion reports, assessments, and evaluations. There is often a sense

that projects are over-analyzed and there is sensitivity to potential criticism concerning government policies and management following projects.

- Project information and ongoing monitoring information can be easily lost during staff changes at the Bank or in the implementing entity.
- Log frames/impact maps are at the core of an impact study and are a key tool to establish the impact storyline. Even if the results framework was relatively clear at the time of project implementation, the story is likely to have changed during the intervening years.
- Attribution of impacts is an extremely difficult and uncertain exercise. There are conceptual challenges (i.e. defining the baseline) as well as practical challenges in obtaining information and credible views on the relative importance of the various factors that may have influenced the outcomes. In particular, it may be almost impossible to get a consensus on the degree of influence and timing of various factors".

Data Collection:

- "It is very difficult to obtain information on actual events (e.g. time-series data on sales) in some countries, let alone data for counter-factual scenarios.
- Notwithstanding the general lack of time series information, much more information is now available on the web – even in developing countries.
- Identifying and gaining access to key informants is essential. However this can be difficult, particularly without a local contact.
- Field missions are essential opportunities to collect opinions as well as documentation that are hard to obtain otherwise. However, it is easy for missions to be diverted from core objectives".

Analysis

- "The lack of certain data makes it difficult to develop credible estimates of impact.
- Understanding the context of the changes is very important to drawing conclusions on attribution.
- In addition to understanding the project and country context, it is important that the analysis be grounded in a thorough understanding of the evolution of the World Bank's own approaches.
- The significance and value of the analysis is greatly enhanced by a comparative study of several projects".

II. Biodiversity Project Impact Evaluation

In addition to the set of impact evaluations of Energy Efficiency projects, the World Bank GEF Office commissioned a study of the **Biodiversity Protection Project in Ecuador.** This study generated a set of lessons relevant to Protected Area and Biodiversity projects in Ecuador, notably the following:

Lessons Related to Global Environmental Benefits

- "Failure to establish key performance indicators related to biodiversity hindered the ability of the project to track its own performance related to the environment, and severely hindered the ability for assessment of impacts or sustained results to take in the future. Beyond measuring environmental aspects in particular, the benefits of a well-developed and utilized system of M&E and its key tools not only include the ability to measure performance in an ongoing manner, but also lessened the project's ability to follow an 'adaptive management' approach during implementation.

- Longer-term impacts in BD conservation and PA management require sustained efforts in awareness-raising at the local level. Such awareness raising and capacity development efforts can translate into direct environmental benefits over time, through improved environmental practices and more sustainable natural resource management. By sustaining awareness-raising at the local level, through education, information dissemination and capacity development programs, changes in attitudes and behavior can occur over time.

- The benefits reaped by the public of well-managed PAs may not be well known. If the public is made more aware of such benefits, like clean water they receive from PAs in the Ecuadorian case, it may serve to raise the public's 'valuation' of the PAs and contribute to raised prioritization by the population at large for their protection and good management. Where civil society is strong, this could potentially contribute as well to enhanced state commitment to PAs".

Lessons Related to Capacity Development and Institutional Strengthening

- "In the context of often-increasing extractive activity in PAs, particular capacities in business administration and management are increasingly required for stakeholders at a variety of levels, from Ministry staff, down to park managers. Such capacity development could improve overall capacity for addressing pressures to PAs and for increasing economic and financial potential.

- A weak and/or unclear enabling environment can greatly exacerbate capacity challenges in the implementation and enforcement of a national system of PAs, as is the case in Ecuador. Addressing the enabling environment through addition/modification of clear

and improved laws, policies, and regulations for both participation in and extraction of resources from, PAs can greatly improve implementation and enforcement capacity overall;

- Poor collaboration, coordination, and communication between the center and the periphery (in the case Ecuador, between the Ministry and the regional, municipal and local level stakeholders in a national system of PAs) further weakens planning, implementation, and enforcement capacity overall. In particular, the sharing of lessons and experience, and the ongoing participation of all relevant stakeholders can enhance capacity in the system overall; and,

- In a context of decentralization, such as the one in Ecuador, local governments gain the right to govern, and make key decisions related to PA management and the environment, but they often still require key training and sufficient resources. Institutional strengthening for decentralized units of government can assist implementation and enforcement overall related to PAs.

- In an unstable context, such as that experienced in Ecuador over the last decade, wrought with political change and high turnover at the government level, it is important to ensure the participation and build the capacity of persons and organizations that will remain over time in the NSPA. This means that project design, as well as capacity development and institutional strengthening must target a wide variety of stakeholders at all levels. Ongoing consultative participatory committees offer an opportunity for needed participation and collaboration.

- When different experiences and lessons to learn from different forms of collaboration at the park management level, including successful examples of co-management between the government, municipalities, NGOs, local communities, and different ways of catalyzing fruitful interactions between different stakeholders in PAs, exist in abundance in the country (or elsewhere), such as it is in the case of Ecuador (and countries like Guatemala and Bolivia), these should be well disseminated and shared, or order to improve learning and practices at all levels.

In terms of financial sustainability in particular, the following lesson is noted:

- The lack of financial sustainability of the NSPA can often be the result of limited political will. With further knowledge on the potential benefits to be reaped by prioritizing BD and PAs in the country, further commitment to its support may emerge".

<u>Lessons Related to Local Beneficiaries</u>

- "Local level education and capacity development programs can have a significant immediate and longer-term effect on local beneficiaries for awareness raising and improved capacities in natural resource management. These may not only spread or have

more cumulative effects, but can also serve to enhance cooperation between PA staff and local communities.

- Follow up actions are necessary at the local level after receiving good participation from local communities (or NGOs for that matter), in order to build on successes, as well as to demonstrate that participants' thoughts and ideas are being put into action by those who make decisions and implement programs. Failure to do so can, over time, result in some significant losses in collaboration, and can serve to undermine to enhanced relations that come about through participatory action at the local level".

Lessons Related to Mainstreaming

The following lesson, pertaining to 'mainstreaming' is noted. However, the general validity of this finding is cast into serious doubt by findings of the GEF EO Local Benefits Study, based on a much larger number of projects.

- "There is an important and mutually beneficial link between BD conservation (and improved PA management) and economic development (and livelihoods) that should be further demonstrated and disseminated. Raised awareness of this link at the national level can contribute to mainstreaming of environmental issues into other sectors in general and into laws and policies in particular. This is particularly potentially potent in Ecuador, where there is significant potential for raising awareness of the PAs' social, environmental and economic value".

Lessons for Project Design and Management

The impact evaluation noted the following lessons:

- "the creation of a distinct project implementation unit had hampered the effectiveness of the project,
- projects, which seek to effect changes at the system level, should be designed with adequate resources, especially with regard to time
- it is important that projects like GEF II (i.e., the second phase of the project) ensure that extremely transparent procedures are followed – there may be a need, in particular, for ongoing independent monitoring of the GEF II project
- GEF II would benefit from promoting further accountability for results".

Lessons for impact assessment of BD projects

The study drew the following lessons on this dimension.

- "It is important, where possible, to select projects for impact assessment that were "state-of-the-art" in terms of monitoring, and which therefore may have generated quantitative

information on global environmental benefits[16]. At a minimum, projects selected should have used some performance indicators during the project, in order to provide some baseline data for the impact assessment.

- It would be beneficial to select projects that only targeted single PAs, or allocate far more than 4 person-weeks to the site visits. This will allow for more fruitful and useful analysis on the ground.

- To test a detailed methodology for impact assessment in Biodiversity, it would be interesting and useful to select a project that had achieved some degree of measurable biophysical results, which could be used for assessing sustainability of results and measuring changes over time.

- The purpose of the site visits and the methodology to be used should not only be clearly communicated to World Bank staff in the country, but also to field staff who will be responsible for arranging the site visits.

- All final evaluations, in particular Independent Evaluations, must be provided to the impact assessment team during the methodology development phase. This will allow whatever available baseline information to be integrated into the key tools to be utilized for the assessment".

[16] The GEF EO team treats this lesson with caution, since it is possible that projects with better data are also better managed, which could promote a pronounced selection bias.

ANNEX B: Linking GEF Project Level Impacts to Global Environmental Status

THE NEED TO LINK PROJECT LEVEL IMPACTS TO GLOBAL IMPACTS

The objective of the GEF is to promote global environmental benefits in the focal areas of biodiversity, climate change, international waters, land degradation, ozone and persistent organic pollutants. The GEF does this by financing the incremental costs of projects that are focused nationally, regionally or globally.

The GEF Evaluation Office intends to carry out a pilot project impact evaluation, which will explore an approach to assessing achievements at the project level in a manner which can contribute to the understanding of the impact of GEF activities at the global level. There are more than 2000 projects in the GEF portfolio, with a combined investment of $5.25 billion dollars (GEF 2005b). It is therefore important and timely to begin to assess the GEF's achievements in a systematic manner.

Although the initial focus will be on impact evaluation at the level of individual projects, the study will place these impacts in the context of global environmental status to help understand the contribution of the GEF at this higher level. This is a significant challenge, given that there are few widely agreed global level indicators in the GEF focal areas. Furthermore, GEF projects do not use a common set of indicators; even within focal areas projects adopt a wide variety of approaches to address the various aspects of global environmental challenges they face.

Defining and measuring global benefits or global environmental status in each of the GEF focal areas continues to be a major challenge, although there has been some progress in this respect in recent years. For example, in the biodiversity focal area, a number of global level assessments and status reports have been conducted and are continuously being updated. These include the IUCN Red List of endangered species, the Global Amphibian Assessment, the Global Mammal Assessment, and the Millennium Ecosystem Assessment.

A number of these global level environmental data sets have been used by the GEF Secretariat in its development of the Global Benefits Indicator (GBI) to be applied in the Resource Allocation Framework (RAF) for the biodiversity focal area. A GBI has also been developed for the climate change focal area, but in a manner distinct from that of the biodiversity GBI.

The GEF EO impact evaluations will attempt to assess impacts from both top-down and bottom-up perspectives, linking indicators of GEF project level impacts to relevant global level environmental status. This will be done through a process of careful identification of appropriate indicators as part of each evaluation, as described below in the relevant focal area sections. In biodiversity for example, both ecosystem and species level indicators may be appropriate depending on the Global Environmental Objective of the project being evaluated. At the project level it will also be important to identify impact indicators for changes in human behavior, although these cannot readily be aggregated at the global level. The biodiversity focal area is given the most comprehensive discussion in this annex, since it is the most developed of the

focal areas in regard to attempts to measure global environmental status. It therefore can be used as an example and to some extent as a model for the international waters and land degradation focal areas. These three focal areas are measured in terms of environmental status, and are tied to biotic factors. The climate change, ozone and POPs focal areas may require different paradigms in terms of how global environmental status is measured, although the same general concept of linking project level impacts to global impacts can be applied. These three focal areas are quantified in terms of quantities of specific anthropogenic inputs to the environment, which subsequently have harmful environmental effects. The international waters focal area crosses these distinctions to some extent, and the type of intervention by an international waters project may affect the impact evaluation approach. Biodiversity and climate change are the only focal areas currently covered by the RAF.

BIODIVERSITY FOCAL AREA

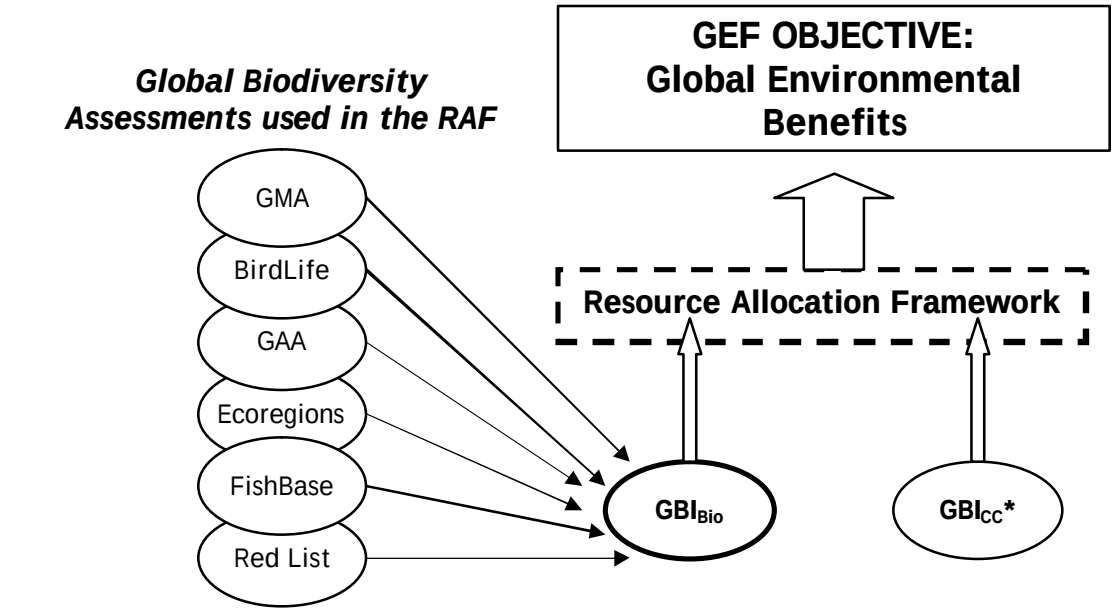
The biodiversity focal area has developed more environmental indicators and global data sets than other focal areas, both within the context of the GEF and at the international level. The discussion below first describes how the GEF Evaluation Office impact evaluations will seek to link project level impacts to global level impacts through the use of the indicator framework developed as part of the RAF. Other relevant ecosystem and species level global data sets in the biodiversity focal area, which could be used within the context of the impact evaluations, are also presented.

Linking Project Level Impacts to Global Environmental Status

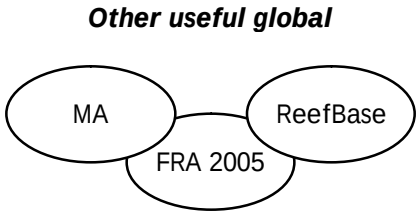
The GEF RAF has developed a GBI for both the biodiversity, GBI_{Bio}, and climate change, GBI_{CC}, focal areas (GEF 2005a). In the biodiversity focal area, this was achieved through the amalgamation of several existing global environmental data sets. These include: WWF Ecoregions, IUCN Red List, Fishbase, and GIS data from the International Food Policy Research Institute. The GBI_{Bio} is disaggregated at the species and ecosystem levels (see Fig. 1). This reflects the practice of the Convention on Biological Diversity (CBD), which defines biodiversity at the ecosystem, species and genetic levels (CBD 1992). The RAF considers that the genetic level of biodiversity is incorporated within the ecosystem and species levels (GEF 2005a) (see Fig. 2).

The GBI_{Bio} is broken down into two factors of ecosystem and species level biodiversity status: a.) representation; and b.) threat level (with respect to species), or quality (with respect to ecosystems) (see Fig. 2). Impact evaluations will focus on the second of these factors since GEF project interventions are targeted at addressing threats to ecosystems and species. GEF projects do not seek to increase the number of species represented in a country beyond the natural level, nor to increase the number of ecosystems within a country.

Fig. 1: Development of the GEF Resource Allocation Framework (GEF 2005a)

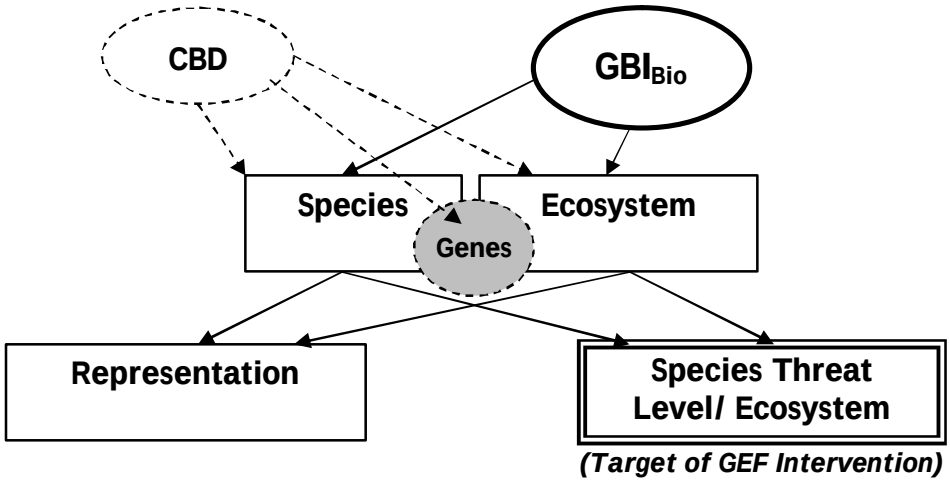


BirdLife – World Bird Database
Ecoregions – WWF identified global ecoregions
Fishbase – Global Fish database
FRA – FAO Global Forest Resources Assessment
GMA – Global Mammal Assessment
GAA – Global Amphibian Assessment
MA – Millennium Ecosystem Assessment
Red List – IUCN Red List of Endangered Species
Reefbase – Global Coral Reef Information System



*GBI_{CC} = Baseline GHG Emissions x (Carbon $Intensity_{1990}$ / Carbon Intensity

Fig. 2: GEF Global Biodiversity Benefits Indicator (based on GEF 2005a)



Although built on widely accepted global level biodiversity scientific data, the RAF is primarily an operational tool to determine potential inputs, not results. To move toward the identification of results, GEF impact evaluations will build on two lessons from the RAF. The first of these is the possibility of assessing links between GEF project environmental objectives and global level data sets. These links can be derived on the basis of the data sets used in the RAF, as well as other relevant global level environmental status assessments that have not been included in the RAF at this initial stage, such as the Food and Agriculture Organization's (FAO) Global Forest Resources Assessment 2005 (FRA 2005). Secondly, the evaluation can take as a starting point the approach of breaking down the biodiversity focal area into ecosystem and species level components.

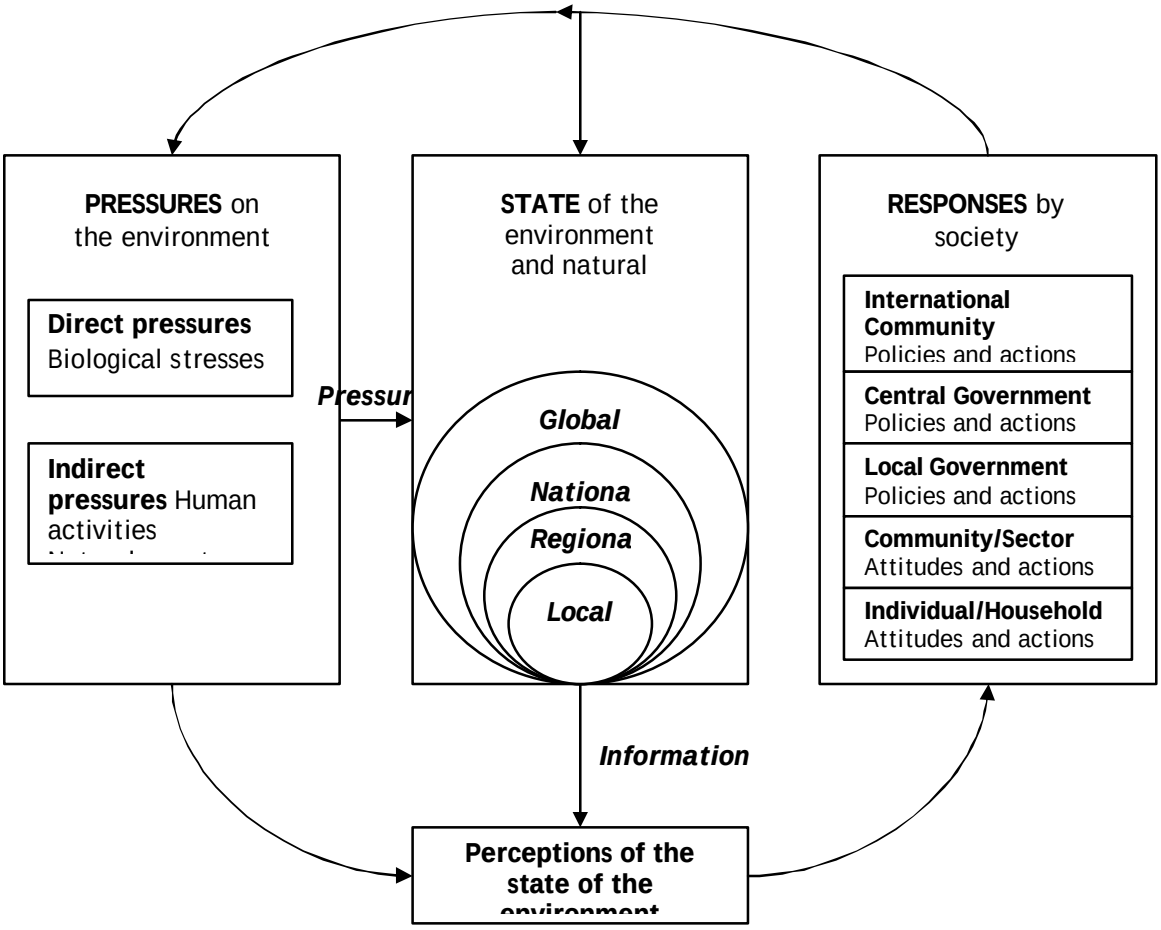
Although the impact evaluations can use species and ecosystem level indicators as a starting point, future evaluations will also need to consider the full range of global biodiversity indicators. For the time being, species and ecosystem level indicators are being developed and considered under the CBD. These include:

- Trends in extent of selected biomes, ecosystems and habitats;
- Trends in abundance and distribution of selected species;
- Change in status of threatened and/or protected species.

Another indicator in use is the coverage of protected areas, although this is seen as an indicator of policy response and not directly of biodiversity status (EASAC 2005).

GEF impact evaluations will make use of species and ecosystem level indicators previously identified by the project in question whenever possible. Many of these indicators will require adjustment so that they can be linked to indicators of global level environmental status. Once potential species and ecosystem level indicators for the project being evaluated have been identified, the evaluations can use a Pressure-State-Response framework to identify impacts to which the project has contributed (through a reduction of threats or pressures, improvement of the state, or improvement of the response). The Pressure-State-Response (PSR) framework was developed by the OECD in the early 1990s and has been widely applied to the understanding of environmental issues. This type of framework is shown in Fig. 3, in a model adapted from the Government of New Zealand.

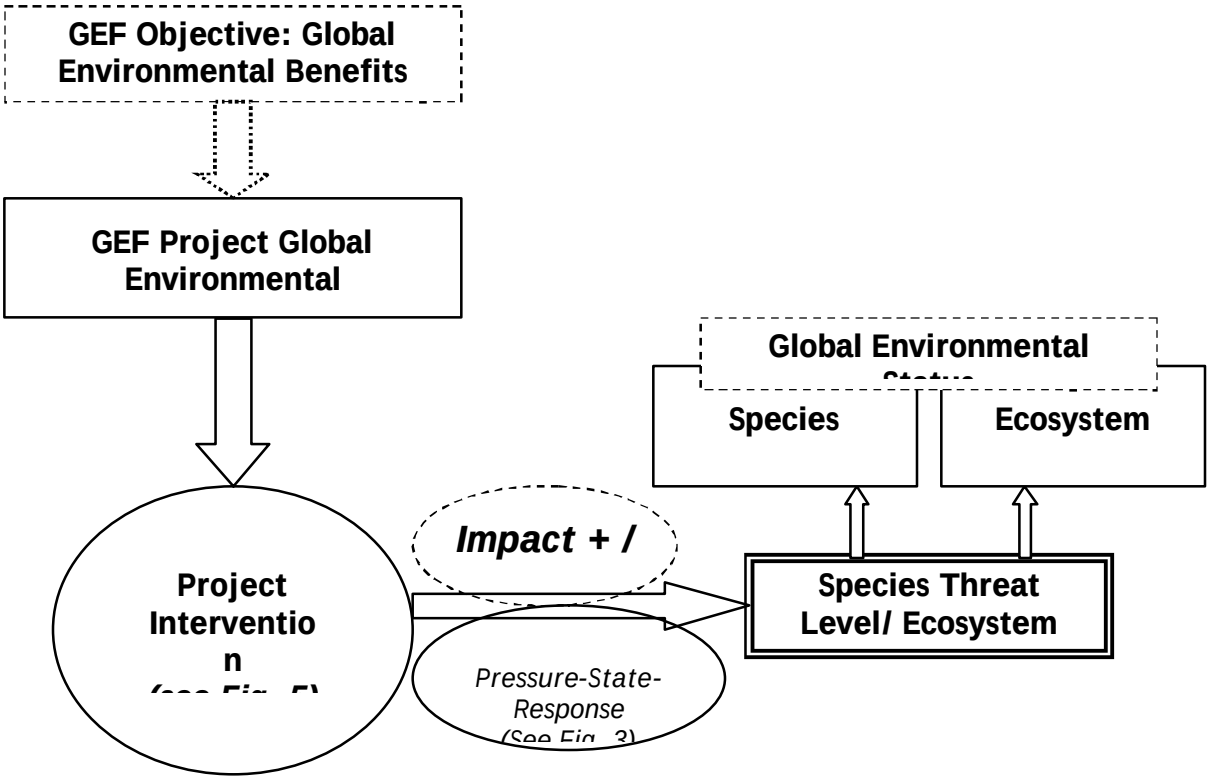
Fig. 3: The OECD Pressure-State-Response Framework (adapted from GONZ 2005)



There are also some "threat analysis" tools that have been developed to help identify where environmental conservation interventions should be targeted. For example, one such tool was used, with mixed success, in the GEF project "Reducing Biodiversity Loss at Cross-Border Sites in East Africa," (UNDP 2004). However, the GEF impact evaluations will only be able to make use of such tools where they have been systematically applied during project implementation. PSR analysis will be conducted during the impact evaluation based on information available at the project level.

The intended pathway of GEF project interventions is shown in Fig. 4 below. GEF projects identify a global environmental objective and, through a set of actions, seek to influence pressures, the state and the response to environmental threats. This theoretically results in a positive (or, in reality, sometimes negative) impact on the environmental status of targeted biodiversity resources, as identified through species or ecosystem level indicators.

Fig. 4: Intended GEF Project Impact Pathway



The general project intervention framework is shown in Fig. 5 below, as adapted from UNDP 2000.

The link from project level impacts to global level impacts is then made by applying the appropriately identified project level species and ecosystem indicators to global level data sets and environmental assessments. This is illustrated in Fig. 6 below. This step is much easier to carry out at this point in the biodiversity focal area because of the large number of scientific global level assessments that have been carried out. The majority of the most widely known assessments and data sets are identified in Fig. 6, but there may be others.

Examples of how this broad approach might be applied are discussed at the end of this paper, on the basis of three projects being considered for the pilot impact evaluation. Adequate project level data must be available for review so that project level impacts can be sufficiently assessed before linking project impacts to global level environmental status. The proposed model represents the ideal situation, but in reality some aspects of data availability will be limited and the evaluation team will have to draw on all available resources. Since time and resources to carry out comprehensive scientific assessments at the project level are not available, GEF impact evaluation teams will utilize "triangulation methods" (see Fig. 7) of evaluation to identify project results.

Fig. 5: The Country Level Impact Assessment Framework/Linkages *(Adapted from UNDP 2000)*

Methodology Evaluation/Achievements

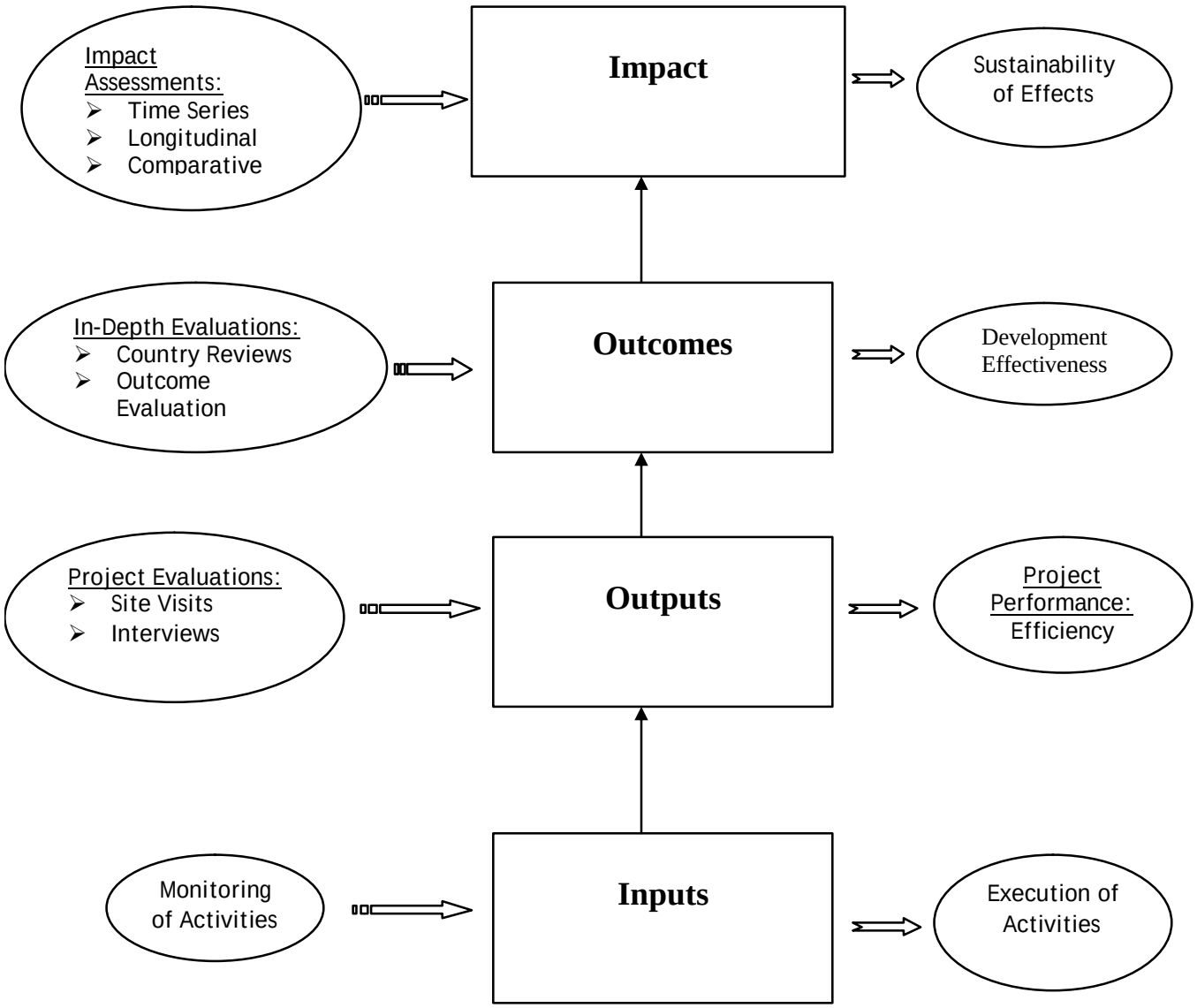


Fig. 6: Proposed Model Linking Project Level Impact to Global Environmental Status

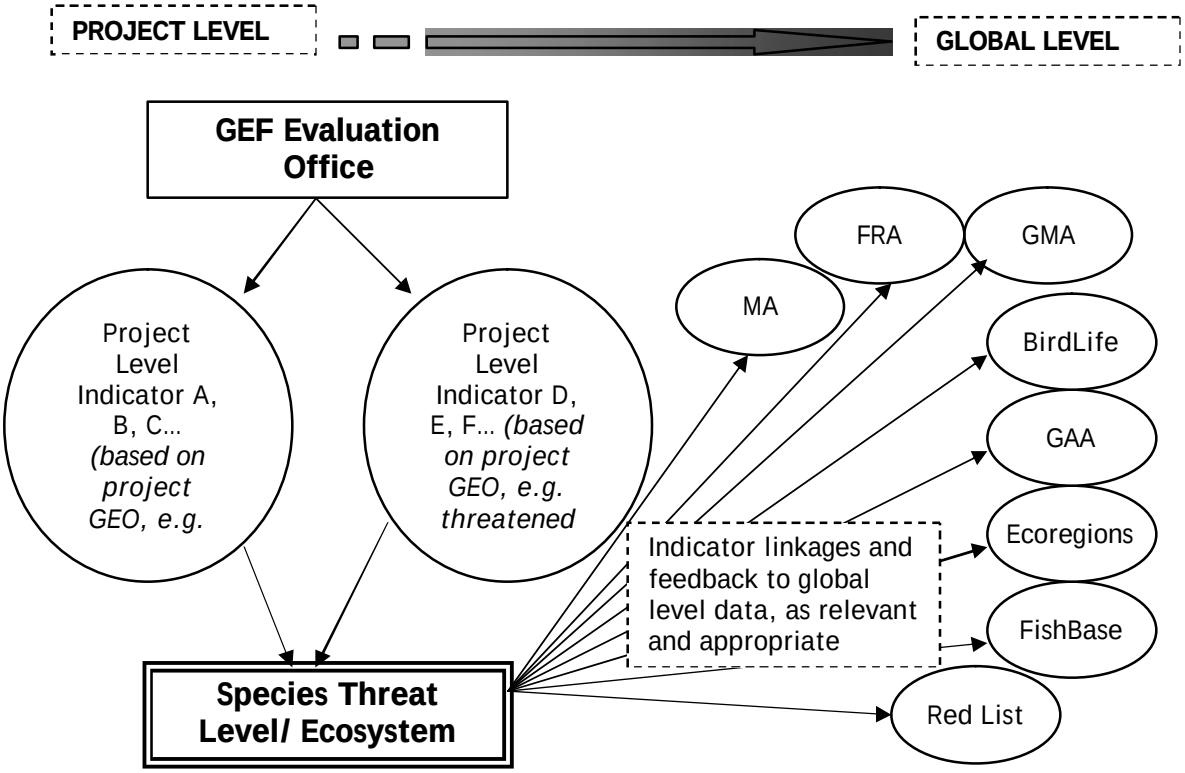
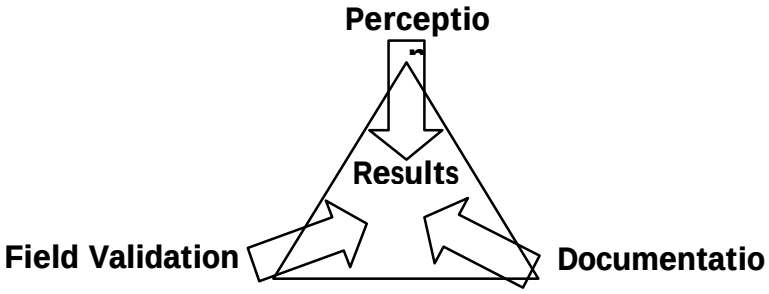


Figure 7: Triangulation Method of Evaluation



Global Biodiversity Assessments (Ecosystem Level)

Millennium Ecosystem Assessment

The Millennium Ecosystem Assessment (MA) was partially funded by the GEF, and was completed in early 2005. A brief overview of the MA is included below.

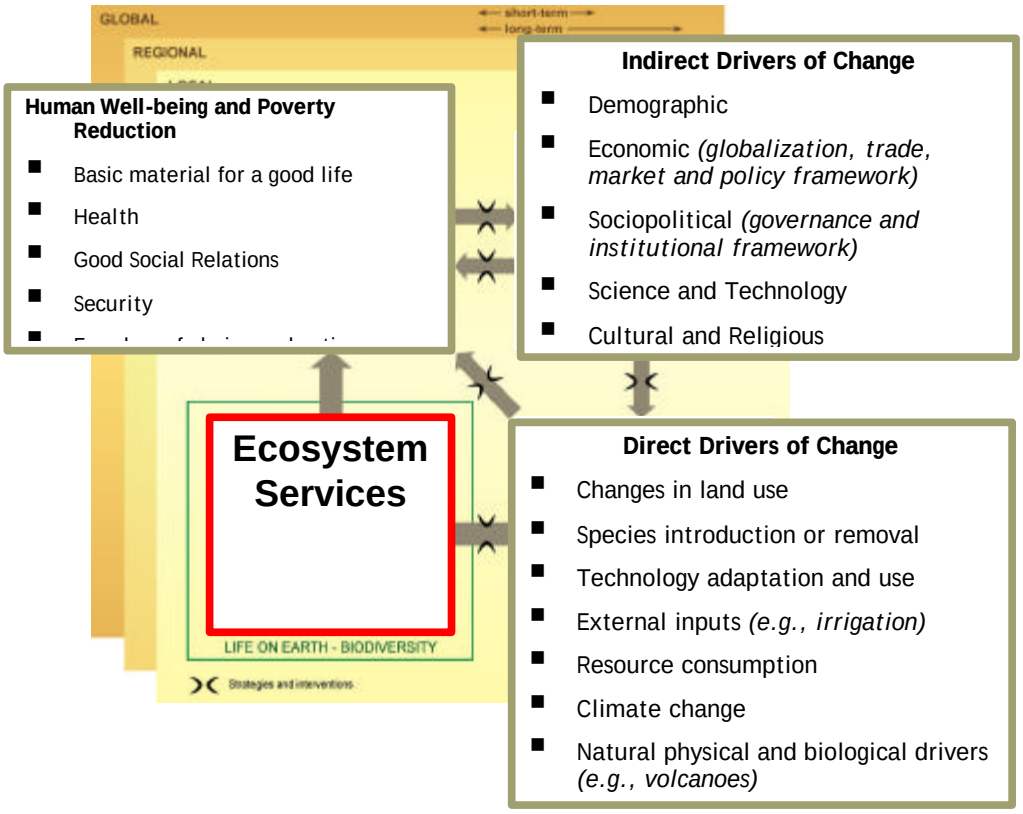
"The Millennium Ecosystem Assessment (MA) is an international work program designed to meet the needs of decision makers and the public for scientific information concerning the consequences of ecosystem change for human well-being and options for responding to those changes. The MA was launched by U.N. Secretary- General Kofi Annan in June 2001 and was completed in March 2005. It will help to meet assessment needs of the Convention on Biological Diversity, Convention to Combat Desertification, the Ramsar Convention on Wetlands, and the Convention on Migratory Species, as well as needs of other users in the private sector and civil society. If the MA proves to be useful to its stakeholders, it is anticipated that such integrated assessments will be repeated every 5– 10 years and that ecosystem assessments will be regularly conducted at national or sub-national scales.

The MA focuses on ecosystem services (the benefits people obtain from ecosystems), how changes in ecosystem services have affected human wellbeing, how ecosystem changes may affect people in future decades, and response options that might be adopted at local, national, or global scales to improve ecosystem management and thereby contribute to human well-being and poverty alleviation. The specific issues being addressed by the assessment have been defined through consultation with the MA users.

The MA synthesizes information from the scientific literature, datasets, and scientific models, and includes knowledge held by the private sector, practitioners, local communities and indigenous peoples. All of the MA findings undergo rigorous peer review. More than 1,300 authors from 95 countries have been involved in four expert working groups preparing the global assessment, and hundreds more continue to undertake more than 20 sub-global assessments. The findings are contained in the fifteen reports listed in the box above.

The MA is an instrument to identify priorities for action. It provides tools for planning and management and foresight concerning the consequences of decisions affecting ecosystems. It helps identify response options to achieve human development and sustainability goals, and has helped build individual and institutional capacity to undertake integrated ecosystem assessments and to act on their findings," (MA 2005a). The MA conception framework is shown below in Fig. 8. The main findings of the MA can be found in the publication "Ecosystems and Human Well-Being: Synthesis" (MA 2005b).

Fig. 8: Millennium Ecosystem Assessment Framework (MA 2005b)



For GEF-EO impact evaluations the most relevant portion of the MA Framework is the Ecosystem Services box. The MA classified ecosystem services into four types: provisioning, regulating, cultural, and supporting. Through the various assessments conducted as part of the MA, the global status of the provisioning, regulating and cultural ecosystem services were evaluated. The assessment of ecosystem services included a wide range of data sources and methods, and the final assessment of the ecosystem services was reached through consensus of the more than 1,300 scientists involved in the MA process. The MA concluded that the conditions of the majority of ecosystem services have been degraded. The results of the assessment are shown below, as drawn from the MA synthesis document (MA 2005b, pg. 7). Impact evaluations could consider whether GEF projects have enhanced (or slowed the degradation of) the various ecosystem services described in the MA. Different types of ecosystem services will have to be quantified in different ways, and this will present methodological challenges.

Table 7. Global Status of Provisioning, Regulating, and Cultural Ecosystem Services Evaluated in the Millennium Ecosystem Assessment.

Status indicates whether the condition of the service globally has been enhanced (if the productive capacity of the service has been increased, for example) or degraded in the recent past. Definitions of "enhanced" and "degraded" are provided in the note below. A fourth category, supporting services, is not included here as they are not used directly by people.

Service	Sub-category	Status	Notes
Provisioning Services			
Food	crops	↑	substantial production increase
	livestock	↑	substantial production increase
	capture fisheries	↓	declining production due to overharvest
	aquaculture	↑	substantial production increase
	wild foods	↓	declining production
Fiber	timber	+/–	forest loss in some regions, growth in others
	cotton, hemp, silk	+/–	declining production of some fibers, growth in others
	wood fuel	↓	declining production
Genetic resources		↓	lost through extinction and crop genetic resource loss
Biochemicals, natural medicines, pharmaceuticals		↓	lost through extinction, overharvest
Water	fresh water	↓	unsustainable use for drinking, industry, and irrigation; amount of hydro energy unchanged, but dams increase ability to use that energy
Regulating Services			
Air quality regulation		↓	decline in ability of atmosphere to cleanse itself
Climate regulation	global	↑	net source of carbon sequestration since mid-century
	regional and local	↓	preponderance of negative impacts
Water regulation		+/–	varies depending on ecosystem change and location
Erosion regulation		↓	increased soil degradation
Water purification and waste treatment		↓	declining water quality
Disease regulation		+/–	varies depending on ecosystem change
Pest regulation		↓	natural control degraded through pesticide use
Pollination		↓[a]	apparent global decline in abundance of pollinators
Natural hazard regulation		↓	loss of natural buffers (wetlands, mangroves)
Cultural Services			
Spiritual and religious values		↓	rapid decline in sacred groves and species
Aesthetic values		↓	decline in quantity and quality of natural lands
Recreation and ecotourism		+/–	more areas accessible but many degraded

Note: For provisioning services, we define enhancement to mean increased production of the service through changes in area over which the service is provided (e.g., spread of agriculture) or increased production per unit area. We judge the production to be degraded if the current use exceeds sustainable levels. For regulating services, enhancement refers to a change in the service that leads to greater benefits for people (e.g., the service of disease regulation could be improved by eradication of a vector known to transmit a disease to people). Degradation of regulating services means a reduction in the benefits obtained from the service, either through a change in the service (e.g., mangrove loss reducing the storm protection benefits of an ecosystem) or through human pressures on the service exceeding its limits (e.g., excessive pollution exceeding the capability of ecosystems to maintain water quality). For cultural services, degradation refers to a change in the ecosystem features that decreases the cultural (recreational, aesthetic, spiritual, etc.) benefits provided by the ecosystem.

[a] Indicates *low to medium certainty*. All other trends are *medium to high certainty*.

World Wildlife Fund Ecoregion Assessment

WWF has undertaken an assessment of the world's ecosystems, identifying each ecoregion and working to understand its importance and threat level. Through this assessment WWF has produced the "Global 200." According to WWF, "WWF's Global 200 is a first attempt to identify a set of ecoregions whose conservation would achieve the goal of saving a broad diversity of the Earth's ecosystems. These ecoregions include those with exceptional levels of biodiversity, such as high species richness or endemism, or those with unusual ecological or evolutionary phenomena," (WWF 2005). The GEF RAF used data on ecoregions from WWF as part of its scientific basis for the development of the $GBI_{Bio.}$

FAO Forest Resources Assessment 2005

"FAO has been coordinating global forest resources assessments every five to ten years since 1946. The Global Forest Resources Assessment 2005 (FRA 2005) was carried out between 2003 and 2005 and involved more than 800 people, including 172 officially nominated national correspondents, their colleagues, an Advisory Group, international experts, FAO staff, consultants and volunteers from around the world. Information has been collated from 229 countries and territories for three points in time: 1990, 2000 and 2005.

FRA 2005 is the most comprehensive assessment of forests and forestry to date – not only in terms of the number of countries and people involved, but also in terms of scope. It examines the current status and recent trends for about 40 variables covering the extent, condition, uses and values of forests and other wooded land, with the aim of assessing all benefits from forest resources. The results are presented according to six thematic elements of sustainable forest management.

FAO worked closely with countries and specialists in the design and implementation of FRA 2005 – through regular contact, expert consultations, training for national correspondents and ten regional and subregional workshops. The outcome is better data, a more transparent reporting process and enhanced national capacity in data analysis and reporting," (FAO 2005).

The FRA 2005 is the most comprehensive global data set on forests, which is one of the key types of ecosystems for GEF. Potentially useful indicators for which the FRA 2005 could provide data are remaining forest area and quality of that area in each country, with the potential to measure at higher resolution, such as at the site level. Also contributing to data on these indicators could be the WRI Global Forest Watch program, which has comprehensive data on a few developing countries, including Brazil, Chile, Indonesia, Russia, Venezuela, and the Central Africa region.

ReefBase

"ReefBase provides country-level data and information in a logical series of themes:

I. Resources: what kind, and how many reefs are there?
II. Status: how are these resources doing?
III. Threats: which negative influences are impacting coral reefs?
IV. Management: what is, or could be done to monitor and manage these resources?
V. Maps and Photos: view maps and photos of coral reefs

VI. References: access key publications and contacts information" (ReefBase 2005).

This database is potentially very useful for GEF impact evaluations, since it provides a comprehensive global data source on one of the most biodiverse types of marine ecosystems, which is also a key ecosystem for the GEF.

Conservation International Hotspots

According to Conservation International, "To qualify as a hotspot, a region must meet two strict criteria: it must contain at least 1,500 species of vascular plants (> 0.5 percent of the world's total) as endemics, and it has to have lost at least 70 percent of its original habitat," (CI 2005). The Hotspots data can be considered at both the species and ecosystem levels. The data sets produced by Conservation International are not aggregated at the global level. However, Conservation International has produced extensive scientific data on each of these areas, which could be of use to GEF impact evaluations when looking at projects in areas identified as Hotspots.

Global Biodiversity Assessments (Species Level)

IUCN Red List of Threatened Species

"The IUCN Red List of Threatened Species provides taxonomic, conservation status and distribution information on taxa that have been globally evaluated using the IUCN Red List Categories and Criteria. This system is designed to determine the relative risk of extinction, and the main purpose of the IUCN Red List is to catalogue and highlight those taxa that are facing a higher risk of global extinction (i.e. those listed as Critically Endangered, Endangered and Vulnerable). The IUCN Red List also includes information on taxa that are categorized as Extinct or Extinct in the Wild; on taxa that cannot be evaluated because of insufficient information (i.e. are Data Deficient); and on taxa that are either close to meeting the threatened thresholds or that would be threatened were it not for an ongoing taxon-specific conservation programme (i.e. Near Threatened)" (IUCN 2005a).

This global database is potentially of great value to GEF impact evaluations as it represents the most comprehensive data source at the species level in the world, and is widely accepted as having scientific validity. Data from this source were used in the development of the RAF. This would likely be the primary source of global data to link to species level project indicators identified under GEF impact evaluations.

Global Amphibian Assessment

"The Global Amphibian Assessment (GAA) is the first-ever comprehensive assessment of the conservation status of the world's 5,743 known species of frogs, toads, salamanders, and caecilians. This website presents results of the assessments, including IUCN Red List threat category, range map, ecology information, and other data for every amphibian species," (GAA 2005).

The GAA could also be useful to GEF impact evaluations in a way similar to the IUCN Red List, as it presents even more comprehensive data on particular taxa. However, most GEF projects have not specifically targeted individual species of amphibians. Since amphibians are known to

be highly sensitive to environmental changes, these data could also be used as a proxy for overall environmental status in some cases.

Global Mammal Assessment

The Global Mammal Assessment (GMA) presents similar data as the GAA, but it has not yet been completed. Preliminary data from this source were used in the development of the RAF. Once completed, the GMA will be a useful data source for GEF impact evaluations, since many GEF projects do target specific mammal populations, particularly keystone species and large, charismatic species that can be identified as flagship species for a particular project, region or country. Initial information on the GMA can be found on the IUCN Species Survival Commission website (IUCN 2005b).

FishBase

FishBase is a comprehensive online database of scientific data about fish species in both freshwater and marine ecosystems. This global data source was used in the development of the RAF for freshwater species. It could also be useful for GEF impact evaluations when looking at biodiversity projects targeting freshwater or marine ecosystems.

Plants

The GEF Secretariat used global data on vascular plants in the development of the RAF. Further inquiries will be made to determine whether the source of this data could be useful for GEF impact evaluations.

BirdLife World Bird Database

"BirdLife has been investing in the development of information management tools to support the activities of the Partnership for many years. This is a fully relational database, known as the World Bird Database (WBDB). The database architecture provides some 120 tables covering in excess of 1,400 data fields. Data are being added continually, and certain tables already hold in excess of 250,000 records.

The World Bird Database provides the information management tool through which the BirdLife Partnership manages, analyses and reports on the breadth of its scientific knowledge - Species, Important Bird Areas (IBAs) and Endemic Bird Areas (EBAs) – much of these data are available through the Data Zone.

With information on some 10,000 species of bird, over 8,000 IBAs and 218 EBAs managed through the WBDB, together with BirdLife's spatial data, multimedia files, other documents and links, the BirdLife Data Zone is truly a valuable information resource," (BirdLife 2005).

Data from BirdLife International's World Bird Database was used in the development of the RAF. This is a valuable resource for GEF impact evaluations.

INTERNATIONAL WATERS FOCAL AREA

The development of appropriate indicators and global environmental databases is less advanced in the International Waters focal area than in the biodiversity focal area. The International

Waters focal area considers transboundary water bodies to be global resources (GEF 2005c). In this sense, GEF relevant global environmental status impact indicators can be identified at the regional level in relation to the identified threats to the transboundary water bodies targeted by GEF projects.

One example that could provide an interesting pilot for the GEF impact evaluations is the Strategic Partnership for Nutrient Reduction in the Danube River Basin and the Black Sea. Within the context of this GEF supported partnership, the various stakeholders have begun to identify relevant indicators of nutrification that could be applied and aggregated at the regional level. This process was recently furthered at a Nutrient Reduction Conference in Tbilisi, Georgia in early October 2005. Supported by the GEF Evaluation Office, the stakeholders in the strategic partnership are considering means of measuring and aggregating environmental results.

LAND DEGRADATION FOCAL AREA

The Land Degradation focal area presents special challenges. It is a relatively recent addition to the GEF portfolio. In addition, comprehensive global level data are generally lacking. According to the MA, "Information on land degradation in drylands is extremely poor. Major shortcomings in the currently available assessments point to the need for a systematic global monitoring program, leading to the development of a scientifically credible, consistent baseline of the state of land degradation and desertification," (MA 2005b).

The GEF will be partially addressing this lack of data through a project called the Land Degradation Assessment in Drylands (LADA). This project is getting underway in 2005, and is expected to go for 4 years. According to the project executive summary, "The project will deliver a global assessment of land degradation, six regional assessments and six national and local assessments that will identify areas at greatest risk and areas where degradation has been successfully controlled. The project will also result in a replicable standardized methodological framework for land degradation assessment and policy guidance that will be disseminated in six sub-regions affected by degradation mitigation and better targeting of investments" (GEF 2005d). Following the completion of this project in a few years time, the GEF should be better able to identify impacts of its projects in the land degradation focal area, and link those impacts to global environmental status with regard to land degradation in drylands.

In the meantime, an interagency working group has been convened, and is working with the United Nations University to develop indicators for the land degradation focal area as a whole. The working group is currently moving towards agreement on a global benefits framework that will then be populated with appropriate indicators. This process is expected to be completed by the end of 2006.

All GEF focal areas, including International Waters and Land Degradation have strategic priorities and targets developed for the anticipated fourth phase of the GEF. However, most of these targets are for coverage or policy response and cannot be directly linked to global environmental status.

CLIMATE CHANGE, OZONE AND POPS FOCAL AREAS

The Climate Change, Ozone and POPs focal areas all require a slightly modified approach to impact evaluation from the approach for the three previous focal areas. Climate Change is potentially the most difficult, because although the ultimate goal is to reduce global GHG emissions, this is often not the direct objective of GEF climate change projects, which more often seek to catalyze markets, or transfer technology. In this case, projects' linkages to global GHG levels are simply impossible to measure at a practical level. Therefore, linking project impacts to global environmental status may not be appropriate for the climate change focal area. It may be necessary to simply identify relevant project level impact indicators (related, for example, to the relevant target market or to technology adoption in the project area).

The 2004 Climate Change Program Study (GEFME 2004) made significant progress in the realm of quantifying impacts of GEF climate change projects. See Chapter 4 and Annex A of this document for further discussion of the methodology involved in determining the impact of the GEF climate change projects. Around the same time, the Climate Change Team of the GEF Secretariat also developed a methodology for calculating GHG reductions or avoided emissions from GEF projects in order to report on the focal area targets established by the GEF for GEF-3. This methodology is described in the document "Reporting on Performance Targets to Be Achieved by Fall 2004," (GEF 2004). This methodology may well represent a starting point for impact evaluations in the climate change focal area.

The RAF calculates the GBI_{CC} through a "Carbon Intensity Adjustment Factor" (GEF 2005a, see Fig. 1) and it is possible that a GEF project's impact could be measured through a change in the Carbon Intensity Adjustment Factor as calculated for the project area, but the methodological challenges in this respect will be considerable.

The Ozone focal area may be somewhat simpler to document. For example, GEF projects may be able to directly identify their contribution to a reduction in ozone depleting substances, and this can then be linked to the overall global level of ozone depleting substances. In time this could be linked to the degradation of the earth's ozone layer, although the long timeframe for the process of ozone depletion and regeneration makes direct linkages from the project to global level difficult.

The POPs focal area resembles the Ozone focal area in many ways operationally, in that there is X amount of POPs in the world that the GEF is working to reduce. The POPs focal area is relatively new, so it will be some time before POPs projects' impacts can be evaluated. This should be relatively straightforward however, and could easily be linked to the global level, assuming that there is a global baseline measuring POPs in existence. However, GEF projects will not be able to remove POPs that have already been disbursed into the environment in any comprehensive manner. POPs have been found in biological tissues in animals throughout the oceans and on all continents.

REFERENCES FOR ANNEX B

BirdLife. 2005. "Data Zone," website at http://www.birdlife.org/datazone/index.html, as accessed November 29, 2005.

Conservation International (CI). 2005. "Hotspots Defined," website at http://www.biodiversityhotspots.org/xp/Hotspots/hotspotsScience/hotspots_defined.xml. As accessed November 29, 2005.

Convention on Biological Diversity (CBD). 1992. "Convention on Biological Diversity." Montreal: Secretariat of the Convention on Biological Diversity.

European Academies Science Advisory Council (EASAC). 2005. "A user's guide to biodiversity indicators," London: The Royal Society.

Food and Agriculture Organization (FAO). 2005. "Global Forest Resource Assessment 2005," website at www.fao.org/forestry/site/fra2005/en. As accessed November 29, 2005.

GEF. 2005a. "The GEF Resource Allocation Framework." GEF/C.27/Inf.8/Rev.1.

GEF. 2005b. "OPS3: Progressing Toward Environmental Results, Third Overall Performance Study of the GEF, Executive Version." Washington, DC: Global Environment Facility.

GEF. 2005c. "Operational Policies in International Waters," website at http://thegef.org/projects/focal_areas/iw/iw_ops.html, as accessed November 29, 2005.

GEF. 2005d. "Project Executive Summary: Land Degradation Assessment in Drylands (LADA).

GEF. 2004. "Reporting on Performance Targets to be Achieved by Fall 2004," GEF/C.24/3.

GEFME. 2004. "2004 Climate Change Program Study." Washington, DC: Global Environment Facility Office of Monitoring and Evaluation.

Global Amphibian Assessment (GAA). 2005. "Global Amphibian Assessment," website at http://www.globalamphibians.org/. As accessed November 29, 2005.

Government of New Zealand (GONZ). 2005. "Quality Planning" website, at http://www.qualityplanning.org.nz/monitoring/intro-pressure-state-response-framework.php. As accessed November 9, 2005.

Millennium Ecosystem Assessment (MA). 2005a. "About the Millennium Ecosystem Assessment," website at http://www.millenniumassessment.org/en/about.overview.aspx? As accessed November 29, 2005.

Millennium Ecosystem Assessment (MA). 2005b. "Ecosystems and Human Well-Being: Synthesis," Millennium Ecosystem Assessment website at http://www.millenniumassessment.org//en/Presentations.aspx. As accessed November 9, 2005.

ReefBase. 2005. Website, at http://www.reefbase.org/. As accessed November 19, 2005.

UNDP. 2000. "DRAFT: Synthesis Report: For the CLIA Methodology Pilots," UNDP Evaluation Office.

UNDP. 2004. "Reducing Biodiversity Loss at Cross-Border Sites in East Africa: Project Terminal Evaluation Report."

World Conservation Union (IUCN). 2005a. "The IUCN Red List of Endangered Species," website at http://www.redlist.org/info/introduction.html. As accessed November 29, 2005.

World Conservation Union (IUCN). 2005b. "The Global Mammal Assessment," website at http://www.iucn.org/themes/ssc/programs/gma/. As accessed November 29, 2005.

World Wildlife Fund (WWF). 2005. "Global 200 Ecoregions," website at http://www.worldwildlife.org/science/ecoregions/g200.cfm. As accessed November 29, 2005.