



OECD Environment Working Papers No. 74

**Monitoring and Evaluation
of Climate Change
Adaptation: Methodological
Approaches**

**Ayesha Dinshaw,
Susannah Fisher,
Heather McGray,
Neha Rai,
Johan Schaar**

<https://dx.doi.org/10.1787/5jxrclr0ntjd-en>

ENVIRONMENT DIRECTORATE

OECD ENVIRONMENT WORKING PAPER No. 74 - MONITORING AND EVALUATION OF
CLIMATE CHANGE ADAPTATION: METHODOLOGICAL APPROACHES

By Ayesha Dinshaw (1), Susannah Fisher (2), Heather McGray (1), Neha Rai (2) and Johan Schaar (3)

- (1) World Resources Institute
(2) International Institute for Environment and Development
(3) Formerly at the World Resources Institute

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Authorised for publication by Simon Upton, Director, Environment Directorate.

JEL classification: H43, O22, Q54

Keywords: Climate change adaptation, monitoring and evaluation, learning.

OECD Environment Working Papers are available at www.oecd.org/env/workingpapers.htm

JT03367755

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.



OECD ENVIRONMENT WORKING PAPERS

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

OECD Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

This series is designed to make available to a wider readership selected studies on environmental issues prepared for use within the OECD. Authorship is usually collective, but principal author(s) are named. The papers are generally available only in their original language - English or French - with a summary in the other.

Comments on Working Papers are welcomed, and may be sent to Environment Directorate, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France or by-mail: env.contact@oecd.org.

OECD Environment Working Papers are published on
www.oecd.org/env/workingpapers.htm

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD 2014

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given.

All requests for commercial use and translation rights should be submitted to rights@oecd.org.

ABSTRACT

This paper explores methodological approaches that can be used to monitor and evaluate climate change adaptation initiatives at the projects and programme levels. It examines approaches that have been used in other areas of development practice to see what lessons have been learned that can inform the development of monitoring and evaluation frameworks targeted at adaptation. The paper focuses on three methodological challenges related to monitoring and evaluation that are particularly relevant for adaptation: i) assessing attribution, ii) establishing baselines and targets, and iii) dealing with long time horizons. The paper also considers the importance of on-going learning in evaluation and the benefit of applying a comprehensive approach to monitoring and evaluation, building on tested practices from participatory methods and social sciences techniques.

JEL Classification: H43, O22, Q54

Keywords: Climate change adaptation, monitoring and evaluation, learning.

RÉSUMÉ

Ce rapport analyse les approches méthodologiques utilisables pour suivre et évaluer les initiatives d'adaptation au changement climatique mises en œuvre au niveau des projets ou des programmes. Il examine les approches adoptées dans d'autres domaines d'action en faveur du développement afin de cerner, parmi les enseignements qui en ont été tirés, ceux qui pourraient éclairer l'élaboration de cadres de suivi et d'évaluation visant l'adaptation. Le rapport met l'accent sur trois enjeux méthodologiques du suivi et de l'évaluation qui s'avèrent particulièrement intéressants du point de vue de l'adaptation : i) évaluer l'attribution, ii) établir des niveaux de référence et des objectifs, et iii) travailler avec des horizons temporels lointains. Il aborde également l'importance que revêt l'apprentissage continu dans l'évaluation, ainsi que l'avantage que présente une approche globale du suivi et de l'évaluation, fondée sur des pratiques éprouvées qui relèvent de méthodes participatives et de techniques des sciences sociales.

Classification JEL: H43, O22, Q54

Mots clés: Adaptation au changement climatique ; suivi et évaluation ; apprentissage.

FOREWORD

This report on “Monitoring and Evaluation of Climate Change Adaptation: Methodological Approaches” is an output from the OECD Task Team on Climate Change and Development Co-operation that is overseen jointly by the Working Party on Climate, Investment and Development (WPCID) of the Environment Policy Committee (EPOC) and the Network on Environment and Development Co-operation (ENVIRONET) of the Development Assistance Committee (DAC).

This report has been written by Ayesha Dinshaw, Susannah Fisher, Heather McGrey, Neha Rai and Johan Schaar, and edited by Nicolina Lamhauge (OECD). The authors gratefully acknowledge comments provided by Simon Anderson (IIED), Abigail Ofstedahl (WRI), and Ian Christoplos (DIIS), Juan Casado-Asensio, Jan Corfee-Morlot, Anthony Cox, Jane Ellis, Megan Grace Kennedy-Chouane and Michael Mullan (OECD) throughout the process

Financial contributions from the UK Department for International Development and the Swiss Development Co-operation Agency are gratefully acknowledged.

TABLE OF CONTENTS

ABSTRACT	3
RÉSUMÉ.....	3
FOREWORD.....	4
LIST OF ACRONYMS	7
EXECUTIVE SUMMARY	8
1. INTRODUCTION.....	10
2. METHOD SELECTION FOR MONITORING AND EVALUATION OF ADAPTATION.....	11
3. MEASURING ATTRIBUTION OF ADAPTATION INITIATIVES	12
3.1 Identifying counterfactuals for monitoring and evaluating adaptation	13
3.2 How did the intervention make it happen?.....	15
4. BASELINES AND TARGETS FOR MONITORING AND EVALUATING ADAPTATION	16
4.1 Techniques for establishing baselines in complex contexts	17
4.2 Setting targets	21
5. ADDRESSING LONG TIME HORIZONS	22
5.1 Current state of long-term monitoring and evaluation	23
5.2 Importance of monitoring over time	24
5.3 Techniques for evaluating over long time horizons	24
6. MONITORING AND EVALUATION FOR ACCOUNTABILITY AND LEARNING.....	26
6.1 Participatory techniques for accountability	27
6.2 Integrating learning into monitoring and evaluation for adaptation.....	29
REFERENCES	32
ANNEX I: SUMMARY TABLE OF APPROACHES AND METHODS	37
ANNEX II: GLOSSARY OF KEY TERMS.....	39

Tables

Table 1.	Categories of relevant monitoring and evaluation methods and approaches.....	12
Table 2.	Participatory/social science techniques for monitoring and evaluating adaptation	27

Figures

Figure 1.	Illustration of single and double loop learning	30
-----------	---	----

Boxes

Box 1. Attribution: Recommendations for monitoring and evaluation.....	13
Box 2. Examples of qualitative evaluation methods used to measure attribution.....	15
Box 3. Baselines and targets: Recommendations for monitoring and evaluation of adaptation.....	17
Box 4. Performing strategic conflict analysis	21
Box 5. Long time horizons: Recommendations for monitoring and evaluation of adaptation	23
Box 6. Accountability and learning: Recommendations for monitoring and evaluation of adaptation	26

LIST OF ACRONYMS

DAC	Development Assistance Committee
DiD	Difference in Difference
DFID	Department for International Development (UK)
FTF	Feed the Future
GIS	Geographic Information Systems
LFA	Limiting Factor Analysis
MDG	Millennium Development Goals
MSC	Most Significant Change
OECD	Organisation for Economic Co-operation and Development
PPCR	Pilot Programme for Climate Resilience
PSM	Propensity Score Matching
RBM	Results Based Monitoring
TAMD	Tracking Adaptation Measuring Development - TAMD
UNDP	United National Development Programme
UNFCCC	United Nations Framework Convention on Climate Change
VAWG	Violence Against Women and Girls
WRI	World Resources Institute

EXECUTIVE SUMMARY

Countries are already affected by current climate variability and increasingly by climate change. In this context, it is important to understand which approaches to adaptation are effective in reducing climate vulnerability and enhancing society's resilience to the effects of climate change. However, the monitoring and evaluation of adaptation interventions is beset with methodological challenges given the uncertain, non-linear, and long-term nature of climate change. These challenges are not unique to adaptation, although they differ from those faced by mitigation. This paper draws on experience gained from the monitoring and evaluation of development projects and programmes to identify practical approaches to addressing these methodological challenges. There is no single solution; rather, it is necessary to combine different qualitative and quantitative methods to monitor and evaluate adaptation. This requires a good understanding of the strengths of each method as well as their underlying assumptions and limitations.

Monitoring refers to an on-going assessment of the intervention and progress made in achieving set milestones and targets. Evaluation, on the other hand, examines whether set objectives were achieved and if they did so in a cost-effective way. This paper focuses on three of the main challenges for monitoring and evaluating adaptation: i) how to assess the attribution of an intervention to adaptation, ii) how to establish baselines and targets for adaptation initiatives, and iii) how to reconcile the longer time horizons associated with climate change with the much shorter assessment cycles. Finally, the paper examines how monitoring and evaluation can address accountability and learning needs.

The **attribution** of a particular development intervention to the final outcome is often assessed through the construction of a counterfactual – i.e. what would have happened in the absence of the intervention. However, since adaptation often is a relatively small component that has been integrated into larger development initiatives, it may be more meaningful to consider the contribution of an intervention to the observed adaptation outcomes (e.g. improved water security, drought resilient crops, and more climate resilient livelihoods). A range of techniques can be used to construct a counterfactual against which the contribution of adaptation interventions can be assessed. When this is not possible, qualitative methods can contribute to a better understanding of why, and the means by which, change has come about, above and beyond simply demonstrating the change.

Appropriate **baselines** for constructing a counterfactual for adaptation can be difficult to identify when the underlying contexts are constantly changing. To overcome this challenge, techniques that have been tested in the context of conflict and fragile states include the reconstruction of baselines and the use of shifting or rolling baselines. A challenge in applying some of these methods is the capacity required to generate the complex computational modelling and to interpret the information. Further, the application of shifting baselines may cause a bias where programme inputs (independent variables) and outputs and outcomes (dependent variables) become correlated with various other factors that may not be controlled or predicted in the original impact assessment model. In this case, there are techniques that can enable the evaluator to separate and quantify the impact of different influencing factors on the final outcome. Further, sequential **targeting** can also help a programme adapt to changing contexts and to ensure that set targets are realistic and can contribute to intended impacts.

The effects of adaptation initiatives will in many cases only become apparent over a **long time horizon** (say 20-50 years) as the effects of climate change are experienced and the efficacy of adaptation can be measured. Yet governments and providers of development co-operation require reporting on the

effectiveness of such initiatives on a much shorter timescale. It is therefore important that the nature and time scales of monitoring and evaluation approaches are adjusted accordingly by for example focusing on intermediate outputs and outcomes in the short term. Although the issue of long time horizons is not unique to the context of adaptation, examples of truly long term evaluations (say 50-100 years after project or programme completion) are rare for development programmes. This can in part be explained by the lack of funding available to do an evaluation long after the end of an intervention. It may also be that an evaluation that comes many years later – when full impacts are seen – is too late to support learning from the experience. Although some lessons may be learned by for example examining approaches used in multi-decadal research, the fact is that long time horizons are intricately entwined with other factors (e.g. population growth, spatial development and urbanisation) that make climate change adaptation complex.

With an increase in the international resources allocated for adaptation, transparency in the allocation of resources is important. At the same time, recipients of climate finance need to be **accountable** to their funders. An approach to improving accountability is to use participatory processes by which stakeholders can be involved in the planning, implementation, monitoring and evaluation of adaptation initiatives. Accountability to beneficiaries and to providers of climate finance can also be improved by integrating **learning** into any programme design, so that a flexible, iterative approach allows periodic reflection and improvement. To ensure accountability and continuous learning, monitoring and evaluation must become an integral component of the project or programme design.

The methods discussed in this paper can start to address the monitoring and evaluation challenges identified; yet none of them is a “silver bullet”. Rather, methods need to be brought together in a way that addresses the specific context and intervention challenges posed by an intervention. One approach is to use evaluation practices such as developmental evaluation and participatory, social science techniques, centred around participation and learning. Only by experimenting with these approaches, documenting lessons learned, and ensuring dissemination of this knowledge can a meaningful evidence base be developed of effective adaptation interventions.

1. INTRODUCTION

1. Monitoring and evaluation are different processes, which work together to assess the performance of an intervention over time. Monitoring refers to an on-going assessment of the intervention and progress made in achieving set milestones and targets. Evaluation, on the other hand, examines if people have become better adapted to climate change as a result of the interventions, and the extent to which they are now more resilient to climate change. Characteristics of climate change, such as uncertainty, non-linearity of climate change patterns, and long-time horizons, pose challenges for monitoring and evaluating. To contend with these characteristics and to ensure continuous learning, adaptation programme design, management and evaluation methods must be open, flexible and adapted to the changing context. To determine the most appropriate methods for monitoring and evaluating adaptation, it is useful to distinguish between interventions that have a simple, complicated or complex design (Quinn Patton, 2011).

2. Simple adaptation initiatives follow a straightforward logic where the causal relationship between inputs, outputs and outcomes is clear. For such interventions, well-established monitoring and evaluation methodologies from development practice may be applied. More complicated adaptation initiatives, that for example entail multiple components or stakeholders over long time horizons, may require a broader portfolio of approaches that assess different aspects of the intervention to determine its overall impact (Rogers, 2008). Many adaptation initiatives are complex. That means that there are fundamental uncertainties about the causal relationship between inputs and outcomes, a causal chain may only become apparent after a climate event, or there may not be an end state at which point the problems have been resolved (Rogers, 2008). In complicated and complex contexts many standard methods for monitoring and evaluation may fail. To overcome this challenge, it can be helpful to use different methods in an integrated framework, based on multiple theories of change that capture each element.

3. Numerous theoretical frameworks for monitoring and evaluating adaptation have emerged in recent years (Ayers et al., 2012; Brooks et al., 2011; 2012; Pringle, 2011; PROVIA, 2013; Spearman and McGray 2011; Villanueva, 2011). These theoretical approaches have been complemented by monitoring and evaluation frameworks developed by dedicated climate funds to assess the impact of their portfolio of activities (e.g. the Adaptation Fund, the Global Environmental Facility (GEF), and the Pilot Programme for Climate Resilience (PPCR) under the Climate Investment Funds (CIF)). Finally, a number of developed and developing countries are also in the process of developing or implementing national monitoring and evaluation frameworks to examine changes in the countries' vulnerabilities to climate change (Lamhauge, forthcoming).

4. Across the different scales, there are a number of common challenges. These include the challenge of attributing observed change to specific activities, setting baselines and targets, and the challenge of assessing the effectiveness of adaptation initiatives with long-term benefits within short- and medium-term evaluation cycles (Lamhauge, Lanzi and Agrawala, 2012). None of these challenges, however, are unique to climate change adaptation, although they differ from those faced by mitigation. This paper explores what lessons the adaptation community can learn from the more established practice of monitoring and evaluating development initiatives. Focusing on the project and programme levels, the paper further aims to examine how existing methodologies and approaches can help to address learning and accountability needs in the field of climate change adaptation.

5. The next section briefly examines how to identify which methods to consider for monitoring and evaluating adaptation. This is followed by three sections that provide options for addressing the specific

methodological challenges. The last section examines how monitoring and evaluation can contribute to accountability and learning needs:

- Section 3. Measuring attributing of adaptation interventions;
- Section 4. Setting baselines and targets for monitoring and evaluating adaptation;
- Section 5. Dealing with long time horizons;
- Section 6. Addressing accountability and learning.

6. The paper provides a range of techniques and examples. These have been selected from a number of sectors that each contributes some insight to addressing the identified challenges. This is not an exhaustive list. Instead, the paper introduces some relevant approaches where published evidence is available of their use in practice. In some cases, the principles underlying the techniques are outlined rather than the details of the technique itself. Further, the applicability of the specific examples, both as stand-alone techniques and in conjunction with each other, will be highly context specific. Annex I highlights which examples are useful across simple, complicated and complex adaptation interventions, as well as which methodological challenges and monitoring and evaluation needs they may help in addressing.

2. METHOD SELECTION FOR MONITORING AND EVALUATION OF ADAPTATION

7. Monitoring and evaluation are different processes, which work together to assess the performance of an intervention over time. Many of the approaches discussed in this paper, however, break down this simple division between monitoring and evaluation. Some techniques incorporate the monitoring of impacts throughout the process, whilst others recommend a continuous iterative evaluation process embedded in the design of the programme. Techniques discussed in this paper then, may apply to either monitoring or evaluation or both depending on the approach taken. Examples are provided in Table 1 with short definitions outlined in Annex 2.

8. All methods have strengths and weaknesses, but they can complement each other for greater clarity. For example, statistical analysis can demonstrate co-variation and dependence between variables but has little to say about why a certain intervention leads to a specific result. Evaluation designs that use different methods offer opportunities for triangulation and complementarity and thus a deeper understanding and conclusions about causality. When applying different methods, however, a careful assessment of the underlying assumptions is required to ensure that the evidence gained from each approach is compatible. For example, to monitor and evaluate an intervention that provided monetary incentives and training to farmers with the aim of increasing their resilience to climate change, a portfolio of methodologies was used. This included quantitative methods (e.g. experiment-related methods, econometrics and statistical analyses) that enumerated outcomes, such as income levels, agricultural productivity or access to services. At the same time, qualitative methods examined how change had come about using surveys, focus group interviews, and participatory techniques. Each methodology had its comparative advantage that provided a more complete picture of the intervention outcome when brought together (Leeuw & Vaessen, 2009).

Table 1. Categories of relevant monitoring and evaluation methods and approaches

Monitoring and evaluation approaches	Examples found in this paper
Overall monitoring and evaluation approaches	Developmental evaluation, longitudinal evaluation, impact evaluation, institutionalised learning
Formal social science methods	Surveys, focus groups, interviews
Econometrics/statistics	Modelling, statistical analyses, stochastic baseline, deterministic baselines, normalisation
Experiment-related methods	Case studies, experimental design, quasi-experimental design, propensity score matching, phased pipeline, purposeful sampling, regression analysis
Participatory methods	Most significant change analysis, beneficiary monitoring, limiting factor analysis, outcome mapping, recall techniques
Iterative methods	Sequential targeting, results based monitoring, theories of change, stepwise approach, contribution analysis, scenario building, rolling baselines, reconstructing baselines

3. MEASURING ATTRIBUTION OF ADAPTATION INITIATIVES

9. To understand whether an intervention has a direct relationship to a specific outcome, impact evaluations are usually applied. What exactly this entails has long been the subject of debate within the evaluation community. Stern et al., (2012) suggest that impact evaluations should:

- evaluate positive and negative, intended and unintended long-term effects on final beneficiaries that result from a development intervention;
- assess the direct and indirect causal contribution claims of the intervention;
- explain how the intervention leads to an effect so that lessons can be learned.

10. Others claim that only through randomised experimental designs similar to those used in clinical trials and biomedical research can the attribution question be resolved (Banerjee & Duflo, 2011). Some counter that such quantitative methods are too narrow, ignore the complex development context that always influences outcomes, and say little about whether an intervention can be successfully replicated beyond the experimental site (see e.g. Ravallion, 2008; Picciotto, 2013; Quinn Patton, 2011; Woolcock, 2009).

11. A challenge when assessing the attribution of adaptation initiatives is the need to understand the counterfactual – what would have happened in the absence of the intervention? This section explores different monitoring and evaluation techniques that can help to: i) establish counterfactuals for monitoring and evaluating adaptation, ii) gauge the attribution of an adaptation intervention when experimental or

quasi-experimental methods cannot be applied, and iii) apply theories of change and how they can be used to articulate assumptions. Examples are provided from the fields of health services, food security, land use change, and peace-building. Recommendations emerging from this section are summarised in Box 1.

Box 1. Attribution: Recommendations for monitoring and evaluation

- When counterfactuals are not available, they can in some cases be inferred through quasi-experimental or participatory methods. But there will also be complex contexts where the nature of climate change is not well understood and it therefore is difficult to establish a counterfactual. In such cases programme designs that map system dynamics and emergent change, while looking for patterns that can be identified through retrospective observations, may be useful for learning.
- In complex contexts it may in some cases be more meaningful to examine the contribution of an intervention to the observed outcome rather than to look for a direct causal attribution.
- To know whether an adaptation intervention can be replicated and scaled up we must understand the mechanisms that leads to the outcome – i.e. the theory of change. Through theories of change, assumptions are made clear from the beginning, and mid-term objectives can feed into long-term goals.

3.1 Identifying counterfactuals for monitoring and evaluating adaptation

12. Investigating impact naturally leads to the counterfactual question: what would have happened in the absence of the intervention? In a strict experimental design, with treatment and control, the question can be answered with great precision for that particular situation, at that time and for that target group. When an experimental design cannot be used, as is often the case with adaptation initiatives given their complicated or complex design, there are a number of quasi-experimental and qualitative methods that can help answer this question. These are particularly useful for adaptation as they offer ways of controlling for external changes by looking at changes in populations both experiencing the external context, but only one of them benefitting from the adaptation intervention.

13. For example, in a project that aims to improve market access for people living in District A, a technique called propensity score matching (PSM) can be used. The evaluators can with this technique construct a comparison group consisting of people living in District B that has similar socio-economic characteristics (e.g. distance to the market, livelihood choices, and social organisation) to District A. Alternatively, a pipeline approach can be used for programmes that apply a phased or staggered implementation design with time lapses between the completion of one component and starting of another. For instance, a pipeline approach was used in the evaluation of a microfinance programme conducted by the United Nations Capital Development Fund (UNCDF) in Nigeria, Malawi, Kenya and Haiti. The evaluation used new clients, defined as those who had not yet received their first loan or those who had received their first loan but had not yet completed a full loan cycle, as the control group, while older clients, defined by those who had been in the programme for at least twenty months, were the treatment group. The treatment group was further disaggregated by location, economic activity, and impact, which enabled evaluators to make more detailed recommendations (White et al., n.d.).

14. In cases where an entire population is affected by an intervention, for example through changes in pricing policies, individuals will not necessarily be affected the same way. For example, poorer households tend to use a larger portion of their income on food than richer households. An increase in food prices will therefore have a much stronger relative effect on their disposable income, which is in turn likely to lead to changes in the composition of their food basket. Such changes can be analysed through statistical regression analysis, where the dependent variable (in this case the proportion of a specific food item or protein content in the food basket) is plotted against an independent variable (e.g. the proportion of the

household budget allocated for food). If there is a close correlation between the two variables, the counterfactual can reflect the extrapolated value of the outcome variable, reflecting the situation when there are no changes in the proportion of the household budget allocated for food (Kandker et al., 2010). In multiple regression analysis the relationship between several independent variables and the outcome variable can be analysed simultaneously. For example, in the case of drinking water facilities, the capacity of the facilities may differ as well as the circumstances in which they are located and the socio-economic circumstances of beneficiaries. While, regression analyses and other quantitative methods can state the mathematical relationship between treatment and control groups, they are often not able to explain how the changes came about.

15. A combination of qualitative and quantitative methods can shed light on the mechanism whereby changes have come about. A technique commonly used in rural development is to ask individuals or communities about the most significant changes they have experienced during the duration of the intervention (IFAD, 2002). If people or communities who have not been subject to the intervention are asked the same questions, they can serve as the counterfactual. In areas where literacy levels are low, this approach can use graphical representation and other tools to illustrate the change. This approach also provides a good understanding of other factors, not directly linked to the intervention, but that nonetheless affect the outcome. Such participatory approaches provide greater contextual depth to the assessment. Further, they may shed light on the likelihood of replicating the intervention in other contexts.

16. Similar information may be derived from semi-structured interviews and focus group discussions that explore changes in attitudes and decision-making processes. The examples summarised in Box 2 illustrate the importance of qualitative and participatory methods in identifying individual and community learning from autonomous actions as well as from external interventions. Without this qualitative information, it is difficult to ensure accountability towards those whose well-being adaptation is meant to enhance.

17. The long period between an intervention and the projected impacts of climate change is a problem many evaluators face when identifying a suitable counterfactual for assessing the impact of adaptation interventions. In some cases, causality simply cannot be identified due to the complex interaction between multiple factors. These factors may only be partly observed and understood. Particularly for longer term interventions with several components, the most suitable approach may be to examine causality through programme designs that map system dynamics and emergent change, while looking for patterns that can be identified through retrospective observations (Quinn Patton, 2011).

18. As described above, attributing observed changes to specific interventions is challenging when interventions are complex or when the adaptation component is a relatively small part of a larger development programme. Other causal factors that may affect the outcome include changes in the policy environment, in markets, external shocks and crises, and other development co-operation interventions. A contribution perspective focuses less on programme interventions as single causes and more on how an intervention interacts with other factors. To overcome this challenge, it can be helpful to use a stepwise approach that examines each link of the results chain in turn. This is, for example, the approach that will be used to monitor and evaluate health systems strengthening (HSS) initiatives (WHO et al., 2010). The objective is to infer causality between trends in health outcomes, coverage and risk behaviours, access and quality of services, and funding. The first step will be to examine the logic that increased resources result in improved quality of and access to health services, which in turn leads to enhanced exposure to interventions. The next step will assess the impact of increased coverage, which depends on the efficacy and effectiveness of interventions, and on contextual factors, such as epidemiology, economic change and political stability. The final step will relate the findings to the overall health impact of the programme, measured in morbidity and mortality rates.

Box 2. Examples of qualitative evaluation methods used to measure attribution

The Office of Evaluation of the International Fund for Agricultural Development (IFAD) evaluated in 2004-2006 rural development projects in The Gambia, Ghana and Morocco (Leeuw and Vaessen 2009). The objective of the evaluations was to set a basis for accountability and to contribute insights for future evaluation design. Given the objectives of the evaluation, it could not simply assess impact; findings needed to be sufficiently grounded in contextual factors and local experience to improve the design and operationalisation of programmes. The methods used included desk reviews, quantitative surveys and focus group discussions with project beneficiaries, control groups, and key informants. Through the use of a combination of different monitoring and evaluation methods, it was possible to cross-check the validity of findings and to identify discrepancies. This sort of triangulation of evidence is especially important when information is scarce.

The global food and finance crises provide an additional example of how qualitative evaluation methods can be applied on a large scale. The 2008 crises caused a gradual deterioration of income and livelihood security in many developing countries. Heltberg et al. (2012) have created a crisis narrative from the perspective of people affected in 17 countries. The narrative included assessments of a range of relief actions and social protection schemes that often are part of, or proposed for, adaptation strategies. For instance, local teams created “listening posts” that allowed communities to describe the impacts of the global crisis on them, in addition to an array of qualitative methodologies tailored to specific country contexts. This provided a better understanding of common trends identified across the countries, as well as the nature of people’s vulnerability and the sources of their resilience.

19. A similar stepwise approach has been used to evaluate large scale interventions such as multi-country health programmes. It has also been proposed for monitoring and evaluating short and long term changes related to health reforms in China. It is important to note that evaluation at this scale requires a solid monitoring system with data on baseline trends for key indicators, in-depth qualitative and quantitative studies, and cost-effectiveness analysis (WHO et al., 2010). When applying a stepwise approach to adaptation initiatives, the importance of on-going monitoring is particularly important for interventions that are implemented over long periods of time and where it can be difficult to secure the financial commitment. Furthermore, it is important to identify criteria of success. Since adaptation is context-specific, and often based on risks that will be mitigated in the future, defining success is not always easy.

20. A contribution perspective also looks at the extent to which, and how, an intervention registers and aligns with observed changes (O’Flynn, 2010). The success of an intervention may lie precisely in its ability to anticipate and take external causal factors into account, and work with them, rather than be surprised by them. A useful way of approaching the attribution versus contribution question is to analyse whether an intervention was a necessary and sufficient causal factor to bring about the desired change. Such contribution analysis may show that in certain cases there are multiple pathways for reaching the desired outcome. After implementation, case studies, correlation studies or contribution analysis can shed light on changes that the intervention may have contributed to.

3.2 How did the intervention make it happen?

21. It is not sufficient to know that an intervention produced a certain result in a specific context if we are interested in replication and scaling up. Would the intervention have delivered the same result to a similar target group elsewhere? To answer this question we need to know how the result came about. This requires an understanding of the theory underlying the intervention, the assumptions that underpin the nature of the intervention, expected interactions with the context, and the impact of external factors on the final outcome. Many complex programmes do not explicitly lay out the underlying assumptions or theories behind their approach, making them difficult to evaluate (Weiss, 1995). As a result, it is not always clear to the stakeholders involved how the intended change will come about. In turn, there is a risk that important intermediate milestones needed to achieve the long-term goal are overlooked.

22. A theory of change entails a critical reflection on how a programme intervention achieves change, the multiple pathways to such change, and the determining factors that can lead to some pathways being chosen over others. This approach encourages critical reflection and supports innovation and flexibility when responding to dynamic contexts. It can support an iterative process of learning during implementation and can be used as a tool to address uncertainty and learning. Theories of change are used in a range of sectors and at different points in the evaluation process. For example, theories of change can be an integral part of programme design and theory, as well as part of a sectoral analysis or national assessment. Theories of change have been used across international development (for reviews see James 2011; Vogel 2012) including complex situations such as fragile states and peace-building (see Government of Denmark, 2013). Establishing local theories of change can include community participation to explore what impact an intervention has had on their livelihoods. If this practice is continued throughout an intervention evaluators can create and test narratives of attribution (Brooks et al., 2013). These can complement monitoring and evaluation design to strengthen the programme assumptions and theory of change that underpin the programme logic.

23. When developing theories of change for adaptation initiatives, a useful first step is to conduct a context analysis focused on the climate vulnerabilities and risks. This can help to shed light on the broader context in which the intervention is implemented, highlighting factors that are driving the problem the intervention aims to address. This analysis can, in the second step, inform the nature of the adaptation intervention. In the third step, a “results hierarchy” is developed that captures all the results the intervention intends to have, beginning with those that are directly under the control of the project and ending with higher-level results dependent on factors that are not under the control of the project. The hierarchy of results will inform a corresponding theory of change, which will help explain how lower-level results feed into higher-level results. This systematic development of a theory of change ensures assumptions are made clear and it improves the likelihood of a successful intervention.

4. BASELINES AND TARGETS FOR MONITORING AND EVALUATING ADAPTATION

24. Carefully designed baselines are necessary to measure performance of adaptation actions. Similarly, targets are required to evaluate whether the intended results have been achieved within the planned timeframe. However, establishing baselines and targets for climate change adaptation can be challenging. This can in part be attributed to the uncertain nature of climate change, making it difficult for programme designers to plan long term outcomes. Gaps in climate change information systems pose an additional challenge.

25. To get a better understanding of approaches that have been applied to overcome similar challenges in other fields, this section draws on examples from peace-building, food and agriculture, cash transfer schemes, Millennium Development Goals (MDGs), and health and nutrition assessments. The remainder of this section explores different monitoring and evaluation techniques that can help to: i) establish baselines in complex contexts where the relevant information may not be readily available, the external environment is shifting, and hazards are irregular, and ii) identify targets that can help to measure the performance of climate adaptation. A few recommendations emerging from this section are summarised in Box 3.

Box 3. Baselines and targets: Recommendations for monitoring and evaluation of adaptation

- In the absence of relevant climate change data, projection techniques and trend extrapolation methods can be used to estimate baselines from past and current trends.
- Comparisons in the context of a shifting baseline for adaptation need to take account of the changing external context and the interaction with project outcomes. There are a range of existing techniques to do this including normalisation and “contextualisation”.
- Target setting is a challenge when tracking long term results and where the exact links between immediate outputs and longer term outcomes and impacts are not fully understood. Monitoring and evaluation of adaptation may factor in sequential short term targeting to measure long term performance and use historical data for scenario-based target setting

4.1 Techniques for establishing baselines in complex contexts

26. Baselines serve two main purposes for monitoring and evaluation: they help to set targets, and they provide a reference point against which performance can be measured (Prag and Clapp, 2011). In a standard evaluation system, baseline data are compared with the same indicator before, during and after implementation to assess how conditions have changed. However, there are three challenges in doing this for adaptation (Brooks et al., 2011):

- **Data availability:** Setting baselines for monitoring and evaluation of adaptation interventions can be a challenge when climate specific information is scattered across different departments, sectors or projects. The cross-sectoral nature of climate change often means that the relevant information may rest with different cross-sectoral ministries and implementers of adaptation interventions. Although some climate-related data can be collected through primary methods, it can be resource intensive to do so at the national level.
- **Complex contexts:** Given the uncertain nature of climate change, the specific, desired long term outcomes of interventions may not be clear. Further, the links between immediate outputs and outcomes and broader impacts, may not be supported with strong evidence. As a result, setting targets is not always straightforward. This can pose a challenge even if baselines are established in the first place. Depending on the nature of climate change, the information that evaluators may wish to evaluate may not necessarily be the same as that considered to be particularly important at the outset. In some cases, this may require the reconstruction of more appropriate baselines at a later stage.
- **Shifting baselines:** Measuring adaptation outcomes against a set baseline can be challenging given the long time-scales of climate change. The changing contexts may require that the baseline is revised to provide a more accurate basis for comparison between what would have happened without the intervention and what actually happened (counterfactual).

27. Practical strategies for estimating or projecting baselines include reconstruction techniques, normalisation, rolling baselines, and the use of a wide range of indicators, including indicators monitoring the changes in the context in which the intervention takes place in addition to project or programme components. Each of these are briefly discussed below.

Estimating or projecting baselines

28. Climate projections can contribute to the establishment of baselines for adaptation interventions. The methods used to project baselines can be simple or complicated depending on the models used. For example, deterministic and stochastic models are useful tools for projecting baselines but they can be fairly complicated to construct. In deterministic models, variables account for the different states observed, and are determined by the parameters in the model. An example is the use of trend extrapolation to provide a single set of projections based on assumptions of average conditions or business as usual conditions. Stochastic models, on the other hand, have an element of randomness to them, and different states are captured by probability distributions rather than by using different sets of variables. Stochastic models can therefore be useful in complex contexts that have shifting baselines as it can consider several alternative futures or scenarios.

29. For example, the Food and Agriculture Policy Research Institute (FAPRI), focusing on the agricultural sector in the US, used a deterministic model to create annual 10 year baseline projections for assessing several domestic and trade policy aspects. The baseline projections were based on assumptions of business as usual conditions. However, this approach only provided a single outlook and was considered inadequate (Blanco-Fonseca, 2010). A stochastic model was therefore used to estimate a more comprehensive baseline that took into consideration 500 different scenarios that vary in the underlying assumptions about factors such as climate, supply and demand, and so on. This approach enabled the model to take into consideration the uncertainties that influence agricultural market developments. Similarly, in the context of adaptation, a stochastic model could help to establish baselines since the approach is based on the premise of uncertain and complex contexts where data availability is an additional barrier. However, the main challenge in using projection methods is the complex computational modelling required and the capacity required to generate and interpret the information.

Reconstructing baselines

30. Evaluators can also draw on recall and mapping techniques to reconstruct baselines for adaptation initiatives when baseline data have not been collected properly ex-ante. Recall techniques ask participants to remember their socio-economic situation and access to services over a certain time period. Mapping techniques, on the other hand, may involve historical Geographic Information System (GIS) maps or participatory mapping with communities. These techniques are commonly used for health, nutrition and social infrastructure interventions, where a range of secondary data sources (e.g. socio-economic surveys and census surveys) are used to construct information about the intervention context, comparison groups, and the condition of the target population prior to programme implementation. This information can also provide reference data for estimating baselines for programme participants and comparison groups. These methods do have limitations in their accuracy and the timeframes for which they can be used, and should where possible be triangulated with other pieces of data.

31. Impact evaluations that use an experimental design require baselines to be created for the treatment group and the control group. In Section 3.1 propensity score matching (PSM) was briefly introduced as a technique used to construct a matched comparison group. Matching reduces the dissimilarities between target and control groups and enhances the reliability of an estimated treatment effect (Rosenbaum and Rubin, 1983). PSM is commonly used when evaluating health and nutrition initiatives. An example is WHO's Integrated Nutrition and Health Project Programme (INHP). The WHO programme used an evaluation technique called Difference in Difference (DiD), which compares impacts between treatment and control groups before and after the implementation of an intervention. In this case, direct comparison between the groups was not possible because the groups were too dissimilar. To overcome the challenge, PSM was used to establish a comparison group which had the same propensity to receive the benefits of the intervention as the target group (Gakhar et al., 2010).

32. Similarly, the Emergency Social Investment Fund of Nicaragua used in 1998 the World Bank's Living Standard Measurement Study (LSMS) data to estimate baselines for project and comparison groups in water and sanitation, health and education projects. PSM was used to enhance comparisons between the two groups (Pradhan and Rawlings, 2002; World Bank, 2010). By providing statistically equivalent baseline comparisons, these methods provided one way of estimating the treatment effects. Even statistically robust techniques such as these have limitations in terms of the assumptions used to create comparison groups and make matches. However, they provide a source of data on how a programme or intervention is performing relative to a reference population. In the context of adaptation this can provide a useful comparator where baselines may be missing or populations are experiencing similar changes to hazards over time.

The use of standardisation and normalisation techniques for monitoring and evaluation

33. Normalisation is a technique commonly used when assessing the impact of nutrition care programmes on health and disease outcomes. In nutrition care programmes, the health and disease outcomes may be influenced by a number of contextual factors in addition to the programme components. To evaluate the impact of such interventions, the contribution of the intervention on health outcomes must be adjusted or standardised against their average or usual trend to make the unequal population more comparable. To standardise the findings, the health and disease outcomes are commonly assessed as deviations from expected impacts under "usual" conditions, defined by data on height, weight and other health related aspects. This standardisation technique enables the evaluator to assess the health outcomes by comparing counts against unequal population areas (Spahn, 2008).

34. A similar approach has been used to evaluate structural change and competitiveness of EU member states. The assessment centred on understanding how aggregate output variations over the business cycle (e.g. idiosyncratic demand and productivity shocks) affect employment, growth and the industrial value added in EU countries. The evaluation used "normalised metrics" to reduce the bias caused by differences in country context. For example the metric on "market share of the industry" was shown in relation to the market share of the country under review. The metrics for output variations on the business cycle – such as idiosyncratic demand and productivity shocks – were also normalised (Janger et al., 2011). This standardisation of metrics allowed for a comparison of the scale of impact on industry performance in different country contexts.

35. In the context of adaptation, shifting climate baselines can lead to endogeneity bias, where the independent variables (programme inputs) and the dependent variable (outputs and outcomes) become correlated with various other factors that are not included in the original impact assessment model. To overcome this challenge, normalisation techniques can be used to separate and quantify the impact of different influencing factors. This may require the inclusion of additional variables that originally were excluded or the application of an instrumental variable (alternative variable) in the regression model. This approach was used to evaluate the impact of the Nicaraguan conditional cash transfer programme, *Red de Protección Social*, on changes in household expenditure on food, improved health care and the nutritional status of children. The assessment model did not consider the individual or household decision to participate as a key explanatory determinant. In the absence of this important variable (i.e. decision to participate), the evaluator attempted to endogenise the "participation decision" by using the "purposive placement of a household within a programme" as an instrument. This provided additional information of the statistical relationship between programme participation and the examined social variables despite the relationship between the various variables on the impact indicators not being fully understood. In the context of climate change where clarity around long term outcomes is lacking, and the choice of appropriate indicators and variables is not always evident, this technique allows evaluators to use robust instruments to provide evidence on causality which would have been difficult to derive from the original baseline.

Monitoring action and interaction

36. Adaptation projects and programmes often take place in complex environments and systems. Changes in contextual factors can prevent an intervention's objectives from being achieved. To avoid this, it is important to monitor contextual indicators and adjust programme implementation in response to this information. This requires an adjustment of the evaluation logic and the programme baseline against which outcomes are monitored and assessed. One approach to such "contextualisation" is to identify a baseline for the context as well as for the intervention, and to monitor both contextual and programme indicators. Several techniques have been developed that monitor changes in complex systems over time to take account of these factors. One of these is using "sentinel indicators" that monitor the risk and complexity of a system and its dynamics (Hargreaves 2010). These indicators are often part of the logical framework or theory of change.

37. When developing "contextualisation" approaches for adaptation, valuable lessons can be learned from projects and programmes in conflict areas. For both sets of initiatives, projects or programmes may be situated in complex contexts where hazards and extreme events can significantly impact outcomes. This may result in shifting baselines and the challenge of attributing programme components to specific outcomes. Using scenarios and likelihood assessments, this approach enables evaluators to understand the contextual dimensions that may interact with the project over its lifetime. For example, programmes that focus on reducing conflicts may use systematic conflict analysis to identify contextual factors that will interact with programme implementation. A good understanding of relevant contextual variables can also inform possible adjustments of the evaluation focus (indicators, baseline) over the lifetime of the programme. There are different methodologies for conducting such an analysis. Three core areas highlighted by Sida (2006) are: i) conflict analysis of drivers and power dynamics, ii) scenario planning, and iii) definition of strategies and options (see Box 4). Although this is specific to conflict, the example highlights how systematic analysis of broader contexts can be done and incorporated into monitoring and evaluation. An OECD study, however, finds that even in areas of conflict, the operationalisation of such "contextualisation" techniques is challenging (OECD 2007).

38. Another example of how contextual monitoring can be operationalised is demonstrated by DFID. When working in conflict situations DFID monitors a range of contextual indicators to support programme delivery. In Nepal, for instance, where DFID provided budget support and contributed to the implementation of sector programmes, qualitative data from a range of sources were used to monitor a set of indicators that were likely to influence programme outcomes. This included indicators on communal violence, human rights abuses, rule of law, and the role of representatives of marginalised groups in political institutions (DFID, 2010). Such approaches are also applicable to adaptation since they explicitly and systematically identify contextual factors that interact with the programme, allowing implementation to continuously adjust and change direction and modalities, ensuring that they remain relevant and also potentially influence the context.

39. A wider set of indicators can be used to monitor longer term social and economic trends. By 2012, and estimated 87% of the global population had been enumerated in the 2010 census round (Balk et al., 2013). This constitutes a potentially rich source of adaptation-relevant data. In many countries censuses include information on people's residence and livelihoods, on household expenditure, farmers' selection of crops, and on nutrition, health and the occurrence of epidemics, as well as school enrolment and drop-out rates. Since censuses are usually carried out at relatively long time intervals and sometimes at a highly aggregated administrative level, they need to be complemented with other types of surveys to inform policy development (Guzman et al., 2013). Data that can be collected and combined at different levels using a step-wise approach include: i) human and social components of adaptation, such as demographic, migration, poverty, asset, tenure, occupation, employment and literacy data; ii) hazard exposure, such as flood and heat wave vulnerability, declining agricultural output and local deforestation; iii) cross-tabulation

of a) and b) with data on female-headed households, elderly-only or adolescent-headed households and migrant-sending and receiving households; iv) integration of environment or climate-specific questionnaires linked to a census, and v) using results to feed back to geography-based policy (Guzman et al., 2013). This however is very resource and time intensive and will therefore depend on countries' resource availability and capacity.

Box 4. Performing strategic conflict analysis

A strategic conflict analysis is a systematic approach that is used to identify contextual factors that interact with programme implementation in a conflict setting. It consists of three steps:

Conflict analysis:

- **Structures:** Identifies key sources of tension and root causes of the conflict as well as factors contributing to peace, such as economic development, equity of distribution, functioning of the political system, respect for human rights and access to natural resources, as well as the ability of institutions to address grievances, prevent opportunistic behaviour, and further interests of confrontational actors.
- **Actors:** Identifies actors at different levels and how they influence and are influenced by the conflict through mapping of interests, attitudes, incentives, culture and relationships of power.
- **Conflict dynamics:** Identifies events, actions and decisions that can serve as trigger factors for conflict, such as rapid economic decline, changes in state cohesion, shifts in control of the central authority and distribution of power.

Scenario analysis: based on the conflict analysis a set of possible scenarios are identified that will influence the evolution of the conflict. The purpose is not to predict but to think through what causes of action and contingencies the programme could select depending on different futures.

Strategies and options: that can positively influence the conflict are then identified, given the donor's comparative advantages and the role of other actors. These are usually actions that promote dialogue, security or structural stability. The latter may include control of arms or programmes that enhance economic development.

Source: Sida (2006), *Manual for peace and conflict analysis: Methods document*, Division for Peace and Security through Development Cooperation. Swedish International Development Cooperation Agency, Stockholm.

4.2 Setting targets

40. Setting targets is another challenging area for climate change adaptation, particularly where the desired final outcomes may not always be clear, given the uncertainty of climate change. When setting targets for adaptation it is therefore important to keep in mind that the ultimate outcomes may be complex and only achieved in the long term. Though intermediate outputs may be easier to identify and measure, it is important to test the links between immediate results and actual achievement of longer-term goals.

41. Sequential targeting is particularly relevant to adaptation, as it involves regular appraisal of goals in a changing context. An example of sequential targeting can be found in DFID programmes designed to reduce Violence against Women and Girls (VAWG). The programmes aim to change social norms and ultimately prevent violence against women in a number of developing countries. However, long term targets to achieve desired change in social norms are difficult to predict. VAWG evaluations therefore focus on sequential targeting whereby interim targets reflect short term milestones and objectives considered important in order to achieve the final outcome (in this case the reduction of violence) (DFID, 2012). For example a VAWG project in Bihar, India, uses short-term targets to assess changes in awareness or knowledge; medium-term targets for changes in attitudes or behaviour, shifts in power and

influence towards women, more support to victims, institutional strengthening; and long-term targets on changing social norms and reduced violence against women and girls. Although this was not specified in the DFID programme, an evaluation could subsequently be used to test the causal links between these different steps in the results chain (e.g. do changes in attitudes actually result in reduced violence), to build the evidence base on how best to achieve necessary intermediate goals, and the cost effectiveness of alternative ways of reaching the overall objective.

42. Scenario based targeting is an alternative approach that can be used when outcomes are unclear and uncertain. It requires a good understanding of the baseline, as well as knowledge of past performance. This approach is used by the US government's Feed the Future (FtF) programme, which seeks to reduce global poverty and hunger with sustainable development impacts (USAID 2012). A tool was developed to facilitate target setting by using a series of national data to run scenarios and to set targets. National data and trends in poverty reduction, GDP growth and nutritional status were used to set target rates for indicators of poverty, nutritional status and stunting. In the context of adaptation, a good understanding of projected climate risks can help identify targets that factor in different scenarios. However, such quantitative techniques often require substantial computational skills and specialised expertise that are not always available. Even determining whether these techniques may yield meaningful findings in a complex vulnerability situation may require data and expert advice. Meanwhile, development and maintenance of knowledge and data systems to support these methods is expensive and can be difficult to sustain.

5. ADDRESSING LONG TIME HORIZONS

43. In many areas of development, investments are made with long term intentions. Building a national education system, developing the institutions and governance of a democratic state, addressing gender inequality – these are all areas where intended outcomes will only emerge in the future. However, unintended results may be early and frequent, for example when an autocratic system releases tensions that had been suppressed, showing that development trajectories are more often non-linear than linear. Within the context of adaptation there are two facets of contending with long time horizons. One is that complex adaptation interventions can take a long time to be implemented (e.g. ecosystem-based adaptation where it takes time for trees to grow or water to refill underground reservoirs). The other is that adaptation results can only be judged as climate change impacts become clearer over time (e.g. farmers with access to drought-resilient crops managing to sustain yields despite increasing incidents of droughts). This underscores the importance of sustained funding and implementation of monitoring systems in long-term adaptation interventions.

44. This section provides an overview of the current state of long-term monitoring and evaluation in other development fields. Further, it explores the importance of long term monitoring in order to obtain adequate data and establish baselines. Finally, it examines different monitoring and evaluation techniques that can help to evaluate change over long time horizons, including modelling, statistical analysis and limiting factor analysis. Key recommendations emerging in this section are summarised in Box 5.

Box 5. Long time horizons: Recommendations for monitoring and evaluation of adaptation

- Since attribution for adaptation monitoring and evaluation becomes even more of a challenge when programmes are implemented over many years, assessing contribution to a general direction of change may be more feasible and useful than establishing direct causal attribution.
- Monitoring and periodic collection of data over the long-term duration of a project is necessary to determine success, and funding should be made available for this aspect of adaptation monitoring and evaluation.
- There is no one right moment for evaluating long term interventions and intermediate solutions and forward planning are needed to both support ongoing learning and adjustment (during implementation) and set the stage for later evaluation. Adaptation interventions will need monitoring and the flexibility for adjustments of the theory of change based on learning and as contexts change and influence outcomes.
- For adaptation efforts that include environmental interventions, evaluation using modelling and statistical analysis may be particularly useful.

5.1 Current state of long-term monitoring and evaluation

45. A review of development programmes yields few examples of evaluations that address the evolution of programmes over a long time. Even evaluations of long-term programmes, such as the “Impact evaluation of drinking water supply and sanitation programmes in rural Benin”, often opt to focus on a particular range of time: “Although the supported programmes have operated since the 1990s and will run until at least 2015, for methodological as well as for practical reasons, the quantitative impact evaluation covers only the period 2008-2010” (Government of Netherlands and Government of Germany, 2011). Another example is the rare occasion when a development agency evaluates a whole scope of programmes over their entire period of implementation, such as the evaluation of support to Vietnam, Laos and Sri Lanka by the Swedish International Development Cooperation Agency (McGillivray et al., 2012). These evaluations tend to use a contribution rather than attribution approach, with a mix of case studies, analyses of longitudinal and thematic datasets, focus group discussions, and semi-structured interviews.

46. A search for long-term results of development impacts from investments in higher education in developing countries found little evidence of systematic evaluations or use of statistical methods. Most of the evaluations identified looked at the results of scholarships. However, they focused mainly on outputs and less on outcomes and impacts, and also lacked counterfactual evidence. A review noted a paucity of independent and robust evaluations and suggested that future longitudinal evaluations of investments in higher education should make use of the kinds of impact evaluation methods presented in Section 3 of this paper. It noted, however, that it is a considerable challenge to define, extrapolate, evaluate and measure impact of a specific higher education intervention, if it is at all possible, given its complex and multi-dimensional context and the long time-lag between intervention and impact (ACU, 2012).

47. The field of ecosystems and natural resources management yielded some information about long term monitoring and evaluation, due in part to the tangible nature of the interventions. In the conservation arena too, some effort has been made to contend with the challenge of evaluating interventions that will have impacts only over the long-term, given that efforts to reinstate species or improve ecosystems cannot manifest in the short term. For example, the official recovery plan for the red cockaded woodpecker describes a detailed series of measures that experts feel will be sufficient to prevent the species’ extinction in the South-eastern United States. However, the team does not expect to remove the woodpecker from the endangered species list until 2075 – nearly 100 years after formation of the original plan (Gullison and Hardner, 2009). Despite these long time horizons, monitoring and evaluation of process indicators and intermediate outputs and outcomes could provide a useful evidence base and inform management of the

intervention over time. If monitoring and evaluation are not integrated early on in the design and implementation of the programme it can be nearly impossible to carry out meaningful, cost effective evaluation later on.

5.2 Importance of monitoring over time

48. Many aspects related to monitoring over the long term are covered by issues raised in previous sections of this paper, especially Section 4 that describes how to set baselines and targets. Monitoring over time is essential to generate the data required for a useful evaluation, but given short time frames and lack of funding, this step is often de-emphasised in monitoring and evaluation. The failure to plan ahead, collect data before and during the programme and monitor relevant indicators makes evaluation more expensive and less effective in the end (because core aspects have to be created from scratch).

49. An example is the Sustaining Lakes in a Changing Environment (SLICE) project, which is the State of Minnesota's Department of Natural Resources' Section of Fisheries' long-term monitoring programme. Over the past several decades, Minnesota's lakes have been suffering as a result of land use changes and climate change. This project aims to monitor these various stressors in order to support sustainable long-term lake management. The monitoring of the lakes includes monitoring of physical, chemical and biological indicators and is taking place in two phases. The first phase included intensive data collection between 2008 and 2012. The second phase, beginning in 2012, consists of an extensive survey with samples being gathered from 24 lakes across the region. The selected lakes are considered to be representative across the state. This form of enduring data gathering is critical for long-term monitoring and can enable the development of climate smart management policies and strategies. However, a critical concern when putting in place such a system is the challenge of sustaining financial support mechanisms over adequately long periods of time (Hitt, 2012). An additional challenge is to be able to draw meaningful conclusions about the direction of travel and the effectiveness of interventions to allow for adjustment before it is too late. This will be a concern in long-term monitoring of adaptation interventions as well.

5.3 Techniques for evaluating over long time horizons

50. Although there are limited examples of long-term evaluations available, there are two particular strategies that may include components relevant for the context of adaptation. For adaptation projects that have to do with physical processes, for flooding or drought, modelling and statistical analysis can be used. For those that focus on capacity building, limiting factor analysis may be helpful in highlighting the factors that evaluators need to keep an eye on.

Modelling and statistical analysis

51. The physical characteristics of ecosystems allow for quantitative monitoring and evaluation. For adaptation efforts that include environmental interventions, evaluation using modelling and statistical analysis may be useful. For example, as part of the Lower Red River Meadow Restoration Project in Idaho, the project team took an approach that included a pre- and post-evaluation (in 1994, and 2000 and 2003, respectively) of the river restoration area. The evaluators used 17 performance indicators, comprising a suite of physical and biological components that interact within the river and wet meadow ecosystems. Indicators were chosen based on their collective potential to quantify and qualify short and long term changes and trends in the general ecological character of the restored area. The team used statistical analysis to identify post-restoration changes. The researchers felt that their monitoring and evaluation had satisfied criteria of determining a successful restoration, but would have liked to have a "control" river which they had not built into the evaluation because that would have doubled budget costs. Additionally, the team would have liked to do a post-restoration evaluation for 10 years following the project but could not secure the required funding (Klein et al., 2007).

52. The methodology and process used in this evaluation, especially the vegetation modelling and topographic data, would be useful for an adaptation process that relates to environmental intervention, i.e. re-greening or soil enrichment. This reiterates the finding noted earlier in the paper, that quantitative methodologies may be useful for evaluating a discrete piece of an adaptation strategy but those results often cannot be extrapolated to the entire adaptation intervention.

Limiting Factor Analysis

53. Perhaps more useful for informing the evaluation of adaptation initiatives is a technique called Limiting Factor Analysis (LFA), also from the biodiversity conservation arena. LFA is a technique that helps providers of development co-operation to develop a common understanding of the key factors that must be assessed, and if necessary (and possible) managed, for a biodiversity conservation project or programme to be viable over the long term. Not to be confused with statistical factor analysis, it is a simple and qualitative method of forecasting, with the goal of identifying currently unmanaged factors that are likely to prevent the grantee from achieving its objectives. Gullison and Hardner (2009) have identified a list of limiting factors relevant to a broad range of project types and ecological systems:

- Scientific understanding that is inadequate to formulate appropriate management actions to sustain the conservation target;
- Public policy that does not support conservation of the target;
- Legislation that does not offer sufficient legal protection to the conservation target;
- Institutional capacity that is inadequate to perform conservation activities;
- Economic pressures that cause destruction of the conservation target;
- Enforcement of laws and regulations that is inadequate to implement the legislation on the ground to protect the conservation target;
- Stakeholder support that is inadequate to conserve the target;
- Short-term funding that is insufficient to establish an adequate level of conservation management, including capital expenditures on equipment and infrastructure;
- Long-term funding that is inadequate to support the recurrent costs of conservation management activities.

54. For each evaluation, this list of factors is customised and, working with providers of development co-operation, the evaluators rank the status of each factor based on when the provider began the evaluation (i.e. a retrospective baseline) and then again at the point of project completion. Following this, providers of development co-operation are asked to identify all entities working to overcoming the identified limiting factors. Development providers must also consider the degree of control or influence they have on the limiting factors. The ultimate goal is to have evaluators consider the “big picture” and assess whether everything necessary is being done to reduce both short- and long-term risks to the project (Gullison and Hardner, 2009).

55. When LFA is done at the beginning of a project or programme it can highlight the potential obstacles to project success. LFA is therefore a valuable tool that can be used by evaluators, development co-operation providers, and other stakeholders to quickly get an understanding of the context in which they

are working. LFA works because it can include factors like long-term financing and policy and legal frameworks, which ultimately determine the ability of a project or programme to manage threats that appear down the road. Gullison and Hardner evaluated two biodiversity projects using LFA to help anticipate or forecast the long-term fate of these projects (2009). This method seemed successful in the two case studies, but shortcomings include subjectivity of development co-operation providers and stakeholders in scoring the extent to which particular factors are impeding conservation at their sites.

6. MONITORING AND EVALUATION FOR ACCOUNTABILITY AND LEARNING

56. Conventional monitoring and evaluation systems are often characterised by an emphasis on being accountable to funding providers and using a linear approach that focuses on input and output results measurements. The limitations of this approaches are now widely recognised (Mosse, 1998, Jackson and Kassam, 1998, Estrella, 2000, Villanueva, 2011). The complexity of adaptation, as well as the position of adaptation finance within the United Nations Framework Convention on Climate Change (UNFCCC), demands greater transparency to those who are intended to benefit from adaptation interventions. This means that consideration must be given, not only to intermediary adaptation outputs and outcomes, but also to the ultimate development impacts. To achieve this objective, monitoring and evaluation frameworks should focus on measuring outcomes and impacts for beneficiaries and engaging them in the process to understand how and why outcomes are achieved, and to increase local accountability for climate finance.

57. This section addresses the issue of accountability and learning by introducing the challenges implementing agencies face of being accountable to both providers of development co-operation and to beneficiaries. Specifically, it examines techniques for increasing accountability, the importance of participatory techniques for adaptation, and how learning can be integrated into teams and programmes to form an iterative feedback loop to continually improve and reassess outcomes. The section provides examples from the fields of health, conflict resolution, education and development planning. Box 6 summarises the main recommendations emerging from this section.

Box 6. Accountability and learning: Recommendations for monitoring and evaluation of adaptation

- The allocation of finance for climate adaptation should be transparent to the beneficiaries of this support, at the same time as implementers of projects and programmes are accountable to providers of development co-operation.
- Participatory techniques -- such as most significant change techniques, citizen scorecards and outcome mapping -- can be successfully used for monitoring and evaluation of adaptation in conjunction with techniques that ensure accountability.
- Informing the public on outcomes and eliciting public participation in monitoring is particularly important for ensuring transparency where outcomes are uncertain and may change over time.
- Integrating iterative learning into monitoring and evaluation of adaptation can be achieved by building an explicit feedback element into the programme, involving beneficiaries and key stakeholders in this process, and institutionalising a learning function within the team.

6.1 Participatory techniques for accountability

58. Monitoring and evaluation can also contribute to accountability through processes that are employed to generate evidence. Some authors note that while monitoring and evaluation processes in areas of conflict can be purely extractive, they also have the potential to be transformative. “By involving respondents in the process of indicator development and analysis, monitors and evaluators can help people understand their own place in – and possibly even their contribution to – a given context” (Saferworld, 2004). This participation is a learning process in itself, at the same time as it enhances transparency.

59. Evidence from development programmes suggests that focusing on participatory outcome-based monitoring and evaluation can be one way of putting the needs of beneficiaries and local stakeholders at the centre of monitoring and evaluation. In addition to ensuring local accountability it also promotes learning at different scales (Smith, Mauramoto and Rassman, 2012). Eliciting feedback from intended beneficiaries can help create a narrative that ultimately informs audiences beyond the evaluators. For instance, Heltberg et al. (2012) conducted an extensive survey after the 2008 global food and financial crisis to provide a narrative from the perspective of those who were affected, and to provide feedback on a range of relief schemes that were implemented after the crisis. Such narratives can contribute to a diverse set of interventions in addition to those that were implemented during the crisis. Table 2 demonstrates the range of participatory and social science techniques examined by Heltberg et al. (2012), highlighting which challenges they address that make them particularly relevant to monitoring and evaluation of adaptation.

Table 2. Participatory/social science techniques for monitoring and evaluating adaptation

Participatory/social science technique	Challenges addressed	Example of usefulness
Survey	Attribution, Baselines, Accountability	Socio-economic surveys and census surveys can provide reference data for estimating baselines
Semi-structured interview	Attribution	Along with other techniques, used to measure contribution of programmes to outcomes
Focus groups	Attribution	Used to elicit changes in farmers' attitudes after a land-use programme was implemented
Creation of a narrative	Attribution	Used to elicit changes in farmers' attitudes after a land-use programme was implemented
Local theories of change	Attribution	Can include participation to understand the intervention impact on local communities
Most significant change	Accountability	Used to evaluate the Community Driven Approaches to Address the Feminisation of HIV/AIDS in India project
Outcome mapping	Accountability	Used to evaluate schools programme in Zimbabwe, using teachers involved in the programme as part of the evaluation process
Scorecards	Accountability	Used in Bangalore to assess urban services and build accountability of service providers

60. Participatory monitoring and evaluation is particularly helpful in complex adaptation contexts to assess changes in attitudes and decision making, to examine the impact and effectiveness of interventions, and to build ownership and recommendations. Participatory data are only one input into monitoring and evaluating these contexts, but offer a simple way of understanding some impacts on livelihoods and household experiences that are complex to capture through a set of indicators. They also capture unintended consequences and impacts if conducted in an open process. For climate change adaptation, one challenge will be to enable participation in the long term, a characteristic of complex adaptation interventions. Participatory approaches to monitoring and evaluation may also need to be combined with

more traditional results frameworks to marry the need for evidence at the international level to accountability and learning within the programme at the local and national level.

Most Significant Change

61. The Most Significant Change (MSC) technique entails asking participants or beneficiaries to tell the stories of the most significant change they have experienced through the programme (Davies and Dart, 2005). These stories are filtered up through staff by a set of criteria and are eventually distilled into a set of top line changes and stories. The stories are accompanied by the reasons the stories were selected. The funders are asked to assess the stories and choose those that exemplify the sort of outcomes they wish to support, as well as the reasons behind their choices. This type of technique can raise unexpected outcomes and makes staff constantly involved in the analysis of outcomes and impacts throughout the programme. The MSC technique was, for example, used to evaluate the DFID supported project on Community Driven Approaches to Address the Feminisation of HIV/AIDS in India (India HIV/Aids Alliance, 2007). The approach highlighted the importance of mobilising the entire community, and not just the key population, since a successful outcome relied on the confidence of the entire community to provide an enabling environment for women. Further, the MSC technique helped promote accountability to beneficiaries by keeping programme managers in touch with ground realities as well as continuous review and re-alignment of the programme assumptions through learning.

62. MSC is well suited to the community and programme level; however, its use at the national level may be limited to being a support tool to acquire community and staff perceptions of change. MSC may be a useful tool for monitoring and evaluating adaptation, as it is an open process that allows unintended consequences and impacts to be considered. It allows project and programme managers to be responsive to changing contexts and long term changes as well as to the specific programme goals.

Beneficiary monitoring and participation

63. Beneficiary monitoring is another example of a participatory approach and can involve techniques such as public perception surveys, citizen scorecards and beneficiary satisfaction indicators in a logframe format. It should be emphasised, however, that beneficiary satisfaction may not equate to greater climate resilience and the relationship between the two is not well understood. The citizen scorecard technique has been used fairly widely to assess public service provision in different development contexts (see Ringold et al 2011 for a review). One example is Bangalore where scorecards were used amongst urban households in both slum and non-slum areas to assess urban services and to build the accountability of the service providers. The scorecard showed widespread dissatisfaction with the services and also documented the amounts people were spending in bribes and other charges to access supposedly cheap services, providing a useful basis for debate around these issues (Paul, 1999). This type of approach could be used at the local level in climate change adaptation to track both provision of climate resilient services (e.g. the ability to reach public services in times of extreme rain), beneficiary satisfaction and potential proxies for vulnerability over time.

Outcome based methods

64. Outcome mapping focuses on identifying changes in the behaviours of individuals, groups and organisations with which programmes work, rather than changes in the physical variables in which the programme is situated (Earl et al., 2001). This is an interesting approach for adaptation, as it focuses on direct aspects of change that do not need to be normalised or controlled for changing external factors. The process involves working with participants to assess changes in the direct sphere of the project. This addresses the challenge that a programme may contribute to development outcomes but is only part of a much wider external context. Outcome mapping involves: i) boundary partners (i.e. those with whom the

project directly works), ii) outcomes defined as changes in behaviours, and iii) graduated progress markers used to identify gradual change.

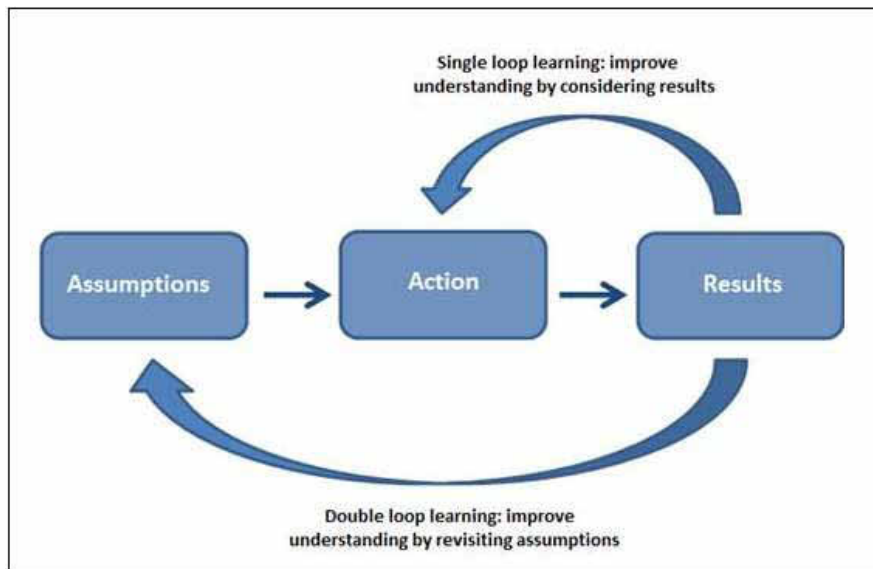
65. In a review of 10 years of outcome mapping it was found that the technique had widespread satisfaction amongst users and has inspired innovative processes of evaluation. The review identified four enabling factors that are needed in order for outcome mapping to succeed: i) a complex intervention, ii) recognition of and willingness to act upon complexity and an understanding of the rationale for outcome mapping application, iii) commitment of one or more champions, and iv) the availability of appropriate technical support (Smith, Mauramoto and Rassman, 2012). This process builds direct accountability towards local stakeholders. Boundary partners involved in the project provide feedback on how the project has been working and take ownership of improving it.

6.2 Integrating learning into monitoring and evaluation for adaptation

66. There is no consensus on what successful adaptation looks like. This is also likely to vary across contexts. It is therefore necessary to test and improve the effectiveness of investments and approaches, using more and better monitoring and evaluation, focusing on learning. Given the complex nature of many adaptation initiatives, complementary monitoring and evaluation efforts require a dynamic and iterative approach to learning and efforts to “institutionalise” the use of the information generated (e.g. by designing programmes with clear objectives, analysing the context and planning ahead for monitoring and evaluation).

67. Organisations or government departments that have reflexive mechanisms in place for learning can improve their monitoring and evaluation designs and performance on an on-going basis. This learning can take place through two interlinked loops: i) single-loop learning where actors learn and change their actions by improving their understanding of the most effective ways for achieving an objective, and ii) double-loop learning that goes beyond changing actions to also testing the assumptions on which these actions or objectives are based (see Figure 1) (Argyris, 1992; Argyris and Schon, 1978; Morris and Lawrence, 2010). Learning from and for adaptation will require double-loop learning (see Figure 1), as both the implementation mechanisms and desirability of programme targets are likely to evolve as contexts change and lessons are learned. Adaptation issues are multi-dimensional and complex, which requires questioning the underlying assumptions, beliefs and learning styles used to inform actions and targets (Tanner et al., 2012).

68. There are several ways to integrate learning into project and programme design and to make learning more integral to monitoring and evaluation than can be found in existing development practice. One way is to build an explicit learning phase into the planning cycles to address the frequently observed problem of activities leading straight into more activities without taking the time to properly learn from those that have already been undertaken. This is linked to the situation of the monitoring and evaluation within the planning cycle. A second approach to integrating learning into monitoring and evaluation is to use participation and involve beneficiaries and key stakeholders in the process, to incorporate changes as well as to learn from the evaluation process.

Figure 1. Illustration of single and double loop learning

Source: Adapted from writing by Argyris and Schon (1978), *Organisational learning: a Theory of Action Perspective*, Addison Wesley, Reading.

69. A third approach is to institutionalise a learning function within the project or programme team. This is the approach taken by USAID in Uganda with the Collaboration, Learning and Adapting (CLA) Plan. To ensure continuing improvements in the country support and to keep their country strategy “a living document”, the Ugandan USAID office has created a unit that will create dialogue around learning, facilitate iterative consultations, explore the wider context and horizon, scan for unexpected events, develop a response plan for unexpected events, and pilot innovative approaches. To achieve this objective, the unit has for example established a cross-mission community of practice on learning, replaced the monitoring and evaluation contractor with a CLA contractor, led stakeholder engagement on “game changers” and the big picture, and conducted an evaluation and re-design of an on-going agricultural project to respond to a strategy. This institutionalises the function of iterative assessments, uptake of learning and facilitating dialogue.

70. Similarly, developmental evaluation is a form of inbuilt, iterative learning that aims to help programme staff use evaluation findings in the development of programmes and throughout the programme cycle (Quinn Patton, 2011). Developmental evaluation has a focus on systems and problems that are complex, characterised by uncertainty and dynamic change. It does not refer to a specific set of methods or tools but rather “a mindset of inquiry”. Rather than performing evaluations based on a pre-determined schedule, developmental evaluations seek to coincide with annual workflow plans, major implementation steps and decisions about the future of programmes. Developmental evaluation generally uses the evaluator as a core member of the programme team who takes full part in developing and instigating evaluative thinking throughout an intervention. Their role is to facilitate discussion and decisions as well as to support data-based decision-making processes (Quinn Patton, 2011). For adaptation, developmental evaluation may need to be combined with formative and summative evaluations of certain programme elements, and may need to be timed for certain stages of the implementation. In order to be of real use this kind of evaluation also needs to be part of the initial stages of programme planning.

71. To facilitate learning from monitoring and evaluation processes, it is important to have sufficient time for projects and programmes to be implemented and for impacts to show. This is a particularly key area for adaptation where it may entail significantly longer timeframes than the ones used for most project and programme evaluations. Learning from experimental programmes needs to be fed into national strategies and plans so that national governments can start to move towards monitoring the resilience and achievement of national strategies with as much evidence as possible of what works in a national context.

REFERENCES

- ACU (2012), “Evaluating impact: measuring the impact of higher education development,” *ACU Spotlight* No 3.
- Argyris, C. (1992), *On organizational learning*, Blackwel, Cambridge.
- Argyris, C. and D. Schon (1978), *Organisational learning: A Theory of Action Perspective*, AddisonWesley, Reading.
- Balk, D., J.M. Guzman, and D. Schensul (2013), “Harnessing Census Data for Environment and Climate Change Analysis”, in *The Demography of Adaptation to Climate Change 2013*, UNFPA, IIED and El Colegio de Mexico, New York, London and Mexico City.
- Banerjee, A. and E. Duflo (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs, USA.
- Blanco-Fonseca, M. (2010), *Literature Review of Methodologies to Generate Baselines for Agriculture and Land Use*, Common Agricultural Policy Regional Impact – The Rural Development Dimension (CAPRI-RD), Bonn.
- Brooks, N. et al. (2011) “Tracking Adaptation and Measuring Development”, *Climate Change Working Paper* No. 1, International Institute for Environmental Development, London.
- Brooks, N., et al. (2013), “An operational framework for tracking adaptation and measuring development”, *Climate Change Working Paper* No. 5, International Institute for Environmental Development, London.
- Care International UK (2012) *Guidance for Designing, Monitoring and Evaluating Peacebuilding Projects: Using Theories of Change*, Care International UK, London.
- CIF (2013), *Pilot Program for Climate Resilience – Results Framework*, Climate Investment Funds, Washington, DC.
- Davies, R. and J. Dart (2005), *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*, MandE, Melbourne.
- DFID (2012), “Violence against Women and Girls- Guidance on Monitoring and Evaluation for Programming on Violence against Women and Girls”, *CHASE Guidance Note Series*, Department for International Development, London.
- DFID (2010) “Working Effectively in Conflict-affected and Fragile Situations”, *Monitoring and Evaluation*, Briefing paper 1, Department for International Development, London.
- Earl, S., F. Carden, T. Smutylo (2001), *Outcome Mapping: Building Learning and Reflection into Development Programs*, International Development Research Centre, Canada.

- Estrella, M. (2000), "Learning from change", in *Learning from Change. Issues and Experiences in Participatory Monitoring and Evaluation*, International Development Research Center, London.
- Gakhar, K.D., N. Kaur, and V. Kapur (2010), "Propensity Score Matching Method in Quasi-Experimental Designs, An Approach to Programme Evaluation of INHP-III", *Discussion Paper 3*, Sambodhi Research & Communications Pvt. Ltd, New Delhi.
- GCF (n.a.), "Green Climate Fund Board takes key decisions on operations and makes progress on 'Essential Eight'", available at: http://gcfund.net/fileadmin/00_customer/documents/pdf/GCF_Press_Release_fin_20140222.pdf.
- GIZ (2012), *Adaptation made to measure: A guidebook to the design and results-based monitoring of climate change adaptation projects*, Gesellschaft für Internationale Zusammenarbeit, Eschborn.
- Government of Denmark, Ministry of Foreign Affairs – Evaluation Department (2013), *Joint Evaluation of The international support to the peace process in Nepal 2006-12*, OECD DAC Evaluation Resource Centre, Paris.
- Government of Germany (2009), *Combating Climate Change: The German Adaptation Strategy*, Federal Ministry for the Environment, Nature, Conservation, and Nuclear Safety, Bonn.
- Government of Netherlands and Government of Germany (2011), *Impact evaluation of drinking water supply and sanitation programmes in rural Benin: The risk of vanishing effects*, OECD DAC Evaluation Resource Centre, Paris.
- Gullison, R. and J. Hardner, (2009), "Using limiting factors analysis to overcome the problem of long time horizons", *Environmental program and policy evaluation: Addressing methodological challenges. New Directions for Evaluation* Vol. 122.
- Hargreaves, M. (2010), *Evaluating System Change: A planning guide*, Mathematica Policy Research Inc, USA
- Heltberg, R., N. Hossain and A. Reva (eds.) (2012), *Living Through Crises. How the Food, Fuel and Financial Shocks Affect the Poor*, the World Bank, Washington, DC.
- Hitt, J. L. (2012), *Sustaining Lakes in a Changing Environment (SLICE): A Long-Term Monitoring and Evaluation Program*, Minnesota Department of Natural Resources, St. Paul.
- IFAD (2002), *Managing for Impact in Rural Development: A Guide for Project M&E*, International Fund for Agriculture and Development, Rome.
- India HIV/Aids Alliance, (2007), *Stories of Significance: Redefining Change: An Assortment of Community Voices and Articulations*, World Bank, Washington, DC.
- Jackson, E.T. and Y. Kassam, (1998), *Knowledge Shared: Participatory Evaluation in Development Cooperation*, Kumarian Press, Connecticut.
- James, C. (2011), "Theory of change review", a report commissioned by Comic Relief, London.
- Janger, J. et al. (2011), *Structural Change and the Competitiveness of EU Member States: Final Report – CR 2011*, European Commission, Brussels.

- Klein, L., S. et al. (2007), “Long-Term Monitoring and Evaluation of the Lower Red River Meadow Restoration Project, Idaho, USA”, *Restoration Ecology*, Vol. 15/2.
- Lamhauge, N. (2014), “National Monitoring and Evaluation of Climate Change Adaptation: Lessons from Developed and Developing Countries”, *OECD Environment Working Paper Series*, No. 73, OECD Publishing, Paris.
- Lamhauge, N., E. Lanzi and S. Agrawala (2012), “Monitoring and Evaluation for Adaptation: Lessons from Development Co-operation Agencies”, *OECD Environment Working Papers*, No. 38, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5kg20mj6c2bw-en>.
- Leeuw, F. and F. Vaessen (2008), *Impact Evaluations and Development. NONIE Guidance on Impact Evaluation*, Independent Evaluation Group, Washington, DC.
- McGillivray, M., A. Pankhurst and D. Carpenter, (2012), *Long-term Swedish Development Cooperation with Sri Lanka, Vietnam and Laos. Synthesis of Evaluation Study Findings*. Swedish International Development Corporation, Stockholm.
- Morris, J. and A. Lawrence (2010), *Learning from Monitoring & Evaluation – a blueprint for an adaptive organisation*, Social & Economic Research Group, Forest Research, Edinburgh.
- Mosse, D. (1998), “Process-oriented approaches to development practice and social research”, in Mosse, D., J. Farrington and A. Rew (Eds.), *Development as a Process: Concepts and Methods for Working with Complexity*, Routledge, London.
- OECD (2007), *Encouraging effective evaluation of conflict prevention and peacebuilding activities: Towards DAC guidance*, OECD Publishing, Paris.
- O’Flynn, M. (2010), “Impact Assessment: Understanding and Assessing our Contributions to Change”, *M&E Paper No. 7*, International Training and Resource Center, Oxford.
- Paul, S. (1999), *Making Voice Work: The Report Card on Bangalore’s Public Services*, World Bank, Washington, DC.
- Picciotto, R. (2013), “The logic of development effectiveness: Is it time for the broader evaluation community to take notice?” *Evaluation* Vol. 19.
- Pradhan, M. and L. Rawlings (2002), “The Impact and Targeting of Social Infrastructure Investments: Lessons from the Nicaraguan Social Fund”, *World Bank Economic Review* Vol. 16, Washington, DC.
- Prag, A. and C. Clapp (2011), “Setting National and Sectoral Baselines”, Prepared for the CCXG/Global Forum on Environment: Seminar on MRV and Carbon Markets by OECD and IEA on 28-29 March 2011, Paris.
- Pringle, P. (2011), *AdaptME: Adaptation monitoring and evaluation*, UK Climate Impacts Programme, Oxford.
- Quinn Patton, M. (2011), *Developmental Evaluation. Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York.

- Ravallion, M. (2008), "Evaluation in the Practice of Development", *Policy Research Working Paper No. 4547*, The World Bank, Washington DC.
- Ringold, D., Holla, A., Koziol, M., and Srinivasan, S. (2011), *Citizens and Service Delivery: Assessing the Use of Social Accountability Approaches in Human Development Sectors*, World Bank, Washington, DC.
- Rogers, P. (2008), "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", *Evaluation*, Vol. 14.
- Rosenbaum, P. and D. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, Vol. 70/1.
- Saferworld (2004), *Conflict-sensitive Approaches to Development, Humanitarian Assistance and Peace Building: Tools for Peace and Conflict Impact Assessment: A Resource Pack*, Saferworld, London.
- Sida (2006), *Manual For Peace And Conflict Analysis: Methods Document, Division For Peace And Security Through Development Cooperation*, Swedish International Development Cooperation Agency, Stockholm.
- Smith, R., J. Mauramoto and K. Rassman (2012), *10 Years of Outcome Mapping Adaptations and Support*, Outcome Mapping Learning Community, Vancouver.
- Spahn, J. M. (2008), "Reliability, validity, sensitivity and usefulness testing of dietetic specific monitoring and evaluation indicators and measurement scales", *Journal of the Academy of Nutrition and Dietetics*, Vol. 112/9.
- Spearman, M. and H. McGray (2011), *Making Adaptation Count: Concepts and Options for Monitoring and Evaluation of Climate Change Adaptation*, Gesellschaft für Internationale Zusammenarbeit, Eschborn.
- Stern, E., N. et al. (2012), "Broadening the Range of Designs and Methods for Impact Evaluations", *DFID Working Paper 38*.
- Tanner, T., M. Lockwood and F. Seballos (2012), *Learning to Tackle Climate Change: The Learning Hub*, International Development Institute, Brighton.
- USAID (2012), "Feed the Future, US Government's Global Hunger and Food Security Initiative", *M&E Guidance Series*, United States Agency for International Development, Washington, DC.
- Villanueva, P. S. (2011), "Learning to ADAPT: monitoring and evaluation approaches in climate change adaptation and disaster risk reduction – challenges, gaps and ways forward", *SCR Discussion Paper 9*, Institute of Development Studies, Brighton.
- Vogel, I. (2012), *Review of the 'Theory of Change'*, UK Department for International Development, London.
- Weiss, C. (1995), *New Approaches to Evaluating Comprehensive Community Initiatives*, The Aspen Institute, Washington, D.C.
- White, H., S. Sinha and A. Flanagan (n.d.), *A Review of the State of Evaluation*, Independent Evaluation Group, Washington D.C.

Woolcock, M. (2009), “Towards a Plurality of Methods in Project Evaluation: A Contextualized Approach to Understanding Impact Trajectories and Efficacy”, *Working Paper 73*, Brooks World Poverty Institute, Manchester.

World Bank (2010), “Reconstructing Baseline Data for Impact Evaluation and Results Measurement”, part of a special series on nuts and bolts of M&E, Washington, DC.

ANNEX I: SUMMARY TABLE OF APPROACHES AND METHODS

		Especially applicable to:			Especially helps address:				
	Methodologies (section where the methodology is described most extensively in parenthesis)	Simple	Complicated	Complex	Attribution	Baselines/ Targets	Long-time horizons	Accountability & learning	Example from
Overall monitoring and evaluation approach	Developmental evaluation			x			x		
	Impact evaluation	x	x		x				
	Longitudinal evaluation	x	x	x	x		x		OECD DAC on donor gender mainstreaming; Sida evaluation of its support to Vietnam, Laos and Sri Lanka
	Institutionalised learning: Timing of evaluations, reflective practice, public disclosure of information			x					USAID Uganda's Collaboration, Learning and Adapting (CLA) Plan; National Health System (NHS) in the United Kingdom
Formal Social Science Methods	Surveys		x	x		x		x	Through extensive surveying a crisis narrative for the 2008 food and finance crisis was created
	Focus group interviews		x	x	x			x	Evaluating land use change by using different methods
	Individual interviews		x	x	x	x		x	Evaluating land use change by using different methods
Econometrics and statistics	Modelling	x	x				x		Lower Red River Meadow Restoration Project in Idaho, USA
	Statistical analysis	x			x	x	x		Lower Red River Meadow Restoration Project in Idaho, USA
	Stochastic baseline					x			Food and Agriculture Policy Research Institute (FAPRI)
	Deterministic baseline					x			Food and Agriculture Policy Research Institute (FAPRI)
	Normalisation		x			x			Evaluating structural change and competitiveness of EU member states

ENV/WKP(2014)12

Experiment Related	Case studies		x	x			x	x	USAID's Caribbean Agricultural Extension Project
	Experimental design w/ randomised controls	x			x				
	Quasi-experimental evaluation	x	x		x		x		
	Propensity score matching	x			x	x			WHO's INH Program, & Emergency Social Investment Fund of Nicaragua
	Phased pipeline	x			x				Health facilities being upgraded
	Regression analysis	x			x				Red de Proteccion Social (RPS), a conditional cash transfer programme in Nicaragua
Participatory	Most significant change analysis			x	x			x	Community Driven Approaches to Address the Feminisation of HIV/AIDS in India", supported by the DFID challenge fund
	Beneficiary monitoring	x	x	x				x	Bangalore where scorecards were used amongst urban households in both slum and non-slum areas to assess urban services and build the accountability of the service providers
	Limiting factor analysis		x	x			x		Biodiversity projects were evaluated using limiting factor analysis
	Reconstructing baselines (recall techniques)		x	x		x			
	Outcome mapping and related techniques			x				x	Education project in Zimbabwe
	Sequential targeting		x			x	x	x	DFID programmes designed to reduce Violence against Women and Girls (VAWG)
Iterative and bigger picture methods	Results based monitoring			x				x	UNDP
	Theories of change		x	x	x		x		CARE International UK's document on peacebuilding
	Stepwise approach		x		x				Can help link trends in health outcomes, coverage and risk behaviours, access and quality of services, and funding
	Contribution analysis			x	x				
	Scenario building			x		x	x		USAID Feed the Future Programme
	Contextualisation			x		x			Work in peace-building and fragile states Systemic Spatial Decision Support Systems used in assessing effectiveness of biodiversity conservation and land use planning
	Rolling baselines		x			x	x		
	Reconstructing baselines		x	x		x			

ANNEX II: GLOSSARY OF KEY TERMS

Term	Definition
Attribution	Attribution is the ascription of a causal link between observed (or expected to be observed) changes and a specific intervention.
Baseline	A baseline is the situation prior to a development intervention, against which progress can be assessed or comparisons made.
Beneficiary Monitoring	Beneficiary Monitoring is a participatory approach and can involve techniques such as public perception surveys, citizen scorecards and beneficiary satisfaction indicators in a logframe approach.
Complex Intervention	A complex problem or system is consequently one where there is both a fundamental uncertainty about causal relationships, where a causal chain may only become apparent after an event, and little agreement about the nature of the problem, how it can be addressed and what a solution will look like.
Complicated Intervention	A complicated intervention can be composed of many parts, and there is therefore less certainty about the outcome, but still agreement about the intervention.
Contribution Analysis	Contribution Analysis focuses less on programme interventions as single causes and more on how an intervention interacts with other aid or non-aid factors and analyse whether an intervention was a necessary and or a sufficient causal factor, along with other factors.
Counterfactual	The ability to demonstrate that changes would not have happened if the intervention had not taken place.
Deterministic Baselines	A deterministic baseline is average conditions or business as usual conditions.
Deterministic Models	Deterministic models are those in which the variables account for the different states observed and are determined by parameters in the model.
Development Evaluation	Developmental evaluation informs and supports innovative and adaptive development in complex dynamic environments. Developmental evaluation brings to innovation and adaptation the processes of asking evaluative questions, applying evaluation logic, and gathering and reporting evaluative data to support project, programme, product, and/or organisational development with timely feedback.
Difference in Differences	Difference in differences compares impacts between treatment and control (comparison) groups both before and after the implementation of an intervention.
Endogeneity	Occurs where independent variables (programme inputs) and the dependent variable (outputs and outcomes) become correlated with various other factors that were either not controlled or predicted in the original impact assessment model.
Experimental Design	Compares the treatment group (e.g. programme participants) against a control group (e.g. non-participants).
Focus Groups	Also referred to as interview or group discussion. An interviewing technique whereby respondents are interviewed in a group setting.
Impact Evaluation	Assesses the direct and indirect contribution of an intervention. Answers the question of whether a specific intervention has a direct causal relationship to a specific outcome.

Limiting Factor Analysis	Limiting Factor Analysis is a technique used to work with the donor and grantees to develop a common understanding of the key factors that must be assessed, and if necessary (and possible) managed, for a project or programme to be viable over the long term.
Logical Framework (Logframe)	A logframe is a tool for structuring the main project activities, outputs and outcomes and established the linkages between them. It is used as both a management tool and as a framework for monitoring and evaluation.
Longitudinal Evaluation	Longitudinal evaluations are evaluations that address the evolution of programmes over time.
Modelling	The construction of physical, conceptual or mathematical simulations of the real world. Models help to show relationships between processes (physical, economic or social) and may be used to predict the effects of changes in land use.
Most Significant Change Analysis	A technique that revolves around asking participants or beneficiaries to tell the stories of the most significant change they have experienced through the programme.
Normalisation	Conceptual procedure in database design that removes redundancy in a complex database by establishing dependencies and relationships between database entities. Normalisation reduces storage requirements and avoids database inconsistencies. Normalisation can help standardise data by different trends (unusual or usual) by adjusting the data against these trends and means. This makes the data derived from different contexts more comparable.
Outcome Mapping	Outcome mapping focuses on identifying changes in the behaviours of the individuals, groups and organisations with which programmes work, rather than changes in the physical variables that may correlate to the development programme objectives.
Phased Pipeline	Phased pipeline is used for programmes that are implemented in a phased or staggered design with time lapses between the completion of one intervention and starting of a new one.
Propensity Score Matching	Propensity score matching is a technique used by programmes to construct a matched comparison group that has the same propensity to receive the intervention benefits as the treatment group.
Purposeful Sampling	A sample in which the individual units are selected by some purposive method. It is therefore subject to biases of personal selection and for this reason is now rarely advocated in its crude form.
Quasi-Experimental Design	Explicitly addresses the validation challenges of attribution and the counterfactual when evaluating the impact of an intervention by comparing intervention participants and some form of non-intervention control or comparator group both before and after the intervention. Different rationales are used to assign control groups but this is undertaken in a non-randomised way.
Randomised Experiments	These experiments involve randomly choosing individuals or units from both the control group and those that have been part of an intervention.
Reconstructing Baselines	An approach to developing baselines is to use secondary administrative data (e.g. national surveys) and practical techniques (e.g. recall and mapping techniques) to reconstruct baselines when indicators change.
Regression Analysis	Regression analysis will show the degree of variation of samples around a linear or non-linear relationship, and thus the statistical significance of the relationship.
Results Based Monitoring	This approach has a strong focus on learning from monitoring and evaluation. This has involved putting strong measures to enhance learning through regular exchange of information, reporting, knowledge products, learning sessions, and evaluation management response systems to incorporate learning into planning and programming.
Rolling Baselines	Collecting baselines during different stages of the programme instead of at one time.

Scenario Building	Scenario building is based on a set of possible scenarios that will influence the evolution of the intervention. The purpose is not to predict but to think through what causes of action and contingencies the programme could select depending on different futures.
Sequential Targeting	Sequential targeting entails establishing targets as short term objectives for achieving an outcome. This would mean setting interim targets or several milestones that relate to expected performance over short intervals.
Simple Intervention	In a simple intervention, there is both certainty about causal relationships and agreement that the intervention is useful.
Statistical Analysis	Analysis that involves collecting, organising and analysing quantitative data.
Stepwise Approach	A stepwise approach begins with inputs and processes, moves through outcomes, outputs, and finally impacts, while accounting for contextual changes along the way.
Stochastic	The adjective “stochastic” implies the presence of a random variable; e.g. stochastic variation is variation in which at least one of the elements is a variate and a stochastic process is one wherein the system incorporates an element of randomness as opposed to a deterministic system.
Stochastic Models	Stochastic models have an element of randomness to them, and different states are not captured by different sets of variables but by probability distributions. Stochastic models consider several alternative futures or scenarios.
Surveys	A survey is an investigation about the characteristics of a given population by means of collecting data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology.
Theories of Change	In its simplest form, a theory of change can be stated as “We believe that if we do x (action) then we will achieve y (progress).” A theory of change seeks to identify the intended activity behind the “if” and the expected change behind the “then”. If the phrase “because” (reason) is added to the end of this sentence the theory of change is further strengthened.
Triangulation	Use of several information sources and different methods simultaneously to generate information about the same topics.