Proposed methodology to assess environmental and social impacts of certification of ecosystem services

Sini Savilaakso and Manuel R. Guariguata

26 October 2013

ForCES®
Forest Certification for Ecosystem Services

CIFOR
CENTER FOR INTERNATIONAL FORESTRY RESEARCH

Contents

Acknowledgements

We thank Stefan Salvador from FSC for sharing his ideas on different approaches towards certification of ecosystem services and for his comments. Mauro Ciriminna, Chris Henschel, Marion Karmann, Alison von Ketteler and Wanggi Jaung provided useful comments. We acknowledge that the indicators presented here are team work and thank all the participants of the indicators workshops in 2012 for their contribution. Special thanks to Que Anh Vu Thi, Adrian Enright and Fabian Noeske in Vietnam, Kalyan Gauli, Megh Adhikari, and Bhishma Subedi in Nepal, Roxanna Ayllon, Ana Young, and Claudia Cuiza in Chile, and Tita Rini, Dita Ramadhani and Aditya Bayunanda in Indonesia for their support in organizing and conducting the indicator workshops. The Global Environment Facility and the Government of Finland provided financial support. Partial funding comes from the CGIAR Research Programme on 'Forests, Trees and Agroforestry: Livelihoods, Landscapes and Governance' led by the Center for International Forestry Research in partnership with Bioversity International, the International Center for Tropical Agriculture and the World Agroforestry Centre.

Introduction

Ecosystem services and goods are the multiple benefits people obtain from ecosystems[1]. The benefits provided by forests include for example, carbon sequestration, prevention of erosion, flood control, and water supply and purification as well as aesthetic beauty. Although humans are fundamentally dependent on these services, they also pose threat to the services through their activities such as deforestation and water pollution. As ecosystem services are at the core of human well-being, it is important to recognize and safeguard them when management decisions regarding natural resources are made.

During the past two decades the value of various benefits ecosystems provide has been increasingly recognized and economic mechanisms have emerged alongside regulation. Forest management certification came into existence after the 1992 Rio conference as a part of a larger drive towards more sustainable resource management and improved livelihoods of the rural communities[2]. At the same time poverty was recognized as one of the key threats to environment and therefore, the integrated conservation and development projects (ICDPs) emerged as an approach to protect the environment through poverty alleviation[3]. More recently payments for environmental services (PES) have arisen as a mechanism to achieve the maintenance of the provision of these services[4]. The idea behind the PES is simple: fair compensation should be provided to local landholders and users who in return adopt management practices that conserve and restore the provision of ecosystem services. Probably the most well known of these schemes is the REDD+[5] which seeks to enhance carbon sequestration potential of forests while at the same time providing livelihood benefits.

The increased recognition of ecosystem services has also meant a rise of new standards to certify initiatives that provide ecosystem services[6]. At the same time the Forest Stewardship Council (FSC) has moved towards more explicit inclusion of ecosystem services into its strategy[7]. As its mission, FSC promotes environmentally appropriate, socially beneficial, and economically viable management of the world's forests[8]. Environmentally appropriate forest management is supposed to ensure that harvesting of timber and non-timber forest products maintains ecological processes, forest productivity and forest's biodiversity. Also, it should help local people and society at large to enjoy long-term benefits and provide incentives for local people to sustain the resources. Finally, the profits from the use of forest resources should not come at the expense of the ecosystem, or affected communities. This all sounds very promising but how effective is FSC in achieving these goals??

As a part of a growing trend of evidence-based policy making calls to prove effectiveness of different interventions have become more frequent and louder[9]. Properly conducted impact evaluations are one answer to those calls as they can provide credible evidence on performance and whether the desired outcomes are achieved. In addition to inform about the effectiveness of the certification intervention to achieve its stated goals, an impact evaluation has other benefits[10]: It provides accountability to those that are affected by the standard system and to those that are supposed to benefit from it; it improves credibility of the standard system if the results of the impact evaluations are openly available; it builds

capacity of the standard system and those involved in the impact evaluation (e.g. forest enterprises); and it informs strategy regarding policy decisions.

At the heart of an impact evaluation is the question "Why?". Why did the changes that are observed happen? Did they happen because of a specific intervention, e.g. change in forest management practices in order to comply with certification requirements, and to what extent are they due to the intervention? Hence, impact evaluation looks at the causal link between the intervention and its desired impacts and how much of the change can be attributed to that specific intervention. They are usually focused on average impacts, e.g. how much the average income of the project participants increased during the project[11]; or did the rate of accidents reduce because certification requires implementation of safety practices; or did the amount of carbon conserved increase because of implementation of better forest management practices? Impact evaluations are carried out at discrete points in time[12] and seek to answer specific questions related to the intervention's implementation and results. As the desired impacts may take long time to materialize, it is common to evaluate progress towards those impacts by looking at the program outcomes and outputs.

As FSC moves towards broader inclusion of ecosystem services (ES) under its certification system and considers different approaches under which the certification of ecosystem services can take place, it is important to have a methodology to assess the impacts of certification of ecosystem services to increase confidence that environmental and social benefits are created as well as to assess the performance of the standard towards its goals. The requirement for impact evaluation is also stated under the criteria 8.2[13] "The organization shall monitor and evaluate the environmental and social impacts of the activities carried out in the management unit, and changes in its environmental condition". Here a methodology to assess the environmental and social impacts of ecosystem service certification is proposed. First the general framework is introduced before a more detailed description of each of the steps.

The methodology and how to use it

The idea of this methodology is to provide a simple four step framework on which an impact evaluation can be based (Figure 1). The first step is to identify whether the site has impact potential, i.e. the potential for certification to make a difference. In other words identifying impact potential means assessing whether there are ecosystem services that need to be conserved or can be enhanced and whether there is willingness to pay or provide nonfinancial compensation for that, for example if the quality of a drinking water source is threatened by forest conversion. The second step is to prepare for an evaluation, for example by creating a theory of change. The third step provides guidance on how to move from expected certification outcomes to the definitions of indicators to describe and measure impact and fourth step discusses issues related to data collection.

Within the third step four increasingly demanding approaches towards monitoring and certification of ecosystem services are considered[14]: 1) Responsible forest management only where the knowledge that a forest is managed according to the FSC Forest Stewardship standard is enough 2) Forest management and monitoring of the provision of ecosystem services, 3) On top of 1 and 2, the services are properly accounted and aspects such as additionality and leakage are addressed, and 4) A reward mechanism for the provision of ecosystem services is included on top of the option 3. Basically demand for and complexity of monitoring increases, the more aspects the model covers (Figure 2). However, at the same time it can be assumed that the level of either monetary or non-monetary compensation increases as the services are properly accounted for.

In the rest of this document the framework is introduced step by step. The overall impact and selected outcomes specified in the FSC's global strategy[15] are used to illustrate the framework and to give guidance on the indicators. Figure 4 in Box 1 shows the key activities and outcomes in the process to certify ecosystem services. Although the use of control sites or subjects is recommended to correctly attribute changes in the status and trends of indicators to an intervention[16], the methodology can also be used without them or mixed methods can be used to gather conclusive information.

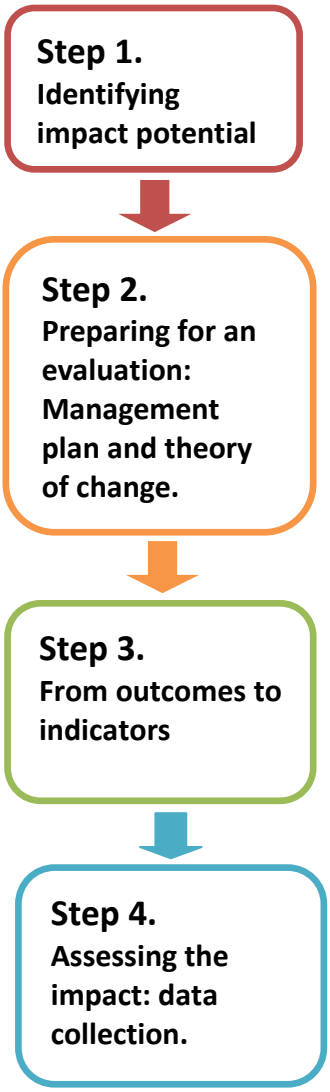


Figure 1. The steps of the impact evaluation methodology.

Step 1. Identifying impact potential

To identify whether certification of ecosystem services has potential to deliver biophysical and social benefits, the following four activities should be conducted:

1. **Identify potential forest users and local communities and engage with them.**
2. **Identify and map potential ES (with the stakeholders).**
3. **Identify threats, cumulative effects and actions that can influence the provision of ES (with the stakeholders).**
4. **Identify potential markets or buyers for the services.**

The purpose of these activities is to investigate whether certification of ES can realistically deliver impact on environmental and social conditions, e.g. there are buyers, the users of the forest are willing to participate in an ES scheme, there are no legal barriers to certification, and there are valuable ES that can be sold. Below they are discussed in more detail. It should be noted that this step can be done in

parallel with the free, prior, and informed consent (FPIC) process (to get agreement for the engagement in related activities) as these activities overlap to a certain extent. FSC provides guidance on how to conduct FPIC.

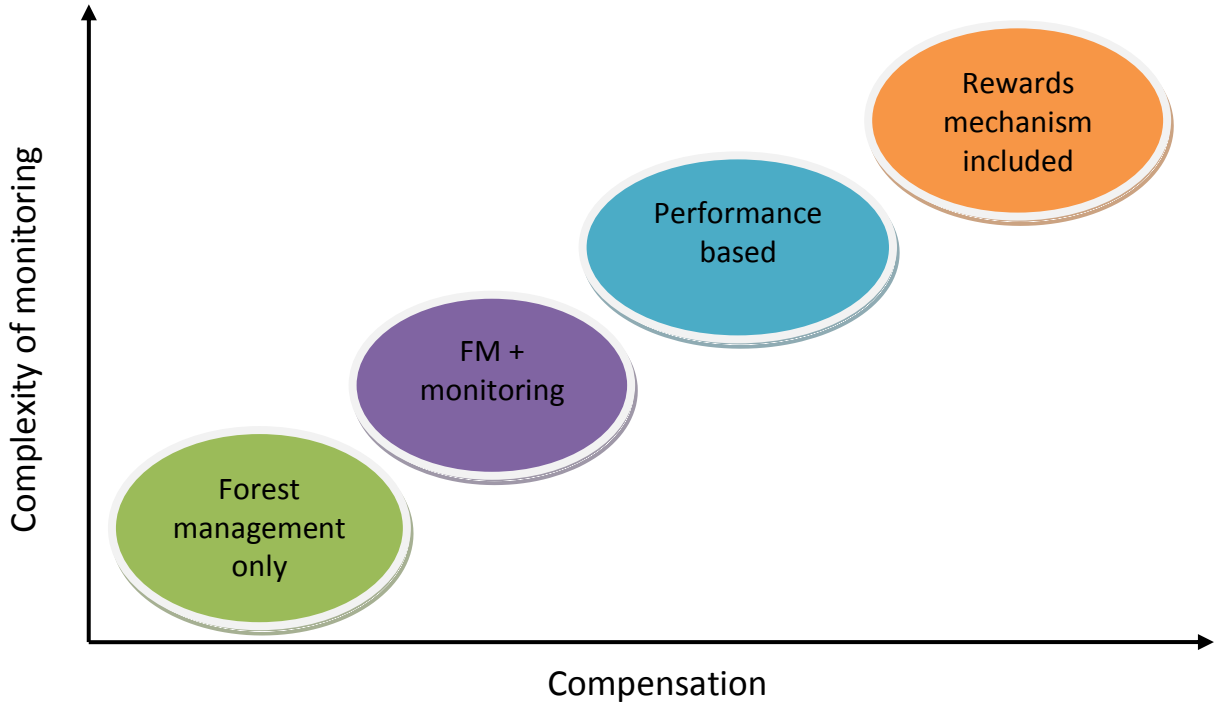


Figure 2. A simplified diagram of the four different approaches for certification of ecosystem services in order of complexity involved in monitoring and assumed rewards. The compensation refers to compensation between different approaches within a project and hence, is not suitable for comparison between projects.

1. Identify potential forest users, local communities, and stakeholders and engage with them.

Regardless whether the planned certification of ES originates from a community or from a private entity, engaging with forest users and local communities early on is crucial. This is recognized by FSC, and the revised FSC Principles and Criteria for Forest Stewardship[17] require engagement with local communities and forest users before a project is implemented as part of a process to obtain free, prior, and informed consent (FPIC)[18]. At this stage also other stakeholders and their role should be identified, e.g. a company that has timber management rights to the area but is not using them at the time. Furthermore, it is important to differentiate between potential sellers of the ES and other users who have no interest in selling the ES but who can influence the ES provision.

There are three questions that should be answered at this stage:

- Who is using the area or has a stake in it?
- What kind of rights do they have regarding the area (formal or informal, land tenure or user rights)?
- Who are the potential sellers among the forest users and other stakeholders?

The explanatory notes for the FSC Principles and Criteria for forest stewardship (FSC-STD-01-001 V5-0 D4-9) provide guidance on identification of communities:

They include those communities, which have affirmed rights to lands, forests and other resources based on long established use, and also those who have not yet done so (from a lack of awareness or empowerment). Also, communities affected by management activities include those neighboring the Management Unit and those that are more distant who may experience negative impacts as a result of activities within the Management Unit.

2. Identify and map potential ES.

Once the forest users and local communities have been identified, the next step is to engage with them for initial identification of ecosystem services and their location[19]. It should be noted that FPIC process should be parallel to identification of the ecosystem services to gauge the willingness of the forest users and local communities to participate in the project. As the concept 'ecosystem services' may not be well known among them, an alternative way to ask about the services is to ask about their use or benefits provided by the area, e.g. what do you use the lake for or what benefits does the area provide? Direct and indirect benefits, such as health benefits from a forest providing erosion control in sandy areas, can subsequently be identified from the answers.

As part of the identification of the services, it is important to map the spatial distribution of sources (areas that generate an ecosystem service) and sinks (landscape features that can absorb, degrade, or deplete a service) of the services[20]. As ecosystem services are not necessarily tied to a certain land cover type, mapping them will help to make informed management decisions on what and where are the possibilities to maintain and enhance the provision of ecosystem services[21].

The questions to be answered at this stage are:

- What are the ecosystem services stakeholders receive from the area?
- What other potential benefits the area provides (e.g. water for users further away)?
- Where are the sources and sinks of the services located**?**

3. Identify threats and opportunities, cumulative effects, and actions that can influence the provision of ES.

Threats are a large component when the potential for impact is defined. Where the services are under medium to high level of threat that can be addressed at reasonable cost the impact potential is large. Whereas where the level of threat is low or where the opportunity costs of avoidance and/or reduction of relevant threats are high, no permanence in the service provision can be expected[22] unless the forest managers or users commit to actively enhance the provision of an ecosystem service that is otherwise not under threat.

Threats take place from small to large scales. Those that happen within defined small areas are easier to address than those that happen at the landscape level. However, risks for threats at all levels should be considered to determine whether those threats can be avoided or reduced.

At the same time opportunities exist or may arise even without threats. For example, improved forest management practices may lead to the situation where more carbon is conserved in the forest than previously. This is a benefit that should be acknowledged and possibly compensated for example through the price of timber or through selling of the carbon credits.

Where there are no threats present, the forest users or managers may still decide actively maintain or enhance the ES provision. In that case it is important to consider the potential of forest management activities to influence the ES provision.

Another thing to consider together with threats is the potential of actions (not necessarily all threats) to create cumulative effects[23] as even low level threats can together have considerable impact on the service provision. There are four types of cumulative effects that can occur[24]:

1. Additive: Impact (A+B) = impact A + impact B
2. Compensatory: Impact (A+B) = impact A - impact B
3. Masking: Impact (A+B) = impact A (or impact B)
4. Synergistic: Impact (A+B) > impact A + impact B i.e. the resulting impact is more than the sum of the parts.

As can be seen the effect of an action can be either positive or negative. As with threats, the scale is important: an impact maybe missed altogether if it is not addressed at an adequate scale. What is an adequate scale will depend on the geographical scope of the activity. The questions to be answered at this stage are:

- Are there threats to the ecosystem services?
- What are the threats?
- What are the opportunities?
- What is the potential of forest management actions to enhance ES provision?
- Do actions at the area have a potential to create cumulative effects?

4. Identify potential markets or beneficiaries of the services.

Final activity when gauging the potential for impact is to identify potential markets or beneficiaries of the certified ecosystem services. The markets can be local or they can be distant, e.g. a consumer in Europe may be willing to pay more for timber that was harvested in a way that does not deplete carbon stocks. A lack of markets does not mean a lack of impact of certification, only that the scope for impact is different regarding some aspects, e.g. without markets monetary benefits may not be realized and hence, the positive impact of them on livelihoods of a local community. The questions that should be answered are:

- Which are the potential benefits?
- Who benefits from the ecosystem service provision?
- Can they or are they willing to compensate for the certified ecosystem service provision?
- What other possible sources of compensation are available, e.g. government support?

Step 2. Preparing for an evaluation: management plan and theory of change.

After it is clear that there are ecosystem services in the area that have potential to be certified, the next steps are:

5. **To prepare a management plan to maintain and enhance the provision of ecosystem services.**
6. **To prepare a theory of change that acts as a roadmap towards the management goals and helps to evaluate impact.**

Below we discuss these shortly.

Management plan

Management plan answers to the question of how to manage the forest to provision ES. Having and implementing a management plan is a compulsory requirement for FSC Principles and Criteria for forest stewardship and hence, probably will be for the certification of ecosystem services as well. Currently a management plan should include (FSC-STD-01-001 V5-0):

- Clearly stated policies and objectives that express the long-term vision, values and objectives for environmentally sound, socially beneficial and economically viable forest management.
- Verifiable targets for assessing progress with achieving the objectives.

The design of the management plan is flexible and depends on the scale, risk, and intensity of the activities. Thus, it can range from a simple few page document to a relatively complex one that includes maps, field guides etc. It should be prepared and approved prior to the start of new management activities taking place and all affected stakeholders should be engaged in management planning and monitoring. Further guidance can be found in the explanatory notes for the FSC principles and criteria for forest stewardship (FSC-STD-01-001 V5-0 D4-9).

Below are questions that can help to integrate ES into forest management plan:

- What ES do you want to provision?
- Do you want to maintain or enhance the ES provisioning?
- What are the current management activities that contribute to the provision of ES?
- What additional activities will be needed to achieve the goals for ES provision?

Theory of change

Theory of change is a document that answers to the question what is the impact that is wanted to be achieved with the management and provision of ES and their certification i.e. why undertake certain kind of forest management to provision ES. It is a causal model that describes relationships between activities, outcomes and impacts (see box 1 for more comprehensive definition[25]) and helps to create more targeted management plan. The preparations of a management plan and a theory of change are interdependent as the management goals informs the theory of change which in turn chart the actions that will be needed in order to achieve those goals (for an example see appendix 1). Whether the desired outcomes and overall impact were achieved will be evaluated later on.

The structure

The form of a theory of change can range from a simple flow chart that lists key things into a complex document that tries to be as specific as possible[26]. One easily understandable approach is to develop a visual diagram that charts pathways from activities to outcomes and impacts (for an example see appendix 1). In addition to outcomes and outputs it is important to include inputs into the theory of change as they can influence the outcomes independently of the intervention, *i.e.* the intervention may be good but if not enough resources are put into implementation, the intervention may fail and no impact can be found [27]. At least as important is to include possible contextual factors that may influence the outcomes but are not directly controlled by the intervention[28], for example, new laws and regulations or market price fluctuations that affect opportunity costs between different land use types. Finally the theory of change should include assumptions on which the causal logic is based as well as the risks (see box 1 for definitions). These can be in a form of separate lists instead of being part of the main diagram.

The components

The components of theory of change can be thought as a pyramid (figure 3). The inputs form the foundation that determines what activities can be undertaken. The activities in turn determine what will be the outputs on a path to outcomes and impacts. All this happens in a real world context that also influences the results. Although the impact pathway is from activities to impacts, the logical sequence to start preparing a theory of change is to think what is the purpose of the intervention? What are the desired long-term impacts that the intervention can help to achieve? For those aiming to be Forest Stewardship Council (FSC) forest stewardship certificate holders the overall goal is defined by FSC as *"world's forests meet the social, ecological, and economic rights and needs of the present generation without compromising those of future generations"* [29]. Although this is the overall impact that any FSC certificate holder should aim for, more specific impact goals can be defined depending on the stakeholder values, such as:

- Increased environmental resilience.
- Improved productivity on surrounding agricultural lands.
- Livelihoods are sustainable in long-term.

- Improved wellbeing of forest workers and forest dependent communities.
- Improved gender equity.

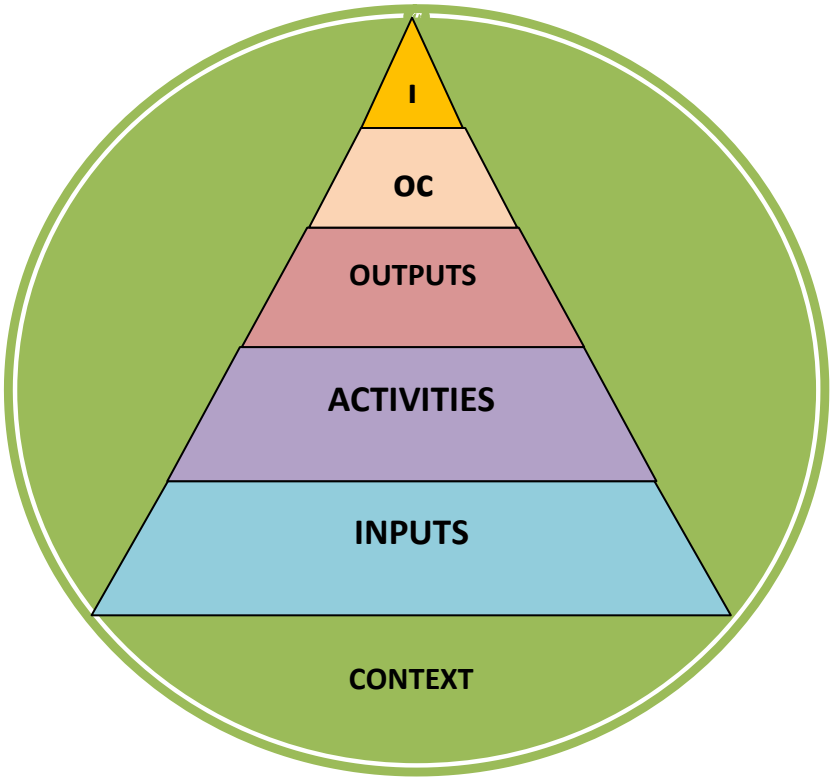


Figure 3. The components of theory of change build on each other and create a pathway from inputs to impacts. In the figure letter 'I' denotes impacts and 'OC' outcomes.

Once the intended and potential unintended impacts are clarified, the next step is to define outcomes that will lead to desired impacts. The outcomes can be modified to be more specific, e.g. pollination services enhanced, and additional outcomes can be added to reflect the management goals, e.g. where the ecosystem service is bundled with timber production, then one of the outcomes could be that timber yields are sustained.

Once the outcomes have been identified, the focus is on what are the outputs that will show progress toward each outcome. For example, to maintain and restore biodiversity, species composition should either remain similar following the activities or in case of restoration move towards desired species composition, e.g. the original species composition that was disturbed by the activities. The outputs and outcomes will be further discussed in the next section as they form the basis for indicators and data gathering.

Box 1. Defining key terms in impact evaluation

Theory of change

Theory of change is a causal model that describes relationships between implementation activities, outcomes and impacts. Thus, it defines key assumptions on how the change is supposed to occur. It also includes outside conditions and influences that can affect the outcomes.

Inputs

Inputs are human, financial and other resources that are available to the program. The focus of an impact evaluation should be on the resources specifically allocated to the program and not on broader conditions that are necessary for the success of the program. For example, money to conduct training on certification requirements to aid successful implementation.

Activities

Measures undertaken to implement certification requirements and goals of management actions. Activities convert the inputs into specific outputs.

Outputs, outcomes and impact

In this document *outputs* are products and services which result from the intervention. They may also include changes resulting from the intervention which are relevant to the achievement of outcomes. *Outcomes* are the short-term and medium-term impacts that result from certified forest management, i.e. the consequences of implementing certification requirements. *Impacts* are defined as the long-term positive and negative changes in the social and environmental situation that result from an accumulation of outcomes, e.g. movement in or out of poverty. The impacts can occur directly or indirectly and they can be either intentional or unintentional. Figure 4 shows the sequence from activities to impacts.

Contextual factors

Contextual factors characterize the socioeconomic, institutional and biophysical setting in which the project is implemented.

Assumptions and risks

The theory of change is built on a*ssumptions* that certain changes result from certain actions. *Risks* can influence the realization of intended results.

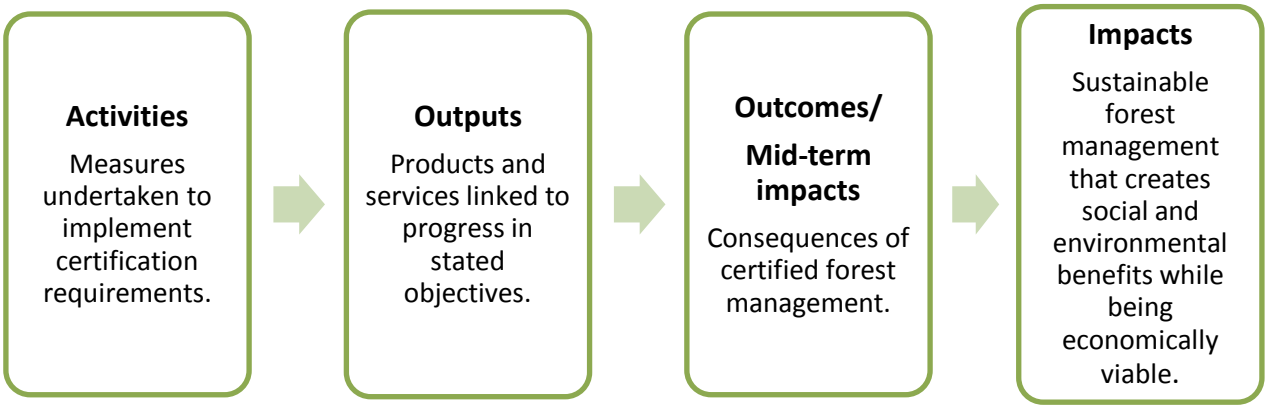


Figure 4. The path from activities to impacts.

The questions below will help to build a theory of change:

- Why do you want to provision the respective ES i.e. what is/are the impact(s) that is desired by provisioning certified ES?
- What are the outcomes (short- and mid-term impacts) that will lead to the desired impacts ?
- How can you achieve what you aim? I.e. what activities are needed to achieve the outcomes?
- What are the measurable outputs from the activities (i.e. what will result from the activities)?
- What resources (e.g. funds, technical assistance) are available to implement the activities?
- What are the contextual factors including barriers (e.g. legal framework, market conditions, and technical knowledge of staff) that may influence the achievement of outcomes and impacts?
- What are the assumptions behind the causal pathways created in the theory of change? E.g. training people leads to improved forest management; planting seedlings will reduce soil erosion.
- What are the risks that may affect the realization of intended results? E.g. conflicts arise between villagers - on land use, on benefit sharing, on power structures..

Step 3. From outcomes to indicators

Once the outcomes and outputs to achieve the impacts are specified, indicators can be derived from them. The purpose of indicators is to help to assess whether the forest management is successful in delivering the desired impacts. For example, if one of the goals for the forest management is to improve water quality (the impact) through decreased soil erosion (the outcome) then an indicator could be the amount of total suspended solids in the water (related to the outcome) or it could be number of surviving seedlings of those planted (output indicator). A word of caution: Although the hypothesis is that the planted seedlings will lead to the outcome of reduced soil erosion, it may not be the case in practice, e.g. the number of surviving seedlings is too low to make a difference or the place to plant them is not correctly chosen. Thus, the output indicator may not always predict outcomes correctly. However, in other cases the link can be tighter, for example monitoring species' populations gives a relatively good estimate whether biodiversity is maintained or enhanced. Whether outcome or output indicator is more appropriate will vary depending on the situation.

The indicators can also differ based on the four approaches towards certification of ES (figure 1). In the "forest stewardship only" model, the focus will be on management activities rather than on specific ES whereas in the other models the ES will also be monitored. For the first model and to a certain extent for the second as well, the forthcoming International Generic Indicators (IGI) (FSC 2013) will give guidance on what are the management indicators linked to the forest stewardship under FSC model. As the "performance based" model addresses aspects such as additionality and leakage, indicators linked to them should be included to evaluate the impacts of them. Finally, in the "reward based" model indicators that address the reward mechanism and its impacts should be included. The idea is not to assess progress in all possible indicators. Rather it is envisioned that the most appropriate indicators for each locality are chosen and progress assessed against them. Appendix 2 gives examples for possible impact indicators as developed in workshops with stakeholders in the pilot countries.

Finally, an important part of impact evaluation will be to record/estimate the opportunity costs of forest protection. Also, the costs towards ES certification should be recorded in order to carry out cost-benefit analyses.

Step 4. Assessing the impact: design of an impact evaluation and data collection

The design of an impact evaluation

The different approaches towards certification of ES introduced earlier (Figure 2) should be considered as they influence the design of an impact evaluation. First, in the forest stewardship approach where the knowledge that forest is managed according to the FSC forest stewardship standard is enough, an impact evaluation with a counterfactual (e.g. control site or group, model)[30] is not strictly necessary. However, it is encouraged until there is large enough body of evidence that FSC certification makes a difference to the state of the forest. Similarly in the second approach where the ES are monitored on top of responsible forest management an impact evaluation with control sites is also not strictly necessary but strongly encouraged to build up the evidence base that FSC certification makes a difference to the provision of ES. In the other two models that are performance based an impact evaluation with counterfactuals is necessary to prove the performance.

Who designs an impact evaluation?

Ideally the impact evaluation should be designed together by buyers and sellers of the services as well as other stakeholders that have interest in the results. These include but are not limited to forest managers, local communities, companies (e.g. water companies), contributing funding agencies, and NGOs that have provided support or have otherwise contributed to the activities. The participatory approach is meant to provide different perspectives and ensure that the most relevant questions that correspond to the needs of the stakeholders are addressed.

When to evaluate the impact? Prospective or retrospective evaluation.

An impact evaluation can either be prospective, which means that it is designed at the same time than the project and built into the project design or it can be retrospective, which means that the treatment and control groups are generated after the intervention happened. The best approach to evaluate the impact of an intervention (e.g. certified forest management) is to include an evaluation into the management plan from the beginning as it is more likely to produce credible and strong results for three reasons[31]. First, baseline data can be collected to evaluate impact later on. Without baseline data prior to intervention it is more difficult to infer impact and attribute it correctly as evaluations are context specific: they describe certain effects under certain circumstances at certain times[32]. Second, there are more options to establish counterfactuals before any activities are carried out. This will ensure that proper variation is generated in the areas or people that are exposed to the project to avoid selection

bias. The avoidance of selection bias is important in order to be able to correctly attribute an impact to an intervention. However, it should be noted here that because certification is voluntary and participants self-select to participate in it, a positive selection bias is likely to occur in many cases[33]. Third, creating a theory of change will put the focus on the intended results.

Establishing counterfactuals

One of the key decisions is to decide whether and what kind of counterfactual is needed. The counterfactual can be an actual area or a group of people or it can be based on modeling. What is the most suitable option will depend on the ES certification approach chosen (i.e. whether it is performance based), fund availability, and the questions the impact evaluation tries to answer.

When counterfactual sites are established, it will be necessary to think how and where to collect data. The questions that should be considered include following:

- Are permits needed?
- Is it necessary to hire a person or persons to do data collection on the counterfactual sites?
- Are the people in the area willing to participate as controls?
- What are the most suitable areas for data collection if randomized data collection is not possible?
- What are possible sources of bias that should be avoided?

Experimental designs and randomization

If prospective evaluation is planned, it is possible to use an experimental design. Experimental designs are impact evaluations with randomization included in them. By nature they are planned prospectively and the results of the intervention are compared to a counterfactual before and after the intervention takes place. There are five main ways to create experimental variation in project designs[34]:

1. *Simple randomization.* The candidates (e.g. areas or people; see the endnote for more comprehensive explanation)[35] are randomly selected to control and one or more treatment groups.

2. *Randomization in oversubscribed projects*. When there are more willing and eligible people to participate than the project can accommodate, the project can use lottery to select the candidates.

3. *Randomized phase-in projects*. When the project will be phased in over time, a lottery can be used to select the order in which the eligible and willing candidates will enter the project. Those candidates that enter the project later will be used as a control. This is probably the most suitable way to create experimental variation in the context of ForCES as all willing participants can in the end participate in the project.

4. *Randomized encouragement*. Instead of randomizing the candidates into treatment and control groups the effort that is used to encourage the candidates to participate in the project is randomized, i.e. some candidates are encouraged more to participate in the project than others.

5. *Discontinuous eligibility criterion.* Rather than randomizing the candidates, the project selects an eligibility criterion, such as cut-off score, that creates control and treatment groups around the cut-off score that are supposedly relatively similar in their characteristics.

Quasi-experimental designs: Before-After and Control-Intervention designs

If experimental project design is not feasible, then the next best approach is to use quasi-experimental[36] methods[37]. The most commonly used quasi-experimental method by project proponents is the Before-After (BA) design whereas in the scientific literature Control-Intervention (CI) designs dominate[38]. In the before-after comparisons data on relevant impact indicators is collected before and after an intervention. Thus, the pre-intervention information forms a control to which data after the intervention is compared. It is assumed that changes detected in the indicators are due to the intervention and thus, conditions in which the intervention takes place stay stable[39]. However, in reality environmental or macroeconomic conditions rarely stay stable over the years, for example market prices fluctuate. Some of the variation can be controlled but this may require extensive data collection on the factors that can influence the outcomes of the project. Another assumption is that the outcomes before treatment are not influenced by anticipation of the intervention. These two assumptions make before-after comparisons most suitable to situations where the causal link between action and impact is simple, for example when the water fetching time is compared before and after building a well in the village.

In the control-intervention designs the outcomes (or outputs) are compared between control and intervention areas/groups after the intervention has already taken place. The intervention areas/groups are those that are directly impacted because of the intervention, e.g. forest management unit where the management practices are changed due to certification. The control areas/groups are those that are not directly influenced (and they should not be indirectly influenced either) by the intervention. The assumption behind control-intervention design is that the control areas/groups are similar enough compared to the intervention areas/groups that the outcomes on both areas/groups would be similar in the absence of the project. In practice it is challenging to find areas/groups that would be a perfect match in all aspects and influenced otherwise similarly by contextual factors expect for the impact of intervention.

To improve the quasi-experimental impact evaluation design before-after and control-intervention designs could be used in combination, i.e. the BACI design referred under the experimental designs. Of course the BACI design itself does not remove the lack of randomization but it adds a layer of information (the before intervention data) that helps to validate the outcomes. In effect, even if randomization took place, the participants in control and impact groups would not be completely similar as there are no mirror images in the real world. However, in expectation and on average the participants would be similar as they were drawn randomly from the same population[40]. Quasi-experimental methods seek to replicate this by selecting a pool of controls that, in expectation and on average, are

comparable to the intervention areas/participants. There are various methods to create the controls[41] and hence, time should be devoted to think what the most suitable method is.

Data collection

Who collects the data?

The forest manager or the management committee should think how the data collection should be done and by whom. If communities are involved in the management activities, then a participatory approach towards data collection is recommended[42]. It is important to remember that impact indicators and compliance indicators may overlap and thus, some of the data for impact evaluation may be collected as part of normal monitoring activities. In a case where counterfactuals are established it is also important that data that is sensitive to variation due to the data collector (e.g. bird counts) is collected by same person (or at least it is checked that the data do not differ significantly) in both control and counterfactual sites to avoid bias.

Sample size

The final step before implementing an impact evaluation is to carefully design data collection. As the quality of impact evaluation depends directly on the quality of data, attention should be paid to the sample sizes to ensure they are adequate. For example, the number of households to be interviewed or the number of hours to count birds.

There is a statistical technique called power calculations to estimate the smallest possible sample size that will allow the differences in outcomes (or outputs) between control and intervention groups to be detected[43]. Usually a larger sample size is needed for quasi-experimental than for experimental designs. Also, larger sample size is required if the outcome indicator is highly variable or a rare event, the minimum detectable effect is small, or impacts are compared between subgroups. Estimating the required sample size will also help to estimate the budget required for data collection.

Timing of data collection

Another issue to consider regarding data collection is seasonality. Thus, the timing of data collection should take into account natural variability of indicators to ensure that the impact is not missed. Depending on the subject, monitoring can either be continuous or it can be point measures taken at certain times. For example, migrant bird species are absent during certain times of a year and hence, monitoring of their population should take place when they are present.

Collecting data on people

If the impact evaluation involves collecting data on people, it should be ensured that those people are protected from any kind of risk that participating in the impact evaluation can cause. The Belmont Report[44] defines commonly accepted principles for ethical subject research as:

1. Respect for persons, i.e. respect individual's autonomy and protect those with diminished autonomy.
2. Beneficence, i.e. do not harm as well as maximize possible benefits and minimize possible harms (secure the individual's well-being).
3. Justice, i.e. fair distribution of benefits and burdens of the research.

There are two important requirements that follow from these principles. First, persons should be given an opportunity to decide whether to participate in the evaluation (especially important regarding the control group) and hence, a free, prior, and informed consent (FPIC) for data collection should be acquired before start of the activities. Second, confidentiality of the information provided should be ensured, especially if a person's wellbeing might be at risk for revealing the information. In practice this means that the data should be coded in a way that hides the identity of the participants and other relevant personal information, such as the name of village where they live.

Questions that should be clarified before data collections start:

- Purpose of data collection?
- What data will be collected?
- Who collects the data?
- Where will the data be collected?
- How will the data be collected (i.e. data collection methods)?
- When will the data be collected?
- Is seasonality taken into account?
- Are the sample sizes adequate?
- Will data be collected on people? If yes, how are their rights ensured?
- How will the data be used – who has access to the info (confidentiality)?
- Will the data and evaluation results be shared with the people affected?
- Who owns the data?

The role of an auditor

In the FSC system the role of an auditor is to evaluate whether the management activities comply with the certification standard. If the capacity of the auditor allows, it would make sense to include assessment of the conducted impact evaluation into the scope of an annual surveillance audit[45]. Basically, the auditor would be provided the documentation for an impact evaluation and he/she would verify whether it is satisfactory in methods and scope. In the performance based model whether performance is at the required level would also be evaluated.

The auditor should also check whether there is a risk in a case of self-evaluation that the results are inflated[46]. Also, the auditor should check that the implementation of activities was properly conducted. It is important to know about the implementation to differentiate between situations where the outcomes where not achieve because the implementation was poor and situations where the outcomes where not achieved because the activities implemented do not lead to the desired outcomes[47].

The questions answered should include following:

- Does the theory of change depict adequately the causal chain for change?
- Who developed and who endorsed the theory of change in question?
- Is it based on correct assumptions?
- Are the risks adequately described?
- Are the contextual factors adequately included?
- Was the selection of counterfactuals proper?
- Were the methods for data collection appropriate?
- Was the scope of the evaluation adequate regarding the evaluation questions?
- Is there a risk that self-evaluation has influenced the results?
- Were the activities properly implemented?

References

Bottrill, M.C., M. Hockings, and H.P. Possingham 2011. *In Pursuit of Knowledge: Addressing Barriers to Effective Conservation Evaluation.* Ecology and Society **16**(2).

Brandon, K.E. and M. Wells 1992. *Planning for people and parks: Design dilemmas.* World Development **20**(4): p. 557-570.

Bruijnzeel, L.A. 2004. *Hydrological functions of tropical forests: not seeing the soil for the trees?* Agriculture, Ecosystems and Environment **104**: p. 185-228.

de Groot, R.S., et al. 2010. *Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making.* Ecological Complexity, **7**(3): p. 260-272.

Ferraro, P.J. 2009. *Counterfactual thinking and impact evaluation in environmental policy*, in *Environmental program and policy evaluation. New Directions for Evaluation.*, M. Birnbaum and P. Mickwitz, Editors. p. 75-84.

Ferraro, P.J. 2012. *Experimental Project Designs in the Global Environment Facility: Designing projects to create evidence and catalyze investments to secure global environmental benefits. A STAP Advisory Document.* Washington, D.C., United States of America.

Ferraro, P.J. and S.K. Pattanayak 2006. *Money for nothing? A call for empirical evaluation of biodiversity conservation investments.* PLoS Biology **4**: p. 482-488.

ForCES. 2013 [cited 2013 25.1]; Available from: http://forces.fsc.org/.

Frondel, M. and C.M. Schmidt 2005. *Evaluating environmental programs: THe perspective of modern evaluation reseach.* Ecological Economics (55): p. 515-526.

FSC 2007. *Strengthening Forest Conservation, Communities and Markets. The Global Strategy of The Forest Stewardship Council*. Forest Stewardship Council, Bonn, Germany.

FSC 2011. *Strategy Paper: Strategic Framework for an FSC Climate Change Engagement*. Forest Stewardship Council, Bonn, Germany.

FSC 2012. *FSC guidelines for the implementation of the right to free, prior and informed consent (FPIC)*. Forest Stewardship Council, Bonn, Germany.

FSC 2013. *FSC International Generic Indicators. FSC-STD.01-004 D1-0 EN.* Forest Stewardship Council, Bonn, Germany.

Gertler, P.J., et al., *Impact evaluation in practice*. 2010, Washington DC: The World Bank. 262.

Greig, L.A., et al. 2003. *Scoping for cumulative effects evaluation. Prepared for Indian and Northern Affairs Canada Environment Directorate, Whitehorse, Yukon Territory*. Richmond Hill, Ontario. p. 51.

ISEAL Alliance 2010. *Code of Good Practice for Assessing the Impacts of Social and Environmental Standards Systems. P041 – Version 1.0.* ISEAL Alliance.

Jagger, P., et al. 2010. *A guide to learning about livelihood impacts of REDD+*. Center for International Forestry Research (CIFOR): Bogor, Indonesia.

Marmorek, D., et al. 2011. *Fraser River sockeye salmon: data synthesis and cumulative impacts.*, in *Cohen Commission Tech. Rep. 6.*, ESSA Technologies Ltd.: Vancouver, B.C. p. 273.

Millennium Ecosystem Assessment 2005. *Ecosystems and Human Well-being: Synthesis.* Washington DC.

Organisation for Economic Co-operation and Development (OECD)/Development Assistance Committee (DAC) 2002. *Glossary of key terms in evaluation and results based management.* The DAC Working Party on Aid Evaluation, OECD, Paris.

Pattanayak, S.K. 2009. *Rough Guide to Impact Evaluation of Environmental and Development Programs*, in *Working Paper No. 40-09*.

Pullin, A. and T. Knight. 2009. *Doing more good than harm - Building an evidence-base for conservation and environmental management.* Biological Conservation **142**(5): p.931-934.

Reisman, J., et al. 2004. *Theory of Change: A Practical Tool For Action, Results and Learning*. Seattle, United States.

Romero, C. et al. 2013. *An overview of current knowledge about the impacts of forest management certification. A proposed framework for its evaluation.* Occasional Paper 91. CIFOR, Bogor, Indonesia.

Steering Committee of the State-of-Knowledge Assessment of Standards and Certification. (2012). *Toward sustainability: The roles and limitations of certification.* Washington, DC: RESOLVE, Inc.

Stern, E., et al. 2012. *Broadening the range of designs and methods for impact evaluations. Report of a study commissioned by the Department for International Development.*, in *Working Paper 38*. Department for International Development, United Kingdom.

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. *The Belmont report: ethical principles and guidelines for the protection of human subjects of research*.

Therivel, R. & Ross, B. 2007. Cumulative effects evaluation: Does scale matter? *Environmental Impact Evaluation Review,* 27**,** 365-385.

Wunder, S., *Payments for environmental services: some nuts and bolts*. 2005. p. 24p.

van Noordwijk, M., Leimona, B., Tomich, T. P., Velarde, S. J., Kallesoe, M, Sekher, M., and Swallow, B. 2007. *Criteria and indicators for environmental service compensation and reward mechanisms: realistic, voluntary, conditional and pro-poor: CES Scoping Study Issue Paper no. 2.* ICRAF Working Paper **37**(37).

Appendix 1. A draft theory of change for Quang Tri, Vietnam (Thuy Tu 1 Village, Vinh Tu Commune)

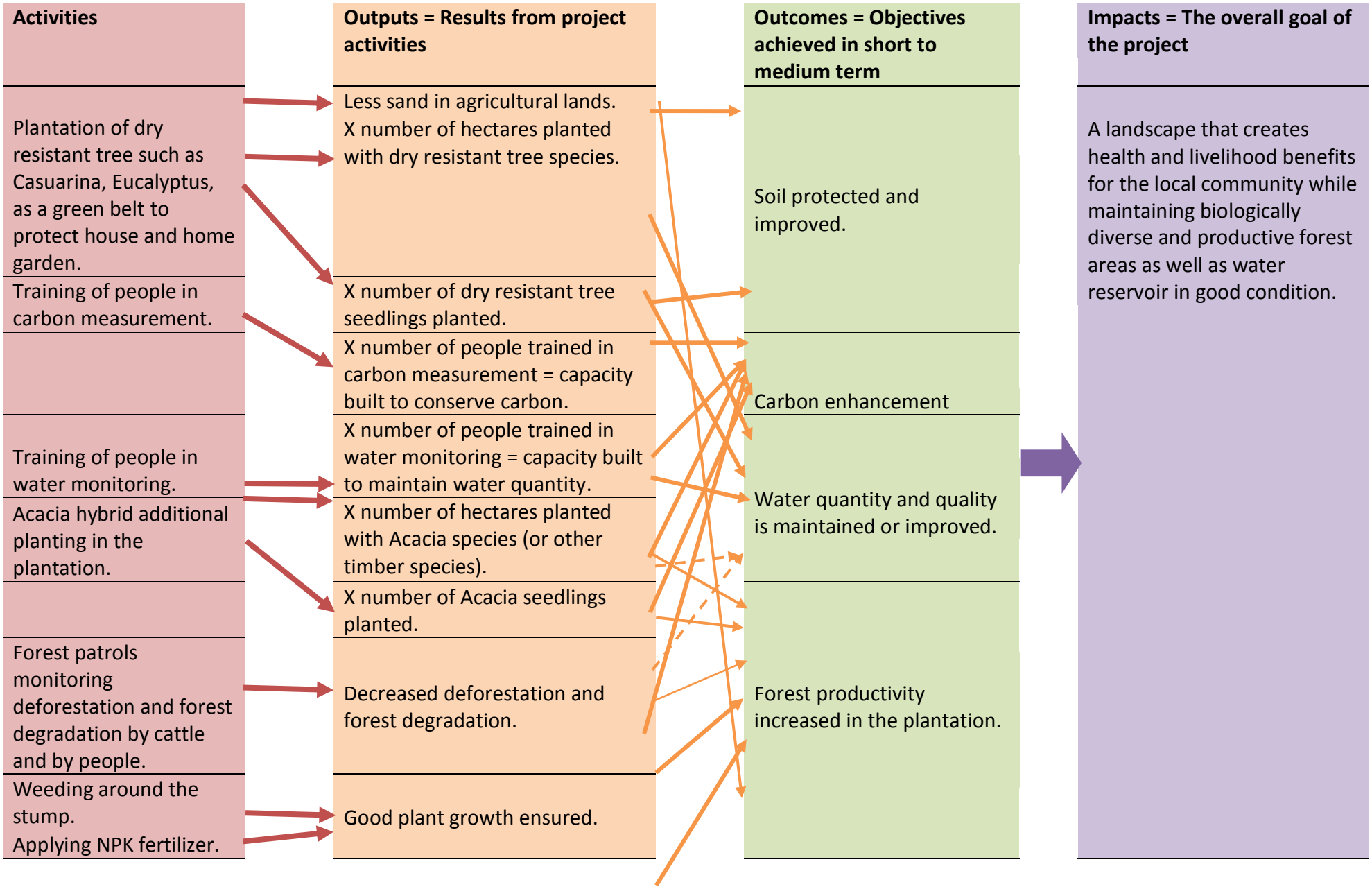
Management goals as stated in the management plan:

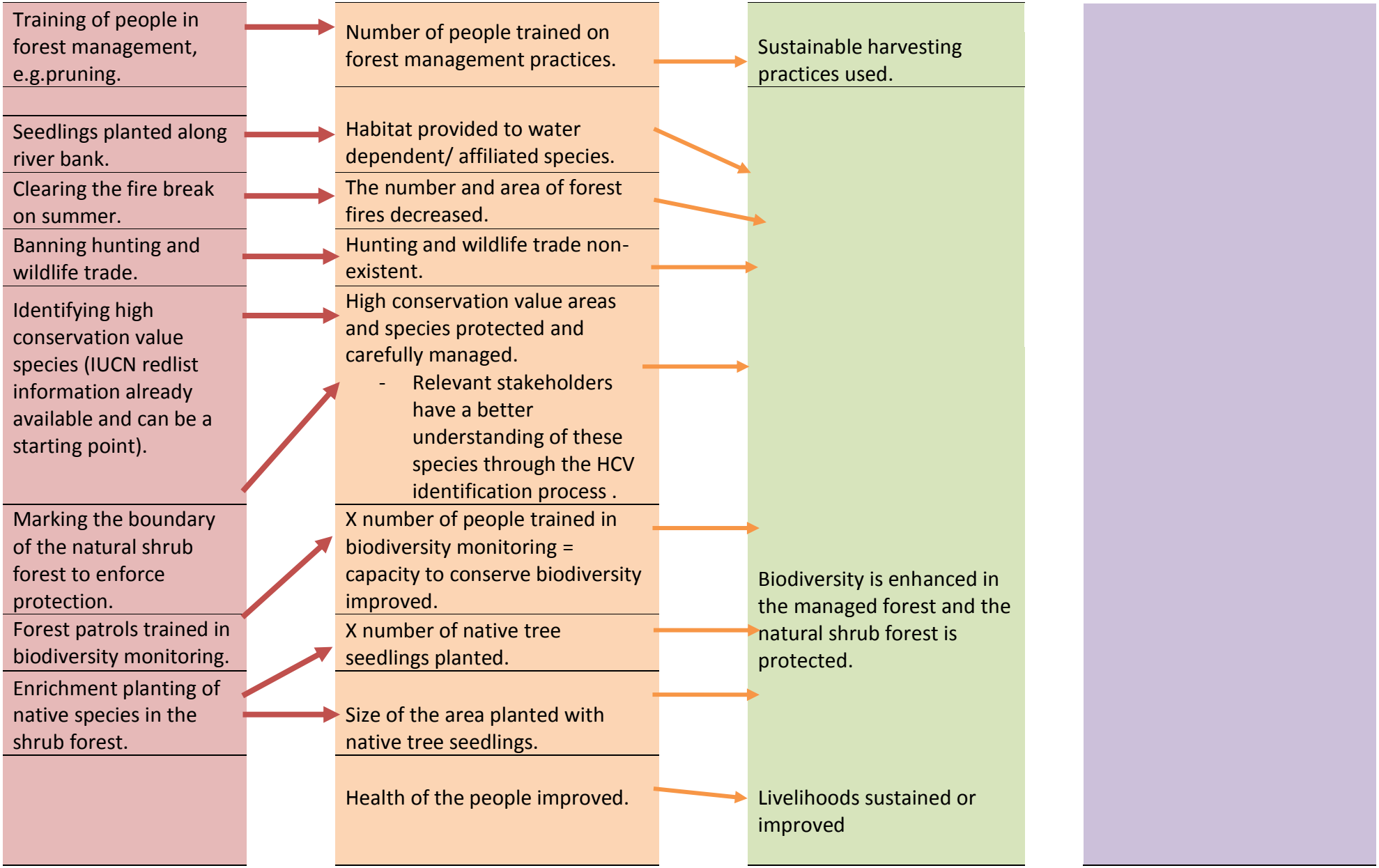
- Produce quality timber: 60% good quality logwood, 40% chipwood, trees not harvested below 10 years of age, harvestable diameter 16-20 cm
- Increase forest productivity and forest health against storms by limiting soil erosion and keeping the soil moisture.
- To ensure good plant growth, eliminate the competitive growth of weeds so that the plants get enough light and nutrition development.
- Ensure the growth of native trees.
- Prevent widespread forest fire.
- Prevent deforestation and forest degradation from cattle and forest destroyer.
- To create a diversity of wildlife resources on the forest plot.

Theory of change i.e. how the overall project impact can be achieved

The draft presented below (Figure 1) is a view of what the overall impact could ideally be and how change can be achieved in Quang Tri based on the knowledge available at the moment. It is only a starting point to develop a more comprehensive theory of change due course of ForCES project to evaluate the impact of the project. The purpose of the theory of change is to help the discussion when activities are planned so the activities listed are just examples of possible things that could be done (e.g. the trainings may not be necessary).

Regarding the question on how to integrate the services in the current management plant, a lot is already there, it just needs to be made explicit. For example, clearing of the fire break will (at least in theory) reduce the number and area of forest fires, hence contributing to carbon enhancement. Hopefully the figure below helps the discussion to make these connections. The information was gathered from the Village Forest Management Plan of Thuy Tu 1 Village, Vinh Tu Commune, from the CIFOR- SNV workshop, and the report entitled 'Mapping of Ecosystem services: Quang Tri and Ha Tinh Province' by Green Field Consulting and Development Co., Ltd (GFC). Context factors, inputs, assumptions or risks have not been included into the example below,


Activities
Outputs = Results from project activities
Outcomes = Objectives achieved in short to medium term
Impacts = The overall goal of the project
Plantation of dry resistant tree such as Casuarina, Eucalyptus, as a green belt to protect house and home garden.
Training of people in carbon measurement.
Training of people in water monitoring.
Acacia hybrid additional planting in the plantation.
Forest patrols monitoring deforestation and forest degradation by cattle and by people.
Weeding around the stump.
Applying NPK fertilizer.
Less sand in agricultural lands.
X number of hectares planted with dry resistant tree species.
X number of dry resistant tree seedlings planted.
X number of people trained in carbon measurement = capacity built to conserve carbon.
X number of people trained in water monitoring = capacity built to maintain water quantity.
X number of hectares planted with Acacia species (or other timber species).
X number of Acacia seedlings planted.
Decreased deforestation and forest degradation.
Good plant growth ensured.
Soil protected and improved.
Carbon enhancement
Water quantity and quality is maintained or improved.
Forest productivity increased in the plantation.
A landscape that creates health and livelihood benefits for the local community while maintaining biologically diverse and productive forest areas as well as water reservoir in good condition.


Training of people in forest management, e.g.pruning.
Seedlings planted along river bank.
Clearing the fire break on summer.
Banning hunting and wildlife trade.
Identifying high conservation value species (IUCN redlist information already available and can be a starting point).
Marking the boundary of the natural shrub forest to enforce protection.
Forest patrols trained in biodiversity monitoring.
Enrichment planting of native species in the shrub forest.
Number of people trained on forest management practices.
Habitat provided to water dependent/ affiliated species.
The number and area of forest fires decreased.
Hunting and wildlife trade non-existent.
High conservation value areas and species protected and carefully managed.
- Relevant stakeholders have a better understanding of these species through the HCV identification process .
X number of people trained in biodiversity monitoring = capacity to conserve biodiversity improved.
X number of native tree seedlings planted.
Size of the area planted with native tree seedlings.
Health of the people improved.
Sustainable harvesting practices used.
Biodiversity is enhanced in the managed forest and the natural shrub forest is protected.
Livelihoods sustained or improved

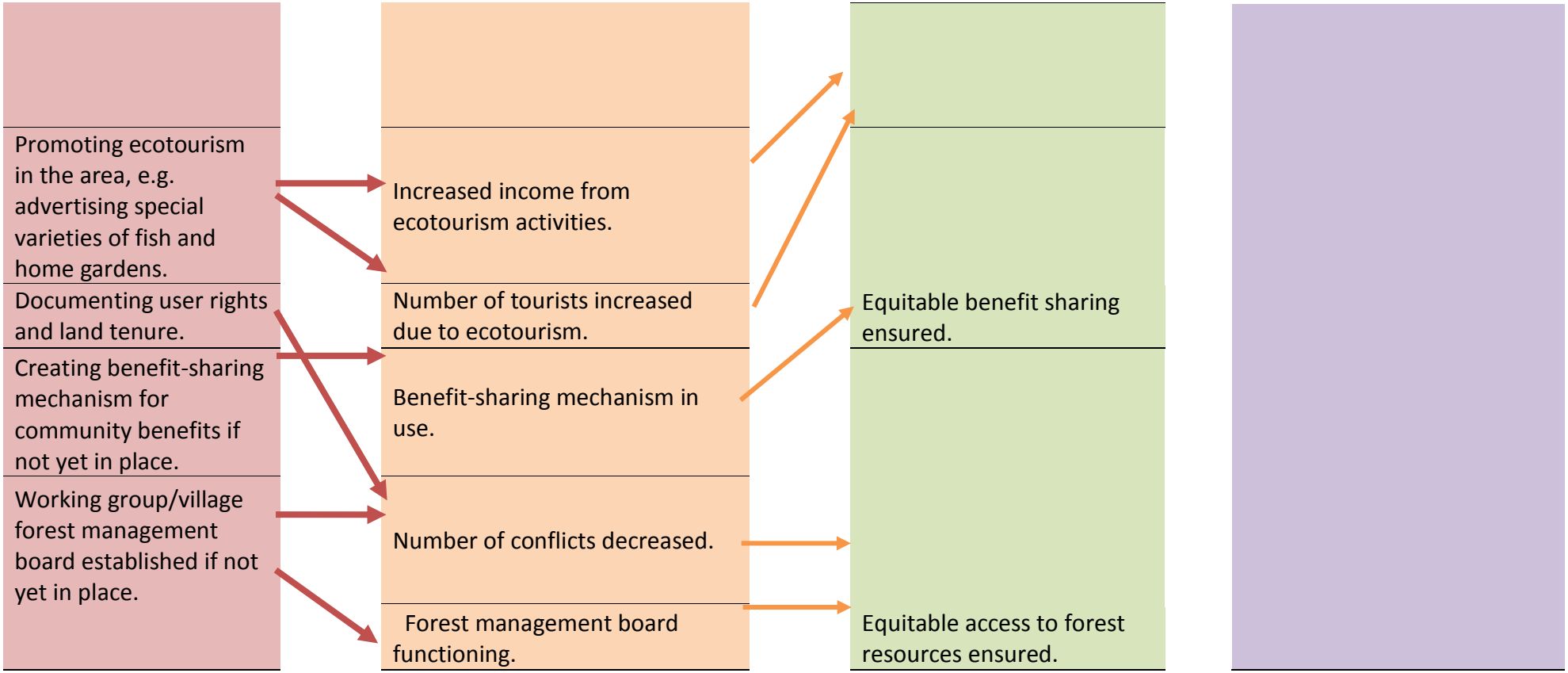


Figure 1. Theory of change for Quang Tri. The arrows represent pathway on how the impact can be achieved. The dashed arrows show a possible pathway but for which a certain amount of uncertainty is linked (hence the impact may not be as indicated in the figure). Not all the arrows are in place but the general idea of the impact pathways is shown.

Appendix 2. Possible impact indicators for different ForCES sites

The indicators presented below are based on indicators developed in participatory manner during the indicator workshops in 2012 and for Vietnam on the follow-up workshop held in July 2013. First are the indicators for each site and in the end are indicators that are either useful for all sites (e.g. equitable access to forest resources, gender) or apply only to specific situations (e.g. rewards and benefit sharing). The indicators can differ between sites for same service. In addition to the presented indicators the opportunity costs of forest protection and plantation management/development should be estimated. Also, the costs towards ES certification should be recorded.

Chile

Carahue: Biodiversity

Species level:

- Population size of endangered species and other valuable species for conservation stable or increasing.
- Population size of key species stable or increasing.
- Species richness maintained over time.
- Species composition maintained over time.
- Populations of invasive species are not present or are decreasing.

Stand level:

- Forest structure maintained over time.

Landscape level:

- Additionality in area conserved
- Connectivity maintained or increasing.
- Forest area maintained or increasing.
- Fragmentation stable or decreasing.
- Land use change does not cause decrease in biodiversity at species level.
- Land use diversity is maintained in diverse landscapes.

Chiloe: Hydrological services

- Turbidity.
- Water flow.
- The condition of vegetation (whether degrading, stable or improving).
- Deforestation rate is either zero or decreasing.
- Forest structure maintained over time.
- The number of incidents related to illegal deforestation and forest degradation.
- The number of native species' seedlings survived of those that were planted.
- The extent invasive species are used as substitute firewood species instead of native species.
- The number of people/households using good management practices.
- The extent erosion prone areas are restored.
- The extent of the buffer zone created around the reservoir to protect erosion prone areas.

Pumalin: Ecotourism

Environmental:

- Conservation of Alerce: number of hectares of Alerce preserved
- Forest structure maintained or improved over time.
- Number and size of annual forest fires
- Forest fire control plan exists and is in use.
- Contingency plan for natural disasters exists.

Economic:

- Number of visitors
- Number of local entrepreneurs related to the park
- Amount of income generated
- User satisfaction

Social:

- Number of conflicts related to e.g resource use and benefit-distribution.
- Number of guides employed.
- Number of work accidents
- System for accident treatment

Indonesia

PT Ratah: Carbon and biodiversity

Carbon:

- Quantity of biomass and forest carbon stock (ton/ha)
    - Above-ground biomass
    - Below-ground biomass
    - Soil carbon
    - Leaf litter
    - Wood debris
- Harvest volume
- Harvesting level
- Deforestation rate is either zero or decreasing.
- Forest structure maintained over time.

Biodiversity:

- Population trends of charismatic and endangered species.
- Variety of species.
- Population trends of lesser known species or species with potential to be charismatic species.
- Distribution of species.
- Extent and state of habitats provided for key species.

Social:

- Number of conflicts.

Lombok: Hydrological services

Environmental:

- Forest cover changes.
- Number of seedlings survived of those planted.
- The area planted.
- Number of springs dried out monthly/annually.
- Water quantity (monthly/annually).
- Chemical content.
- Turbidity.
- Pathogens and bacteria (E. Coli).

Social:

- Number of conflicts.

West Kalimantan: Ecotourism

Biodiversity:

- Population size and trends of key species (e.g. high conservation value species) stable or increasing.
- Number of sightings of key species per day.
- Number of bird species and their population trends.
- Number of dragon fish catch per area.
- Honey production per year.
- Deforestation rate zero or decreasing.
- Forest fragmentation rate zero or decreasing.
- Invasive species controlled or eradicated.
- Number and area of forest fires stable or decreasing.

Social and economic:

- Number of guides employed.
- Number of visitors who came to see a traditional event.
- Income from honey, by gender.
- Income from fish tourism, by gender.
- Income from other ecotourism activities, by gender.

Nepal

Charnawati: Carbon

- Quantity of biomass and forest carbon stock (ton/ha).
    - Above-ground biomass
    - Below-ground biomass

    - Soil carbon
    - Leaf litter
    - Wood debris
- Deforestation rate is either zero or decreasing.
- Forest structure maintained over time.

Gaurishankar: Ecotourism/recreation

Visitors:

- Number of visitors.
- Number of trekking routes.
- Expenditure per visitor.
- Satisfaction of visitors on different facilities and services.
- Value/price rating of visitors.
- Minimum safety conditions in place.
- Number of accidents, lootings, thefts.
- Complaints against inhumane behavior (e.g. towards the guides).

Environmental:

- Extent of pollution in key tourism destinations.
- Extent of deforestation and land use change.
- Number of poaching cases filed.
- Measures to control landslides and physical infrastructure in place.
- Measures for management of biological hotspots/key species habitats.
    - The condition of biological hotspots maintained or improved.
    - Populations of key species stable or increasing.

Local communities/local economy:

- Involvement of local communities to tourism activities.
- Income generated from tourism activities.
- Number of locally owned tourism related business enterprises per number of total tourism enterprises.
- Extent of use of local products.
- Number of days of local employment from tourism business.
- Extent of application of environmentally friendly measures on infrastructure construction.
- Number of cultural villages / natural heritage declared.
- Use of building code of construction.

Vietnam

Huong Son: Carbon with biodiversity and hydrological services

- The amount of total dissolved and suspended solids in water decreased compared to a baseline.
- The amount of carbon conserved due to improved harvesting practices and forest protection.
- The amount of carbon stored due to reforestation/afforestation activities, e.g. biomass of survived seedlings.

- Area set aside (no extractive activities) for conservation purposes.
- Area managed primarily for conservation purposes.
- Sightings or population trends of key species, e.g. species valuable for ecotourism, national breeds, HCV or endangered species.
- Species composition shows no change compared to a baseline and/or species composition changes towards desired state.
- The income of laborers.
- Hours of employment from forestry.

Quang Tri: Soil services and hydrological services

- The percentage of the border that is clearly marked**.**
- The area of the natural forest deforested (in hectares).
- The number of incidents related to illegal forest degradation, deforestation, illegal hunting and wildlife trade per month (based on the forest patrol records). Data should be aggregated at the end of each month.
- The amount of sand/topsoil in agricultural lands/residential areas.
- The area harvested per year (in hectares).
- The area replanted per year (in hectares).
- The number of native tree seedlings survived of those that were planted.
- The number and area of forest fires per year.
- The area under two seasonal paddy cultivation per year (in hectares).
- Status of the forest allocation process and rights (clarified (e.g. communicated clearly) and/or strengthened)
- The number of people trained in FSC and ES certification.
- Water quantity and quality, e.g. the amount of suspended solids in water decreased compared to a baseline and/or counterfactual area and changes in water level.
- The improvement of livelihoods due to provision of ecosystem services (e.g. fruit tree seedlings received as compensation -> fruit sold and the money used to send children to school).

Indicators shared by the sites depending on specific situations

Rewards (monetary and non-monetary):

- Benefit distribution:
    - Type of the mechanism.
    - Who among those certified receives benefits (e.g. do all or just part)?
    - Which entity receives benefits (e.g. community fund or individuals).
    - Percentage of benefits shared among different groups (women, poor, disadvantaged groups).
    - Percentage of benefits shared between community and government.
    - Composition of user committee (women, poor, disadvantages groups).
    - The amounts distributed.
    - The type of benefits distributed (e.g. cash, seedlings, etc.).

- Evidence of impacts of benefit distribution on, for example:
  - Creation of alternative/additional income sources
  - Education opportunities
  - Financial benefits (micro-credits)
  - Infrastructure provided (electricity, road, transportation, etc.)
- Amount of money invested in forest management activities.
- Amount of money invested in socio-economic activities.
- Existence of income generation activities of poor, women, disadvantage groups.
- Perceived change in wellbeing.
- Perceived change in economic situation since certification.

Equitable access to forest resources:

- Type of rights (e.g. informal/formal; user rights/tenure) exist regarding land and forest products (timber, NTFPs, and game)
- The status of rights (clarified (e.g. communicated clearly) and/or strengthened)
- Perceived security of tenure rights
- Number of conflicts reduced or kept at low levels.

Equal participatory rights:

- Public consultation (including) with potentially affected local stakeholders was properly conducted.
  - How and by whom the stakeholders were identified.
  - The percentage of the groups present in the area that participated in the consultations (i.e. representativeness).
  - The way (e.g. posters, invitation letters) how the identified local stakeholders were consulted.
  - The number of days between when the notice of stakeholder consultation was given and the actual consultation.
  - The way local stakeholders had an opportunity to influence decision making. Note: there may be no way to influence the decision making.
  - The existence of negotiation support to groups that have weak bargaining power (existent/non-existent/exists partially)
  - The share of willing stakeholders that had an opportunity to participate in the project. If not all could participate, what was the mechanism to determine who can participate?
  - Sharing of the outcome of consultations (Was the outcome shared locally? With what type of feedback? If needed on particular issues, how follow-up occurred?).

Gender:

- The share of women and/or other minority groups represented in the decision making bodies.
- The power of women and/or other minority groups to influence decision making (i.e. do they have influence on decisions?).
- The level of knowledge about forest management among women.
- The level of awareness among men regarding women's participation.
- Range of rights to forests and trees held by women and men.

- The number of sexual harassment cases reported.
- The number of gender discrimination cases reported.

Notes

[1] The Millennium Ecosystem Assessment (2005) categorizes the ecosystem services to *provisioning services* such as food, water, timber, and fiber; *regulating services* that affect climate, floods, disease, wastes, and water quality*; cultural services* that provide recreational, aesthetic, and spiritual benefits; and *supporting services* such as soil formation, photosynthesis, and nutrient cycling. Thus, goods are included as a type of service.
[2] The FCS founding assembly was held in 1993 and the FSC was established as a legal entity in 1994. PEFC was founded in 1999. The exception is ISO, which was established already in 1947.
[3] Brand and Wells 1992.
[4] Wunder 2005.
[5] REDD stands for "reducing emissions from deforestation and forest degradation in developing countries".
[6] e.g. The Climate, Community & Biodiversity (CCB) Standards released in 2005.
[7] See for example the FSC Strategy Paper (2011) on FSC climate change engagement. Also, the principle 6 of the FSC Principles and Criteria for Forest Stewardship on environmental values and impacts states that "The Organization* shall* maintain, conserve and/or restore ecosystem services* and environmental values* of the Management Unit*, and shall* avoid, repair or mitigate negative environmental impacts".
[8] FSC 2007.
[9] E.g. Ferraro and Pattanayak 2006, Pullin and Knight 2009. There are several initiatives currently underway that try improve the evidence base for conservation and development, e.g. the Conservation Measures Partnership http://www.conservationmeasures.org/, the International Initiative for Impact Evaluation http://www.3ieimpact.org/en/about/, the Collaboration for Environmental Evidence http://www.environmentalevidence.org/index.htm, and the Real World Evaluation http://www.realworldevaluation.org/Home_Page.html.
[10] ISEAL Alliance 2010.
[11] Gertler et al. 2010.
[12] Although the evaluation takes place in a discrete point of time, continuous monitoring may be needed to collect informative data for the evaluation purpose.
[13] FSC-STD-01-001 V5-0 D5-0.
[14] These business models were presented by Stefan Salvador from FSC at the Annual Project Manager's Meeting of the ForCES project on October 30, 2013.
[15] FSC 2007.
[16] Ferraro 2009, Stern et al. 2012, Gertler et al. 2010.
[17] FSC-STD-01-001 V5-0 D4-9.
[18] As explained in the FSC guidelines for the implementation of the right to free, prior and informed consent (FPIC) (FSC 2012) the community must be given the opportunity to form their opinion and make their decision before the proposed activity or project starts, and also before a final decision is taken that such activity or project will be implemented. In the ideal situation, the community must be asked about the initial idea before the project or activity is developed in detail.
[19] More ecosystem services may be found in the course of time as the range of ES users can be wider than initially thought.
[20] Sinks are biophysical features that can deplete service flows as defined by Bagstad et al. 2012 Spatial dynamics of ecosystem service flows: A comprehensive approach to quantifying actual services. Ecosystem services. DOI http://dx.doi.org/10.1016/j.ecoser.2012.07.012
[21] de Groot et al. 2010.
[22] van Noordwijk et al. 2007.

[23] The cumulative effects are defined as changes to the environment that are caused by an action in combination with other past, present and future human actions (Hegmann et al. 1999. Cumulative Effects Evaluation Practitioners Guide. Prepared by AXYS Environmental Consulting Ltd. and the CEA Working Group for the Canadian Environmental Evaluation Agency, Hull, Quebec). The 'environment' in the definition is meant in a broad sense (not just biophysical) as cumulative effects can also be social, for example decline of traditional skills or changes in social structures (Therivel and Ross 2007).
[24] Greig et al. 2003 in Marmorek et al. 2011.
[25] The definitions are from Bottrill et al. 2011, ISEAL Alliance 2010, Jagger 2010, OECD 2002, and Pattanyak 2009.
[26] Reissman et al. 2004.
[27] Pattanayak 2009.
[28] Jagger et al. 2010.
[29] FSC 2007.
[30] Counterfactual is needed to assess what would have happened without the intervention. Because the actual state cannot be observed, it can be estimated through a counterfactual which can be established through modeling, observing outcomes at a control sites, or using a control group (Jagger et al. 2010).
[31] Gertler et al. 2010.
[32] Ferraro 2012.
[33] Romero et al. 2013.
[34] For further information on creating randomization see Ferraro 2012.
[35] The candidates in these designs are either the actors whose behavior the project tries to modify (individuals, households, or areas) or an aggregation of these actors at higher level (larger areas, villages, administrative units).
[36] The term quasi-experimental refers to a project design in which assignment to *treatment* and *control* groups is not controlled by the project implementers, but which can, under certain assumptions, allow one to infer *causal effects* of the project.
[37] Ferraro 2009.
[38] Jagger et al. 2010.
[39] Frondel and Schmidt 2005.
[40] Jagger et al. 2010.
[41] For a detailed discussion see Gertler et al. 2010.
[42] Further resources http://www.participatorymethods.org/. Also, it should be noted that there are risks in self-evaluation, for example inflated results on positive impacts. See Steering committee of the state-of-knowledge assessment of standards and certification 2012.
[43] Gertler et al. 2010.
[44] The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research 1979.
[45] Although an entirely independent evaluation would be the ideal situation, for the system to be practical an impact evaluation verified by auditor is probably the most feasible solution.
[46] When participants collect data on themselves and on their activities, there is a risk for inflated results or falsified reports. See Steering committee of the state-of-knowledge assessment of standards and certification 2012 for examples and further discussion.
[47] Steering committee of the state-of-knowledge assessment of standards and certification 2012