GREEN CLIMATE FUND | Independent Evaluation Unit

TRUSTED EVIDENCE.
INFORMED POLICIES.
HIGH IMPACT.

LEARNING-ORIENTED REAL-TIME IMPACT ASSESSMENT (LORTA)

IMPACT EVALUATION BASELINE REPORT FOR FP087: BUILDING LIVELIHOOD RESILIENCE TO CLIMATE CHANGE IN THE UPPER BASINS OF GUATEMALA'S HIGHLANDS

February 2022

GREEN CLIMATE FUND
INDEPENDENT EVALUATION UNIT

Learning-Oriented Real-Time Impact Assessment (LORTA)

IMPACT EVALUATION BASELINE REPORT FOR FP087: BUILDING LIVELIHOOD RESILIENCE TO CLIMATE CHANGE IN THE UPPER BASINS OF GUATEMALA'S HIGHLANDS

02/2022

CONTENTS

TABLES

FIGURES

PREFACE

In 2018, the Independent Evaluation Unit initiated the Learning-Oriented Real-Time Impact Assessment (LORTA) Programme, in collaboration with the Center for Evaluation and Development (C4ED), project teams funded by the Green Climate Fund (GCF), local evaluation teams and academics. The LORTA programme incorporates state-of-the-art approaches for impact evaluations to measure results and raise awareness about the effectiveness and efficiency of GCF projects. This impact evaluation of the Participatory Integrated Climate Services for Agriculture was designed to align with the LORTA approach for measuring causal impacts, as implemented in Malawi.

The LORTA programme has a twofold aim: (a) to embed real-time impact evaluations into funded projects so GCF project task managers can quickly access accurate data on the project's quality of implementation and likelihood of impact; and (b) to build capacity within projects to design high-quality data sets for overall impact measurement. The purpose of the impact evaluations is to measure the change in key result areas of the GCF that can be attributed to project activities. The LORTA programme informs on the returns of GCF investments and helps GCF projects track implementation fidelity. The IEU selects projects/programmes in coordination with the Secretariat, which further participates in the implementation of LORTA for learning purposes.

Acknowledgements

We thank key contributors from the International Union for Conservation of Nature (IUCN), including Tania Ammour of the IUCN Regional Office for Mexico, Central America and the Caribbean, Raquel Sigüenza of the IUCN Guatemalan Office, and Pedro Pineda of the University Rafael Landívar, for their vital support. We also thank the entire team of Sintaxis for their hard work in collecting data, including Daniela Herrera Duarte, Jonathan Samuel Reynoso Duarte, Silvia Patricia Villatoro Canahui and Mario Noel Mansilla (the specialists); Alejandra Fabiola Morales Rodas, Juan Pablo Lima Rodríguez and Hesler Agusto Morales Rodas (the supporting team); and Rebeca Cabrera Escobar, Eva Maribel Capriel Herrera, Lucy Antonieta Ixmucané Capriel Herrera, Manuel Esteban Tambriz Catinac, Yuly Nohely Cocón Mucia, Herlinda Guarcax Churunel, Dolores Mendoza Sosof, Elida Marleny Noemí Ramírez, Sonia Yaneth Ramírez Ramírez, Luz Aurora Ramos Vásquez, Josefina Benicia Sánchez López and Norberto Felipe Yojcom Puac (the team of enumerators).

At C4ED, we thank Esther Heesemann for her work on this project up until the end of the baseline data collection, and Giulia Montresor for her valuable feedback.

The authors would also like to thank Deborah Kim of the Independent Evaluation Unit (IEU) for her comments.

LIST OF AUTHORS

The authors of this approach paper on Building livelihood resilience to climate change in the upper basins of Guatemala's highlands are (in alphabetical order by last name):

FULL NAME	AFFILIATION
Dafne Edith Domínguez	International Union for Conservation of Nature Guatemala
Jakob Gärtner	Center for Evaluation and Development
Adolfo Ottoniel Monterroso Rivas	International Union for Conservation of Nature Guatemala
Martin Prowse	Independent Evaluation Unit, Green Climate Fund
Viviana Urueña	Center for Evaluation and Development

FOREWORD

This document is the baseline report for the impact evaluation of the project FP087 “Building livelihood resilience to climate change in the upper basins of Guatemala’s highlands”. The overall goal of this project is to contribute to reducing the impact of climate change in the western highlands of Guatemala, by improving the management of ecosystems and water in the watershed, as well as improving the capacity of social and institutional responses to climate events. The project focuses on the maintenance of the hydrological cycle, by promoting water and soil management and conservation practices and ensuring the restoration and provision of ecosystem services that contribute to adaptation. The project seeks to achieve three outputs. The first one is to develop integrated climate-smart watershed management; the second one is to promote community-led watershed management systems through grant facilities; and the third output is to provide climate-related information to farmers and target stakeholders for improved watershed management.

This baseline report outlines the theory of change of the intervention as well as the evaluation questions and indicators. It also describes how the data collection was conducted. The report also describes the baseline characteristics of beneficiaries and shows the results between the beneficiary and comparison households that were part of the evaluation’s treatment and control groups. The report was drafted by IUCN, the accredited entity for this project, and by C4ED. The GCF IEU wrote and revised several sections of the draft report and provided feedback and valuable comments on the evaluation design, indicators and data analysis, that were duly included in this final version.

ABBREVIATIONS

C4ED	Center for Evaluation and Development
CSA	Climate-smart agriculture
DiD	Difference-in-differences
EWS	Early warning system
FCG	Guatemalan Foundation for Environment and Natural Resources Conservation
GCF	Green Climate Fund
IARNA	Institute of Research and Outreach on the Natural Environment and Society, Guatemala
ICC	Intra-cluster correlation
IEU	Independent Evaluation Unit
INAB	National Forest Institute, Guatemala
INSIVUMEH	National Institute of Seismology, Volcanology, Meteorology and Hydrology, Guatemala
IUCN	International Union for Conservation of Nature
KMO	Kaiser-Meyer-Olkin
LORTA	Learning-Oriented Real-Time Impact Assessment
MAGA	Ministry of Agriculture, Livestock and Food of Guatemala
MARN	Ministry of Environmental and Natural Resources of Guatemala
MDES	Minimum detectable effect size
PCA	Principal component analysis
PINPEP	Forestry Incentive Program for Small Landowners with Forestry or Agroforestry Vocation *(Programa de Incentivos Forestales para Poseedores de Pequeñas Extensiones de Tierra de Vocación Forestal o Agroforestal),* Guatemala
PROBOSQUE	Promotion of the establishment, rehabilitation, restoration, management, production, and protection of forests in Guatemala *(Fomento al establecimiento, recuperación, restauración, manejo, producción y protección de los bosques en Guatemala)*
ToC	Theory of change
URL	University Rafael Landívar (*Universidad Rafael Landívar)*, Guatemala

EXECUTIVE SUMMARY

The baseline assessment of the project FP087: "Building livelihood resilience to climate change in the upper basins of Guatemala's highlands" presented in this document, outlines the theory of change of the intervention as well as the evaluation questions and indicators; it also describes how the data collection was conducted. Moreover, the report describes and compares the baseline characteristics of beneficiaries and shows the results of comparisons between beneficiary and comparison households in treatment and similar nearby micro watersheds (control areas).

The FP087 project focuses on reducing the impacts of climate change on the hydrological cycle in target watersheds in Guatemala's highlands through improved land-use practices, community grants and an early warning system. The project aims to improve the resilience and livelihoods of the population, strengthen their capacity for adaptation, and reduce their exposure to climate risks through the use and application of tools, information and practices that are either completely or partially climate related. The accredited entity for this project is the International Union for Conservation of Nature (IUCN), with offices at the regional level, headquarters in Gland, Switzerland, and a national office in Guatemala. The project is co-financed by the GCF and Korea International Cooperation Agency, and by the Government of Guatemala (in-kind and cash funding).

The baseline allows an impact evaluation of the following two subcomponents: C1.1: micro watershed management and ecosystem-based adaptation training; and C3: the early warning system.

The main research questions to be answered by the impact evaluation, derived from the theory of change, are as follows:

1) To what extent are farmers in the intervention area more resilient / less vulnerable to extreme weather events?
2) To what extent did the intervention lead to better awareness and knowledge of climate-smart agriculture by farmers?
3) To what extent did the intervention lead to the implementation of activities related to climate-smart agriculture by farmers?
4) To what extent did the intervention lead to the diversification of crops by farmers?

To answer these questions, two indices were created for this baseline report:

- A resilient and diversified livelihoods index (*Índice de medios de vida diversificados y resilientes*)
- A responsiveness index (*Índice de Capacidad de Respuesta*)

Both of these indices are composed of numerous indicators that cover multiple dimensions. The resilient and diversified livelihoods index measures the economic resilience of livelihoods for individuals who are exposed to climate shocks. The responsiveness index captures different levels of the response capacity of households and communities to the effects of climate change.

Due to the non-random selection of the treated communities, only quasi-experimental methods can be used to estimate the project impacts. After consultations with the project team, the difference-in-differences (DiD) with matching method was selected as the most robust method to evaluate the impact of the project. The proposed methodology makes it possible to identify the effects of the intervention on the indices and on the individual indicators. The unit of analysis for this baseline report is the household, as all activities target households. To measure the impacts of the two interventions outlined above, endline data will be collected on households in both treatment and control watersheds.

Baseline data collection was conducted during Q1 and Q2, 2021. During data-collection preparation, the questionnaire and training material were reviewed by IUCN and the Learning-Oriented Real-Time Impact Assessment (LORTA) team composed of staff from the GCF Independent Evaluation Unit (IEU), and the Center for Evaluation and Development (C4ED). IUCN and the LORTA team also carried out daily monitoring of data-collection activities. The total sample size is 1,486 households, distributed into 21 treated (758 households) and 13 control (728 households) micro watersheds.

As expected, comparisons between treatment and control households show both groups are very similar: households and communities have almost the same vulnerabilities, gaps and needs regarding improving their resilience. For instance, socioeconomic characteristics are very balanced between both groups: on average, households are considered as poor, and education and gender variables show no significant differences. Conversely, agricultural practices do present some variability: treatment households showed more experience of agroforestry systems and silvopastoral systems than control households. There is another systematic and significant difference between both groups: on average, treatment households showed more knowledge on responses to climate change than control group households. However, despite 17 per cent of all households having faced climate-related shocks, both groups expressed the same limited perception of knowledge on climate change events and consequences.

The baseline data-collection exercise and analysis highlighted a number of challenges: for example, due to the ongoing COVID-19 pandemic new communities had to be selected during the fieldwork to replace the communities that chose not to participate in the survey. Moreover, there is also a risk of spillovers from treatment to control micro watersheds, and the potential for confounding interventions within the limited number of treatment or control micro watersheds that are part of this evaluation. These challenges will be monitored carefully as project interventions scale up and we move towards and beyond endline data collection.

I. CONTEXT

A. Regional context

Guatemala is a developing country classified as an upper-middle-income nation, bordering Mexico in the North, with Belize to the east, and El Salvador and Honduras to the south. Despite its low but steady economic growth in the past decades, the country still suffers from high rates of poverty and malnutrition, especially in rural areas.

Figure 1. Map of Central America

Project implementation is in the central and western highlands of Guatemala. The population in the highlands is made up mostly of indigenous Mayan people (84 per cent of households identified themselves as indigenous) and comprises a variety of linguistic communities. The largest share of the population in this area consists of members of the K'iche' community, followed by the Kaqchikel, Mam and Spanish communities. Individuals living in the highlands are mainly small-scale agricultural producers who rely on natural resources for subsistence and their livelihoods. The vulnerability of these families is exacerbated by unsustainable agricultural practices, such as deforestation, land degradation and, to a lesser extent, slash-and-burn agriculture. Such agricultural practices decrease soil fertility until the soil becomes infertile, requiring the acquisition of new land, which can lead to deforestation.

The highlands of Guatemala are between 1,800 and 3,300 metres above sea level, with annual minimum and maximum temperatures between 10°C and 18°C, and an average of 15°C. Deforestation in the highlands contributes to soil erosion, the alteration of water flows in the watershed, loss of habitat for flora and fauna and, in general, changes in the provision of ecosystem services. Climate variability is modifying the rainfall and temperature patterns in the western highlands. In some years, an excessive level of humidity has been reported, and in others, the highlands have suffered from drought conditions, a phenomenon influenced by the El Niño Southern Oscillation (Giorgi, 2006; Aguilar and others, 2005). Changes are also reflected in the average number of dry and wet days in the highlands, such as a higher number of dry days followed by a limited number of days with higher rainfall intensity (International Union for Conservation of

Nature [IUCN], 2021). The main threats derived from climate variability and climate change in the project area have been identified as follows:

- **Increase in the intensity of precipitation:** causes natural hazards of a geomorphological nature, such as processes of surface erosion, mass movements, river floods and changes in channels and alluvial plains, which can affect housing, infrastructure and inhabitants
- **Change in the duration of dry periods:** fewer consecutive days with rain, which are interrupted by longer dry periods
- **Increase in temperature:** both during the day and at night, which means there is a greater demand for evapotranspiration and, as a result, an increased demand for water
- **Increase in frosts:** cold periods tend to be more intense and, in some cases, more frequent, both during the day and at night

B. Policy context

Responses to climate change in Guatemala are regulated by the National Climate Change Policy and the National Development Plan: K'atun Our Guatemala 2032. Specific legal instruments, laws and regulations also help to define mitigation and adaptation programmes and priority projects. For example, projects are aligned with the National Action Plan on Climate Change, which defines national priorities. For adaptation, these include the use of nature-based solutions, and the conservation of ecosystem goods and services. A further priority is the use of new or improved tools or technologies for communicating climate risks such as floods or forest fires.

The PROBOSQUE[1] law focuses on adaptation actions through conservation and aims to increase the country's forest cover with the creation and application of an incentive programme. This law focuses on commercial plantations, natural forests for production or protection, restoration and agroforestry systems. A further forest incentive programme focused on smallholders is PINPEP,[2] which is based on its own law and regulations. Figure 2 illustrates the linkages between the National Development Plan 2032 and these associated policy frameworks and laws.[3]

Figure 2. Laws and regulations that directly support the actions of the project

[1] PROBOSQUE: Promotion for the establishment, recuperation, restoration, management, production and protection of forests in Guatemala (*Fomento al establecimiento, recuperación, restauración, manejo, producción y protección de los bosques en Guatemala).* PROBOSQUE is a law approved in 2020 in Guatemala to protect the forests in the country.
[2] PINPEP: Forestry Incentive Program for Small Landowners with Forestry or Agroforestry Vocation *(Programa de Incentivos Forestales para Poseedores de Pequeñas Extensiones de Tierra de Vocación Forestal o Agroforestal).* It is an initiative from the National Forest Incentive Programme.
[3] In addition, this project acts in accordance with the National Policy of Integral Rural Development in order to achieve sustainable human development in rural areas.

II. PROJECT (INTERVENTION) DESCRIPTION

The FP087 project will restore 22,500 hectares through sustainable land-use systems that will improve the provision of ecosystem services, primarily linked to the management and conservation of water and soil, as well as watershed management. The land uses to be promoted include agroforestry and silvopastoral systems, natural forest management systems and restoration-focused forest plantations. The expected direct beneficiaries are 132,000 people, of which at least 30 per cent will be women, with a special emphasis on the participation of single-parent female-led households. The project will also work directly with Mayan indigenous peoples from the linguistic communities Kaqchikel, K'iche' and Mam, as well as with rural youth groups.

The impact that is sought in these communities is to improve the resilience and livelihoods of the population, strengthening their capacity for adaptation, as well as reducing their exposure to climate risks through the use and application of tools, information and practices that are either completely or partially climate related. The project is co-financed by the GCF and Korea International Cooperation Agency, the accredited entity for this project is the IUCN who collaborates closely with the Ministry of Environment and Natural Resources (MARN), as the GCF's National Designated Authority in Guatemala. The project is implemented by the IUCN Guatemala Office, alongside the Foundation for the Conservation of Natural Resources and the Environment in Guatemala (FCG) and the Rafael Landívar University as executing entities. Additionally, the project works closely with the Ministry of Environmental and Natural Resources (MARN), Ministry of Agriculture and Livestock (MAGA), National Forest Institute (INAB) and National Institute of Seismology, Volcanology, Meteorology and Hydrology (INSIVUMEH) which provide in-kind and cash financing to the project alongside technical capacities and infrastructure. Moreover, the project will provide equipment for the operation of an EWS in the Western Highlands, and the technical capacities of INAB will be strengthened in relation to the incentive programme for ecological restoration and the provision of ecosystem services related to climate change adaptation.

The project consists of three project components that are implemented at community and watershed levels. In particular, the first component builds on existing institutional structures from the MAGA, including the Rural Development Learning Centers and agricultural extension workers, municipal forestry offices / environment units, the environmental education decentralized services of the MARN, and the local forestry extension support of the INAB. The second component channels funding from the Green Climate Fund (GCF) and the Korea International Cooperation Agency directly to community-based organizations through a medium and small grants facility, which will be managed in collaboration with the Guatemalan Foundation for Environment and Natural Resources Conservation (FCG). The third component focuses on climate information and the implementation of an EWS directed to farmers and communities. This is implemented in collaboration with the University Rafael Landívar through its Institute of Research and Outreach on Natural Environment and Society (IARNA).

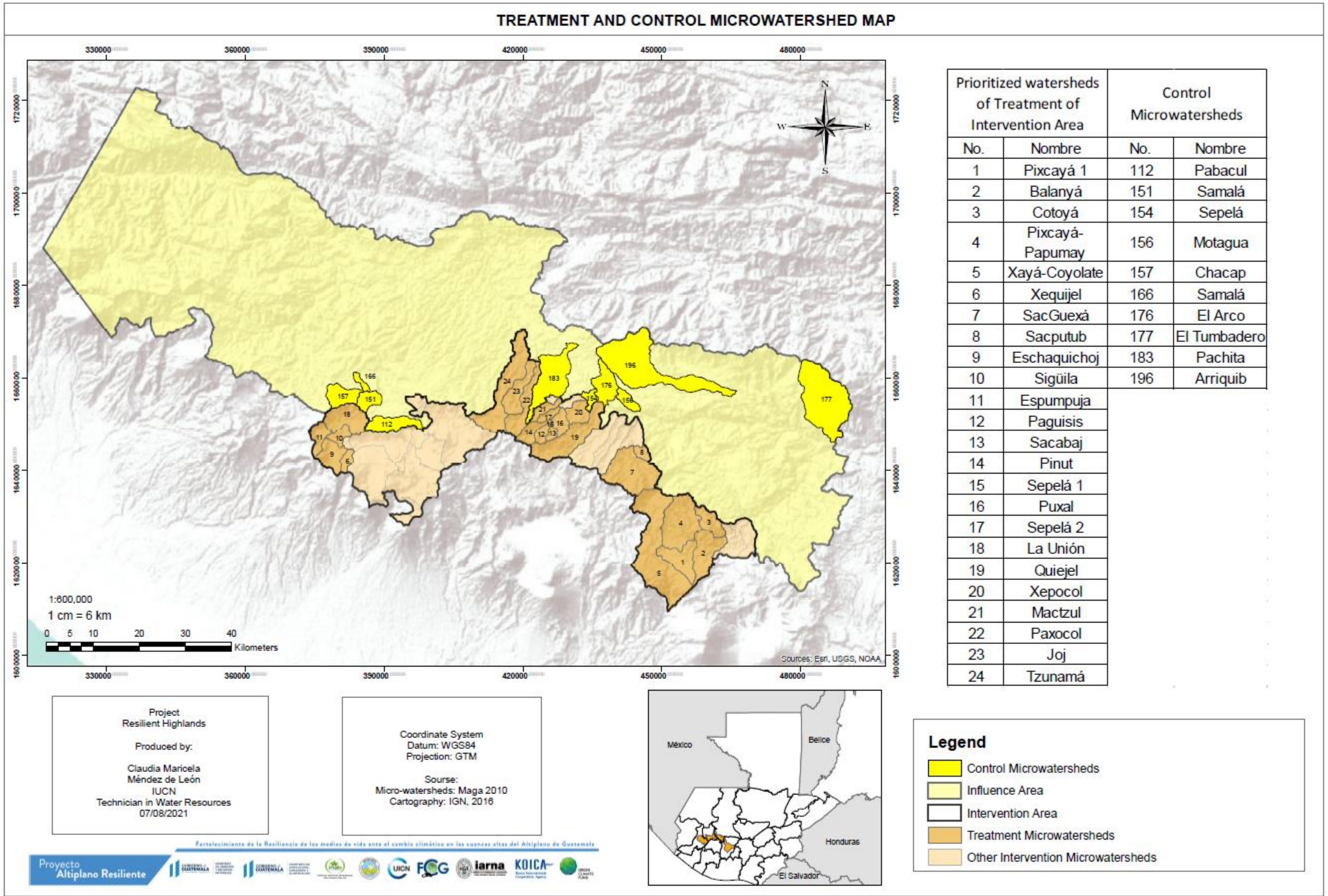
The project area is divided into two parts: the area of influence and the area of intervention. The area of influence of the project is in the upper watersheds of the western highlands of Guatemala and totals 7,673 km2 and 334 micro watersheds (see Figure 3, area in light yellow). Although the project will not implement direct actions in the entire area of influence, the outcomes and impacts of the project (including the ecosystem-based approach to adaptation, climate-smart watershed management, and the early warning system actions) can be replicated within the area of influence.

The project actions will focus specifically on the recharge areas of four watersheds – Motagua, Coyolate, Samalá and Chixoy – making up what is called the project area of intervention (see Figure

3, area in orange). The project area of intervention covers a total of 1,468 km2 of 48 micro watersheds. The actions of the three components will be implemented in this targeted area. However, due to financial constraints, the impact evaluation covers the four watersheds but only 21 of the micro watersheds. As mentioned above, the “prioritized” micro watersheds were selected because of their high level of vulnerability. More specifically, the project prioritized 24 micro watersheds that cover 858 km2 of the intervention area, in which specific watershed management activities will be carried out, such as the formation of micro watershed committees, governance activities and the promotion of forest management plans, among others. However, during the baseline data collection, not all 24 prioritized micro watersheds were included, and data were collected in 21 prioritized micro watersheds (treatment) and additional 14 micro watersheds outside of the area of intervention (which form the control areas).

The project implementation period is seven years from 8 April 2020, the effective date of the funded activity agreement. In order to achieve the main goals, the project has ensured the development of a series of activities that are expected to be implemented over that period. In 2020, the project started enhancing the capacity of extension workers (from INAB, MAGA, etc.). This training will continue until 2026. During 2021, the project concentrated its efforts on designing the EWS and on developing the micro watershed development plans. In Q1 2022, farmers will start receiving capacity training, and early warning information will start to be transferred to the communities. The main deliverables and some important milestones expected over time are detailed in Figure 4.

Figure 3. *Map of treatment and control micro watersheds*



Prioritized watersheds of Treatment of Intervention Area		Control Microwatersheds	
No.	Nombre	No.	Nombre
1	Pixcayá 1	112	Pabacul
2	Balanyá	151	Samalá
3	Cotoyá	154	Sepelá
4	Pixcayá-Papumay	156	Motagua
5	Xayá-Coyolate	157	Chacap
6	Xequijel	166	Samalá
7	SacGuexá	176	El Arco
8	Sacputub	177	El Tumbadero
9	Eschaquichoj	183	Pachita
10	Sigüila	196	Arriquib
11	Espumpuja		
12	Paguisis		
13	Sacabaj		
14	Pinut		
15	Sepelá 1		
16	Puxal		
17	Sepelá 2		
18	La Unión		
19	Quiejel		
20	Xepocol		
21	Mactzul		
22	Paxocol		
23	Joj		
24	Tzunamá		

Figure 4. Timeline for main intervention, results and activities

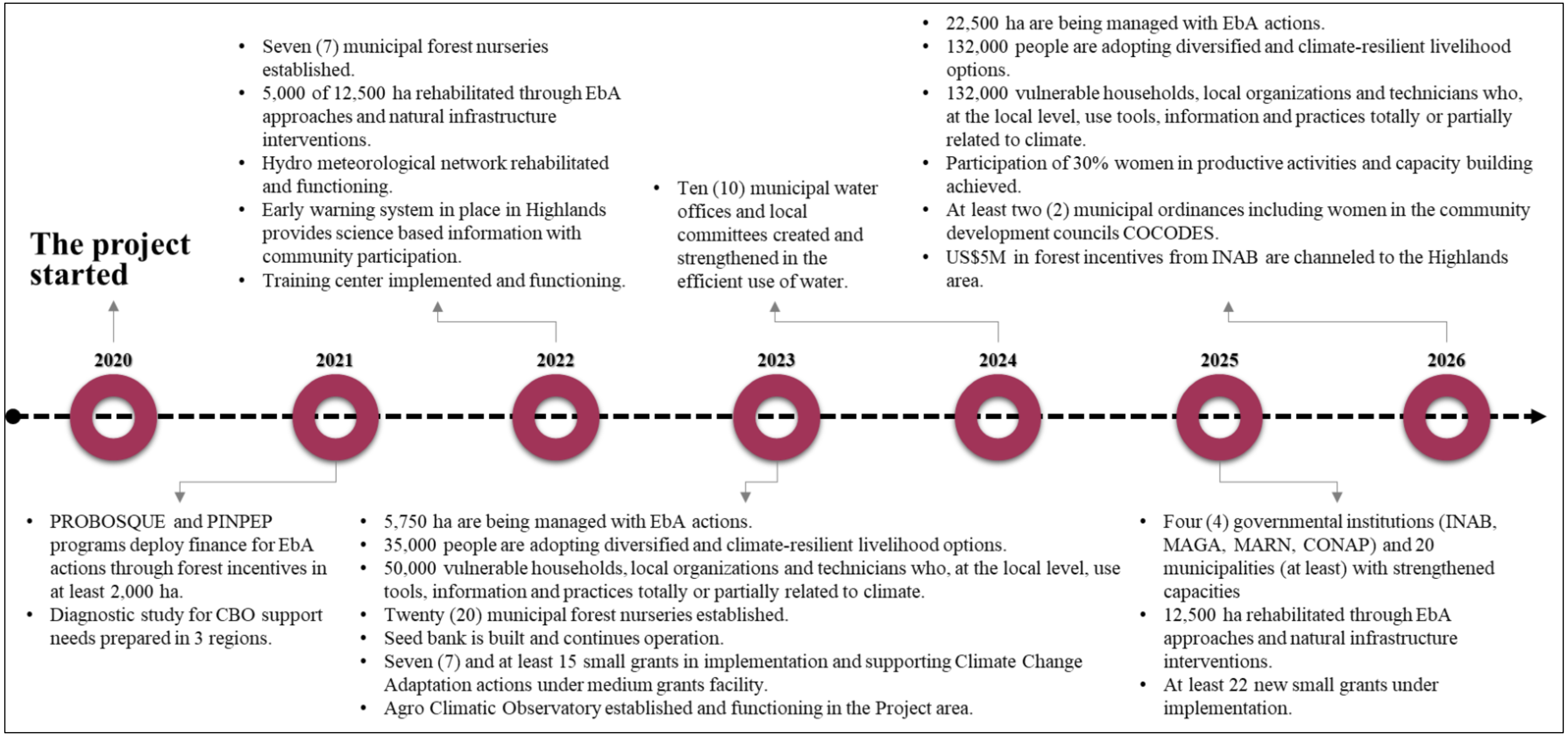
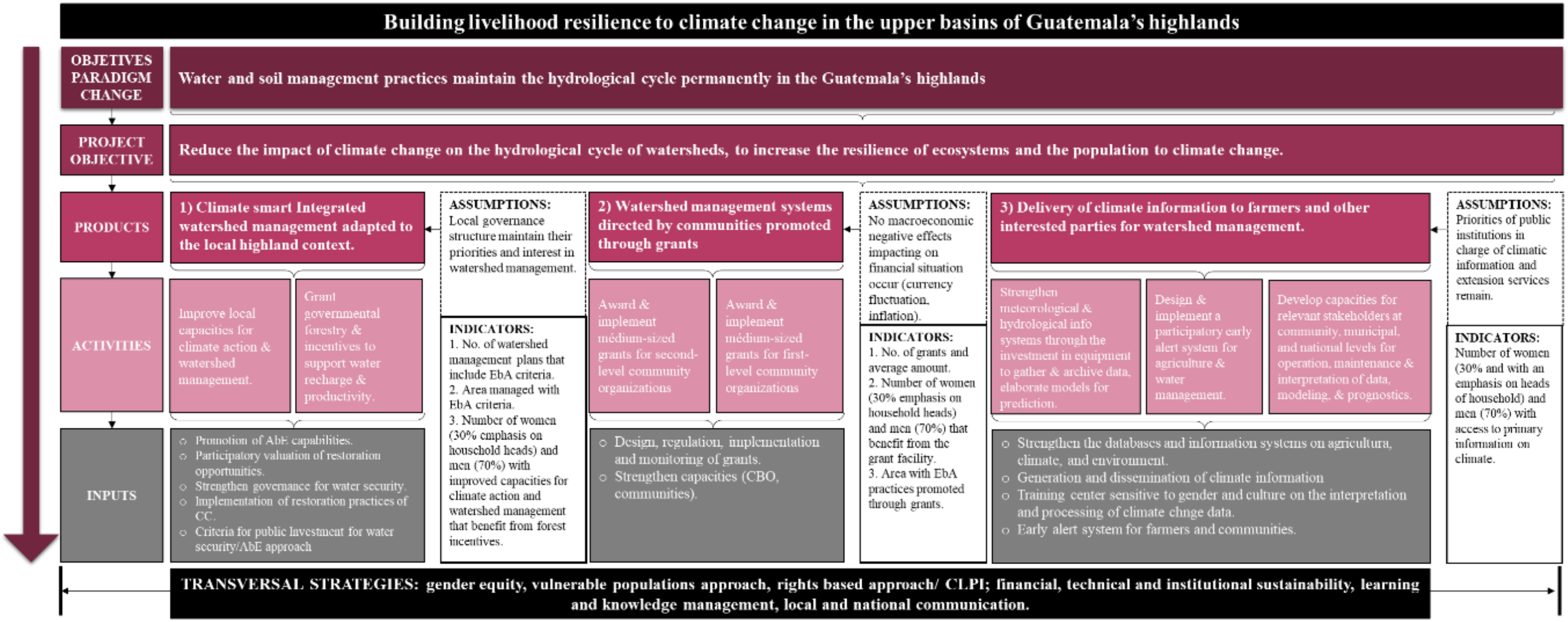


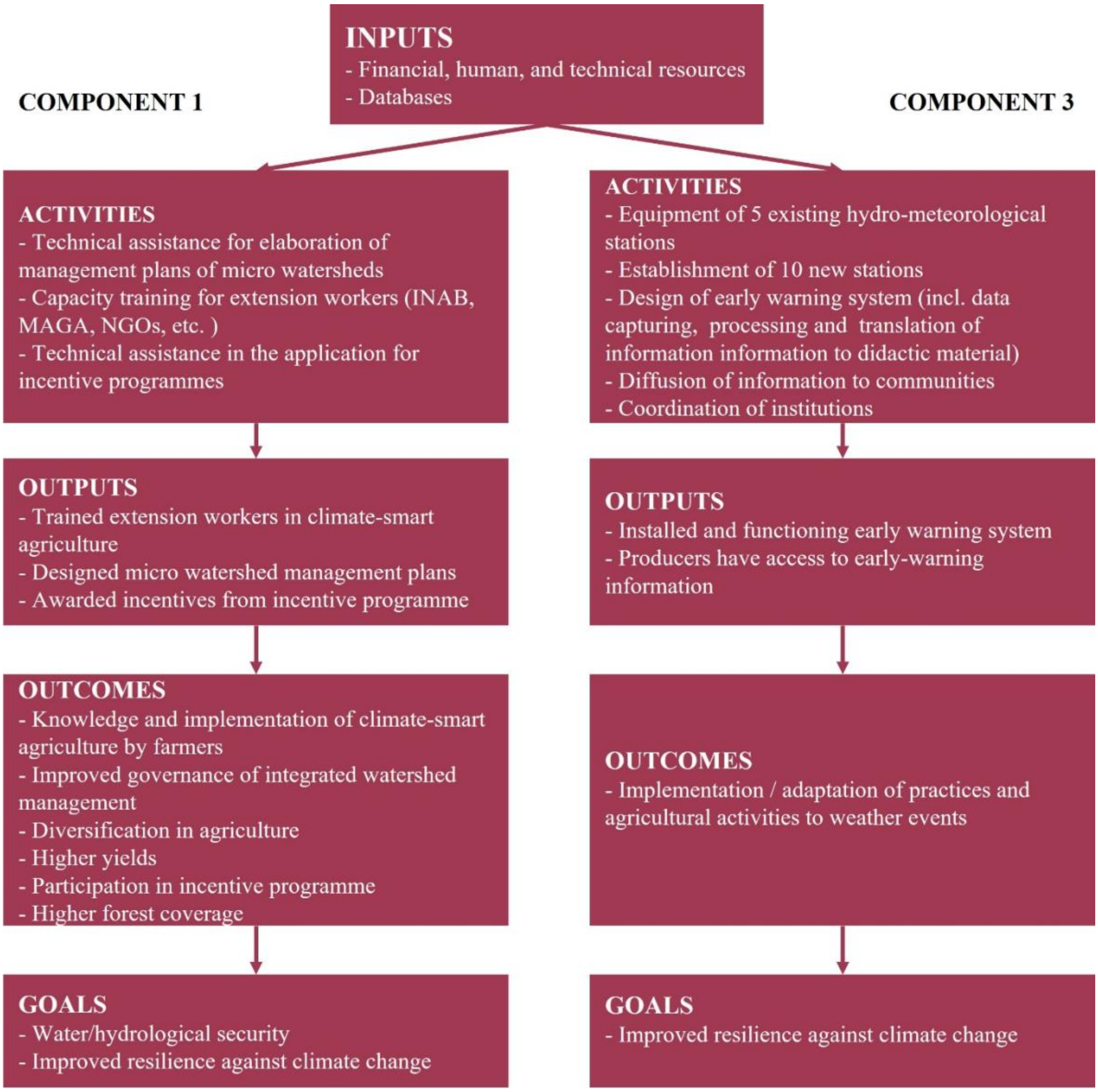
Figure 5. Theory of change of the project intervention



Theory of change

The project is complex in its activities and targets multiple levels (households and micro watersheds) in the intervention area (see Figure 5). Because of this complexity, the impact evaluation will be conducted on subcomponent C1.1: micro watershed management and ecosystem-based adaptation training and component C3: early warning systems. Both parts are related to households, which are our unit of analysis. Subcomponent C1.1 and component 3 have the potential to affect all households in the project area, in particular those with farming activities. Furthermore, the activities will be fairly homogeneous across micro watersheds, which facilitates identifying key evaluation outcomes. The theory of change (ToC) for the first project subcomponent and the third component is depicted in Figure 6. These ToCs rely on several assumptions that were discussed extensively during the impact evaluation design workshop in Guatemala and during meetings that followed the inception visit, as a basis for the design of the baseline.

Figure 6. Theory of change, C1.1 and C3



III. EVALUATION QUESTIONS AND INDICATORS

A. EVALUATION QUESTIONS

The main research questions to be answered by the impact evaluation, derived from the ToC, are listed below:

1) To what extent are farmers in the intervention area more resilient / less vulnerable to extreme weather events?
2) To what extent did the intervention lead to better awareness and knowledge of climate-smart agriculture (CSA) by farmers?
3) To what extent did the intervention lead to the implementation of activities related to CSA by farmers?
4) To what extent did the intervention lead to the diversification of crops by farmers?

To answer these questions, two indices were created. These indices speak to each of the evaluation questions and will be estimated in this baseline (before the start of the intervention) and after the intervention. Key indicators will also be compared between treatment and control groups. Table 1 maps each evaluation question to the corresponding indicators. Some indicators speak to more than one of the evaluation questions.

The creation of the two main evaluation indices was based on the experience of the IUCN team, as well as on existing literature. Moreover, there are benefits in conducting the impact evaluation analysis based on few outcome variables to avoid problems from testing multiple hypotheses (as the likelihood of finding significant differences between treatment and control groups increases as one increases the number of hypotheses). If these indices had not been created, then corrections for testing multiple hypotheses (e.g. Bonferroni, Benjamin-Hochberg) would need to be applied. In addition, adding different variables makes outcome measurements less noisy because random errors are cancelled out.

Nevertheless, it is important to keep in mind that merging different variables into an index has the downside that the index will fail to explain what might be driving the results. As highlighted above, comparisons of key impact indicators will also be conducted to unpack what may be driving any observed differences.

Table 1. Evaluation matrix

EVALUATION QUESTION	INDICATORS
1) To what extent are farmers in the intervention area more resilient/less vulnerable to extreme weather events?	Multidimensional poverty, Climate change vulnerability, EWS in place at the community, Use of EWS information at the community
2) To what extent did the intervention lead to better awareness and knowledge of CSA by farmers?	Household perception of climate change risk, Household knowledge on responses to climate change effects
3) To what extent did the intervention lead to the implementation of activities related to CSA by farmers?	Water accessibility and quality, Collecting forest products, Household knowledge on responses to climate change effects, Agricultural diversification
4) To what extent did the intervention lead to the diversification of crops by farmers?	Agricultural diversification

B. EVALUATION INDICES

As mentioned above, this impact evaluation aims to measure the effects of the ecosystem-based adaptation training and the implementation of the EWS on two indices:

- A resilient and diversified livelihoods index *(Índice de medios de vida diversificados y resilientes)*, related to subcomponent 1.1
- A responsiveness index *(Índice de Capacidad de Respuesta)*, related to component 3

We now explain the composition of each index and offer a brief explanation of how each indicator was measured.

1. Resilient and diversified livelihoods index

This index measures resilience that is essentially economic: the resilience of livelihoods for individuals who are exposed to climate shocks. While this index gives an approximation of how vulnerable and resilient a household is, by proxying how well a household could cope with a climate-related shock, it does not predict how fast they could recover from it. The composition of this index was based on previous work by members of the IUCN team (see for example IUCN and CIAT, 2019) and was thoroughly discussed with the LORTA team. The structure of the index also matches available literature as it includes the indicators that capture the definition of economic resilience (see for example Food and Agriculture Organization, 2010). Table 2 shows the different indicators (and subcomponents) that compose this index, and their weights and characteristics.

2. Responsiveness index

The responsiveness index captures household and community capacity to perceive and respond to the effects of climate change. The index was tested in other projects implemented in the Guatemalan highlands by IARNA (see URL-IARNA, 2020), and adapted to this evaluation by IUCN Guatemala. It includes agricultural and natural resource management practices at the plot level, as well as community response capacity. This index is composed of four indicators and measures how responsive a household is towards climate risks and future shocks. This index also takes into account the existence and use of EWS at the community level. Hence, the indicator captures variation not only among households but also among communities. Table 3 describes the characteristics of the indicators that compose this index.

Table 2. Description of the resilient and diversified livelihoods index

Indicator	Subcomponents	Indicator type	Weighting
Multidimensional poverty (URL-IARNA, 2020)	– Education: Sum of education years completed by female and male head of household – Life quality: Points given to type of roof, floor, water access and quality – Income: Points given based on income and assets (economic strata, household size, TV, motorbike, car, etc.) – Food security: Household classification into four	Continuous (values between 0 and 100) Higher values indicate a lower poverty level	Education score 20; Life quality score 25; Income score 30; Food security score 25 = Total score 100. The score is given based on the performance of each indicator; for instance, if education = 20, then the household obtains 20 points in this subcomponent. The points attributed for each subcomponent were added and based on the total number the household was classified in:

INDICATOR	SUBCOMPONENTS	INDICATOR TYPE	WEIGHTING
	categories: not food insecure, low food insecure, medium food insecure and high food insecure		– Very poor (0–30) – Poor (31–66) – Not poor (67–80) – Not poor at all (81–100)
Water accessibility and quality	Water accessibility	Discrete (values from 0 to 2) Higher values indicate better accessibility	N/A
	Water quality	Discrete (values from 0 to 2) Higher values indicate better quality	N/A
Collecting forest products	Number of products that are collected	Continuous (values between 0 and 1) Higher values indicate that more products are collected (out of 8 options that are provided)	N/A
Climate change vulnerability	Not affected by a climate shock	Binary Lack of money or food was due to a climate shock Lack of money or food was not due to a climate shock Higher value (=1) indicates that the household was not affected by a climate shock.	N/A
Agricultural diversification	Diversification index (Simpson's diversity index)	Continuous (values between 0 and 1) Higher values indicate more diversification	

Table 3. Description of responsiveness index

INDICATOR	DESCRIPTION	VARIABLE TYPE
Household perception of climate change risk	Proportion of answers (out of 10) marked with a "yes" regarding households' perceptions/beliefs of damages that climate change can cause them (crop loss, plagues, less agricultural productivity, etc.)	Continuous (values from 0 to 1) Higher values indicate higher perceptions of risk towards climate change.
Household knowledge on responses to climate change effects	Proportion of options mentioned by the participant (out of 13) regarding households' knowledge on measures/strategies to minimize the negative effects from climate change (crop diversification, water storage, soil conservation, etc.)	Continuous (values from 0 to 1) Higher values indicate more knowledge.
EWS in place at the community	Average of answers marked with a "yes". The average was calculated per community.	Continuous (values from 0 to 1) Higher values indicate that more people in the community

INDICATOR	DESCRIPTION	VARIABLE TYPE
		know about an EWS that is in place
Use of EWS information at the community	Average of answers marked with a "yes". The average was calculated per community.	Continuous (values from 0 to 1) Higher values indicate that more people in the community use the EWS that is in place

IV. EVALUATION STRATEGY AND DESIGN

Randomization was not an option as the intention of the project is to target the most vulnerable micro watersheds (those where quality and availability of water is relatively low). We refer to these as “prioritized micro watersheds”, whereas control micro watersheds not part of the intervention area are referred to as “non-prioritized micro watersheds”. Due to the non-random selection of the treated communities, only quasi-experimental methods could be used. Matching or DiD methods were viable options. A combination of DiD with matching method was selected to evaluate the impact of the subcomponents, as it represents a more robust method than applying either technique on its own.

The DiD technique makes it possible to estimate treatment effects via the comparison of changes in outcomes over time between a treated and a control group. A key assumption of the DiD technique in identifying treatment effects is that in the absence of the programme project, the outcomes for the two groups would have evolved in a similar fashion over time. This is known as the “parallel trends” assumption. Evidence can be provided on pre-project trends of relevant outcomes or impacts in treatment and control groups.[4] The DiD method controls for observed and unobserved time-invariant characteristics. Importantly, time-varying differences are not controlled for with this method and, if present, would undermine the unbiased estimation of the treatment effects. In addition to the DiD, micro watersheds and/or farmers in the treatment group will be matched with micro watersheds and/or farmers in the control group, based on observable baseline (pre-programme) characteristics (i.e. before project implementation). By ensuring balance in baseline characteristics between treatment and control group, matching increases the credibility of the parallel trend assumption.

The aforementioned design was proposed during Phase I of the LORTA programme, and after the baseline data collection it is still valid as it was possible to allocate watersheds between treatment and control groups. Figure 7 shows the composition of treatment and control groups based on the intervention. In order to reduce differences among treatment and control groups, and hold other components of the intervention constant, this evaluation focuses on households that do not receive any type of incentives from the project. The proposed methodology will make it possible to identify the effects of the intervention on the indices and indicators mentioned in the previous section. To measure the impacts, both baseline and endline data on outcomes are needed. However, attrition in both treatment and control groups will be a challenge for this impact evaluation. This issue has already been raised by the data-collection firm, given that they faced difficulties in finding and encouraging household participation during the baseline data collection, especially for the control group. We will provide more detailed information about this and other challenges in part C of this section.

4 The project team has collected panel data on treatment and control micro watersheds to test the parallel trend assumption.

Figure 7. Design C1.1 and C3

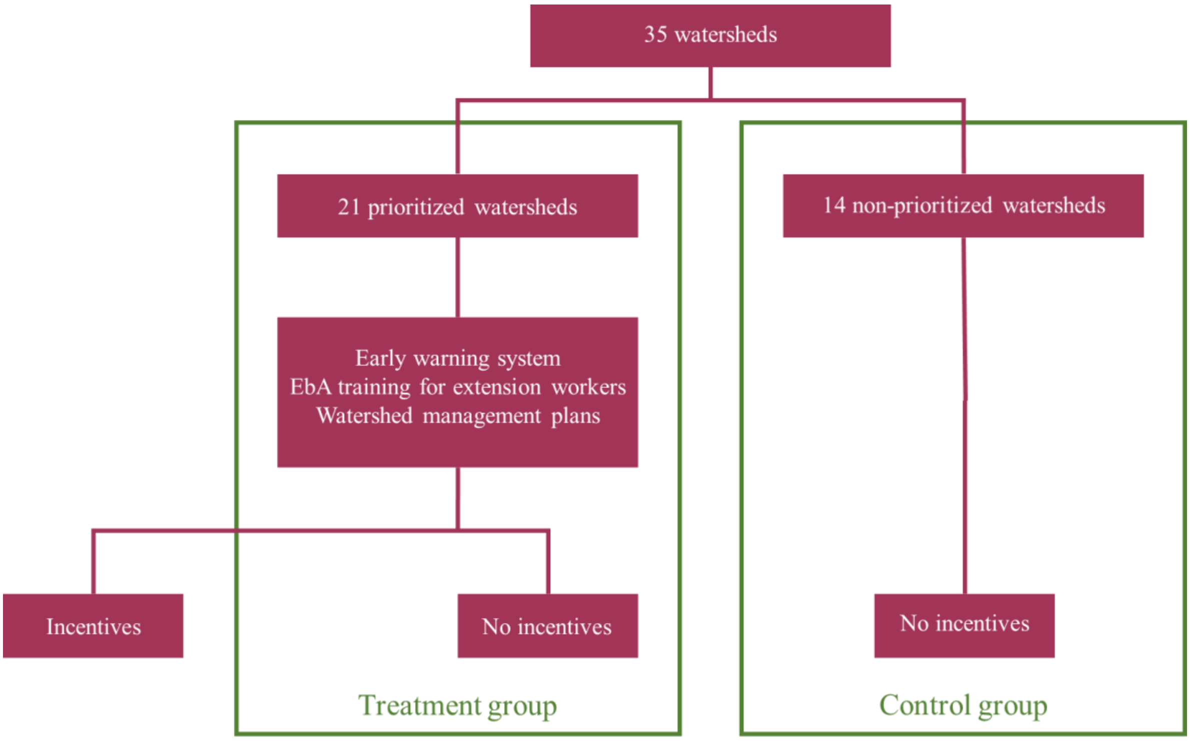
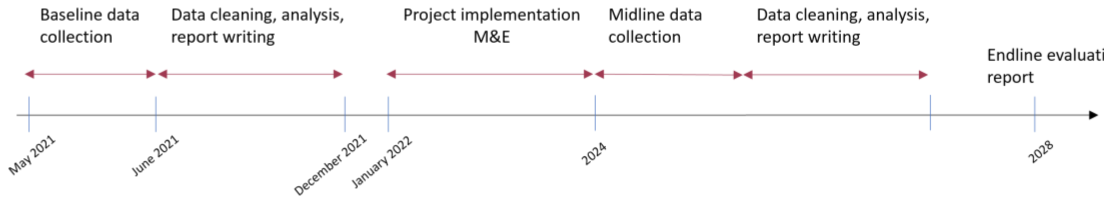


Figure 8 shows the evaluation timeline. Baseline data collection started in March 2021 and lasted until June 2021. The process of data cleaning, analysis and report writing was completed in December 2021. Although the project started in April 2020, the implementation of field activities related to the component C1.1 started on Q2 2021 (which mostly consisted of staff training sessions) and will start on Q1 2022 for component C3.[5] In Q1 2022, farmers will start receiving capacity training, and early warning information will start to be transferred to the communities. Midline data collection is planned for Q1 2024 and the endline report is planned for 2028, approximately seven years after the start of the intervention.

Figure 8. Evaluation timeline



A. Description of the units for decision making, the intervention and for analysis

The selection of the evaluation sample followed a two-stage cluster sampling approach. First, treated micro watersheds were randomly selected from a list of prioritized micro watersheds and the control micro watersheds were randomly selected from a list of non-prioritized micro watersheds. Second, a sample of households was randomly selected within each treated or control micro watershed.

[5] The total implementation period of the project is seven years. In 2020, the project started enhancing the capacity of extension workers (from INAB, MAGA, etc.) and designing the EWS. This training will continue until 2026.

All the micro watersheds are organized or clustered into four groups based on weather /ecological conditions, land use, social, cultural and political characteristics, and while they present similar characteristics to each other, there is still a high degree of variability between groups that needs to be considered in the analysis. Because of this, the present analysis accounts for such differences by clustering standard errors at the micro watershed level. For the endline report, the analysis will include heterogeneity analysis across groups (as it is possible that the intervention has different effects between groups).

Table 4. Distribution of respondents across groups and treatment allocation

Treatment allocation	Group				Total
	1	2	3	4	
Control	66	328	152	182	728
Treatment	65	287	192	214	728
Total	131	615	344	396	1,486

All activities that are part of our two interventions target households, and surveys are conducted with household heads (the survey questionnaire is available on request). An important element of the baseline assessment is subgroup analysis, where impacts of the interventions are expected to be assessed for single-parent and dual-parent households. In Guatemala, many vulnerable households have a single parent, frequently an indigenous woman, which is partly due to emigration to the United States.[6]

B. Sample size and power calculations

1. Before baseline data collection

Before baseline data collection, power calculations were conducted to determine the sample size needed to detect impact. Several parameters were needed to perform power calculations. Power calculations were performed for the outcome indicator of maize yields.[7] We sought a significance level of 5 per cent and a power of 80 per cent. In line with another agricultural study conducted in Guatemala (Hellin, Cox and López-Ridaura, 2017), we assumed a baseline mean value of 1.7 t/ha and a standard deviation of 1.105. Furthermore, we considered an average cluster size of 10 households (assuming cluster level to be the community level). Last, we considered a minimum detectable effect size (MDES) of 10 per cent, which is a conservative estimate. Patt, Suarez and Gwata (2005) found yields for farmers in Zimbabwe to increase by 19 per cent when applying forecast information. According to research carried out by Arslan (2015), CSA in the form of intercropping could lead to an increase in yields of at least 20 per cent. We repeated the power calculations with intra-cluster correlation coefficients (ICCs) of 0.05, 0.1 and 0.2, with different sample sizes and R^2 (0 per cent and 30 per cent) to see the trade-off between sample size and MDES.[8]

Table 5 reports results showing that to receive an MDES equivalent to a 10–12 per cent change in the outcome variable, a sample size of 1,700 is needed given an ICC of 0.1. This corresponds to a change in maize yields of 0.17 t/ha with respect to the baseline mean. The table shows that a higher

[6] The gender or other characteristics of the respondent cannot be determined from the survey data as all questions relate to household characteristics.
[7] We selected maize as the outcome because it is the most important crop grown in the western highlands of Guatemala (Hellin, Cox and López-Ridaura, 2017).
[8] Given a constant cluster size, varying the sample sizes is identical to varying the number of clusters.

ICC increases the required sample size to detect a statistically significant treatment effect, whereas a higher MDES would require a lower sample size.

Table 5. Power calculations for maize yield

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	BASELINE MEAN	BASELINE STD. DEVIATION	ICC	SAMPLE SIZE (TOTAL)	R^2	SIZE OF GROUP	SAMPLE SIZE IN C	SAMPLE SIZE IN T	MDES (IN T/HA)
1	1.7	1.105	0.05	1,300	30%	10	650	650	0.173
2	1.7	1.105	0.05	1,300	0%	10	650	650	0.207
3	1.7	1.105	0.1	900	30%	10	450	450	0.238
4	1.7	1.105	0.1	900	0%	10	450	450	0.285
5	1.7	1.105	0.1	1,700	30%	10	850	850	0.173
6	1.7	1.105	0.1	1,700	0%	10	850	850	0.207
7	1.7	1.105	0.2	2,600	30%	10	1300	1300	0.170
8	1.7	1.105	0.2	2,600	0%	10	1300	1300	0.203

Based on the power calculations shown in Table 5, as well as for financial and time reasons, IUCN decided to use a sample size of 1,500. The initial sample size included 750 households each for the prioritized micro watersheds and the control micro watersheds, resulting in a total of 1,500 households. After baseline data collection, a total of 1,486 households were interviewed, of which 758 households were in prioritized micro watersheds and 728 in the control micro watersheds.

2. AFTER BASELINE DATA COLLECTION

We updated power calculations using parameters from the actual baseline data.[9]

The formula for the updated power calculations is

$$MDES = \left(t_{1-\kappa} + t_{\frac{\alpha}{2}}\right) * \sqrt{1+\rho(m-1)} * \sqrt{\frac{1}{P(1-P)}\frac{\sigma^2}{N}} * (2(1-r))$$

where $t_{1-\kappa}$ and $t_{\frac{\alpha}{2}}$ are t-statistics representing the required power and level of statistical significance.

(by convention, we seek a power of 80 per cent and a statistical significance alpha of 5 per cent), ρ is the ICC, m is the average number of households per cluster, σ^2 is the variance, N is the total sample size. Finally, r represents the extent to which the outcomes are autocorrelated between two waves of data collection – that is, baseline and endline data collection. The term $(2(1-r))$ is added to the “usual” power calculation equation in order to adjust the sample size from a “simple” randomized control trial to a DiD specification (McConnell and Vera-Hernández, 2015).

The power calculations are repeated for different values of autocorrelation because this parameter can only be calculated from the data after the endline data collection. The outcomes used for the calculations are the two constructed indices described above. The ICC (column (3)), the mean value (column (4)) and the standard deviation (column (5)) are calculated from the baseline data.[10] The

[9] The power calculations conducted before data collection did not account for 2(1-*r*). At this point in time, the evaluation design, DiD, was not yet selected.
[10] The ICC was calculated by using Stata command “loneway”.

mean value and standard deviation are taken from the respective index for the treatment group.[11] Columns (6), (7) and (8) report the total sample sizes and the sample sizes for control and treatment groups, respectively. The average number of households per cluster m is 42.457 (for 35 clusters).[12] The deviation from the before calculation stems from the different cluster assumptions: community and micro watershed level.

Table 6 shows that a higher autocorrelation in the outcome variable, holding all other parameters fixed, results in a smaller MDES – that is, with a higher level of autocorrelation we will be able to detect a significant treatment effect for smaller effect sizes. Column (9) reports the MDES as absolute numbers with regard to the baseline mean.

Table 6. Power calculations for both indices with different values of autocorrelation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Autocorrelation	Outcome	ICC	Baseline mean	Baseline std deviation	Sample size (total)	Sample size in C	Sample size in T	MDEs
1	0.25	Index 1	0.036	0	1	1,486	728	758	0.344
2	0.5	Index 1	0.036	0	1	1,486	728	758	0.229
3	0.75	Index 1	0.036	0	1	1,486	728	758	0.115
4	0.25	Index 2	0.129	0	1	1,486	728	758	0.549
5	0.5	Index 2	0.129	0	1	1,486	728	758	0.366
6	0.75	Index 2	0.129	0	1	1,486	728	758	0.183

C. Challenges encountered with the research design and in data collection

A limitation of the impact evaluation is that the selection of the priority micro watersheds was determined purposively, based on different vulnerability criteria (e.g. economic, social), which prevented the use of an experimental design. Moreover, as previously mentioned, one of the challenges of this impact evaluation is attrition. The baseline questionnaire asked participants about their willingness to participate again in a similar survey. Based on their answers, 10.2 per cent of participants in the treatment group and 8.8 per cent of those in the control group reported that they would not like to participate again. The survey collected contact information from each participant to minimize attrition, yet we expect an attrition rate in the order of 10 per cent.

Another threat to the evaluation is the risk of spillovers from treated to control micro watersheds, particularly likely in case of CSA knowledge and early warning messages. We cannot rule out that personal relations between community members or extension workers across treatment and control areas would lead to information-sharing and behavioural adaptation of farmers. Moreover, IARNA commented in the LORTA inception workshop that the design of the EWS in the treatment area is part of a national plan to improve meteorological services countrywide. Over the course of a few years, the investments are likely to reach the control group as well. Such spillovers might positively affect the outcomes in the control group over time, which would lead to an underestimation of the

[11] Since both indices do not show any significant differences between treatment and control group, the selection does not change the results.
[12] Ideally, the number of households would be equally distributed across clusters. However, in our sample a different number of households were interviewed per cluster (minimum 9, maximum 118). McConnell and Vera-Hernández (2015) show that only a wide variation in cluster sizes would lead to a large increase in required sample size for a given MDES.

treatment effect. Also, there could be other initiatives and projects working in the controlled micro watersheds, which could positively influence those communities that are not the priority of this project. Due to the limited number of micro watersheds that have been randomly selected, this may affect the evaluation at the endline. Hence, additional questions will be added in the endline questionnaire about this, and other initiatives present in the micro watersheds.

With the objective to facilitate the implementation of the surveys in the communities, IUCN communicated with the local authorities from the municipalities, the Community Development Councils and community leaders to share information about this process and request authorization. They provided formal letters addressed to the authorities when necessary. In most cases, community leaders were contacted by phone to minimize in-person visits as much as possible (given the COVID-19 pandemic). Nevertheless, in some cases, in-person visits were coordinated with the objective of presenting the project and explaining the purpose of the surveys.

During the community visits, the surveyor teams were accompanied by community leaders from the Community Development Councils, municipality or governmental institution. In those cases where communities did not show interest, for whatever reason, we respected their right not to participate in the survey. In such cases, the team of surveyors were asked to retire from the community and replace that community with another. The replacement of communities was carried out randomly within the same micro watershed (control or treatment). About 36 per cent of the communities initially chosen to participate in the survey were replaced. Some communities did not give permission to enter. The most common reasons noted were as follows: to prevent COVID-19 contagion; that the survey was not immediately of direct benefit to them; and that local administrative changes complicated the management of the survey in some way.

During household visits, community borders were verified to ensure that the survey was implemented within the correct jurisdiction, as well as to ensure the team kept the safest areas possible with the least amount of risk. Due to the absence of a sampling frame, the selection of households to be interviewed was done randomly, in an alternate zigzag pattern along community streets.

Because of community replacements, it was necessary to carry out a second field phase to complement the missing surveys and comply with the established sample size. Moreover, in each of the selected communities, the aim was to conduct a maximum of 10 interviews. However, in some communities, it was necessary to conduct a larger number of surveys due to the small number of communities that are part of a specific micro watershed (either treatment or control). Because of these changes, and as mentioned above, the number of prioritized and non-prioritized watersheds slightly changed from what was initially planned (24 prioritized micro watersheds). The actual allocation is 21 prioritized and 14 non-prioritized watersheds. Even though data collection went through the above-mentioned challenges, we do not believe that this threatens the internal validity of the baseline assessment or the impact evaluation.

An additional and important challenge for the data collection was the ongoing COVID-19 pandemic. All the people working in the field were tested for COVID-19 to minimize the risk of contagion to the team as well as community members. Additionally, team members were provided with equipment such as face masks and sanitizers to prevent infection. Illustrations were used to convey understanding about the application of preventive measures, and talks were held with the team to create awareness and train them on the use of preventive measures. For the data collection, enumerators were selected from the broad area where the project is being carried out. Nine out of 12 surveyors were women, and all enumerators spoke at least one Mayan language of the territory: Kaqchikel, K'iche' and Mam.

D. DATA AND QUALITY ASSURANCE

During the preparation of data collection, the questionnaire and training materials were reviewed by the LORTA team and IUCN. Questionnaires went through pre-test and pilot phases. During data collection, C4ED and IUCN carried out daily monitoring of data-collection activities with the objective of identifying and correcting atypical values according to the following criteria:

- Location using GPS coordinates of the completed interviews
- Duration of each completed interview
- Proportion of response options “refused to respond” and “does not know” was utilized to identify possible anomalies
- Verification of logical combinations of responses – for example, cross-checking ages with type of family member, education level with illiteracy, electrical system with Internet
- Distribution of trigger questions to identify that enumerators had not shortened the survey

In order to achieve high-quality data, during data collection we identified valid and invalid interviews based on the above criteria. If anomalies were detected in completed questionnaires, IUCN and Sintaxis were promptly informed and asked to correct them.

Based on the daily number of completed interviews, we calculated progress thresholds, which helped to foresee the duration of data collection and to identify enumerators that deviated from the threshold. This information was shared with IUCN and Sintaxis during regular monitoring and progress calls.

E. SOFTWARE AND CODE

The Kobo ToolBox application (free version) was used to develop the survey and generate the data set. The tool made it possible to systematize the survey data in real time. This application has a user manual, which specifies the steps for its use according to the type of user. The KoboCollect application was installed on each of the surveyor’s devices.

To conduct quality checks during the data collection, the statistical software Stata was used. This tool was used to check for duplicates and outliers, monitor key variables, prevent data fabrication, and track progress and enumerator performance. As indicated in section D, daily data checks were performed by C4ED, as well as monitoring of GPS and interview duration. Once the data collection was finalized, the data-collection firm Sintaxis conducted further checks and data cleaning using Microsoft Excel. The data were organized into rows (participants’ answers) and columns (questions), which enabled a matching process to identify outliers for correction or discarding. The use of each software and the outputs derived are described in Table 7.

Table 7. List of software used for the baseline data collection

SOFTWARE	PURPOSE	PROJECT OBJECTS DERIVED
Kobo ToolBox	Data collection	Excel file with data collected
Microsoft Excel	Data cleaning	Excel file with cleaned data
Stata	Data quality checks, construction of indicators and main analysis	Do files with the description of the quality-check process as well as do files for the construction of indicators and main analysis

As data cleaning was conducted by the data-collection firm (Sintaxis) using Microsoft Excel, there is no code available to trace the changes from raw to clean data. To check for potential data manipulation, we compared different socioeconomic variables across the two data sets (raw and clean data). As shown in Table 8, we do not find differences in main socioeconomic characteristics, which supports the view that the fidelity of the data is high. Overall, there are 1,486 observations in the raw and cleaned data sets (merged using a unique ID field).

Each variable included in the data set is described in a codebook (available on request) and is associated with the corresponding question from the survey. However, since the cleaning was done with Microsoft Excel, the variables are not labelled, hence the importance of reviewing the codebook for the analysis.

Table 8. Checks between raw and cleaned data

Variable name	Raw data (mean value)	Cleaned data (mean value)
Age mother (continuous)	44.62544	44.62544
Age father (continuous)	47.30317	47.30317
Completed primary education mother (0–1)	0.4219381	0.4219381
Completed primary education father (0–1)	0 .4764468	0 .4764468
# of adults living in household (continuous)	3.318304	3.318304
Share of Mayan people (0–1)	0.8445491	0 .8445491
Plot size (continuous)[a]	4.564441	4.564441

Notes: [a] In this table, plot sizes are not converted to hectares. Only the raw numbers from the two data sets are reported as this exercise only checks for consistency between the two data sets.

F. Ethics

Before starting the surveys with the selected participants, the data-collection firm obtained participants' consent. When asking for consent, the enumerator provided information about him/herself, the company they represent and the purpose of the survey (to collect data in the community to know how the environment has impacted their life).

For the baseline data collection, no ethical clearance with a local or international research body (e.g. an institutional review board) was requested. In Guatemala, there is no institution that can provide ethical clearance for the study, and the project team does not have the funding necessary to obtain such clearance from other institutions outside the country. Yet, given the steps taken before the start of data collection (well-organized enumerator training following protocols and pilot sessions) and the handling of data (no names are displayed in the analysis or cleaned databases, use of aggregate instead of individual data, and use of ID numbers) there is no strong reason to be concerned regarding the ethics of the present study.

Given that the survey did not include sensitive questions on topics such as child labour or domestic violence, the integrity of participants was not threatened. The questions in the survey were related to socioeconomic variables (education, income, assets, food security) and participants' access to natural resources, knowledge of agroforestry and agricultural practices, use of warning systems, and risk perception of climate change. The intervention does not involve sensitive topics as the information received is intended to enhance land management and does not involve or generate any risks to participants.

An additional way in which the integrity of participants was ensured was that participants were able to stop taking part in the survey at any time if they were feeling uncomfortable. The enumerators also communicated that the survey would take about 40 minutes and that all information provided during the survey will remain anonymous and confidential. Data were collected from all households that gave consent to participate in the survey. Likewise, as part of this agreement, the participants were informed that their personal information will be stored in a confidential manner.

V. PRESENTATION OF RESULTS

The results section is divided into two parts. First, we describe the sample socioeconomic characteristics and conduct balance checks between treatment and control groups. Second, we explain how the two indices (the resilient and diversified livelihoods index and the responsiveness index) were estimated. In particular, we describe the method used to aggregate the indicators in the two indices. Thereafter, we present balance checks between treatment and control groups for the indices and their components.

A. BALANCE CHECKS FOR DESCRIPTIVE STATICS

To conduct the balance checks for the descriptive statistics we implemented two different strategies. First, we regressed a set of socioeconomic and outcome variables (all of them continuous variables) with the treatment variable as an independent variable. The logic behind this approach is to test whether there is a correlation between treatment assignment and the selected dependent variable. Moreover, by conducting a regression-based approach (which simulates a t-test), we can check whether the sample is balanced, meaning that the average values of each variable tested are not significantly different for the treatment and control groups. If necessary, this approach also allows for further variables to be controlled for. Second, we ran Chi-square tests to check the balance for binary variables because running a linear regression with this type of variable as the dependent variable is not the most preferred estimation approach.

We ran the linear regression shown in equation (1):

$$Y_h = \alpha_0 + \beta_1 T_h + \varepsilon_h$$

In the equation, Y_h represents the variable of interest for each household h – in this case, socioeconomic variables, indicators and final indices.[13] T_h denotes the treatment variable that takes the values: 0 if household is in the control group and 1 if household is in treatment group, and ε_h is the error term. We cluster standard errors at the micro watershed level. Following equation (1) we test whether $\beta_1 = 0$.

[13] We did not conduct balance tests exclusively on the indices, as otherwise we could have been potentially underestimating imbalances across groups.

Table 9 presents socioeconomic characteristics and the balance checks for socioeconomic variables using the linear regression approach. As shown in the table, the sample comprises households in which parents have an average age between 44 and 47 years old and where there are on average two children and three adults. In addition, households own an average of 0.4 hectares of land, with maize being the most dominant crop (as a share of total land) at around 0.2 hectares under this crop. These figures are similar to those presented by Hellin, Cox and López-Ridaura (2017) in the western highlands of Guatemala. All of the socioeconomic characteristics included in the table are balanced between the treatment and control groups.

Table 9. Balance, socioeconomic characteristics (continuous variables)

Variable	N/[Cluster]	Control mean/SE (1)	N/[Cluster]	Treatment mean/SE (2)	Difference (1–2)
T-test					
Age respondent	728	40.794	758	40.046	0.748
	[14]	[0.738]	[21]	[0.826]	
Age mother respondent	700	44.687	731	44.566	0.121
	[14]	[0.686]	[21]	[0.827]	
Age father respondent	615	47.385	645	47.225	0.161
	[14]	[0.624]	[21]	[1.011]	
Household size: Children	728	2.310	758	2.522	-0.212
	[14]	[0.131]	[21]	[0.105]	
Household size: Adults	728	3.212	758	3.421	-0.209
	[14]	[0.110]	[21]	[0.076]	
Household size: Elders	728	0.420	758	0.344	0.076
	[14]	[0.037]	[21]	[0.030]	
Plot size total (ha)[a]	702	0.407	742	0.390	0.016
	[14]	[0.018]	[21]	[0.012]	
Plot size maize (ha)[a]	692	0.228	729	0.202	0.026
	[14]	[0.012]	[21]	[0.007]	

Notes: Variables included in the table are continuous. The value displayed in the last column is the mean difference based on a linear regression with standard errors clustered at the micro watershed level. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level. Missing values originate from respondents who answered “do not know” or “refuse to answer”.
[a] The values of plot sizes were converted to hectares using a conversion factor averaged on the department level. The unit of plot sizes (*cuerdas*) varies across and within departments.

The results of the Chi-square test for the binary socioeconomic characteristics are reported in Table 10. From these results, we observe that in about 45 per cent of households the mother or father completed primary school and that the largest share of households (on average, 82 per cent) hold two adults. In the remaining households, around 14 per cent are single-parent households led by women and 3 per cent are single-parent households led by men (Table A - 1). About half of the women in single-parent households are not formally educated (52 per cent) and only 35 per cent have completed primary education. This picture improves for dual-parent households, where 42 per cent of women are not formally educated and where 45 per cent have completed at least primary education (Table A - 2). In Table 10 we observe that, on average, 94 per cent of households have electricity and 84 per cent own at least one cellular phone. Regarding productive activities, we observe that almost all households produce annual crops such as maize, beans or rice (on average, 95 per cent) rather than perennial crops such as fruit trees (on average, only 8 per cent of households produce these crops).[14] Moreover, just over two thirds of sample households farm livestock for sale or consumption.

[14] Annual crops are those that can be harvested once or over the course of one season or year and that afterwards die. Hence, these crops have a life span of about one year or less. Perennial crops have a longer life span, can be harvested once or more per year without dying and are part of the landscape.

Importantly, we observe significant differences in agricultural practices (with a higher proportion of treatment households growing annual crops and a lower proportion growing permanent crops) and livestock raising (which was more common in control households). We also find that a higher share of treatment households owns a cellular phone. These differences may be driven by the characteristics of the four groups that cluster the 35 micro watersheds (see Table 4). Hence, the socioeconomic characteristics that are not balanced will be included in the endline report analysis as controls.

Since we observe significant differences on productive activities, we conducted further checks on other related variables of interest. Given that one of the main objectives of the intervention is to train household heads on practices that increase households' resilience towards climate change, we conducted balance checks on initial knowledge about agricultural and conservation practices. The practice that has been most implemented by households is agroforestry, the most frequent type being the planting of trees on a contour (located at the edge or limit of their land or parcel) and trees intercropped with annual crops. On average, 85 per cent of households have applied agroforestry on their plots. In addition, the majority of households have applied soil conservation techniques and agricultural practices against climate change. A minority of households have also implemented silvopastoral systems, the most frequent being live barriers with grasses and live fences in the pasture area. As shown in Table 11, a significantly larger proportion of households in the treatment group has applied all of these agricultural practices compared to households in the control group. These differences are significant at the 1 per cent level, which indicates that groups are not comparable before the intervention starts.

Overall, a greater proportion of treatment households grow annual crops; own a cellular phone; already apply silvopastoral systems, agroforestry and soil conservation techniques; and have already implemented agricultural practices against climate change. A greater proportion of control households produce perennial crops and farm livestock. While the observed significant differences do not threaten the evaluation design, it is important to keep in mind that households seem not to be comparable in their initial knowledge and application of strategies to mitigate the negative effects of climate change.

Table 10. Balance, socioeconomic characteristics (binary variables)

Variable	N	Control mean/SE (1)	N	Treatment mean/SE (2)	P-value
Chi-square-test					
Completed primary school: mother	728	0.430	758	0.414	0.540
		[0.018]		[0.018]	
Completed primary school: father	728	0.484	758	0.470	0.593
		[0.019]		[0.018]	
Type of household, 0 – monoparental; 1– biparental	728	0.826	758	0.823	0.906
		[0.014]		[0.014]	
Electricity	725	0.934	758	0.953	0.119
		[0.009]		[0.008]	
Cellular phone	723	0.831	757	0.864	0.080*
		[0.014]		[0.012]	
Annual crops	728	0.934	758	0.968	0.002***
		[0.009]		[0.006]	

VARIABLE	N	CONTROL MEAN/SE (1)	N	TREATMENT MEAN/SE (2)	P-VALUE
Perennial crops	728	0.106	758	0.057	0.001***
		[0.011]		[0.008]	
Livestock farming	677	0.713	678	0.665	0.055*
		[0.017]		[0.018]	

Notes: Variables included in the table take only the values of 1 or 0. The value displayed for chi-square-tests are p-values. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level. Missing values originate from respondents who answered "do not know" or "refuse to answer".

Table 11. Balance, variable associated with climate change (binary variables)

VARIABLE	N	CONTROL MEAN/SE (1)	N	TREATMENT MEAN/SE (2)	P-VALUE
Chi-square-test					
Application of any silvopasture system	716	0.264	756	0.351	0.000***
		[0.016]		[0.017]	
Application of any agroforestry system	719	0.829	757	0.875	0.014***
		[0.014]		[0.012]	
Application of any soil conservation technique	724	0.552	757	0.736	0.000***
		[0.018]		[0.016]	
Implemented agricultural practices against climate change	692	0.480	732	0.586	0.000***
		[0.019]		[0.018]	

Notes: Variables included in the table take only the values of 1 or 0. The value displayed for chi-square-test are p-values. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level. Missing values originate from respondents who answered "do not know" or "refuse to answer".

B. INDICES ESTIMATION AND BALANCE CHECKS

The two indices (the resilient and diversified livelihoods index and the responsiveness index) were estimated using principal component analysis (PCA) and the Anderson index methodology (see Anderson, 2008). Both indices were calculated with the statistical software Stata. A condition of estimating these indices is that all indicators point in the same direction.[15] We now explain each of the methods, outline their advantages and disadvantages, and explain which method was selected for estimating the index.

PCA generates new variables that are linear functions of those in the original data set. These "new variables" maximize variance and are uncorrelated with each other (principal components). Usually, the principal component with an eigenvalue larger than 1 is used to predict the index. This eigenvalue is a proxy for the amount of variance that the component captures from the data. Since the aim is to obtain an informative index, the component with an eigenvalue larger than 1 is selected as it captures the maximum possible information from all the variables that compose the index.

[15] This can be done manually with PCA (by changing the direction of the variables) or with a coding option ("flip") for the Anderson index.

Although PCA is helpful in compressing a lot of information in a limited number of components, there are several disadvantages for this methodology. First, it requires variables to be continuous and normally distributed.[16] Second, if observations are missing data, then they are dropped from the estimation, reducing the sample size. Last, in some cases more than one principal component has an eigenvalue larger than 1. Usually, under this scenario, the proportion of variation is less than 50 per cent in the first principal component, meaning that it is not capturing half of the information.

The Anderson index is based on a generalized least squares estimation and can be explained as the weighted mean of a set of standardized variables/indicators.[17] This statistical method can be used with all variable types: binary, ordinal, categorical or continuous. It first standardizes all variables (which makes it easier to interpret treatment effects) and then squeezes them, by assigning different weights to variables, into a single value. Under this methodology, highly correlated indicators receive less weight than uncorrelated indicators. Moreover, it takes into account all data but ascribes lower weight to indicators with missing values. This method also replaces missing information with the mean of the normalized indicator (zero) (which is sometimes considered as a shortcoming).

Based on the characteristics of each indicator and of the data set that was collected and collated by Syntaxis, IUCN and the LORTA team, we believe that the Anderson index is the preferred method to estimate the two indices mentioned above. One prominent reason is the high number of missing values in some of the indicators, especially those related to income and agricultural production. However, to have a robust approach that supports our decision, we performed the Kaiser-Meyer-Olkin (KMO) test to assess the suitability of PCA. As shown in Table A - 3 and Table A - 4 in Appendix 1, for both indices the KMO has a value lower than 0.59, indicating that the selected group of variables are not ideal for PCA.[18] However, as robustness checks in the estimation of the two indices, we performed both the PCA and the Anderson methods, including and excluding binary variables, and compared whether the estimated indices under each method are correlated or not. The correlation tables (

[16] Polychoric PCA allows the use of binary variables alongside continuous variables.

[17] With the standardization, the mean values of each variable are 0 and the standard deviation is 1. This is useful as it helps to compare data that have different metrics.

[18] Following the test specifications, the following ranges based on Kaiser (1974) indicate the different categories for suitability: 0.00 to 0.49 unacceptable; 0.50 to 0.59 miserable; 0.60 to 0.69 mediocre; 0.70 to 0.79 middling; 0.80 to 0.89 meritorious; 0.90 to 1.00 marvellous.

Table A - 5, Table A - 6, and Table A - 7) are presented in Appendix 2. According to the tables, the estimations are highly correlated at 1 per cent significance level. Hence, we are confident that with the Anderson index methodology we are able to get a better estimate of the indices.

Table 12 shows the descriptive statistics of the resilient and diversified livelihood index, calculated with the Anderson methodology, as well as of components that underlie the index (see Table 2 for details). Table 12 shows that households obtained an average of 40 points in the multidimensional poverty index, which indicates that the sample is, on average, composed of poor households. This result is in line with the report from URL and IARNA (2020), in which the authors find that, on average, households in Guatemala can be classified as poor.[19] In the single-parent households led by women, 67 per cent can be classified as poor and 33 per cent as extremely poor (not shown). For dual-parent households we observe a slightly better picture, in which 81 per cent are classified as poor and 17 per cent as extremely poor (not shown). When considering only the income variable, 60 per cent of the households have a monthly income of less than Q 1,000.[20] In the case of single-parent households led by a woman, 71 per cent receive a monthly income lower than Q 1,000 (not shown). Regarding the water accessibility variable, the principal source of water in the households surveyed is through a home connection (75 per cent of the households), followed by those with their own well (16 per cent – not shown). Two per cent of households collect water from surface sources (river, lake or spring), and 1 per cent collect rainwater (not shown).[21]

In addition, we observe that household members collect few forest products for sale or consumption (on average, they collect 15 per cent of the products that were mentioned to them during the survey).

Regarding household vulnerability to the effects of climate change, on average, 83 per cent of households stated that they did not lack money or food due to climate change but due to other reasons (i.e. unemployment).[22] Moreover, 61 per cent of households reported that during the last 12 months they have lacked food or money in the house (not shown). Importantly, 17 per cent of households stated they were lacking food or money at home due to the negative effects of meteorological events (17 per cent). We also observe that there is a low-to-medium diversification of agricultural products, as the Simpson-Herfindahl index has an average value of 0.39.[23]

Based on the insignificant differences between treatment and control groups, shown in the last column of the table, we conclude that, on average, the groups are comparable on the characteristics with the first index.

Table 12. Balance, resilient and diversified livelihood index

Variable	N/[Cluster]	Control mean/SE (1)	N/[Cluster]	Treatment mean/SE (2)	Difference (1–2)
T-test					
Resilient and diversified livelihood index	726	-0.001	757	0.001	-0.001
	[14]	[0.068]	[21]	[0.050]	

[19] It is not possible to do a 1–1 comparison of both multidimensional poverty indicators because the study by the university includes other dimensions such as health, and also gives different weights to the other dimensions. However, based on their classification, households with 30–67 points can be classified as poor. The average number of points that households in Guatemala obtained was 66.
[20] According to the 2021 OANDA converter, this is less than USD 130.
[21] We did not include these detailed numbers in the tables but wanted to provide more information in the text, so the reader has more knowledge on the household characteristics.
[22] For the indicator, we assumed a 0 for the households that did not lack food or money in the last 12 months.
[23] Following the work of Bellon and others (2020), we characterize this value of the Simpson-Herfindahl index as low to medium. As shown in the paper, low diversification is considered when the index has values lower than 2.

Multidimensional poverty index [a]	706	40.449	742	39.315	1.134
	[14]	[1.216]	[21]	[0.915]	
Education	726	9.221	755	8.684	0.537
	[14]	[0.536]	[21]	[0.493]	
Life quality	711	14.679	748	14.842	-0.163
	[14]	[0.396]	[21]	[0.413]	
Income	722	9.136	756	8.406	0.730
	[14]	[0.430]	[21]	[0.286]	
Food security	725	7.388	756	7.341	0.047
	[14]	[0.139]	[21]	[0.128]	
Water accessibility [b]	725	1.219	757	1.127	0.092
	[14]	[0.068]	[21]	[0.045]	
Water quality [b]	722	1.119	755	1.118	0.001
	[14]	[0.040]	[21]	[0.019]	
Collecting forest products [a]	717	0.148	755	0.154	-0.006
	[14]	[0.007]	[21]	[0.008]	
Non-vulnerable to climate change [c]	712	0.826	750	0.841	-0.015
	[14]	[0.015]	[21]	[0.018]	
Agricultural diversification [a] (Simpson-Herfindahl index)	702	0.375	742	0.398	-0.023
	[14]	[0.026]	[21]	[0.017]	

Notes: The value displayed in the last column is the mean difference based on a linear regression with standard errors clustered at the micro watershed level. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level. Missing values originate from respondents who answered “do not know” or “refuse to answer”. [a] Continuous. [b] Discrete. [c] Binary. Table 2 explains the indicators in detail.

Table 13 shows the descriptive statistics and balance checks for the second index (see Table 3 for details of the index and underlying indicators). In general, the indicators suggest that there is limited perception and knowledge of climate change events and consequences. On average, households reported knowing only 18 per cent of the impacts of climate change from a list of 10 items that were listed in the survey. Moreover, households stated awareness of only 9 per cent of possible responses to climatic impacts from a list of 13 measures/strategies mentioned to them. Lastly, households reported limited awareness and use of community-level EWS (on average 9 per cent and 7.4 per cent, respectively). Contrary to our expectations, given the significant differences observed in agricultural practices, we do not observe significant differences between treatment and control groups for almost all indicators as well as for this index. Knowledge of climate change responses is the only indicator with a significant difference between the groups (at the 10 per cent level), with the treatment group having, on average, greater awareness than the control group. As mentioned above, initial differences do not threaten the evaluation design because the DiD and matching will control for them.

Table 13. Balance, responsiveness index

VARIABLE	N/[CLUSTER]	CONTROL MEAN/SE (1)	N/[CLUSTER]	TREATMENT MEAN/SE (2)	DIFFERENCE (1–2)
T-test					
Responsiveness index (Anderson index)	728	-0.050	758	0.048	-0.097
	[14]	[0.087]	[21]	[0.081]	
CC risk perception, proportion of “yes” answers [a]	728	0.172	758	0.182	-0.010
	[14]	[0.008]	[21]	[0.008]	
Knowledge response on CC events, proportion of “yes” answers [a]	716	0.084	753	0.099	-0.015*
	[14]	[0.007]	[21]	[0.005]	
Existence of EWS, community average [b]	728	0.093	758	0.089	0.003
	[14]	[0.009]	[21]	[0.011]	
Use of EWS information, community average [b]	728	0.075	758	0.073	0.001
	[14]	[0.008]	[21]	[0.008]	

Notes: The value displayed in the last column is the mean difference based on a linear regression with standard errors clustered at the micro watershed level. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level. Missing values originate from respondents who answered “do not know” or “refuse to answer”. a Continuous. b Discrete. Table 3 explains the indicators in detail.

VI. DISCUSSION

The baseline data present relevant and accurate information that will be vital in the assessment of the impact of subcomponent 1.1 on micro watershed management and ecosystem-based adaptation as well as component 3 on EWS. The balance tests for the two outcome indices and sub-indicators show no systematic differences between treatment and control groups.

The baseline data have, however, highlighted significant differences between treatment and control households in terms of initial household characteristics. On average, our evaluation sample is composed of households in which household heads have an average age between 44 and 47 years of age, have two children, own 0.4 hectares of land and grow maize on around half of this land. Almost all households grow annual crops (such as maize, beans or rice), around 8 per cent grow perennial crops (such as fruit trees) and around two thirds farm livestock for sale or consumption. Around 17 per cent are single-parent households, mainly headed by women who have less education than women in dual-parent households. On average, 94 per cent of households have electricity and 84 per cent own at least one cellular phone.

Turning to the differences between the two groups, a greater proportion of treatment households grow annual crops; own a cellular phone; already apply silvopastoral systems, agroforestry and soil conservation techniques; and have already implemented agricultural practices against climate change. A greater proportion of control households produce perennial crops and farm livestock.

Are these results in line with expectations? There are two possible explanations for the significant differences that are observed. One possibility is that interventions that are being implemented – micro watershed management, ecosystem-based adaptation training and EWS – are attracting beneficiaries that are better connected with service providers (such as public agricultural extension agencies and NGOs) and have had a greater degree of interaction with these types of projects previously. The second possibility is that the differences reflect the nature of the groups of micro watersheds (which have been clustered according to similar conditions). As Table 4 highlights, a different number of treatment and control households were surveyed in each group and the observed differences may simply reflect these different environmental conditions.

These two possible explanations do not threaten the evaluation design, because socioeconomic characteristics that are not balanced will be included in the impact evaluation estimates as controls and systematic differences between clusters of micro watersheds (groups) will be examined. Moreover, any initial differences in outcome variables are accounted for by the DiD design.

An important element of the baseline, and future evaluation, is a subgroup analysis where the main outcome variables of the impact intervention are assessed for single-parent household and dual-parent households. As explained above, in Guatemala many vulnerable households have a single parent, frequently an indigenous woman, which is partly due to emigration to the United States. Women in single-parent households have less education than their married counterparts and are poorer, with 71 per cent receiving a monthly income lower than Q 1,000 (compared to 60 per cent of households overall). These baseline characteristics confirm the project’s decision to prioritise interventions for this group, and to assess the impact of the interventions on this subgroup (although, due to the limited sample size, these findings will be underpowered).

Turning to the two outcome indices, the baseline data found only one significant difference between treatment and control groups: on average, greater awareness of climate change responses in the treatment group. Overall, though, respondents from both the treatment and control groups displayed a low level of awareness and knowledge about climate change impacts and responses.

The baseline results illustrate that the area under study is not prepared to adapt to climate change and that there is a low level of awareness about the effects climate change can have on livelihoods. These findings align with the study by URL and IARNA (2020), who also find that households are not ready to confront climate challenges.

VII. CHALLENGES AND SHORTCOMINGS

As highlighted above, the baseline encountered a number of challenges. We now summarize these and the shortcomings of the evaluation strategy. In brief, they are the balance of treatment and control groups; attrition of households between baseline and endline; access to communities and participation in the baseline survey; the COVID-19 pandemic; and the construction of indices.

While the baseline data on socioeconomic characteristics offer evidence that the treatment and control groups do not vary significantly, we do find significant differences in terms of households' productive activities and the agricultural practices that households apply on their plots. The characteristics that are not balanced between treatment and control groups will either be part of the matching propensity score or be added in the estimation as controls. Also, time-variant unobservables (e.g. motivation) might threaten the evaluation design if they affect selection into the treatment.

Another challenge to the impact evaluation is attrition of households between baseline and endline data collection. It is common for households to move. Some households, especially elderly households, pass away. The baseline data we have here indicates an attrition rate of 10 per cent, which should be considered as a lower bound for the anticipated attrition rate.[24] Attrition can have implications for the analysis and can threaten the validity of the impact evaluation. When households disappear between baseline and endline, two potential threats might affect the DiD method and the matching technique: the average outcomes at endline are not measured over the same set of households as in the baseline, and attrition might differ across treated and control households. The first threat is less problematic if attrition is independent of the treatment. In such a case, treatment estimates will be unbiased (as attrition balances out between treatment and control groups) but the lower sample size will reduce power. A smaller sample size also reduces the chances that for each treated household two or more untreated households are available for matching procedures.[25]

If treatment itself leads to attrition, this could undermine the evaluation design due to selection bias. When attrition rates differ between treatment and control areas, an estimation bias is induced. In addition, the assumption of common support between groups could potentially not hold anymore.[26] On the other hand, attrition rates might increase more in control than treatment areas if households perceive that they did not benefit from any project activities.

A closely related challenge is when communities are unwilling to participate in the data-collection exercise. As we saw in the baseline data collection, some communities chose not to take part in the survey or in the intervention. As explained above, around 36 per cent of communities initially chosen for the survey did not participate due to the COVID-19 pandemic, local administrative changes or a lack of immediate benefit from the exercise. In such cases, the replacement of communities was carried out respectfully and randomly within the same (control or treatment) micro watershed. The dynamics of the communities were considered and respected to achieve a broad participation – for example, by considering the frequency of fair days and market days and the fluency of the team in local Mayan languages. The replacement of communities led to a second

[24] The baseline questionnaire asked participants about their willingness to participate again in a similar survey. Based on their answers, 10.2 per cent of participants in the treatment group and 8.8 per cent of those in the control group reported that they would not like to participate again.
[25] In order to increase the chances of matching an individual from the treatment group, we ideally sample at least two households from the comparison group for each treatment household.
[26] There are bounding approaches at hand that would still report biased results but could lead to some information about bounds of treatment effects.

baseline field phase to reach the desired sample size. In some communities, we conducted a larger number of surveys due to the small number of communities that were part of a specific micro watershed. Overall, the baseline survey completed questionnaires with 21 prioritized and 14 non-prioritized watersheds.

A challenge that underlay the data-collection exercise and, indeed, the communication between all members of the LORTA team, was the ongoing COVID-19 pandemic. As explained above, all possible measures were taken to minimize the risk of contagion to the team as well as community members (through the use of face masks, sanitizers, training sessions on the use of preventive measures). Illustrations were used to convey information about the application of preventive measures, and talks were held with the team to create awareness and train them on the use of preventive measures.

The final challenge discussed in this section relates to the construction of indices, specifically the presence of missing values. As mentioned above, we use the Anderson index as it accounts for them and does not use only the observations that have all information. However, for the endline report we will consider using an imputation method to replace missing values and avoid losing power. As a first step we performed some checks on whether the missing values are randomly distributed between treatment and control groups. Table 14 shows that there are significant differences for the two indices. In particular, the table shows that households in the control group were more prone not to report information. When conducting the impact evaluation estimates, it will be important to compare results with and without the imputation of missing values.

Table 14. Systemic differences in missing values

Variable	N/[Cluster]	Control mean/SE (1)	N/[Cluster]	Treatment mean/SE (2)	Difference (1–2)
Missing values index 1	728	0.091	758	0.053	0.038**
	[14]	[0.016]	[21]	[0.008]	
Missing values index 2	728	0.016	758	0.007	0.010*
	[14]	[0.004]	[21]	[0.002]	

Notes: The value displayed in the last column is the mean difference based on a linear regression with standard errors clustered at the micro watershed level. Standard errors in brackets. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level.

VIII. CONCLUSION

This baseline report has outlined the ToC of the intervention and the evaluation questions and indicators. It has also described how the data collection was conducted. The total sample size is 1,486 households, distributed into 21 treated (758 households) and 14 control micro watersheds (728 households). The report has explained that, after consultations with the project team, a DiD with matching approach was selected as the most robust method to evaluate the impact of the project. The report also described the baseline characteristics of beneficiaries and showed the results of comparisons between beneficiary and comparison households in treatment and similar nearby micro watersheds. The baseline report has also compared treatment and comparison households across two indices:

- A resilient and diversified livelihoods index (*Índice de medios de vida diversificados y resilientes*)
- A responsiveness index (*Índice de Capacidad de Respuesta*)

As expected, comparisons between treatment and control households show both groups are very similar: households and communities have almost the same vulnerabilities, gaps and needs regarding improving their resilience. Socioeconomic characteristics are very balanced between both groups: on average, households are considered as poor, and education and gender variables show no significant differences. Conversely, agricultural practices do present some variability: treatment households showed more experience of agroforestry systems and silvopastoral systems than control households. There is another systematic difference between both groups: on average, treatment households showed more knowledge of responses to climate change than control group households. Both groups expressed the same limited perception of knowledge of climate change events and consequences.

These baseline survey results have highlighted two key features that are important for project and programme implementation. The first of these is the widespread experience of households with agroforestry (around 85 per cent of households), soil conservation measures (about 64 per cent of households) and silvopastoral systems (around 30 per cent of households). Overall, more than half of the households have already implemented some agricultural practices against climate change. Because of this, the IUCN team is considering working with a broader range of interventions that are suited to the different productive systems in the area, especially those that complement local knowledge and enable diversification of livelihoods. These interventions will support agricultural practices that are in accordance with local ancestral knowledge of productive systems. This is expected to increase uptake and use and to dovetail closely with existing practices.

The second key feature of the baseline survey with direct relevance for project and programme implementation comes from the responsiveness index. Here, respondents reported that only 9 per cent were aware of community EWS and only 7 per cent used EWS information. This is surprisingly low, and the project will focus its resources on enhancing capacities in the communities to increase awareness and of these systems with an explicit gender dimension.

Finally, it is important to highlight the complementarity between the project's monitoring and evaluation system (according to the indicators identified in the FAA signed between GCF and IUCN) and the LORTA process. For example, the estimation of baseline data reported to the GCF Secretariat benefited from the LORTA team’s work on this impact evaluation baseline (through the design of the questionnaire, monitoring of fieldwork, estimation of indicators). For IUCN, the cost of estimating the indicators for GCF Secretariat reporting was reduced, since they were incorporated in a single household survey. More broadly, this report offers a wider view of the characteristics of beneficiary and control households in the Western Highlands and provides invaluable data to support the project during implementation.

REFERENCES

Aguilar, E., and others (2005). Changes in precipitation and temperature extremes in Central America and northern South America 1961–2003. *Journal of Geophysical Research*, vol. 110(D23).

Anderson, Michael (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association*, vol. 103, No. 484, pp. 1481–1495.

Arslan, Aslihan, and others (2015). Climate smart agriculture? Assessing the adaptation implications in Zambia. *Journal of Agricultural Economics*, vol. 66, No. 3, pp. 753–780.

Bellon, Mauricio, and others (2020). To diversify or not to diversify, that is the question. Pursuing agricultural development for smallholder farmers in marginal areas of Ghana. *World Development*, vol. 125, No. 104682.

Food and Agriculture Organization of the United Nations (2010). Measuring Resilience: A Concept Note on the Resilience Tool. Food Security Information for Decision Making: Concept note. Available at *https://www.fao.org/resilience/resources/resources-detail/en/c/317275/.*

Giorgi, F. (2006). Climate change hot-spots. *Geophysical Research Letters*, vol. 33, No. L08707. Available at *https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2006GL025734.*

Hellin, Jon, Rachael Cox and Santiago López-Ridaura (2017). Maize diversity, market access, and poverty reduction in the western Highlands of Guatemala. *Mountain Research and Development*, vol. 37, No. 2, pp. 188–197.

International Union for Conservation of Nature (2021). Síntesis de los cambios en el clima identificados y las amenazas a los sistemas productivos en el área priorizada del proyecto (Synthesis of identified climate changes and threats to production systems in the project's prioritized area). Consultancy report.

Kaiser, Henry (1974). An index of factor simplicity. *Psychometrika 39*, pp. 31–36.

McConnell, Brendon, and Marcos Vera-Hernández (2015). Going beyond simple sample size calculations: a practitioner's guide. IFS Working Paper W15/17. Institute for Fiscal Studies.

Patt, Anthony, Pablo Suarez and Chiedza Gwata (2005). Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe. PNAS 35, pp. 12623–12628.

Universidad Rafael Landívar and Institute of Agriculture, Natural Resources and Environment (2020). Informe de la línea base del proyecto "Construyendo redes de investigación-acción para el desarrollo territorial y la adaptación al cambio climático en Guatemala". Report presented to the Swedish Agency for Development (ASDI). Available at *http://www.infoiarna.org.gt/nuestro-trabajo/proyectos/proyecto-redes-de-investigacion-accion/.*

Appendix 1. THEORIES OF CHANGE FOR SPECIFIC PROJECT COMPONENTS

Component 1.1: Integrated Climate-Smart Watershed Management – Training

1) Inputs: While the budget is provided by IUCN through GCF funding, inputs of human and technical resources are provided by different Guatemalan agencies/organizations (INAB, MAGA, NGOs, etc.).
2) Activities: The actions at the core of this component are training of extension agents (training of trainers) as well as technical and legal assistance to develop management plans of micro watersheds in line with climate change. In addition, information about the incentive programmes will be disseminated, and extension agents will provide technical assistance for the application process as well as support for the implementation of ecosystem-based adaptation measures before and after the incentives are awarded.
3) Outputs: Through the inputs and activities of the component, extension workers are trained in practices of CSA, micro watershed management plans are designed, and more farmers have access to the incentives.
4) Outcomes: These consist of enhanced knowledge and application of CSA practices, improved governance of integrated watershed management and the participation of farmers in the incentive programme. This will lead to a diversification of crops, higher yields and increased forest cover.
5) Goals: The main goals behind this component are higher water security of the communities - and in particular farmers – as well as improved resilience against climate change.

Component 3: Climate-related information provided to farmers and other target stakeholders for watershed management - EWS

1) Inputs: While the budget is provided by IUCN through GCF funding, the input in terms of human and technical resources comes from INSIVUMEH and URL-IARNA.
2) Activities: The activities all refer to the gathering, analysis, translation and dissemination of climate information. Five existing weather stations will be complemented with the right equipment to ensure precise climate information. In addition, 10 new hydrological and meteorological stations will be established. A comprehensive EWS for the Highlands will be designed, including the distribution methodology of the information and recommendations for actions.
3) Outputs: Through the inputs and activities of this component, the EWS is in place and the information and recommendations related to sustainable practices are distributed to the communities.
4) Outcomes: If the target group benefits from the elements described in the output stage, more producers/farmers will be informed on weather-related events and coping productive coping and conservation management strategies. This will trigger the implementation and adaptation of CSA practices and higher and less volatile yields.
5) Goals: The main goal behind this component and programme is improved resilience against climate change.

Appendix 2. TABLES

Table A - 1. Household type

VARIABLE	N	FREQUENCY
Biparental	1,225	82.44
Monoparental		
Female	206	13.86
Male	35	2.36
Other	20	1.35

Table A - 2. Mother's education by household type

HOUSEHOLD	EDUCATION		
	NO EDUCATION	COMPLETED PRIMARY	TOTAL
Biparental	498	555	**1,225**
Monoparental	109	72	**206**

Table A - 3. PCA KMO test, index 1

VARIABLE	KAISER-MEYER-OLKIN (KMO)
Multidimensional poverty index	0.6033
Water accessibility	0.5426
Water quality	0.5493
Collecting forest products	0.5047
Agricultural diversification (Simpson-Herfindahl index)	0.5055
Non-vulnerable to climate change	0.5810
Overall	**0.5465**

Table A - 4. PCA KMO test, index 2

VARIABLE	KAISER-MEYER-OLKIN (KMO)
CC risk perception, proportion of "yes" answers	0.5255
Knowledge response on CC events, proportion of "yes" answers	0.5281
Existence of EWS, community average	0.5006
Use EWS information, community average	0.5011
Overall	**0.5021**

Table A - 5. Corelation table PCA and Anderson, index 1 (all variables)

	PCA METHOD
PCA method	1.0000
Anderson method	0.5761***

Notes: The estimation of both indices takes into account all continuous and binary variables that compose index 1. PCA index was standardized to make it comparable with the Anderson index. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level.

Table A - 6. Corelation table PCA and Anderson, index 1 (only continuous variables)

	PCA METHOD
PCA method	1.0000
Anderson method	0.5018***

Notes: The estimation of both indices takes into account only the continuous variables that compose index 1. PCA index was standardized to make it comparable with the Anderson index. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level.

Table A - 7. Corelation table PCA and Anderson, index 2 (all variables)

	PCA METHOD
PCA method	1.0000
Anderson method	0.7299***

Notes: The estimation of both indices takes into account all variables that compose index 2 as they are all continuous. PCA index was standardized to make it comparable with the Anderson index. ***, ** and * indicate significance at the 1, 5 and 10 per cent critical level.

Independent Evaluation Unit
Green Climate Fund
175 Art center-daero, Yeonsu-gu
Incheon 22004, Republic of Korea
Tel. (+82) 032-458-6450
ieu@gcfund.org
https://ieu.greenclimate.fund

GREEN CLIMATE FUND | Independent Evaluation Unit