GREEN CLIMATE FUND | Independent Evaluation Unit

October 2026

# Learning-Oriented Real-Time Impact Assessment Programme (LORTA)

# Impact Evaluation Baseline Report for SAP021: Community-Based Landscape Management for Enhanced Climate Resilience and Reduction of Deforestation in Critical Watersheds


**Summary**

This report presents the baseline findings of an impact evaluation of GCF SAP021: Community-Based Landscape Management for Enhanced Climate Resilience and Reduction of Deforestation in Critical Watersheds in Timor-Leste. Implemented by the Japan International Cooperation Agency (JICA), the project aims to strengthen climate resilience, reduce deforestation, and improve rural livelihoods through community-based natural resource management, participatory land-use planning, and climate-resilient agricultural practices.

The impact evaluation, supported by the IEU's LORTA programme, examines the effects of these interventions on forest conservation, agricultural practices, food security, and household livelihoods. The baseline survey, conducted in September–October 2024, covered 1,002 households across 50 villages (25 treatment and 25 comparison). This report documents pre-intervention conditions and establishes a foundation for assessing the project's impacts at endline.

© 2026 Green Climate Fund Independent Evaluation Unit
175, Art center-daero
Yeonsu-gu, Incheon 22004
Republic of Korea
Tel. (+82) 032-458-6450
Email: ieu@gcfund.org
https://ieu.greenclimate.fund

All rights reserved.

*First Edition*

This report is a product of the Independent Evaluation Unit at the Green Climate Fund (GCF/IEU). It is part of a larger effort to provide open access to its research and work and to make a contribution to climate change discussions around the world.

While the IEU has undertaken every effort to ensure the data in this report is accurate, it is the reader's responsibility to determine if any and all information provided by the IEU is correct and verified. Neither the author(s) of this document nor anyone connected with the IEU or the GCF can be held responsible for how the information herein is used.

Rights and permissions
The material in this work is copyrighted. Copying or transmitting portions all or part of this report without permission may be a violation of applicable law. The IEU encourages dissemination of its work and will normally grant permission promptly. Please send requests to ieu@gcfund.org.

Citation
The suggested citation for this report is:
Independent Evaluation Unit. *Impact Evaluation Baseline Report for SAP021: Community-Based Landscape Management for Enhanced Climate Resilience and Reduction of Deforestation in Critical Watersheds*. Songdo, Republic of Korea: Independent Evaluation Unit, Green Climate Fund, 2026.

Credits
*Head of the GCF Independent Evaluation Unit*: Andreas Reumann
*Evaluation task lead:* Susumu Yoshida, Evaluation Specialist – Implementation Science, Independent Evaluation Unit
*Editing*: Toby Pearce
*Layout*: Giang Pham
*Cover design:* Therese Gonzaga
*Cover image*:

A FREE PUBLICATION

Printed on eco-friendly paper

## List of authors

The authors of the LORTA report are (in alphabetical order of the surnames):

| **Full name** | **Affiliation** |
|---|---|
| Hiroaki Asaoka | Japan International Cooperation Agency |
| Ankhbayar Delgerchulun | Center for Evaluation and Development |
| Johanna Gather | Center for Evaluation and Development, Mannheim University |
| Yuta Hayakawa | Japan International Cooperation Agency |
| Susumu Yoshida | Independent Evaluation Unit, Green Climate Fund |

# Table of Contents

List of authors ..... c
Abbreviations ..... f
Executive summary ..... 1
I. Context ..... 2
II. Project description ..... 4
III. Theory of change ..... 5
IV. Evaluation questions and indicators ..... 8
V. Evaluation strategy and design ..... 10
5.1. Units of intervention, sampling, and analysis ..... 11
5.2. Sample size and power calculations ..... 12
5.3. Data collection and quality assurance ..... 12
5.4. Challenges encountered with the research design and in data collection ..... 15
5.5. Ethics ..... 15
5.6. Data processing and analytical software ..... 16
VI. Presentation of results ..... 17
6.1. Village-level summary statistics ..... 17
6.1.1. Economic activities ..... 18
6.1.2. Experience of natural disaster ..... 18
6.1.3. Village-level environmental violations ..... 19
6.2. Household-level summary statistics ..... 20
6.2.1. Income and economic activities ..... 21
6.2.2. Forest conservation practices ..... 22
6.2.3. Resilience and food security ..... 23
6.3. Village regulations ..... 24
6.4. Agricultural practices and knowledge ..... 25
VII. Discussion ..... 27
VIII. Challenges and shortcomings ..... 28
IX. Conclusion ..... 29
Appendix 1. Sampled *sucos* and *aldeias* by treatment and comparison groups ..... 30
Appendix 2. Propensity score estimation and covariate balance assessment ..... 32
Appendix 3. Evaluation matrix ..... 34
References ..... 36

# List of Tables

Table 1: List of software used during evaluation lifecycle....16
Table 2: Village demographic characteristics and primary water sources....17
Table 3: Village-level economic activities....18
Table 4: Village-level occurrence of natural disasters....19
Table 5: Village-level environmental violations....20
Table 6: Household summary statistics....21
Table 7: Household income and economic activities....22
Table 8: Shifting cultivation and forest conservation practices....23
Table 9: Resilience and food security....23
Table 10: Familiarity and compliance with village regulations....24
Table 11: Agricultural practices and knowledge....25
Table 12: Agricultural production values....26
Table 13: List of *sucos* and *aldeias*....30
Table 14: Evaluation matrix....34

# List of Figures

Figure 1: Project theory of change....6
Figure 2: Timeline of evaluation and project implementation activities....11
Figure 3: Data collection points in selected *sucos* in the Hatulia post-administration unit....14
Figure 4: Distribution of the propensity score....32
Figure 5: Standardized mean difference of matching variables....33

# Abbreviations

| | |
|---|---|
| **C4ED** | Center for Evaluation and Development |
| **CCVA** | Climate change vulnerability assessment |
| **CF** | Community forest |
| **CBNRM** | Community-based natural resource management |
| **CAPI** | Computer-assisted personal interviewing |
| **CSI** | Coping Strategy Index |
| **DiD** | Difference-in-differences |
| **FFS** | Farmer field school |
| **FIES** | Food Insecurity Experience Scale |
| **FCOTI** | Fundação Carbon Offset Timor |
| **GIS** | Geographic information system |
| **GHG** | Greenhouse gas |
| **HFC** | High-Frequency Check |
| **IRB** | Institutional Review Board |
| **JICA** | Japanese International Cooperation Agency |
| **LUCF** | Land use change and forestry |
| **MALFF** | Ministry of Agriculture, Livestock, Fisheries and Forestry |
| **NDC** | Nationally determined contribution |
| **NRM** | Natural resource management |
| **NGO** | Non-governmental organization |
| **PLUP** | Participatory land use planning |
| **PSM** | Propensity score matching |
| **QGIS** | Quantum GIS |
| **RADS** | Research and Development Solutions |
| **VR** | Village regulations |

# Executive summary

Timor-Leste faces considerable environmental and socio-economic challenges, including deforestation, forest degradation, and climate vulnerability, which threaten livelihoods, food security, and biodiversity. In response, the SAP21 project "*Community-Based Landscape Management for Enhanced Climate Resilience and Reduction of Deforestation in Critical Watersheds*", implemented by the Japanese International Cooperation Agency (JICA), aims to promote sustainable land management and climate-resilient agricultural practices through community-based natural resource management (CBNRM). This initiative seeks to empower local communities by formalizing participatory land-use planning, enforcing village-level regulations, and implementing sustainable farming practices to mitigate environmental degradation. The impact evaluation follows a quasi-experimental design, using a propensity score-based approach to compare treatment and control villages, ensuring a rigorous assessment of the project's effects on deforestation, climate resilience, and community livelihoods. Baseline data collection was conducted across 50 villages (25 treatment, 25 control), surveying 1,002 households to capture socioeconomic, agricultural, and environmental indicators, in September and October 2024. The survey employed computer-assisted personal interviewing using KoboToolbox, ensuring real-time data entry and quality control.

The baseline results reveal systematic differences at the household level between treatment and control groups across various socioeconomic, environmental, and agricultural indicators. While village-level characteristics appear largely balanced, household-level disparities may influence programme outcomes and require careful consideration in the impact evaluation. Control households have a higher share of working members than treatment households, while treatment households have a higher share of retired members. Additionally, household heads in the treatment group are significantly more likely to have completed secondary education than those in the control group. While total household income is similar between groups, treatment households receive significantly higher remittances, indicating greater access to external financial support. Conversely, control households engage in a higher number of income-generating activities, suggesting greater economic diversification. Control households exhibit higher knowledge and adoption of climate-resilient farming techniques, including agroforestry, intercropping, organic manure use, and terracing, with all differences statistically significant except for pest and weed control. This suggests that control households may have a pre-existing advantage in agricultural capacity. Households in the treatment group are more likely to practice shifting cultivation and have larger planned areas for shifting cultivation, with significant differences observed. However, control households show greater involvement in forest conservation activities, indicating a potentially stronger conservation ethic at baseline.

Control households report higher food insecurity and greater reliance on coping strategies, with significant differences in the Coping Strategy Index (CSI) and Food Insecurity Experience Scale (FIES). Control households more frequently consume seed stock, purchase food on credit, reduce meal frequency, and rely on friends for help, suggesting higher vulnerability to food shortages at baseline. Significant differences exist in agricultural production values between groups. Control households report higher total agricultural production values, particularly for coffee, roots, and livestock, while treatment households report higher values for fruit and fish production. The large and statistically significant gap in coffee production may be structurally driven by geographic and ecological factors.

While village-level characteristics appear balanced, differences at the household level in education, income sources, agricultural practices, land use, food security, and production patterns suggest important pre-existing variations that may affect intervention outcomes. These differences must be accounted for in the impact evaluation through robust statistical adjustments to ensure an accurate assessment of the programme's effectiveness.

# I. Context

1. Timor-Leste, a small island nation, has faced considerable challenges since its 2002 independence, including rebuilding infrastructure, strengthening civil administration, and decreasing poverty. The country's GDP grew by 11.4 per cent in 2008, before growth moderated and fell below 10 per cent from 2010 onwards.[1] Offshore oil and gas production has boosted fiscal revenue and enabled most of the developmental progress in the past two decades. However, petroleum revenues have declined substantially as production from the Bayu-Undan field has approached depletion.[2] The Asian Development Bank projected GDP growth of 3.1 per cent in 2024 and 3.9 per cent in 2025, supported by robust private consumption driven by consumer credit, government transfers, personal remittances and tourist arrivals.[3] In 2014, 42 per cent of the population lived under the national poverty line.[4] Agriculture remains a major source of livelihoods in Timor-Leste, with more than half of the population engaged in agriculture, many of whom rely on subsistence farming. However, the sector remains characterized by low productivity and limited commercialization.

2. Timor-Leste is highly vulnerable to the impacts of climate change. Changes in rainfall patterns, extreme rainfall events, drought, sea level rise and associated coastal hazards pose risks to communities, livelihoods and ecosystems, including through impacts on food and water security, agricultural productivity, biodiversity and fisheries. To address these risks, Timor-Leste's nationally determined contribution (NDC) outlines commitment areas in climate risk governance, nature-positive growth and transition, low-carbon development, climate change adaptation and resilience building.[5] Timor-Leste ranked 129$^{th}$ out of 180 countries in the 2022 Environmental Performance Index, which evaluates nations on 40 indicators related to environmental health, ecosystem vitality, and climate change mitigation efforts.[6] According to Climate Watch, Timor-Leste emitted 6.19 MtCO2e in 2021, representing a 19 per cent decrease from 2011. The energy sector accounted for 76 per cent of emissions, followed by agriculture at 15 per cent and land use change and forestry (LUCF) at 7 per cent.[7]

3. Deforestation and forest degradation are important sources of greenhouse gas (GHG) emissions and present challenges to Timor-Leste's climate mitigation efforts. According to Global Forest Watch, Timor-Leste lost 32.7 kilohectares (kha) of tree cover between 2001 and 2023, equivalent to a 4.5 per cent decrease from the 2000 level and 17.6 MtCO2e in associated emissions. Annual tree-cover loss remained below 1.5 kha over the most recent five-year period. Shifting cultivation was the dominant driver of tree-cover loss, accounting for an average of 70 per cent of annual loss between 2001 and 2023.[8] Timor-Leste's NDC (2022–2030) identified overgrazing, farmland expansion, growing demand for fuelwood and charcoal, urban expansion, and logging among the main drivers of deforestation and forest degradation.[9] Shifting cultivation, monoculture and other unsustainable land management practices can contribute to soil degradation and erosion and reduce the capacity of landscapes to retain water, with implications for agricultural productivity and vulnerability to floods and droughts.

4. CBNRM is a decentralized approach that involves local communities directly in the management of natural resources. It seeks to integrate livelihood and conservation objectives by enabling communities to actively participate in resource-use planning, management and decision-making and to benefit from the sustainable management of natural resources. Effective CBNRM also recognizes

[1] Asian Development Bank, "Key Indicators Database, Timor-Leste, GDP (% annual change)."
[2] International Monetary Fund, *Democratic Republic of Timor-Leste: 2021 Article IV Consultation*.
[3] Asian Development Bank, *Asian Development Outlook September 2024*.
[4] World Bank, *Timor-Leste Economic Report: Moving Beyond Uncertainty*.
[5] Timor-Leste, *Nationally Determined Contribution Timor-Leste 2022–2030*.
[6] Yale Center for Environmental Law & Policy, "2022 Environmental Performance Index: Timor-Leste."
[7] Climate Watch, "Timor-Leste Country Profile: Climate Watch Data."
[8] Global Forest Watch, "Timor-Leste Dashboard."
[9] Timor-Leste, *Nationally Determined Contribution Timor-Leste 2022–2030*.

the importance of local institutions, traditional practices, and community knowledge. Social-ecological systems with a longer tradition of CBNRM and greater local buy-in appeared more resilient to environmental challenges, while governance challenges such as power imbalances and inequitable distribution of benefits could undermine resilience.[10] More broadly, effective natural resource management can contribute to environmental and livelihood outcomes. These practices focused on reducing pollutants, managing watersheds, and promoting sustainable land use, displaying a substantial positive effect on water quality, but barriers such as financial and technical constraints must be addressed for full effectiveness.[11] Despite these challenges, CBNRM provides an inclusive and participatory approach to natural resource management that can strengthen community involvement in conservation and resource management.[12]

[10] María del Mar Delgado-Serrano et al., "Influence of Community-Based Natural Resource Management Strategies in the Resilience of Social-Ecological Systems," *Regional Environmental Change* 18, no. 2 (2018).
[11] Mohammed Elmadani et al., "Systematic Review of the Impact of Natural Resource Management on Public Health Outcomes: Focus on Water Quality." *Resources* 13, no. 9 (2024).
[12] Munalula Melinda et al., "Community-Based Natural Resource Management: Exploring the Benefits and Challenges," *International Journal of Humanities, Art and Social Studies* 5, no. 3 (2020).

# II. Project description

5. The SAP021 "*Community-Based Landscape Management for Enhanced Climate Resilience and Reduction of Deforestation in Critical Watersheds*" project, implemented by JICA, aims to promote a proven CBNRM approach in the most food-insecure forest-adjacent upland communities in Timor-Leste. The CBNRM approach places participatory development of natural resource management (NRM) and climate-resilient agricultural practices at its core. The project's objective is to strengthen the capacity of vulnerable upland communities and key government and non-government field officials to mitigate forest degradation and deforestation. This will be achieved by implementing the CBNRM mechanism within four priority watersheds: Laclo, Comoro, Tafara, and Caraulun. The CBNRM mechanism spans approximately 3.5 years per cycle and consists of three components shown in Figure 2: (i) participatory land use planning (PLUP) with climate change vulnerability assessment (CCVA), (ii) subsequent monitoring of the PLUP result, and (iii) the implementation of microprogrammes. It begins with PLUP with CCVA over 4–5 months to help communities identify their resources and vulnerabilities. This will lead to the creation of a future land use plan and village regulations (VRs), formalized through the traditional *Tara Bandu* ceremony – a traditional ceremony that formalizes custom-based NRM rules with State recognition.[13] This tradition is recognized in Timor-Leste's environmental law as a custom that is an integral part of the culture, regulating the relationship between people and the environment.[14] The ceremony has been restored since the country's independence, after being abolished and replaced by national laws during the Indonesian occupation. *Tara* means to hang, and *Bandu* means prohibition in Tetum, and it has demonstrated its effectiveness as an institutional mechanism for protecting forests, wildlife, water sources, sacred sites, and community property rights.[15] Monthly monitoring of VR enforcement, supported by external bodies, addresses incidents like forest fires, illegal exploitation, and free-grazing animals. Parallel microprogrammes, such as sustainable upland farming with sloping agricultural techniques, tree planting, and small-scale cottage industry development, promote shifts to climate-resilient livelihoods.

6. The CBNRM mechanism has been tested and proven effective over 15 years of JICA's work in Timor-Leste. The initial study (2005–2010) focused on watershed management in the Laclo and Comoro River Basins. Building on this, Phase 1 (2010–2015) established CBNRM models in six Aileu villages and strengthened the capacity of local officials for implementation. Phase 2 (2016–2022) expanded efforts, creating a national CBNRM roadmap, enhancing stakeholder coordination, and building capacity among broader groups, including government officials, local non-governmental organizations (NGOs), and development partners. The success of CBNRM initiatives led the Timor-Leste government, specifically the General Director of Forest of the Ministry of Agriculture, Livestock, Fisheries and Forestry (MALFF), to develop a CBNRM roadmap (2021–2030) over 14 priority watersheds as part of implementing the General Forest Regime Law which endorsed community forest (CF) and CBNRM as the main policy tool for sustainable forest management in Timor-Leste. This proposed project marks the initial phase of the national programme and targets four out of the 14 priority watersheds outlined in the CBNRM roadmap.

7. The project aims to reduce GHG emissions by addressing deforestation and forest degradation while enhancing carbon sequestration through reforestation, potentially cutting 4.4 MtCO2e over the project lifespan. This impact will be achieved by decreasing unsustainable practices like shifting cultivation, illegal exploitation, and free overgrazing while promoting sustainable, climate-adaptive livelihoods. This approach enables local farmers to use their land more productively and generate alternative income sources without harming existing forests. The project aims to directly benefit 74 villages,

[13] José Pinto Casquilho and Xisto Martins, "On the Influence of Customary Law 'Tara Bandu' in Community-Based Forest Management at the Mountains of Timor-Leste."
[14] Timor-Leste, *Decreto-Lei N.o 26/2012: Lei de Bases do Ambiente*.
[15] Marlon Cardinoza, "Reviving Traditional NRM Regulations (Tara Bandu) as a Community-Based Approach of Protecting Carbon Stocks and Securing Livelihoods," In *Carbon Forestry: Who Will Benefit?*

accounting for 34 per cent of the areas targeted in the CBNRM roadmap by 2030, and indirectly empower approximately 148,140 people, or 12 per cent of Timor-Leste's population. These 74 villages, situated within key watersheds, have been tentatively identified as target areas.

8. This project plans to make a considerable contribution to Timor-Leste's NDC by implementing two of the three key LUCF mitigation actions and three of the nine priority adaptation measures. Specifically, the project will (i) tackle the primary causes of forest degradation, such as unsustainable land use and shifting cultivation, to lower $CO_2$ emissions in the LUCF sector, (ii) drive policy transformation by embedding CBNRM as a core mitigation strategy in national policies and laws, and (iii) promote sustainable, climate-resilient livelihoods that are adaptable, community-accepted, and reduce reliance on forest resources. The project consists of four components, each containing specific activities. These activities are closely connected to achieving the project outputs, as illustrated in the section III.

9. Component 1: Establishment of a community-driven NRM system
   (a) Participatory assessment and planning
   (b) Participatory sustainable NRM with village regulations
   (c) Enhancement of local coordination for sustainable NRM at the sub-watershed level

10. Component 2: Implementation of microprogrammes/farmer field schools (FFS)
   (a) Assistance in the transformation of conventional farming and livelihood practices into low-emission and sustainable ones
   (b) Development of an incentive mechanism for reforestation
   (c) Promotion of CF as a sustainable forest management scheme
   (d) Capacity-building for MALFF field officials responsible for the target watersheds on CBNRM

11. Component 3: Institutional and capacity development for scale-up of CBNRM/CF beyond the target areas
   (a) Issuance of new government order or circular for promotion of CBNRM/CF
   (b) Capacity building of MALFF and NGO officials working in non-target watersheds
   (c) Institutionalization of the project outputs to enhance sustainability and scalability
   (d) Effective knowledge sharing among key stakeholders, particularly lawmakers

12. Component 4: Impact assessment
   (a) Establishment of baseline conditions for impact assessment
   (b) Evaluation of the project impacts with established methodologies for impact assessment
   (c) Development of technical references for assessment and evaluation of similar projects in future

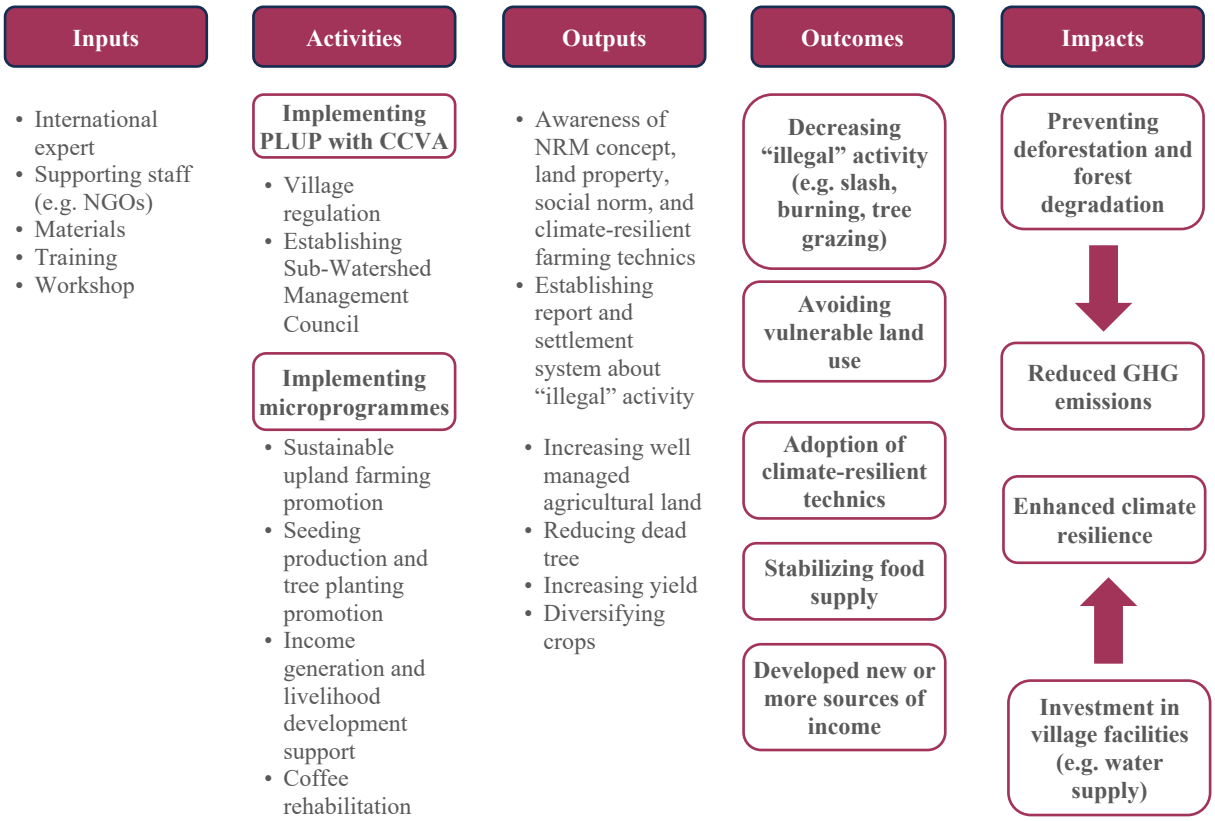
The evaluation, an activity under component 4, will focus on activities under components 1 and 2 related to the implementation of NRM and the microprogrammes/FFS.

## III. Theory of change

13. The theory of change for the CBNRM project (Figure 1) presents a structured pathway for achieving targeted impacts through interconnected steps, starting with inputs, progressing through activities, and resulting in measurable outputs, outcomes, and long-term impacts. The flow highlights a logical

progression, emphasizing the integration of participatory methods and sustainable practices for improving climate resilience, reducing deforestation, and enhancing community livelihoods.

**Figure 1: Project theory of change**




*Source*: Authors' elaboration based on project documentation.

### Inputs and activities

14. The foundational inputs include engaging international experts, supporting staff such as NGOs, and providing materials, training and workshops. These resources support key activities, such as PLUP and CCVA. This process involves engaging local communities to assess their land use patterns and identify climate vulnerabilities, fostering a participatory approach to resource management. Additionally, establishing VR and sub-watershed management councils empowers local governance structures to oversee and manage natural resources sustainably. In parallel, project component 2 promotes food security and livelihood diversification through microprogrammes. Activities include sustainable upland farming practices, seed production, tree planting, income-generating initiatives, and rehabilitating coffee production. These efforts aim to transform traditional livelihood practices into sustainable, climate-resilient alternatives.

### Outputs

15. The outputs generated from these activities include increased community awareness of NRM concepts, land property rights, social norms, and climate-resilient farming techniques. Reporting and settlement systems for illegal activities, such as slash-and-burn agriculture and free grazing, are established to reinforce compliance and promote sustainable practices. Other outputs involve better-managed agricultural land, increased crop yields, diversified crops, and improved community capacity for sustainable livelihoods.

### Outcomes and impacts

16. The project aims for transformative outcomes positively influencing the community's behaviour and practices. This includes reducing unsustainable activities, adopting climate-resilient techniques, stabilizing food supplies, and creating alternative income sources. The participatory approach fosters a sense of ownership and accountability among community members, enhancing their resilience to external and environmental challenges. This leads to tangible impacts, including preventing deforestation and forest degradation, reducing GHG emissions, and enhancing overall climate resilience within the community.

17. The impacts of these activities and outcomes are transformative, contributing to broad-scale goals such as preventing deforestation and reducing forest degradation. The project's impacts extend to reducing deforestation, preventing forest degradation, and lowering GHG emissions. The project contributes to long-term climate resilience and enhanced living conditions in forest-adjacent communities by improving resource management and promoting sustainable livelihoods. The landscape's carbon sequestration capacity is enhanced. Investments in critical village infrastructure, such as water supply systems, further support these efforts and ensure lasting benefits for the targeted areas.

# IV. Evaluation questions and indicators

18. The purpose of this evaluation is to assess the impact of SAP21 on climate resilience, deforestation reduction, and community welfare within Timor-Leste's critical watersheds. The evaluation aims to provide structured, evidence-based insights through a series of specific, measurable, attributable, realistic, and time-bound (SMART) indicators aligned with each key evaluation question. The evaluation is guided by the overarching question of **the extent to which a community-based landscape management approach enhances climate resilience and reduces deforestation among communities in the targeted areas of Timor-Leste**. To answer this, the evaluation examines several focused questions that drive the evaluation. Each question is paired with targeted indicators to ensure a precise and comprehensive assessment.

EQ1. To what extent are participatory land use and forest management planning and village regulations effective in promoting sustainable, climate-resilient practices?

19. To answer this evaluation question, indicators measuring land use change will be used. These indicators capture shifts in agricultural practices, such as reductions in slash-and-burn agriculture, increases in conservation agriculture methods, adoption of reforestation practices, and modifications to grazing patterns. Increased rates of conservation agriculture and reforestation suggest positive shifts towards sustainable land use and reduce slash-and-burn and grazing on vulnerable land indicate compliance with project goals.

EQ2. What changes in awareness and adherence to village-level natural resource management regulations are observed among households in treatment areas compared to control areas?

20. This evaluation question measures the percentage of households that know about village-level NRM regulations (awareness) as well as the percentage of households adhering to the established NRM regulations (compliance). High levels of awareness and compliance indicate that the project is successfully communicating and enforcing regulations aimed at sustainable land and resource management.

EQ3. What is the level of community satisfaction and perceived effectiveness of local governance structures among households in treatment areas compared to control areas?

21. This question looks into the rate of respondents' satisfaction with governance structures or processes on a scale from very dissatisfied to very satisfied. Households rate their satisfaction with local governance processes using a Likert scale (e.g. 1–5). Aggregate scores give an overall sense of community perception. Higher satisfaction ratings suggest that the project is effectively enhancing the community's trust in and satisfaction with local governance, reflecting the perceived legitimacy and effectiveness of local governance efforts.

EQ4. How do sustainable upland farming practices affect household income and food security in treatment areas relative to control areas?

22. This evaluation question will be answered through the investigation of average income per household from all sources. The household survey used for data collection captures income data from multiple sources (e.g. agricultural activities and non-farm income). These are summed to calculate the monthly income for each household and then averaged across treatment and control groups for comparison. Increases in average household income suggest that sustainable farming practices positively impact economic resilience, providing households with more stable and diversified income streams.

EQ5. What is the level of adoption of climate-resilient cropping techniques among households in treatment areas, and how has this adoption influenced agricultural productivity and resilience to climate change?

23. Adoption rates of climate-resilient techniques will look at the percentage of households practising climate-resilient agricultural practices, such as drought-resistant crops. Adoption rates are constructed

by asking households about specific climate-resilient practices they have adopted, with responses aggregated to calculate the percentage of households using each technique. High adoption rates and improved productivity suggest that climate-resilient practices have enhanced agricultural efficiency and resilience, contributing to the community's overall adaptability to climate variability.

EQ6. To what extent has income generation and livelihood development support contributed to household income diversification in treatment areas compared to control areas?

24. The number of different income sources per household count different income-generating activities engaged in by each household. Household surveys capture the range of income sources (e.g. agriculture, wage labour, small businesses). Responses are coded to reflect income diversity and summed to calculate the total number of sources per household. A greater number of income sources indicates increased diversification, suggesting that the project is successfully enhancing households' economic resilience by reducing reliance on a single income source.

EQ7. What changes have occurred in household food security and resilience as a result of project interventions?

25. The food security status of the household is measured by the frequency of food shortages, the FIES, and the CSI. Surveys ask households about their experiences with food shortages and use standardized scales (FIES and CSI) to quantify food security. FIES scores are calculated based on responses to questions about food availability, while CSI measures coping mechanisms used in times of shortage. Lower frequency of food shortages and improved FIES/CSI scores suggest enhanced food security and resilience, indicating that the project has positively impacted households' ability to cope with food-related risks.

# V. Evaluation strategy and design

26. The project evaluation follows a quasi-experimental design using propensity score matching (PSM) to construct a comparison group for estimating programme impacts. Because treatment villages were selected based on the timing of interventions within each watershed, as well as watershed-level eligibility criteria described in the CBNRM roadmap – such as forest cover and exposure to climate-related risks (e.g. flooding, erosion, or drought) – rather than through random assignment, systematic differences may exist between treatment and control areas. PSM helps address this issue by matching treated units with comparable non-treated units based on observable characteristics, thereby improving the credibility of the counterfactual. However, because PSM only accounts for observed characteristics, additional robustness checks will be conducted during the endline analysis.

27. PSM is selected as the primary method because treatment villages were selected based on specific criteria rather than random assignment.[16] This introduces geographic, environmental, and socio-economic heterogeneity between treatment and non-treatment areas. Under these conditions, constructing a matched comparison group helps address observable differences and improves comparability between treatment and comparison households along characteristics that may influence project participation and outcomes.

28. While PSM provides an important tool for addressing selection bias in observational settings, PSM alone may not fully eliminate bias arising from unobserved factors. Therefore, the evaluation also considers complementary estimators, such as inverse probability weighted regression adjustment (IPWRA) and, difference-in-differences (DiD) using the panel structure of the data. These approaches provide additional robustness checks and strengthen causal inference where underlying assumptions, such as common support and parallel trends are satisfied.

29. A two-stage propensity score approach is used to construct the comparison group. First, village-level propensity scores were estimated using secondary data available prior to fieldwork (e.g. population size, land area, forest cover) for treatment villages and a broader pool of untreated villages. Villages within the region of common support were identified, and control villages were randomly selected from this subset to improve comparability with treatment villages on observable pre-intervention structural characteristics.

30. Second, household survey data were collected within the selected villages. The baseline dataset provides detailed pre-treatment household characteristics that will be used to estimate household-level propensity scores and construct matched treatment and comparison samples for the impact analysis. The same households will be followed at endline to assess programme impacts.

31. SAP21 comprises two primary treatment components that will be evaluated for their impact on household and village outcomes:

    (a) The **NRM and land use planning** component will implement community-based land use plans and establish village regulations aimed at sustainable resource management.

    (b) Additionally, **microprogrammes**, that is, FFS, will promote climate-resilient agriculture and reforestation, aiming to enhance food security and reduce deforestation-linked income reliance.

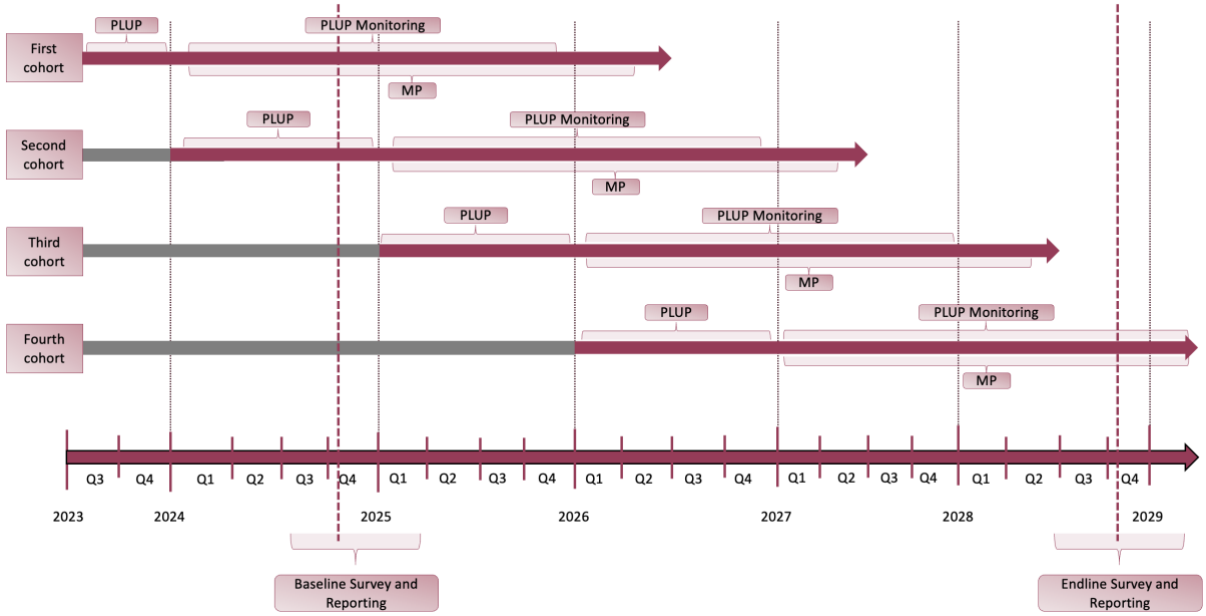
32. By investigating these two areas of intervention, the evaluation will isolate the effects of NRM planning and sustainable agriculture on various project outcomes, such as household income, food security, and community governance structures.

33. A timeline for evaluation and project activities has been established to guide implementation and follow-up data collection. Key points include baseline data collection (2024), phased rollout of PLUP

[16] Climate vulnerability, forest degradation, alignment with national watershed priorities.

and FFS activities, and scheduled endline data collection, which will capture the impact of interventions. This timeline ensures that evaluation activities are aligned with SAP21's project implementation, maintaining a separation between pre-intervention measurements and post-intervention evaluation.

**Figure 2: Timeline of evaluation and project implementation activities**




34. Through this planned evaluation design, SAP21 aims to generate reliable, evidence-based insights into the effects of community-based climate resilience strategies on households and communities in Timor-Leste, setting a foundation for long-term monitoring and policy recommendations.

## 5.1. Units of intervention, sampling, and analysis

35. The village is the primary unit of interventions, as community governance structures at this level shape and enforce policies and practices under the project. Interventions target villages and households within watershed regions, where villages serve as the primary units of intervention. For analysis, households are the primary unit, and outcomes are aggregated to capture community and household-level impacts.

36. Treatment villages were defined as those scheduled to receive the initial project intervention between 2024 and 2025, corresponding to the second and third cohorts shown in Figure 2. Control villages were selected from neighbouring watersheds, such as Be Lulic and Loes, with comparable geographic and environmental characteristics. This approach was intended to improve comparability between treatment and control villages under budget constraints to cover at national level. The detailed procedure for selecting treatment and control villages is described below.

37. To construct a pre-intervention sampling frame and ensure comparability between treatment and control villages, administrative, demographic, and environmental data were compiled at the village level. Spatial village boundary data from the Humanitarian Data Exchange of the United Nations Office for the Coordination of Humanitarian Affairs were combined with 2022 village-level

household counts from the National Institute of Statistics Timor-Leste census. Geographic characteristics, including village area (km²) and elevation, were derived using geographic information system (GIS) tools. Environmental variables were constructed using a 2010 forest classification dataset provided by the MALFF.

38. Within each selected village, households were sampled using a random procedure to ensure a representative household sample for the baseline survey.

## 5.2. Sample size and power calculations

39. The target sample size for the evaluation was approximately 1,000 households. The sample was distributed across the selected villages, with 20 households surveyed per village, allocated across multiple sub-villages (*aldeias*) within each village. This decision resulted from careful consideration of budgetary constraints and logistical feasibility. Given the project's objectives and expected effect sizes, the evaluation team sought to ensure that the sample size would allow for sufficient statistical power to detect meaningful impacts, while also remaining within budgetary limits.

40. Power calculations are typically conducted to determine the minimum sample size required to detect a meaningful effect size with a desired level of statistical power (usually 80 per cent) and significance (often 5 per cent). These calculations are particularly straightforward in the context of randomized controlled trials (RCTs), where random assignment minimizes selection bias and allows for more direct estimation of treatment effects. In RCTs, intra-cluster correlation values and other sample characteristics can be used to estimate the minimum detectable effect with a higher level of precision. However, in this study, we are employing PSM to create a comparison group that mimics randomization based on observed covariates. Because PSM is a quasi-experimental design, it introduces additional considerations for estimating treatment effects, and thus, power calculations derived from RCT formulas may only serve as rough approximations. The values provided here should be interpreted with caution as they are approximations, relying on assumptions such as balanced covariates and adequate common support between treatment and comparison groups. In practice, PSM's effectiveness in mimicking randomization depends on the quality of matching, sample size, and the representativeness of the treatment and control groups across covariates, which may influence the realized power of the study.

## 5.3. Data collection and quality assurance

41. The data collection process took place from October 2 to 30, 2024 and was carried out by a local partner NGO Fundação Carbon Offset Timor (FCOTI). Data collection covered 25 treatment villages across four target watersheds: Comoro, Laclo, Caraulun, and Tafara, alongside 25 control villages in neighbouring or comparable watersheds, such as Be Lulic and Loes. Data were collected in a dual-level design at both household and village levels.

42. The household survey questionnaire emphasized several key variables of interest, including agricultural production, land-use pattern changes, knowledge and adoption rates of climate-resilient agricultural practices, and food security. It also asked about employment, income, and expenditure and explored perceptions of village governance and trust. These variables align with the theory of change outlined in section III and the evaluation matrix introduced in section IV of this report.

43. The village survey questionnaire also addressed village infrastructure, NRM, gender equivalence, and natural disasters. In addition to the household survey, a village-level questionnaire was administered in each of the 50 sampled villages. The village survey was conducted with key informants, typically the village chief or other knowledgeable community representatives, to collect information on village infrastructure, NRM, governance arrangements, and exposure to natural hazards. These interviews were conducted once per village by trained enumerators using the same computer-assisted personal

interviewing (CAPI) system to ensure consistency and facilitate integration with the household-level dataset.

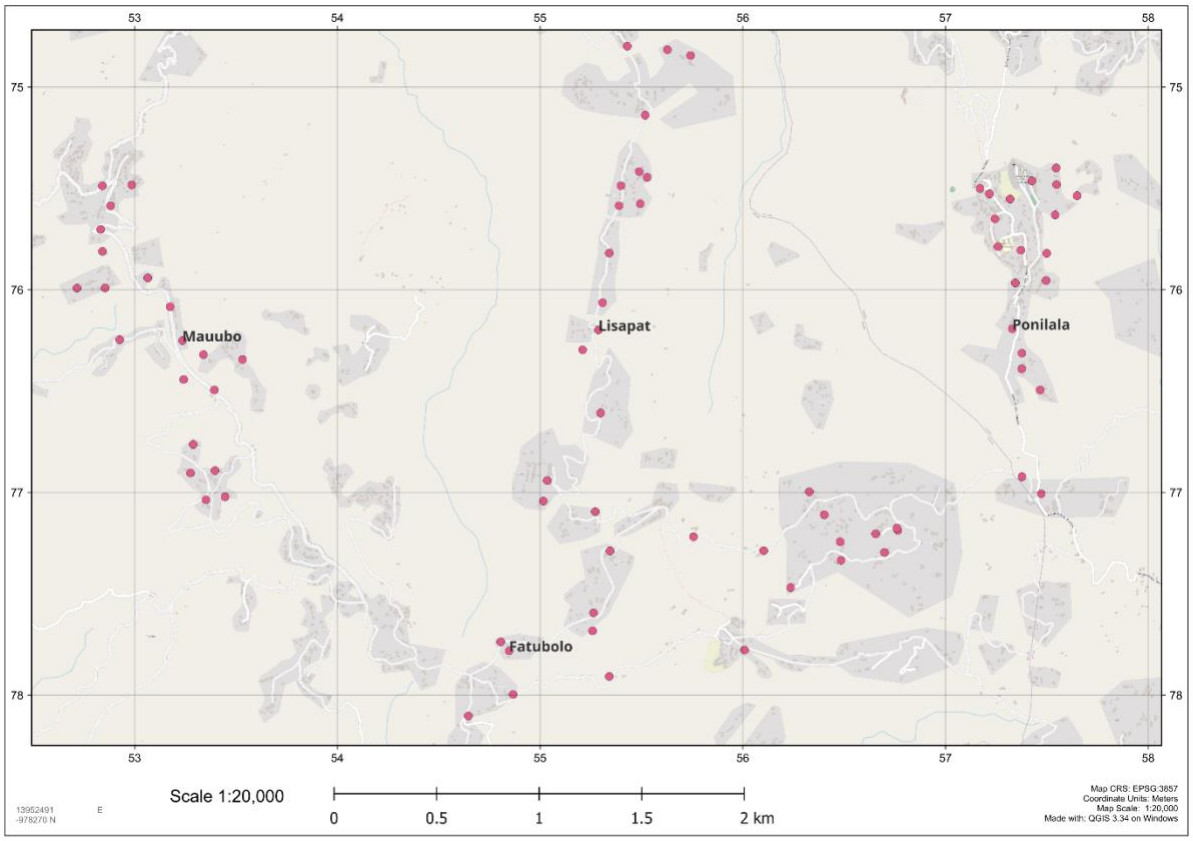
44. Collecting accurate data on agricultural production posed challenges, as respondents often struggled to quantify their crop harvests. To address this, the survey team, based on pilot results and observations, decided to ask respondents how many sacks of crop they harvested by different sizes of sacks. During the enumerators' training and pilot phase, measuring land size posed a challenge for some enumerators. To address this, extensive training sessions were conducted, and experienced enumerators were paired with less experienced ones to enhance skills and ensure accurate data collection. This approach led to noticeable improvements as data collection progressed.

45. Within each selected village, the survey team employed a random sampling approach. Enumerators first identified several *aldeias* within each village and then randomly selected households within these *aldeias*.

46. To operationalize the household sampling strategy, a GIS-based procedure was used to identify target *aldeias* within the 50 selected villages. Using the most recent spatial data sets provided by FCOTI, 291 *aldeias* were initially identified within the study villages. To ensure feasibility under fieldwork time and accessibility constraints, a 2-kilometre buffer was established around each village centre, and only aldeias intersecting this buffer were retained, reducing the sampling frame to 235 *aldeias*. From this subset, 105 target *aldeias* were randomly selected using Quantum GIS (QGIS). A small number of *aldeias* located in predominantly urban settings were excluded and replaced based on accessibility, population distribution, density, and size. The final sample of approximately 1,000 households was then distributed across the selected *aldeias*, with around 6–8 households per *aldeia* and a maximum of about 20 households per village.

47. To achieve a diverse sample within each *aldeia*, a random-walk method was used, starting from a central point and maintaining a minimum 10-minute walking distance between selected households. Data accuracy was further enhanced by using geo-fencing features within QField, an open-source mobile GIS application, ensuring that enumerators remained within the *aldeia* boundaries during data collection. However, some *aldeias* in which households are located in a linear manner or along the main road created a challenge for enumerators to adhere to the sampling approach and keep to the 10-minute walking distance between selected households (Figure 3).

**Figure 3: Data collection points in selected *sucos* in the Hatulia post-administration unit**




*Source*: Authors' elaboration based on project documentation.
*Note*: *suco* = [treatment and control] village.

48. The survey instrument, consisting of two questionnaires featuring approximately 400 and 150 questions across various thematic areas for households and village heads, respectively, was developed to capture a comprehensive range of data. Translation of the instrument into Tetum, one of Timor-Leste's official languages, was conducted with the assistance of JICA and FCOTI, while tablets displayed both English and Tetum versions to facilitate a clear understanding for enumerators during data collection. Before full implementation, a pilot survey was conducted in the last week of September 2024, which collected data from 53 households and two village heads outside of the target area, alongside a three-day training session for enumerators. Both questionnaires went through another round of revision afterward based on the observations and results from the pilot and training.

49. To ensure data quality and accuracy, field supervisors reviewed the data collected by the enumerators at the end of each day before uploading them to the server and sharing them with other partners. High-Frequency Check (HFC), developed by the Center for Evaluation and Development (C4ED) team based on the IPAcheck package from Innovations for Poverty Action, was conducted concurrently with data collection. This ensured consistent data quality monitoring. The C4ED team performed HFCs on 10 occasions, focusing on providing feedback and enhancing the data collection process, gradually reducing in frequency as improvements were made. The results and feedback of HFCs are frequently shared with the survey team, and detailed reports supported by Excel databases are delivered to highlight any issues. FCOTI reviewed HFC results and addressed identified concerns.

50. HFC involved multiple checks to ensure data integrity and consistency. These included verifying the uniqueness of survey IDs, checking for duplicate survey variables, monitoring data missingness,

identifying outliers in numeric variables, and validating constraint adherence in numeric data. In addition, dashboards for surveys and enumerators were created to provide continuous monitoring and feedback on data collection quality.

## 5.4. Challenges encountered with the research design and in data collection

51. One of the primary challenges was the selection and accessibility of villages. Although efforts were made to ensure minimal differences between treatment and control villages, pre-existing variations in exposure to conservation initiatives and socioeconomic conditions posed difficulties in achieving full comparability. Additionally, some villages were geographically remote and difficult to access, which delayed data collection and, in certain cases, restricted randomization in household selection. Coordination with local authorities also took longer than expected in some areas, as permissions and cooperation had to be established before enumerators could proceed.

52. The data collection process, which utilized CAPI via the KoboToolbox platform, presented further logistical and technical issues. While the use of digital tools aimed to enhance accuracy and efficiency, challenges arose in ensuring standardized responses across enumerators. Despite training sessions, some interviewers struggled with the consistent interpretation of survey questions, particularly for open-ended responses and qualitative indicators. Language barriers further complicated the process, as some questions, originally designed in English and translated into Tetum, led to misinterpretations or required additional clarification. Technical difficulties with tablets, including battery failures, syncing errors, and global positioning system malfunctions, occasionally disrupted the smooth collection of data.

53. The quasi-experimental approach remains valid, yet baseline imbalances in socioeconomic and environmental conditions must be carefully controlled during analysis. Additional robustness checks may also be required to mitigate biases resulting from attrition, data inconsistencies, and measurement errors. Despite these challenges, the baseline data provides a strong foundation for impact evaluation. Future phases will incorporate enhanced training for enumerators, refined measurement tools, and improved coordination with local stakeholders to minimize potential biases and strengthen the validity of the study's findings.

## 5.5. Ethics

54. Data collection was subject to ethical approval and underpinned by a commitment to integrity, honesty and competence. Participation in data collection was voluntary; all informants were asked to give informed consent. All interviews were conducted in private and confidentially, and anonymity was assured. Consent could be revoked at any time without any repercussions.

55. The evaluation team submitted an application for ethical approval to Research and Development Solutions (RADS), based in Islamabad, Pakistan, outlining our planned actions and adherence to ethical standards. The Institutional Review Board (IRB) of RADS convened on September 27, 2024, and granted approval for the ethical review, with the reference number RADS/IRB-C4ED/10-2024/061.

56. All data are saved in SharePoint, a cloud storage service provided by Microsoft. The strict protocols were in place for those who could access data, that is, only the designated investigative team members. Additionally, data was de-identified. Personal identifying information was separated and stored in an alternate dataset.

## 5.6. Data processing and analytical software

57. The research team used three main software tools. To optimize data collection and improve the efficiency and accuracy of the survey process, both the household- and village-level survey questionnaires were digitized and programmed for use with KoboToolbox, a widely utilized mobile data collection platform through which all data was collected. Integrating digital tools for CAPI was essential to the survey design, enabling real-time data entry, reducing human errors, and streamlining data management.

58. QField, built using QGIS software, supports accurate navigation and data gathering by displaying essential geographic features such as Google imagery, village boundaries, and household locations. QField is utilized by enumerators and field supervisors to accurately locate potential households and ensure adherence to the designated sampling strategy.

59. After this, Stata was used to clean the collected data, for example to identify and trim outliers and create indexes and indicators, as introduced in section IV. Following the data-cleaning process, the matching households and villages were estimated using Stata. The code used to clean the data and run the matching technique was recorded in Stata do-files. Table 1 lists all employed software.

**Table 1: List of software used during evaluation lifecycle**

| SOFTWARE | PURPOSE | OUTPUTS |
|---|---|---|
| KoboToolbox | Data collection | Baseline dataset |
| QField Mobile GIS Application | Mapping selected households in the targeted *aldeia* | Locating and reaching the target sample populations |
| Stata | Cleaning, creating indexes and indicators, conducting matching | Cleaned dataset<br>Various indexes and indicators<br>Matched households and villages |

*Source*: Authors' elaboration.

# VI. Presentation of results

60. This section presents descriptive baseline statistics for treatment and control groups prior to project implementation. The analysis is organized by level of observation, beginning with village-level characteristics and followed by household-level outcomes. The purpose is to assess baseline comparability and identify pre-existing differences that should be accounted for in the endline impact analysis.

## 6.1. Village-level summary statistics

61. Village-level statistics provide an overview of structural, environmental and institutional conditions across treatment and control villages (*sucos*). These results provide the empirical starting point for understanding structural differences between groups prior to any intervention and inform which covariates must be accounted for in the endline matching strategy. Table 2 summarizes demographic composition and primary water sources. Some differences are observed in water source usage, although only reliance on open dug wells is statistically significant.

**Table 2: Village demographic characteristics and primary water sources**

| | Control (N = 25) | Treatment (N = 25) | Total (N = 50) | Mean difference |
|---|---|---|---|---|
| Demographic characteristics | | | | |
| Total population | 3,625.800 (1,916.508) | 3,545.040 (2,337.798) | 3,585.420 (2,116.027) | –80.760 |
| Total households | 805.640 (435.355) | 758.040 (442.726) | 781.840 (435.217) | –47.600 |
| Share of males in the village | 0.489 (0.081) | 0.504 (0.015) | 0.497 (0.058) | 0.015 |
| Share of females in the village | 0.493 (0.096) | 0.497 (0.013) | 0.495 (0.068) | 0.004 |
| Share of children in the village | 0.150 (0.166) | 0.121 (0.212) | 0.136 (0.189) | –0.029 |
| Source of water | | | | |
| Piped gravity water | 0.560 (0.507) | 0.333 (0.482) | 0.449 (0.503) | –0.227 |
| Springs (natural) | 0.880 (0.440) | 0.958 (0.204) | 0.918 (0.344) | 0.078 |
| River | 0.160 (0.374) | 0.125 (0.338) | 0.143 (0.354) | –0.035 |
| Reservoir | 0.040 (0.200) | 0.000 (0.000) | 0.020 (0.143) | –0.040 |
| Well (open dug well) | 0.120 (0.332) | 0.000 (0.000) | 0.061 (0.242) | –0.120* |

*Source*: Authors' calculations based on baseline survey data.
*Notes*: Information obtained through village-level survey. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1.

### 6.1.1. Economic activities

62. Economic activities at the village level differ moderately between treatment and control areas, reflecting the ecological and livelihood structures of the respective watersheds. Control villages report higher engagement in aquaculture and forestry-related activities, suggesting stronger ties to resource extraction sectors and possibly different natural resource pressures. In contrast, treatment villages report a higher presence of service-sector activities (16 per cent compared to 4 per cent in control areas), which may indicate slightly more diversified or non-agricultural income opportunities. Livestock keeping is common in both groups but somewhat more prevalent in treatment villages. These differences matter because livelihood composition shapes households' exposure to climate and market risks and may influence how communities interact with natural resources. For example, forestry- and aquaculture-intensive villages may respond differently to new resource regulations than villages with stronger service-oriented or mixed economic profiles. While none of the differences are statistically significant, they highlight the underlying livelihood heterogeneity that the matching strategy aims to adjust for at endline.

**Table 3: Village-level economic activities**

| Main economic activities | Control (N = 25) | Treatment (N = 25) | Total (N = 50) | Mean difference |
|---|---|---|---|---|
| Fishery | 0.080 (0.277) | 0.040 (0.200) | 0.060 (0.240) | –0.040 |
| Aquaculture | 0.240 (0.436) | 0.080 (0.277) | 0.160 (0.370) | –0.160 |
| Livestock (incl. poultry) | 0.600 (0.500) | 0.720 (0.458) | 0.660 (0.479) | 0.120 |
| Forestry | 0.480 (0.510) | 0.280 (0.458) | 0.380 (0.490) | –0.200 |
| Processing industry | 0.080 (0.277) | 0.160 (0.374) | 0.120 (0.328) | 0.080 |
| Major commercial trade/retail | 0.040 (0.200) | 0.080 (0.277) | 0.060 (0.240) | 0.040 |
| Tourism | 0.280 (0.458) | 0.160 (0.374) | 0.220 (0.418) | –0.120 |
| Services | 0.040 (0.200) | 0.160 (0.374) | 0.100 (0.303) | 0.120 |
| Hunting | 0.040 (0.200) | 0.000 (0.000) | 0.020 (0.141) | –0.040 |

Source: Authors' calculations based on baseline survey data.
*Notes*: Information obtained through village-level survey. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1.

### 6.1.2. Experience of natural disaster

63. Villages in both treatment and control groups report exposure to a range of natural hazards, which aligns with the climatic vulnerability of upland areas in Timor-Leste. Most differences are small, but two patterns stand out. First, control villages report a higher incidence of strong winds and riverbank erosion. These are both indicators of environmental stress that can affect land use, infrastructure, and agricultural stability. Second, treatment villages report slightly lower exposure to several hazards, although the differences are generally not large enough to suggest systematic imbalance. These patterns indicate a need to account for environmental risk in the evaluation, as disaster exposure can influence agricultural decisions, adoption of conservation practices, and the effectiveness of CBNRM interventions. It also reflects the project's own targeting logic: treatment villages were selected partly based on vulnerability and watershed priorities, and some variation in hazard exposure is therefore expected.

**Table 4: Village-level occurrence of natural disasters**

| Experience of natural disaster | Control (N = 25) | Treatment (N = 25) | Total (N = 50) | Mean difference |
|---|---|---|---|---|
| Flooding | 0.720 (0.458) | 0.520 (0.510) | 0.620 (0.490) | –0.200 |
| Rainfall | 0.880 (0.332) | 0.800 (0.408) | 0.840 (0.370) | –0.080 |
| Drought | 0.760 (0.436) | 0.800 (0.408) | 0.780 (0.418) | 0.040 |
| Landslide | 0.920 (0.277) | 0.840 (0.374) | 0.880 (0.328) | –0.080 |
| Forest fire | 0.560 (0.507) | 0.480 (0.510) | 0.520 (0.505) | –0.080 |
| Strong wind | 1.000 (0.000) | 0.880 (0.332) | 0.940 (0.240) | –0.120* |
| Riverbank erosion | 0.840 (0.374) | 0.680 (0.476) | 0.760 (0.431) | –0.160 |

Source: Authors' calculations based on baseline survey data.
*Notes*: Information obtained through village-level survey. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1.

### 6.1.3. Village-level environmental violations

64. The baseline data on village-level environmental violations reveal some differences between the treatment and control groups, although none of these differences appear to be statistically significant. Table 5 reports mean occurrence rates for five types of violations along with standard deviations in parentheses. These statistics provide an initial picture of environmental challenges in the study areas before the intervention begins. The findings indicate that illegal logging and illegal hunting are more frequently reported in treatment villages than in control villages. Specifically, the mean incidence of illegal logging is 32 per cent in treatment villages compared to 20 per cent in control villages, reflecting a mean difference of 12 percentage points. Similarly, illegal hunting is reported in 32 per cent of treatment villages, compared to just 16 per cent of control villages, corresponding to a 16-percentage-point difference. While these differences suggest that treatment villages may have faced greater pressures from unsustainable resource extraction, they are not statistically significant, meaning that they could be due to random variation rather than systematic differences between the groups. Conversely, slash-and-burn practices appear to be slightly more common in control villages, where 76 per cent of villages reported the practice, compared to 68 per cent in treatment villages. This represents a mean difference of –8 percentage points, suggesting that control villages may have a marginally higher reliance on this traditional but environmentally harmful agricultural technique. However, this difference is also not statistically significant, indicating that slash-and-burn remains a widespread issue across both groups.

65. Table 5 also shows that illegal construction and forest fires occur at comparable levels in both treatment and control villages, with only slight differences. Illegal construction is reported in 20 per cent of treatment villages and 16 per cent of control villages, a 4-percentage point difference. Likewise, forest fires are observed in 36 per cent of treatment villages versus 28 per cent of control villages, an 8-percentage point difference. As with the previous variables, neither of these differences is statistically significant, meaning they do not necessarily represent a systematic imbalance between the groups at baseline. The lack of statistically significant differences between treatment and control villages suggests that, at baseline, environmental violations are relatively balanced across the two groups. However, the higher prevalence of illegal logging and hunting in treatment villages suggests that these areas may have been experiencing greater pressures on natural resources before the intervention. This aligns with the project's goal of targeting areas where deforestation and resource

exploitation are significant concerns. Similarly, the relatively higher rate of slash-and-burn in control villages may indicate that these areas rely more on traditional farming methods, which could influence their responsiveness to the intervention's agricultural sustainability components.

**Table 5: Village-level environmental violations**

| | CONTROL (N = 25) | TREATMENT (N = 25) | TOTAL (N = 50) | MEAN DIFFERENCE |
|---|---|---|---|---|
| Illegal logging | 0.200 (0.408) | 0.320 (0.476) | 0.260 (0.443) | 0.120 |
| Illegal hunting | 0.160 (0.374) | 0.320 (0.476) | 0.240 (0.431) | 0.160 |
| Slash-and-burn | 0.760 (0.436) | 0.680 (0.476) | 0.720 (0.454) | –0.080 |
| Illegal construction | 0.160 (0.374) | 0.200 (0.408) | 0.180 (0.388) | 0.040 |
| Forest fire | 0.280 (0.458) | 0.360 (0.490) | 0.320 (0.471) | 0.080 |

Source: Authors' calculations based on baseline survey data.
*Notes*: Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1.

## 6.2. Household-level summary statistics

66. The household-level results describe the socioeconomic, environmental and behavioural characteristics of sampled households before project implementation. These indicators relate directly to the pathways in the theory of change, including income diversification, food security, sustainable agricultural practices, and natural resource use.

67. Table 6 provides baseline descriptive statistics for key household-level characteristics across treatment and control groups. The results include mean values for each group, the total mean, and the mean differences between groups, with standard deviations reported in parentheses. On average, households in both the treatment and control groups have approximately seven members, with no statistically significant difference between the groups. Yet, a statistically significant difference is observed in the share of household members engaged in work, with the control group having a higher share (41.0 per cent) compared to the treatment group (38.8 per cent). The treatment group also has a higher share of retired members (6.5 per cent) compared to the control group (5.0 per cent). The share of household heads with completed secondary education is significantly higher in the treatment group (32.1 per cent) compared to the control group (16.8 per cent), with a large and significant mean difference of 0.154.

68. Households in the treatment group have lived in their current location for fewer years (30.3 years) than those in the control group (36.2 years). Households in the treatment group report significantly larger access to agricultural land (0.86 hectares) compared to the control group (0.72 hectares), but no significant differences are observed between the groups in terms of engagement in selling agricultural products or other business activities. These results suggest that while most household-level characteristics are balanced between groups, some significant differences in education, agricultural land access, and the share of retired and working household members highlight potential contextual variations that may influence outcomes.

**Table 6: Household summary statistics**

| | CONTROL (N = 501) | TREATMENT (N = 501) | TOTAL (N = 1,002) | MEAN DIFFERENCE |
|---|---|---|---|---|
| Total HH members | 6.976 (2.828) | 7.034 (2.831) | 7.005 (2.828) | 0.058 |
| Share of HH members in education | 0.400 (0.193) | 0.420 (0.191) | 0.410 (0.192) | 0.019 |
| Share of HH working members | 0.410 (0.174) | 0.388 (0.163) | 0.399 (0.169) | –0.023** |
| Share of HH retired members | 0.050 (0.111) | 0.065 (0.117) | 0.058 (0.114) | 0.015** |
| Age of HH head | 50.230 (13.428) | 49.295 (14.516) | 49.762 (13.983) | –0.934 |
| The HH head is married | 0.834 (0.372) | 0.870 (0.336) | 0.852 (0.355) | 0.036 |
| Completed primary education dummy | 0.246 (0.431) | 0.259 (0.439) | 0.252 (0.435) | 0.014 |
| Completed secondary education dummy | 0.168 (0.374) | 0.321 (0.467) | 0.245 (0.430) | 0.154*** |
| Years lived in current location | 36.201 (18.520) | 30.271 (18.593) | 33.273 (18.782) | –5.930*** |
| Hectares of accessible land for agriculture | 0.715 (0.508) | 0.863 (1.042) | 0.789 (0.823) | 0.148** |
| Engages in selling agricultural products | 0.884 (0.320) | 0.838 (0.369) | 0.861 (0.346) | –0.046 |
| Engages in other business activities | 0.269 (0.444) | 0.255 (0.437) | 0.262 (0.440) | –0.014 |

Source: Authors' calculations based on baseline survey data.
*Notes*: HH = household. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

### 6.2.1. Income and economic activities

69. Table 7 displays the baseline balance statistics for income and economic activities. Most variables are balanced between the two groups, indicating comparability at baseline. However, significant differences are observed in two variables. First, remittance income is significantly higher in the treatment group (USD 183.05) compared to the control group (USD 92.82), and the difference is statistically significant. This suggests that treatment households may have better access to external financial support through remittances. Second, the total number of income-generating activities per household is higher in the control group (2.93) compared to the treatment group (2.47), with the difference also statistically significant. While some variables, such as veterans' pension income and monthly income from family livestock activities, show large differences in means between the groups, these differences are not statistically significant, suggesting that they may be due to random variation rather than systematic imbalance.

**Table 7: Household income and economic activities**

Unit: USD

| | CONTROL (N = 501) | TREATMENT (N = 501) | TOTAL (N = 1,002) | MEAN DIFFERENCE |
|---|---|---|---|---|
| Total household income in a typical month | 482.676 (378.428) | 468.831 (455.218) | 475.838 (417.963) | –13.84 |
| Monthly income from family crop production (winsorized) | 218.095 (258.917) | 198.531 (332.227) | 209.521 (293.263) | –19.57 |
| Monthly income from family livestock activities (winsorized) | 170.935 (250.802) | 229.780 (273.671) | 197.178 (262.623) | 58.85 |
| Monthly income from non-agricultural self-employment (winsorized) | 90.725 (136.685) | 136.093 (169.842) | 114.354 (155.977) | 45.369* |
| Agricultural wage income per month (winsorized) | 91.258 (120.907) | 85.755 (108.243) | 87.942 (113.095) | –5.50 |
| Public transfer income per month (winsorized) | 163.741 (102.931) | 185.881 (97.958) | 175.061 (100.915) | 22.14 |
| Remittance income per month (winsorized) | 92.821 (79.062) | 183.047 (211.935) | 131.838 (157.451) | 90.225*** |
| Veterans Pension income per month | 135.538 (96.068) | 272.062 (425.521) | 187.548 (276.360) | 136.52 |
| Other income per month (winsorized) | 102.476 (99.230) | 122.365 (98.889) | 108.508 (99.346) | 19.89 |
| Wage employment income per month (winsorized) | 172.088 (157.492) | 189.906 (184.413) | 180.727 (170.979) | 17.82 |
| Total number of income-generating activities engaged in by each household | 2.932 (1.201) | 2.467 (1.278) | 2.700 (1.261) | –0.465*** |

Source: Authors' calculations based on baseline survey data.
*Notes*: Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

### 6.2.2. Forest conservation practices

70. Table 8 shows forest conservation practices used by households in the treatment and control villages. Households in the treatment group are more likely to practice shifting cultivation compared to the control group, and this difference is statistically significant. They also report larger planned areas for shifting cultivation and are more likely to have previously practiced shifting cultivation, both of which are significant differences. In contrast, involvement in forest conservation is higher in the control group than in the treatment group, and this difference is also statistically significant.

71. A notable difference can be seen in the forest area cleared for cultivation, which is substantially larger in the treatment group. However, this difference is not statistically significant, suggesting it may be due to random variation. Other variables, such as growing crops with shifting cultivation and forest areas cleared for cultivation, do not show significant differences.

**Table 8: Shifting cultivation and forest conservation practices**

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Practices shifting cultivation | 0.168 (0.374) | 0.354 (0.479) | 0.261 (0.439) | 0.186*** |
| Grows crops with shifting cultivation | 0.917 (0.278) | 0.893 (0.310) | 0.900 (0.300) | –0.024 |
| Plans for shifting cultivation area | 1.524 (0.526) | 1.757 (0.525) | 1.682 (0.536) | 0.233** |
| Previously practiced shifting cultivation | 0.046 (0.209) | 0.102 (0.304) | 0.070 (0.256) | 0.057** |
| Forest area cleared for cultivation (ha) | 0.608 (0.446) | 0.558 (0.454) | 0.575 (0.450) | –0.049 |
| Involved in forest conservation | 0.699 (0.459) | 0.549 (0.498) | 0.625 (0.484) | –0.150*** |

Source: Authors' calculations based on baseline survey data.
*Notes*: ha = hectares. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

### 6.2.3. Resilience and food security

72. Table 9 presents baseline conditions on resilience and food security. Households in the control group report higher values for the CSI and FIES, indicating greater reliance on coping strategies and higher levels of food insecurity than the treatment group, with both differences being statistically significant. Control households also report more frequent days of consuming seed stock, relying on friends and families for help, reducing meal frequency, and purchasing food on credit, all of which are significantly higher than in treatment households. In addition, control households more frequently report worrying about food shortages and experiencing limited food variety.

73. Differences for some variables, such as days skipping entire meals and days prioritizing working household members over others, are small and not statistically significant. Overall, these patterns suggest that control households experienced greater food insecurity and relied more heavily on coping strategies than treatment households at baseline.

**Table 9: Resilience and food security**

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Coping Strategy Index (CSI) | 6.659 (5.845) | 4.467 (4.548) | 5.563 (5.348) | –2.192*** |
| Food Insecurity Experience Scale (FIES) | 4.118 (2.003) | 2.812 (1.852) | 3.465 (2.035) | –1.305*** |
| Days using less preferred food (last 7 days) | 1.457 (1.370) | 1.297 (1.637) | 1.377 (1.510) | –0.160 |
| Days with limited meal portions | 0.729 (1.102) | 0.393 (0.963) | 0.561 (1.048) | –0.335*** |
| Days reducing meal frequency | 0.495 (0.944) | 0.232 (0.864) | 0.363 (0.914) | –0.263*** |
| Days reliant on friends for help (last 7 days) | 0.601 (0.870) | 0.339 (0.693) | 0.470 (0.797) | –0.261*** |
| Days purchasing food on credit (last 7 days) | 0.707 (0.767) | 0.315 (0.580) | 0.511 (0.707) | –0.391*** |
| Days sending children to eat with neighbours | 0.118 (0.341) | 0.052 (0.365) | 0.085 (0.354) | –0.066*** |
| Days prioritizing working HH members over others | 0.134 (0.415) | 0.154 (0.715) | 0.144 (0.584) | 0.020 |

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Days consuming seed stock (last 7 days) | 1.413 (1.421) | 0.968 (1.219) | 1.191 (1.342) | –0.445*** |
| Days gathering wild food (last 7 days) | 0.679 (1.011) | 0.582 (1.272) | 0.630 (1.149) | –0.097 |
| Days needing financial help | 0.259 (0.614) | 0.076 (0.320) | 0.168 (0.498) | –0.184*** |
| Days skipping entire meals | 0.068 (0.275) | 0.074 (0.452) | 0.071 (0.374) | 0.006 |
| Worry about food shortage | 0.896 (0.305) | 0.687 (0.464) | 0.791 (0.406) | –0.210*** |
| Lack of healthy food | 0.717 (0.451) | 0.599 (0.491) | 0.658 (0.475) | –0.118** |
| Limited variety of food | 0.733 (0.443) | 0.451 (0.498) | 0.592 (0.492) | –0.281*** |
| Had to skip a meal | 0.627 (0.484) | 0.357 (0.480) | 0.492 (0.500) | –0.269*** |
| Ate less than desired | 0.601 (0.490) | 0.287 (0.453) | 0.444 (0.497) | –0.313*** |
| Household ran out of food | 0.305 (0.461) | 0.271 (0.445) | 0.288 (0.453) | –0.034 |
| Felt hungry but did not eat | 0.154 (0.361) | 0.104 (0.305) | 0.129 (0.335) | –0.050* |
| Went a full day without eating | 0.086 (0.280) | 0.056 (0.230) | 0.071 (0.257) | –0.030 |

Source: Authors' calculations based on baseline survey data.
*Notes*: Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

## 6.3. Village regulations

74. Table 10 shows the baseline balance statistics for familiarity and compliance with village regulations. Both groups exhibit very similar levels of familiarity and compliance with village regulations, and no statistically significant differences are observed. The differences in means for both variables are minimal, suggesting that treatment and control households had comparable knowledge and adherence to village regulations at baseline.

**Table 10: Familiarity and compliance with village regulations**

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Familiarity with village regulations | 4.048 (0.379) | 4.044 (0.406) | 4.046 (0.392) | –0.005 |
| Compliance with village regulations | 4.072 (0.327) | 4.052 (0.369) | 4.062 (0.349) | –0.020 |
| Satisfaction with decision-making inclusivity | 3.984 (0.483) | 4.008 (0.418) | 3.996 (0.452) | 0.024 |
| Trust in community members | 4.060 (0.341) | 4.062 (0.399) | 4.061 (0.371) | 0.002 |
| Perceived effectiveness of village regulations | 4.048 (0.326) | 4.050 (0.309) | 4.049 (0.318) | 0.002 |

Source: Authors' calculations based on baseline survey data.
*Notes*: Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

## 6.4. Agricultural practices and knowledge

75. Table 11 presents baseline balance statistics for knowledge and practices related to agriculture. Control households consistently demonstrate higher levels of knowledge across all agricultural techniques, including agroforestry practices, intercropping, organic manure use, pest and weed control, crop residue use, and terracing. These differences are statistically significant for all the variables, except for "uses pest and weed control". The overall results indicate a more substantial adoption of agricultural practices among control households. This could suggest a potential advantage in agricultural capacity for control households at the start of the intervention.

**Table 11: Agricultural practices and knowledge**

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Knows agroforestry practices | 0.836 (0.370) | 0.765 (0.424) | 0.801 (0.399) | –0.071** |
| Knows improved seed planting | 0.958 (0.201) | 0.895 (0.307) | 0.927 (0.260) | –0.063*** |
| Knows intercropping practices | 0.958 (0.201) | 0.858 (0.349) | 0.908 (0.289) | –0.100*** |
| Knows organic manure use | 0.872 (0.334) | 0.798 (0.402) | 0.835 (0.371) | –0.074** |
| Knows pest and weed control | 0.870 (0.337) | 0.723 (0.448) | 0.797 (0.402) | –0.147*** |
| Knows crop residue use | 0.912 (0.283) | 0.737 (0.441) | 0.825 (0.380) | –0.175*** |
| Knows terracing practices | 0.766 (0.424) | 0.579 (0.494) | 0.673 (0.469) | –0.187*** |
| Practices agroforestry | 0.905 (0.294) | 0.778 (0.416) | 0.844 (0.363) | –0.127*** |
| Uses improved seeds | 0.944 (0.231) | 0.907 (0.290) | 0.926 (0.261) | –0.036* |
| Practices intercropping | 0.931 (0.253) | 0.842 (0.365) | 0.889 (0.314) | –0.089*** |
| Uses terracing practices | 0.755 (0.431) | 0.562 (0.497) | 0.672 (0.470) | –0.192*** |
| Uses pest and weed control | 0.885 (0.320) | 0.870 (0.336) | 0.878 (0.327) | –0.014 |
| Uses organic manure | 0.870 (0.337) | 0.769 (0.422) | 0.822 (0.383) | –0.100** |
| Uses crop residue | 0.891 (0.312) | 0.814 (0.390) | 0.856 (0.351) | –0.077** |

Source: Authors' calculations based on baseline survey data.
*Notes*: Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

76. The baseline results in Table 12 reveal significant differences in agricultural production values between treatment and control groups for several key categories. The control group consistently shows higher production values for total agricultural products, roots, coffee and livestock, while the treatment group reports higher values for fruits and fish. These differences suggest potential pre-existing disparities in agricultural production capacity and activity between the two groups, which should be accounted for in subsequent analyses to ensure accurate estimation of programme impacts. One key driver of these differences appears to be coffee production, which shows a large and statistically significant gap, with the control group reporting substantially higher estimated coffee values (USD 622.95 versus USD 301.09, $p < 0.01$). Given that coffee production is regionally concentrated and may only be implemented in certain villages, this imbalance could be structurally driven by geographic and ecological conditions.

77. It is important to note that these values represent total production values, not adjusted for land area or production volume. Therefore, observed differences may not only reflect differences in productivity or crop selection but also variations in the total amount of cultivated land across treatment and control groups. This highlights the need to account for differences in cultivated land area and household-level agricultural engagement in subsequent analyses to ensure accurate estimation of programme impacts.

**Table 12: Agricultural production values**

Unit: USD

| | Control (N = 501) | Treatment (N = 501) | Total (N = 1,002) | Mean difference |
|---|---|---|---|---|
| Total value of agricultural products | 1,100.602 (1,055.972) | 810.479 (831.138) | 957.064 (961.913) | –290.122** |
| Gross value of maize harvest | 204.962 (184.938) | 205.938 (300.324) | 205.571 (262.372) | 0.976 |
| Estimated value of grain | 538.056 (616.067) | 214.545 (274.968) | 415.345 (531.937) | –323.510 |
| Estimated value of vegetables | 263.468 (309.715) | 274.151 (379.688) | 269.175 (348.548) | 10.683 |
| Estimated value of roots | 174.315 (160.018) | 112.780 (143.319) | 138.049 (153.136) | –61.535*** |
| Estimated value of legumes | 262.311 (319.592) | 181.188 (240.669) | 213.190 (276.684) | –81.124 |
| Estimated value of fruits | 89.052 (196.388) | 172.380 (274.507) | 144.113 (253.488) | 83.328 |
| Estimated value of coffee | 622.954 (652.032) | 301.091 (417.595) | 489.146 (587.923) | –321.864*** |
| Estimated value of spices | 78.000 (204.245) | 77.810 (112.109) | 77.900 (160.309) | –0.190 |
| Estimated value of harvested trees | 409.909 (645.212) | 200.000 (346.571) | 274.484 (479.959) | –209.909 |
| Estimated value of NTFPs | 205.000 (264.134) | 186.250 (278.601) | 190.938 (266.353) | –18.750 |
| Estimated livestock value | 517.465 (483.191) | 406.868 (404.704) | 463.387 (449.581) | –110.597 |
| Estimated fish value | 28.750 (15.478) | 81.786 (65.622) | 70.000 (62.048) | 53.036* |

*Source*: Authors' calculations based on baseline survey data.
*Notes*: NTFP = non-timber forest products. Standard deviation in parentheses. The table reports balance statistics for baseline covariates across treatment and control groups. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors clustered at the *suco* level.

# VII. Discussion

79. The baseline results reveal a number of structural differences between treatment and control areas that matter for how the evaluation should be interpreted. At the village level, livelihood structures differ in ways that are consistent with watershed geography: control *sucos* show greater involvement in aquaculture and forestry-related activities, while treatment *sucos* have a somewhat stronger presence of service-sector activities. These patterns matter because the types of economic activities that dominate a village shape both exposure to climate and market risks and how households interact with natural resources. Villages with stronger forestry or aquaculture profiles may also respond differently to new resource management rules than villages with more diversified income structures. Exposure to natural hazards is broadly comparable, although slightly higher levels of strong winds and riverbank erosion in control areas point to localized structural vulnerabilities that could influence land use decisions over time. Differences in reported environmental violations are not statistically significant but still suggest slightly higher pressures from illegal logging and hunting in treatment villages and a marginally higher prevalence of slash-and-burn in control villages.

80. More pronounced differences emerge at the household level. Treatment households have higher levels of secondary education and greater access to agricultural land. These traits may facilitate the uptake of agricultural training or NRM practices. Control households, on the other hand, show broader income diversification and engage in more income-generating activities, which can act as informal buffers against shocks. Agricultural knowledge also differs substantially: control households consistently report higher familiarity with agroforestry, intercropping, terracing, crop residue management, and other sustainable techniques. These differences may reflect previous exposure to extension services or local agroecological conditions and imply that treatment households have more room for improvement, while control households may already be closer to saturation for certain practices. Land-use patterns show similar divergence: treatment households are more engaged in shifting cultivation, which is one of the major contributors to deforestation, while control households report higher involvement in forest conservation activities. Taken together, these patterns indicate that the two groups face different baseline pressures and opportunities in terms of land use, agricultural capacity, and environmental management.

81. Food security indicators reveal another consistent gap. Control households report higher food insecurity and rely more frequently on short-term coping strategies. Food insecurity shapes how households prioritize labour, how willing or able they are to change agricultural practices, and how resilient they are to environmental or economic shocks. These baseline vulnerabilities will influence how households respond to the intervention and must be taken into account when interpreting downstream outcomes.

82. Overall, the baseline data confirm that treatment and control areas are not identical and that several differences could independently influence projected outcomes. This is not unexpected given the targeting strategy and the non-random assignment of treatment. It does, however, show the need for a careful analytical strategy. The observed differences, particularly in education, agricultural knowledge, income diversification, food security, and land-use pressures, justify the use of matching and weighting approaches to adjust for observable imbalance, as well as the use of robust estimation methods such as inverse probability weighted regression adjustment, and where feasible, DiD. Ensuring that these baseline characteristics are properly accounted for at endline will be critical for isolating the effects of the intervention from pre-existing variation in capacity, vulnerability, and environmental conditions.

# VIII. Challenges and shortcomings

83. Several challenges and shortcomings were identified during this phase of the evaluation, leaving key considerations for the endline analysis.

84. First, the observed differences highlight the heterogeneity of communities and households within Timor-Leste, emphasizing the importance of a rigorously implemented quasi-experimental evaluation approach at endline. While control *sucos* were selected to achieve similarity with treatment villages based on specific criteria, notable deviations remain at the household level, particularly regarding agricultural practices targeted by the intervention.

85. In this study, some baseline differences between treatment and control groups are anticipated due to the non-random selection of *sucos* for treatment and control, rather than household-level randomization or randomized treatment assignment. While households were randomly sampled within villages to ensure representativeness, the village selection process was not random. Treatment villages were chosen based on specific eligibility criteria, including vulnerability to climate change, GHG emissions reduction potential, and alignment with national policies, as outlined in the funding proposal. In contrast, control villages were selected using matching based on secondary data, including population size, land area, land area squared, and dense forest area. These criteria do not fully align with the treatment villages' selection process, resulting in systematic differences between the groups, both observed and unobserved.

86. Second, challenges in accurately measuring outcomes of interest arise from several factors, including the capacity of enumerators and the recall bias of respondents. Enumerator capacity can affect the quality and consistency of data collection, particularly in complex or technical areas such as estimating agricultural yields, measuring land size, or understanding detailed household economic activities. Inadequate training or limited experience among enumerators may lead to inaccuracies, inconsistencies, or difficulties in probing respondents for reliable information.

87. Recall bias, on the other hand, poses a significant challenge, especially when respondents are required to provide retrospective data about agricultural practices, income sources, or food security experiences. Memory lapses, or a lack of precise records, can result in incomplete, skewed, or inconsistent data. This issue is further compounded in contexts where respondents are not accustomed to maintaining detailed records or participating in surveys.

88. These challenges highlight the importance of investing in rigorous enumerator training, employing robust data validation protocols, and utilizing tools such as memory aids, structured recall techniques, and real-time data collection methods (e.g. CAPI) to mitigate measurement errors.

89. Lastly, for the microprogramme, the specific beneficiaries within the treatment *aldeias* were not identified at the baseline stage. Instead, a random sampling approach was employed, meaning the actual number of treated households included in the baseline sample remains unknown until later stages. This uncertainty poses potential challenges for the evaluation, as the effective sample size of treated households may impact the statistical power needed to accurately estimate the microprogramme's impact. Ensuring adequate power will depend on identifying and verifying the proportion of treated households in the sample, which may require adjustments in the analysis or additional data collection to strengthen the reliability of impact estimates.

# IX. Conclusion

90. This baseline report documents the initial household and village conditions across the 25 treatment and 25 control *sucos* surveyed in October 2024 and provides the reference point for all subsequent impact estimation. While treatment and control areas share a number of structural similarities, the results also confirm systematic differences in agricultural knowledge, food security, livelihood diversification, land access, and land-use pressures. These differences are not unexpected given the project's targeting criteria and the non-random assignment of treatment, but they are consequential: they represent sources of observable imbalance that must be addressed analytically at endline.

91. The initial matching of control *sucos*, conducted using secondary village-level indicators, achieved broad geographic and demographic comparability but could not fully replicate the criteria used to select treatment villages. Nor did it incorporate the household-level characteristics now available through the baseline survey. As a result, some meaningful baseline differences remain. These baseline variables will form the basis for refined matching and weighting at the analysis stage. This ensures that the counterfactual is constructed using pre-intervention characteristics only, avoiding contamination of the comparison group. The unadjusted balance tables included in this report establish the factual pre-intervention distributions of key variables and help indicate which covariates require explicit adjustment through PSM, inverse probability weighting, and regression-based methods at endline. Where appropriate, DID may complement these approaches, provided that parallel trends can be demonstrated.

92. Several implications follow directly from the baseline results:

    (a) Baseline differences in education, agricultural knowledge, land access, food security, and income diversification must be controlled for at endline to avoid biased estimates of project effects.

    (b) Variables showing substantive imbalance will be incorporated into the matching and adjustment models to strengthen covariate balance and improve the credibility of causal inference.

    (c) The heterogeneity observed across *sucos* highlights the need for a transparent, well-justified quasi-experimental approach that fully leverages the information collected at baseline.

93. Taken together, the baseline data provide a solid and necessary foundation for the endline evaluation. By explicitly adjusting for pre-existing differences and grounding the analytical strategy in the patterns documented here, the evaluation will be well positioned to generate credible and policy-relevant evidence on the effectiveness of the CBNRM interventions in Timor-Leste.

# Appendix 1. Sampled *sucos* and *aldeias* by treatment and comparison groups

**Table 13: List of *sucos* and *aldeias***

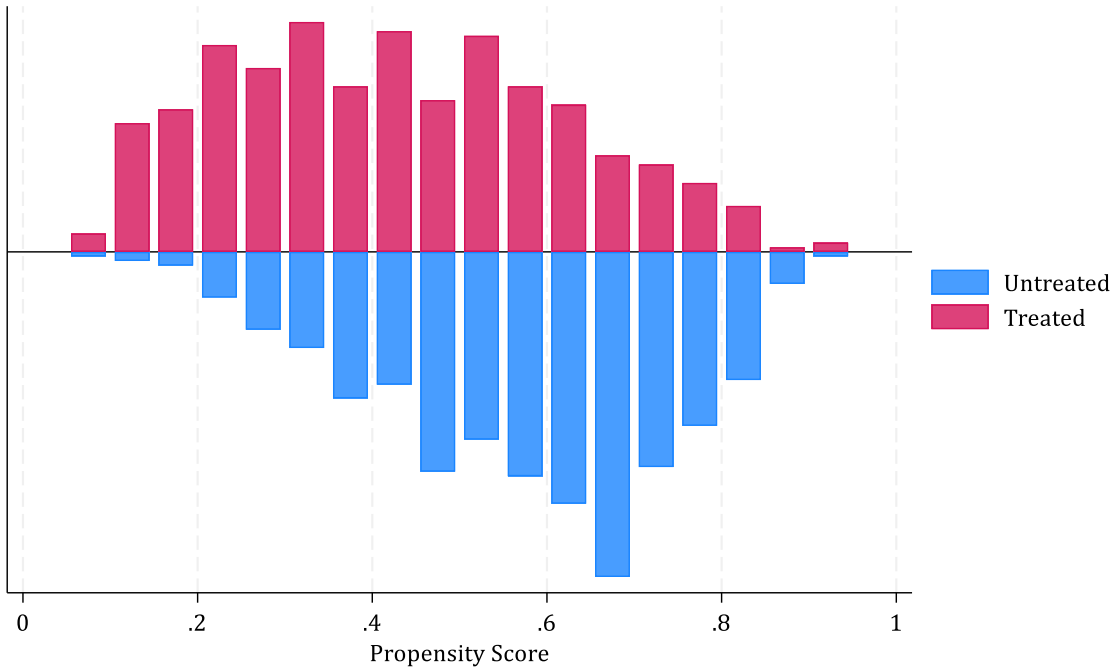
| N | *Suco* | *Aldeia* | Treatment |
|---|---|---|---|
| 1 | Aissirimou | Aituhularan | T |
| 2 | Aissirimou | Bessilau | T |
| 3 | Fahiria | Fahiria | T |
| 4 | Fahiria | Sarin | T |
| 5 | Fahiria | Sidole | T |
| 6 | Fatubossa | Coulau | T |
| 7 | Fatubossa | Fatubossa | T |
| 8 | Fatubossa | Liclaucana | T |
| 9 | Lahae | Eralolo | T |
| 10 | Lahae | Lacasori | T |
| 11 | Lahae | Lahae | T |
| 12 | Seloicraic | Colihoho | T |
| 13 | Seloicraic | Lio | T |
| 14 | Seloimalere | Cotbauru | T |
| 15 | Seloimalere | Hularema | T |
| 16 | Seloimalere | Kabasfatin | T |
| 17 | Acumau | Aimerahun | T |
| 18 | Acumau | Lerolica | T |
| 19 | Tulataqueo | Roluli | T |
| 20 | Tulataqueo | Samalete | T |
| 21 | Cassa | Civil_1 | C |
| 22 | Cassa | Civil_2 | C |
| 23 | Leolima | Groto | C |
| 24 | Leolima | Luro | C |
| 25 | Leolima | Suro Craic | C |
| 26 | Aituto | Aihou | T |
| 27 | Aituto | Lientuto | T |
| 28 | Aituto | Russulau | T |
| 29 | Horaiquic | Cartolo | T |
| 30 | Horaiquic | Lau Hel | T |
| 31 | Maubisse | Hato Luli | T |
| 32 | Maubisse | Lequi Tei | T |
| 33 | Maubisse | Teli Tuco | T |
| 34 | Atabae | Fatubesi | C |
| 35 | Atabae | Helesu | C |
| 36 | Colimau | Atublogo | C |
| 37 | Colimau | Manunia | C |
| 38 | Colimau | Tegoabe | C |
| 39 | Maliubu | Atu Lara | C |
| 40 | Maliubu | Hatu Udu | C |
| 41 | Datorua | Fatulidun | T |
| 42 | Datorua | Hali Laran | T |
| 43 | Fohoren | Fatuc Bitic Laran | T |
| 44 | Fohoren | Fatuc Laran_1 | T |
| 45 | Fohoren | Sadahur | T |
| 46 | Lactos | Cacaut | T |
| 47 | Lactos | Fatuc Laran_2 | T |
| 48 | Ogues | Baulela | T |
| 49 | Ogues | Foho Rua | T |
| 50 | Baboileten | Liabe | C |
| 51 | Baboileten | Raelau | C |
| 52 | Batumano | Batumano | C |
| 54 | Laclo | Sorate | C |
| 55 | Laclo | Tapomea | C |
| 56 | Laubono | Biabote | C |
| 57 | Leimealeten | Ahigara | C |
| 58 | Leimealeten | Orbeto | C |
| 59 | Malabe | Malabe | C |
| 60 | Paramin | Airae | C |
| 61 | Paramin | Ilat | C |
| 62 | Humboe | Borohei | C |
| 63 | Humboe | Hatali | C |

| N | *Suco* | *Aldeia* | Treatment |
|---|---|---|---|
| 64 | Ponilala | Cota Heu | C |
| 65 | Ponilala | Eroho | C |
| 66 | Coliateleotelo | Hau Hei | C |
| 67 | Coliateleotelo | Manulete | C |
| 68 | Fatubessi | Peregrinacao | C |
| 69 | Fatubessi | Sao Francisco | C |
| 70 | Fatubolo | Fatubolo | C |
| 71 | Lisapat | Hatupae | C |
| 72 | Lisapat | Lauraehou | C |
| 73 | Mauubo | Caisoru | C |
| 74 | Mauubo | Grotu_2 | C |
| 75 | Catraicaraic | Manturai | C |
| 76 | Catraileten | Merique | C |
| 77 | Eraulo | Darudu | C |
| 78 | Eraulo | Madede | C |
| 79 | Hatugau | Cailiti | C |
| 80 | Hatugau | Hatugau | C |
| 81 | Hatugau | Sauria | C |
| 82 | Lauana | Grotu_1 | C |
| 83 | Lauana | Roulo | C |
| 84 | Fatuquero | Caitarahei | T |
| 85 | Liho | Camalpun | T |
| 86 | Liho | Hi | T |
| 87 | Railacoleten | Manuponihei | T |
| 88 | Railacoleten | Tuileso | T |
| 89 | Lissadila | Cai Cassa | C |
| 90 | Lissadila | Darulema | C |
| 91 | Lacumesac | Tahagamu | T |
| 92 | Batara | Balulin | T |
| 93 | Batara | Fatuha | T |
| 94 | Babulo | Lapuro | T |
| 95 | Babulo | Searema | T |
| 96 | Babulo | Turon | T |
| 97 | Grotu | Coli Dassi | T |
| 98 | Grotu | Dato Rae | T |
| 99 | Grotu | Leo Dato | T |
| 100 | Holarua | Fatuco | T |
| 101 | Holarua | Uru Fu | T |
| 102 | Letefoho | Ailuli | T |
| 103 | Letefoho | Cotalala | T |
| 104 | Letefoho | Ladiqui | T |
| 105 | Tutuluro | Dalun | T |
| 106 | Tutuluro | Sosemera | T |

# Appendix 2. Propensity score estimation and covariate balance assessment

To assess the feasibility of household-level matching, a logistic regression model was estimated using treatment assignment as the dependent variable and baseline household and village characteristics as predictors. These included demographic structure, education, land access, agricultural knowledge and practices, exposure to disasters, and selected village-level characteristics. Variables were selected based on theoretical relevance to treatment assignment and potential influence on key outcomes. An iterative process was used to identify a specification with adequate overlap and improved balance. Figure 4 illustrates the overlap in propensity scores between treatment and control groups. The considerable overlap indicates that matching is feasible, with most observations falling within the region of common support.

**Figure 4: Distribution of the propensity score**




*Source*: Authors' calculations based on baseline survey data.

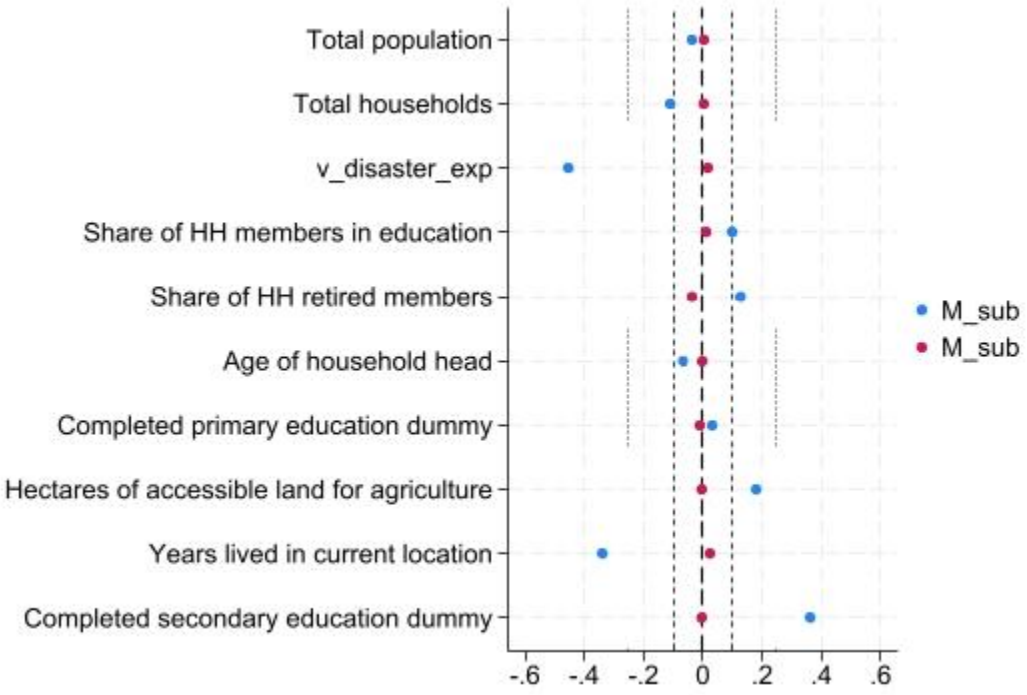
Figure 5 below presents the standardized mean differences for key covariates before and after applying propensity score weighting. Standardized mean differences are a common measure used to assess the balance of covariates between treatment and control groups in observational studies. The dashed lines in the figure represent two critical thresholds: 0.25, which indicates an acceptable level of imbalance, and 0.1, which suggests a more effective bias reduction after adjustment. In this context, the blue ("raw") values show the differences between treatment and control households as they appear in the original data. The red ("weighted") values reflect the same differences after applying weights derived from the estimated propensity score. These weights were obtained through the inverse probability of a household being in its observed group: households that look more similar across groups receive higher weights, while households that look very different receive lower weights. Applying these weights does not change the underlying data; it adjusts how much each observation contributes to the comparison. This reweighting makes the distribution of the matching variables more

comparable between treatment and control groups, which is why the weighted standardized differences are substantially smaller.

Before applying weighting, most covariates fall within the acceptable 0.25 threshold, suggesting a reasonable balance between treatment and control groups. However, a few key covariates show notable imbalances, meaning there were substantial differences between the two groups at baseline. Specifically, three covariates exhibit standardized mean differences greater than 0.25 before weighting. Households in the treatment and control groups had different levels of past disaster exposure, which may influence their vulnerability and adaptive strategies. There was also a considerable difference in the length of time households had resided in their current village, which could impact their familiarity with local resource management practices and decision-making structures. The proportion of household heads who had completed secondary education was noticeably different between treatment and control groups, potentially affecting household knowledge, decision-making, and livelihood choices.

These initial imbalances indicate that the treatment and control groups were not perfectly comparable in some key aspects before adjustment, which could introduce bias into the impact evaluation. However, after applying propensity score weighting, the balance improves significantly. The red dots in the figure represent the weighted covariates, all of which now fall well within the 0.1 threshold. This demonstrates that the weighting process was effective in reducing bias and ensuring comparability between treatment and control groups. By addressing these initial imbalances through matching and weighting techniques, the evaluation design strengthens the validity of causal inferences that will be drawn at the endline. Ensuring covariate balance is essential for isolating the true effects of the intervention and minimizing the influence of pre-existing differences between the groups.

**Figure 5: Standardized mean difference of matching variables**




*Source*: Authors' calculations based on baseline survey data.

# Appendix 3. Evaluation matrix

Table 14 maps each data analysis plan outcome indicator to its source variables in each survey wave and documents the construction method.

**Table 14: Evaluation matrix**

| EVALUATION QUESTION | INDICATORS | UNIT OF ANALYSIS | DATA SOURCE |
|---|---|---|---|
| 1. To what extent are participatory land use and forest management planning (PLUP) and village regulations effective in promoting sustainable, climate-resilient practices? | • Adoption of conservation agriculture practices<br>• Adoption of reforestation practices<br>• Reduction in slash-and-burn agriculture<br>• Changes in grazing practices | Household | Household survey |
| 2. What changes in awareness and adherence to village-level natural resource management (NRM) regulations are observed among households in treatment areas compared to control areas? | • Awareness of village NRM regulations<br>• Compliance with village NRM regulations | Household | Household survey |
| 3. What is the level of community satisfaction and perceived effectiveness of local governance structures among households in treatment areas compared to control areas? | • Satisfaction with village governance<br>• Trust in community members<br>• Perceived effectiveness of village regulations | Household | Household survey |
| 4. How do sustainable upland farming practices affect household income and food security in treatment areas relative to control areas? | • Total household income<br>• Income from agricultural activities<br>• Income from non-farm sources | Household | Household survey |
| 5. What is the level of adoption of climate-resilient cropping techniques among households in treatment areas, and how has this adoption influenced agricultural productivity and resilience to climate change? | • Adoption of drought-resistant crops<br>• Adoption of climate-resilient agricultural practices<br>• Agricultural production values | Household | Household survey |
| 6. To what extent has income generation and livelihood development support contributed to household income diversification in treatment areas compared to control areas? | • Number of income sources per household<br>• Participation in income-generating activities | Household | Household survey |

| EVALUATION QUESTION | INDICATORS | UNIT OF ANALYSIS | DATA SOURCE |
|---|---|---|---|
| 7. What changes have occurred in household food security and resilience as a result of project interventions? | • Food Insecurity Experience Scale (FIES)<br>• Coping Strategy Index (CSI)<br>• Frequency of food shortages | Household | Household survey |

## References

Asian Development Bank. "Key Indicators Database: Timor-Leste."

Asian Development Bank. *Asian Development Outlook September 2024: Growth in Asia and the Pacific*. Asian Development Bank, 2024.

Asian Development Bank. *Tax Effort and Tax Potential in Timor-Leste: Assessing Post-Pandemic Fiscal Challenges and Priorities*. Asian Development Bank, 2024. http://dx.doi.org/10.22617/SPR240293-2.

Cardinoza, Marlon. "Reviving Traditional NRM Regulations (Tara Bandu) as a Community-Based Approach of Protecting Carbon Stocks and Securing Livelihoods." In *Carbon Forestry: Who Will Benefit?* edited by Daniel Murdiyarso and Hety Herawati. Center for International Forestry Research, 2005.

Climate Watch. "Timor-Leste Country Profile: Climate Watch Data."

Delgado-Serrano, María del Mar, Elisa Oteros-Rozas, Isabel Ruiz-Mallén, Diana Calvo-Boyero, Cesar Enrique Ortiz-Guerrero, Roberto Ivan Escalante-Semerena, et al. "Influence of Community-Based Natural Resource Management Strategies in the Resilience of Social-Ecological Systems." *Regional Environmental Change* 18, no. 3 (2018): 581–592.

Elmadani, Mohammed, Evans Kasmai Kiptulon, Simon Klára, and Máté Orsolya. "Systematic Review of the Impact of Natural Resource Management on Public Health Outcomes: Focus on Water Quality." *Resources* 13, no. 9 (2024): 122. https://doi.org/10.3390/resources13090122.

Global Forest Watch. "Timor-Leste Dashboard."

International Monetary Fund. *Democratic Republic of Timor-Leste: 2021 Article IV Consultation—Press Release; Staff Report; and Statement by the Executive Director for Timor-Leste*. IMF Country Report No. 21/152. International Monetary Fund, 2021.

International Monetary Fund. *Democratic Republic of Timor-Leste: 2023 Article IV Consultation—Press Release; Staff Report; and Statement by the Executive Director for the Democratic Republic of Timor-Leste*. IMF Country Report No. 24/56. International Monetary Fund, 2024.

Melinda, Munalula, Brivery Siamabele, Valentine Kalonje, and Matildah Kaliba. "Community Based Natural Resource Management: Exploring the Benefits and Challenges." *International Journal of Humanities, Art and Social Studies* 5, no. 3 (2020).

Pinto Casquilho, José, and Xisto Martins. "On the Influence of Customary Law 'Tara Bandu' in Community-Based Forest Management at the Mountains of Timor-Leste." 2021.

Timor-Leste, Ministry of Justice. *Decreto-Lei N.º 26/2012: Lei de Bases do Ambiente*. July 4, 2012.

Timor-Leste. *Nationally Determined Contribution Timor-Leste 2022–2030*. 2022.

World Bank Group. *Timor-Leste Economic Report: Moving Beyond Uncertainty*. World Bank Group, 2019. https://doi.org/10.1596/31706.

Yale Center for Environmental Law & Policy. "2022 Environmental Performance Index: Timor-Leste."