BASELINE SURVEY DESIGN PROTOCOL

Inclusive Aquaculture for East Africa:

Enabling Small-Scale Participation

(Tilapia and Catfish Production in Kenya and Tanzania)

June 2026

1. STUDY OVERVIEW

1.1 Study Title and Purpose

This protocol outlines the design, methodology, and implementation procedures for a baseline survey to assess the current state of small-scale tilapia and catfish pond farming in selected counties of Kenya and districts of Tanzania. The baseline will establish a reference point against which changes in farming practices, production performance, and sector development can be measured at endline (3-4 years).

The survey is conducted as part of Gatsby Africa's programme 'Inclusive Aquaculture for East Africa: Enabling Small-Scale Participation'. The programme aims to test and refine commercially viable production models for small-scale fish farmers.

Central Research / Learning Question

What is the level of uptake of production models, or practices within them, for tilapia and catfish pond farmers in the programme focus regions of Kenya and Tanzania?

Specific Research / Learning Questions for each species:

1. Current practices (behaviours): What production practices do farmers currently employ across feed regimes, pond fertilisation, fingerling sourcing & stocking, pond preparation & water management, disease management & biosecurity, harvest practices, and sales/marketing?

Current performance: production information including: number of fish stocked, number of fish harvested, number of production cycles per year, and the average price of fish at point of sale (by weight or unit)?

2. Scale of production & decision-making: Number & size of ponds, experience and plans for pond farming, labour use and pond management (gender disaggregated)?
3. Knowledge and information sources: What are farmers' current sources of technical advice, training, information-seeking behaviours and record-keeping practices?
4. Importance of pond farming & fish: What is the relationship between fish farming and other household livelihood activities? What is the pattern of household fish consumption?

Rationale for 'Baseline the Practices, Not the Model' Approach

The programme will define commercially viable production models through testing during implementation. Since these models do not yet exist in final form, and we do not know whether they will emerge as cohesive packages or separable practices, the baseline cannot be designed around the ‘production model’.

Therefore, this baseline adopts a 'components' approach: capturing detailed data across all key practice dimensions at sufficient granularity that when models are eventually defined, we can retrospectively construct meaningful comparisons. This approach:

Works regardless of whether the eventual model is tightly specified or expressed as acceptable ranges
Reveals how far farmers had to travel to reach adoption
Provides a richer picture of partial and varied uptake beyond binary adoption indicators

Link to Theory of Change / Analytical Framework

The survey design is grounded in the Building Blocks Framework. The framework recognises that systems change when actors change behaviours. The baseline survey focuses on quantitatively measuring current behaviours (practices) and performance (production outcomes). Questions on rationale, blockers, and enablers for pond farmers will be addressed through the qualitative baseline survey aspects while these components for other value-chain actors will be assessed in other activities.

Component	What it means
1. Performance (change)	What's produced? The output of behaviours in terms of, for example, quantity, quality, productivity, price, and timing. Performance change is distinct from behaviour change: an actor can do something differently (behaviour) without doing it better (performance).
2. Behaviour (change)	Who's doing what? The specific thing an actor does differently; something you could, in principle, actually observe in the real world.
3. Rationale	Why would actors change? Whether the change is genuinely in the actor's interest, given real costs and returns. Without adequate rationale, actors won't change even when every other obstacle is removed. Rationale is an analytical judgment, not just what actors tell you.
4. Blockers	What blocks them from changing? What prevents change despite adequate rationale? Four types: capacity constraints (lacking skills or systems), resource constraints (lacking finance or inputs), information constraints (not knowing how to change), and regulatory constraints (rules that prevent it).
5. Enablers	What has helped actors overcome blockers. These can come from the programme, other development actors, government, the private sector, or the actors themselves.

1.2 Study Design and Methodological Approach

The baseline surveys employ a mixed-methods design, collecting and analysing quantitative and qualitative data concurrently to answer the research / learning questions.

Quantitative Component:

- Cross-sectional household survey of pond farmers
- Structured questionnaire
- Addresses: 'What is happening?' and 'How much?'
- Provides statistically representative estimates

Qualitative Component:

- Focus Group Discussions (FGDs) with pond farmers

Addresses: 'Why?' - exploring rationale, blockers, and enablers

Complements quantitative findings with depth

Note: Activities by Gatsby will cover other value chain actors

The study will use a repeated cross-sectional design (baseline and endline) with selection of farmers to survey conducted independently for each round. The core purpose of the survey is to measure whether improved practices diffuse through the broader farming population over time and therefore independent population-level samples is appropriate.

2. TARGET POPULATION

2.1 Geographic Scope

The baseline will be conducted in programme focus areas only:

Kenya: 6 – 8 Counties from across Western, Nyanza, Central and Eastern Regions (to be confirmed)

Tanzania: 3 Districts (Bagamoyo, Morogoro, Songea)

2.2 Population Definition and Inclusion Criteria

Farmer Type Categorisation

Four categories of pond farmers:

Category 1: Commercially-Oriented Producing Farmers (program target)

Uses lime, fertilizer and feed

Expects / Receives income from fish farming

Category 2: Non-Commercial Producing Farmers (Subsistence)

Rarely uses inputs – no lime, fertilizer or regular feed

Fish raised primarily for family consumption, any surplus given away

Category 3: Inactive Farmers (Ponds Not in Use)

Farmers with ponds who have stopped pond fish production

Category 4: Never-Farmers

Farmers without ponds who have never engaged in pond fish farming

Category 1 and 2 farmers are the target population for the study and will be included in the survey. However, no distinction will be made prior to or during data collection as to which of the two categories they fall under. **This categorisation will be done at the analytical stage.**

Species Coverage

The survey will primarily cover Tilapia, but the same questions will also be asked for Catfish if the pond farmer is harvesting both species. The survey will capture what species each farmer produces and at what scale. This approach:

- Acknowledges that the current population of catfish farmers in both Kenya and Tanzania is small
- Recognises catfish is often grown as secondary species alongside tilapia
- Captures whatever catfish activity exists as a subset of broader tilapia pond farmer population

Note that at endline the species coverage may include both Tilapia and Catfish as focus species, depending on the changes to numbers of catfish farmers.

2.3 Sampling Frame Development

A sampling frame of Tilapia pond farmers in the program focus areas will be used to select farmers to survey.

Tanzania

A list of 1,371 farmers across the three districts is available and considered fit for purpose (i.e. that it includes both commercially-orientated and non-commercial / subsistence Tilapia producing pond farmers and any missing farmers are 'missing at random').

Kenya

A list of ≈7,000 farmers (from a total population of ≈25,000) is available. This list is believed to be broadly representative of the tilapia pond farmer population, ranging from non-commercial / subsistence to commercially-orientated, although some likely hidden bias is acknowledged. Third-party implementers will be asked to obtain the lists and assess them for quality and coverage (e.g. discussions with administrative officers, asking farmer referrals during data collection etc.).

3. SAMPLE SIZE AND POWER ANALYSIS

3.1 Quantitative Sample Size

Estimation Approach

Sample size has been calculated to ensure sufficient precision for estimating key primary outcome indicators. The sample size was calculated for a single-timepoint (baseline, endline) estimate of the key indicators with an acceptable margin of error.

Primary Outcome Indicators

Adoption Indicators (Binary/Categorical):

- Proportion using improved feed regime
- Proportion practicing pond fertilisation
- Proportion sourcing fingerlings from commercial hatcheries

Performance Indicators (Continuous):

Total annual harvest per farm (kg)

Total revenue

Sample Size Parameters

Confidence level: 95% (α = 0.05)

Desired precision: ±5 percentage points

Expected adoption / performance:

Expected adoption indicator percentages range from 5 – 50% is expected for the adoption indicators and therefore 50% was used to produce the most conservative sample size estimate.

Expected performance indicators used data from previous studies to estimate sample size. These produced sample size lower than those of a 50% adoption percentage.

Finite population adjustment: applied to Tanzania only

Non-response / Data loss: 5% (N-adjusted = N / (1-0.05))

Using Cockran's formula[1] for a 1-sample binomial sample size with p = 0.50, margin of error = 0.05 and Z-value = 1.96 gives an unadjusted sample size of N = 385.

- **Kenya** – adjusted for 5% non-response / data loss to give adjusted **N = 405**.
- **Tanzania** – adjusted for finite population correction gives N = 301 then further adjustment for 5% non-response / data loss to give adjusted **N = 317**

The sample size for each country will be distributed across the focus counties / districts proportional to size (number of pond farmers in the sampling frame for each county / district).

Gender Considerations

Proportional representation of men and women will be interviewed for the survey. This is based on the understanding that focus is on estimating adoption / performance across the whole population of pond farmers rather than comparative gender analysis. Stratification would be warranted if models were expected to work differently for male vs female managers, but this is not the case.

3.2 Qualitative Sample Size

Details and justification for the sample size for focus group discussions (FGD) will be agreed with the 3rd-party implementors and will depend on the objectives of the FGD. An initial proposal is:

1 – 2 FGD per county / region

6-10 participants per FGD

Disaggregation: Gender-segregated; mix of farmer types

Potentially recruit from quantitative survey sample

[1] Cochran, W.G. (1977). Sampling Techniques (3rd ed.). New York: John Wiley & Sons.

4. SAMPLING PROCEDURE

4.1 Quantitative Sampling Strategy

Proposed sampling procedure for Tilapia pond farmers for survey:

- Group farmers by county / district
- Randomly select X+y farmers within each county / district

  X = proportional to farmer numbers;

  y = additional sample in case of not fulfilling inclusion criteria or non-consent

Confirm / Apply inclusion criteria: Pond must have fish currently growing OR harvested in past 12 months

Respondent Selection Within Farm

The survey targets the **primary pond manager**, defined as the person (or people) who makes day-to-day management decisions.

Enumerators will:

Ask farm to identify who is primarily responsible for pond management

- Probe actual task responsibility using decision-making questions (stocking, feed purchases, feeding, harvest timing, sales)

If 2+ people share at least 2 of 4 core decisions, classify as 'joint' management and interview all

4.2 Qualitative Sampling

To be finalised in discussions with third-party implementor: Purposive sampling with maximum variation. Draw FGD participants from quantitative survey sample. Ensure diversity across farm scale, practice profiles, performance levels, and species. Separate FGDs by gender. Continue until theoretical saturation achieved.

5. DATA COLLECTION TOOLS AND METHODS

5.1 Quantitative Survey Instrument

The survey should contain the following modules, details of specific variables to be proposed by the 3rd-party implementors, considering what can be reliably measured, and agreed with the Gatsby team.

Questionnaire Modules

- Respondent Identification: including primary pond manager identification
- Household Demographics: roster, primary manager(s) demographics, assets
- Livelihood Activities: All income sources, relative importance of fish farming
- Pond Infrastructure and Species: Number/size/type of ponds, species stocked, years farming, future plans
- Feed Regime: Types used, proportions, expenditure, feeding decisions
- Pond Fertilisation: Whether practiced, fertilizer types, frequency
- Fingerling Sourcing: Source, number stocked, price, size, stocking density

- Pond Preparation and Water Management: Draining, liming, water quality monitoring, exchange practices
- Disease Management: Mortality experiences, responses, preventative measures, service access
- Production Performance: Cycles completed, harvest quantity/timing/duration, fish size, survival rate, harvest decisions
- Sales and Marketing: Proportion sold/consumed, sales channels, product form, prices, contracts, feed cost %
- Knowledge and Training: Formal training, information sources, record-keeping, learning pathways
- Labour and Gender: Number involved, family vs hired, task division, income control
- Household Fish Consumption: Frequency, quantity, changes, species preferences

5.2 Qualitative Research Tools

Focus Group Discussion Guide

FGD guide to be developed in parallel with questionnaire by 3rd-party implementors. Proposed themes based on the Building Blocks Framework are:

- Introduction and Context: Ice-breaker, general perceptions of fish farming
- Rationale for Practice Choices: Why farmers do what they do - feed regime, stocking, harvest timing
- Blockers and Constraints: Capacity, resource, information, market constraints
- Enablers and Support: Help received, informal knowledge sharing, desired support
- Gender Dimensions: How became involved, challenges faced, control over decisions/income
- Future Outlook: Plans, what needs to change for viability