Quantifying the public benefits of agroforestry

An algorithm blueprint for LocalScale

Research synthesis, measurement design and implementation specification

Prepared 21 October 2026 | Version 1.0 | Based on the four supplied research PDFs

Recommended product: an auditable ledger of additional outcomes for identified beneficiaries, supported by physical measurements and uncertainty ranges. Add a transparent multicriteria dashboard for comparison and monetary values only where a defensible valuation pathway exists.

Agroforestry can rebuild soil function, habitat and carbon stocks while supporting livelihoods. The magnitude and even direction of a benefit depend on the prior land use, climate, tree species, age, management and comparison system. A practice label or tree count cannot establish the externality delivered.

The essential distinction

A benefit to a farmer is valuable, but it is not automatically a positive externality. Higher private yields and lower input bills belong in a farm viability account. Cleaner downstream water, habitat conservation, climate mitigation and community knowledge spillovers may benefit people beyond the farm. The algorithm must identify that recipient and the causal pathway.

What this document provides

Design layer	Where to find it
Research and benefit logic	Pages 2-4: evidence strength, beneficiaries and boundaries
Measurement and calculation	Pages 5-9: indicators, attribution, service equations and worked example
Comparison and implementation	Pages 10-13: scoring, data contract, monitoring and pilot plan
Evidence trail	Page 14: references and source-to-design map

Status: proposed design, not a validated calculator, carbon-credit methodology or payment standard. Equations, sample schedules, thresholds and example values are LocalScale design proposals unless explicitly linked to a source. No external prices or universal benefit rates are assumed.

What the supplied research supports

Source / type	Useful finding	Constraint on algorithm design
[R1] Lunelli et al., 2013 Narrative literature review	Organizes indicators into soil, water, biodiversity, economic and sociocultural dimensions; supports accessible, participatory assessment.	A taxonomy and indicator proposal, not pooled effect sizes. Its illustrative radar chart uses hypothetical data. Water quantification was sparse.
[R2] Phelpin et al., 2023 Conference abstract	Proposes ELECTRE Tri-nC and ELECTRE III to combine qualitative and quantitative criteria for agroforestry evaluation.	Describes intended work, not results, calibrated criteria or validated weights. Trees may also harbor pathogens.
[R3] Comolli et al., uploaded 2024 version Research Square preprint / field case	Describes a 10 ha trial established in 2010 within a larger yerba mate system; reports species-dependent yields, soil improvement and pest observations.	Single regional production context. Referenced supplementary tables and methods are not in the supplied PDF. Carbon estimates are described as work in progress.
[R4] Willmott et al., 2024 Systematic review	45 comparative studies covering seven categories. Many benefits, neutral results and primary-crop yield trade-offs; outcomes depend on comparator.	Content analysis of significance directions, not an effect-size meta-analysis. Coffee/cocoa and a limited set of locations dominate; multiple papers can represent one site.

Interpretation for LocalScale

Use the papers to select mechanisms, candidate indicators and likely trade-offs. Do not turn study counts or positive/neutral/negative classifications into benefit magnitudes. In [R4], a neutral comparison means no statistically significant difference was found; it does not prove equivalence or zero impact. Its plotted numerical scores were intended for visual comparison only. [R4, PDF pp. 3-5]

Organic agroforestry combines tree integration and organic management. Its comparison with a conventional monoculture cannot isolate the effect of trees. Keep these interventions separate in both the data model and the causal interpretation. [R4, PDF pp. 3, 9-12]

Research confidence and project confidence differ: a plausible mechanism from a review cannot replace a measured project outcome. Conversely, a well-measured local result can be credible without being transferable globally.

Translate benefits into beneficiary outcomes

Pathway	Potential external recipient / endpoint	Boundary and evidence need
Trees, litter and roots → less soil loss	Downstream residents, reservoirs or water users: less delivered sediment	Retained topsoil is primarily a farm benefit. Estimate the fraction of avoided erosion that would reach a receptor. [R1, pp. 2-3]
Biomass and soil carbon → lower atmospheric GHG burden	Global population: additional net tCO2e removal or avoided emissions	Measure changes in stocks and emissions against a baseline; account for losses and leakage. [R3, pp. 11-13; R4, pp. 7, 11]
Habitat structure and native species → conservation	Public and future generations: native habitat condition and species persistence	Separate planted diversity from wild native biodiversity. Species richness alone is insufficient. [R1, p. 3; R4, pp. 9-10]
Reduced runoff or chemical loss → cleaner water	Downstream users: reduced sediment, nutrient or pesticide exposure	Inputs and infiltration are proxies; delivered pollutant loads and exposure are outcomes. [R4, pp. 7, 12]
Pollinators, natural enemies → crop support	Neighboring farms: measured gains in pollination or avoided pest damage	Own-farm gains are private. Predator counts alone do not prove crop protection. [R4, p. 12]
Shade and diverse tasks → wellbeing	Workers and nearby residents: reduced heat exposure or improved conditions	Record working hours, exposure and labor terms. Do not infer health improvement from canopy cover. [R3, pp. 11-12]
Participation and knowledge → social resilience	Community members: usable knowledge, collaboration or shared food access	Record actual reach and adoption with consent. Income, training attendance and jobs are not automatically net social benefits. [R1, pp. 4-6; R4, pp. 10-11]

Record private and public effects together, but separately

Maintain three accounts: **farm viability** (profit, yields, labor, income risk); **external outcomes** (effects on other recipients or shared ecological assets); and **adverse effects** (including those on vulnerable groups). Some worker and household outcomes are partly contractual or private; classify the specific uncompensated spillover rather than assigning the whole category to externalities.

Set the accounting boundary before calculating

Unit, intervention and comparator

Use a georeferenced parcel × reporting interval × outcome × beneficiary as the basic record. Report both total project outcomes and per-hectare rates. Record trees, crop/livestock components, planting date, species, survival, density, pruning, fertilizer, pesticides and irrigation. Distinguish new conversion, enrichment and continued maintenance of an existing system.

Name the realistic alternative: continued conventional monoculture, organic monoculture, conventional agroforestry or another locally credible use. Use intact habitat as an ecological reference when useful, but do not credit conversion of native habitat as restoration. Existing agroforestry has ongoing service flows; only its incremental improvement or credible avoided loss is additional to the chosen baseline.

Spatial and temporal scope

Parcel boundaries suit inventory and soil measurements. Water effects require a catchment and downstream receptor; biodiversity needs landscape context; pollination needs receiving fields. Climate accounting may require off-site inputs and displaced production. Record these linked boundaries explicitly.

Maintain annual realized accounts and a separate forecast by year. A 20-year scenario can support planning if labeled as a LocalScale scenario choice, with establishment costs, maturation, harvest, mortality and disturbance represented. Never book future carbon uptake as a current outcome. Stock at a date and service flow during a year must have distinct units.

Keep treatment effects identifiable

Comparison	Interpretation
Organic agroforestry vs conventional monoculture	Effect of the whole combined package; trees and organic management are confounded.
Agroforestry vs monoculture, same input regime	Closer to the marginal effect of tree integration, subject to design and site differences.
Four matched groups: organic/conventional × trees/no trees	Can estimate separate effects and their interaction, if sufficient independent farms and comparable histories exist.

Before enrollment: document land-use history, tenure and participation rights, prior habitat condition, beneficiary consent where relevant, and other projects claiming the same outcome. Record any native-habitat conversion, invasive species, labor harm or displaced production as adverse effects. Thresholds and consequences must be agreed before scoring.

Minimum indicator dictionary: physical outcomes

These are proposed operational indicators. A proxy can trigger monitoring; it cannot silently become a quantified externality. Frequencies are pilot starting points and must reflect local variability and detection limits.

Outcome / unit	Measurement and baseline	Cadence / externality rule
Net climate benefit tCO2e/ha/year and total	Permanent tree plots (species, diameter, height, survival); locally applicable biomass model. Soil carbon concentration, bulk density, coarse fragments and depth; fuel, fertilizer and livestock emissions in both systems.	Annual tree and activity inventory; soil remeasurement initially every 3-5 years, revised by power analysis. Report interval-average soil stock change, not an invented annual observation.
Avoided delivered sediment t/year at receptor	Comparable erosion plots or calibrated erosion model, linked to sediment delivery and stream monitoring. Record slope, rainfall, soil and ground cover.	Event-based and seasonal sampling. Avoided erosion on the parcel becomes a downstream outcome only after delivery is assessed.
Avoided nutrient loads kg N or P/year at receptor	Flow and representative concentrations at relevant outlets; paired catchments or a calibrated transport model. Separate nitrogen and phosphorus.	Wet/dry seasons and storms; flow-weighted load. A lower concentration with higher discharge may not mean a lower load.
Water availability / flood moderation m3 by season; peak flow m3/s	Gauged flow, precipitation, abstraction and relevant weather controls. Identify downstream demand and timing.	Continuous logger where feasible. Higher infiltration does not prove higher annual water yield; tree water use may offset gains.
Native habitat condition ha × condition; native occupancy	Mapped habitat extent with a predefined 0-1 condition rubric; standardized native taxa surveys, effort and detection tracking. Separate planted/exotic species.	Annual habitat audit and repeated matched-season surveys. Interpret habitat units as local indicators, not interchangeable biodiversity credits.
Neighboring crop service yield or damage change per recipient field	Matched receiving fields, distance gradients or exclusion trials for pollination/pest control; crop, weather and input controls.	Key flowering/pest periods and harvest. Count a service only when linked to a receiving-field endpoint; avoid overlap with private harvest revenue.

Social outcomes, viability and low-cost collection

Indicator	How to quantify	Interpretation / limitation
Heat exposure and wellbeing	Paired measurements of a locally suitable heat-exposure metric during comparable tasks; person-hours exposed, work/rest patterns and anonymous worker feedback.	Shade and air temperature are screening variables. Clinical health claims need an exposure-response basis and appropriate expertise.
Community knowledge spillover	Consented recipient roster, independently confirmed adoption or retained skills, follow-up interviews and comparison group where feasible.	Attendance is an activity. Count unique beneficiaries and sustained outcomes; avoid rewarding repeat attendance as new reach.
Food access / security	Consistent, locally validated survey instrument, seasonal follow-up and a relevant comparison group; report household and community populations separately.	Own-household food output is mainly a private benefit. Community effects require evidence of improved access, affordability or shared provision.
Employment and labor quality	Paid and unpaid hours, earnings, safety incidents, age/sex-disaggregated participation where consented, opportunity cost and displacement.	Jobs and wages are not equal to net welfare gain. More work can mean overwork; account for forgone work and distribution.
Farm viability and production	Net margins including establishment, maintenance and family labor; crop-specific yields per hectare of whole system, total food output and income variability.	Report separately from externality value. Avoid adding unlike crop weights; use product vectors, calories where relevant, or land-equivalent ratios with comparable sole-crop references.

Three evidence tiers

Tier 1 - screen: geotagged photos, tree survival, ground-cover transects, simple erosion rubric, farmer observations and input logs. These support practice verification and prioritization. [R1, PDF pp. 2, 5-6]

Tier 2 - estimate: standard field sampling plus locally calibrated models, documented comparator and repeated observations. Report physical estimates with model uncertainty.

Tier 3 - verify: independent checks, stronger counterfactual design and recipient-level outcomes. Use for externally communicated high-confidence claims or potential payments after applicable requirements are established.

Do not translate Tier 1 ratings such as "less erosion" into tonnes using an arbitrary multiplier. Preserve ordinal observations and their anchored definitions. Use a subset of independently measured farms to assess how well inexpensive indicators track outcomes.

Estimate what changed because of the intervention

Preferred hierarchy for the counterfactual

Use randomized rollout where feasible; otherwise matched controls with repeated before/after measurements. Match on prior land use, crop, rainfall, slope, soil, farm scale and management. Keep controls outside likely spillover zones, or model their exposure. Cluster inference by independent farm or catchment: multiple plots within one farm do not create independent farms.

$$D_i = s_i \times [(Y_{project,after} - Y_{project,before}) - (Y_{control,after} - Y_{control,before})]$$

D is a direction-adjusted difference-in-differences estimate in the original outcome unit; s = +1 where more is beneficial, and -1 where less is beneficial (for example, sediment load). This is a period change, not automatically a per-year flow. For stock change, divide by the elapsed years only to obtain an interval-average rate.

This estimator relies on credible parallel trends and no unmodeled concurrent shocks or treatment spillovers. Inspect multiple pre-intervention observations where available; adjust for weather, crop age and management. If data are insufficient, label the result modeled or provisional rather than causally verified.

Move from ecological change to external service

$$Q_i = f_i(D_i, \text{area, transport, recipient exposure}) - L_i$$

f is the documented service-specific mapping into an external endpoint; L is off-site leakage or displacement in the same unit. If D is already a recipient-level total, f may be the identity. Do not multiply by area or a delivery fraction a second time. Store negative Q values: an algorithm for positive externalities must still reveal harm.

When stronger controls are unavailable

Fit a calibrated baseline model using comparable farms, management histories and environmental covariates; test it on held-out farms or years. Use a before/after comparison alone only as weak evidence because weather, market changes and soil recovery may explain the difference. Do not apply an unexplained "attribution percentage" on top of a causal estimate.

Additionality and contribution: test whether the improvement exceeds business as usual. Record LocalScale funding and other interventions separately from the agroforestry effect. An improvement can be caused by agroforestry without having been caused by LocalScale. Any allocation of a shared claim must have an explicit agreement and must sum to no more than the verified total.

Service modules and calculation rules

Climate: net stock change plus avoided emissions

$$Q_{climate} = (44/12) \times (\Delta C_{project} - \Delta C_{baseline}) + (E_{baseline} - E_{project}) - L$$

C is carbon stock in tonnes C across included biomass, dead organic matter and soil pools; ΔC is change over the reporting interval. E is tCO2e emissions over that same interval. Use one documented gas-conversion basis. Exclude CO2 fluxes already represented in stock changes. Record harvest and mortality transfers between pools; deadwood is not permanent storage. Imported organic matter requires checking source losses to avoid claiming a transfer as new atmospheric removal.

Soil stock calculation for each layer: **SOC (t C/ha) = carbon fraction × bulk density (t/m3) × depth (m) × 10,000 × (1 - coarse-fragment fraction)**, where bulk-density and fragment definitions are consistent. Prefer comparable soil mass where management changes density. Sum layers; distinguish organic matter percentage from carbon percentage. Soil concentration alone cannot establish sequestration.

Report realized net change, uncertainty and reversal exposure separately. Any future reserve or buffer withheld for a payment program is a separate claim adjustment with a disclosed rule, not a measured physical reduction. Following fire, harvest or tree loss, update stocks and the cumulative ledger; do not erase losses.

Water and sediment: measure loads and delivery

$$\text{Avoided delivered sediment} = \text{area} \times (\text{erosion}_{baseline} \times SDR_{baseline} - \text{erosion}_{project} \times SDR_{project})$$

Erosion is t/ha/year; SDR is the fraction delivered to the specified receptor. For measured water-quality loads, integrate concentration × discharge × duration with unit conversions. Do not replace this with average concentration × average flow when storm covariance matters. For altered flows, report seasonal gains and losses and downstream consequences, not a universal positive water score.

Habitat: retain ecological meaning

$$\Delta \text{ habitat units} = \Sigma\ [\text{area}_{project} \times \text{condition}_{project} - \text{area}_{baseline} \times \text{condition}_{baseline}]$$

Use mutually exclusive habitat polygons and the same locally defined condition rubric (0-1). Record native composition, structure, connectivity and reference condition. Report survey outcomes beside habitat units; do not add them as separate quantities of biodiversity. These units are not proof of species persistence or permission to offset native-habitat loss.

Worked example: one project-year

Entirely hypothetical. The following inputs demonstrate the proposed calculation logic; none is a coefficient, price or finding extracted from the supplied papers. Assume a 10 ha project with a credible comparator and a one-year interval.

Module	Illustrative calculation	Result
Climate	Incremental carbon stock gain = 12 t C for the whole project. Project emissions exceed baseline emissions by 3 tCO2e; leakage = 2 tCO2e.	12 × 44/12 - 3 - 2 = **39 tCO2e/year**
Sediment	Baseline erosion = 5 t/ha/year; project erosion = 2. Delivery fraction = 0.20 for both.	10 × (5 × .20 - 2 × .20) = **6 t/year** avoided at the downstream receptor
Habitat condition	The same 10 ha has baseline condition 0.30 and project condition 0.45 under one local rubric.	10 × (.45 - .30) = **1.5 incremental habitat units**, a stock-like condition measure
Private farm margin	Incremental annual farm net margin after labor and maintenance = 2,000 illustrative currency units (CU).	Report **2,000 CU/year** in farm viability; do not add it to externality value.

Optional monetary layer

Assume hypothetical marginal social values of 50 CU/tCO2e and 15 CU/t sediment delivered to this receptor. The latter represents eligible downstream avoided damage, with no overlap with carbon value. Then the **monetized subset is 39 × 50 + 6 × 15 = 2,040 CU/year**. Habitat and community outcomes remain unpriced; 2,040 is not the total value of the project.

A carbon market price is not automatically the social damage avoided by a tonne. A payment or credit sale is also not an extra ecological benefit. If using social values, document geography, beneficiary scope, currency, price year, valuation basis and uncertainty. Do not substitute replacement cost unless it represents a necessary, effective and least-cost substitute with relevant demand.

An illustrative sensitivity envelope

If net climate benefit ranges from 25 to 45 tCO2e and sediment benefit from 3 to 9 t/year, holding the assumed prices fixed gives **1,295 to 2,385 CU/year**. These are scenario bounds, not a statistically estimated confidence interval. Actual uncertainty must also cover the baseline, carbon model, transport, measurement and unit values.

Do not add 39 tCO2e, 6 tonnes of sediment and 1.5 habitat units into a physical total. Preserve the outcome vector even when a common currency or decision score is shown.

Aggregation without hiding trade-offs

Keep four outputs distinct

1. Outcome ledger: signed physical quantities and beneficiary populations. **2. Evidence panel:** provenance, uncertainty and missingness. **3. Decision dashboard:** normalized criteria and preferences. **4. Optional valuation:** a priced subset, shown separately from transfers and farm profit. A high decision score is not a measured quantity of externality.

A transparent MVP score, if a single comparison is needed

$$z_i = 100 \times \mathrm{clip}(D_i / T_i, -1, +1) \qquad S = \Sigma\, w_i\, z_i \qquad \Sigma\, w_i = 1$$

D is the beneficial-direction incremental outcome and T is a positive, predefined meaningful target in the same unit and boundary. Scores range from -100 to +100; zero means no estimated incremental improvement. Choose targets by local ecological relevance and stakeholder deliberation, document them before observing results, and retain uncapped physical quantities. Do not use the best farm in the current sample as a moving target.

Define domain weights first, then distribute each domain weight among its criteria. More indicators must not automatically mean more weight. Use comparable contexts, ages and boundaries. Keep confidence separate from performance: multiplying by evidence quality can make well-documented harm appear less harmful.

Missingness and non-compensation

Do not replace missing values with zero or silently renormalize weights. Report coverage $C = \Sigma$ weights with valid data. As an **illustrative pilot rule**, show no overall point score if $C < 0.80$ or a required safeguard is unassessed. Otherwise show the known weighted subtotal plus a full-score interval [subtotal - 100(1-C), subtotal + 100(1-C)]. Call this a missing-data bound, not a confidence interval.

Predefine safeguard failures that prevent a favorable overall classification, such as native-habitat conversion, severe worker harm or credible downstream water harm. Their thresholds require local agreement; a carbon gain cannot automatically compensate for those harms. Show unresolved assessments as pending, not passed.

MCDA extension

The ELECTRE methods proposed in [R2] offer a later route to sorting or ranking with explicit preference, indifference and veto thresholds. The abstract supplies no ready-to-use parameters. Implement the full method with appropriate expertise, test sensitivity to weights and thresholds, and allow incomparability where evidence or preferences do not support a ranking.

Uncertainty, valuation and overlap controls

Propagate uncertainty rather than disguising it

Track measurement error, sample variation, baseline uncertainty, model error and valuation uncertainty separately. For nonlinear modules, use repeated draws from defensible input distributions; retain shared weather, soil and site effects so correlated indicators do not appear falsely independent. Report median, a defined interval (for example 5th-95th percentiles), probability of positive effect and the main uncertainty drivers.

When distributions cannot be justified, provide transparent low/base/high scenarios. A confidence interval spanning zero means the sign is unresolved at that precision; a non-significant estimate is not automatically zero. Do not apply arbitrary discounts and call the result a confidence bound.

A service-endpoint overlap ledger

Potential overlap	Required treatment
Soil carbon, litter and "soil quality"	Include carbon pools once in climate accounting. Soil condition can remain a dashboard criterion, but do not monetize it again for the same carbon benefit.
Erosion, clean water and water treatment	Erosion reduction is a mechanism; sediment-related avoided treatment or damage is an endpoint. Use one valuation pathway per affected endpoint, or justify distinct damages and populations.
Pollination, pest control and harvest	Choose attributable receiving-field welfare or production benefit. Do not also add the same yield gain under another service category or farm margin.
Carbon stored and carbon sold	Removal is the ecological outcome; sale proceeds and subsidies are transfers. Keep a claims registry for any allocations across programs.
Habitat, recreation and cultural value	Some values are distinct, others include each other. Keep unpriced indicators unless beneficiary-specific valuation and separability are justified.

Time and distribution in optional valuation

$$\text{PV of monetized subset} = \Sigma_{t=0..T} \Sigma_{\text{eligible } i} [Q_{i,t} \times v_{i,t}] / (1 + r)^t$$

Use consistent real or nominal unit values and discount rates, with a declared base year and horizon. Test more than one rate; keep physical flows undiscounted. Show adverse external effects alongside gains. Call a value "net social benefit" only if relevant real implementation and opportunity costs are included, with transfers excluded and private benefits treated consistently. Disaggregate who gains and who bears costs; aggregate value can conceal local losses.

Implementation contract for the algorithm

Record	Minimum fields
Project and context	project_id; parcel_id; geometry; area_ha; land-use history; climate/soil strata; crop/tree species; tree age; management; start_date; participant/tenure status
Observation	observation_id; indicator_id; parcel or receptor ID; timestamp/interval; value; unit; sampling depth/effort; method_version; device/lab; source file; measured/modeled flag; QA flags
Baseline and causal model	baseline_id; comparator type; control IDs; covariates; pre-period; estimator_version; assumptions; validation statistics; spillover scope
Outcome and recipient	outcome_id; service_endpoint_id; beneficiary_group; spatial/time scope; signed quantity; unit; uncertainty; leakage; evidence tier; realized/forecast; data coverage
Decision and valuation configuration	target_version; weights; safeguard rules; missing-data rule; price basis/year/currency; discount scenario; overlap exclusions; approved claims and allocations
Audit and provenance	source reference/page; calibration data version; run_id; input hashes; code/config versions; reviewer; change log; access/consent restrictions

Calculation pipeline

01 Validate: reject invalid units, duplicate records, impossible dates and inconsistent boundaries. Preserve raw observations. **02 Estimate baseline:** fit or load the versioned comparator and evaluate its quality. **03 Compute effects:** retain signs and the stock/flow distinction. **04 Map services:** apply recipient exposure, delivery and leakage once. **05 Propagate uncertainty:** retain shared sources of variation. **06 Check overlaps and safeguards:** flag unpriced, unverified or blocked claims. **07 Produce outputs:** physical ledger first, then score and optional valuation. **08 Archive:** freeze all inputs, configuration and review records.

State logic and verification cases

Use explicit states: **measured**, **modeled**, **proxy-only**, **missing**, **not applicable** and **failed QA**. A missing datum is null, not zero. Not-applicable criteria need a predefined applicability rule and disclosed treatment. Outputs inherit the weakest relevant evidence and disclose mixed inputs.

Required tests: unchanged project/control gives zero; higher project emissions reduce climate benefit; zero sediment delivery gives zero downstream sediment benefit; a change in reporting area cannot change a per-ha rate; harvest transfers are not counted twice; missing critical data blocks favorable classification; duplicate claims are flagged; rerunning identical inputs reproduces identical results.

Pilot, validate and release in stages

Stage 1 | Agree the question and freeze the specification

Choose one landscape and production system, one explicit comparator and a narrow set of outcomes. Co-design beneficiary definitions, indicator rubrics and safeguards with farmers, workers, local ecological specialists and downstream users. Establish access, consent and data governance. Start with climate, delivered sediment where a relevant receptor exists, native habitat condition and farm viability; retain social outcomes even if initially unpriced.

Stage 2 | Establish baselines and a feasible sample

Map parcels and receiving areas; reconstruct land-use and management histories. Select independent treatment/control farms or catchments and stratify by tree age, prior land use, slope and soil. Run a pilot variance study, then size the sample using the minimum detectable change, expected clustering and monitoring budget. No universal farm count is justified by these four papers. Collect matched-season baselines before new implementation where possible.

Stage 3 | Run shadow accounting before financial claims

A proposed first 12-month cycle can test collection workflows, unit handling, event-based water sampling and annual tree inventories. It cannot establish long-term soil carbon, mature-tree benefits or resilience to rare shocks. Continue multi-year observations and scheduled soil remeasurement. Run low-cost proxies alongside reference measurements to quantify bias, burden and missingness.

Stage 4 | Validate prediction, causality and usability

Validation question	Proposed release evidence
Do modules estimate observed outcomes?	Held-out farms/years; bias and error by context; interval calibration; sensitivity to baseline, transport and allometry.
Can small farms participate fairly?	Completion rates and monitoring cost per farm; burden and error by farm size, region and gender where consented. Do not penalize scarce data as poor performance.
Are decisions robust?	Rank/classification stability under plausible targets and weights; explanations of reversals; independent safeguard review.
Can a result be audited?	Trace one result to raw data and source assumptions; reproduce the run; confirm overlaps, loss updates and claim allocations.

Release gate and ownership

Release screening dashboards first. Release quantified service estimates only for calibrated contexts and documented uncertainty. Any payment use requires a separate, fit-for-purpose verification and claims policy. Assign a scientific owner for modules, a field lead for sampling, an engineering owner for provenance and units, and a community reviewer for consent and distributional effects. Publish model limitations and a correction process with each version.

References and source-to-design map

Page references below use the page number in the supplied PDF, including cover pages. This report paraphrases the supplied material. Papers cited inside those sources were not independently reviewed. No supplementary files beyond the four uploaded PDFs were available.

[R1] Indicator dimensions and participatory assessment

Lunelli, N. P.; Ramos, S. F.; Oliveira Júnior, C. J. F. (2013). **Agroflorestas e externalidades / Agroforestry and externalities.** Revista Verde de Agroecologia e Desenvolvimento Sustentável, 8(5), 163-170. Supplied file: 2013_agroflorestas_externalidades.pdf.

PDF pp. 1-3: framing, soils, sparse water indicators and biodiversity context. PDF pp. 4-6: social/economic benefits, accessible indicators and the proposed multidimensional model. LocalScale application: category selection and low-cost collection; no empirical calibration from the hypothetical radar chart.

[R2] Multicriteria decision support

Phelpin, O.; Macary, F.; Denaix, L. (2023). **Multicriteria decision support for evaluating new agricultural systems in agroforestry.** Abstract, 95th EWG on MCDA, Jaén, Spain, 13-15 April 2023. Supplied file: 20230201_Phelpin-EWG95-MCDA-abstract-discussion.pdf.

PDF p. 1: proposed ELECTRE Tri-nC / ELECTRE III evaluation, mixed criteria and potential pathogen trade-off. LocalScale application: transparent stakeholder preferences and non-compensatory decision rules. The scoring formula in this report is a separate proposed MVP, not ELECTRE.

[R3] Integrated yerba mate agroforestry field case

Comolli, L. R., et al. (2024 uploaded version). **A Sustainable Integrated Agroforestry System.** Research Square preprint, posted 6 May 2024. DOI: 10.21203/rs.3.rs-4333865/v1. Supplied file: A_Sustainable_Integrated_Agroforestry_System.pdf.

PDF pp. 5-10: trial design, tree associations, harvest and management observations. PDF pp. 11-13: soil tests, work environment and pending carbon quantification. PDF pp. 15-16: project funding and author links to the producing business. PDF p. 20 locates Tables 2-4 in unavailable supplementary files. LocalScale application: species/age/management fields and testable pathways, not transferable effect sizes.

[R4] Comparative evidence and its limitations

Willmott, A.; Riar, A.; Saj, S.; Armengot, L.; Cicek, H.; Kiboi, M.; Singh, A.; Grass, I.; Cotter, M. (2024). **The ecological and socioeconomic sustainability of organic agroforestry: a systematic review.** Agroforestry Systems, 98, 2933-2949. DOI: 10.1007/s10457-024-01064-w. Supplied file: The_ecological_and_socioeconomic_sustainability_of.pdf.

PDF pp. 3-7: 45 studies, comparator definitions, directional synthesis and geographic/crop concentration. PDF pp. 9-12: native and functional diversity, labor, yield interpretation, carbon and water trade-offs. LocalScale application: comparator-specific inference, separate farm viability, uncertainty and explicit adverse effects.

Design provenance: causal estimators, carbon and sediment accounting equations, schema, scoring rule, hypothetical example and pilot release gates are proposed analytical design choices. They require local calibration and specialist review; the four sources do not constitute a validated implementation standard.