

A Harmonised Cross-Modal Accuracy Assessment of Forest Canopy Height and Biomass Estimation Methods

A Systematic Meta-Analysis (2020–2026)

Aaron Smiles

Nuvelen, UK

Extended Abstract · July 2026

Status: pre-results extended abstract. Written to stand in while Phase 1 screening and harmonisation are in progress. Every quantitative finding is marked [TBC] — no result is asserted until the harmonised database is complete and manually verified.

Problem

Accuracy figures for satellite- and airborne-derived forest canopy height and aboveground biomass (AGB) are reported across widely incompatible conventions. Studies differ in the error metric used (RMSE, relative RMSE, R^2 , bias, and combinations thereof), in the reference-data regime against which error is computed (field plots, airborne lidar, terrestrial laser scanning, or another satellite product), and in how — or whether — results are stratified by forest type.

The consequence is that a practitioner choosing a sensor stack cannot compare a reported 4.2 m RMSE in a boreal conifer stand against a reported R^2 of 0.71 in tropical dense forest, because neither the metric nor the reference basis nor the biome is held constant. Published accuracy is abundant; comparable published accuracy is not.

Existing reviews do not resolve this. They are organised by sensing modality and report each study's numbers as the original authors framed them, which reproduces the incompatibility rather than removing it. A table of scraped figures is not an answer to “which modality combination should I deploy in this biome.”

Approach

This meta-analysis introduces a harmonisation protocol and applies it to the 2020–2026 literature. The protocol has four components.

1. **Inclusion and exclusion criteria.** Publication window 2020–2026; presence of quantitative accuracy metrics; a reference-data regime that can be mapped onto a common basis. Studies reporting only qualitative agreement, or validating against a reference of materially lower quality than the retrieval, are excluded with a recorded reason.
2. **Metric normalisation.** Reported accuracy is converted to a common basis: RMSE in metres for canopy height, RMSE in tonnes per hectare for AGB, plus relative RMSE, R^2 , and bias. Where a study reports insufficient information to normalise, it is excluded rather than imputed.

3. **Stratification.** Results are grouped by forest type: tropical dense, temperate deciduous, boreal conifer, and plantation.
4. **Within-stratum comparison only.** Modalities are compared against each other only inside a single stratum. This is the control against the apples-versus-oranges risk that cross-biome aggregation otherwise introduces, and it is the reason the protocol yields a prescriptive rather than a descriptive result.

Modality classes assessed are optical multispectral, spaceborne and airborne lidar, SAR, and fused multi-modal pipelines.

Corpus

An initial sweep identified approximately 165 candidate studies organised by variable $\times$ modality $\times$ forest type, expanded to 182 unique identifiers for bulk ingestion. 154 items are currently indexed in the working reference collection; a documented triage of 28 failed imports and roughly 22 mis-imported records drives the Phase 1 fix-up.

Target screening volume is 200–250 studies, with an anticipated exclusion rate of 40–60% for missing metrics or incompatible reference data, yielding an expected 80–150 usable studies — conventional scope for a meta-analysis of this kind. Table extraction is semi-automated using a PDF parsing pipeline, with manual verification checkpoints at every stage; no extracted value enters the harmonised database without a human check against the source PDF.

Expected contributions

1. **A reusable harmonisation protocol.** The normalisation and stratification rules are published in full so that other groups can apply them to future retrievals and extend the harmonised database.
2. **Stratified accuracy tables.** Normalised error by modality within each forest type: [TBC — pending completion of Phase 1].
3. **A prescriptive empirical finding.** A mapping from forest type to the modality combination that minimises normalised error in that stratum: [TBC]. This is the output that distinguishes the paper from a wrapper table — it tells a practitioner what to deploy where, with a stated error expectation.

Limitations

The protocol cannot repair incomparability that originates in the reference data itself. Where two studies validate against field plots of materially different quality, or against airborne lidar of different point density, normalisation makes their metrics arithmetically comparable without making them epistemically equivalent; these cases are flagged in the database rather than silently pooled.

Coverage is also expected to be uneven across strata, with tropical dense forest and plantation likely under-represented relative to temperate and boreal sites — a limitation the analysis will quantify rather than paper over.

Intended venue: IEEE JSTARS, or *Remote Sensing of Environment* as alternative. Follow-on to the canopy growth state-of-the-art survey, which supplies the verified seed corpus.