Aboveground Biomass of Restoration Plots

Drone–satellite predictive modeling in north-central Mato Grosso, Brazilian Amazon

Method Note & Results · July 2026 Prepared by: Andrew Boyd on Behalf of Instituto Centro de Vida (ICV)

Summary

Twelve direct-seeding and Assisted Natural-Regeneration (ANR) restoration plots (3–4 years old) carry an estimated aboveground biomass of 4–51 Mg/ha (median ~24 Mg/ha), predicted from drone canopy structure and multispectral imagery using a model trained on spaceborne LiDAR (GEDI) over ecologically analogous secondary forest.

At the plots' low stature the plot-scale uncertainty is about ±30 Mg/ha. Relative to same-age and same-height natural secondary forest, the plots fall within the normal range — above the class median and near or above its (right-skewed) mean — consistent with the trajectory of unassisted natural regeneration for their stage.

The estimates are best used for screening, benchmarking and prioritizing field validation, not for high-confidence carbon accounting. All values are aboveground dry biomass, not carbon.

1 Project and objective

This study estimates aboveground biomass (AGB) for a cluster of forest-restoration plots (0.8–3.7 ha; direct seeding and assisted natural regeneration; 3–4 years old at the time of survey) in north-central Mato Grosso, at the southern edge of the Brazilian Amazon. No field biomass or allometry measurements were available, so the estimate is derived entirely from drone data (multispectral imagery and a photogrammetric canopy height model) combined with satellite reference data. The reference is restricted to **secondary forest** so that the model is calibrated on regrowth comparable to the plots rather than on mature forest.

The approach has two parts: a **regional model** that learns how canopy structure and greenness relate to AGB across secondary forest in the region, and **application to the plots**, in which the drone-measured structure of each plot is passed to that model to predict its biomass. This note documents the methodology and the resulting estimates.

All values in this report are aboveground dry biomass in Mg/ha, not carbon. Converting to carbon stock would require applying a biomass-to-carbon fraction (commonly ~0.47–0.50); this report does not convert to carbon or CO_2e.

2 Reference data

GEDI L4A (2019) — the AGB reference. NASA's GEDI spaceborne LiDAR provides a field-calibrated, footprint-level AGB reference at 25 m resolution that is better able to resolve low-biomass variation than wall-to-wall products, which saturate over young regrowth. Only high-quality footprints (beam sensitivity ≥ 0.95) over mapped secondary vegetation were used.

MapBiomas Collection 9 secondary-vegetation age. Each footprint carries the stand age contemporaneous with its acquisition year, so reference points are labelled by successional stage without temporal mismatch.

GLAD canopy height (2019). A 30 m Landsat-derived canopy height model provides the height predictor. It was preferred over alternatives that either saturate over these stands or are statistically circular with the GEDI target.

Sentinel-2 NDVI (2019). Dry-season NDVI provides the greenness predictor, matched to the drone multispectral product.

3 Method — building the training set

Three choices keep the reference data internally consistent and the model appropriate to the plots.

3.1 Temporal consistency

The satellite predictors are year-matched to each footprint, but the canopy-height layer is available only for 2019. Pairing 2019 height with GEDI footprints from later years would mismatch the dominant predictor against stands that had since grown. Training is therefore restricted to **2019 GEDI**, so height, greenness and AGB all describe the same year. Because the model learns a **time-invariant structure-to-AGB relationship**, it applies to drone structure acquired later — the stands grew, but the relationship between a given structure and its biomass did not.

3.2 Secondary-forest purity

A 25 m GEDI footprint can straddle the secondary-forest boundary, mixing in pasture or mature forest. To avoid this contamination, each footprint's **secondary-forest purity** — the fraction of secondary cover in its immediate neighborhood — is computed, and footprints below 0.8 are discarded. Roughly 40% of footprints are fully within secondary forest and about half are partial, so this filter removes a substantial source of edge contamination.

3.3 Ecologically analogous region

Restricting the reference to the immediate study area leaves too few clean footprints for a stable fit. The pool is therefore expanded across the state to **climatically analogous secondary forest** — footprints whose precipitation and elevation fall within the study area's own range (precipitation ~1950–2230 mm; elevation ~230–400 m). This yields **22,331 footprints** while keeping the reference ecologically comparable to the plots.

3.4 Parsimonious predictors

Because the plots are clustered in a single location, climate and terrain are effectively constant across them and cannot distinguish one plot from another; tested directly, precipitation, elevation and slope did not improve plot-scale prediction. The model therefore uses only four **drone-measurable** predictors — canopy height (mean and 90th-percentile), canopy roughness (height standard deviation), and NDVI

— in a random forest fit with spatial-block cross-validation so that reported accuracy is not inflated by spatial autocorrelation.

4 Accuracy — and why it depends on structure

A central finding is that error **scales with the size of the stand**. A single footprint-scale RMSE (~49 Mg/ha) is dominated by tall stands that the plots do not resemble. Broken out by canopy height — the quantity the model actually uses — the error in the young-regrowth regime is far smaller:

Canopy-height band	mean AGB (Mg/ha)	RMSE (Mg/ha)	Successional stage
0–6 m (restoration plots)	18	**~30**	Young regrowth
6–12 m	34	~41	Early succession
12–20 m	58	~53	Mid-succession
20–100 m	83	~57	Mature secondary

The restoration plots are short (mean canopy height 0.3–6.7 m), placing them in the top row, where the plot-scale RMSE is **about 30 Mg/ha**. This is the relevant uncertainty for these plots because the model predicts from their *measured structure*, and that structure is low. Figure 1 shows the full relationship: absolute error rises steadily with canopy height (~30 to ~58 Mg/ha) while relative error falls, and the plots occupy the two lowest-error classes.

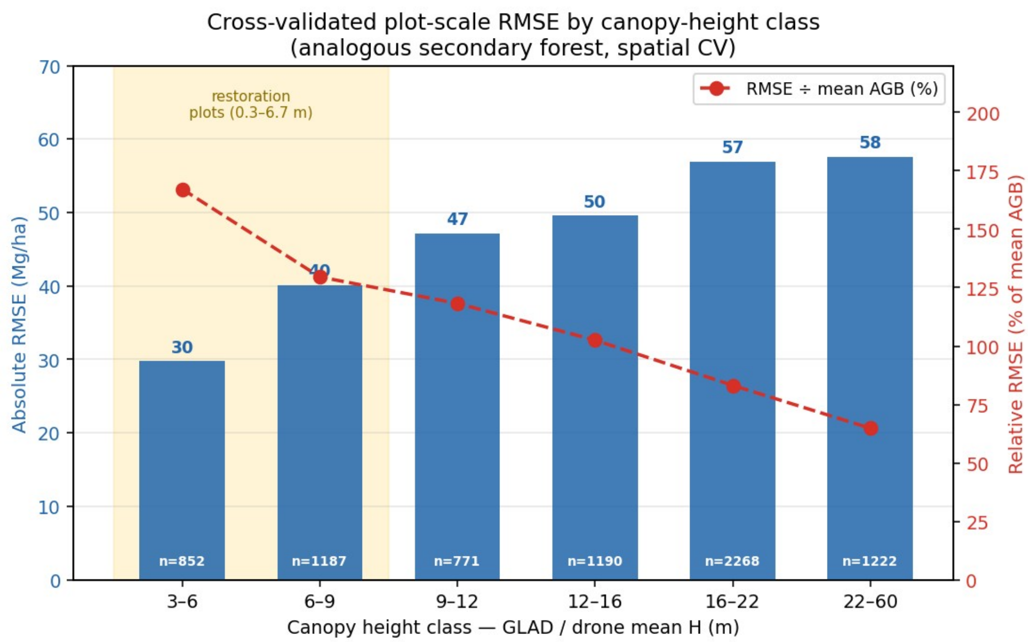

Figure 1. Cross-validated plot-scale RMSE by canopy-height class (blue, left axis) and relative RMSE as a percentage of mean AGB (red, right axis). Error scales cleanly with height — the quantity the model uses; the restoration plots (shaded, 0.3–6.7 m) fall in the lowest-error classes.

4.1 A counterpoint — RMSE by stand age

Stratified instead by stand *age* (Figure 2), the error is roughly flat at ~44–55 Mg/ha across all ages and does **not** fall for young stands. Age and biomass are only loosely coupled: a young age class contains a wide spread of structures, including vigorous tall regrowth well above these plots. In relative terms the young classes are the least accurate (RMSE exceeding 100% of mean AGB). As a *category*, therefore, young secondary forest carries about 50 Mg/ha of scatter; the plots occupy the low-error part of that range only because their *measured* structure is low. Both stratifications are reported so that the accuracy claim does not rest on the more favorable one alone.

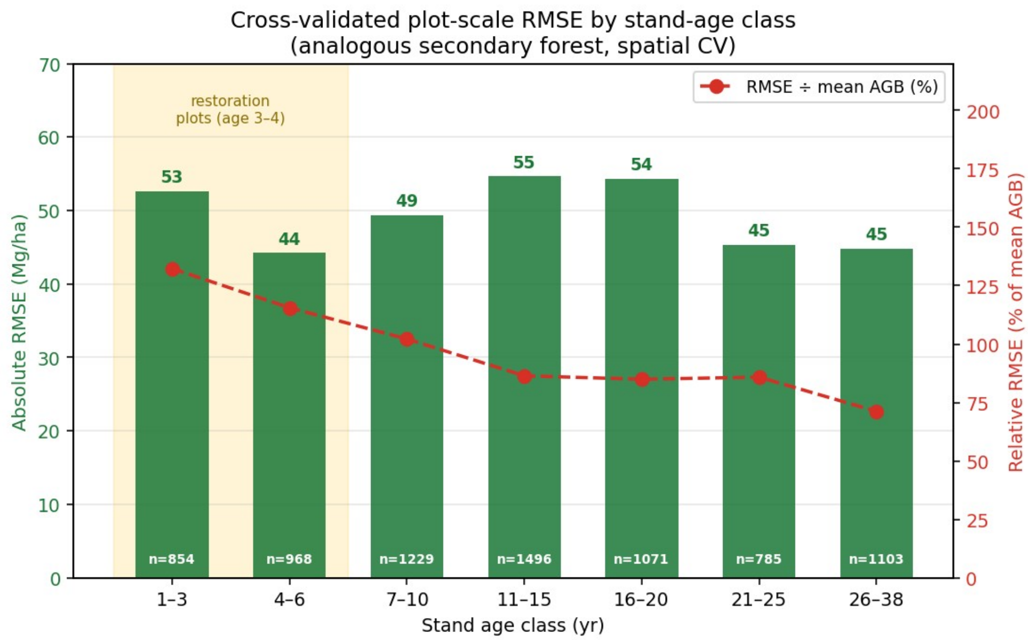

Figure 2. Cross-validated plot-scale RMSE by stand-age class (green, left axis) and relative RMSE as a percentage of mean AGB (red, right axis). Error is roughly flat in absolute terms and worst in relative terms at young ages; the restoration plots fall in the shaded age-3–4 range but at lower structure than the class average.

The residual scatter is genuine structure-to-AGB uncertainty, not noise in the reference itself (the GEDI per-footprint uncertainty is only ~13 Mg/ha); it would be reduced mainly by locally measured allometry.

5 Results — per-plot AGB

Applying the model to each plot's drone structure gives the estimates below (12 plots). Values span **4–51 Mg/ha, with a median of about 24 Mg/ha** — values consistent with 3–4-year regrowth.

Plot	mean H (m)	AGB (Mg/ha)	± RMSE	Municipality
Chácara São Jorge	1.6	**4.0**	±30	Nova Monte Verde
Sítio Piracicaba	1.7	**4.9**	±30	Alta Floresta
Sitio Quatro Irmãos	0.3	**9.5**	±30	Alta Floresta
Anacã	3.3	**18.4**	±30	Alta Floresta
Sítio Nossa Senhora Aparecida	2.1	**21.4**	±30	Paranaíta
Fazenda Sonho Meu	2.9	**21.6**	±30	Nova Monte Verde
Sítio Mutum	2.8	**26.7**	±30	Nova Monte Verde
Sitio Santos Reis	3.8	**27.7**	±30	Alta Floresta
Sítio Santa Clara	3.6	**27.9**	±30	Alta Floresta
Sítio Estância Fantasia II	3.0	**29.0**	±30	Nova Monte Verde
Sítio Rosângela	3.2	**36.0**	±30	Nova Bandeirantes
Fazenda Santo Antônio	6.7	**50.6**	±43	Alta Floresta

Because the plot-scale uncertainty (~±30 Mg/ha) is comparable to the estimates themselves at the low end, the results are most reliable for the **aggregate distribution** — the median, the range, and comparison to reference populations — rather than for precise ranking among individual low-biomass plots. In particular, the shortest plots (mean height below 2 m) sit at the low-biomass detection floor; their individual values (~4–9 Mg/ha) are effectively near zero within measurement noise, and small differences between them should not be over-interpreted.

6 Comparison to natural secondary forest

Two views place the plots against secondary forest drawn from the same analogous reference. The first compares each plot to secondary forest of the same **age** (Figure 3); the second, to secondary forest of the same **canopy height** (Figure 4).

6.1 By stand age

In Figure 3, the left panel compares each plot's estimate (green, ±RMSE) with the analogous secondary forest at that plot's age (blue band = class mean ±1 SD), with the class median (~15 Mg/ha) and mean (~37 Mg/ha) drawn as vertical reference lines; the right panel shows the plots against the age-class median and interquartile range.

How this is read depends on a key property of the data. Young secondary-forest AGB is strongly right-skewed: at 3–4 years the class mean is ~37 Mg/ha but the median is only ~15, because a minority of vigorous stands pulls the mean well above the typical stand (standard deviation ≈ 50). The two summaries therefore point in opposite directions — the plots lie above the median (9 of 12) but below the mean (1 of 12). The spread reconciles them: every plot falls within one standard deviation of the age-class mean, so the plots are statistically indistinguishable from the normal range of same-age natural regeneration.

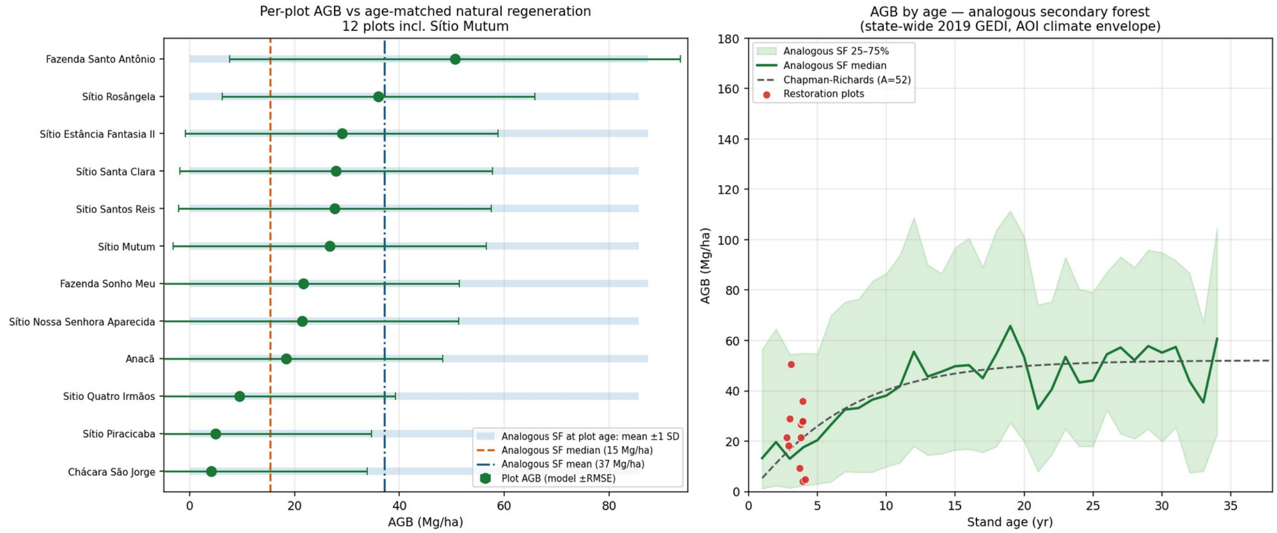

Figure 3. Left: per-plot AGB (green = model, ±RMSE) against the analogous secondary-forest mean ±1 SD at each plot's age (blue band); orange dashed and blue dash-dot lines mark the secondary-forest median (~15) and mean (~37). Right: AGB versus stand age for analogous secondary forest (median, 25–75% band, growth-curve fit) with the restoration plots. The mean far exceeds the median at young ages because the distribution is right-skewed.

6.2 By canopy height

Because canopy height is the model's dominant predictor, a second comparison places each plot against secondary forest of similar height (Figure 4). All twelve plots exceed the secondary-forest median for their height class, and eight of twelve exceed the height-class mean. This comparison is partly endogenous — since height drives the prediction, each plot's AGB is drawn toward the height-to-AGB relationship — so it does not independently demonstrate superior performance. What it does indicate is that the drone-derived greenness and roughness signals do not place the plots *below* comparable secondary forest of the same height; a plot sitting above the height-class mean carries a denser or greener canopy than a typical stand of that height. The three shortest plots fall in height classes represented by relatively few reference footprints, so their benchmarks are correspondingly uncertain.

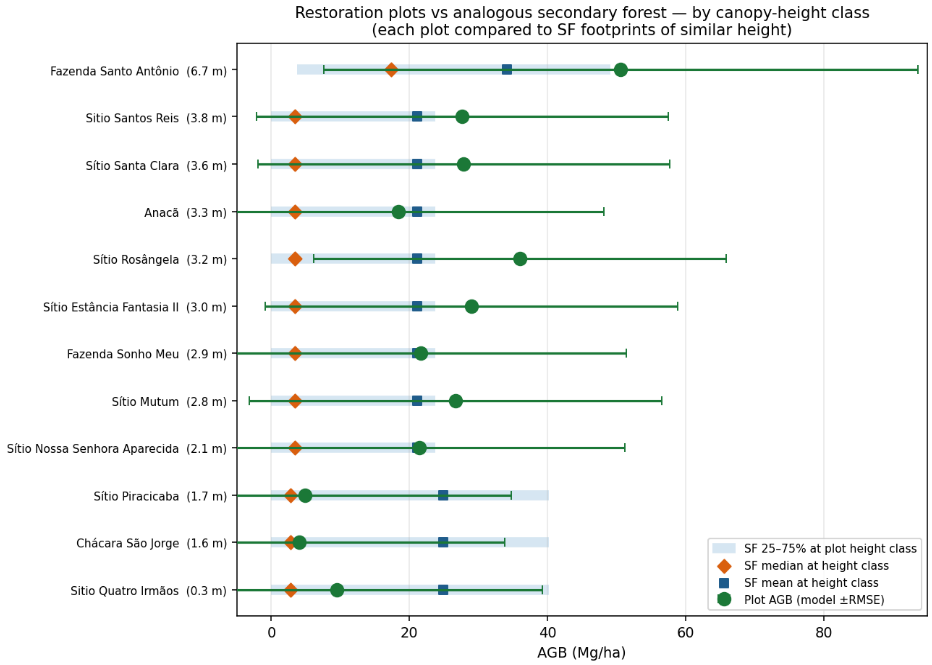

Figure 4. Each restoration plot's AGB (green, ±RMSE) compared to analogous secondary forest of similar canopy height: secondary-forest median (orange diamond), mean (blue square) and 25–75% range (band) for the plot's height class. Plots are ordered by canopy height.

7 Limitations

Height-source transfer. The model is trained on satellite (Landsat-derived) canopy height but applied to drone photogrammetric height; the two agree only weakly at the plot scale. They are also several years apart. No empirical calibration between them was attempted by design: any apparent offset between the two height sources would combine sensor bias, spatial scale, and real vegetation growth, so correcting it statistically would risk removing the biological signal of interest. The drone structure is therefore used directly, and this transfer is the main residual source of uncertainty.

Low-biomass relative error. In absolute terms the plot-scale error (~30 Mg/ha) is modest, but relative to a mean AGB near 18 Mg/ha it is large — a general feature of satellite biomass estimation at the low end, visible in the relative-error curves of Figures 1 and 2. This is why individual low-biomass plots should be read as order-of-magnitude estimates rather than precisely ranked values.

Biomass, not carbon. Results are aboveground dry biomass. A carbon-stock or CO_2e figure would require a biomass-to-carbon fraction (commonly ~0.47–0.50) and is outside the scope of this note.

No field validation. Absolute accuracy is anchored to GEDI, itself a model. A small field-allometry campaign is the single change that would most tighten the estimates, confirm their absolute level, and allow the drone's cover and roughness signals to enter the model more fully.

8 Conclusion

The 12 restoration plots hold an estimated median of about 24 Mg/ha of aboveground biomass (range ~4–51), with a plot-scale uncertainty of about ±30 Mg/ha at their current stature. The results support the interpretation that the plots are developing within the expected range of young secondary forest, with no evidence of systematic underperformance relative to comparable regrowth by either age or height. The estimate rests on a temporally consistent, purity-filtered, ecologically analogous GEDI reference and a parsimonious set of drone-measurable predictors.

These estimates are suitable for screening, benchmarking, and prioritizing field validation, but not for high-confidence carbon-credit quantification without local allometric measurements. Locally measured allometry is the recommended next step: it would validate the absolute level, allow conversion to carbon, and let the drone's full structural signal contribute to the model.