Gualán — co-registration of flight YS-20240123-211849

Date performed: 2026-07-28 **Performed with:** PDAL 2.10 (`filters.transformation`), numpy/rasterio co-registration

Problem

The Gualán site was flown three times on 2024-01-23. Two flights (204440, 220238) were successfully re-processed in YellowScan CloudStation (`SITES/Gualan/Updated_data/export/Gualan.las`, 265.9 M points, two flights merged and distinguished by `point_source_id` 1001–1006 and 2001–2013). The third flight (211849) could not be re-processed because its original raw data files were no longer available, so only the earlier export (`YS-20240123-211849-20240207-082517-F001.las`) exists.

That flight carries a large vertical blunder — previously identified in `DATA_QA_findings.md` as roughly +13 m and attributed to an uncorrected (autonomous) base-station height entered in POSPac. It was therefore re-levelled by co-registration to the two trusted flights rather than by re-processing.

Method

1. Class-2 (ground) returns of both datasets were binned to a 1 m mean-elevation grid over their common window (exact per-cell binning; no radius smearing). Both clouds already carry YellowScan ground classification (~80 % of returns).
2. The 3-D translation was solved by a three-stage brute-force search over (Δx, Δy) — ±12 m at 1 m, then 0.2 m, then 0.02 m — minimising the NMAD of the ground-surface difference, with Δz taken as the median residual at each candidate. This is free of slope/aspect sign conventions.
3. A Nuth & Kääb (2011) slope/aspect regression was run independently. It did not converge as tightly (NMAD 0.27 m vs 0.056 m) and was **not** adopted; the brute-force solution was used.
4. Only a translation was solved. This is the physically correct model for a base-station coordinate blunder, and the residual tilt check below confirms no rotation or scale term is warranted.

Solution applied

	Δx (m)	Δy (m)	Δz (m)
Translation applied to flight 211849	**−4.460**	**−0.920**	**−13.307**

Horizontal magnitude 4.55 m. The vertical component agrees with the −13.17 m independently estimated in `DATA_QA_findings.md`.

Error metrics (ground surface, 1 m cells, 457,814–465,065 cells)

	median	NMAD	RMSE	std
Before	−13.298 m	0.634 m	13.324 m	0.794 m
After	**+0.0001 m**	**0.056 m**	**0.094 m**	**0.094 m**

Residual NMAD of 5.6 cm is consistent with the site's own strip-to-strip agreement (8.2–10.5 cm 3-D RMSE from CloudStation), i.e. the alignment is as good as the internal consistency of the survey allows.

Validation

Spatial stability — the shift was re-solved independently in 3×3 spatial blocks. Three blocks held enough overlap:

block	cells	Δx	Δy	Δz	NMAD
0,1	18,135	−4.62	−0.92	−13.331	0.057
0,2	274,675	−4.42	−0.90	−13.313	0.063
1,2	132,160	−4.50	−0.92	−13.297	0.041

Spread (1σ): Δx 0.082 m, Δy 0.009 m, Δz 0.014 m. A single rigid translation describes the offset everywhere it can be measured.

Residual tilt — a plane fit to the post-alignment residuals gives dZ/dX = −1.26 cm/km and dZ/dY = −5.69 cm/km. Over the ~2 km overlap this is ~1 cm; de-trending improves NMAD only from 5.47 cm to 5.21 cm. No rotation/scale term is justified.

Slope dependence — median residual is ≤1 cm in every slope bin; NMAD grows from 4.2 cm (0–5°) to 12.0 cm (30–45°) as expected from point-density and ground-classification behaviour on steep terrain, with no systematic bias.

slope	cells	median	NMAD
0–5°	93,060	+0.13 cm	4.23 cm
5–10°	146,946	+0.02 cm	4.61 cm
10–20°	137,689	−0.24 cm	6.20 cm
20–30°	53,317	+0.38 cm	8.66 cm
30–45°	10,693	+1.01 cm	12.01 cm

End-to-end check on the delivered files — the written LAZ products were re-gridded from scratch and differenced: flight 211849 vs 220238 over 475,963 cells gives median −0.10 cm, NMAD 5.49 cm, RMSE 12.17 cm.

Important caveats

- **The alignment is anchored almost entirely to flight 220238.** The three flights form a SW→NE chain and 211849 is the middle one. Its overlap with 220238 is 475,963 ground cells (~0.48 km²) — this is what constrains the solution.
- **The tie to flight 204440 is negligible**: only 131 ground cells (a ~13 × 16 m patch at X 249056–249069, Y 1673227–1673243), at the outer swath edge of both flights. Over that patch the aligned 211849 sits **+39.6 cm** above 204440 (consistent across point-count thresholds, NMAD 5–6 cm).

This sample is far too small and too edge-located to justify adjusting anything, but it is recorded here because it is a real, repeatable signal.

- **Flights 204440 and 220238 do not overlap each other at all** (204440 max Y 1673243 < 220238 min Y 1673886). Their mutual consistency therefore cannot be verified from the data; both are trusted on the basis of having used the correctly-processed base station, not on the basis of a measured tie.
- The 2 undocumented extra bytes per point in the original 211849 LAS (no ExtraBytes descriptor VLR) were dropped on rewrite.

Reproducibility

Scripts used: `grid_ground.py` (ground binning), `coreg.py` (solver), `validate.py` (block/tilt/slope validation), `make_uploads.py` (transform + write), `make_dtm.py` (DTM), `verify_products.py` (end-to-end check).