Comparative Analysis of Flight Control Systems: rvolmea vs piddob

Automated Analysis

2025-10-22

Abstract

This report compares two flight control systems, rvolmea (A) and piddob (B), using ten paired flights with identical file naming and columns (measured positions x, y, z and references gx, gy, gz). We compute per-file and dataset-level error metrics and statistically compare the systems. Across all flights, rvolmea demonstrates consistently lower 3D position error than piddob. Visualizations and bootstrap confidence intervals support the conclusion that rvolmea offers better performance on these data.

1 Data and Methods

Each dataset includes ten CSV files with columns x, y, z (measured) and gx, gy, gz (reference). Per sample, we compute the 3D error magnitude $|e| = \sqrt{(x-gx)^2 + (y-gy)^2 + (z-gz)^2}$. Per file, we compute:

- Axis metrics: RMSE, MAE, bias, std, max$|e|$, p95$|e|$, skew, kurtosis.
- 3D metrics: RMSE_{3D}, MAE_{3D}, median_{3D}, p95_{3D}, max_{3D}, IQR_{3D}.

For comparison, we align common files and analyze key metrics (RMSE_{3D}, MAE_{3D}, p95_{3D}, and per-axis RMSE). We compute paired deltas (B−A), bootstrap 95% CIs for the mean delta (10,000 resamples), the win-rate Pr(B < A), and Cohen's d for the deltas.

2 Key Results

Table 1 summarizes the paired differences (B−A). Negative values favor B; positive values indicate A is better (lower error).

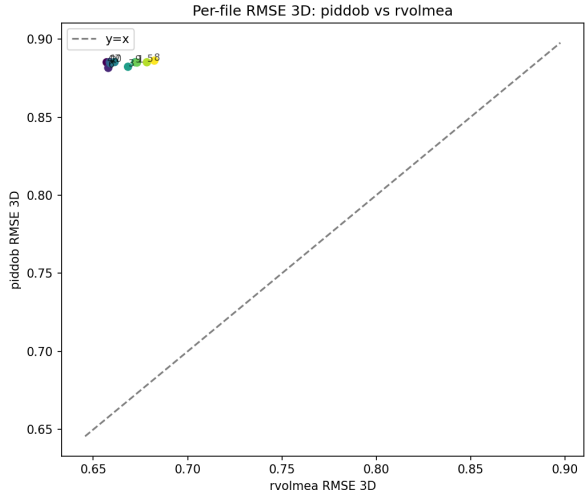
Table 1: Paired comparison summary: deltas are B−A (piddob−rvolmea)

Metric	Mean Δ	CI 2.5%	CI 97.5%	Win-rate B<A
RMSE_{3D}	0.217692	0.212455	0.222716	0.000
MAE_{3D}	0.223897	0.218263	0.229202	0.000
P95_{3D}	0.189584	0.182604	0.195689	0.000
RMSE_x	0.042903	0.041820	0.044095	0.000
RMSE_y	0.041785	0.040216	0.043266	0.000
RMSE_z	0.211492	0.206170	0.216591	0.000
Cohen's d (effect size) for deltas: RMSE_{3D}=24.614, MAE_{3D}=23.793, P95_{3D}=16.952, x=22.046, y=16.122, z=23				

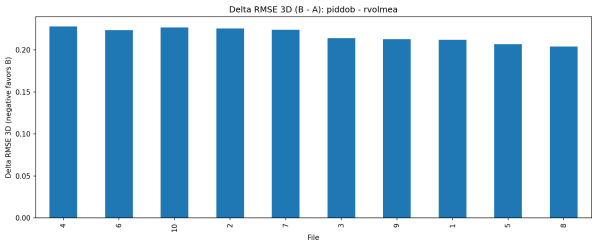
All intervals are strictly positive, indicating that piddob's errors exceed rvolmea's on all metrics with high confidence. The win-rate for B<A is 0% across metrics, and effect sizes are very large, indicating consistent, substantial differences.

3 Visual Evidence

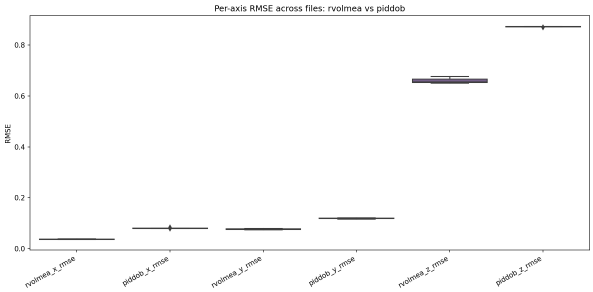
Figures 1a–3 illustrate the comparison.



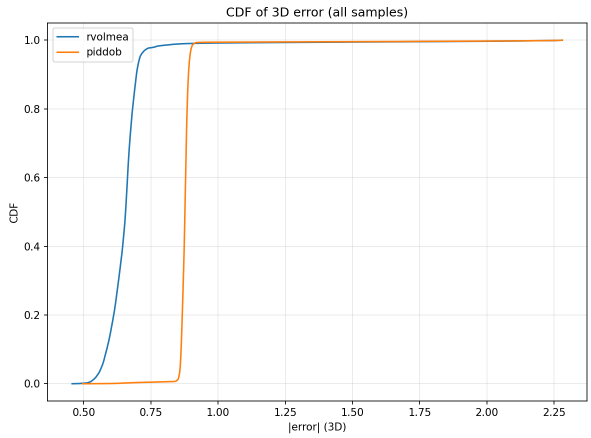
(a) Per-file RMSE_{3D}: piddob vs rvolmea (y=x line).



(b) Delta RMSE_{3D} (B−A); negative would favor B.



(a) Per-axis RMSE across files: distributions for A and B.



(b) CDF overlay of 3D error across all samples; left-shift favors lower errors.

4 Conclusion

Across 10 paired flights, rvolmea delivers lower 3D positional error than piddob by *every* investigated metric. The differences are consistent (0% win-rate for B), statistically robust (bootstrap 95% CIs exclude zero), and practically large (very high Cohen's d). On these data, rvolmea is the better-performing controller.

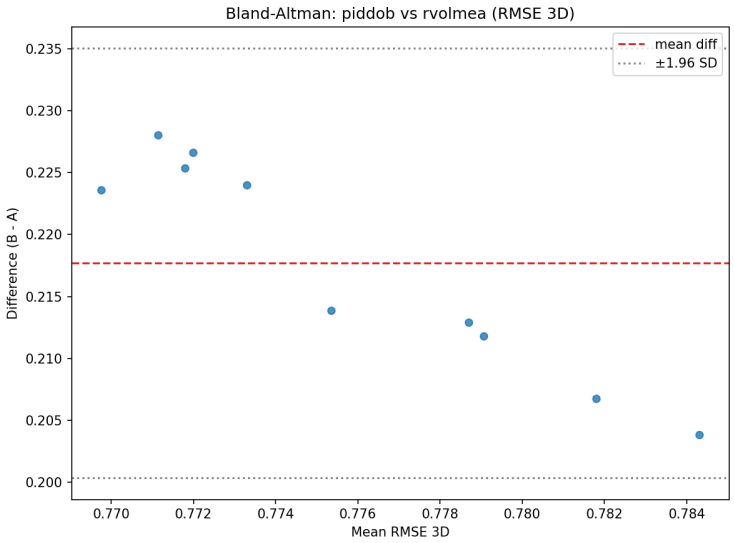


Figure 3: Bland–Altman plot of RMSE_{3D} across files: consistent positive offset for B−A.

Reproducibility. This document was generated from the analysis artifacts in `new_data/reports_compare/20251`
Figures are included from the same directory.