

UWB sensor accuracy and stability

Engineering characterization of archived observations. 20 vector chart plates and detailed metrics.

Held-out point-mean RMSE: **17.71 cm raw** to **16.15 cm corrected**. Absolute mean error improved in **5 of 8** point-anchor cases. A2 worsened at both held-out points. The correction remains an engineering candidate.

Anchor	Raw mean RMSE (cm)	Corrected (cm)	Cases improved
A1	16.25	12.36	2/2
A2	27.34	28.00	0/2
A3	11.68	9.84	1/2
A4	10.36	2.99	2/2

50,748 parsed ranges across 16 actual captures plus one empty preflight attempt. The 10 selected captures contain 33,854 ranges; 33,669 are eligible for range statistics. The six problem captures remain archived and are shown in the acquisition inventory.

Scope: nominal-reference range error, within-capture dispersion, tails, temporal structure, cycle coverage and an appendix with 3,492 sampled Local Compute X,Y snapshots. No new RF acquisition, model fit, firmware change or server correction.

Reference limitation: positions and equal antenna heights were operator-confirmed. No independent reference survey or reference-distance uncertainty is available. Results characterize these archived conditions; they do not establish traceable absolute accuracy.

How to interpret the measurements

Keep accuracy, dispersion, completeness and position-solver performance separate.

Signed error and bias. $e = y - u$. Here u is the operator nominal range. Bias = mean(e). Positive bias means over-reading; negative bias means under-reading.

Dispersion and RMSE. Sample SD uses $n-1$. MAE = mean(abs(e)); RMSE = sqrt(mean(e^2)). Exact identity: $\text{RMSE}^2 = \text{bias}^2 + ((n-1)/n) \cdot \text{SD}^2$. A stable offset may give small SD and large error.

Tails and percentages. P95/P99 are empirical percentiles of absolute sample error (linear interpolation). They are not confidence intervals or guaranteed tolerance bounds. Illustrative 5/10/20/50 cm bands are not acceptance specifications.

Eligibility and weighting. Use original canonical-valid unique, finite $0 < \text{range} \leq 100$ m observations. Keep 30 invalid and 155 before-status selected records outside metrics. No additional positive-outlier trimming. Summarize each point-anchor cell first.

Training and held-out data. The frozen affine $y = a \cdot u + b$ uses only eight P1-P8 means per anchor. Use $(y-b)/a$ for offline correction. No refit here. P9 and P1R are held out of fitting, but P9 raw statistics had already been viewed before the original fit; validation is not blind.

Dependence and time. Use tag acquired_ms within separate captures and boot pairs. Do not interpolate. ACF uses consecutive eligible cycles only, with 180-220 ms timing increments; CSV exports pair counts. No iid confidence bands, effective-N claim or $\text{SD}/\sqrt{n}$ uncertainty is reported.

Held-out error metrics for individual readings

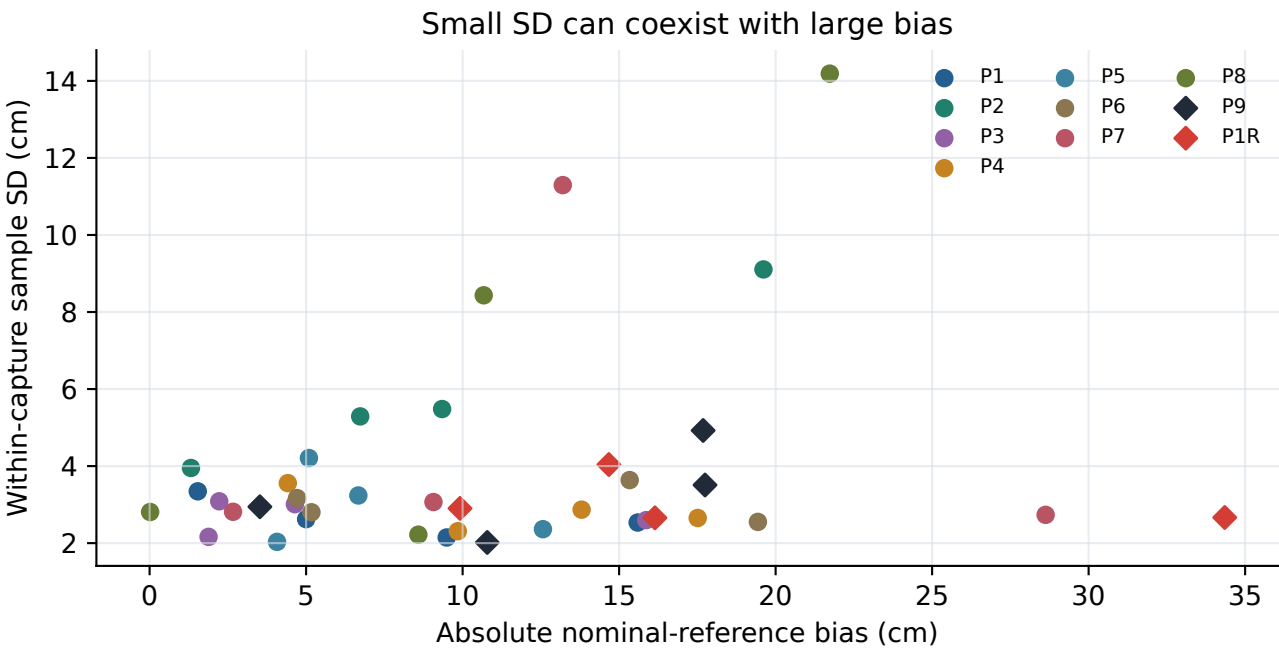
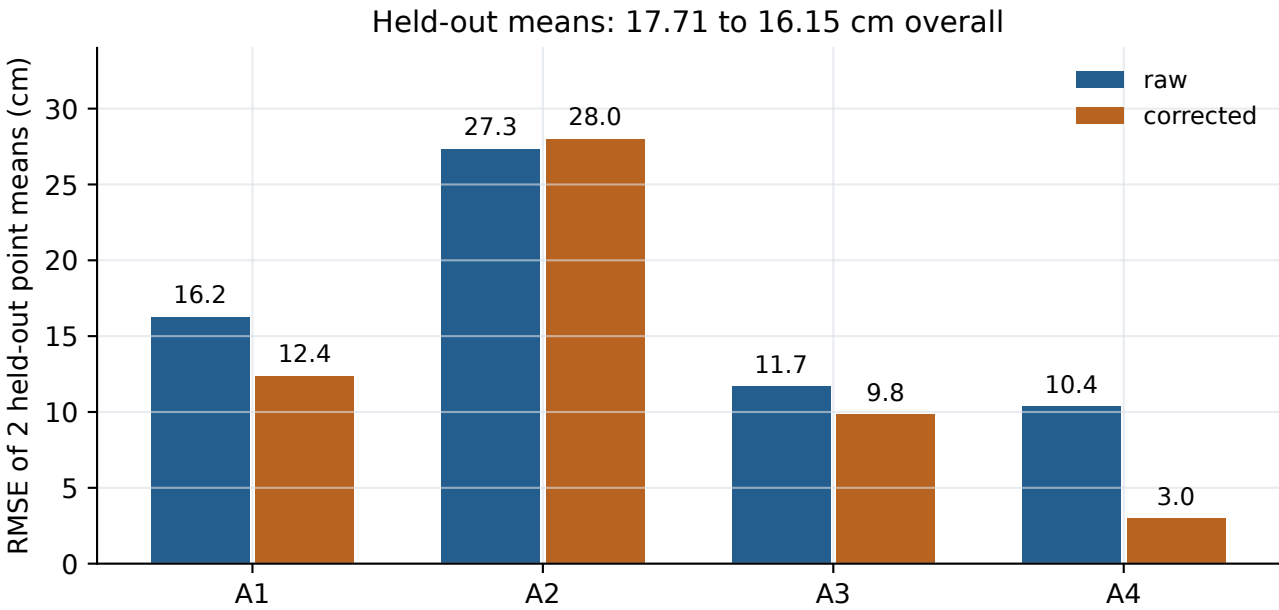
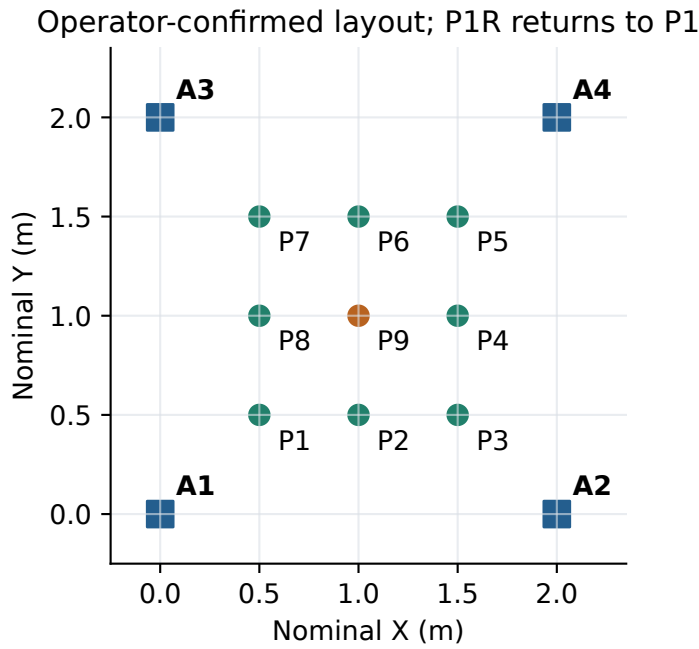
All values below are cm. P95 and max refer to absolute errors; SD refers to within-capture sample dispersion.

Point / anchor	Method	n	Bias	SD	MAE	RMSE	P95	Max
P9 / A1	raw	862	+17.68	4.93	17.69	18.36	25.14	35.46
P9 / A1	corrected	862	+13.20	4.84	13.26	14.06	20.52	30.67
P9 / A2	raw	834	+17.74	3.51	17.74	18.09	23.73	27.01
P9 / A2	corrected	834	+19.00	3.42	19.00	19.31	24.83	28.03
P9 / A3	raw	805	+3.52	2.94	3.86	4.59	8.25	14.35
P9 / A3	corrected	805	+8.73	2.85	8.73	9.18	13.30	19.20
P9 / A4	raw	864	-10.79	2.02	10.79	10.97	13.81	20.37
P9 / A4	corrected	864	-1.76	2.15	2.22	2.78	4.98	11.99
P1R / A1	raw	856	+14.67	4.04	14.67	15.22	19.84	23.59
P1R / A1	corrected	856	+11.46	3.97	11.52	12.13	16.54	20.23
P1R / A2	raw	800	+34.35	2.67	34.35	34.45	38.47	42.69
P1R / A2	corrected	800	+34.75	2.60	34.75	34.84	38.76	42.88
P1R / A3	raw	777	-16.14	2.66	16.14	16.36	20.65	24.87
P1R / A3	corrected	777	-10.85	2.57	10.85	11.15	15.20	19.29
P1R / A4	raw	845	-9.92	2.90	9.92	10.33	14.61	23.05
P1R / A4	corrected	845	+3.84	3.09	4.14	4.93	8.84	11.85

Every row uses only that capture and anchor. The headline RMSE on the overview page instead uses eight mean errors with equal weights. The two quantities answer different questions.

01 UWB range accuracy and repeatability

P1-P8 estimate the frozen model; P9 and P1R evaluate it. All errors use operator nominal ranges.



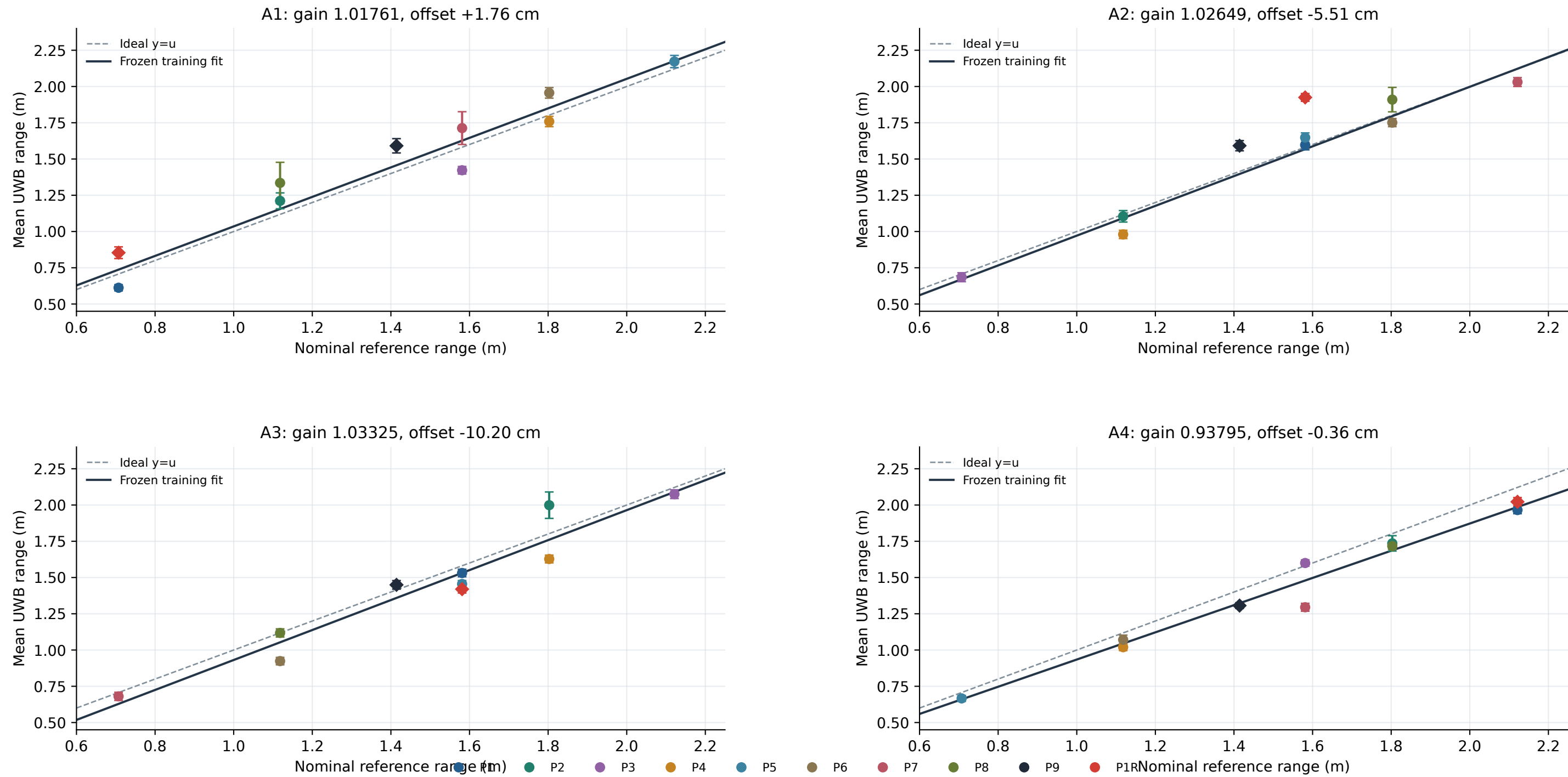
Analysis population

50,748 parsed ranges across 16 captures
33,854 raw ranges in 10 selected captures
33,669 canonical-valid unique ranges for statistics
30 basic-invalid; 155 before-status exclusions
6 problematic captures retained separately
3,492 unique sampled position snapshots
Model unchanged; no new fit or on-board correction

Interpretation. Bias measures mean error against the nominal reference. SD describes dispersion within a capture. Neither independently establishes traceable absolute accuracy. Primary validation gives equal weight to eight point-anchor means.

02 Range response across the measured geometry

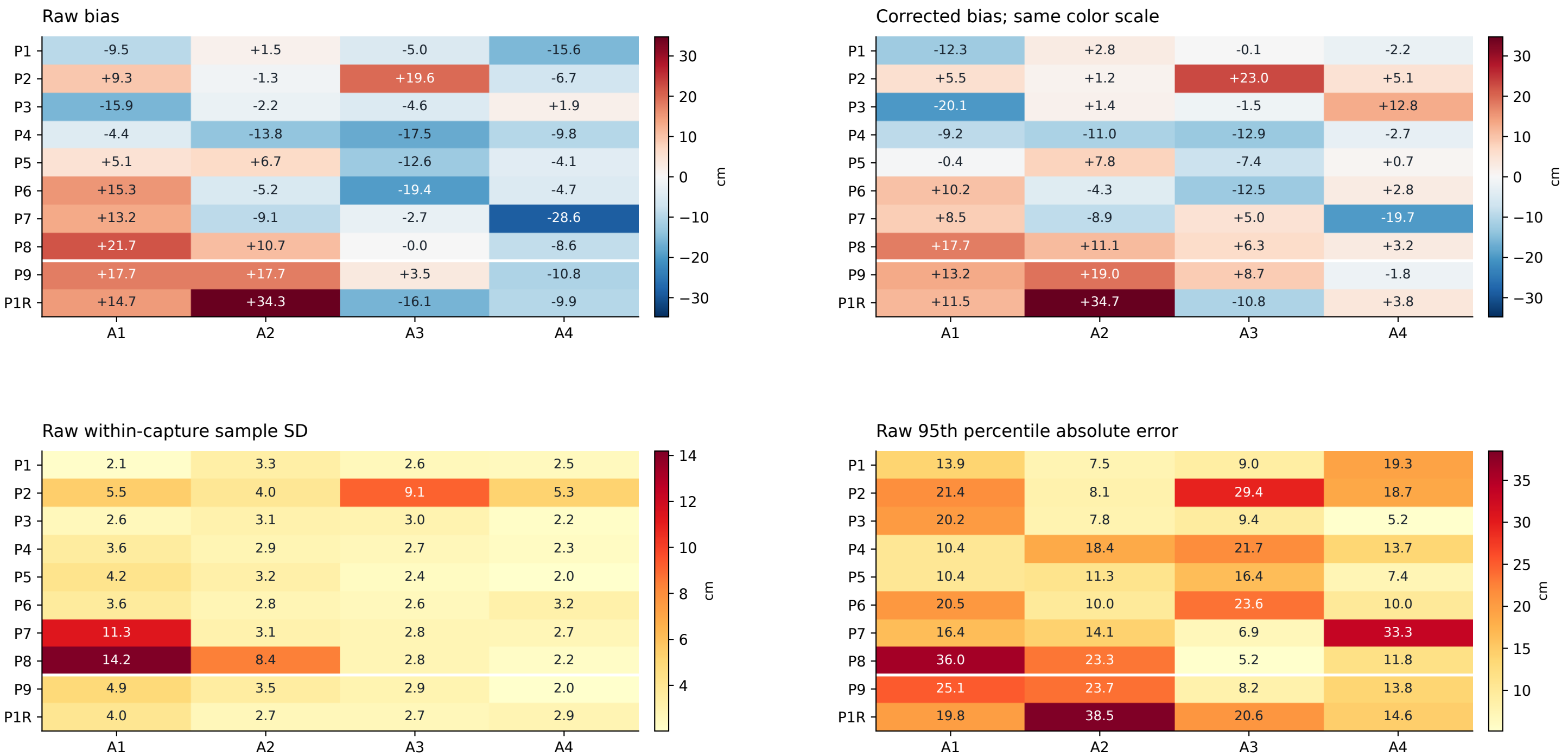
Error bars are within-capture sample SD. Frozen $y=a*u+b$; inverse correction is $(y-b)/a$.



Interpretation. Only the eight P1-P8 point means per anchor were used for the existing fit. Held-out diamonds are not fitted here. No extrapolation beyond the tested nominal ranges is supported.

03 Point-by-point error and dispersion

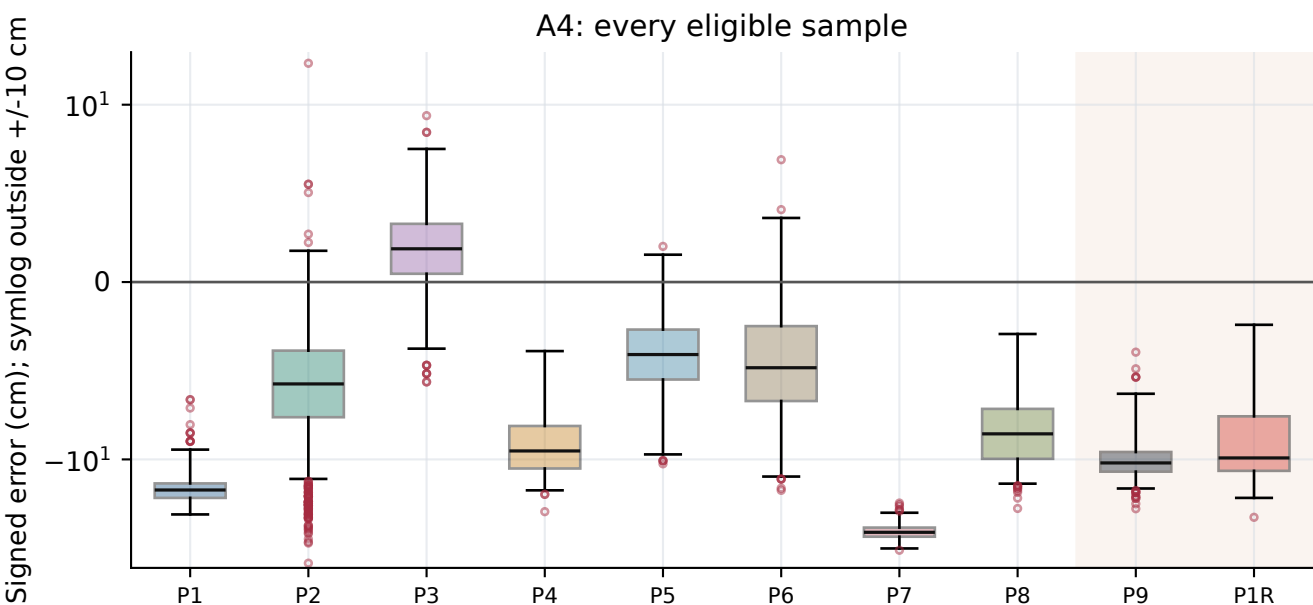
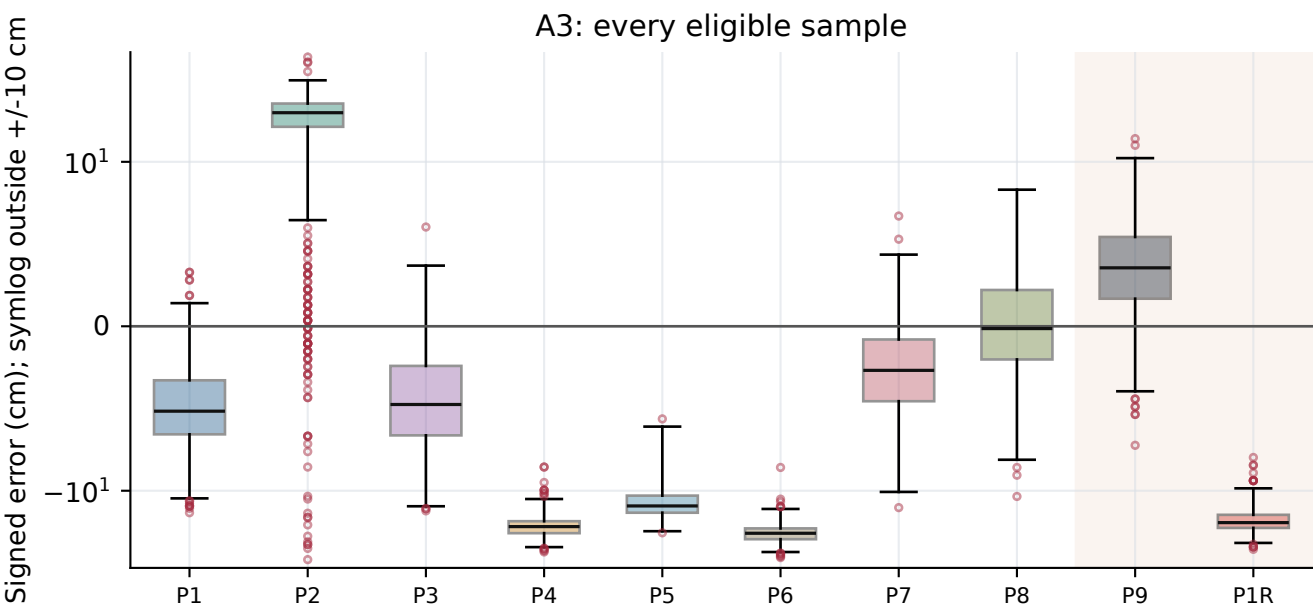
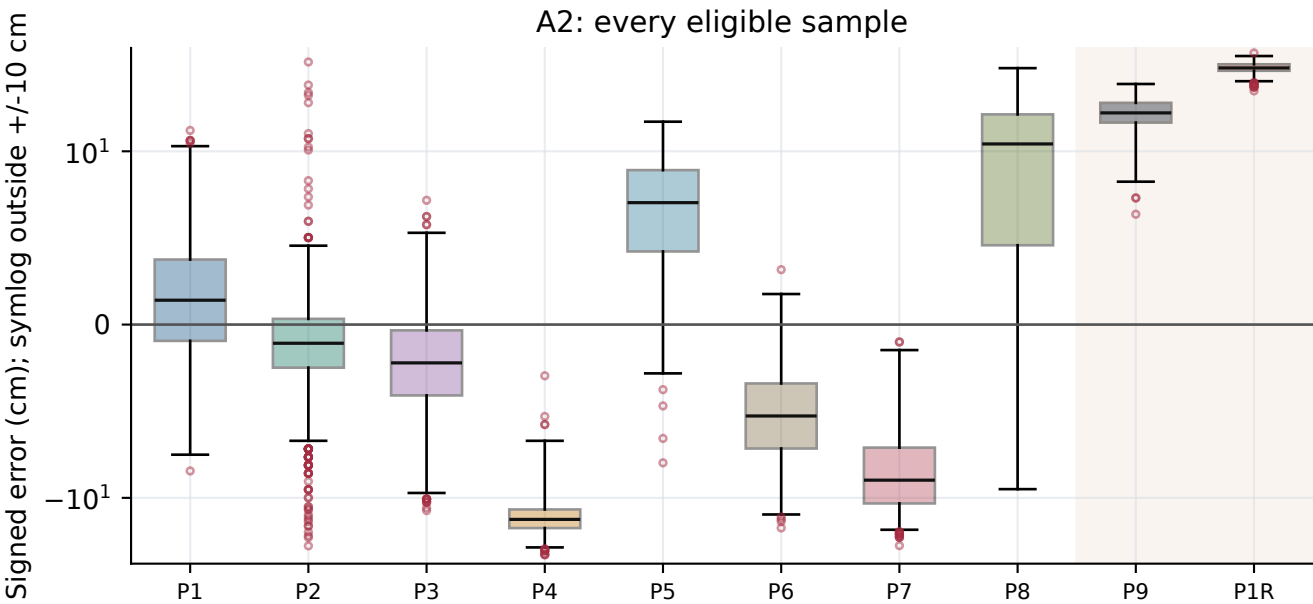
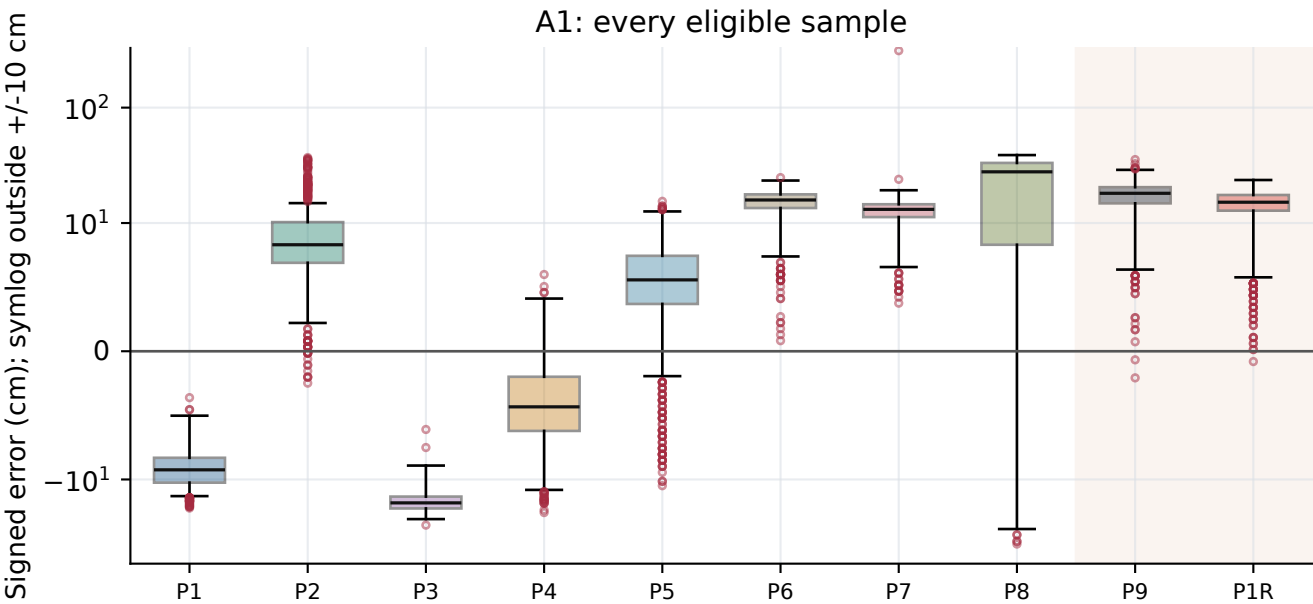
White separator isolates P9 and P1R. Every cell uses its own received canonical-valid observations.



Interpretation. P95 is an empirical quantile of absolute sample errors, not a 95% confidence interval, tolerance bound or guarantee. Positive outliers remain in the statistics.

04 Raw error distributions without tail trimming

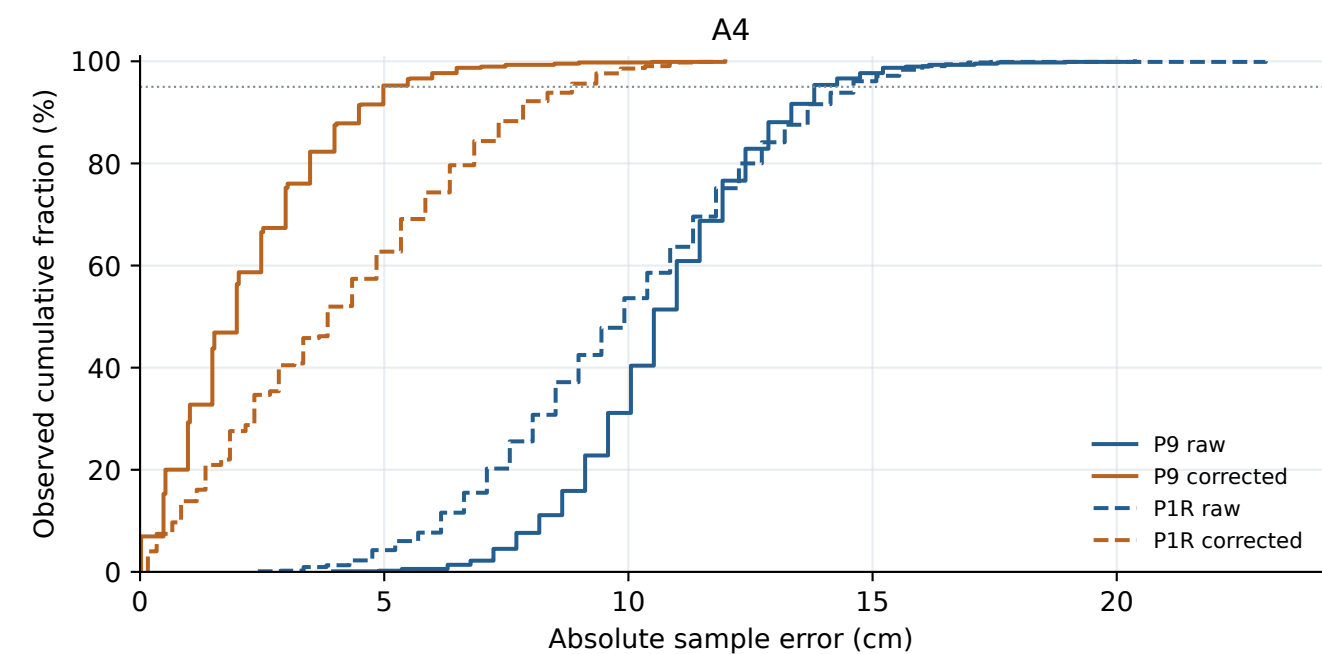
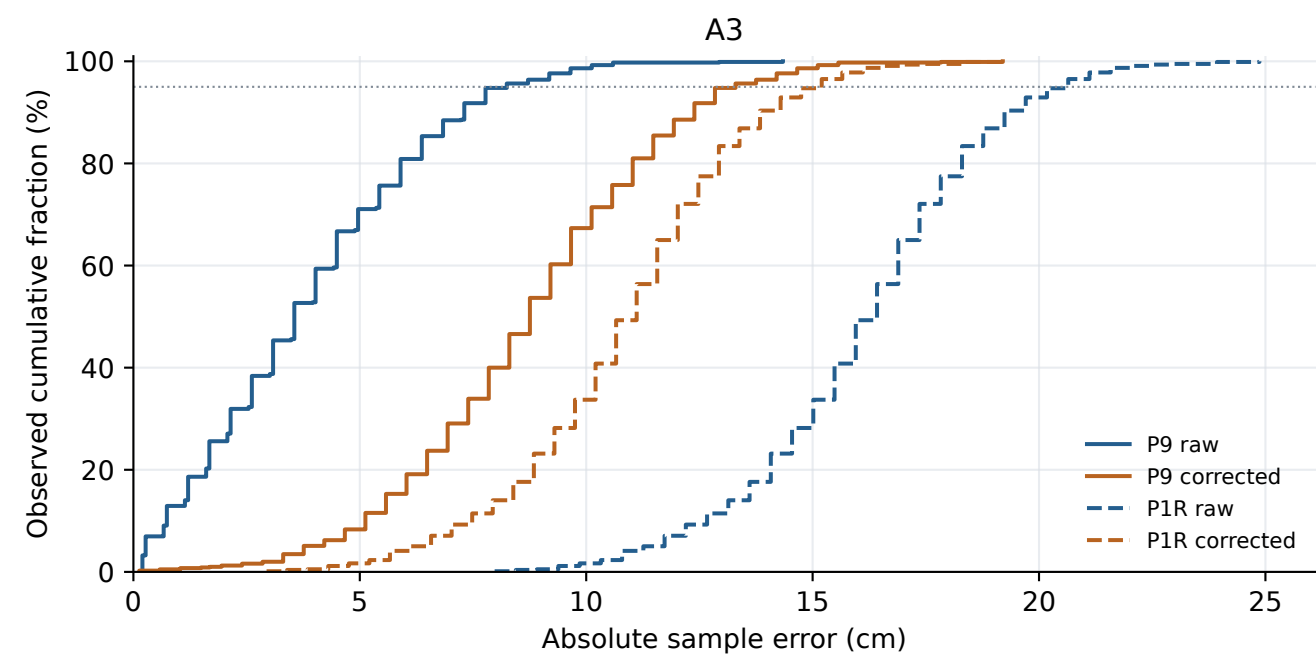
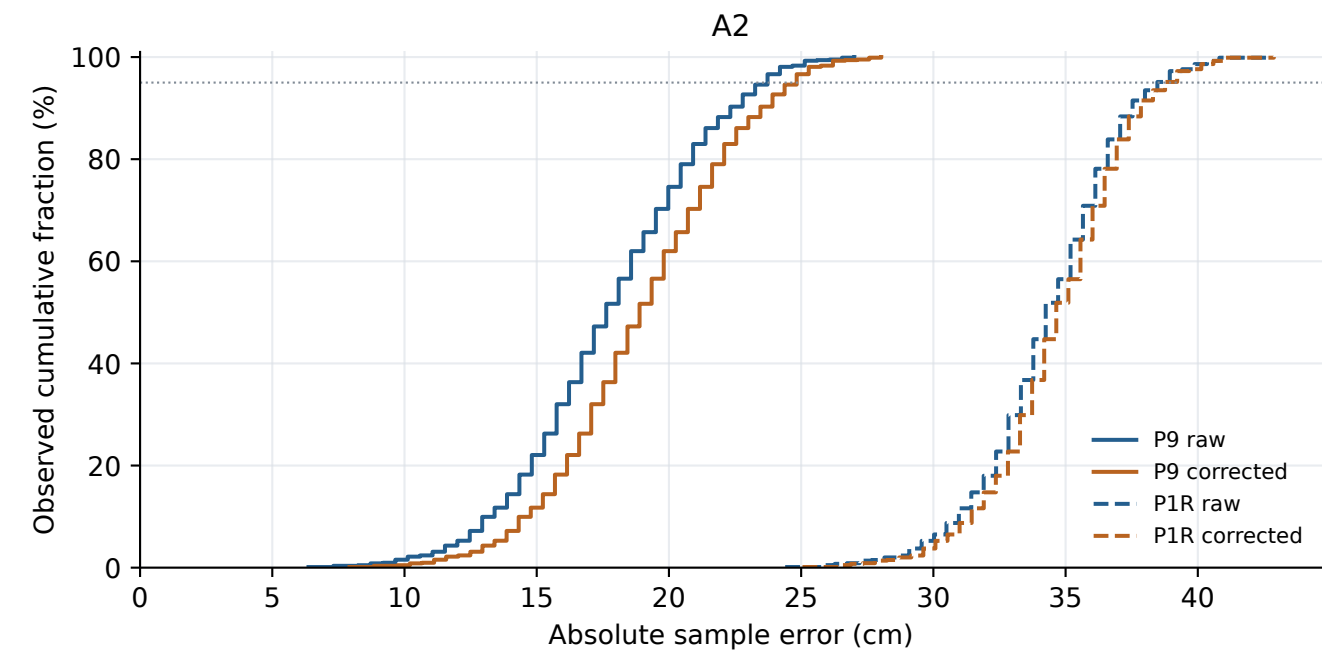
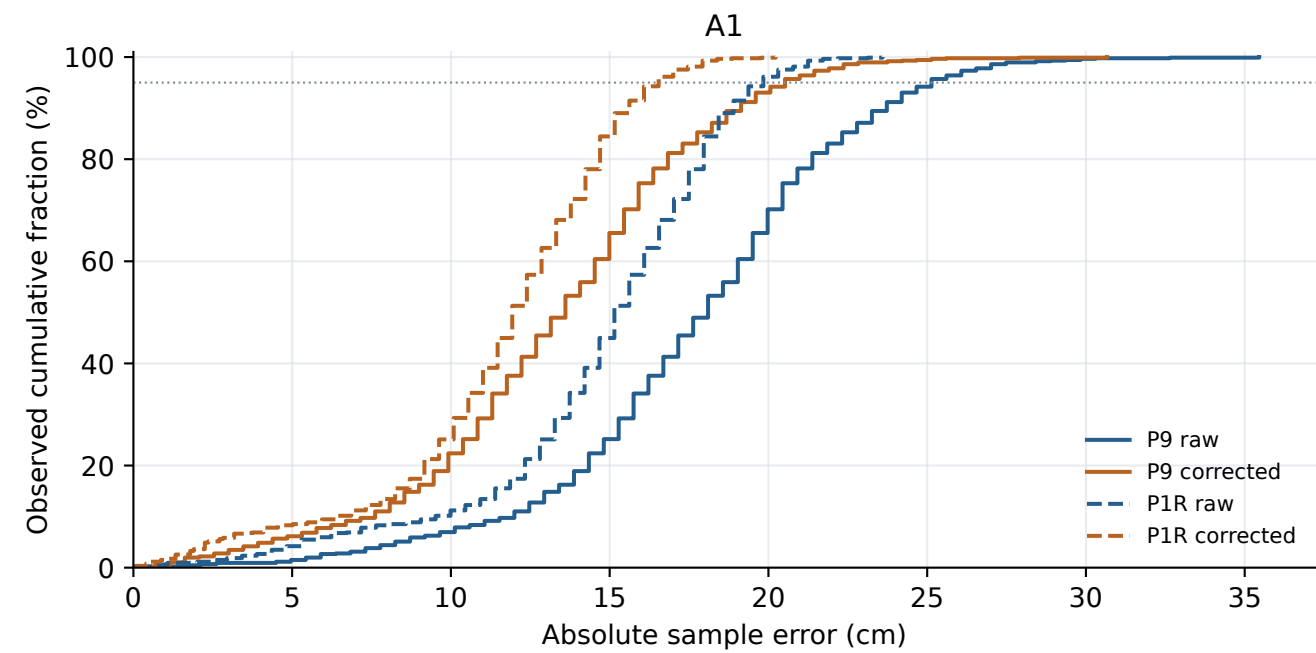
Box: median and quartiles. Whiskers: 1.5 IQR. All beyond-whisker samples remain plotted and included.



Interpretation. The symmetric logarithmic axis retains large errors while keeping the central region readable. The 1.5-IQR convention is visual only; it is not an exclusion rule.

05 Held-out single-sample error distributions

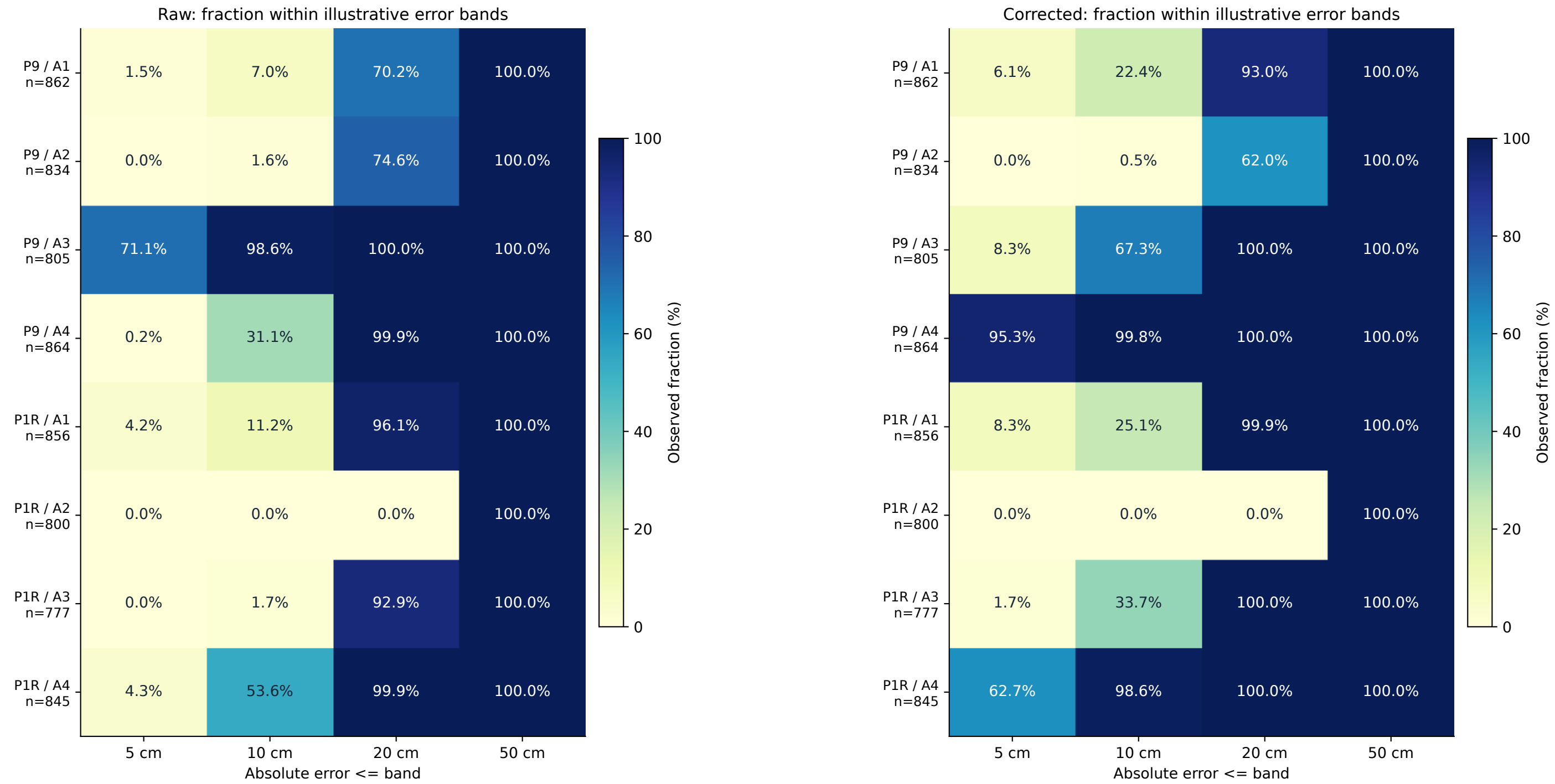
P9 solid; P1R dashed. A curve further left means smaller errors for that capture.



Interpretation. Each CDF has its own empirical denominator. Samples are serially dependent and these curves are descriptive. No held-out data are used to refit or select a new correction.

06 How often were errors within a chosen band?

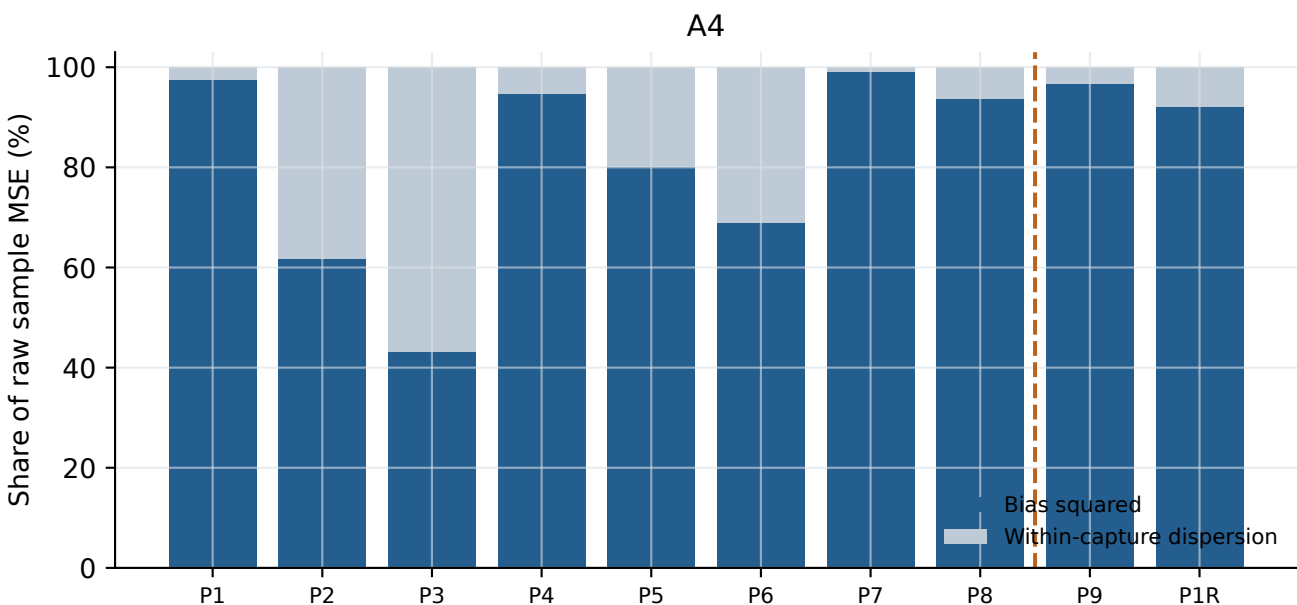
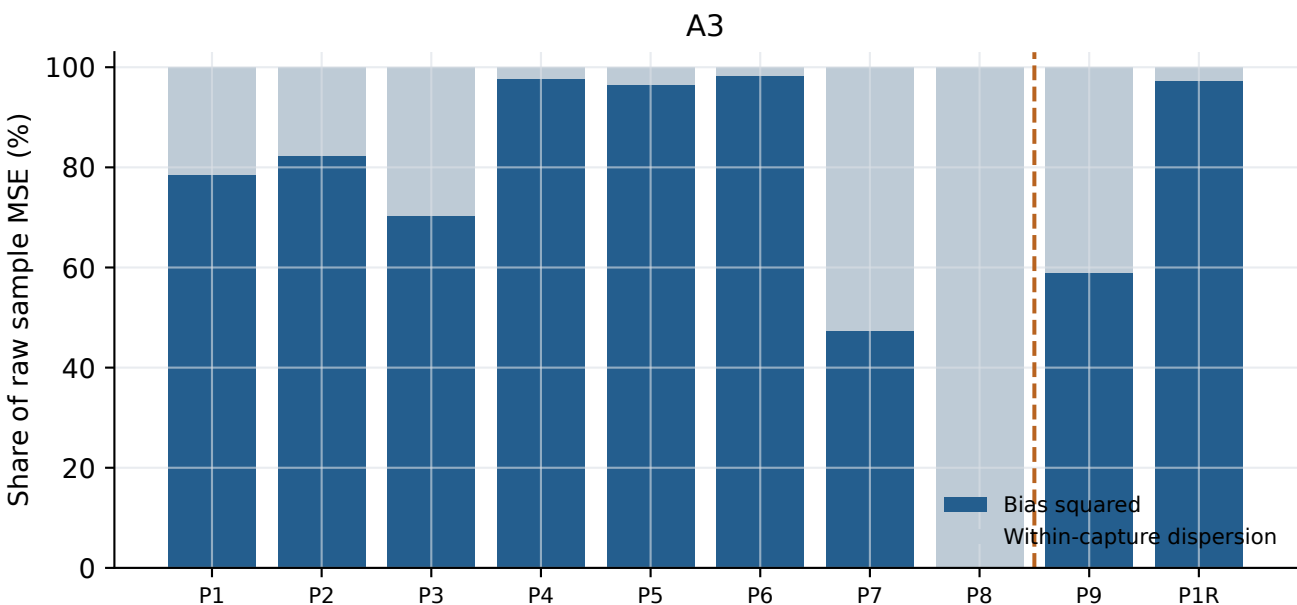
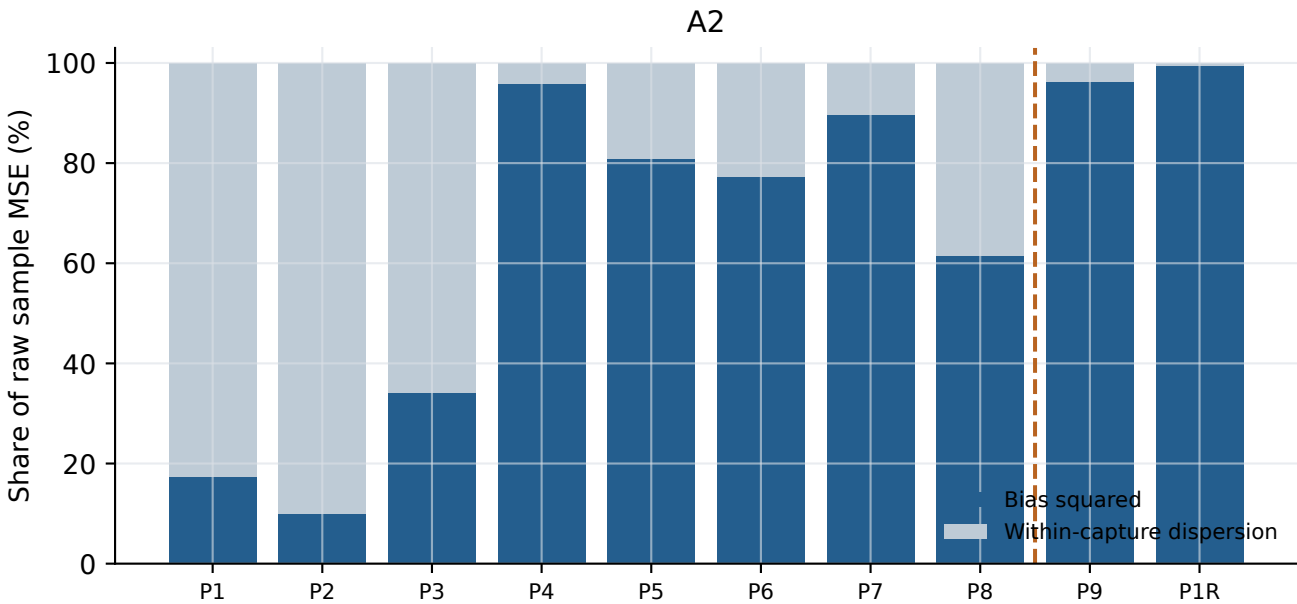
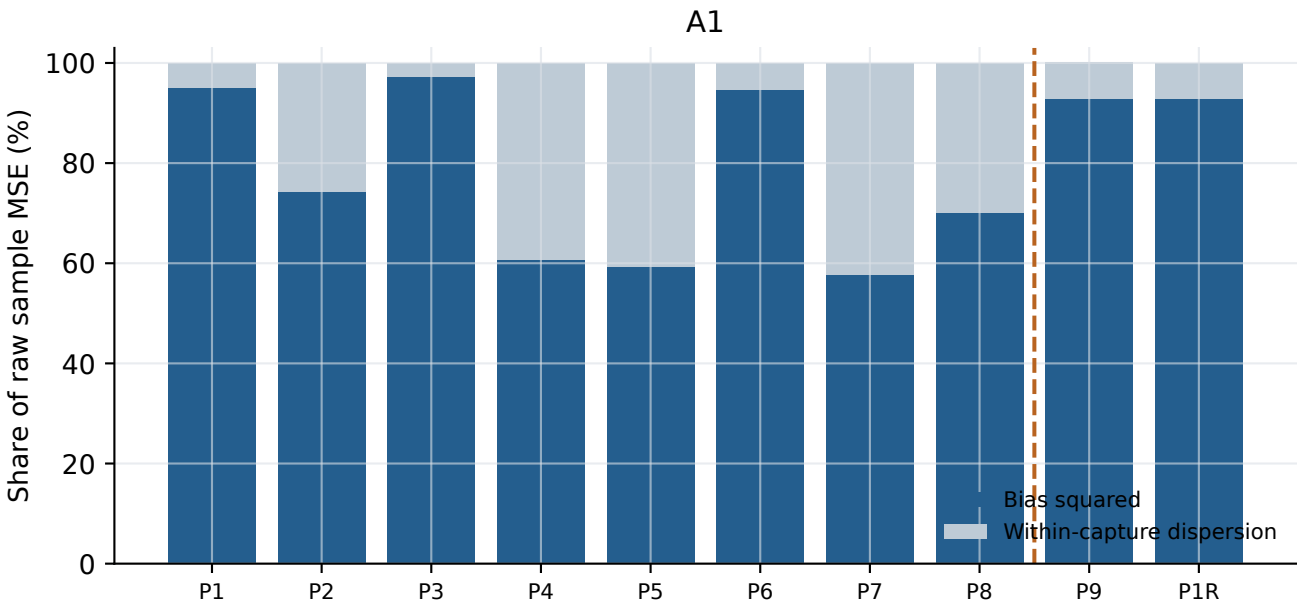
Illustrative 5, 10, 20 and 50 cm bands are comparison aids. No acceptance specification was supplied.



Interpretation. Percentages are conditional on received canonical-valid unique ranges; they exclude invalid and unavailable measurements. They are not availability-adjusted success probabilities or safety guarantees.

07 How much error comes from bias versus spread?

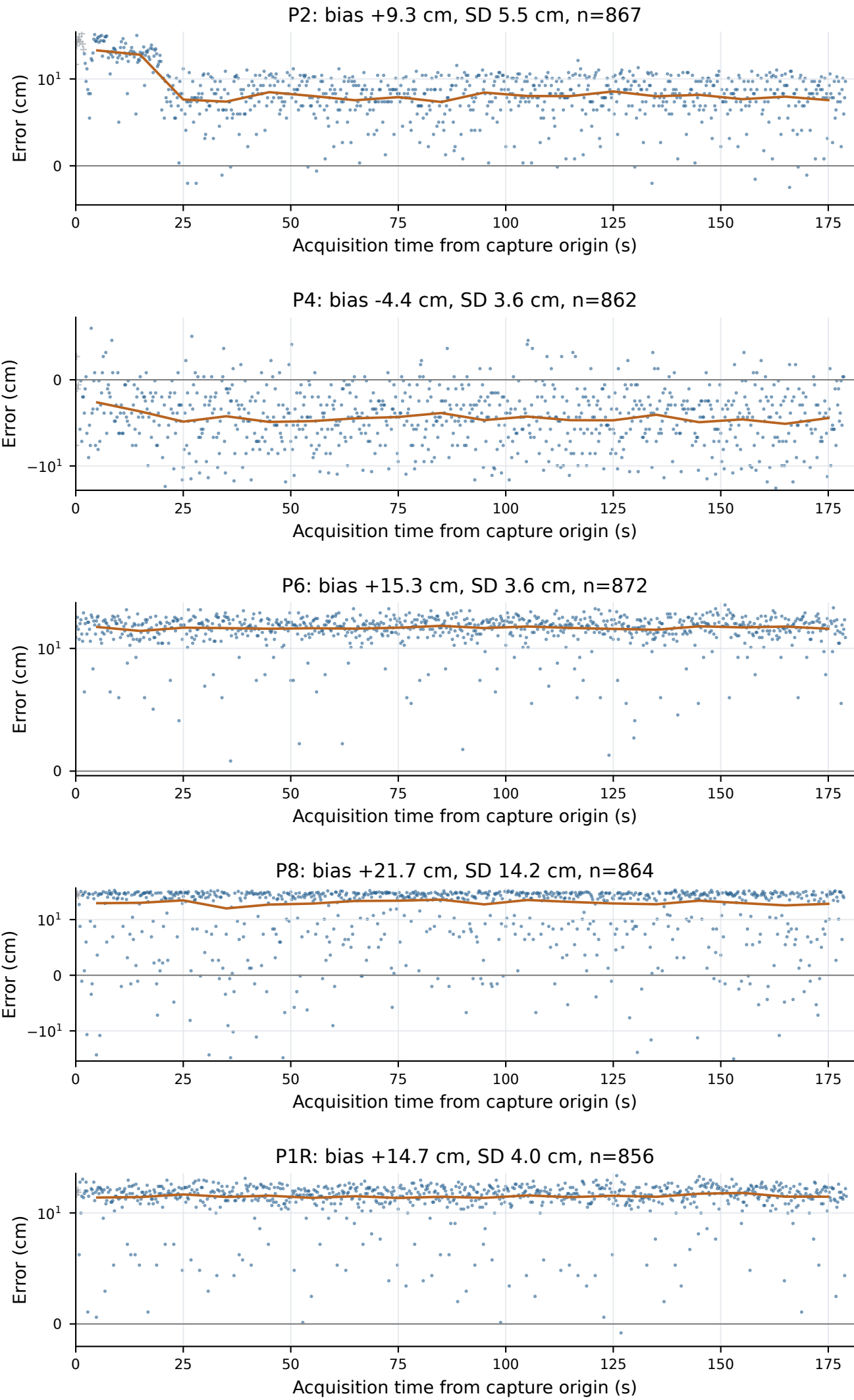
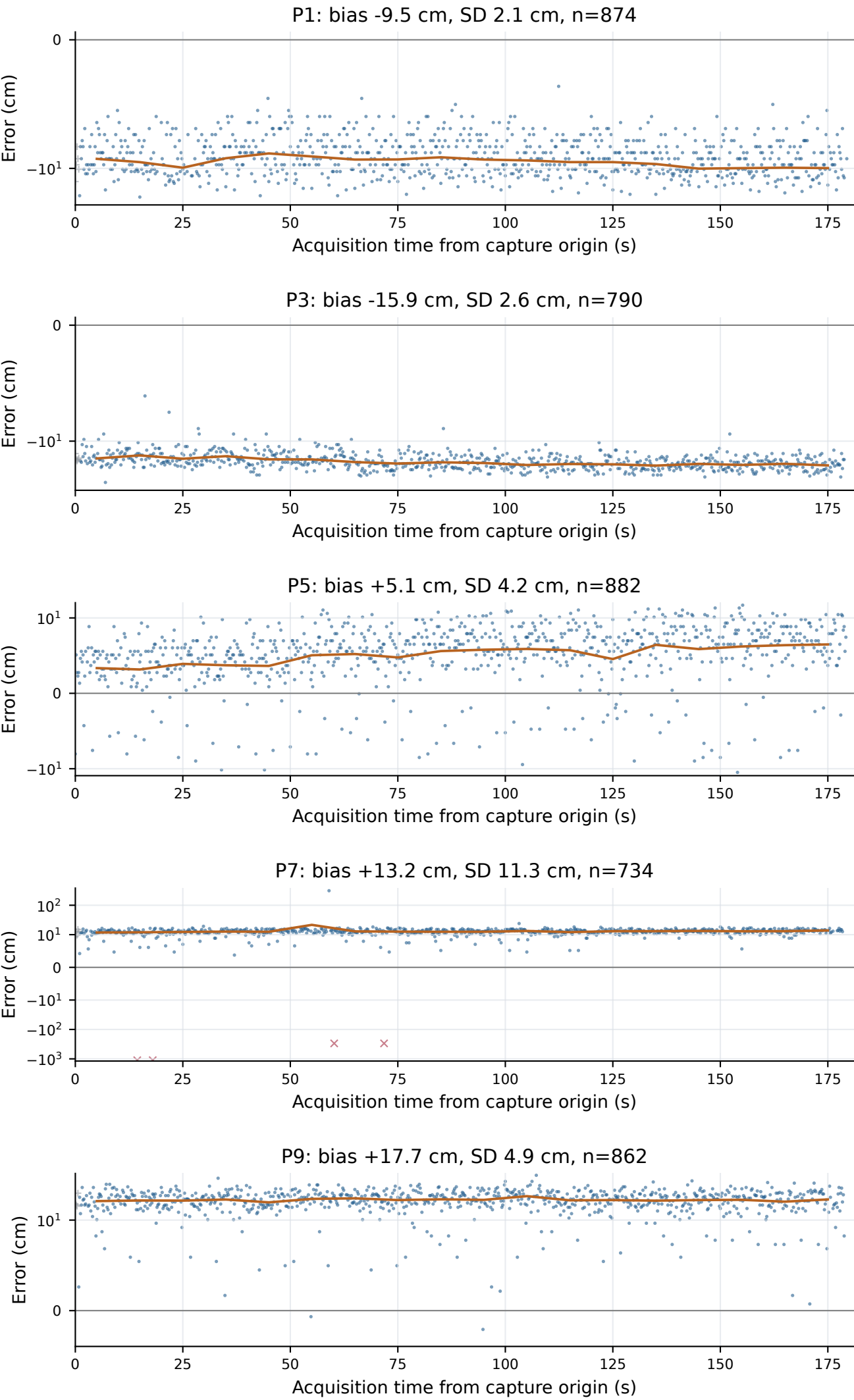
Exact finite-sample identity: $RMSE^2 = bias^2 + ((n-1)/n) * sample_SD^2$.



Interpretation. This is an algebraic decomposition against the nominal reference, not a causal partition of hardware errors. A large bias share explains why averaging alone may leave substantial offset.

08 A1: complete received time histories

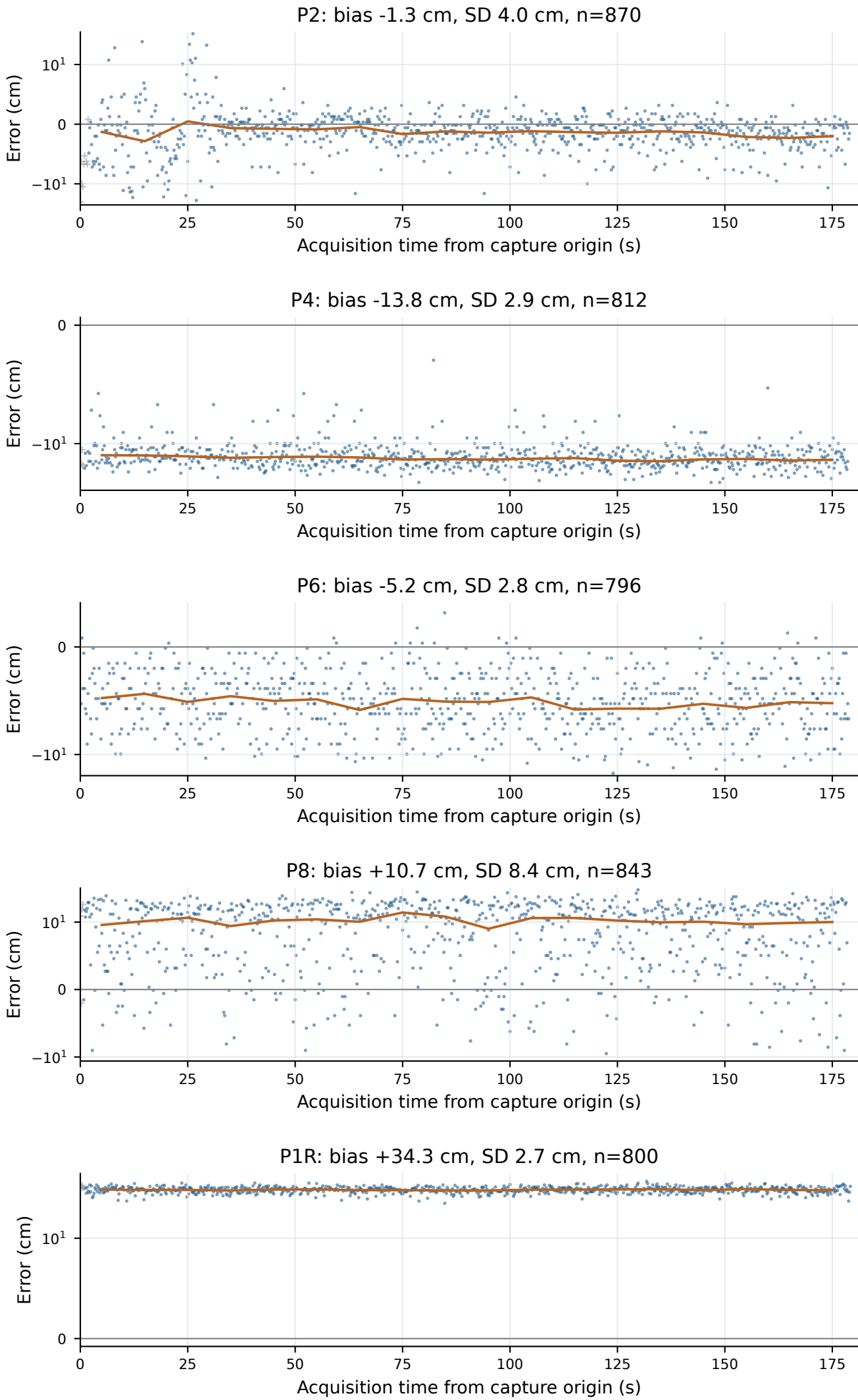
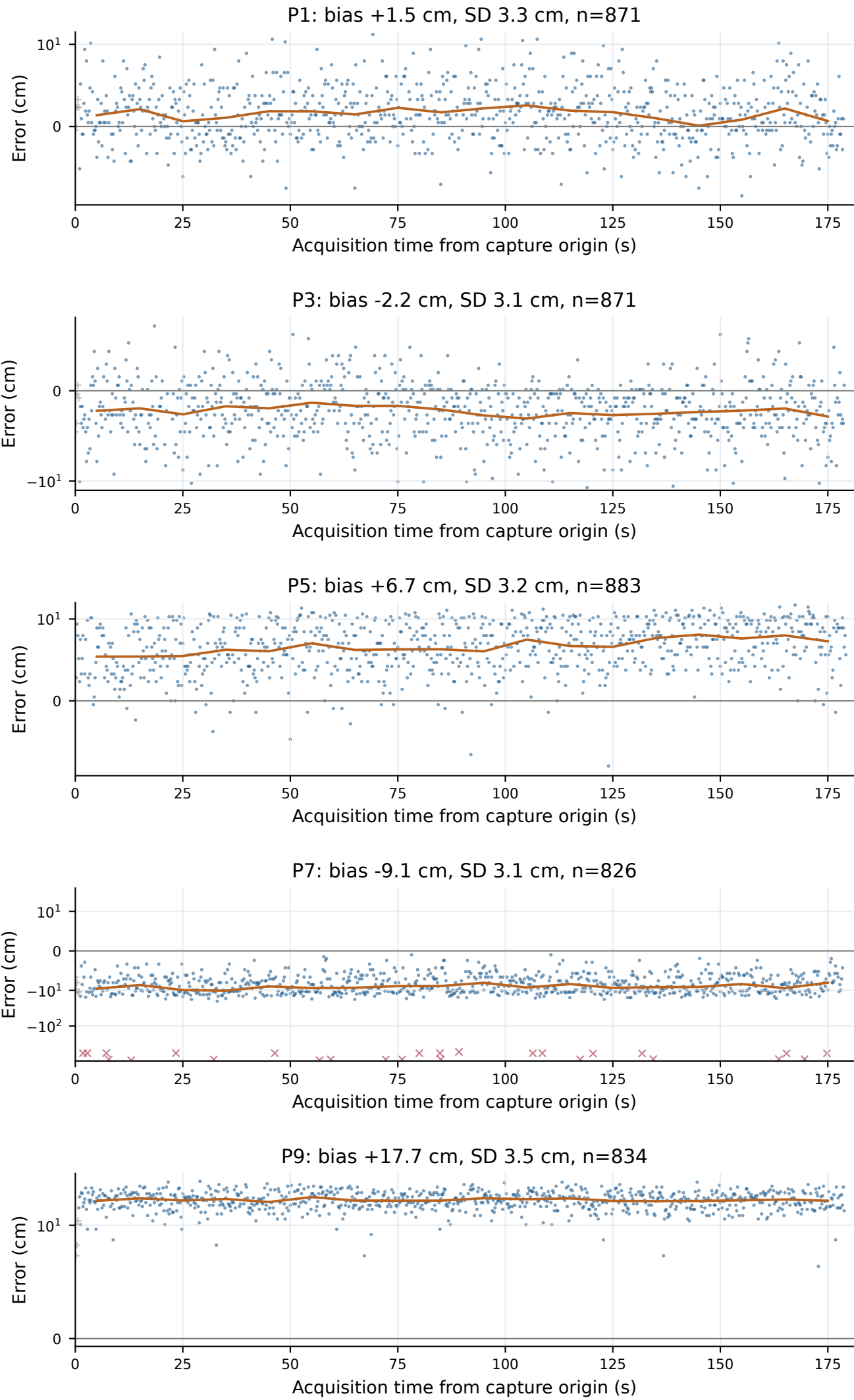
Blue: statistical population. Red x: basic-invalid, excluded from metrics. Gray +: before-status. Orange: 10 s mean.



Interpretation. Times use tag acquisition clocks within each capture. No samples are interpolated. Panels use symlog axes and individual limits. Orange bins summarize existing samples and do not reconstruct missing data.

09 A2: complete received time histories

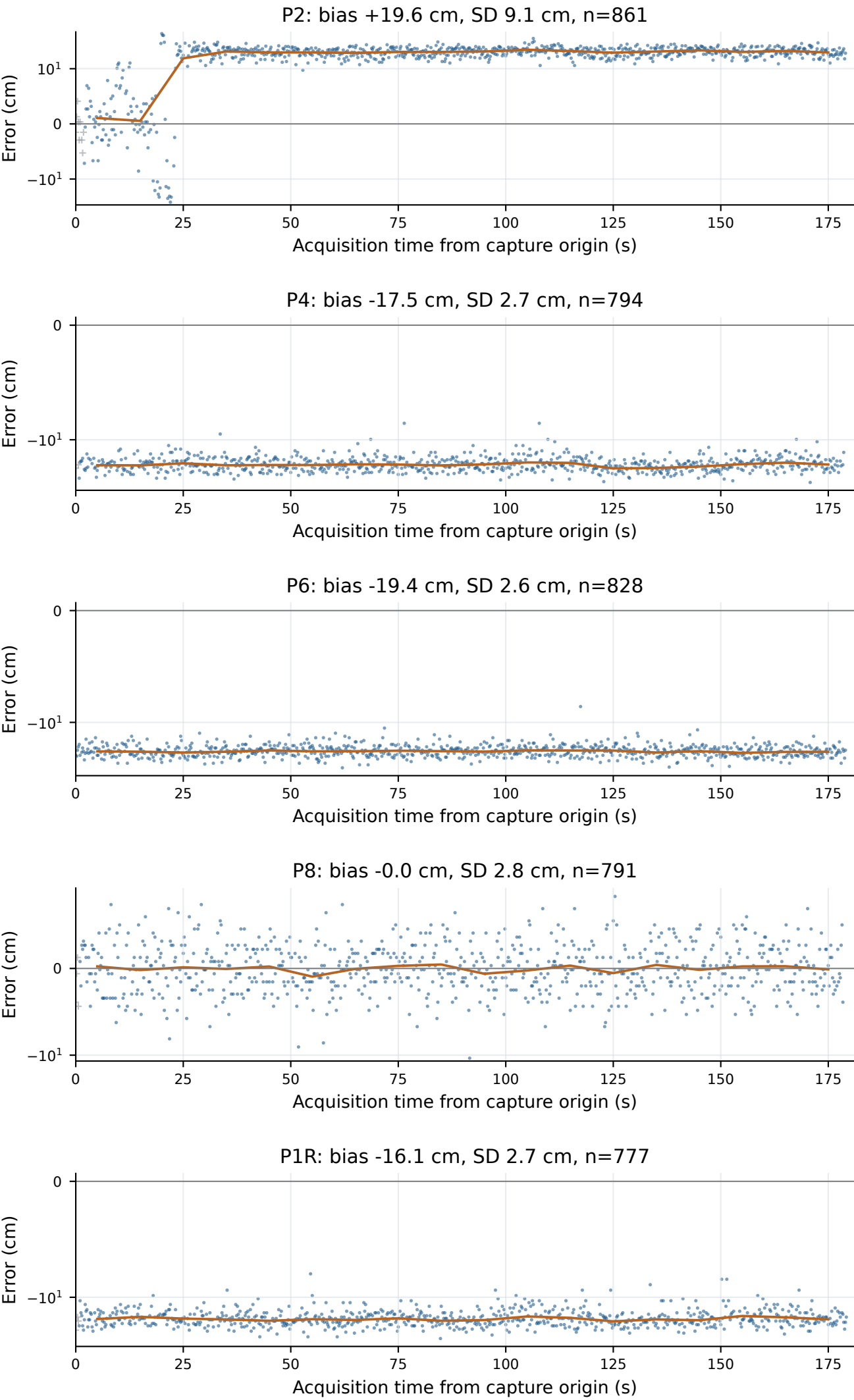
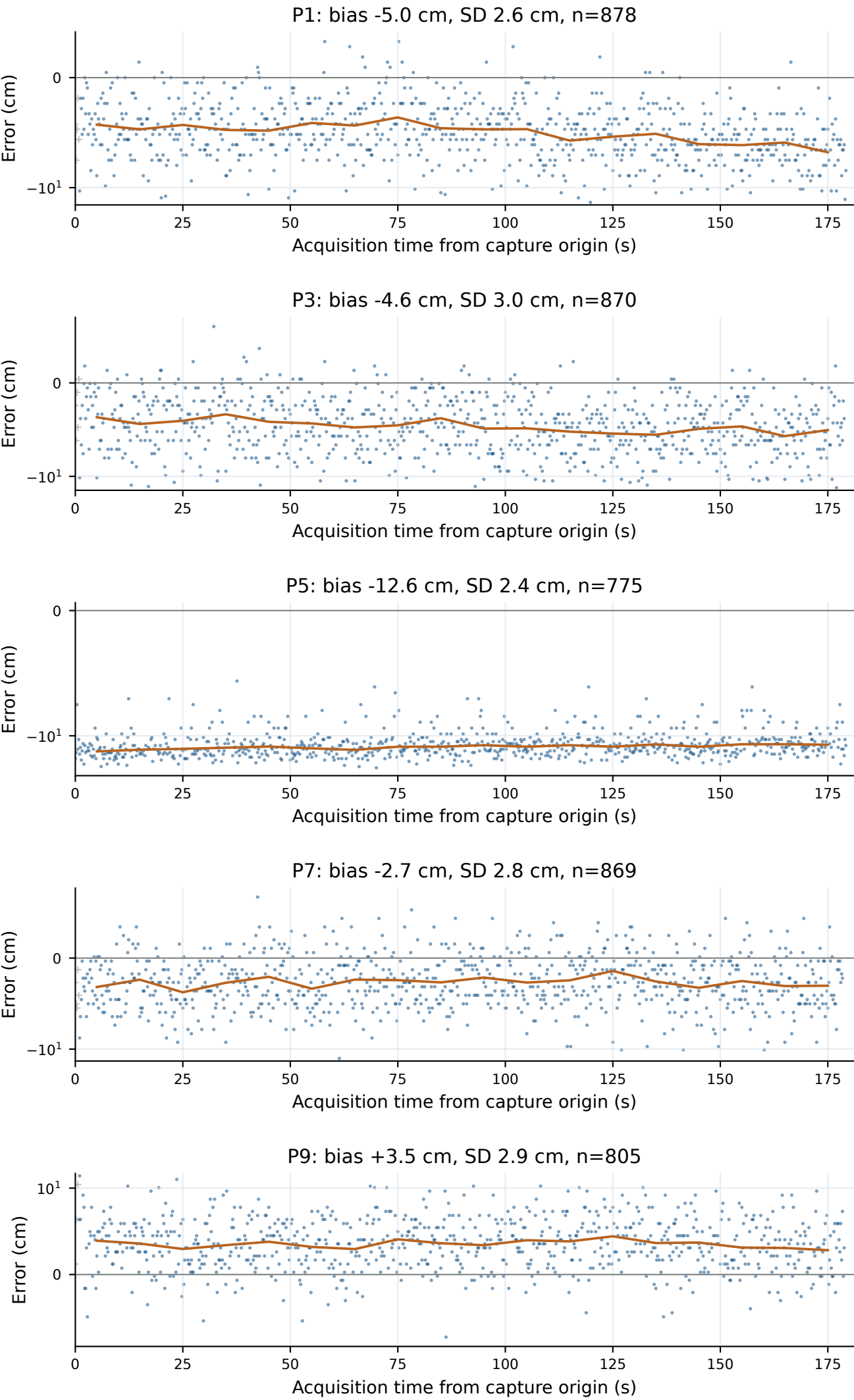
Blue: statistical population. Red x: basic-invalid, excluded from metrics. Gray +: before-status. Orange: 10 s mean.



Interpretation. Times use tag acquisition clocks within each capture. No samples are interpolated. Panels use symlog axes and individual limits. Orange bins summarize existing samples and do not reconstruct missing data.

10 A3: complete received time histories

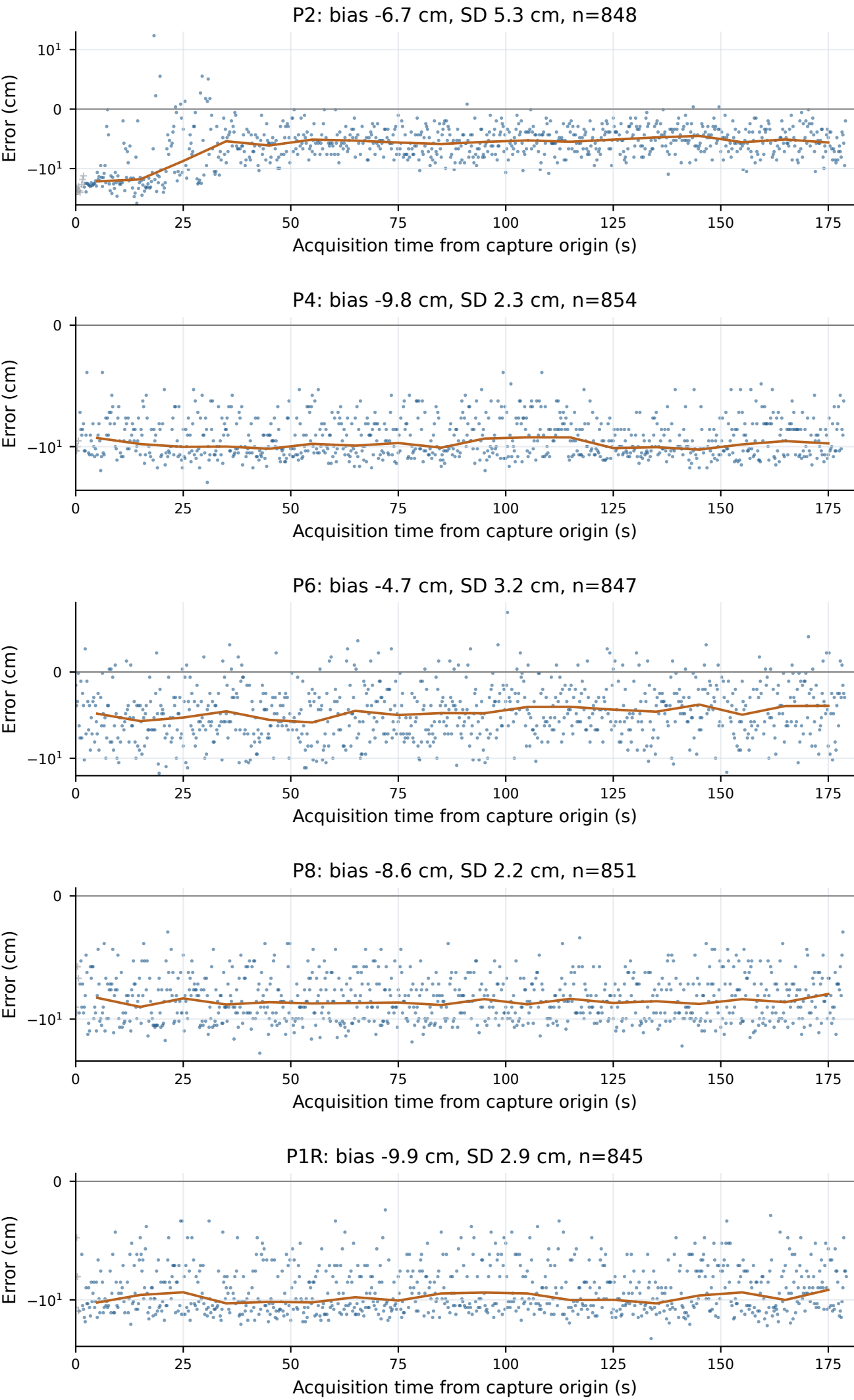
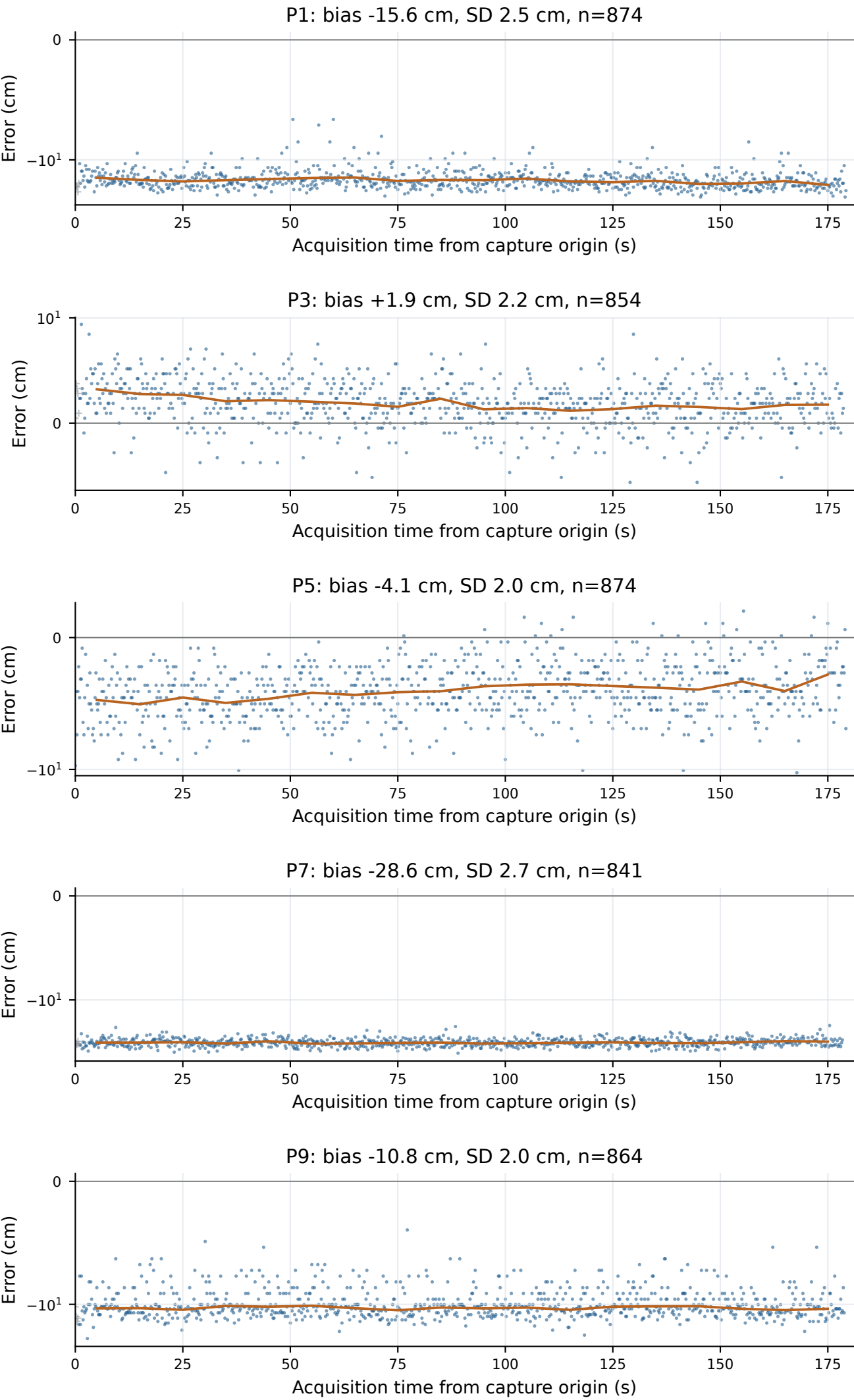
Blue: statistical population. Red x: basic-invalid, excluded from metrics. Gray +: before-status. Orange: 10 s mean.



Interpretation. Times use tag acquisition clocks within each capture. No samples are interpolated. Panels use symlog axes and individual limits. Orange bins summarize existing samples and do not reconstruct missing data.

11 A4: complete received time histories

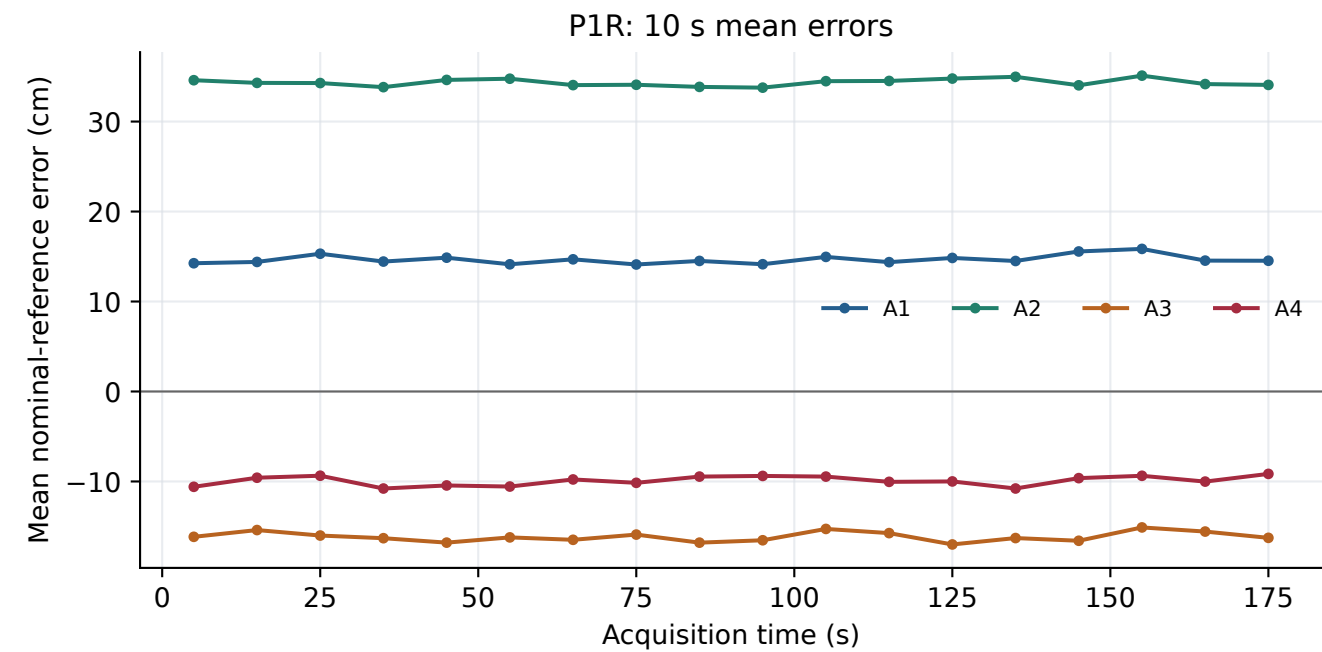
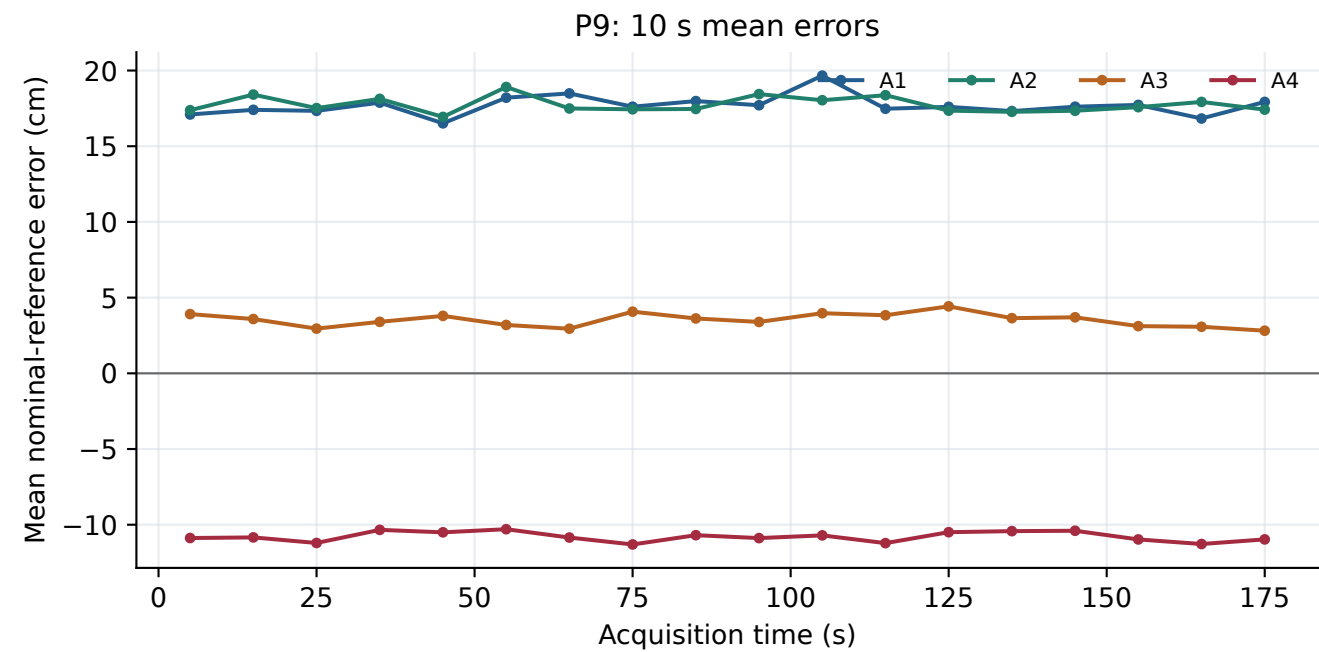
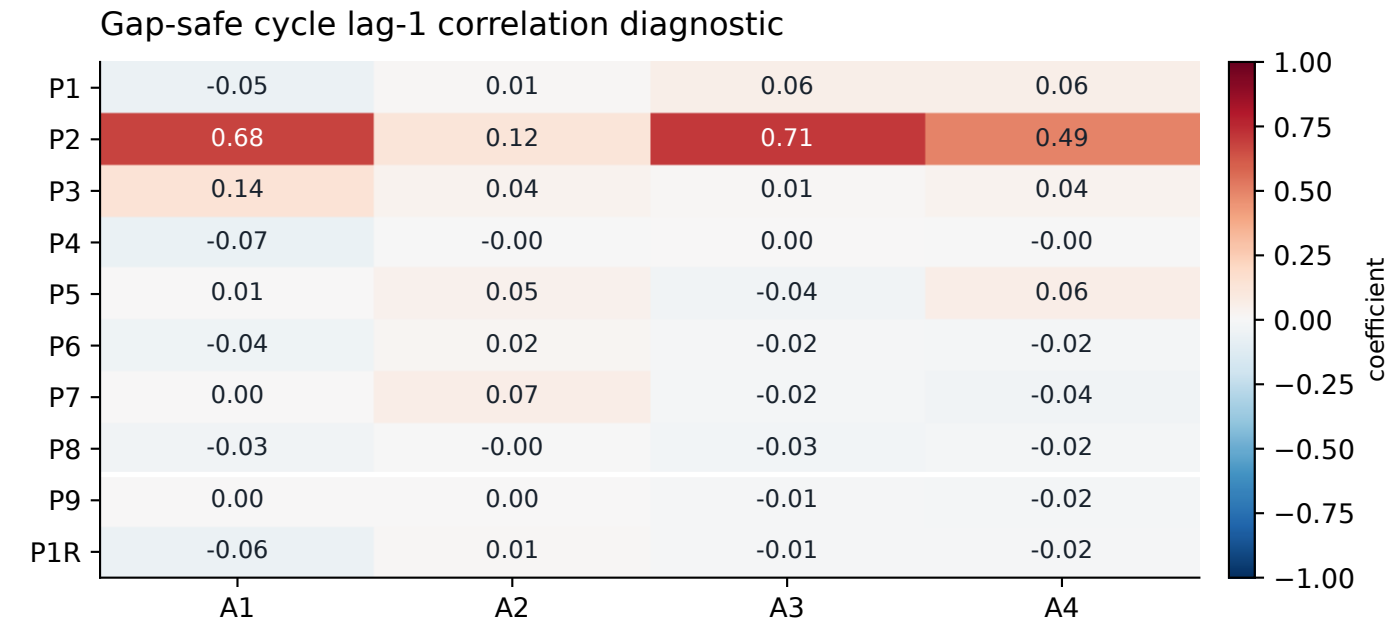
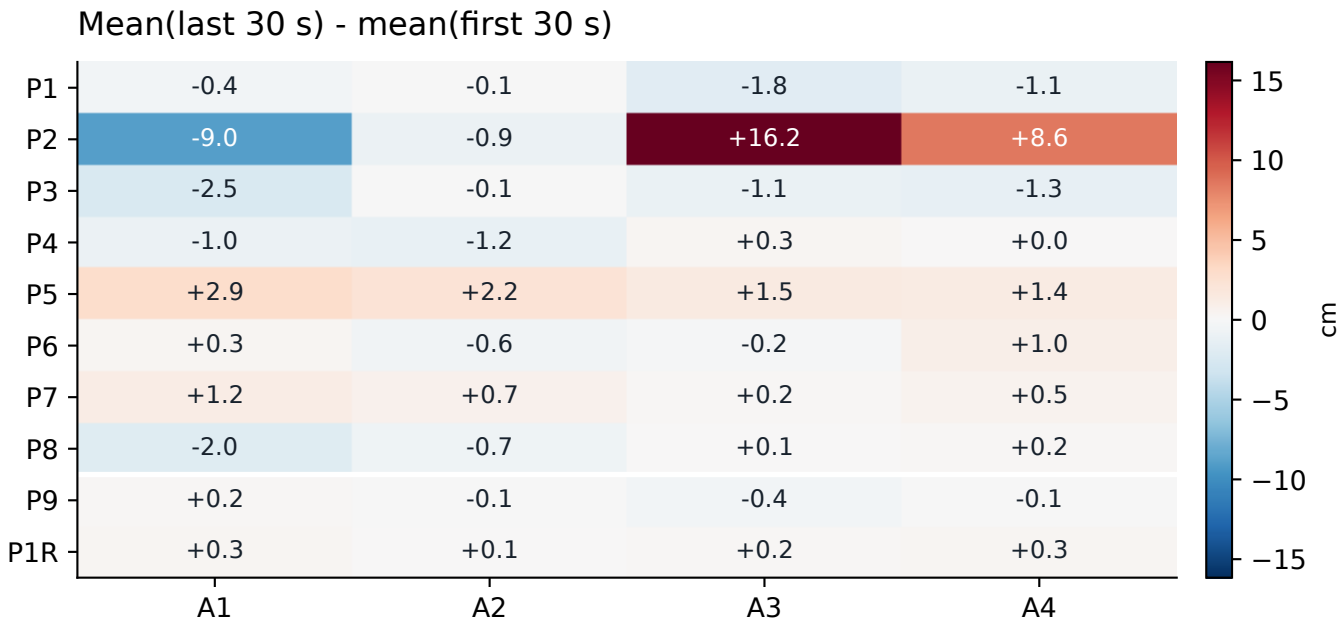
Blue: statistical population. Red x: basic-invalid, excluded from metrics. Gray +: before-status. Orange: 10 s mean.



Interpretation. Times use tag acquisition clocks within each capture. No samples are interpolated. Panels use symlog axes and individual limits. Orange bins summarize existing samples and do not reconstruct missing data.

12 Stability over each 180 s stationary capture

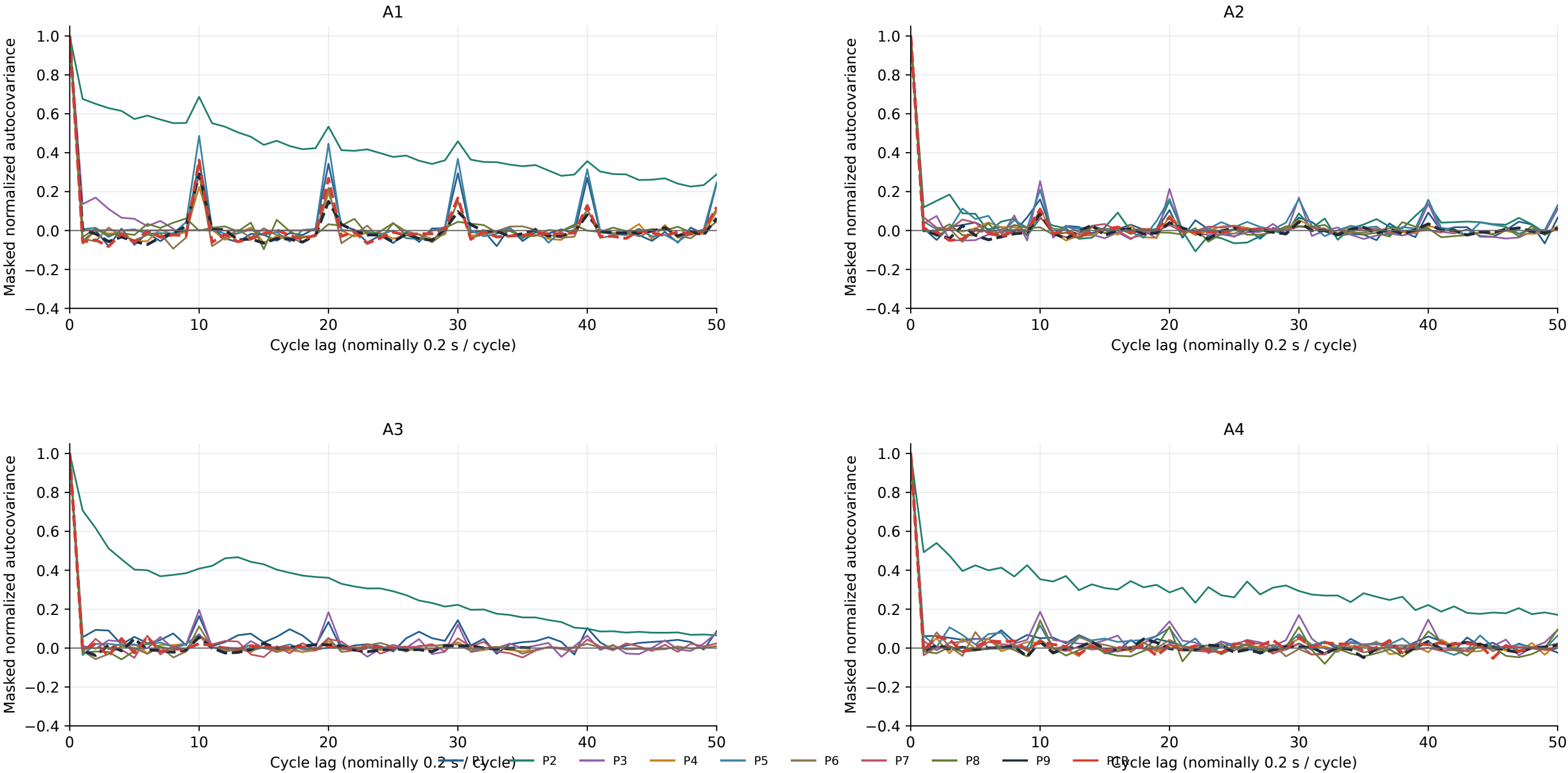
First/last windows are fixed at <30 s and >=150 s from the capture acquisition origin; populations are exported.



Interpretation. The signed first/last change is descriptive. It does not isolate temperature, drift, multipath or supply effects. Autocorrelation indicates why sample count alone does not establish independent repeats or a standard error.

13 Temporal dependence without bridging missing cycles

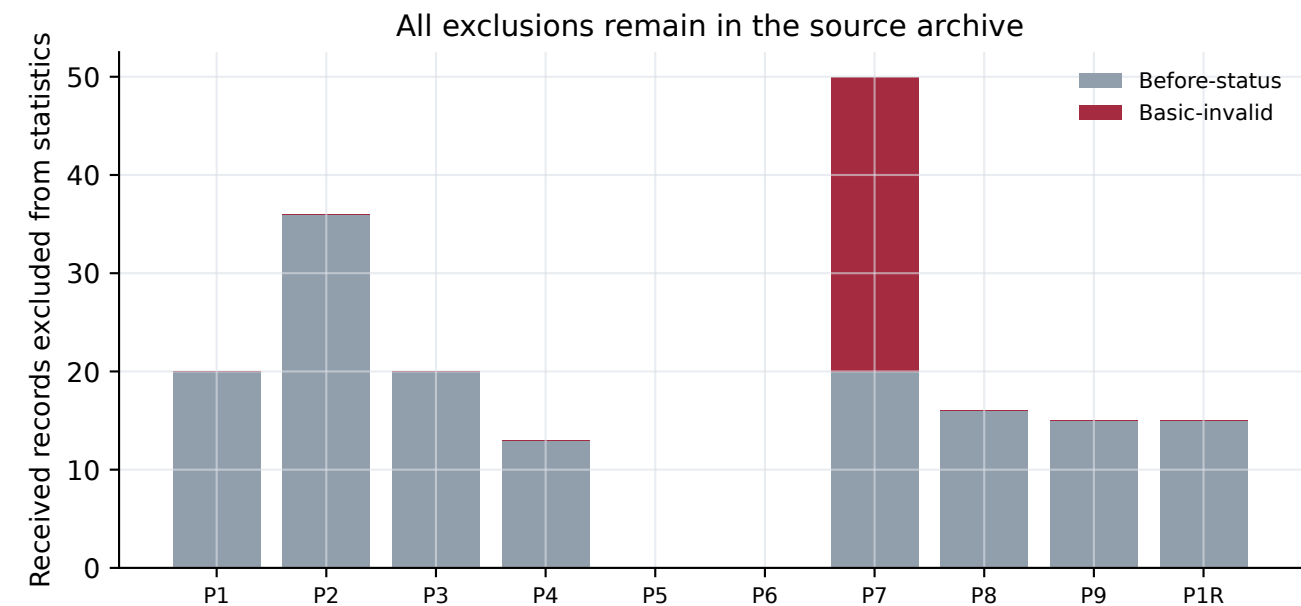
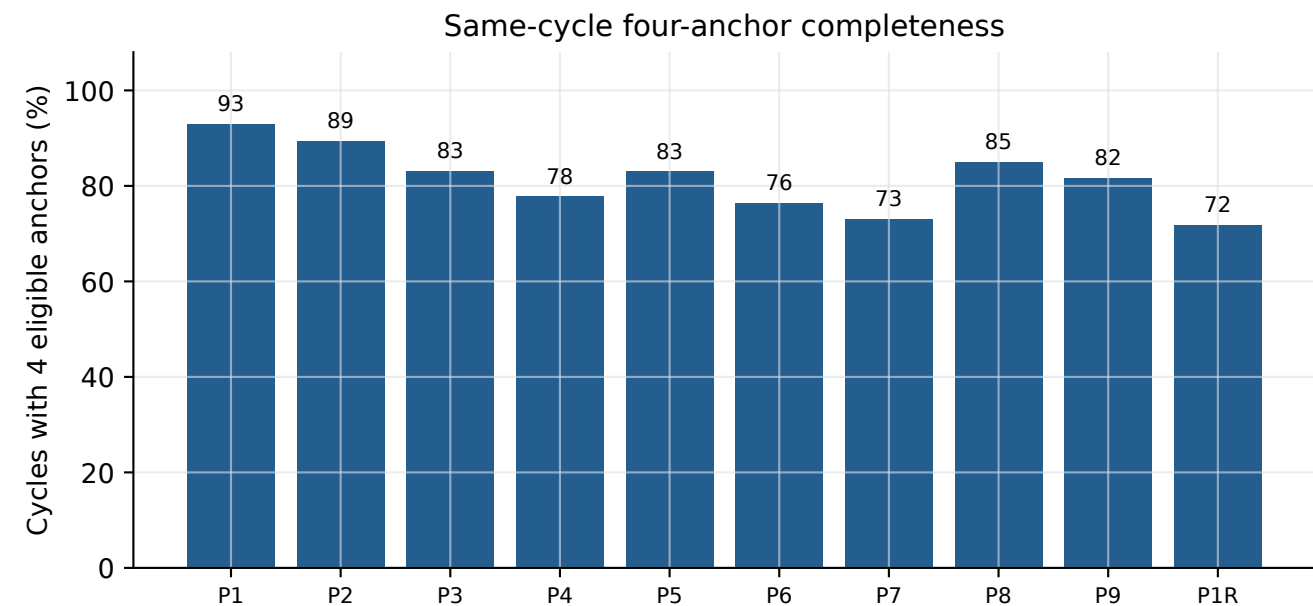
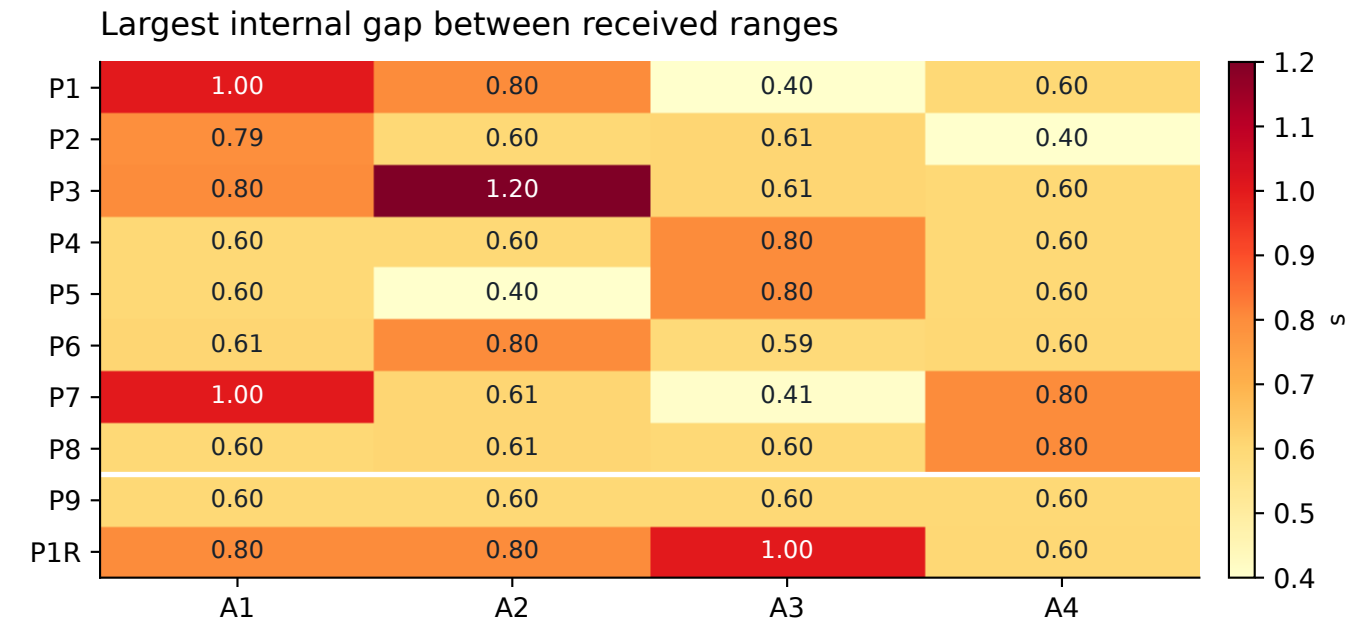
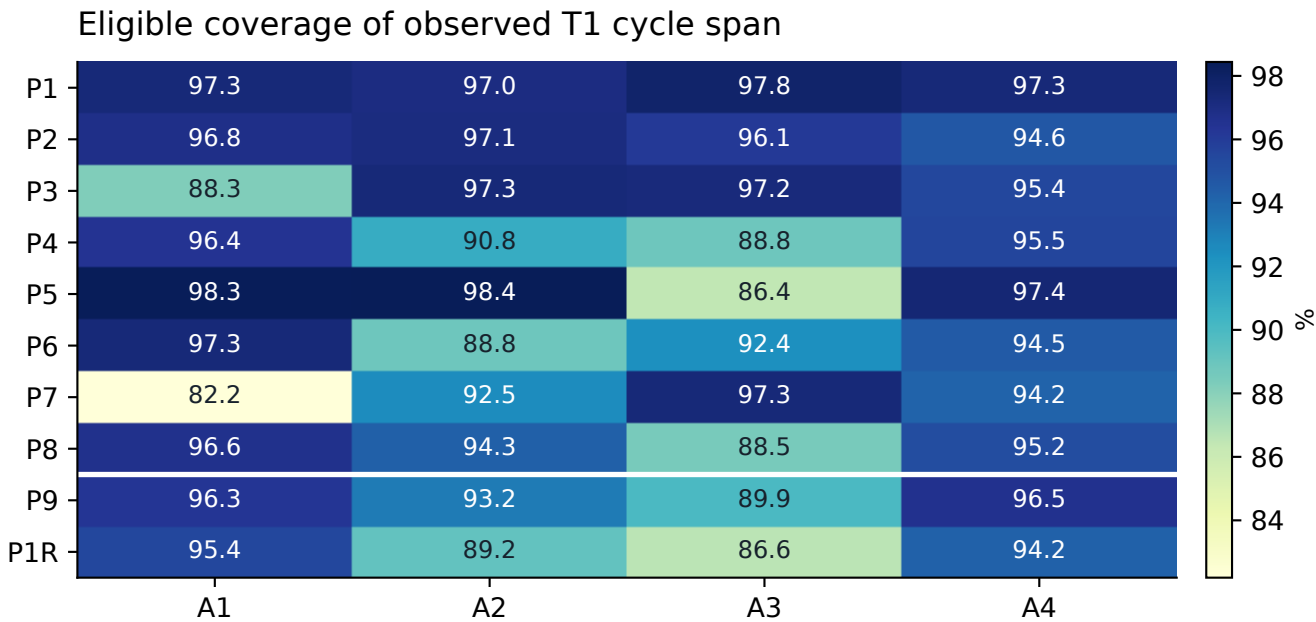
Run-demeaned covariance uses consecutive valid-cycle segments, 180-220 ms spacing. Lags with fewer than 30 pairs are omitted.



Interpretation. Normalization uses total squared deviations of that capture; missing-cycle cross-products are absent. This masked ACF is a descriptive diagnostic. No iid confidence bands, effective sample sizes or $\sqrt{n}$ uncertainty claims are made.

14 Data completeness and gaps

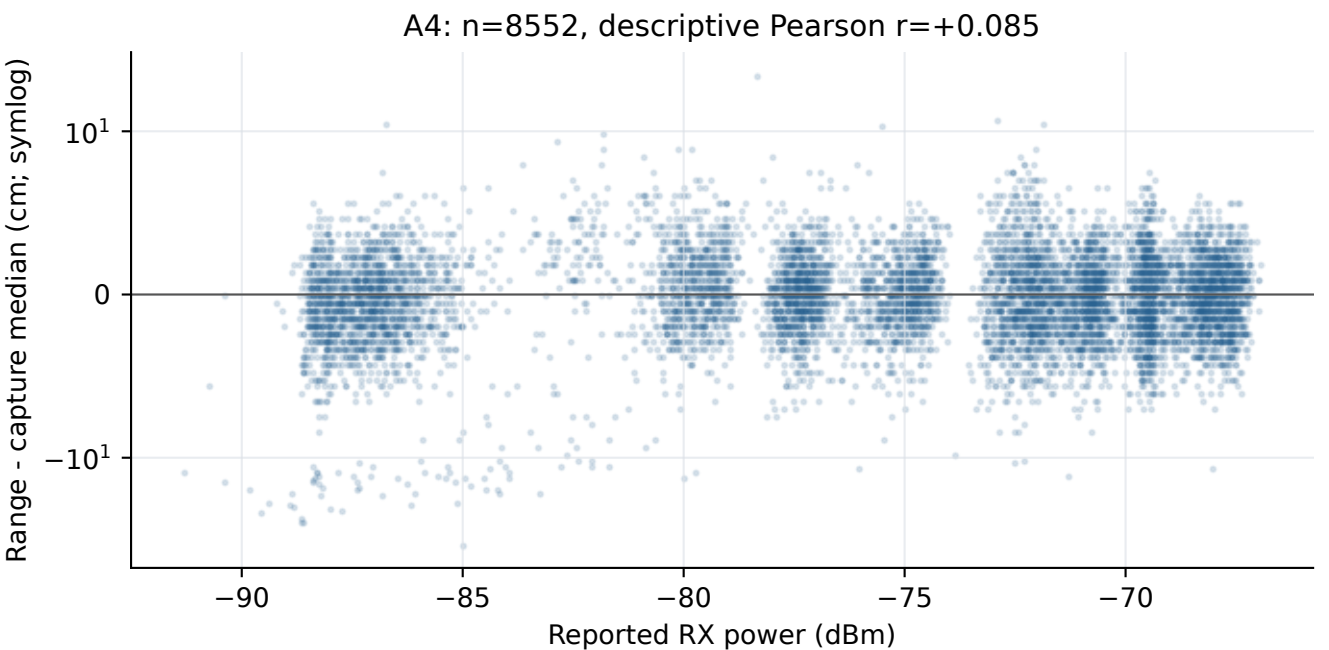
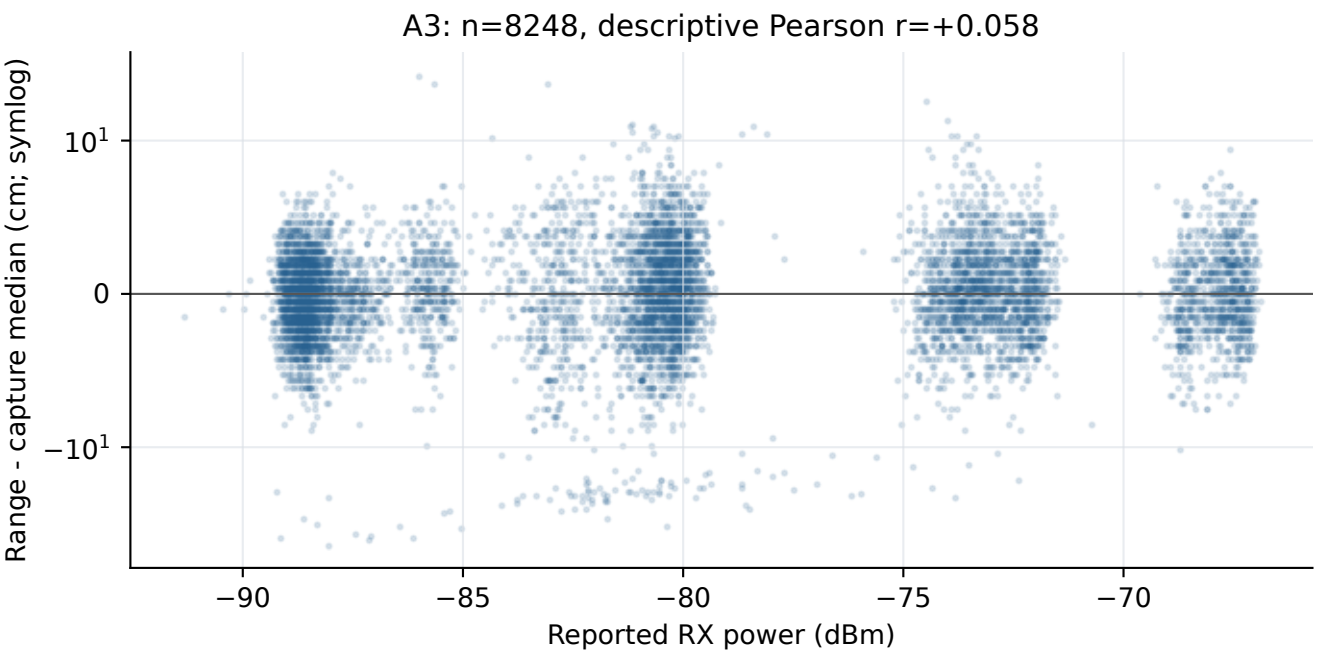
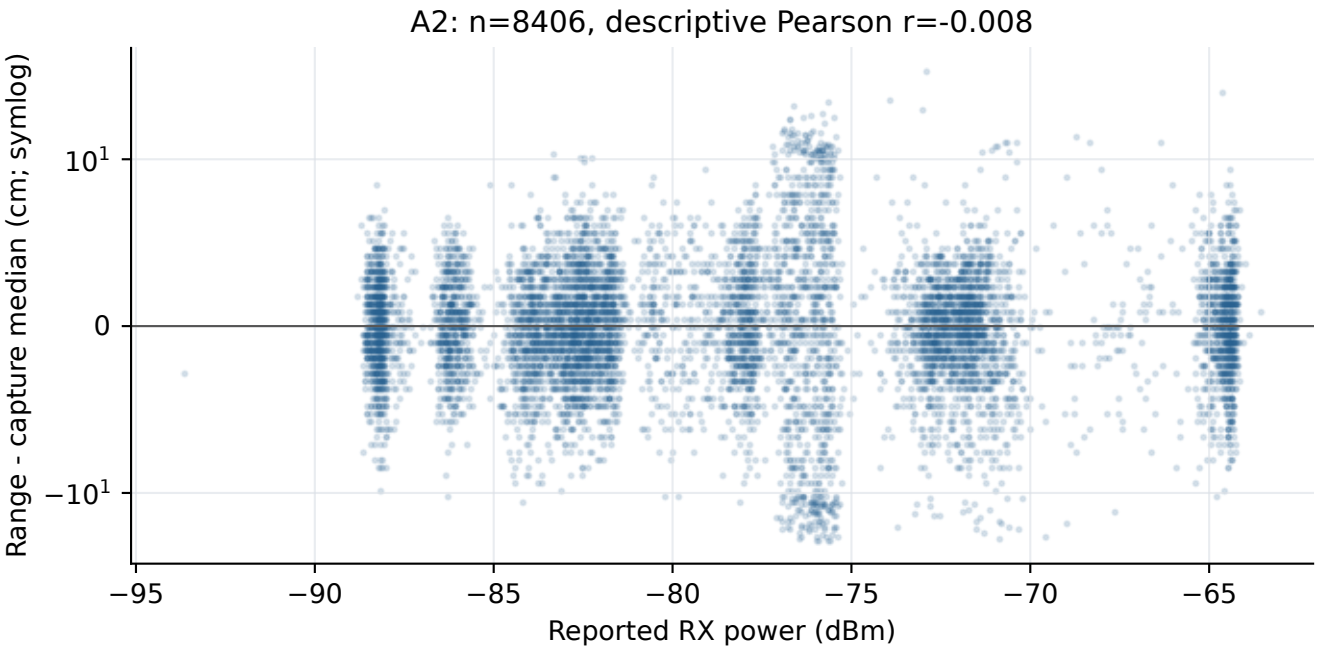
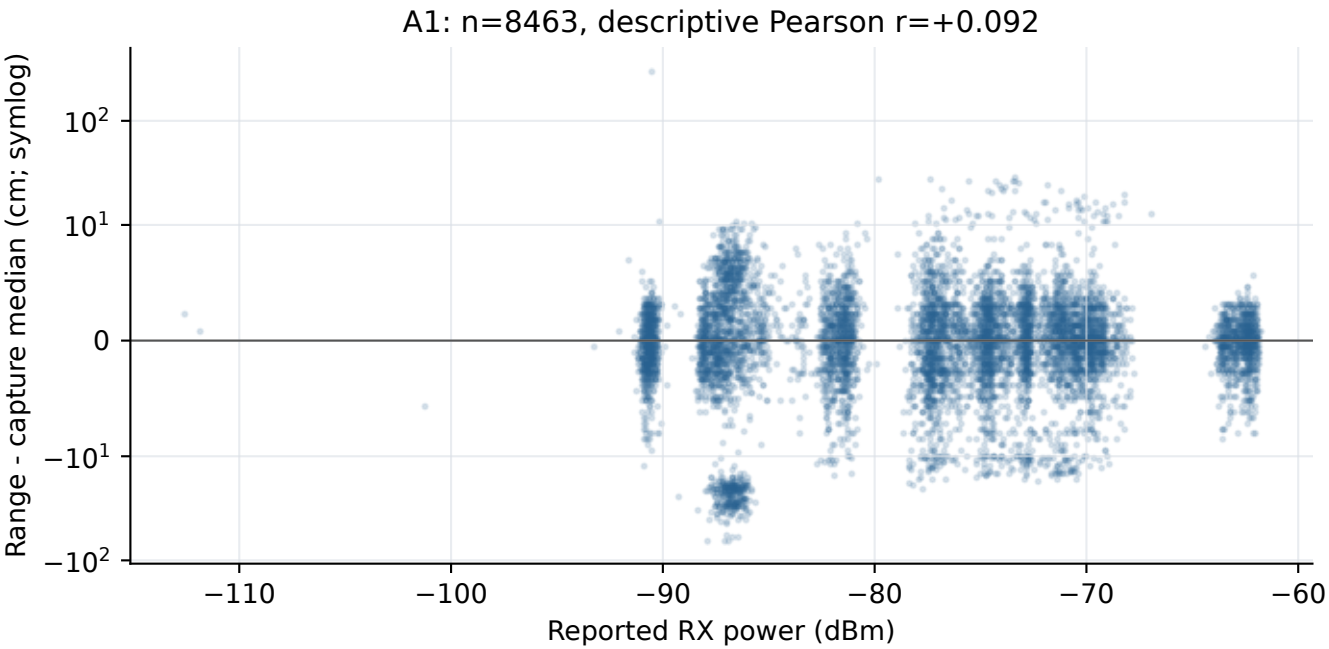
Cycle coverage denominator: every integer cycle between the earliest and latest received T1 cycle in each selected capture.



Interpretation. Coverage is not radio packet-delivery ratio: a missing report may arise anywhere in ranging, forwarding, serial or status acceptance. Maximum gap excludes leading/trailing capture silence. Four-anchor completeness is not XY accuracy.

15 RX power and deviations within each capture

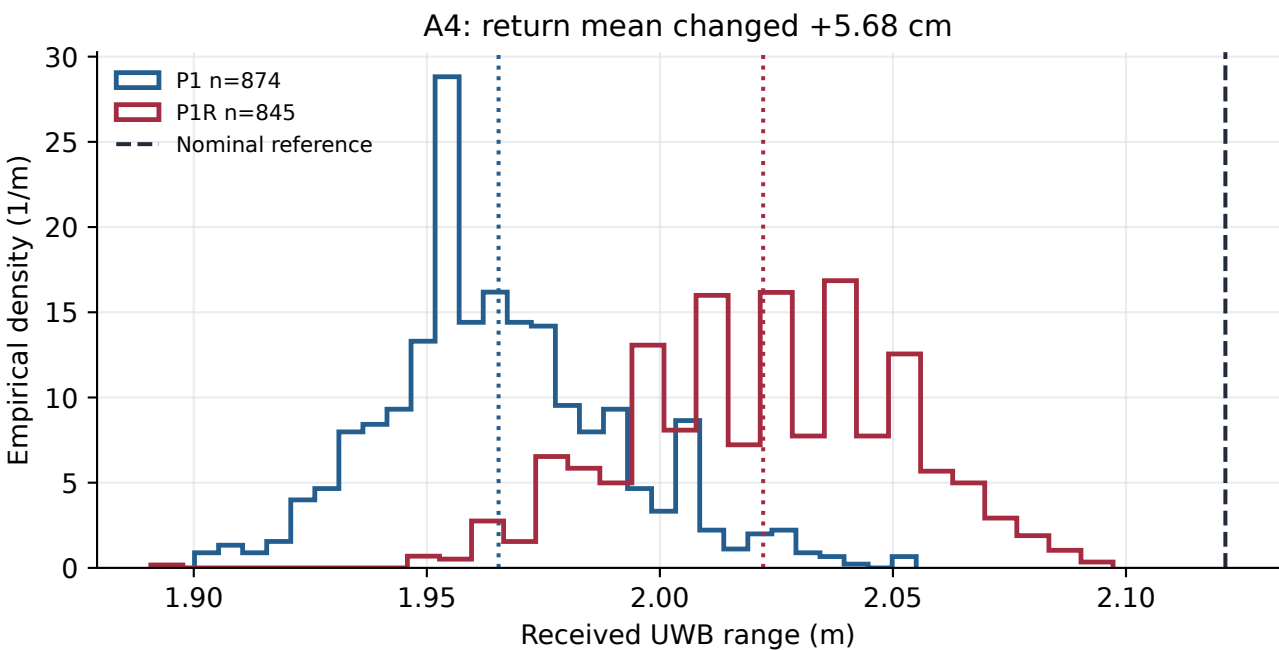
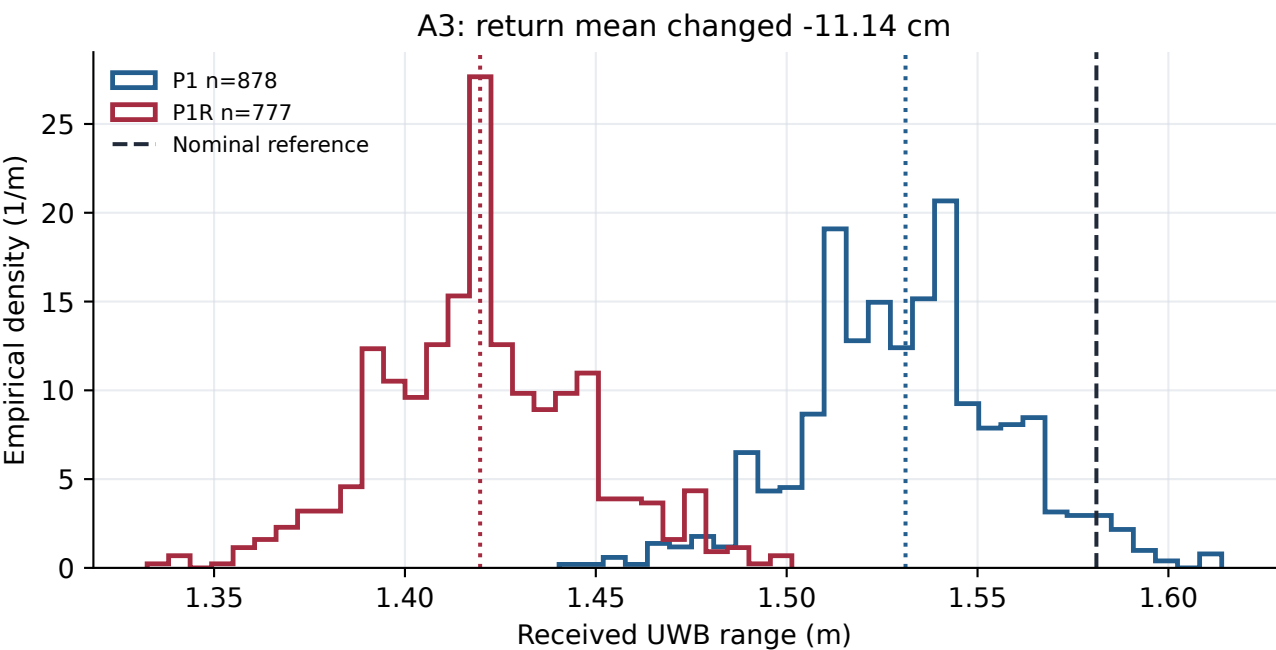
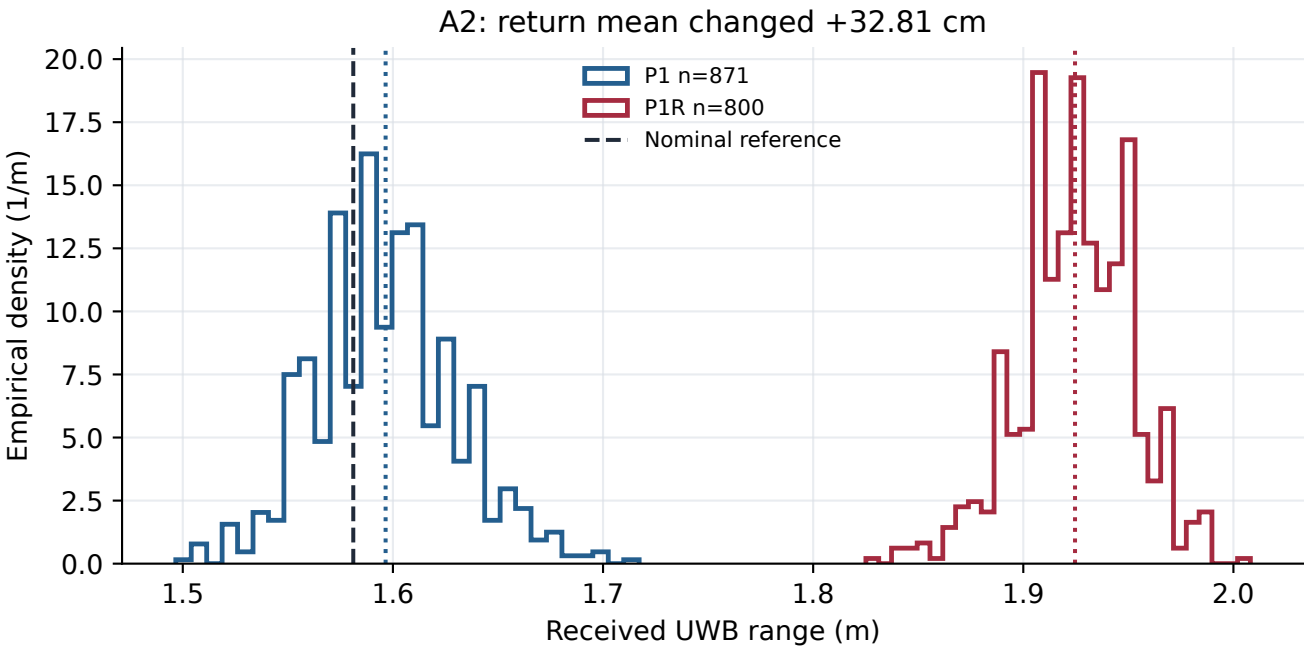
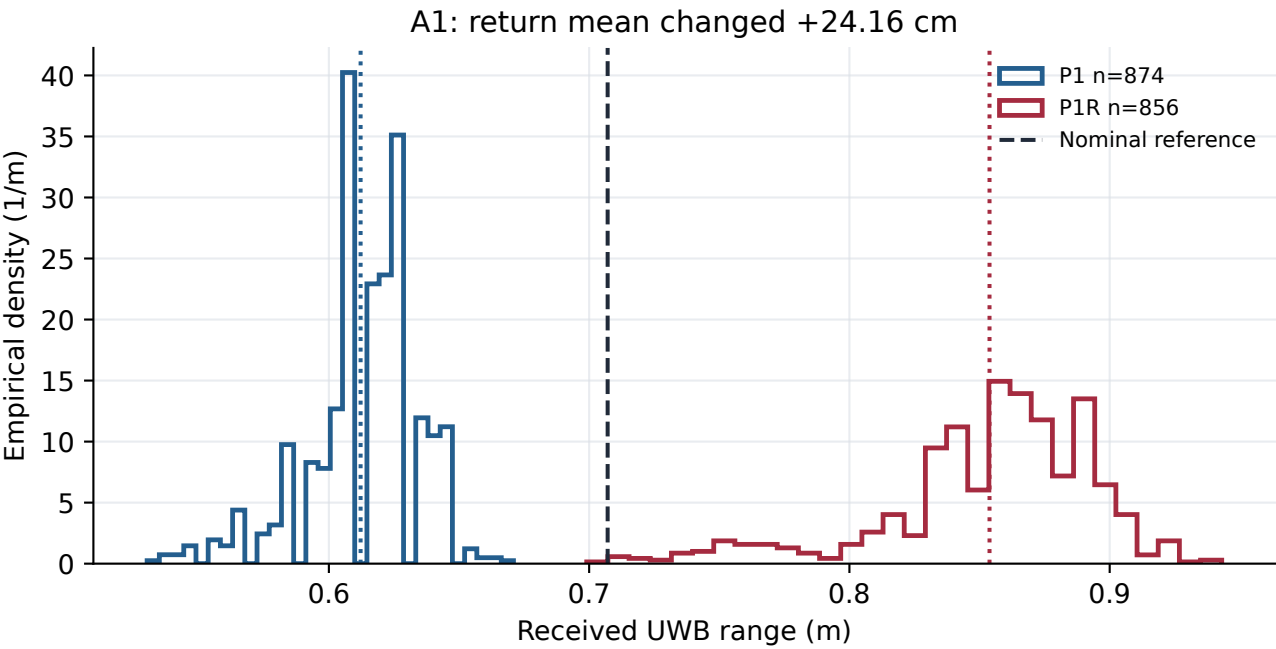
Each range is centered by its own point-anchor capture median. All finite RX power and eligible samples are retained.



Interpretation. Reported RX power is a diagnostic covariate, not an independent distance truth or an NLOS classifier. Centering removes capture offsets but does not control all confounding. Correlations do not establish a cause.

16 Same nominal P1, different returned readings

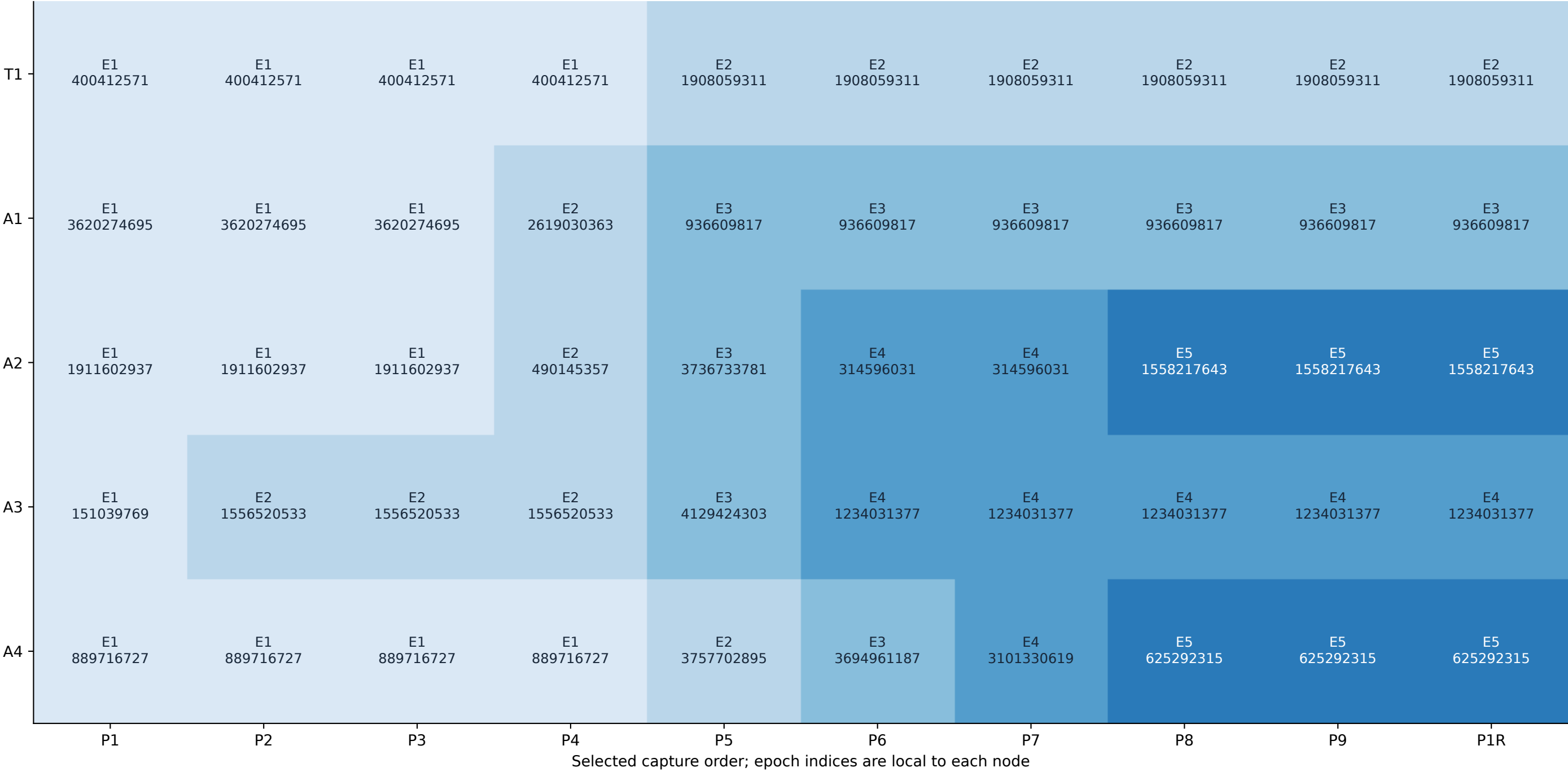
T1 and every anchor have different boot epochs between P1 and P1R; reset, power and board-fall events occurred between captures.



Interpretation. The comparison demonstrates a difference under changed conditions. It cannot isolate time drift, reset sensitivity, handling, antenna orientation or geometry as the cause. There is one return comparison, not a reproducibility study.

17 Boot epochs across the selected measurements

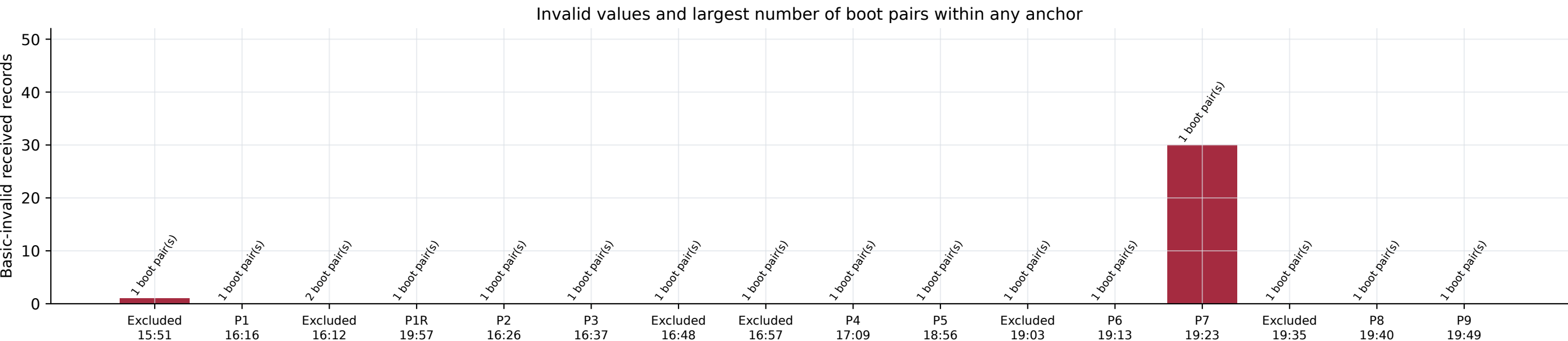
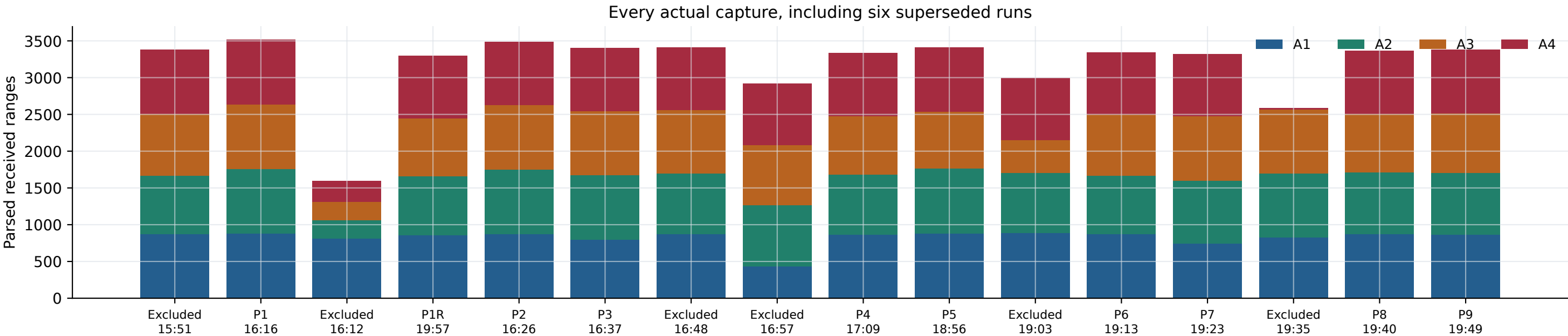
Each cell gives a local epoch index and the full boot identifier. Equal local indices across different nodes do not imply a shared boot.



Interpretation. No selected capture combines multiple boot pairs for an anchor. Boot changes between captures are preserved. An observed boot change without an operator event does not establish why a node restarted.

18 Problematic captures remain visible in the evidence

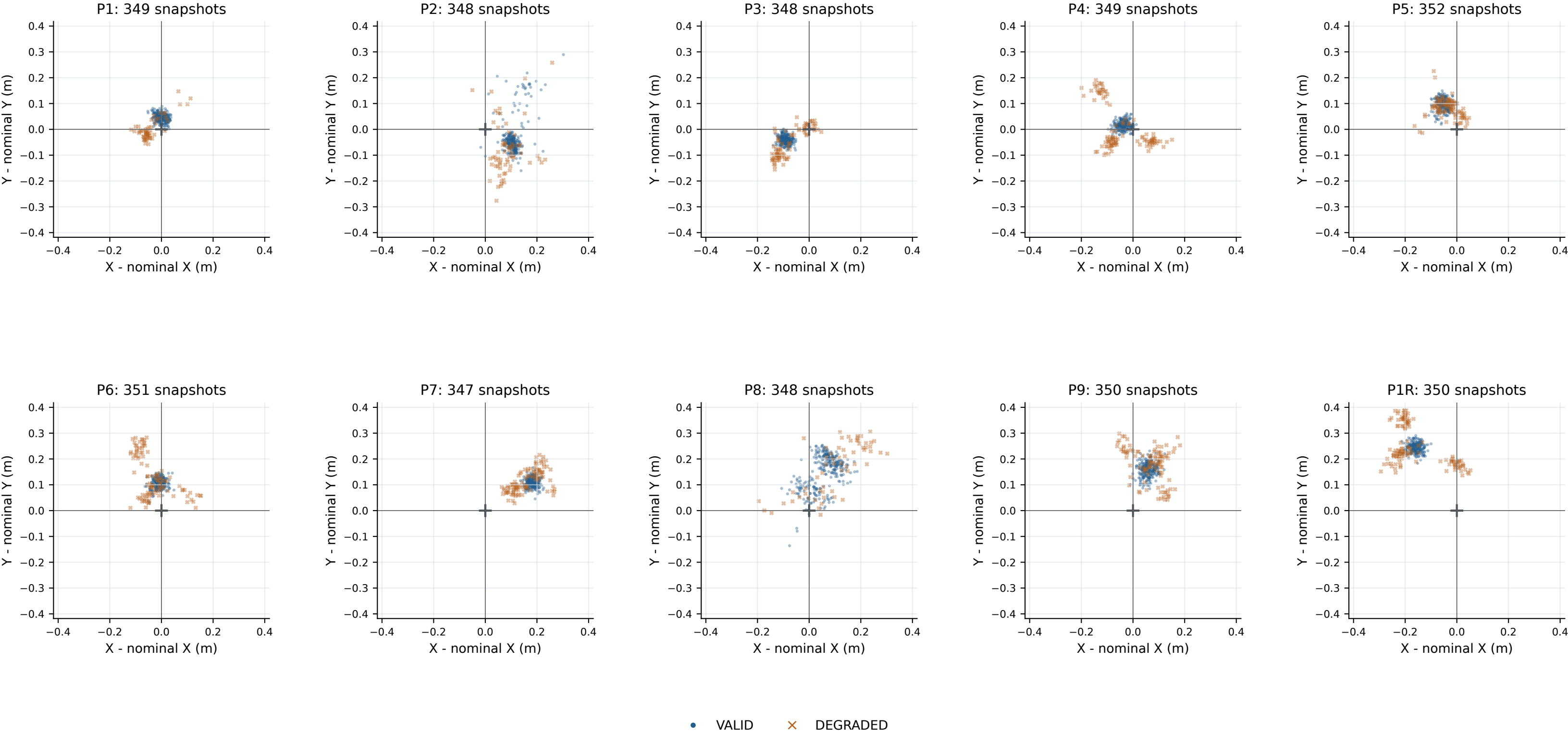
Total: 50,748 ranges. One additional preflight attempt had zero samples. Excluded runs are not added to calibration accuracy metrics.



Interpretation. Exclusion reasons include operator-reported link problems, obstruction or anchor movement, terminal report silence and reset recovery. Their raw measurements remain archived. Counts are not accuracy rankings across changed conditions.

19 Local Compute: sampled position offsets

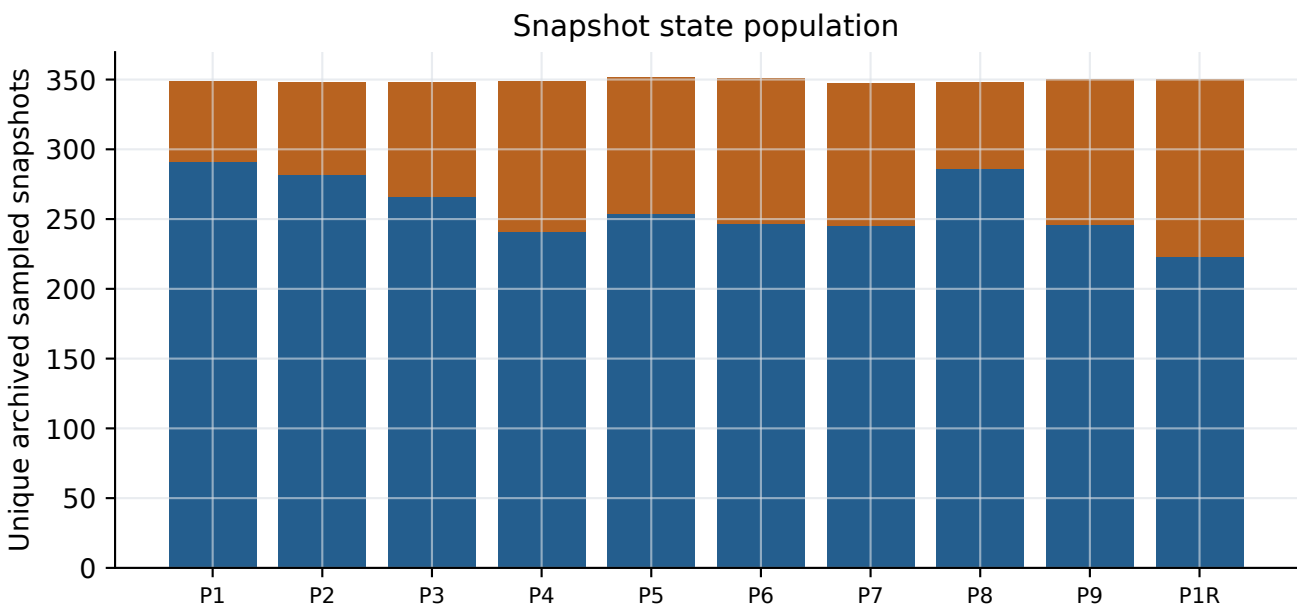
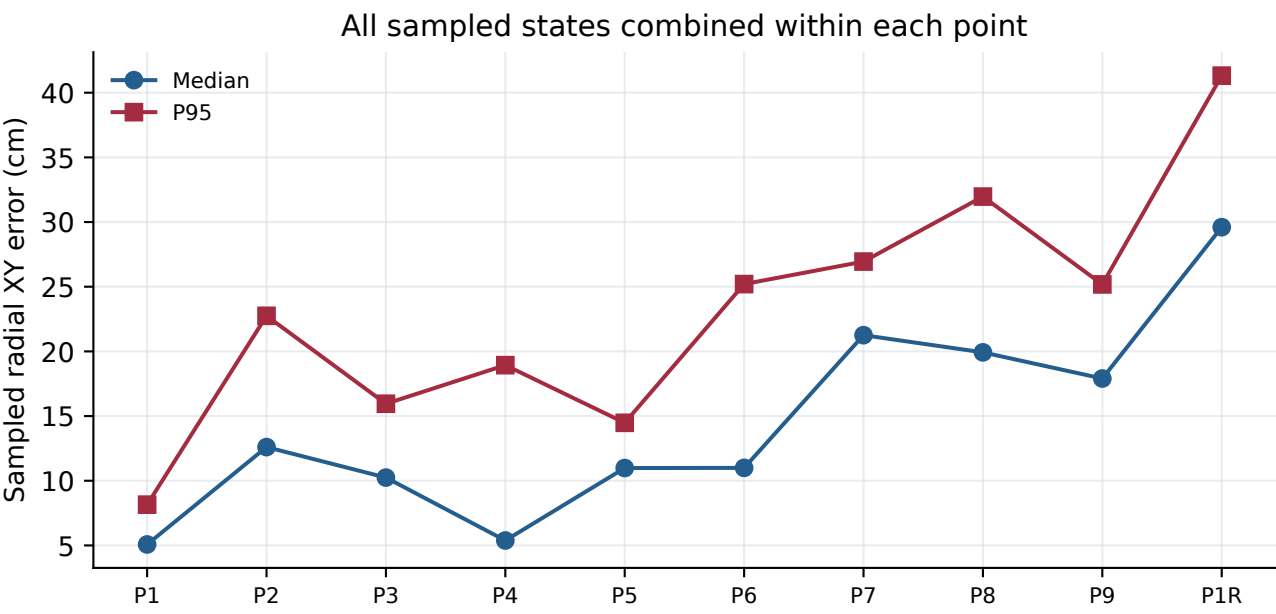
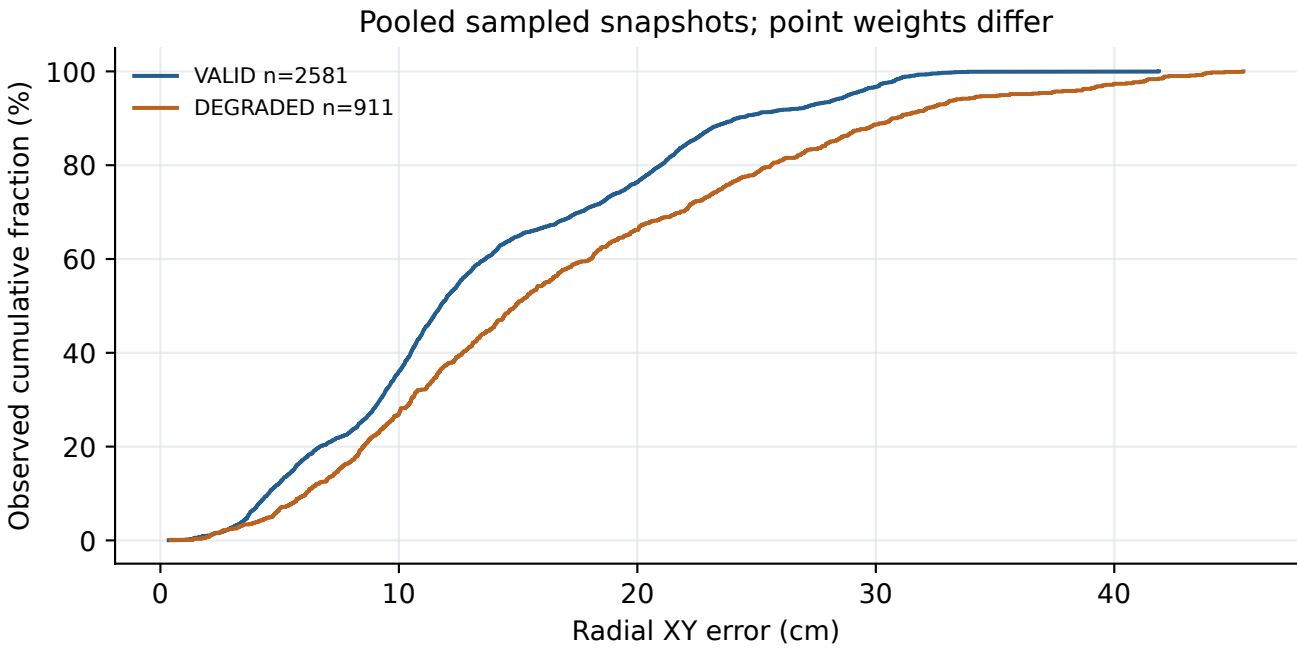
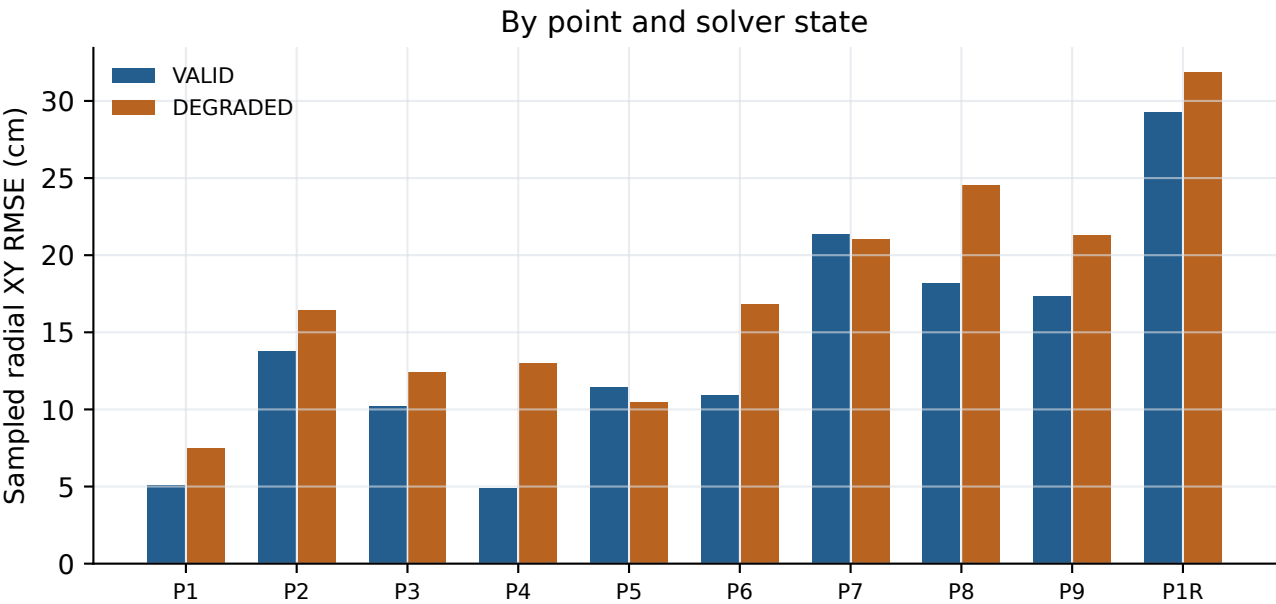
All 3,492 archived unique snapshots; common X/Y limits. These are solver outputs before the offline affine correction.



Interpretation. The plus sign is the operator nominal location. VALID and DEGRADED are solver states, not metrological acceptance. These are sampled API snapshots, not an exhaustive sequence of all computed fixes or an availability statistic.

20 Position estimates and range measurements have different errors

Radial error = $\sqrt{(X-X_{ref})^2+(Y-Y_{ref})^2}$. No corrected-XY result is claimed; the server was not recalibrated.



Interpretation. State-conditioned curves can differ because point mix, geometry and available anchors differ. Do not interpret a pooled VALID/DEGRADED difference as a causal benefit of the state label. The offline range correction was not applied to these positions.

Uncertainty budget and follow-up measurements

Detailed graphics do not replace reference metrology or independent validation.

Within-capture noise. Measured SD, MAD, tails, time bins and ACF are available. Serial dependence and the limited 180 s observation window prevent treating all received samples as independent repeats.

Reference and geometry. The uncertainty of reference X,Y,Z, the measurement instrument and antenna phase centers has not been measured. Do not insert zero for these missing quantities or quote a combined expanded uncertainty.

Between-capture changes. P1 return means changed by +24.16, +32.81, -11.14 and +5.68 cm for A1-A4. Both tag and anchor epochs changed; reset, power and board-fall history confound causal attribution. One return pair is not a reproducibility study.

Controlled next test. Survey antenna phase-center geometry, record orientation, power and temperature, randomize repeated placements across multiple controlled epochs, and reserve independent test data. Set range error and availability requirements before measuring.

Static versus dynamic. The tag was stationary at each point. These data do not identify dynamic poles, zeros, bandwidth or a time-varying true-distance transfer function. No NASA qualification or metrological certification is asserted.

Method references (accessed 24 Oct 2026):

NIST: analysis of bias: <https://www.itl.nist.gov/div898/handbook/mpc/section4/mpc45.htm>

NIST: autocorrelation plot: <https://www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm>

NIST: measurement uncertainty: <https://physics.nist.gov/cuu/Uncertainty/basic.html>

NIST: uncertainty budgets: <https://www.itl.nist.gov/div898/handbook/mpc/section5/mpc56.htm>

Source commit: 2a7311ce0ae67b056a75781e4c48b3654f976960

Frozen model SHA256: be80f26030710537fdf4e1841accf5480785edf5563c10b2420f78b884b10805