Supplementary Material §S2 — Per-station and per-cycle data, reproducibility tables

Companion to JFM_paper_arXiv_v2_5_16 §3.3, §3.4, §4.8 and Appendix C. Updated with ρ_faktor = 0.2843 production data, 2026-04-27.

This supplement provides the individual numerical values underlying the summary statistics reported in the main paper. Each table is independently reproducible from the corresponding archived run directory using the recipe in Appendix C.

S2.1 Production-run configuration matrix

Fixed across all four runs: ρ_faktor = 0.2843 (1D-channel-calibrated localisation parameter, used unchanged in 3D), M = 32 ensemble members, ρ = 1.50 multiplicative inflation, N = 12 LETKF assimilation cycles, mesh 2.56×10^6 cells (ESI OpenFOAM v2412 periodic-hill tutorial).

Run	σ	Seed	RANS baseline	Cycle-12 RMSE	Reduction	Archive
Run D-clean v2	0.05	42	0.35596	0.30086	15.5% (failure)	archive_runD_clean_v50_20260427/
Run F	0.20	42	0.37594	0.06419	82.9%	archive_runF_v50_20260427/
Run E	0.30	42	0.44799	0.04715	89.5%	archive_runE_v50_20260426/
Run E2	0.30	123	0.35254	0.05713	83.8%	archive_runE2_v50_20260426/

Run	Outcome
Run D-clean v2	Spread-damping failure mode (cycle-2 minimum RMSE 0.16, monotone degradation cycles 3–12 to terminal RMSE 0.301)
Run F	Monotone convergence (intermediate σ, σ-regime plateau confirmation)

Run E	Monotone convergence (primary σ = 0.30 result)
Run E2	Monotone convergence (independent-seed replication of Run E)

Note on RANS baseline variability. The four production runs report RANS baseline RMSE values ranging from 0.353 to 0.448 — a 27% spread. This reflects scotch-decomposition non-determinism: in the absence of an explicit scotch-seed lock, each decomposePar invocation produces a different cell-to-rank mapping, propagating to slightly different RANS solutions and thereby to different starting points for the LETKF correction. The same source of variability is responsible for the N = 10 baseline distribution (§S2.5, std 0.045 around mean 0.371). Cross-run determinism would require an additional scotch-seed lock and is identified as a future-work item; bit-for-bit reproducibility within a single run is preserved by the documented seeds for the GRF perturbation and the orchestrator RNG.

All runs use 9 observation stations at $x/h \in \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0\}$ sampled against Krank, Kronbichler & Wall (2018) DNS at $Re_H = 5600$. The localisation radius scales as $\rho(x) = 3\, l_C(x)\, [1 + 0.2843 \cdot \ln(\max(1, Re_l(x)/100))]$; bug-fix commit trail referenced in Appendix C.3.

S2.2 Cycle-12 RMSE per station — Run E (σ = 0.30, seed 42)

Station x/h
0.5
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
Mean

Global ensemble spread at cycle 12: 0.00855 (M = 32 members, streamwise velocity standard deviation across the 9 observation stations). Spread/obs_err ratio = 0.171.

S2.3 Cycle-12 RMSE per station — Run F (σ = 0.20, seed 42)

Station x/h
0.5
1.0
2.0
3.0

4.0

5.0

6.0

7.0

8.0

Mean

Global ensemble spread at cycle 12: 0.00782. Spread/obs_err ratio = 0.156 (used as the plateau midpoint in §3.4 σ-regime analysis).

S2.4 Full RMSE convergence trajectories

Per-cycle global mean RMSE (across the 9 observation stations), all four production runs:

Cycle
0 (RANS baseline)
1
2
3
4
5
6
7
8
9
10
11
12

Run D-clean v2 reaches a cycle-2 minimum RMSE = 0.161 (the only cycle below the N=10 baseline mean) before degrading monotonically through cycles 3–12 to terminal RMSE = 0.301 — the spread-damping failure-mode signature documented in §3.4 of the main paper. Runs F, E, and E2 converge monotonically over all 12 cycles.

Per-cycle global ensemble spread (M = 32 members):

Cycle
1
12

Spread is approximately stable across all 12 cycles for every run, consistent with bounded ensemble dynamics under multiplicative inflation ρ = 1.50. The σ = 0.05 run does not show a *spread* collapse below the σ = 0.30 runs — the failure mode is in the *RMSE* trajectory, not the spread itself.

S2.5 N = 10 baseline variance (σ = 0, ten independent seeds)

The intrinsic variance of the k–ω SST baseline RANS solution under seed perturbation:

Run #
1
2
3
4
5
6
7
8
9
10
Mean
Sample standard deviation
Coefficient of variation
Min / max

Run E at cycle-12 RMSE 0.04715 lies (0.37149 − 0.04715) / 0.04501 = **7.21 standard deviations below the baseline mean**.

This baseline variance reflects the same source of variability as the run-to-run RMSE difference between σ = 0.30 replications (Run E vs Run E2): scotch-decomposition non-determinism produces different cell-to-rank mappings across runs, which propagates to slightly different RANS solutions and thereby to slightly different LETKF-corrected RMSE.

S2.6 σ-regime falsification of linear damping

The four production runs at σ ∈ {0.05, 0.20, 0.30, 0.30} provide a direct test of the linear-scaling spread-damping prediction. Under a linear model where ensemble spread scales linearly with σ, anchored at the Run E measurement (σ = 0.30), the predicted spread/obs_err at σ = 0.20 would be (0.20/0.30) × 0.171 = 0.114. The measured value at Run F (σ = 0.20) is 0.156 — 37% higher than the linear prediction.

A power-law fit log(spread) = α · log(σ) + β over the three measured σ-points (Run D-clean v2, Run F, mean of Run E and Run E2) yields:

σ
0.05

0.20

0.30

0.30

0.30

α (power-law exponent)

0.152

The measured exponent $\alpha = 0.15$ is markedly sub-linear (a linear law would have $\alpha = 1$). This is consistent with a plateau-shaped operating window (over the validated $\sigma \geq 0.20$ regime, varying σ by 50% changes spread by less than 12%) and inconsistent with a knife-edge transition.

S2.7 Spread-error calibration (Task #21, Run E cycle 12)

Across the 9 observation stations at cycle 12 of Run E, the global ensemble spread (0.00855 m/s) and the global mean per-station RMSE (0.04715 m/s) yield an aggregate spread/RMSE ratio of 0.181. Under the per-station spread-error calibration of Task #21 (computed using full per-station ensemble standard deviations), the mean spread/RMSE ratio is approximately 0.18, with worst-case station ~0.12 and best-case ~0.36. All stations report ratio < 1, indicating severe under-dispersion, with monotonic improvement of the ratio across cycles 1 → 12.

The §5 UQ claim that ensemble spread provides a calibrated per-cell uncertainty estimate is therefore quantitative but bounded: the framework underestimates posterior variance by approximately a factor of three at the worst-resolved station and a factor of five at the best, requiring inflation correction or post-hoc rescaling for engineering use.

The Run E2 spread-error calibration ratio is 0.173 (computed locally from archive_runE2_v50_20260426/results/summary.json), within 1% of Run E's 0.171 spread/obs_err, confirming seed-robustness of the under-dispersion signature.

Status of S2 at v2.5.16 submission

Section	Status
S2.1 Run config matrix (incl. RANS-baseline-variability note)	✓ complete
S2.2 Run E per-station	✓ all 9 stations populated from archive_runE_v50_20260426/
S2.3 Run F per-station	✓ all 9 stations populated from archive_runF_v50_20260427/
S2.4 Convergence trajectories	✓ all 4 runs × 12 cycles populated
S2.5 N = 10 baseline	✓ all 10 individual values + scotch-non-determinism connection

S2.6 σ-regime falsification	✓ updated with $\sigma^{0.152}$ power-law (corrected from $\sigma^{0.140}$ in v2.4 series)
S2.7 Spread-error calibration	✓ Run E aggregate populated; Run E2 ratio confirmed 0.173

Bottom line: All headline numbers locked from v50 production runs ($\rho_{faktor} = 0.2843$ throughout). The σ-sweep falsification narrative (sub-linear scaling, plateau operating window, $\sigma_{critical} \in (0.05, 0.20]$) is reproduced with the 1D-calibrated parameter, validating the "single calibrated parameter held fixed across all cases" claim of the abstract.

End of S2 supplementary material.