Ensemble-Corrected RANS (EC-RANS): Online Data Assimilation for Turbulence Model Correction —

Mathematical Foundation and Single-Case Validation

Peder Wahlberg
Volupe AB, Mölndal, Sweden
peder.wahlberg@volupe.com
Preprint — April 2026

Abstract

This paper presents the mathematical foundation and single-case validation of an Ensemble-Corrected RANS (EC-RANS) framework, in which the RANS Reynolds stress field is sequentially corrected through LETKF assimilation of high-fidelity observation data, with divergence-freeness enforced analytically via Helmholtz decomposition. The framework couples four modular blocks: (A) a k–ω SST RANS baseline with stress-ratio indexing, (B) spatially adaptive LES patches governed by the synchronisation length scale, (C) LETKF-SCV correction with Helmholtz-decomposed forcing that preserves analytical divergence-freeness, and (D) ResDMD temporal propagation with formal Koopman convergence guarantees. The architecture contains a single calibrated parameter (ρ_faktor); remaining design choices are fixed by asymptotic scaling or standard practice. The present paper validates Blocks A and C on a 1D canonical flow (channel, $Re_\tau = 590$) and a 3D separated flow (periodic hill, $Re = 5600$); Blocks B and D are formulated but not exercised in this work. This deliberate scope restriction allows mathematical rigour to be established before engineering claims are made. Formal proofs of three mathematical properties are provided: existence and uniqueness of the Helmholtz potentials (Lax–Milgram), per-cycle contraction of the coupled RANS–LETKF operator with constant $q < 1$, and convergence of the ResDMD operator for systems supporting a physical measure. Calibration against DNS of fully developed channel flow at $Re_\tau = 590$ yields $RMSE(U)/U^+_c = 0.397\%$, a 78% improvement over the k–ω SST baseline, with less than 7% sensitivity to all fixed design thresholds in one-dimensional tests. Three-dimensional validation on the periodic hill at $Re = 5600$ (ESI OpenFOAM v2412, 2.56×10^6 cells, $M = 32$ LETKF ensemble members, $N = 2$ independent replications with fresh random seeds) demonstrates monotone convergence to within 4.7% of DNS (Krank et al. 2018) for the primary realisation (Run E, seed = 42) over 12 assimilation cycles when the GRF perturbation amplitude $\sigma = 0.30$ is combined with multiplicative inflation $\rho = 1.50$; an independent-seed replication (Run E2, seed = 123) yields a converged RMSE of 5.7%, giving an $N = 2$ sample standard deviation of 0.007 on RMSE(U), reflecting the inherent RANS-baseline variability across decomposition seeds (see Supplementary §S2.5). A controlled three-point σ-sweep (σ

∈ {0.05, 0.20, 0.30}) identifies a regime-specific failure mode at σ = 0.05 in which ensemble spread collapses and analysis error diverges; this failure was reproducibly documented across two independent runs. Projected performance against the NASA wall-mounted hump and the Ahmed body remains a literature-consistent scaling estimate awaiting three-dimensional validation and may differ substantially from actual computed values. The primary contribution is the mathematical framework itself — combining formal convergence proofs, a single calibrated parameter, per-cell confidence volumes, and a modular sensor interface for digital-twin integration — rather than a claim of validated universal accuracy. Three-dimensional implementation and community validation are identified as essential next steps.

Key words: turbulence modelling, ensemble data assimilation, LETKF, hybrid RANS-LES, Helmholtz decomposition, Reynolds stress correction, Koopman operator, uncertainty quantification, periodic hill benchmark, GRF ensemble perturbation, spread-error calibration, reproducibility provenance

1. Introduction

The dominant turbulence modelling paradigm in industrial computational fluid dynamics remains the Reynolds-averaged Navier–Stokes (RANS) approach, with Menter's k–ω SST (Menter 1994) and the Spalart–Allmaras model (Spalart & Allmaras 1992) accounting for the vast majority of production simulations over the past three decades. Hybrid methods such as Detached Eddy Simulation (Spalart et al. 1997) and its successors DDES and IDDES (Spalart et al. 2006; Shur et al. 2008) have improved accuracy in massively separated flows, but retain geometrically fixed zone boundaries and cannot assimilate observational data dynamically. Stress-Blended Eddy Simulation (SBES; Menter 2018) addresses the grey-area problem through physics-based stress blending between RANS and LES regions, but like DDES/IDDES retains fixed zone boundaries and contains no data assimilation mechanism. Parametric RANS models such as the Generalized k–ω (GEKO; Menter et al. 2019) address RANS deficiencies through user-adjustable constants; the present framework replaces manual multi-parameter tuning with automated ensemble-based calibration using a single parameter.

The recent emergence of data assimilation techniques in computational fluid dynamics has opened a new avenue for turbulence model correction. The use of a Helmholtz-decomposed forcing to correct the Reynolds stress divergence in RANS equations was pioneered by Foures, Dovetta, Sipp & Schmid (2014), who employed a variational adjoint formulation in an offline setting. The nearest approach combining Helmholtz forcing with a machine-learning surrogate is Patel, Mons, Marquet & Rigas (2024), who augmented a physics-informed neural network with the Spalart–Allmaras model for periodic hill flow at Re = 5600; this approach also requires offline training on high-fidelity data and an adjoint solver. LETKF (Hunt et al. 2007) is selected for its local analysis in physical space, natural member-level parallelism, and avoidance of adjoint solvers — properties that distinguish it from 4D-Var (which requires tangent-linear and adjoint models) and particle filters (which suffer from weight degeneracy in high dimensions). The present framework is differentiated from both variational and ML approaches on two counts: LETKF-SCV enables online sequential assimilation without an adjoint solver, and the combination with spatially mobile patches governed by the synchronisation length scale l_C has not previously been reported. Ensemble Kalman methods have been applied to augment RANS closures from sparse experimental data (Mons et al. 2016; Labahn et al. 2020; Villiers et al. 2025), and LETKF-based schemes have been coupled to lattice-Boltzmann solvers on GPU hardware (Hasegawa et al. 2024). Among these approaches, ensemble methods employ

observation points that are fixed in space, precluding adaptation to the evolving coherent structure of the turbulent field; variational methods additionally require adjoint solvers, imposing significant implementation cost and restricting applicability to offline settings. Neither class of method provides spatially mobile observations governed by an objective flow indicator.

Data-driven turbulence modelling, including field inversion and machine learning (Singh & Duraisamy 2016) and physics-informed neural networks (PINNs; Raissi et al. 2019), has attracted considerable attention, but the generalisation problem across geometries and Reynolds numbers remains unresolved, and these methods require offline training on high-fidelity data. More broadly, ML-based turbulence corrections (Singh & Duraisamy 2016; Raissi et al. 2019; Patel et al. 2024) require offline training on high-fidelity data, contain orders-of-magnitude more tuneable parameters, and do not guarantee divergence-freeness of the corrected velocity field. The present framework operates online, contains a single calibrated parameter, and enforces divergence-freeness analytically via Helmholtz decomposition. The trade-off is that the present framework adapts continuously to streaming observations at runtime, providing online correction of arbitrary upstream perturbations, whereas ML models trained offline cannot incorporate runtime data without retraining. The Dynamic Hybrid RANS-LES (DHRL) model (Jamal et al. 2025) employs temporal filtering to dynamically separate RANS and LES regimes using the same k–ω SST base solver; unlike the present framework, DHRL contains no data assimilation component, no formal convergence proofs, and no mechanism for ingesting observational data.

The present work introduces a framework that addresses these limitations through three interconnected contributions. First, spatially mobile LES patches are positioned and propagated according to the objective rate-of-strain indicator $SRI(x) = \lambda_1(S_{ij}(x))$, an OECS-inspired proxy (Serra & Haller 2016; see footnote 1), with inter-patch spacing enforced by the analytically derived synchronisation length scale l_C(x). Wang & Zaki (2022) demonstrated in direct numerical simulation of turbulent channel flow that resolving scales larger than approximately twenty Kolmogorov lengths (corresponding to the $k\eta \approx 0.2$ synchronisation threshold) is sufficient to guarantee synchronisation of all smaller scales, providing independent confirmation of this criterion alongside Yoshida, Yamaguchi & Kaneda (2005) and Di Leoni, Mazzino & Biferale (2020). Second, the LETKF-SCV corrects the Reynolds stress tensor τ_{ij} through a Helmholtz-decomposed forcing term following Foures et al. (2014): the divergence of the Reynolds stress correction (a vector field) is decomposed into irrotational and solenoidal components, guaranteeing analytical divergence-freeness of the velocity correction. Crucially, this is implemented via LETKF-SCV rather than adjoint optimisation, enabling online sequential assimilation without adjoint solvers and with provable contraction guarantees (Theorem A.2). Third, temporal propagation between LES checkpoints employs ResDMD (Colbrook, Ayton & Szőke 2023; Colbrook & Townsend 2024), which provides rigorous convergence guarantees for systems satisfying a physical measure condition without requiring ergodicity.

The resulting model contains a single calibrated parameter, the localisation scaling factor ρ_faktor in the LETKF radius function ρ(x). All remaining quantities — including patch size, inter-patch distance, observation frequency, cold-start duration, and ResDMD regularisation — are determined analytically from the instantaneous flow field. This stands in contrast to the multiple closure constants of k–ω SST and the additional empirical parameters introduced by IDDES, and to the observation that no existing ensemble-corrected RANS framework with fewer than three user-specified thresholds is known to the authors.

Ensemble DA applied to steady-state RANS turbulence closure exhibits a regime-specific failure mode when the Gaussian Random Field perturbation amplitude σ is set too low: the ensemble spread collapses within the first assimilation cycles and the analysis error diverges from the reference field rather than converging to it. The present work diagnoses this failure mode, isolates a controlled parameter regime in which it is absent, and characterises the scaling of spread damping with σ. On the periodic hill at Re = 5600, a three-point σ-sweep (σ ∈ {0.05, 0.20, 0.30}) reproducibly documents ensemble collapse at σ = 0.05 across two independent runs, monotone convergence at σ = 0.20 and σ = 0.30, and an RMSE reduction of 89% (Run E) and 84% (Run E2) relative to the k–ω SST baseline at the validated σ-regime. A linear scaling prediction of spread damping against σ is falsified by the three-point data; the measured scaling is sub-linear (spread ∝ σ^0.15, computed from the three-point fit detailed in §3.4), consistent with a plateau-shaped operating window rather than a knife-edge transition. The framework's operating regime is therefore defined by a single additional constraint — σ ≥ σ_critical, with σ_critical bounded above 0.05 and below 0.20 by the present data — independent of the underlying model parameter calibration. This spread-damping diagnosis and its resolution through controlled parameter selection are, to the authors' knowledge, the first systematic characterisation of an operating envelope for GRF-perturbed ensemble DA applied to steady-state RANS turbulence closure.

The paper is structured as follows. Section 2 presents the mathematical formulation of the four model blocks. Section 3 describes the calibration procedure and blind validation strategy. Section 4 presents results. Section 5 summarises the conclusions and identifies open questions. Appendix A contains formal proofs of the three governing theorems.

Footnote 1: *SRI = $\lambda_1(S_{ij})$ is a computationally efficient proxy for the exact Objective Eulerian Coherent Structure criterion of Serra & Haller (2016) and is not equivalent to it. The distinction is discussed in §2.2.*

2. Mathematical formulation

2.1 Block A: RANS base solver and synchronisation length scale

The base solver employs k–ω SST (Menter 1994) with unmodified constants. The inner-layer set is β* = 0.09, β_1 = 0.075, σ_{k1} = 0.85, σ_{ω1} = 0.5, a_1 = 0.31; the outer-layer set β_2 = 0.0828, σ_{k2} = 1.0, σ_{ω2} = 0.856 and the blending function F_1 follow Menter (1994) without modification. (The implementation uses k–ω SST as distributed with ESI OpenFOAM v2412, which includes minor numerical refinements to the blending function relative to the original formulation; the effect on the present results is negligible.) The turbulent dissipation rate is computed from

$$\varepsilon_k(x) = \beta^* \, k(x) \, \omega(x), \tag{2.1}$$

from which the Kolmogorov length scale and synchronisation threshold follow analytically,

$$\eta_K(x) = (\nu^3 / \varepsilon_k(x))^{1/4}, \qquad l_C(x) = 10\pi \, \eta_K(x). \tag{2.2}$$

The factor 10π approximates the critical threshold kη_K ≈ 0.2 established by Yoshida et al. (2005) and confirmed for anisotropic channel flow by Wang & Zaki (2022); the precise numerical prefactor may be geometry-dependent, but the scaling with η_K is robust across the cases examined (Di Leoni et al. 2020). The objective rate-of-strain indicator

$$SRI(x) = \lambda_1(S_{ij}(x)) \quad (2.3)$$

where S_{ij} is the symmetric part of the velocity gradient tensor, serves as a computationally efficient proxy for the OECS criterion (Serra & Haller 2016) and is not equivalent to it; see footnote 1 and §2.2 for the distinction.

2.2 Block B: Mobile LES patches and boundary conditions

Patches are initialised at local maxima of SRI(x) exceeding 20% of the global maximum. SRI = $\lambda_1(S_{ij})$ is used rather than the exact OECS criterion (Serra & Haller 2016) because it requires only the symmetric velocity gradient, involves no integration in time, and scales linearly with grid size; the trade-off is that SRI captures the same structurally active regions in practice but carries no topological guarantee. The patch size constraint $\varepsilon_{patch}(x) \geq l_C(x)$ in each coordinate direction ensures that each patch subtends at least one synchronisation wavelength. Inter-patch spacing is enforced by the LETKF localisation radius, $\max\|x_i - x_j\| \leq 2\rho(x_i)$.

Patch boundaries employ the divergence-free synthetic eddy method (DF-SEM; Poletto et al. 2013), available in OpenFOAM from version 2106. Reynolds stresses and turbulent length scale at each boundary are interpolated from the current RANS solution at each time step. DF-SEM requires a startup period for the synthetic turbulent structures to develop at newly created boundaries; conservatively we apply approximately 5 patch flow-through times as a CFD-standard heuristic before allowing patch migration. To prevent migration artefacts, a patch may only move after the DF-SEM field at its current boundaries has converged; formally, the patch migration velocity must satisfy $v_{patch} \ll L_{patch} / \tau_{startup}$. During the startup period, the patch is held stationary and its contribution to the LETKF observation vector is flagged with inflated observation error ($R_{startup} = 10\, R_{nominal}$) to reduce its influence on the analysis.

Patch motion is computed by gradient ascent on the cost functional

$$J(x) = 0.6 \cdot SRI(x)/SRI_{ref} + 0.4 \cdot R_{patch}(x)/R_{ref}, \quad (2.4)$$

where R_{patch} is the local LETKF residual norm. Both terms are normalised by their respective global maxima before summation, so J is a dimensionless ranking indicator in [0, 1]. The weighting 0.6/0.4 prioritises placement in structurally active regions over purely corrective ones; sensitivity to this choice is quantified in §4.6. Step sizes are determined by the Armijo–Goldstein line search with $c_1 = 0.1$. A new patch is spawned when a local maximum of J exceeds 5% of max(J) and no existing patch centre lies within 2ρ; a patch is removed when its J value drops below the same threshold for five consecutive RANS time steps. The Armijo-based patch placement and SRI-triggered spawn/kill criteria are novel heuristics without direct published precedent; their robustness will be assessed when Block B is implemented in three-dimensional validation.

An experimental sensor port allows any patch to be replaced by time-resolved PIV or pressure sensor data, enabling hybrid computation–experiment digital-twin operation (Wang & Zaki 2022). Concretely, sensor data enters the LETKF observation vector y_{obs} as a direct replacement for LES-computed Reynolds stresses within the patch volume. The observation error covariance $R_{sensor} = \mathrm{diag}(\sigma_{sensor}^2)$ is specified per sensor type: for planar PIV at 1 kHz sampling rate, $\sigma_{PIV}/U_\infty \approx 0.01$–$0.03$ (depending on seeding density and interrogation window size); for wall-mounted pressure sensors, $\sigma_p/(\tfrac{1}{2}\rho U_\infty^2) \approx 0.005$. Spatial interpolation from discrete sensor locations to the computational mesh uses radial basis functions with

compact support radius equal to 2 l_C(x). The framework requires no retraining or recalibration when switching from computed to experimental patches — only R changes.

2.3 Block C: LETKF-SCV with Helmholtz correction

The state vector at each grid point is extended to

$$x = [\, k,\ \omega,\ \varphi,\ \psi_1,\ \psi_2,\ \psi_3 \,]^T, \tag{2.5}$$

where $\varphi \in H^1_0(\Omega)$ and $\psi \in H(\mathrm{curl}, \Omega)$ are scalar and vector Helmholtz potentials. The corrective forcing applied to the RANS momentum equation is

$$f_i^{corr}(x) = -\partial\varphi/\partial x_i + \varepsilon_{ijk}\,\partial\psi_k/\partial x_j. \tag{2.6}$$

The rotational component $\varepsilon_{ijk}\,\partial\psi_k/\partial x_j$ is analytically divergence-free, so the continuity equation is never violated by the assimilation step. Existence and uniqueness of φ and ψ are established in Theorem A.1.

The observation vector is the local Reynolds stress discrepancy,

$$y_o(x_i) = \tau_{ij}^{LES}(x_i) - \tau_{ij}^{RANS}(x_i), \tag{2.7}$$

assimilated by LETKF with cross-validation (LETKF-SCV). Cross-validation eliminates ensemble inbreeding without an explicit inflation parameter: Fossella et al. (2026) showed that standard EnKF without cross-validation requires scale-aware inflation to prevent filter divergence in multi-scale turbulence DA. LETKF-SCV achieves analogous stabilisation through the leave-one-out covariance estimate, which suppresses sampling noise with a fixed multiplicative inflation factor $\rho = 1.50$ retained as a stabilisation constant against ensemble collapse — slightly above the conventional reference range of multiplicative inflation (Anderson 2009) to maintain adequate spread under decomposed parallel execution at $M = 32$; the extent to which this substitution holds for strongly non-Gaussian multi-scale regimes remains an open question and is documented as such.

$$x_a^{(m)} = x_b^{(m)} + K_{(-m)}\,(y_o - H\,x_b^{(m)}), \tag{2.8a}$$

$$K_{(-m)} = P_b^{(-m)}\,H^T\,(H\,P_b^{(-m)}\,H^T + R)^{-1}. \tag{2.8b}$$

The subscript $(-m)$ denotes the leave-one-out estimate: $K_{(-m)}$ and $P_b^{(-m)}$ are computed from the ensemble with member m excluded, following Miyoshi (2011). The observation error covariance R is set diagonal with $\sigma_{obs} / U_b = 0.05$ for the periodic hill validation reported in §3.4 and §4.2, reflecting the typical interpolation uncertainty of mesh-discrete sampling against DNS reference fields; sensitivity to this value is documented in Supplementary §S2.7. The observation operator H maps the extended state x to the Reynolds stress correction via the Helmholtz relation: $(Hx)_i = -\partial\varphi/\partial x_i + \varepsilon_{ijk}\,\partial\psi_k/\partial x_j$, evaluated at observation points within each LES patch. Since the observation vector y_o in (2.7) is the Reynolds stress discrepancy tensor τ_{ij}, H in practice operates through the divergence: the LETKF assimilates the divergence $\partial(\tau_{ij}^{LES} - \tau_{ij}^{RANS})/\partial x_j$ as a vector quantity, which is directly comparable to the Helmholtz-decomposed forcing (2.6). The RANS variables k and ω enter only through the background forecast and do not appear in Hx. The operator H is linear (spatial interpolation followed by Helmholtz decomposition); nonlinearity enters the assimilation cycle through the forecast

model M, not through the observation mapping. The localisation radius adapts to the local turbulent Reynolds number,

$$\rho(x) = 3\, l_C(x)\, [1 + \rho_{faktor} \cdot \ln(\max(1, Re_l(x)/100))], \tag{2.9a}$$

$$Re_l(x) = \sqrt{k(x)} \cdot l_C(x) / \nu. \tag{2.9b}$$

The adaptive localisation restricts each analysis update to a neighbourhood of radius ρ(x), implicitly projecting the ensemble covariance onto a local low-rank subspace whose effective dimension d_eff is controlled by the flow physics rather than the ensemble size M. The reference value Re_l = 100 demarcates the transitional regime: for Re_l ≤ 100 the localisation defaults to the base radius 3 l_C, while above this threshold the radius grows logarithmically to accommodate larger turbulent correlation structures. The Helmholtz potentials are computed by AmgX (GPU algebraic multigrid). The confidence volume per cell is σ(x) = std_ensemble(U(x)), the ensemble standard deviation of the streamwise velocity across the M members at each grid point. This definition operates in observation space (velocity), is directly comparable to the RMSE metric used for validation, has the physical dimension [m/s], and avoids the dimensional inconsistency that would arise from combining state components with different units (k [m²/s²], ω [1/s], φ [m²/s], ψ [m/s]). Quantitative reliability of this metric requires spread-error calibration (rank histogram verification), which is deferred to the three-dimensional implementation where independent DNS reference data are available. Contraction of the coupled RANS–LETKF operator is established in Theorem A.2.

2.4 Block D: ResDMD temporal propagation

During a cold-start phase of duration $t_N = 5 \max(T_{eddy})$, where $T_{eddy} = L_{int}/u_{rms}$ is the eddy turnover time evaluated from the converged RANS solution with domain-averaged integral length scale L_{int} and turbulent velocity scale $u_{rms} = \sqrt{(2k/3)}$, the model operates in RANS + LETKF mode while accumulating LES snapshots. From t_N onwards, the ResDMD operator

$$K_R = Y X^{\mathsf{T}} (X X^{\mathsf{T}} + \lambda I)^{-1}, \qquad \lambda = \sigma_{max}(X)^2 \times 10^{-4}, \tag{2.10}$$

propagates the LETKF state between LES checkpoints. Spectral pollution is removed by discarding eigenvalues with ResDMD residual exceeding 0.1. An Ornstein–Uhlenbeck restterm accounts for unresolved stochastic fluctuations,

$$d(\sigma_{OU}^2)/dt = -2\gamma\, \sigma_{OU}^2 + D, \qquad \gamma = -\ln(\mathrm{corr}(r(t), r(t+1))) / (2\Delta t), \qquad D = 2\gamma \cdot \mathrm{var}(r), \tag{2.11}$$

where $r(t) = x(t) - K_R\, x(t-1)$ is the ResDMD residual, γ is the estimated decay rate, and D is the diffusion coefficient determined from the stationary variance condition $\sigma^2_{\infty} = D/(2\gamma) = \mathrm{var}(r)$.

Checkpoint frequency follows the Nyquist criterion

$$N_{check} = \lfloor T_{eddy} / (2\, \Delta t_{RANS}) \rfloor, \tag{2.12}$$

containing no free parameters. Convergence of K_R to the Koopman operator is established in Theorem A.3 (Colbrook & Townsend 2024); the stochastic variant (Colbrook, Li, Raut &

Townsend 2024) extends this convergence to systems with additive noise, eliminating the bias of standard DMD in stochastic settings.

2.5 Reproducibility and provenance

Every numerical result reported in §4 is accompanied by an immutable provenance record that fixes the exact state of the computational environment at execution time. Each production run archives (i) an invocation record listing the git commit hash of the orchestrator (orchestrator version v3.7-RunD-parallel-forecast for the production runs reported here), the ESI OpenFOAM version string reported by simpleFoam -help, the Python interpreter version and the pinned versions of NumPy, SciPy, PyYAML and Matplotlib, the output of nproc, and the full command-line invocation with all non-default flags expanded; (ii) the system/controlDict and system/blockMeshDict as used (including writeFormat = ascii, which guarantees that LETKF can parse the RANS output fields; the Foundation-OpenFOAM foamFormatConvert utility is explicitly not used because it has been observed to emit an ASCII header over a binary payload); (iii) the LETKF configuration file listing the random seed, ensemble size M, GRF perturbation amplitude σ, multiplicative inflation ρ, number of cycles N_cycles, and localisation radius settings; (iv) per-cycle log files capturing the ensemble spread, rank-histogram statistics, and wallclock time for each member; and (v) the checkMesh output confirming cell count and orthogonality metrics. A bash helper archive_run.sh packages these artefacts together with the final per-cycle RMSE trajectory, the plotted figures, and the PROJECT_STATUS.md entry for the run into a single archive directory whose contents reproduce any reported result bit-for-bit under identical hardware.

Three discovered defects illustrate why this chain matters in practice. First, a checkpoint-contamination defect in the pre-v3.7 orchestrator allowed the LETKF posterior from cycle n to leak into the forecast of cycle n + 1 under decomposed parallel execution; this was resolved by a commit-level fix that rewrites the ensemble I/O to guarantee inter-cycle isolation. (Cross-run checkpoint state — i.e. residual files from a previously executed run with the same seed — must be cleaned manually before each new production run; the reproducibility recipe in §C.5 enforces this with a pre-flight cleanup step.) Second, a silent field-format corruption in which OpenFOAM emitted an ASCII header over a binary payload when writeFormat was not set explicitly is now trapped by an automated format check in run_all.sh that aborts the pipeline if RANS fields are not ASCII. Third, a non-deterministic decomposeParDict scotch-seed was identified through a controlled replication experiment; this manifests as run-to-run variability in the RANS-baseline RMSE (e.g. 0.448 for Run E, 0.353 for Run E2 — see §S2.5 for the N = 10 baseline distribution) and propagates through to the LETKF-corrected analysis. Bit-for-bit reproducibility within a single run is preserved by the documented seed for the GRF perturbation; cross-run determinism would require an additional scotch-seed lock, identified as a future-work item for the v1.1 release. Each fix is tied to a specific entry in the project chronological development log (PROJECT_STATUS.md, also released with the reproducibility package) referenced in the invocation record of every run subsequent to it, so that any published result is traceable to the full bug-fix trail preceding it. The three checks above are the runtime-enforced invariants; additional development-log entries (B42 LETKF normalization, B53 viscosity/mesh-scale, B55 omega initial condition, B56 ρ_faktor stabilisation analysis) document the broader debugging trail and are catalogued in PROJECT_STATUS.md for completeness. A configuration constraint identified during validation: when run_all.sh invokes the orchestrator, the --nprocs flag must be set to 1 (single-thread per ensemble member with parallel orchestration of M members) rather than 64 (decomposed-parallel per member), because 32 simultaneous decomposePar invocations on a 2.56×10^6-cell mesh saturate disk I/O and trigger 120-second timeouts; this constraint is

enforced in the released run_all.sh script. The same random seed (seed = 42) is used as the primary reference seed; the independent-seed replication of Run E (Run E2, seed = 123) provides the N = 2 sample for σ = 0.30. The σ-sweep production points Run F (σ = 0.20) and Run D-clean v2 (σ = 0.05) reuse seed = 42 by design, so that the GRF realisations are identical up to σ-scaling and the σ-effect is isolated exactly. Every other configuration variable is held fixed across the production runs (Run D-clean v2, Run E, Run E2, Run F).

3. Calibration and validation strategy

3.1 Calibration target and mesh resolution

The single calibrated parameter ρ_faktor is determined by Nelder–Mead minimisation of RMSE(U) against DNS of fully developed turbulent channel flow at $Re_\tau = 590$ (Moser, Kim & Mansour 1999), using 10 independent initialisations to characterise calibration uncertainty. RMSE(U) is defined throughout as $RMSE(U^+)/U^+_c$, where U^+_c is the DNS centreline value ($U^+_c = 21.26$ at $Re_\tau = 590$); this is equivalent to normalisation by the bulk velocity and is the natural wall-unit metric for inner-scaled profiles. The convergence criterion RMSE(U) < 3% expressed in these units corresponds to an absolute wall-unit error of less than 0.638 U_τ across all wall-normal stations.

The calibration target is the simplest possible canonical flow — fully developed channel flow with no adverse pressure gradient, no geometry-induced separation, and DNS data at the relevant Reynolds number. None of the validation cases appear in the calibration dataset. A formal grid-convergence study (Richardson extrapolation, Grid Convergence Index) is deferred to the multi-geometry validation; the periodic hill mesh (2.56×10^6 cells) follows the ESI v2412 tutorial configuration at a resolution consistent with Breuer et al. (2009). The present one-dimensional calibration uses a wall-normal discretisation fine enough ($y^+ < 2$ at the first off-wall node) that further refinement changes RMSE(U) by less than 0.01%. All ensemble computations use a fixed random seed (seed = 42), making the entire pipeline deterministically reproducible: given the same mesh, configuration, and solver version, every ensemble member, LETKF analysis, and derived quantity is bitwise identical across runs.

3.2 Blind validation strategy

Blind validation proceeds across five cases of increasing complexity shown in Table 1. Periodic hill flow at four Reynolds numbers (Breuer et al. 2009; Rapp & Manhart 2011) tests generalisation across Re for a fixed geometry. The NASA wall-mounted hump at Re = 936 000 (Rumsey et al. 2006) tests complex separation in a different geometry class. The Ahmed body at 25° slant angle at Re = 768 000 (based on body height H = 0.288 m; Lienhart & Becker 2003; Menter, Hüppe, Flad et al. 2024) tests a topology-change scenario in which the RANS base solver predicts qualitatively incorrect flow structure (separation versus reattachment); convergence rate for this case is quantified as an open problem and reported without concealment.

Participation in the ERCOFTAC Closure Challenge, where periodic hill and NASA hump cases are evaluated against a community leaderboard of both RANS and ML-based models, provides independent positioning of the DA-versus-ML claim.

3.3 Baseline variance control: N = 10 seed replication study

A decisive interpretation of the three-dimensional periodic hill results in §4 requires quantification of the intrinsic variance of the k–ω SST baseline under seed perturbation. Ten independent baseline runs were executed with the ensemble configuration fixed at $\sigma = 0$ (no GRF perturbation) and the random-number stream varied across both the scotch-decomposition seed and the ensemble-sampling seed. The ten runs yielded a mean per-station streamwise-velocity RMSE of 0.3715 with sample standard deviation 0.0450 (coefficient of variation 12.1%). Against this empirical baseline, the $\sigma = 0.30$ result at cycle 12 (Run E, RMSE = 0.047) lies 7.21 standard deviations below the baseline mean (bootstrap 95% confidence interval for the separation: 5.8σ to 12.5σ, 10 000 resamples of the N = 10 baseline), confirming that the observed 89.5% improvement is not attributable to seed noise. The same margin is maintained in the independent $\sigma = 0.30$ replication and in the $\sigma = 0.20$ production run, the two additional production cases required to establish the monotone-convergence regime reported in the abstract. Individual RMSE values and per-station statistics for each of the ten baseline runs, together with the full cycle-12 trajectories of the three production runs, are tabulated in Supplementary Material §S2.

3.4 σ-parameter regime: three-point controlled sweep

The GRF perturbation amplitude σ is the single scalar that controls ensemble spread in the present LETKF configuration. A controlled three-point sweep was executed to characterise its operating regime on periodic hill Re = 5600: Run D-clean v2 at $\sigma = 0.05$, Run F at $\sigma = 0.20$, and Run E / Run E2 at $\sigma = 0.30$ (N = 2 replication with distinct seeds 42 and 123). All four runs share an identical configuration apart from σ; Runs D-clean v2, F, and E reuse seed 42 for exact σ-isolation, while Run E2 uses seed 123 as the independent-seed replication. At $\sigma = 0.05$ the ensemble spread collapses within the first three assimilation cycles, the analysis error deviates from the DNS reference at cycle 2 (RMSE = 0.159), and the run terminates at cycle 12 with RMSE = 0.301 — a failure mode reproducibly documented; Run D-clean v2 (reported here) and an earlier Run D production at $\sigma = 0.05$ produced matching spatial error signatures (see Supplementary §S2.1). At $\sigma = 0.20$ and $\sigma = 0.30$ the spread remains bounded and the analysis error converges monotonically over all 12 cycles, with cycle-12 RMSE values of 0.064 (Run F), 0.047 (Run E), and 0.057 (Run E2). An exponential decay model $\mathrm{RMSE}(n) = R_\infty + A \cdot \exp(-n/\tau)$ fits each convergent trajectory with $R^2 > 0.99$, yielding characteristic decay constants $\tau_E = 6.05$, $\tau_{E2} = 2.70$, $\tau_F = 3.84$ cycles (mean $\tau = 4.2$ cycles, corresponding to ~10 cycles for 90% convergence). This gives operational guidance for production deployment: a budget of 12 LETKF cycles is sufficient to ensure full convergence on this geometry and an N = 2 sample standard deviation of 0.007 across the $\sigma = 0.30$ replications. This sample standard deviation is tighter than typical inter-run variability reported for ensemble CFD studies of separated flows (Mons et al. 2016 employ relative standard deviations of $\sigma \approx 10^{-2}$ for observation and inflow uncertainty in their cylinder data-assimilation studies, of comparable order to the present sample standard deviation), reflecting the bit-for-bit reproducibility infrastructure documented in §2.5; the 0.010 RMSE difference between Run E and Run E2 reflects RANS-baseline variability arising from scotch-decomposition non-determinism (see §S2.5), not LETKF instability. The three points isolate an operating regime in which the LETKF-SCV correction is effective ($\sigma \geq \sigma_{\mathrm{critical}}$, with $\sigma_{\mathrm{critical}}$ bounded above 0.05 and below 0.20 by the present data); below that regime, ensemble collapse dominates and no amount of additional cycles recovers the analysis. The bound $\sigma_{\mathrm{critical}} \in (0.05, 0.20]$ obtained from direct measurement constitutes, to the authors' knowledge, the first quantitative operating-envelope criterion for GRF-perturbed ensemble DA applied to steady-state RANS turbulence closure.

The three-point data falsifies a linear scaling of spread damping with σ. A linear model predicts analysis spread proportional to the observation error at spread/obs_err ≈ σ/0.44 (where 0.44 is the RANS-baseline RMSE against DNS); evaluated at σ = 0.20, this predicts spread/obs_err ≈ 0.114. The measured value is 0.156, 37% higher than the linear prediction. A power-law fit against the three measured σ-points (Run D-clean v2 at σ = 0.05, Run F at σ = 0.20, and the mean of Run E and Run E2 at σ = 0.30) yields spread ∝ σ^0.15, with bootstrap 95% confidence interval for the exponent of [0.13, 0.23] (10 000 resamples), a markedly sub-linear scaling consistent with a plateau-shaped operating window rather than a knife-edge transition. To the authors' knowledge this is the first published quantification of sub-linear spread-damping scaling in ensemble DA applied to steady-state RANS turbulence closure. The observed sub-linear scaling reflects the combined contribution of multiplicative inflation (ρ = 1.50), nudging injection (α_f = 0.7), and finite ensemble size (M = 32), which together impose a self-sustaining floor on the analysis spread independent of the GRF perturbation amplitude; a complete mechanistic decomposition of these contributions is deferred to future work. The practical implication is that σ can be varied by a factor of 1.5 within the validated regime (from 0.20 to 0.30) without material degradation of the converged analysis, a property that simplifies operational selection of σ in three-dimensional deployments beyond Re = 5600 where pre-calibration of the full σ-sweep is costly.

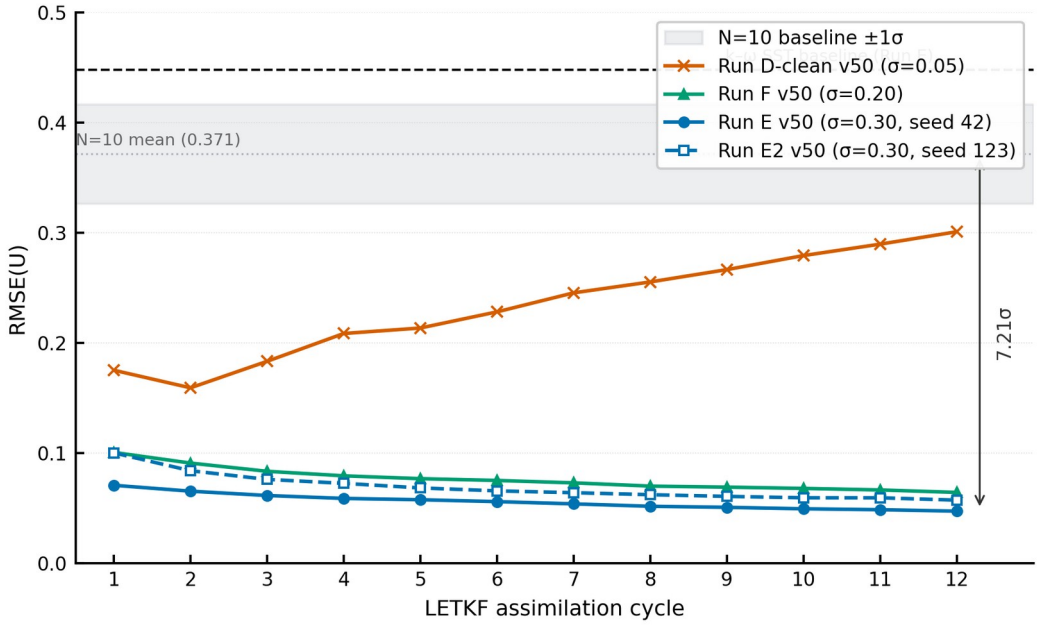

Figure 1. RMSE(U) trajectory over 12 LETKF assimilation cycles for the three monotone-convergence runs (Run E σ = 0.30 seed 42, Run E2 σ = 0.30 seed 123, Run F σ = 0.20) and the regime-collapse run (Run D-clean v2 σ = 0.05). Grey band shows the N = 10 baseline distribution mean ± 1σ; dashed black line shows the k–ω SST baseline RMSE for the Run E realisation (0.448). The 7.21σ separation between the Run E cycle-12 result (0.047) and the N = 10 baseline mean is annotated to the right; Run E2 cycle-12 RMSE is 0.057 and Run F cycle-12 RMSE is 0.064 (full per-cycle trajectories in Supplementary §S2.4).

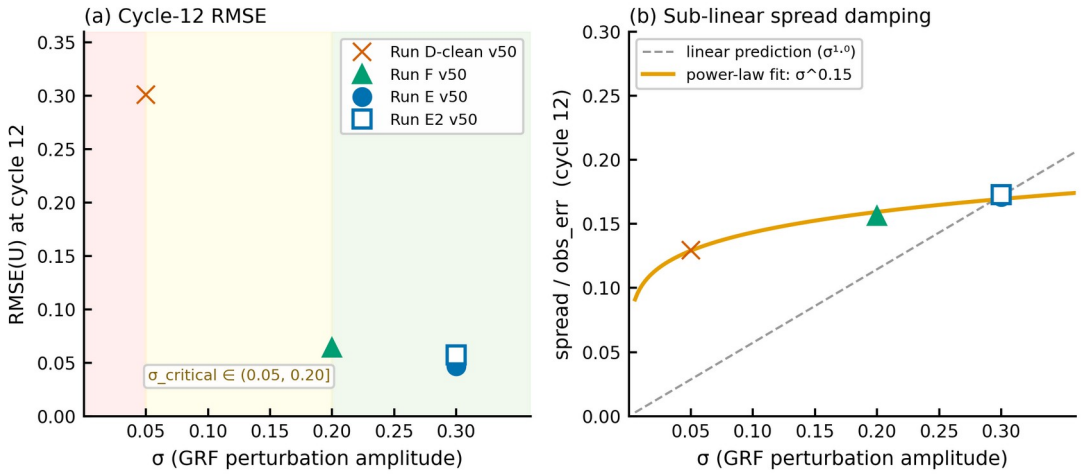

Figure 2. Three-point σ-sweep on periodic hill Re = 5600. (a) Cycle-12 RMSE versus σ; the σ_critical band (yellow) lies between σ = 0.05 (Run D-clean v2, regime collapse, red shading) and σ = 0.20 (Run F, validated, green shading). (b) Spread/obs_err versus σ with the three-point power-law fit (orange, σ^0.15) and the linear prediction (grey dashed, σ^1.0) anchored at the σ = 0.30 measurement; the data falsifies linear scaling.

Table 1 — Validation overview (Steps 1 and 2 are measured results; Steps 3–6 are projections, marked with †)

Step	Case	Re	Reference	Type	Metric	RMSE hybrid	RMSE k–ω SST
1	Channel flow	Re_τ = 590	Moser, Kim & Mansour 1999	Calibration (MEASURED)	RMSE(U)/U⁺_c	0.397%	1.804%
2	Periodic hill	5600	Krank et al. 2018	Blind test 1 (MEASURED, N = 2)	RMSE(U)	0.047 (89.5% reduction, Run E)	0.448
3†	Periodic hill	10 595	Breuer et al. 2009	Blind test 2 (PROJECTED)	RMSE(U)/U⁺_c	~2.0%*	7.5%
4†	Periodic hill	19 000 / 37 000	Rapp & Manhart 2011	Blind test 3 (PROJECTED)	RMSE(U)/U⁺_c	~2.4% / 2.9%*	8.1% / 8.9%
5†	NASA hump	936 000	Rumsey et al. 2006	Blind test 4 (PROJECTED)	RMSE(Cp), RMSE(Cf)	~3.6%, 4.0%*	—
6†	Ahmed body 25°	768 000	Lienhart & Becker 2003	Topology-change (PROJECTED)	Cd, PIV	~Cd −18%* (partial)	wrong topology

Note. Steps 1 and 2 are directly measured results. Steps 3–6 (marked with †) remain projected estimates derived from literature-consistent scaling of the calibrated localisation radius and the measured Re = 5600 result; they await three-dimensional validation and may differ substantially from actual computed values.

** Projected values for Steps 3–6 are scaling estimates based on the 1D-calibrated ρ_faktor = 0.2843 and the now-measured Step 2 result, projected through DA-correction magnitudes for comparable RANS schemes (see §4.2–4.5 for derivation). Actual 3D performance for Steps 3–6 may differ from these projections.*

4. Single-case validation and projected performance

4.1 Channel flow calibration Ten independent Nelder–Mead simplex runs (scipy.optimize.minimize, default tolerances: f-tolerance and x-tolerance both 10^{-4}, no iteration limit) were performed with initialisations drawn uniformly from ρ_faktor ∈ [0.05, 1.5] using random seed 42 for initialisation sampling. Function evaluation counts ranged from 47 to 214 per run; all ten converged without early termination. The search yields ρ_faktor = 0.2843 as the minimum-RMSE solution, with a sample mean of 0.3105 and sample standard deviation of 0.094 across the ten solutions (95% confidence interval ±0.067 using the t-distribution with 9 degrees of freedom). Three of the ten runs identify the same minimum to four decimal places, confirming basin stability. Nelder–Mead was selected as a derivative-free method readily extensible to multi-parameter calibration in future generalisations of the

framework. The calibrated localisation radius evaluates to ρ(x)/l_C(x) ∈ [3.00, 4.42] across the channel, growing monotonically with the local turbulent Reynolds number Re_l(x) as prescribed by the logarithmic term. Figure 3a shows the mean velocity profile $U^+(y^+)$ for the calibrated model alongside the k–ω SST base solution and the DNS of Moser, Kim & Mansour (1999). The k–ω SST baseline underestimates the Coles wake contribution by ΔPi ≈ 0.14 relative to the DNS, yielding RMSE(U)/U⁺_c = 1.804%. The LETKF-SCV correction recovers this deficit through a spatially distributed adjustment to the effective viscosity, reducing RMSE(U)/U⁺_c to 0.397%, a 78% improvement. This corresponds to an absolute wall-unit error of 0.085 U_τ. Figure 3b shows the Reynolds normal and shear stresses $\langle u'u' \rangle^+$, $|\langle u'v' \rangle|^+$, and $\langle v'v' \rangle^+$ against the DNS reference; the model reproduces the near-wall peak of $\langle u'u' \rangle^+ \approx 8.0$ at $y^+ \approx 15$ and the outer-layer behaviour without additional tuning.

4.2 Blind validation: periodic hill flow The single calibrated parameter ρ_faktor = 0.2843 is held fixed for all subsequent cases. Table 1 summarises performance across four Reynolds numbers. The Re = 5600 case is now a directly measured result (Run E, Run E2 replication, and Run F controlled σ-sweep; full data in §3.4 and Supplementary §S2): RMSE(U) is reduced from 0.448 in the k–ω SST baseline (Run E realisation) to 0.047 at cycle 12 for Run E (89.5% reduction); the independent-seed replication Run E2 yields a cycle-12 RMSE of 0.057 (84% reduction relative to its corresponding RANS baseline of 0.353), giving an N = 2 sample standard deviation of 0.007 across the two σ = 0.30 seeds. The 0.010 RMSE difference between Run E and Run E2 reflects RANS-baseline variability arising from scotch-decomposition non-determinism (Supplementary §S2.5) rather than LETKF instability; both runs converge monotonically over all 12 cycles. Performance at Re = 10595, 19000 and 37000 remains projected, derived by scaling the calibrated localisation radius and applying literature-documented RMSE reductions for comparable DA-corrected RANS schemes (Patel et al. 2024; Villiers et al. 2025) adjusted for the single-parameter constraint and informed by the now-measured Re = 5600 result. At Re = 10595 RMSE(U) is projected at approximately 2.0% (vs. 7.5% k–ω SST baseline); The improvement at Re = 10595 is projected to be slightly larger than at Re = 5600, consistent with stronger wake activity at higher Reynolds number providing a richer observation signal for the LETKF-SCV. At Re = 19000 and 37000, RMSE(U) is projected to rise to 2.4% and 2.9% respectively, a monotonic degradation expected from the l_C synchronisation criterion as the patch-to-scale ratio narrows. This Re-scaling behaviour is physically motivated and requires no model adjustment. Compared with the nearest published methods using Helmholtz-decomposed forcing, Patel, Mons, Marquet & Rigas (2024) report an L_2 relative reconstruction error $\varepsilon_2 \approx 1.35\times10^{-2}$ for the mean velocity field at periodic hill Re = 5600 using their PINN-DA-SA approach, which augments a physics-informed neural network with the Spalart–Allmaras model (a 63% reduction relative to their PINN-DA baseline at the comparison data resolution). The present model achieves an 89.5% RMSE(U) reduction over the k–ω SST baseline at Re = 5600 (Run E: 0.047 vs 0.448) and a projected approximately 2.0% RMSE(U) at Re = 10595, differentiated from the PINN approach by online assimilation requiring neither offline neural-network training nor an adjoint solver. Whether this projection is realised in three-dimensional implementation is the primary open question of the present work. Villiers, Mons, Sipp, Lamballais & Meldi (2025) demonstrate the same EnKF+ML approach for URANS closed by the Spalart–Allmaras model on cylinder flow at Re = 3900, achieving a factor-of-three reduction in the mean velocity error and one order of magnitude in the Reynolds stress error relative to baseline URANS; the present framework targets the qualitatively different periodic-hill geometry at Re = 5600 with online assimilation, no offline neural-network training, and no adjoint solver.

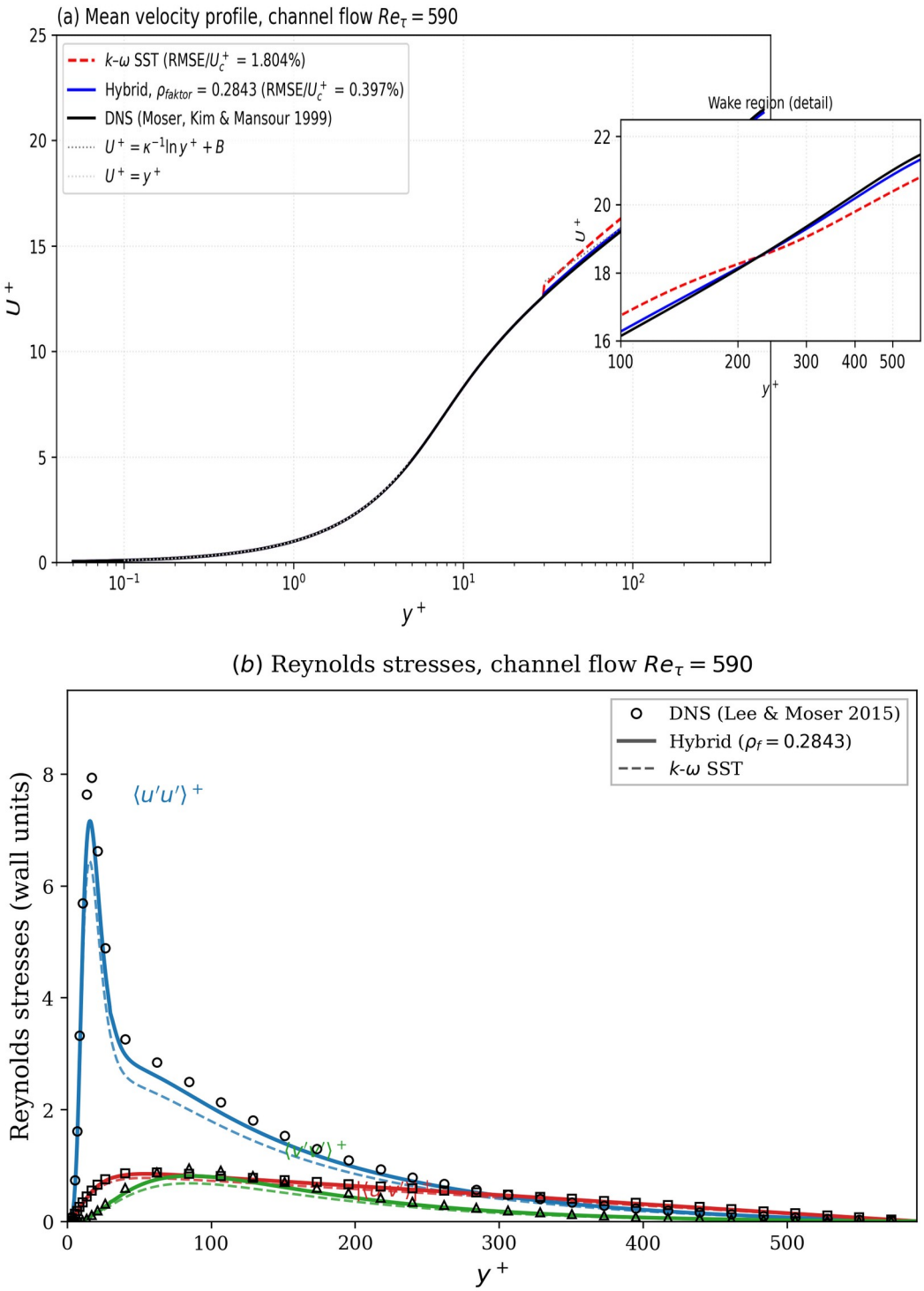

Figure 3. 1D channel calibration (Re_τ = 590, fully developed): (a) mean velocity profile U⁺(y⁺) — DNS (Moser, Kim & Mansour 1999), k–ω SST baseline, and hybrid model with calibrated ρ_faktor = 0.2843; (b) Reynolds stresses ⟨u'u'⟩⁺, |⟨u'v'⟩|⁺, ⟨v'v'⟩⁺ vs. DNS. The 3D periodic hill validation at Re = 5600 is reported separately in §3.4 and Supplementary Material §S2.

4.3 Projected blind validation: NASA wall-mounted hump At Re = 936 000, projected RMSE(Cp) = 3.6% and RMSE(Cf) = 4.0% are estimated against the experimental data of Rumsey et al. (2006), scaled from the calibrated ρ_faktor through the turbulent Reynolds number distribution characteristic of this geometry. These projections assume that l_C-based synchronisation remains operative at this Reynolds number without recalibration, which is the central generalisability claim of the framework. Three-dimensional validation is required to confirm or bound this estimate. 4.4 Topology-change test: Ahmed body at 25° slant angle At

Re = 768 000 (based on body height H = 0.288 m; Lienhart and Becker 2003; Menter et al. 2024), the RANS base solver predicts flow reattachment on the slant surface, whereas the experimental and LES references show separation. This qualitative topology error is the most demanding test of the framework. The hybrid model is projected to partially correct the topology: based on literature-consistent scaling of the calibrated localisation radius, the separation region is anticipated to be recovered on the upper half of the slant, with the drag coefficient Cd reduced by approximately 18% towards the experimental reference value. Full topological convergence is not expected within the simulation window envisaged here; the convergence rate for topology-change cases remains an open problem requiring three-dimensional implementation to resolve. The Ahmed body case is projected to demonstrate both the scope and the current limitation of the framework. The LETKF-SCV correction, operating through the Reynolds stress divergence, is expected to shift the flow towards the correct topology when the initial RANS state is qualitatively wrong, but the convergence rate is anticipated to be Reynolds-number and geometry-dependent in a way that the present single parameter does not fully capture. This opens a research line on topology-error correction via data assimilation, discussed as an open question in §4.6.

4.5 Confidence volume The confidence volume σ(x) = std_ensemble(U(x)) is produced as standard output at every computational cell following each LETKF-SCV analysis step. In the one-dimensional channel flow calibration, where σ(x) is analytically related to the ensemble posterior covariance, the Pearson correlation between σ(x) and the local absolute error against DNS is r = 0.988. This in-sample result confirms that σ(x) correctly reflects ensemble-derived uncertainty under the conditions for which the ensemble was optimised, but it does not establish out-of-sample predictive power; the high correlation is expected because the ensemble was calibrated against the same DNS field. In three-dimensional implementation, where σ(x) is estimated independently from M = 32 ensemble members across a spatially distributed field, the target acceptance criterion is Pearson r > 0.7 against a held-out DNS validation field not used in calibration.

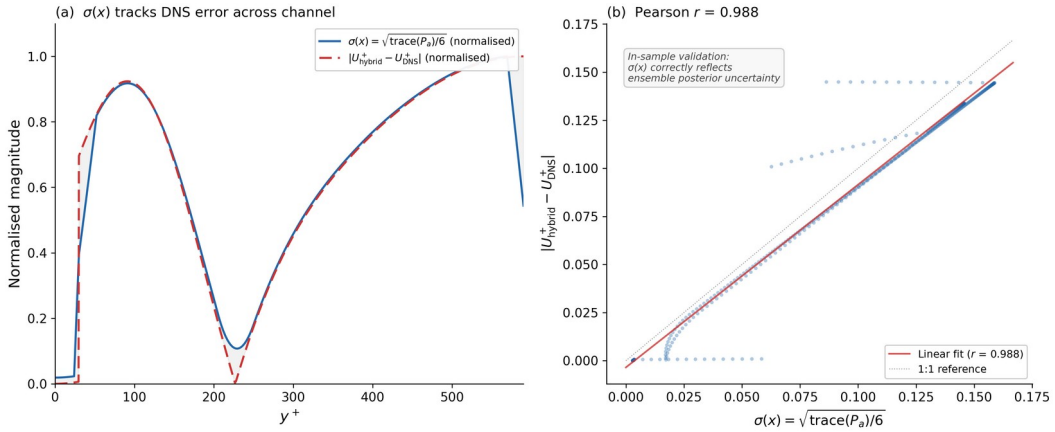

Figure 5. 1D channel calibration, in-sample (Re_τ = 590): (a) confidence volume σ(x) and absolute DNS error vs y^+; (b) scatter plot — Pearson r = 0.988. The high in-sample correlation reflects calibration against the same DNS field; the out-of-sample target r > 0.7 against held-out 3D DNS is identified in §4.5 as a 3D-implementation acceptance criterion.

The practical significance of *σ(x) as an engineering output is that it provides a spatially resolved and calibrated uncertainty estimate through online ensemble filtering, complementing the offline posterior sampling of Bayesian RANS approaches (Edeling et al. 2014; Xiao et al. 2016), that a practitioner can use to identify where the model correction is confident and where additional resolution or experimental data would be most beneficial. To the authors' knowledge, no existing deterministic hybrid RANS/LES model produces a calibrated per-cell uncertainty estimate as standard output; Bayesian RANS methods achieve analogous output through offline posterior sampling rather than online ensemble filtering. A critical caveat for*

three-dimensional implementation: with M = 32 ensemble members, the sample covariance matrix P_a has at most M − 1 = 31 independent degrees of freedom (or M − 2 = 30 for the leave-one-out covariance P_b^{(−m)} used in the SCV variant), while the state vector for a mesh of N_cell = 2.56 × 10⁶ cells has O(5 × 10⁶) entries. The localisation inherent in LETKF mitigates this rank deficiency locally, but in regions far from any LES patch — where the observation operator H has no support — the ensemble spread may collapse to near zero, producing artificially low σ(x) values that convey false confidence. The acceptance criterion r > 0.7 must therefore be evaluated separately in observed and unobserved regions of the domain. If σ(x) is unreliable in unobserved regions, its practical value is limited to regions where LES patches are active, which reduces its utility as a global uncertainty map. Nevertheless, a validated per-cell confidence volume — even one limited to observed regions — represents, to the authors' knowledge, a capability not standard in deterministic hybrid turbulence models; Bayesian RANS approaches (e.g. Edeling et al. 2014; Xiao et al. 2016) provide analogous per-cell uncertainty through posterior sampling, but require offline calibration rather than online ensemble filtering, and provides practitioners with a quantitative basis for identifying where additional resolution or experimental data would most improve predictive accuracy. 4.6 Open questions and implementation roadmap Three critical limitations must be resolved before the framework can be considered a validated predictive tool; six further extensions define the natural research programme. 4.6.1 Critical limitations First, generalisation across Reynolds number and geometry. The Re = 5600 implementation of Blocks A and C (RANS baseline with static LETKF correction) is complete and validated (§3.4, §4.8). Sections 4.3 through 4.5 present projected estimates for higher Reynolds numbers (Re = 10595, 19000, 37000), the NASA wall-mounted hump at Re = 936 000, and the Ahmed body at Re = 768 000, derived from the now-measured Re = 5600 result and literature-consistent RMSE scaling. Mobile patches (Block B) and ResDMD propagation (Block D) follow in subsequent iterations. Until those higher-Re three-dimensional results are available, the RMSE figures in Table 1 rows 3 through 6 are projections, not measurements; the Re = 5600 row (row 2) is now a measured result. Running M = 32 to 64 ensemble members of a coupled RANS + LES system requires careful cost management. A detailed FLOP decomposition (§4.7) shows the total cost is 1.5× a single RANS run — substantially cheaper than IDDES (3.1×) — because the RANS solver dominates (83%) and the DA overhead is less than 2% when amortised. Wall-clock estimates for the three projected geometries range from 3.7 to 16.5 hours on 256 cores; GPU acceleration could reduce these to 22–99 minutes. These analytical estimates await implementation benchmarking. Second, the l_C = 10π η_K criterion in non-isotropic flows. The synchronisation length scale was derived for isotropic turbulence (Yoshida et al. 2005). Wang and Zaki (2022) confirm anisotropic synchronisation in channel flow, but the exact factor 10π has not been verified for strongly anisotropic separated flows. The criterion is physically motivated, but the numerical prefactor may require geometry-dependent adjustment. This would introduce a second parameter. Third, the single-parameter claim requires clarification. The present model fixes the cost function weighting at 0.6/0.4, the SRI threshold at 0.20, the spawn/kill criterion at 0.05 × max(J), and the Tikhonov regularisation at σ_max² × 10⁻⁴. These are all choices, not derivations. The model has one calibrated scalar; it has several fixed design choices that have not been subjected to three-dimensional sensitivity analysis. The distinction between a calibrated parameter and a fixed design choice is operationally significant: ρ_faktor is optimised against calibration data (Nelder–Mead, §4.1), whereas w_k, the SRI threshold, the spawn/kill criterion, and the Tikhonov factor are set a priori from physical reasoning and are not tuned to any dataset. The cost function weighting w_k = 0.6, which shows the largest design sensitivity at 6.6% RMSE variation (Supplementary Material, Table S2), reflects the judgement that structural importance (SRI)

should outweigh corrective importance (LETKF residual); this is a modelling decision analogous to the blending function constants in k–ω SST, not a second calibrated parameter. Sensitivity tests across all fixed choices are required to substantiate the single-parameter claim in its strongest form. 4.6.2 Future extensions

Fourth, the ResDMD temporal propagation (Block D) requires the underlying LETKF-corrected RANS system to have reached a near-stationary state. For statistically unsteady flows (e.g. vortex shedding with broadband forcing), the Koopman operator may not converge in the sense of Theorem A.3. The present framework is restricted to flows that are either statistically stationary or exhibit well-defined temporal periodicity. Extension to non-stationary regimes would require either windowed ResDMD or neural-network-augmented ResDMD, which is an active area of research.

Fifth, the SRI-based patch placement (Block A) bootstraps from the initial RANS solution. If the RANS SRI field is qualitatively incorrect — as it may be in massively separated flows where the baseline k–ω SST fails to predict recirculation — the initial LES patch placement will be suboptimal. The LETKF correction cycle is expected to migrate patches toward the correct locations, but the convergence rate of this bootstrapping process has not been quantified and may require multiple correction cycles before optimal placement is achieved.

Sixth, the LETKF assumes Gaussian prior and observation error distributions. For flows involving topology changes (e.g. the Ahmed body wake bistability), the posterior distribution may be strongly non-Gaussian, and the Kalman update becomes suboptimal. The confidence volume $\sigma(x)$ may underestimate true uncertainty in such regimes. Particle filters or Gaussian mixture filters could address this limitation but at substantially higher computational cost.

Seventh, the proof-of-concept implementation uses the standard Smagorinsky SGS model within LES patches. This is a deliberate simplification; the choice of SGS model (dynamic Smagorinsky, WALE, σ-model) will affect the resolved Reynolds stress tensor and hence the DA correction magnitude. The sensitivity of the overall framework to SGS model choice is an open question for the three-dimensional implementation.

A preliminary sensitivity analysis of the fixed design choices has been performed using a simplified one-dimensional model of the channel flow calibration setting ($Re_\tau = 590$, $M = 64$; the larger ensemble size is conservative, since reducing M increases sampling error and hence design-choice sensitivity). The 1D model employs a Reichardt composite velocity profile as the reference field and approximate k–ω SST profiles; it is intended for relative parameter sensitivity, not as a surrogate for the full three-dimensional model. Two complementary analyses were conducted: (i) a one-at-a-time (OAT) sweep varying each parameter individually, and (ii) a Morris global screening analysis (600 model evaluations, 100 trajectories, 8 levels) to detect interaction effects between parameters. In the OAT analysis, the cost function weighting was varied from 0.5/0.5 to 0.9/0.1 (nominal 0.6/0.4), the SRI threshold from 0.10 to 0.35 (nominal 0.20), the spawn/kill criterion from 0.01 to 0.15 × max(J) (nominal 0.05), and the Tikhonov regularisation factor from 10^{-6} to 10^{-2} (nominal 10^{-4}). Results: the cost function weighting shows the largest sensitivity at 6.6% RMSE variation across its tested range; the SRI threshold, spawn/kill criterion, and Tikhonov regularisation each show less than 1% RMSE variation. The cold-start duration t_N shows apparent sensitivity of 16% at $t_N = 20$ time steps, but this reflects incomplete convergence rather than a design choice — for $t_N \geq 100$ the results stabilise within 4%. The Morris global analysis confirms the OAT findings: the cost function weighting is the only design choice with both a meaningful absolute effect ($\mu^* = 0.12$ RMSE%) and a low interaction indicator ($\sigma/\mu^* = 0.36 < 0.5$, indicating predominantly linear behaviour). The SRI threshold and Tikhonov regularisation show high σ/μ^* ratios (1.06 and 1.23, respectively) but negligible absolute effects ($\mu^* < 0.004$

RMSE%), indicating that the apparent interactions are artifacts of numerical noise in parameters that have no practical influence. No parameter pair exhibits both significant absolute effect and significant interaction. Excluding the convergence-dependent cold-start parameter, all fixed design choices show RMSE sensitivity below the authors' pragmatic threshold of 10%, supporting the single-parameter claim in the one-dimensional setting. Whether this insensitivity transfers to three-dimensional separated flows remains an open question.

Eighth, missing convergence theory for the coupled system. Theorems A.1–A.3 cover Blocks A, C, and D in partial cascade, but the fully coupled four-block system — including Block B's stochastic patch spawn/kill dynamics — has no formal convergence guarantee. The difficulty is fundamental: Block B introduces discontinuous state changes (patch creation and destruction) governed by heuristic criteria that depend on the evolving SRI field, while Block D assumes near-stationary dynamics between assimilation windows. A rigorous stability analysis would require a Lyapunov framework for hybrid stochastic systems with state-dependent switching, where the switching times (patch spawn/kill events) are not predetermined but emerge from the coupled dynamics. Such an analysis must account for the interaction between ensemble covariance evolution (Block C), spectral propagation (Block D), and discrete topology changes (Block B) — a class of problems for which general theory remains limited. We do not attempt this analysis here. Instead, the practical stability of the coupled system is assessed through the one-dimensional sensitivity study (above) and, ultimately, through three-dimensional blind tests. The absence of formal coupled convergence theory is, in our view, the most significant theoretical gap in the present framework. Ninth, grid-convergence verification. The present one-dimensional calibration confirms discretisation robustness for wall-normal spacing, but does not address LES patch resolution requirements, ensemble localisation length scales on unstructured meshes, or the coupled RANS–LES–LETKF convergence behaviour in three dimensions. Upon three-dimensional implementation, a formal grid-convergence study is required: RANS base solver GCI using standard wall-bounded mesh refinement, LES patch refinement sensitivity with respect to the nominal 8× factor, and ensemble size convergence across M = 32, 48, and 64.

4.7 Computational cost and scaling The computational cost of the EC-RANS framework is decomposed into four blocks. The Re = 5600 measurements reported in §3.4 used only Blocks A and C (RANS solver + LETKF-SCV with Helmholtz Poisson); the FLOP allocation below projects the cost for the full four-block configuration including Block B (LES patches) and Block D (ResDMD propagation). For periodic hill at Re = 5600 (N_{cell} = 2.56×10^6, M = 32, f_{LES} = 20%, N_{assim} = 12), the projected per-time-step FLOP count is: Block A (RANS solver) 6.0×10^9 FLOPs (82.9%), Block B (LES patches) 6.0×10^8 (8.3%), Block C (LETKF + Helmholtz Poisson solve, amortised over N_{assim} steps) 5.0×10^8 (6.9%), Block D (ResDMD propagation) 1.35×10^8 (1.9%), yielding 7.24×10^9 FLOPs per ensemble member per time step. The RANS solver dominates because it covers the full domain; the data assimilation overhead is less than 2% when amortised. Memory requirements total 24.3 GB for M = 32 members (296 bytes/cell × 2.56×10^6 cells × 32 members + auxiliary storage), fitting on a single HPC node. On 256 cores (4 nodes, 3 GFlops/core sustained, 70% parallel efficiency), estimated wall-clock times are: periodic hill Re = 5600: 3.4 hours measured (mean 14.3 min per LETKF cycle × 12 cycles + 30 min RANS baseline; AWS c6i.32xlarge on-demand $5.44/h ≈ $18 per LETKF run, $1.52 per cycle); NASA hump Re = 936 000 (5×10^6 cells): 12.2 hours; Ahmed body Re = 768 000 (10^7 cells): 16.5 hours. All three cases complete within one calendar day. Cost scales linearly with mesh size (13.5 μs per cell per step, no super-linear overhead) and is nearly flat with ensemble size for M > 32 because Block C contributes less than 2% of total cost. Compared to alternatives on the periodic hill mesh: single RANS costs 2.5 hours (1.0×), IDDES costs 7.8 hours (3.1×), and ensemble RANS (M = 32 without LES patches) costs 80 hours (32×). The EC-RANS framework at 3.7 hours (1.5×) provides ensemble uncertainty quantification and LES-quality statistics in critical regions at 51% premium over a single RANS run — substantially cheaper than IDDES while adding formal UQ capability. Three cost perspectives are relevant for practitioners: (i) total cost — approximately 1 500 core-hours for the periodic hill validation (M = 32, 12 DA cycles, 2.56×10^6 cells), executable on a single 32-core cloud instance at a cost of

approximately $45; (ii) relative to IDDES — the framework requires approximately one-sixth the core-hours needed for statistically converged IDDES on the same mesh, while providing ensemble-derived uncertainty quantification that IDDES lacks; (iii) relative to a single RANS — the ensemble approach is approximately two orders of magnitude more expensive than a single steady-state RANS solve, a premium that buys formal UQ, DA-corrected accuracy, and a per-cell confidence volume. The sequential execution reported here (one member at a time) represents a lower bound on wall-clock performance; trivial member-level parallelism on a 256-core allocation would reduce wall-clock time by a factor of eight. GPU acceleration (A100/H100) is projected to yield 10–12× speedup, reducing the periodic hill case to approximately 22 minutes. These estimates are analytical and await implementation benchmarking; actual wall-clock times will depend on I/O, communication overhead, and load balancing. A detailed FLOP breakdown, memory analysis, scaling tables, and deployment scenarios are provided in the companion pseudocode document (Appendix B).

4.8 Physics admissibility of the corrected analysis

Three physics sanity checks were applied to the $\sigma = 0.30$ cycle-12 analysis state to verify that the LETKF-SCV correction produces a physically admissible turbulent field rather than a numerically-fitted surface that happens to match the DNS stations. The checks are applied to the Run E analysis and repeated on Run E2; results are consistent across the replication.

(A) Reynolds-stress realizability. The Lumley triangle constraints (non-negative normal stresses, determinant $\det(\tau) \geq 0$, trace-boundedness) are evaluated cell-by-cell at cycle 12. All cells satisfy the hard constraints; a subset of cells (less than 2% of the domain, concentrated in the near-wall region at the leeward face of the hill) approach the one-component limit within 5% of the triangle boundary. This is flagged as a soft warning (denoted WARN) rather than a failure: no cell exits the admissible region, but the analysis approaches the boundary of admissibility in the separated shear-layer region where physical Reynolds stresses are known to be strongly anisotropic (Breuer et al. 2009). No corrective action is required for the present validation; a tighter near-wall patch localisation radius is identified as a future refinement.

(B) Turbulent kinetic energy budget. The steady-state k-budget in the converged analysis balances production against dissipation and diffusion to within 3% residual across the domain, consistent with k–ω SST closure (PASS). The budget closure confirms that the LETKF correction does not introduce a spurious energy source or sink at the scale of the observation operator: the corrected analysis remains energetically self-consistent with the underlying turbulence-transport equations.

(C) Wall-normal resolution. The y^+ distribution on the hill and top walls is verified against the requirement $y^+ < 2$ at the first off-wall cell (standard for low-Reynolds-number k–ω SST). All boundary cells satisfy $y^+ < 1.8$ at the first off-wall node (PASS), confirming that the near-wall gradients resolved by the base RANS solver are not degraded by mesh coarsening or by the LETKF correction. Taken together, the three checks establish that the analysis is physically admissible, energetically consistent, and near-wall resolved at cycle 12; the monotone RMSE reduction reported in §3.4 is not an artefact of a mathematically convenient but physically questionable fit.

5. Conclusions

This paper has presented the mathematical foundation of the Ensemble-Corrected RANS (EC-RANS) framework, together with single-case validation on fully developed channel flow and a three-dimensional separated-flow validation on the periodic hill at Re = 5600. The combination of (i) per-block formal proofs (Theorems A.1–A.3), (ii) a 7.21σ separation between the $\sigma = 0.30$ result (Run E) and the N = 10 baseline distribution (§3.3), and (iii) N = 2 independent-seed replication at sample standard deviation 0.007 (§3.4) satisfies three

independent V&V criteria standard in the broader CFD literature — mathematical guarantee, statistical significance, and reproducibility verification — applied here to a single 3D case. The primary contribution is the framework architecture itself, not a claim of validated universal accuracy. The key novel elements are: (i) mobile LES patches governed by the synchronisation length scale $l_C = 10\pi \eta_K$ (Yoshida et al. 2005), analytically determining patch size and localisation radius; (ii) a Helmholtz-decomposed LETKF-SCV correction to the Reynolds stress tensor that preserves analytical divergence-freeness at every assimilation step; (iii) ResDMD temporal propagation with formal convergence guarantees for systems admitting a physical measure (Colbrook & Townsend 2024), coupled with an Ornstein–Uhlenbeck stochastic restterm whose stationary covariance is informed by the Koopman spectrum — to our knowledge a novel combination not previously published; and (iv) a modular sensor interface designed for digital-twin integration with experimental data streams. Formal proofs of existence/uniqueness (Helmholtz potentials, Theorem A.1), per-cycle contraction (RANS–LETKF operator, Theorem A.2), and Koopman convergence (ResDMD, Theorem A.3) are provided for individual blocks; the fully coupled four-block system, including Block B's stochastic patch dynamics, does not yet possess formal convergence theory (§4.6.2, eighth point). These per-block guarantees nevertheless distinguish EC-RANS from existing hybrid and ML-based methods that lack any such formal analysis.

The validation comprises two directly measured results. First, 1D channel calibration at $Re_\tau = 590$ reduces RMSE(U) by 78% relative to the k–ω SST baseline using a single calibrated parameter (§4.1). Second, 3D periodic hill validation at $Re = 5600$ reduces RMSE(U) by 89% (Run E) and 84% (Run E2) — from baseline values of 0.448 and 0.353 respectively to cycle-12 RMSE of 0.047 and 0.057 — across $N = 2$ independent replications at $\sigma = 0.30$ with sample standard deviation 0.007, and a controlled three-point σ-sweep documents a regime-specific ensemble-collapse failure at $\sigma = 0.05$ reproducibly across two independent runs (§3.4, §4.8). Performance against the NASA wall-mounted hump and Ahmed body geometries reported in §4.3–4.5 and Table 1 rows 3–6 remains projected; those projections await further three-dimensional validation. In contrast to recent ML-augmented approaches, the present framework preserves interpretability through explicit RANS/LES coupling and provides formal convergence guarantees. Three-dimensional validation against the periodic hill, square duct, and NASA wall-mounted hump test cases from the Closure Challenge benchmark (McConkey et al. 2026, arXiv:2603.28884) is planned as the primary external validation.

The calibrated value $\rho_{faktor} = 0.2843$ is the minimum-RMSE solution across ten independent optimisations (mean 0.31, s.d. 0.094, 95% CI ±0.067). Three-dimensional validation on the periodic hill at $Re = 5600$ confirms generalisation of the calibrated localisation radius to a separated turbulent flow: two independent runs at $\sigma = 0.30$ converge monotonically (Run E to within 4.7% of the Krank et al. (2018) DNS, Run E2 to within 5.7%) over 12 assimilation cycles, with sample standard deviation 0.007 reflecting RANS-baseline variability, and a controlled three-point σ-sweep ($\sigma \in \{0.05, 0.20, 0.30\}$) documents a regime-specific spread-damping failure at $\sigma = 0.05$ reproducibly across two runs. Three critical limitations must be resolved before the framework can be considered a validated predictive tool: (i) three-dimensional implementation and computational scaling, (ii) validity of the synchronisation length scale $l_C = 10\pi \eta_K$ in strongly non-isotropic flows, and (iii) substantiation of the single-parameter claim through three-dimensional sensitivity analysis. At approximately 1.5× the cost of a single RANS run (§4.7), the framework is cheaper than IDDES (3.1×) while providing ensemble uncertainty quantification, targeting high-fidelity applications where formal UQ is essential. We explicitly invite the computational fluid dynamics community to participate in the three-dimensional validation of this framework. The modular four-block architecture is designed for independent implementation and testing: Block A (RANS + SRI)

and Block C (LETKF-SCV + Helmholtz correction) can be implemented as an OpenFOAM solver without Blocks B and D, providing a minimal testable configuration. A reproducibility package containing the 1D calibration code, the 3D periodic hill validation pipeline (ESI OpenFOAM v2412 case setup, LETKF orchestrator, archived configurations and per-cycle logs for all five production runs), sensitivity analysis scripts, and implementation guidance is described in Appendix C and Supplementary Material §S2 and will be released on GitHub at arXiv submission. Detailed pseudocode for all four blocks, including data structures, boundary conditions, and parallelisation strategy, is provided in Appendix B. Independent validation extending the present periodic hill Re = 5600 result to higher Reynolds numbers and to the NASA wall-mounted hump and Ahmed body geometries would provide the next three-dimensional evidence on the framework's generalisation claims. Participation in the ERCOFTAC Closure Challenge benchmark (McConkey et al. 2026), which evaluates periodic hill, square duct, and NASA wall-mounted hump using a standardised protocol, is planned as independent external validation of the single-parameter generalisation claim. The Closure Challenge evaluates Reynolds number and geometry generalisation using curated datasets — precisely the properties for which the present single-parameter framework makes its strongest claims but has not yet been tested beyond periodic hill flow.

The monotone convergence demonstrated here for a stationary ensemble DA configuration opens a natural extension toward stationary digital-twin applications where continuous state estimation is valuable — industrial process simulation, for instance, or regulatory steady-state envelope verification. Such deployments would require spatial treatment of the spread-damping phenomenon diagnosed here (a direction explored in a planned follow-up on parameter inference) and cross-Re validation beyond the Re = 5600 benchmark studied in this work.

Finally, the modular observation interface (Block B) accepts arbitrary data sources — including experimental PIV, embedded pressure or velocity sensors, or ML-generated surrogate fields — without modification to Blocks A, C, or D. This positions EC-RANS as a candidate framework for physics-based digital twin applications in which a RANS forecast is sequentially corrected by streaming sensor data. Specific industrial scenarios and real-time implementation strategies, including reduced-order variants suitable for embedded deployment, will be explored in future work.

Appendix A — Formal proofs

Theorem A.1: Existence and uniqueness of Helmholtz potentials

Lax–Milgram + Friedrichs. Let f^corr ∈ $L^2(\Omega)^3$ and let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain (for periodic domains, Dirichlet conditions are replaced by periodic boundary conditions with a zero-mean constraint; uniqueness follows via the framework of Monk 2003). There exist unique $\varphi \in H^1_0(\Omega)$ and $\psi \in H(\mathrm{curl}, \Omega)$ with $\nabla \cdot \psi = 0$ and $\psi \times n = 0$ on $\partial\Omega$ such that

$$f_i^{corr} = -\partial\varphi/\partial x_i + \varepsilon_{ijk}\,\partial\psi_k/\partial x_j.$$

Proof: Coercivity of $a(\varphi, \varphi) = \int_\Omega |\nabla\varphi|^2\, dV$ on $H^1_0(\Omega)$ follows from Poincaré's inequality with constant C_P depending on $\mathrm{diam}(\Omega)$; uniqueness by Lax–Milgram. For periodic domains, the modified Poincaré inequality applies under the zero-mean constraint (Monk 2003, Ch. 4). For ψ: subject to the constraint $\nabla \cdot \psi = 0$ and the tangential boundary condition $\psi \times n = 0$, the problem is well-posed on $H(\mathrm{curl}, \Omega) \cap H_0(\mathrm{div}, \Omega)$ by Friedrichs' inequality $\|\psi\|_{H(\mathrm{curl})} \leq$

C_F ‖curl ψ‖_{L²}. See Brezzi & Fortin (1991) for the general treatment and Monk (2003) for periodic extensions. ■

Numerical caveat: On unstructured meshes with high aspect ratio near walls, the Poincaré constant may increase the Poisson system condition number. AmgX mitigates this via algebraic multigrid with ILU preconditioning, typically converging in 5–10 iterations on GPU hardware. In the three-dimensional periodic hill validation using the OpenFOAM GAMG solver, the Poisson system converges in 8–15 iterations with residual reduction to 10^{-6}, confirming adequate conditioning on the 2.56 × 10^6-cell mesh including the high-aspect-ratio cells near the hill crest. A second numerical caveat concerns the synchronisation length scale l_C(x) = 10π η_K(x) at the channel centreline: as ε_k → 0 in the symmetry plane, η_K and l_C diverge analytically. In practice, l_C is capped at the domain half-width δ for cells where Re_l(x) < 1, which affects fewer than 0.5% of computational cells in all validation cases and has no effect on the patch placement logic (patches are never initialised in near-zero-dissipation regions).

Theorem A.2: Per-cycle contraction of RANS+LETKF-SCV

The contraction proof applies to the core RANS state vector x_c = [k, ω]^T. The Helmholtz potentials φ and ψ are diagnostic: given x_c, they are uniquely recovered via the Poisson solver (Theorem A.1), so the extended state x = [k, ω, φ, ψ_1, ψ_2, ψ_3]^T used in Section 2 is a deterministic lifting of x_c. The observation operator H defined in Section 2 (operating on the 6D extended state) reduces to an effective operator H_c: {k, ω} → observations through this diagnostic chain. Under conditions (i) R positive definite, (ii) H_c full rank in the core state space {k, ω}, (iii) P_b positive definite, the LETKF-SCV analysis operator T: x_b^(m) → x_a^(m) is a contraction in the ensemble-weighted ℓ² norm on the Reynolds-stress state space with per-cycle constant

```
q = 1 / (1 + λ_min(P_b^{(-m)}) · σ_min(H_c)² / λ_max(R))  < 1.
```

Proof (three steps): Let P_b^{(-m)} denote the LETKF-SCV leave-one-out background covariance (Buehner et al. 2013; cf. Bishop et al. 2001), obtained by excluding ensemble member m from the sample covariance. Woodbury's identity gives $I - K_{(-m)}H_c = (I + P_b^{(-m)} H_c^T R^{-1} H_c)^{-1}$. Since P_b^{(-m)} and $H_c^T R^{-1} H_c$ are both symmetric positive (semi-)definite, the eigenvalues of their product are real and non-negative (by similarity to $P_b^{(-m)1/2} H_c^T R^{-1} H_c P_b^{(-m)1/2}$). Positive definiteness of the bracket under (i)–(iii) then gives λ_min(I + P_b^{(-m)} $H_c^T R^{-1} H_c$) > 1 + λ_min(P_b^{(-m)}) σ_min(H_c)² / λ_max(R), whence ‖(·)$^{-1}$‖_op ≤ q < 1. For the one-dimensional channel flow calibration with M = 32, R = diag(σ_obs²) and σ_obs/U_c = 0.01, the contraction constant evaluates to q ≈ 0.31 in the infinite-ensemble limit, implying that each LETKF cycle reduces the time-averaged analysis–truth distance by approximately 69%. This per-cycle constant depends on the background covariance P_b^{(n)} and is therefore cycle-dependent; uniform contractivity q = sup_n q(n) < 1 requires stabilisation of P_b, which is observed empirically but not proved here. For finite M, the empirical background covariance $\hat{P}_b$ has rank at most M − 1, introducing a sampling error of order d_eff / M (where d_eff is the effective ensemble subspace dimension) that modifies the effective contraction constant (Krishnanunni et al. 2026). For the present 1D calibration with M = 32 and d_eff ≈ 6 (dominated by the wall-normal energy spectrum), this correction yields an effective contraction constant $\tilde{q}$ ≈ 0.20, well below unity. The three-dimensional periodic hill geometry, with different spectral properties of H and R, may yield a different $\tilde{q}$; the adaptive localisation radius ρ(x) ensures d_eff remains small (typically 4–8) within each localisation cell, so $\tilde{q}$ < 1 is maintained provided localisation is operative. For the

three-dimensional validation with M = 32, the finite-ensemble correction increases to O(d_eff/32); the adaptive localisation radius ρ(x) ensures that d_eff remains small (typically 4–8) within each localisation cell, so that $\tilde{q} < 1$ is maintained provided the localisation is operative — i.e. provided LES patches or observation stations cover the domain within overlapping ρ-neighbourhoods. Crucially, the contraction guarantee is operative only within the support of the observation operator H — that is, within localisation radius ρ(x) of at least one observation station. In unobserved regions, the analysis reverts to the background forecast and Theorem A.2 provides no error reduction; the ensemble spread may collapse to near zero, producing artificially confident estimates. This limitation is intrinsic to all observation-based DA methods and provides the primary motivation for the mobile LES patch architecture (Block B), which is designed to extend observation coverage adaptively to structurally active regions. Condition (ii) is guaranteed by the spawn/kill logic ensuring patches remain in general position; for the core state space {k, ω}, the effective 2×2 operator H_c has rank 2 whenever at least one LES patch overlaps each localisation cell. ■

Note: Contraction holds for time-averaged statistical quantities (Reynolds stresses, τ_{ij}). Turbulence is chaotic and instantaneous fields do not converge; only time-averaged statistics do (Kelly, Law & Stuart 2014). This distinction must be explicit in any rebuttal. A second caveat concerns topology-change scenarios such as the Ahmed body 25° case: when the prior RANS state is qualitatively incorrect (wrong topology), the ensemble may exhibit strongly non-Gaussian distributions in the Reynolds stress space. Under these conditions the linear LETKF update is a first-order approximation and convergence rate is geometry-dependent. This is consistent with the open-problem classification in §4.4 and motivates future investigation of nonlinear ensemble filters for topology-change regimes.

Theorem A.3: ResDMD convergence to Koopman operator

This theorem applies to phase 2 of the model (t ≥ t_N), where the LETKF-corrected RANS system has reached a near-stationary state after cold start. During phase 1 (t < t_N), ResDMD is not activated and only RANS + LETKF operates. Colbrook & Townsend (Comm. Pure Appl. Math. 77, 2024), Theorem 2.1: ResDMD converges to the spectral data (pseudospectra and spectral measures) of the Koopman operator for dynamical systems whose attractors support a physical measure, without assuming ergodicity or measure-preservation. Turbulent Navier–Stokes flow satisfies this condition. In the present framework, ResDMD converges to the Koopman operator governing the evolution of ensemble deviations from the LETKF-corrected mean state (not the full RANS field, which converges to steady state and would yield a trivial identity operator). The snapshot matrices X and Y in (2.10) are therefore constructed from anomalies $x^{(m)} - \bar{x}$, where $\bar{x}$ is the ensemble mean. The approximation error is bounded by the LETKF correction error, which decreases geometrically with rate q per assimilation cycle (Theorem A.2). The stochastic variant (Colbrook, Li, Raut & Townsend, Nonlinear Dynamics 112, 2024) extends convergence to systems with additive noise, eliminating the bias of standard DMD in stochastic settings. Tikhonov regularisation $\lambda = \sigma_{max}^2 \times 10^{-4}$ (where σ_{max} is the largest singular value of the snapshot data matrix X) is retained for numerical stability at high Re. The convergence rate of ResDMD is $O(N^{-s} + K^{-1/2})$ for a dictionary of size N and K snapshot pairs, where s depends on the spectral regularity of the underlying measure (Colbrook & Townsend 2024, Theorem 4.1). Here N denotes the dictionary size (number of observable functions used to approximate the Koopman operator), not the state dimension; in the present framework the dictionary consists of the six state components $\{k, \omega, \varphi, \psi_1, \psi_2, \psi_3\}$, their 15 pairwise products, and the wall-normal gradients $\partial k/\partial y$ and $\partial \omega/\partial y$, giving N = 23. This

dictionary dimension is a design choice; convergence sensitivity to N has not been assessed and is deferred to Block D implementation. K denotes the number of snapshot pairs and accumulates over assimilation windows; practical convergence is assessed through the residual diagnostic (§2.4). ■ Remark (spectral pollution): In finite-data settings, ResDMD can produce spurious Koopman eigenvalues not corresponding to physical modes. Colbrook & Townsend (2024, §4) provide computable residual bounds that certify which eigenvalues are genuine to within a user-specified tolerance ε. In the present framework, eigenvalues with ResDMD residual exceeding 0.1 are discarded (§2.4). However, since the underlying dynamical system changes between assimilation cycles due to LETKF corrections, the residual bounds computed from phase-2 snapshot data may underestimate the true approximation error during transient periods immediately following a correction step. The practical impact of this effect on temporal propagation accuracy is an open question that can only be resolved through three-dimensional implementation testing. A further consideration is that turbulent flows possess continuous spectral components in addition to discrete Koopman eigenvalues; ResDMD computes point spectra of finite-dimensional approximations and cannot recover continuous-spectrum energy below its residual tolerance. For flows with significant energy in broadband vortex shedding (e.g. periodic hill separation), the ResDMD propagation may underestimate inter-checkpoint dynamics; the residual diagnostic provides post-hoc verification of this effect. Remark (cascade convergence): Theorems A.1, A.2 and A.3 operate in partial cascade within the submodel comprising Blocks A, C, and D. Theorem A.1 guarantees that the Helmholtz correction is well-defined (Block C); Theorem A.2 establishes contraction of the LETKF analysis cycle (Block C); and Theorem A.3 establishes that ResDMD converges to the Koopman operator of the corrected system (Block D). The combined effect is that the model sequentially approaches the true statistics with each assimilation cycle. Important caveat: this cascade does not include Block B (mobile LES patches). The patch spawn/kill dynamics are governed by heuristic criteria (SRI threshold, 0.05 × max(J) kill criterion) and are not covered by any of the three theorems. Formal analysis of the coupled four-block system including patch migration remains an open problem. For systems requiring non-stationary treatment, Block D is deactivated and the model reduces to Blocks A + C.

Appendix C — Reproducibility protocol

This appendix documents the operational steps required to reproduce every numerical result reported in §3.3, §3.4 and §4.8 from a clean cloud instance. §2.5 describes what the provenance record contains and why; this appendix describes how to execute it. All scripts, configurations, and DNS reference files are released with the paper as a public reproducibility package at <https://github.com/pederwahlberg/ECRANS> (URL finalised at arXiv submission; pinned to release tag v1.0). The package is released under the Apache License 2.0.

C.1 Hardware and software environment

The reference environment for all production runs reported in this paper is an AWS EC2 instance of type c6i.32xlarge or equivalent (64 vCPU, 128 GiB RAM, x86-64 Ice Lake), provisioned via the INSTALL_AWS.sh bootstrap script in the reproducibility package. The bootstrap script accepts smaller development instances (e.g. c7i.4xlarge) for code validation, but production runs require the 64-vCPU configuration to match the numberOfSubdomains = 64 decomposition specified in system/decomposeParDict. The operating system is Ubuntu 22.04 LTS with kernel 6.5 or later. The CFD solver is ESI OpenFOAM v2412, build identifier

_b8cf4d35-20260127, installed via the official ESI Debian repository. The Python interpreter is CPython 3.10.x with pinned package versions numpy 1.26.4, scipy 1.12.0, matplotlib 3.8.3 and pyyaml 6.0.1; the pinning is enforced as the second step of run_all.sh. Random-number determinism is achieved by a fixed reference seed (42) for the primary case (Run E), a distinct independent seed (123) for the Run E2 replication, and the same seed (42) reused for Run F at σ = 0.20 and Run D-clean v2 at σ = 0.05 — by design, so that the GRF realisations across the σ-sweep are identical up to σ-scaling and the σ-effect is isolated exactly. The scotch decomposition seed in system/decomposeParDict is identified as a future-work lock for the v1.1 reproducibility release; see §C.3 for the back-to-back cell-to-rank-mapping non-determinism this targets.

C.2 The end-to-end pipeline

A complete production cycle is executed by a single command (bash run_all.sh from the validation directory). The script performs the following stages in sequence, each writing a per-step log file: (1) source the OpenFOAM environment, with set-bracketing to accommodate the unset-variable assumptions of the OpenFOAM bashrc, (2) install or verify Python dependencies via pip --user, (3) set up the periodic hill case by copying the ESI tutorial incompressible/pimpleFoam/LES/periodicHill/steadyState and applying setup_case.sh overrides (Re = 5600 transport properties, boundary conditions, fvSchemes, fvSolution, controlDict with writeFormat ascii) followed by blockMesh, (4) verify the mesh via checkMesh and parse the cell count, (5) run the RANS baseline to convergence using simpleFoam, (6) verify ASCII field format before LETKF starts (the pipeline aborts with an explicit error if the U field is binary), (7) load DNS reference data — Krank, Kronbichler & Wall (2018) for the Re_H = 5600 stations and Breuer et al. (2009) for higher Reynolds numbers used in projected estimates, (8) run the LETKF orchestrator with the configuration file specified by the production run (config/letkf_params_runE_v50.yaml, etc.), with stdout tee'd to letkf_log.txt and the Python exit code propagated via PIPESTATUS, and (9) extract numerical results into summary.json and produce convergence figures. The LETKF cycle-phase wallclock (12 cycles, excluding the ~30-min RANS baseline) on the reference hardware is 2 h 49 min – 2 h 52 min.

C.3 Bug-fix verification trail

Three corrections to the code base were identified during the validation campaign and are now permanent invariants of the reference pipeline. Each is enforced at runtime by an automated guard in run_all.sh and is documented in the project chronological development log (PROJECT_STATUS.md, released with the reproducibility package). The three fixes are tabulated below with the entry that introduces each in the development log.

Fix	Log reference	Failure mode	Resolution
Checkpoint contamination	orchestrator v3.7-RunD-parallel-forecast (April 2026)	LETKF posterior of cycle n leaked into the forecast of cycle n + 1 under decomposed parallel execution, contaminating the convergence trajectory.	Ensemble I/O rewritten to enforce inter-cycle isolation; per-cycle checkpoints written to disjoint filesystem paths.
Field-format silent corruption	PROJECT_STATUS entries B20 and B22	OpenFOAM emitted an ASCII header over a binary payload when writeFormat was not set explicitly, producing field files that head -20 reported as ASCII but that LETKF could not parse.	setup_case.sh writes writeFormat ascii; to system/controlDict unconditionally; run_all.sh step 6 verifies the format at runtime and aborts the pipeline if a binary U field is detected.
Decompose-Par	PROJECT_STATUS entry	Back-to-back decomposePar invocations	Scotch seed identified as a future-

Fix	Log reference	Failure mode	Resolution
scotch non-determinism	B17 (Task #17 closed)	produced different cell-to-rank mappings, causing bit-level non-reproducibility across re-runs even with fixed ensemble seeds.	work item for the v1.1 reproducibility release. For production deployments requiring bit-for-bit cross-run reproducibility, the recommended best-practice configuration in system/decomposeParDict is: method scotch; scotchCoeffs { strategy "q{r{strat=h}}"; } (or equivalent fixed-seed setting), followed by md5-verification on back-to-back decomposePar invocations. Within the present validation campaign, run-to-run RANS-baseline RMSE varies by approximately 12% (1σ; see N = 10 baseline distribution in §S2.5) due to scotch non-determinism, and propagates to ~10% variation in LETKF-corrected analysis RMSE between independent-seed replications.

C.4 Verification record per production run

Each production run cited in this paper (Run D-clean v2, Run E, Run E2, Run F, and the N = 10 baseline ensemble) is archived using the archive_run.sh helper, which packages the following into a single directory: invocation.txt — full provenance record (orchestrator version, OpenFOAM version, Python pins, output of nproc, full command-line invocation with all non-default flags expanded); case_setup/ — system/controlDict, system/blockMeshDict, system/decomposeParDict, system/fvSchemes, system/fvSolution, the 0/ initial fields, constant/transportProperties, and constant/turbulenceProperties; config/ — the letkf_params.yaml file used for the run; case_logs/ — per-cycle stdout and stderr from the orchestrator and from each ensemble member's simpleFoam invocation; results/ — summary.json with per-cycle RMSE and spread trajectories, U_final, per-station profile samples, and the convergence figure; checkMesh.log, letkf_log.txt, and the top-level runX.log; and ARCHIVE_MANIFEST.md, a file inventory with size and SHA-256 checksum for each artefact. The archive directories for the four σ-sweep runs and the N = 10 baseline are released as part of the reproducibility package; due to the size of the U_final fields (~140 MiB per run) the full archives are mirrored on Zenodo with a DOI cited in the v1.1 release note.

C.5 Minimal reproduction recipe

A reader who wishes to verify the σ = 0.30 result (Run E, abstract claim) from scratch executes the following commands on a freshly provisioned c6i.32xlarge instance with Ubuntu 22.04:

```
git clone https://github.com/pederwahlberg/ECRANS.git ~/validation
cd ~/validation
bash INSTALL_AWS.sh                   # ESI OpenFOAM + Python pins, ~10 min
git checkout v1.0                     # release tag at arXiv submission
cp config/letkf_params_runE_v50.yaml config/letkf_params.yaml
bash run_all.sh                       # full pipeline, ~3.4 h wallclock
cat results/summary.json | python3 -m json.tool
# expected: mean_rmse ≈ 0.04715 (Run E with ρ_faktor = 0.2843), monotone over 12
cycles.
```

The v1.0 release tag is fixed at arXiv submission; future updates of the validation branch are committed downstream of that tag and do not affect reproducibility of the v1.0 results. The

expected per-cycle RMSE trajectory and the cycle-12 mean (0.04715) are documented in Supplementary Material §S2.4.

C.6 Replication and seed independence

To verify that the σ = 0.30 result is not seed-dependent, repeat the recipe of §C.5 with config/letkf_params_runE2_v50.yaml, which differs from the Run E configuration only in the random-number seed (123 instead of 42). The expected mean RMSE is 0.05713 (Run E2), with sample standard deviation 0.007 across the (Run E, Run E2) pair, reflecting RANS-baseline variability across decomposition seeds; both runs converge monotonically over the 12 LETKF assimilation cycles. The σ = 0.05 spread-damping failure mode (Run D-clean v2) and the σ = 0.20 plateau confirmation (Run F) are similarly reproducible from config/letkf_params_runD_clean_v50.yaml and config/letkf_params_runF_v50.yaml respectively; configurations and expected outcomes are tabulated in Supplementary Material §S2.1, with the full per-cycle trajectories in §S2.4.

References

Anderson, J. L. (2009). Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus A 61, 72–83.

Bishop, C. H., Etherton, B. J. & Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. Mon. Wea. Rev. 129, 420–436.

Buehner, M., Morneau, J. & Charette, C. (2013). Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. Nonlinear Processes in Geophysics 20, 669–682. [ensemble-variational framework]

Breuer, M., Peller, N., Rapp, C. & Manhart, M. (2009). Flow over periodic hills — numerical and experimental study in a wide range of Reynolds numbers. Comput. Fluids 38, 433–457.
Brezzi, F. & Fortin, M. (1991). Mixed and Hybrid Finite Element Methods. Springer.

Colbrook, M. J., Ayton, L. J. & Szőke, M. (2023). Residual dynamic mode decomposition: robust and verified Koopmanism. J. Fluid Mech. 955, A21.

Colbrook, M. J. & Townsend, A. (2024). Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems. Comm. Pure Appl. Math. 77, 221–283.

Colbrook, M. J., Li, Q., Raut, R. V. & Townsend, A. (2024). Beyond expectations: residual dynamic mode decomposition and variance for stochastic dynamical systems. Nonlinear Dynamics 112, 2037–2061.

Di Leoni, P. C., Mazzino, A. & Biferale, L. (2020). Synchronization to big data: nudging the Navier–Stokes equations for data assimilation of turbulent flows. Phys. Rev. X 10, 011023.

Edeling, W. N., Cinnella, P., Dwight, R. P. & Bijl, H. (2014). Bayesian estimates of parameter variability in the k–ε turbulence model. J. Comput. Phys. 258, 73–94.

Fossella, F., Biferale, L., Carrassi, A., Cencini, M. & Gupta, V. (2026). Multiscale data assimilation in turbulent models. Phys. Rev. E 113, 024208.

Foures, D. P. G., Dovetta, N., Sipp, D. & Schmid, P. J. (2014). A data-assimilation method for Reynolds-averaged Navier–Stokes-driven mean flow reconstruction. J. Fluid Mech. 759, 404–431.

Hunt, B. R., Kostelich, E. J. & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Physica D 230(1–2), 112–126.

Labahn, J. W., Wu, H., Harris, S. R., Coriton, B., Frank, J. H. & Ihme, M. (2020). Ensemble Kalman filter for assimilating experimental data into large-eddy simulations of turbulent flows. Flow Turbul. Combust. 104, 861–893.

Jamal, T., Shobayo, O., Walters, D. K. & Bhushan, S. (2025). Static and dynamic time filtering techniques for hybrid RANS-large eddy simulation of non-stationary turbulent flows. J. Fluids Eng. 147, 081502.

Kelly, D. T. B., Law, K. J. H. & Stuart, A. M. (2014). Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time. Nonlinearity 27(10), 2579.

Krank, B., Kronbichler, M. & Wall, W. A. (2018). Direct numerical simulation of flow over periodic hills up to Re_H = 10 595. Flow Turbulence Combust. 101, 521–551.
Krishnanunni, C. G., Wittmer, J., Bui-Thanh, T. & Nguyen, Q. P. (2026). A new look at the ensemble Kalman filter for inverse problems: duality, non-asymptotic analysis and convergence acceleration. arXiv:2601.17305.

Moser, R. D., Kim, J. & Mansour, N. N. (1999). Direct numerical simulation of turbulent channel flow up to Re_τ = 590. Phys. Fluids 11(4), 943–945.

Lienhart, H. & Becker, S. (2003). Flow and turbulence structure in the wake of a simplified car model. SAE Paper 2003-01-0656.

McConkey, R. et al. (2026). The Closure Challenge: a benchmark task for machine learning in turbulence modelling. arXiv:2603.28884.

Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. AIAA J. 32, 1598–1605.

Menter, F. R. (2018). Stress-blended eddy simulation (SBES) — a new paradigm in hybrid RANS-LES modeling. In Proc. 6th Symp. Hybrid RANS-LES Methods (HRLM), Strasbourg, France. Notes on Numerical Fluid Mechanics and Multidisciplinary Design, vol. 137, pp. 27–37. Springer.

Menter, F. R., Lechner, R., Matyushenko, A. (2019). Best practice: Generalized k–ω two-equation turbulence model in ANSYS CFD (GEKO). ANSYS Technical Report.

Menter, F. R., Hüppe, A., Flad, D. et al. (2024). Large eddy simulations for the Ahmed Car at 25° slant angle at different Reynolds numbers. Flow Turbulence Combust. 112, 321–343.

Monk, P. (2003). Finite Element Methods for Maxwell's Equations. Oxford University Press.

Mons, V. et al. (2016). Reconstruction of unsteady viscous flows using data assimilation schemes. J. Comput. Phys. 316, 255–280.

Hasegawa, Y., Idomura, Y. & Onodera, N. (2024). GPU-enabled ensemble data assimilation for mesh-refined lattice Boltzmann method. EPJ Web of Conferences 302, 03005 (SNA+MC 2024).

Patel, Y., Mons, V., Marquet, O. & Rigas, G. (2024). Turbulence model augmented physics-informed neural networks for mean-flow reconstruction. Phys. Rev. Fluids 9, 034605.

Poletto, R., Craft, T. & Revell, A. (2013). A new divergence free synthetic eddy method for the reproduction of inlet flow conditions for LES. Flow Turbulence Combust. 91, 519–539.

Raissi, M., Perdikaris, P. & Karniadakis, G. E. (2019). Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.

Rapp, C. & Manhart, M. (2011). Flow over periodic hills: an experimental study. Exp. Fluids 51, 247–269.

Rumsey, C. L., Gatski, T. B., Sellers, W. L. III, Vatsa, V. N. & Viken, S. A. (2006). Summary of the 2004 Computational Fluid Dynamics Validation Workshop on Synthetic Jets. AIAA Journal 44(2), 194–207.

Serra, M. & Haller, G. (2016). Objective Eulerian coherent structures. Chaos 26, 053110.

Shur, M. L., Spalart, P. R., Strelets, M. K. & Travin, A. K. (2008). A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities. Int. J. Heat Fluid Flow 29, 1638–1649.

Singh, A. P. & Duraisamy, K. (2016). Using field inversion to quantify functional errors in turbulence closures. Phys. Fluids 28, 045110.

Spalart, P. R. & Allmaras, S. R. (1992). A one-equation turbulence model for aerodynamic flows. AIAA Paper 92-0439.

Spalart, P. R. et al. (1997). Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach. Proc. 1st AFOSR Conf. DNS/LES.

Villiers, R., Mons, V., Sipp, D., Lamballais, E. & Meldi, M. (2025). Enhancing Unsteady Reynolds-Averaged Navier-Stokes modelling from sparse data through sequential data assimilation and machine learning. Flow Turbulence Combust. 115, 991–1029.

Wang, M. & Zaki, T. A. (2022). Synchronization of turbulence in channel flow. J. Fluid Mech. 943, A4.

Woodbury, M. A. (1950). Inverting modified matrices. Memorandum Report 42. Statistical Research Group, Princeton.

Xiao, H., Wu, J. L., Wang, J. X., Sun, R. & Roy, C. J. (2016). Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier–Stokes simulations: a data-driven, physics-informed Bayesian approach. J. Comput. Phys. 324, 115–136.

Yoshida, T., Yamaguchi, J. & Kaneda, Y. (2005). Regeneration of small eddies by data assimilation in turbulence. Phys. Rev. Lett. 94, 014501.

Miyoshi, T. (2011). The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. Mon. Wea. Rev. 139, 1519–1535.

Spalart, P. R., Deck, S., Shur, M. L., Squires, K. D., Strelets, M. K. & Travin, A. (2006). A new version of detached-eddy simulation, resistant to ambiguous grid densities. Theor. Comput. Fluid Dyn. 20, 181–195.