S Simon Kuhn

Journal Pre-proof



Data-driven Transport Modeling for Tracer Injection and Mixing Time Screening in Stirred Tank Reactors

Asif Shaik, Swantje Pietsch-Braune, Ryan Rautenbach, Christian Weiland, Michael Schlüter, Stefan Pirker, Stefan Heinrich

PII: S0009-2509(26)02191-3
DOI: https://doi.org/10.1016/j.ces.2026.125476
Reference: CES 125476

To appear in: *Chemical Engineering Science*

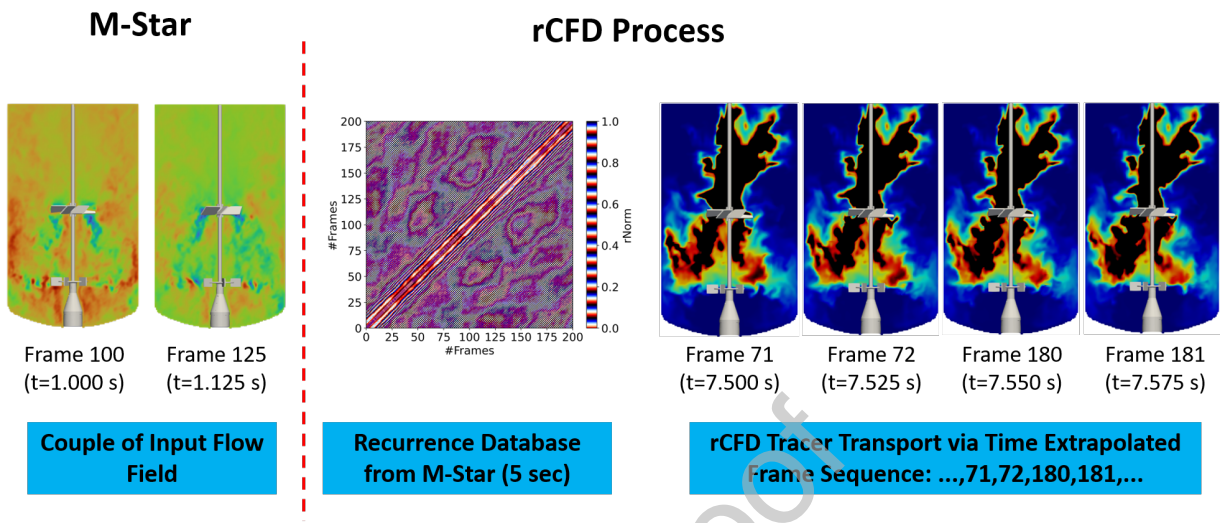
Received date: 29 July 2026
Revised date: 19 October 2026
Accepted date: 28 October 2026

Please cite this article as: Asif Shaik, Swantje Pietsch-Braune, Ryan Rautenbach, Christian Weiland, Michael Schlüter, Stefan Pirker, Stefan Heinrich, Data-driven Transport Modeling for Tracer Injection and Mixing Time Screening in Stirred Tank Reactors, *Chemical Engineering Science* (2026), doi: https://doi.org/10.1016/j.ces.2026.125476

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Graphical abstract



Journal Pre-proof

Journal Pre-proof

Highlights

- First implementation of transport-based rCFD for mixing time in STR
- Single LB-LES flow field is the sole input to the recurrence database
- One second of process time simulated in 3 s of wall-clock time
- Design space with 90 tracer injection location analysed in 2.25 hours

Journal Pre-proof

Journal Pre-proof

Data-driven Transport Modeling for Tracer Injection and Mixing Time Screening in Stirred Tank Reactors

Asif Shaik[a], Swantje Pietsch-Braune[a], Ryan Rautenbach[b], Christian Weiland[b], Michael Schlüter[b], Stefan Pirker[c], Stefan Heinrich[a]

[a]*Institute of Solids Process Engineering and Particle Technology, Hamburg University of Technology, Denickestraße 15, Hamburg, 21073, Germany*
[b]*Institute of Multiphase Flows, Hamburg University of Technology, Eißendorfer Straße 38, Hamburg, 21073, Germany*
[c]*Department of Particulate Flow Modelling, Johannes Kepler University, Altenbergerstraße 66a, Linz, 4040, Austria*

Journal Pre-proof

Abstract

Transient CFD of liquid mixing in stirred tanks is too expensive for systematic design studies, and mixing time predictions deviate from experiment even at scale-resolving resolution. Transport-based recurrence CFD (rCFD) decouples scalar transport from the flow solve, replaying a short pre-computed flow database through cell-to-cell displacement maps, so that transport level corrections can be applied without repeating the momentum solution. This is exploited here for the 95% mixing time in a 30 L dual-impeller stirred tank reactor, using a 4.5 s LB-LES flow database from M-Star as the sole input. The rCFD first results inherit the over-prediction of its input flow field (33.1 ± 1.4 s, against 30.1 ± 3.0 s for M-Star and 28.1 ± 2.7 s for Ansys Fluent at comparable spatial resolution, and 25.0 ± 3.0 s measured by conductivity probes in experiment part). Calibrating a single face-swap diffusion parameter in rCFD transport model against one probe measurement brings the prediction to 24.5 ± 1.0 s, recovering experimental agreement without re-running the flow simulation. The rCFD simulation requires approximately 3 s of wall-clock time per second of process time (423× M-Star, 1200× Ansys Fluent) and is almost insensitive to a fivefold increase in extrapolation time step. Ninety injection locations spanning radius, height, and azimuth are therefore screened from the single database in 2.25 h, against estimates of 40 and 113 days for the reference solvers, resolving a 3.6-fold spread in mixing time (13.6–48.8 s) in which injection height and baffle-relative azimuth dominate. The main objective of this work is a fast, spatially and temporally resolved description of scalar transport; the 95% mixing time is used as the quantitative metric for validating it against sparse experimental data.

Keywords: Mixing time, Scalar transport, Data-driven modelling, Time extrapolation, Computational speed-up

1. Introduction

Mixing is a fundamental unit operation across numerous manufacturing sectors, playing a critical role in liquid-phase reactions, heat and mass transfer, emulsification, and gas–liquid

Journal Pre-proof

contacting [1, 2, 3]. In the chemical and process industries, mechanically agitated vessels are essential for producing specialty chemicals, polymers, and conducting heterogeneous reactions [4, 5, 3]. The pharmaceutical and bio-pharmaceutical sectors rely heavily on liquid mixing for fermentation, cell culture, and bioreactor media preparation, where mechanically stirred tanks dominate batch production [1, 6, 3]. Additionally, mixing operations are integral to petroleum processing for drilling fluids and gas treating, as well as pulp and paper manufacturing for pulp blending and bleaching [3].

Liquid–liquid and gas–liquid mixing operations employ a diverse range of equipment, from static mixers and micro-mixers for fine-scale applications [7, 8] to bubble columns and gas-lift reactors for large-scale gas–liquid contacting [9, 10, 11, 12]. Among these, mechanically agitated stirred tank reactors (STRs) remain the preferred choice for both chemical and biochemical processes due to their operational flexibility and well-characterized hydrodynamics [13, 14, 15, 16], and form the basis of the present work.

Achieving adequate mixing in an STR is critical, as it directly governs reaction yield, product selectivity, and process consistency [17, 18]. Insufficient mixing leads to the gradients, which eventually leads to poor yield and selectivity [19, 20]. Process parameters such as impeller speed, feed location, and reactor geometry significantly influence the transition from segregated to well-mixed conditions, yet predicting this transition remains challenging without robust quantitative metrics [21, 22]. Therefore, accurate characterization of mixing state through residence time distribution analysis, mixing time measurements, and non-ideality indices is essential not only for optimizing existing processes but also for establishing reliable design criteria for industrial-scale reactors.

Mixing time, defined as the duration required for a tracer to achieve uniform distribution within a specified tolerance band, can be determined through both experimental and computational approaches [23, 24]. Experimentally, tracer injection methods coupled with conductivity, pH, or optical measurements provide direct quantification of global mixing time, while advanced imaging techniques enable extraction of local mixing time distributions to identify poorly mixed regions [24, 25]. Computational fluid dynamics (CFD) complements these measurements by resolving three-dimensional velocity and concentration fields, offering detailed insight into flow structures, dead zones, and recirculation patterns that are difficult to capture experimentally [26, 27]. Furthermore, approaches such as Mean Age Theory (MAT) provide computationally efficient alternatives to full transient simulations while maintaining good agreement with experimental observations [28]. Passive tracer transport on frozen mean flow fields offers a computationally efficient approach for mixing characterization, wherein the time-averaged flow solution is computed once and subsequently used for tracer simulations. Wu et al. [28] showed that frozen-flow transport requires only a fraction of the computational effort of fully transient simulations, but also that the resulting mixing times can deviate considerably from experiment. This is particularly pronounced in multi-impeller vessels: for dual-impeller tanks, mixing times obtained on frozen Reynolds-Averaged Navier-Stokes (RANS) (MRF, k-ϵ) flow fields exceeded the measured values by approximately 25-90%, whereas Large Eddy Simulation (LES) gave the closest agreement with experiment, albeit at roughly 50 times the computational cost [29].

Despite their complementary strengths, both experimental and computational approaches face significant limitations in characterizing mixing in stirred tank reactors. Experimental techniques are often constrained by limited spatial resolution, optical access requirements,

Journal Pre-proof

and probe interference, making comprehensive three-dimensional flow characterization particularly challenging in opaque or gas-sparged systems [30]. RANS-based simulations of multi-impeller STRs using the standard k-ϵ model have been shown to over-predict the mixing time [29]. This has been attributed to the averaged representation of all turbulence fluctuations, including the large eddies responsible for mixing; LES, which resolves these large-scale structures, achieved considerably better agreement with experiment [29]. Notably, macro-instabilities, which can account for up to 30% of the fluctuating kinetic energy in this region, are not resolved by frozen-flow Moving Reference Frame (MRF) approaches. Scale-resolving methods such as LES recover these dynamics but require correspondingly fine spatial resolution and small time steps; consequently, simulating even a few seconds of physical mixing time can require days of computational effort per operating condition, rendering systematic parametric studies impractical in feasible time period for industrial applications [31, 32]. This computational burden presents a substantial barrier to using transient CFD as a routine design tool, highlighting the need for accelerated simulation strategies that preserve predictive fidelity while dramatically reducing computation time.

To address these computational constraints, Lichtenegger and Pirker [33] developed recurrence-based CFD (rCFD), a time-extrapolation framework that capitalizes on the inherent pseudo-periodicity observed in many industrial flows. Building upon the recurrence analysis concepts established by Eckmann et al. [34], this approach identifies recurring flow patterns within a brief monitoring window (τ_{rec}) of a pre-computed flow simulation, subsequently constructing long-term flow evolution by concatenating these recurring segments through recurrence matrices rather than performing continuous, computationally intensive calculations [33, 35]. The rCFD framework encompasses two distinct implementations: flow-based rCFD, which reconstructs complete flow field sequences for subsequent solution of transport equations [35, 36, 37], and transport-based rCFD, which directly addresses scalar transport via Lagrangian memory shift operations [38, 39, 40]. The present study employs transport-based rCFD, wherein continuous flow displacements are mapped onto discrete cell-to-cell memory shift patterns [41]. Governed by recurrence statistics, these shift operations propagate scalar quantities such as species concentration or temperature without recalculating the underlying flow fields, yielding computational acceleration of several orders of magnitude at a fraction of the cost of the underlying flow simulation.

Several accelerated simulation strategies have been proposed to address the computational burden of transient CFD for mixing applications. Reduced Order Models (ROMs), such as those based on Proper Orthogonal Decomposition, achieve speed-up by projecting the governing equations onto a low-dimensional subspace constructed from high-fidelity snapshots [42]. While ROMs can be effective for problems with smooth parametric dependence, they require a dedicated offline construction phase for each new geometry or operating condition, and their accuracy degrades when the query lies outside the training manifold. Machine-learning surrogate approaches, such as coarse-grid CFD with data-driven error correction [43], offer flexible acceleration but similarly depend on representative training datasets and do not inherently preserve the physical transport mechanisms of the underlying flow. Transport-based rCFD occupies a complementary position: it requires no training phase beyond a single pre-computed flow field, preserves convective transport physics directly through precomputed cell-to-cell displacement mappings, and is immediately deployable for parametric scalar transport studies once the recurrence database is established. Its primary

Journal Pre-proof

constraint that predictions are anchored to the operating condition represented in the input database is shared in spirit by all data-driven acceleration methods, but the absence of a separate training infrastructure makes rCFD particularly practical for industrial applications where a validated flow field is already available.

While transport-based rCFD has demonstrated promising agreement with full CFD simulations [38, 44], its potential extends beyond mixing characterization to coupling with reaction kinetics models for rapid process simulation and parametric optimization. However, rigorous validation against experimental measurements remains essential before widespread adoption in academia or industry.

The present work addresses this need by employing transport-based rCFD to investigate mixing behavior in an STR, systematically comparing predictions against experimental observations. Two open questions in the development and application of transport-based rCFD, with liquid mixing in stirred tank reactors as the demonstration case. First, rCFD mixing time predictions have not previously been validated against experimental measurements. Second, parametric injection location optimization in STRs requires many independent scalar transport simulations when performed with traditional CFD tools; whether rCFD can deliver equivalent insight at substantially reduced computational cost has not been demonstrated in liquid mixing systems. Both questions are investigated in a 30 L dual-impeller STR. Validation is performed against M-Star (v3.12.21) Lattice Boltzmann simulations conducted in the present work and experimental conductivity probe measurements, as well as Ansys Fluent (version 2024 R1) finite-volume predictions reported by Rautenbach et al. [45] for the same reactor configuration, providing a multi-method benchmarking framework.

2. Methods

The section provide the stirred tank reactor configuration and operating conditions, followed by the M-Star Lattice Boltzmann simulation setup, the rCFD framework, data conversion from LBM to rCFD format, and the monitoring strategy and mixing time criterion applied throughout all simulations are defined.

2.1. Stirred Tank Reactor Configuration

The present numerical study is based on a geometrically fixed, mechanically agitated, fully baffled STR with a nominal working volume of $V_{\mathrm{reac}} = 30$ L. The reactor geometry, impeller configuration, probe arrangement, and tracer injection strategy are identical to those employed in the experimental and numerical study of Rautenbach et al. [45], whose experimentally measured mixing times serve as the validation reference for the present work. A schematic overview of the configuration is provided in Figure 1.

The vessel is equipped with a centrally mounted dual-impeller system consisting of a pitched blade turbine (PBT) and a Rushton turbine (RT). The tank is filled with deionized water ($\rho_{\mathrm{f}} = 998.2\ \mathrm{kg\,m^{-3}}$) to the nominal working height and operated under single-phase, isothermal flow conditions at a constant impeller rotational speed of $n = 200$ rpm, ensuring direct comparability between the experimental reference data and the numerical results presented here.

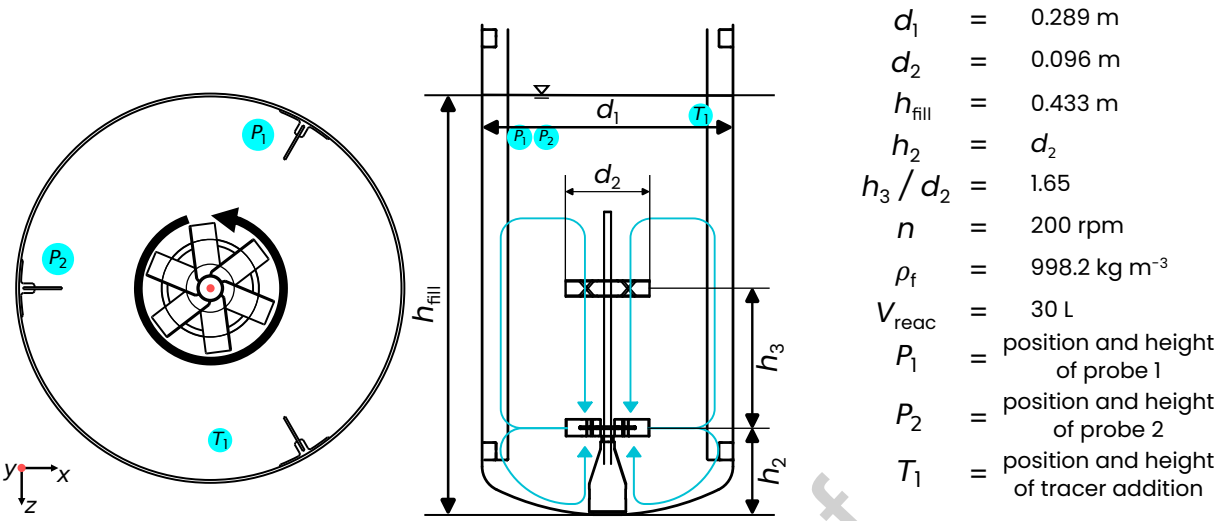


Figure 1: Schematic of the 30 L stirred tank reactor configuration adopted in the present study, showing the impeller arrangement, tracer injection location, probe positions, and key geometric dimensions as per Rautenbach et al. [45]

2.2. M-Star (LBM) Simulation

LES of the stirred tank flow were performed using the commercial Lattice Boltzmann solver M-Star. The numerical configuration closely follows the setup established in Rautenbach et al. [45].

The Lattice Boltzmann simulations employ a D3Q19 lattice model with a single-relaxation-time Bhatnagar–Gross–Krook (BGK) collision operator and explicit time integration. Turbulence closure is provided by the Smagorinsky–Lilly subgrid-scale model with a Smagorinsky constant coefficient $C_S = 0.1$, and a fixed Courant number of Co = 0.1. The impeller geometry is represented using the voxelized immersed boundary approach implemented in M-Star, in which local solid volume fractions are assigned to lattice voxels and updated dynamically to account for wall motion.

In the present work, M-Star simulations are conducted at a resolution of $4.7 \cdot 10^6$ lattice nodes, which serves as the sole source of flow field data for the rCFD time-extrapolation framework. The M-Star simulation is run until the flow field reaches a statistically stationary state, after which the instantaneous three-dimensional velocity field is recorded over a window from $t = 25$ s to $t = 29.5$ s, at a sampling frequency of $f_{\text{out}} = 200$ Hz (temporal resolution $\Delta t_{\text{CFD}} = 5.0 \cdot 10^{-3}$ s), yielding 900 velocity field snapshots. These snapshots constitute the flow-field database from which the rCFD recurrence statistics are constructed. All M-Star simulations were executed in single precision on a single NVIDIA A100-PCIE-40GB GPU.

2.3. Recurrence CFD (rCFD)

This section provides a structured description of the transport-based rCFD methodology as implemented in the present work. For comprehensive theoretical details, the reader is referred to the work of Lichtenegger and Pirker [35, 38].

rCFD (version 26.3) is a data-driven framework that enables long-timescale prediction of transport processes without explicitly solving the Navier–Stokes equations during the extrapolation phase. It relies on a precomputed flow field database exhibiting pseudo-repetitive

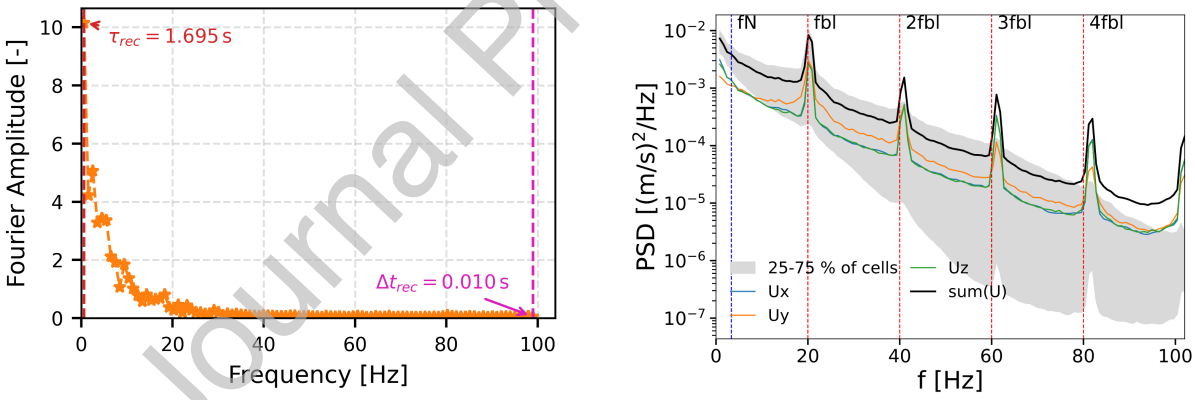
dynamics, from which the long-term system trajectory is reconstructed by reusing dynamically similar flow states.

In contrast to conventional CFD, where scalar transport is governed by partial differential equations solved at each time step, rCFD employs a displacement-based formulation in which scalar quantities are advected through precomputed cell-to-cell (C2C) mappings derived from the velocity field. Further details of the C2C construction are provided in Section Database Generation: Cell-to-Cell Transport Mapping.

2.3.1. Recurrence Statistics and Time Scales

The applicability of rCFD requires that the underlying flow exhibits pseudo-recurrent dynamics over a finite monitoring window. Two characteristic time scales govern the construction of the recurrence database: the recurrence horizon τ_{rec}, which must be long enough to capture at least one complete cycle of the slowest relevant flow structure, and the temporal resolution Δt_{rec}, which must be fine enough to resolve the fastest dynamically significant motion.

Both scales are determined from frequency-domain analysis of transient velocity signals. Welch power spectral densities (Hann window, 50 % overlap) of the three velocity components are computed for all cells of a horizontal plane through the PBT impeller ($y = 0.04$ m, $\approx$ 9 800 cells) and for approximately 500 probe cells distributed across the domain (Figure 2a). A detailed characterization of the flow dynamics at these locations is presented in Section Nonlinear Dynamics and Flow Recurrence Analysis.



(a) Dominant frequency observed at a cell away from impellers near wall of the STR.

(b) Power spectral density of a 2D plane at mid-section of ruhston turbine at the bottom of the STR.

Figure 2: Characteristic time scales of rCFD.

At the high-frequency end, the spectra of the impellers exhibit a distinct peak at the blade-passing frequency $f_{\mathrm{bl}} = N_b N/60 = 20$ Hz ($N_b = 6$ and $N = 200$ rpm), followed by its harmonics $2f_{\mathrm{bl}}$ to $4f_{\mathrm{bl}}$ with monotonically decreasing amplitude. No peak occurs at the shaft frequency $f_N = N/60 = 3.33$ Hz, as expected for an impeller with identical blades. The blade-passing signature is confined to the impeller swept region and its discharge stream and decays with distance from the impeller. At the low-frequency end, the spectra are governed by a slow pseudo-periodic motion at $f_{\mathrm{min}} = 0.59$ Hz.

Journal Pre-proof

Journal Pre-proof

The smallest characteristic frequency defines the minimum recurrence horizon, $\tau_{\mathrm{rec}} \geq 1/f_{\mathrm{min}} = 1.70$ s. The temporal resolution must resolve the fastest dynamically significant motion, the blade passage; by the Nyquist criterion this requires $\Delta t_{\mathrm{rec}} < 1/(2f_{\mathrm{bl}}) = 0.025$ s. The chosen $\Delta t_{\mathrm{rec}} = 0.005$ s samples each blade passage five times and resolves f_{bl} and $2f_{\mathrm{bl}}$. The M-Star flow field database, sampled at $f_{\mathrm{out}} = 200$ Hz ($\Delta t_{\mathrm{CFD}} = 0.005$ s, Nyquist frequency 100 Hz), resolves the blade-passing frequency and its first four harmonics. The actual database window of 4.5 s, approximately $2.65\,\tau_{\mathrm{rec}}$, is used to ensure robust coverage of the pseudo-repetitive flow structure; the adequacy of this window is confirmed by the recurrence quantification analysis presented in Section Nonlinear Dynamics and Flow Recurrence Analysis. The sensitivity of the predictions to the C2C mapping interval is assessed in Appendix A; strict time-step independence is achieved by the propagator-moments formulation of Lumetzberger et al. [41], to which the present implementation will be transferred in future work.

2.3.2. Recurrence Matrix Construction

The similarity between flow states at different time instants is quantified using a normalized recurrence norm. In the present implementation, the velocity magnitude field $\mathbf{u}$ is used as the recurrence quantity. The normalized recurrence measure between two snapshots at times t and t' is defined as:

$$\mathcal{R}(\mathbf{u};\, t, t') = \frac{1}{N^{(\mathbf{u})}} \int_{\Omega} \left\| \mathbf{u}(\mathbf{r}, t) - \mathbf{u}(\mathbf{r}, t') \right\|^2 \, d\mathbf{r}, \tag{1}$$

where Ω represents the whole domain and the normalization factor is:

$$N^{(\mathbf{u})} = \max_{t,\, t'} \int_{\Omega} \left\| \mathbf{u}(\mathbf{r}, t) - \mathbf{u}(\mathbf{r}, t') \right\|^2 \, d\mathbf{r} \tag{2}$$

scales the measure to the interval [0, 1], where zero indicates identical states and unity indicates the maximally dissimilar pair within the database.

Evaluating this measure for all pairs of stored velocity snapshots yields the recurrence matrix $\mathcal{R}(t, t')$. The matrix is symmetric with a zero diagonal, since each snapshot is perfectly similar to itself. Small off-diagonal values identify pairs of dynamically similar flow states. During trajectory reconstruction, described in Section Flow Reconstruction via Recurrence Stitching, the recurrence matrix is queried row-wise: given the current flow state at time t, the next segment is selected as the stored snapshot t' that minimizes $\mathcal{R}(t, t')$, ensuring dynamic consistency between successive segments. The construction of the recurrence database is illustrated in Figure 3.

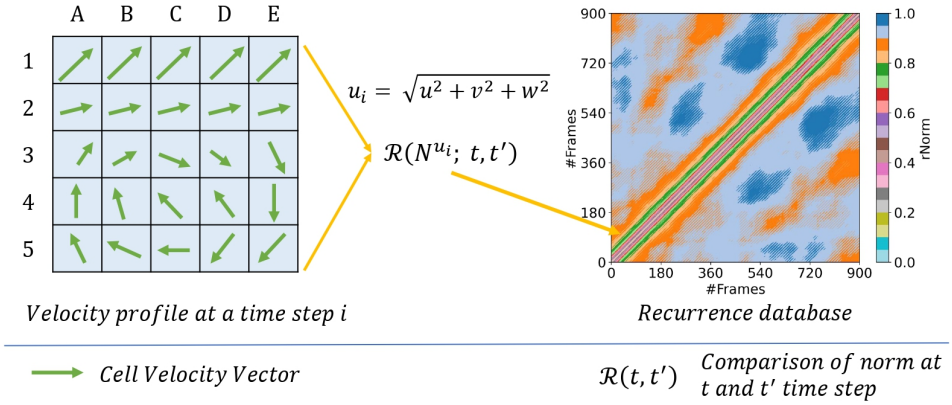


Figure 3: Recurrence matrix $\mathcal{R}(t, t')$ quantifying pairwise velocity field similarity across the M-Star snapshot database; high-recurrence regions identify flow states suitable for frame sequencing.

2.3.3. Database Generation: Cell-to-Cell Transport Mapping

In the present implementation, the cell-to-cell (C2C) transport mappings are constructed using an auxiliary scalar tracking approach within Ansys Fluent, applied to the converted M-Star velocity field. This procedure estimates fluid displacement directly from convected scalar fields, without requiring numerical integration of Lagrangian trajectories.

Three passive scalar fields (ϕ_x, ϕ_y, ϕ_z) are initialized in each computational cell to store the initial spatial coordinates of that cell in the respective directions. These scalars are treated as user-defined scalars (UDS) with negligible diffusivity ($\approx 10^{-10}$ m^2 s^{-1}), ensuring that their evolution is governed purely by convection.

As the simulation advances to the next time step (Δt_{CFD}), the scalar fields are transported with the local velocity field. Since each scalar stores the initial cell coordinate, the difference between the current scalar value and its initial value yields directly the displacement of fluid originating from that cell. The destination cell is then identified as the cell whose centroid is nearest to the displaced coordinate, establishing a discrete source-to-destination mapping that defines the convective transport operator for that time step.

This process is illustrated in Figure 4. The procedure is repeated for every cell and every stored velocity snapshot, producing a database of C2C mappings indexed by time. Each entry records the source-destination cell pair, which is applied directly during the time-extrapolation phase without further flow field computation.

Journal Pre-proof

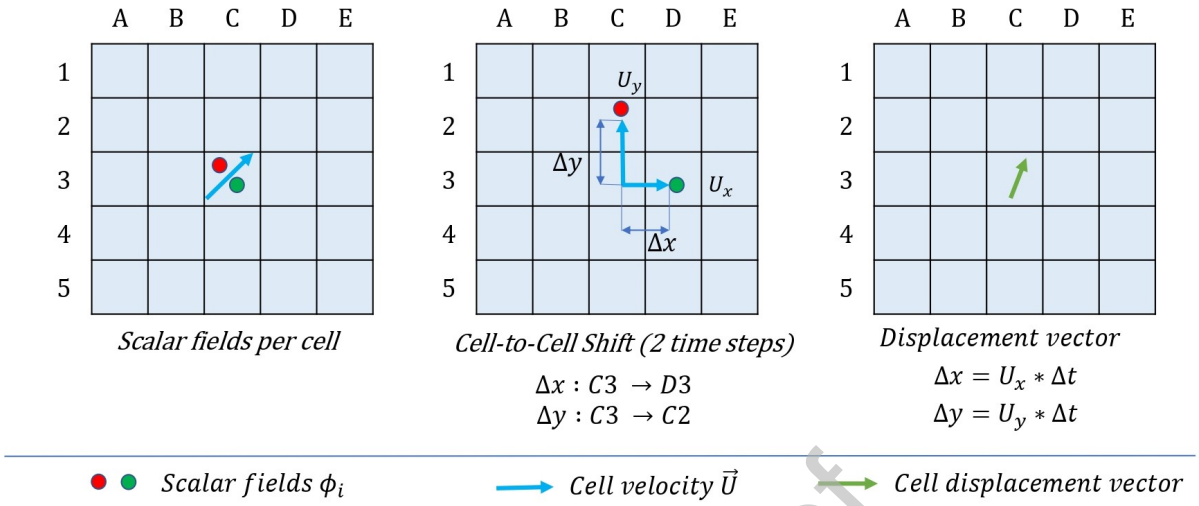


Figure 4: Cell-to-cell (C2C) mapping based on scalar tracking. Passive scalar fields store initial positions and are convected with the flow, enabling reconstruction of displacement between source and destination cells over a time interval Δt_{CFD}.

Unlike direct velocity-based displacement methods, the scalar tracking approach inherently accounts for the integrated flow history over the snapshot interval Δt_{CFD} and avoids the interpolation errors associated with explicit trajectory integration. The resulting mappings represent purely advective transport; diffusive effects are introduced separately through the face-based exchange mechanism described in the following section.

2.3.4. Numerical Diffusion via Face-Based Concentration Exchange

The purely advective C2C transport does not account for sub-grid diffusive mixing absent from the discrete displacement operators. To represent this effect and maintain a physically consistent scalar field, a face-based concentration exchange mechanism is introduced at each rCFD time step.

This mechanism is governed by the dimensionless parameter f_{swap}, which controls the fraction of the local concentration difference exchanged between neighboring cells per time step. It does not correspond directly to a molecular or turbulence diffusivity, it carries no physical units of diffusion but serves as a calibration parameter that represents the net effect of sub-grid mixing unresolved by the C2C displacement field. Its value is determined by calibration against experimental observations, as described in Section Mixing Time Sensitivity to Temporal Injection.

The exchange is applied across each internal face shared between neighboring cells c_0 and c_1. The effective exchange volume is bounded by the smaller of the two adjacent phase volumes:

$$V_{\mathrm{flip}} = \min(V_{c_0}\,\alpha_{c_0},\; V_{c_1}\,\alpha_{c_1})\,, \tag{3}$$

where V denotes cell volume and α the local phase volume fraction. This volume-limiting step ensures that the exchange cannot exceed the physically available fluid volume at the interface.

Journal Pre-proof

The driving concentration difference is taken as half the interfacial difference:

$$\Delta c = \frac{c_{\text{high}} - c_{\text{low}}}{2}, \tag{4}$$

where c_{high} and c_{low} are the concentrations in the higher and lower-concentration cells, respectively. The factor of $\frac{1}{2}$ ensures that the exchanged mass moves both cells toward their interfacial mean without overshooting it. If the full difference $c_{high} - c_{low}$ were transferred (for $f_{swap} = 1$ and equal cell volumes), the concentrations of the two cells would simply be swapped, displacing the local gradient rather than smoothing it.

The mass exchanged across the face per time step is:

$$\Delta m_{\text{diff}} = f_{\text{swap}}\, V_{\text{flip}}\, \Delta c. \tag{5}$$

This mass is removed from the higher-concentration cell and added to the lower-concentration cell. The updated cell concentration is:

$$c_i^{n+1} = c_i^n + \frac{\sum_{\text{faces}} \Delta m_{\text{diff}}}{V_i\, \alpha_i}, \tag{6}$$

where the sum runs over all internal faces of cell i, with sign determined by the local concentration gradient direction. This formulation guarantees bounded scalar transport, controlled gradient smoothing, and strict local mass conservation.

It is noted that the face-swap calibration described above is specific to the present implementation of transport-based rCFD. In a recent extension of the framework, Lumetzberger et al. [41] incorporated diffusion directly into the C2C mapping by tracking the zeroth, first, and second moments of the scalar fields originating from each cell, yielding a propagator kernel that carries both convective and diffusive contributions within the mapping itself.

2.3.5. Flow Reconstruction via Recurrence Stitching

Long-time flow evolution is reconstructed by sequentially connecting dynamically similar segments from the precomputed database. At each step, a segment of duration $\Delta t_i = \Delta t_{\text{rCFD}}$ is selected and the corresponding C2C mapping is applied to advance the scalar field. Starting from an initial database time $t_0^{(b)}$, which corresponds to the tracer injection instant, the next segment start is chosen as:

$$t_{i+1}^{(b)} = \arg\min_{t'} \mathcal{R}(t_i^{(e)}, t'), \tag{7}$$

where $t_i^{(e)} = t_i^{(b)} + \Delta t_{\text{rCFD}}$ is the end of the current segment. This selects the stored snapshot whose flow state most closely resembles the current end state, ensuring dynamic consistency between successive segments. The reconstructed trajectory is not required to be temporally continuous in the database, the algorithm may jump forward or backward among stored snapshots but each transition is constrained to be dynamically consistent by the recurrence criterion. This procedure generates an arbitrarily long extrapolated trajectory by reusing the finite database without enforcing strict periodicity. The concept is illustrated in Figure 5.

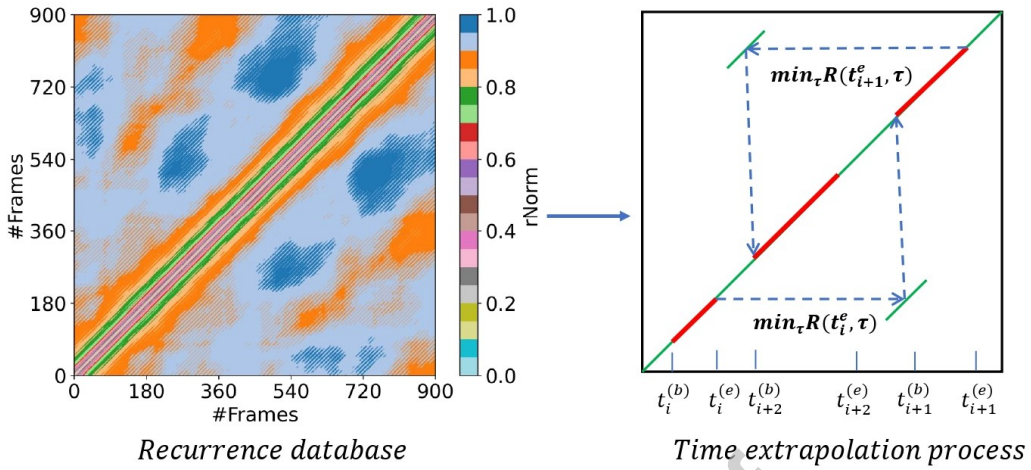


Figure 5: Recurrence-based frame stitching: sequential selection of database snapshots to reconstruct a continuous flow trajectory for scalar transport extrapolation.

2.3.6. Mass Conservation

Cell-to-cell mappings do not in general guarantee strict local mass conservation, since the discrete source-to-destination assignment may introduce small numerical drift in the total scalar inventory. A global conservation check is therefore applied after each C2C displacement step to enforce scalar balance before the face-swap diffusion exchange is performed.

The conservation criterion requires that the sum of all species concentrations over the full computational domain remains equal to the total fluid mass present at initialization. For a domain initially containing a fluid mass m_{total}, this condition is expressed as:

$$\sum_i c_i^n \, V_i \, \alpha_i = m_{\text{total}} \quad \forall \, n, \tag{8}$$

where c_i^n is the concentration in cell i at step n, V_i is the cell volume, and α_i is the phase volume fraction. In multiphase configurations, species belonging to each phase are required to satisfy the same condition independently, ensuring phase-wise as well as global conservation. In the present single-phase study, the condition reduces to conservation of the total injected tracer mass across all cells at every extrapolation step.

If a drift from the reference total is detected, cell concentrations are uniformly rescaled to restore the global balance, ensuring that no scalar mass is artificially created or destroyed during the extrapolation. The correction is applied globally across the full computational domain; extension to sub-domain level, applicable in configurations with internal source or sink terms, is supported by the framework but not employed in the present study.

2.3.7. Time Extrapolation of Scalar Transport

The time extrapolation phase combines the operations established in the preceding sections into a repeated per-step cycle. At each rCFD time step of duration Δt_{rCFD}:

1. The recurrence stitching algorithm (Section Flow Reconstruction via Recurrence Stitching) selects the C2C mapping from the database whose source flow state most closely matches the current extrapolated state, as quantified by the recurrence matrix $\mathcal{R}(t, t')$.

Journal Pre-proof

2. The selected C2C mapping is applied to displace scalar quantities from source to destination cells, advancing the concentration field by purely advective transport.
3. The face-based diffusion exchange (Section Recurrence CFD (rCFD)) is applied across all internal faces, introducing sub-grid scalar smoothing controlled by f_{swap}.
4. A global mass conservation correction (Section Mass Conservation) adjusts the total scalar inventory to maintain balance.

This cycle repeats until the specified process time is reached. The framework does not impose a constraint that Δt_{rCFD} must equal Δt_{CFD}; the recurrence database is queried at the frequency determined by Δt_{rCFD}, with the corresponding C2C mapping applied at each step. The sensitivity of mixing time predictions to the choice of Δt_{rCFD} is examined in Appendix Section Effect of Time Step Width on Mixing Time and Computational Cost.

As no Navier–Stokes equations are solved at any extrapolation step, the computational cost scales with Δt_{rCFD} rather than with the underlying flow physics. The frame selection step introduces a stochastic element: since the recurrence criterion identifies a best-matching stored state rather than a uniquely determined successor, the sequence of frames assigned across independent runs under identical injection conditions may differ. This produces bounded run-to-run variability in predicted scalar transport histories, the magnitude of which is quantified in Section Mixing Time Sensitivity to Temporal Injection.

2.3.8. Real World Informed Calibration

Data-driven acceleration methods, including rCFD, achieve their computational efficiency by replacing explicit solution of conservation equations with precomputed mappings derived from a reference simulation. This substitution introduces an inherent trade-off: the input flow field cannot capture every physical mechanism at every scale, and the discrete C2C transport operator does not include molecular diffusion or sub-grid scalar mixing by construction. Experimental calibration provides a principled means of compensating for these gaps by incorporating external physical information via experimental observations, theoretical correlations, or known process constraints directly into the simulation framework.

In the present study, this role is fulfilled by the face-swap diffusion parameter f_{swap}. As established in Section Recurrence CFD (rCFD), f_{swap} controls the rate of inter-cell scalar smoothing and compensates for the under-resolution of velocity gradients in the coarse input flow field, which reduces effective turbulence dispersion relative to the physical system, which is not represented in the simulated passive scalar. Calibrating f_{swap} against conductivity probe measurements anchors the rCFD predictions to experimental reality without requiring a repeat of the underlying flow simulation.

This approach is represented schematically in Figure 6, which illustrates the interaction between mechanistic simulation, data-driven extrapolation, and experimental grounding. It is acknowledged that calibration against a single probe location constrains the spatial transferability of the selected value; the implications of this limitation are discussed in Section Mixing Time Sensitivity to Temporal Injection and the Conclusions.

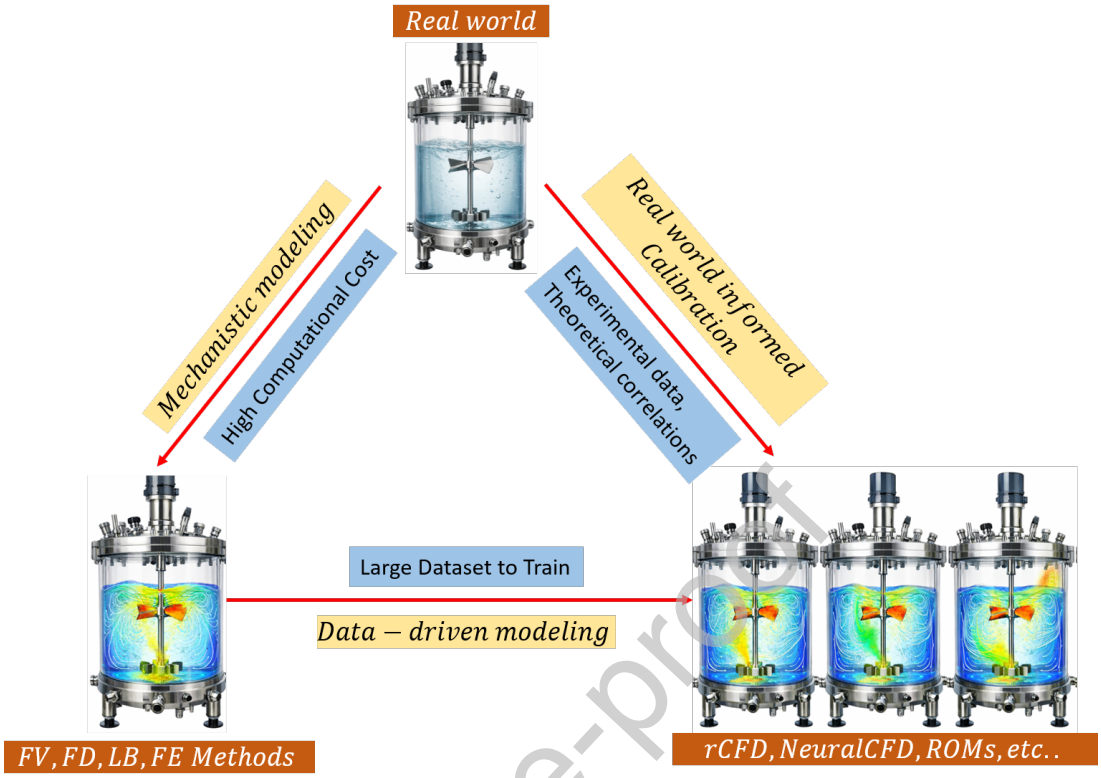


Figure 6: Relationship between physical problem, conventional simulation, and data-driven modelling, illustrating the computational bottleneck that motivates the rCFD framework.

2.4. Data Processing Pipeline

In the conventional rCFD workflow, the C2C mapping database is constructed from velocity fields generated directly by Ansys Fluent. The present work departs from this because the input flow field originates from M-Star, a Lattice Boltzmann solver that stores data on lattice nodes rather than the cell-centered format required by Ansys Fluent. A four-step conversion pipeline was therefore established to bridge the two frameworks; the technical details are provided in Appendix Data Processing Pipeline and Conversion Fidelity, where the fidelity of the node-to-cell conversion is also assessed quantitatively.

The M-Star simulation was run for a total of 30 s, with the final 4.5 s ($t = 25$ to 29.5 s) sampled at 200 Hz, yielding 900 velocity field snapshots for the rCFD database. Tracer mixing is monitored simultaneously at three locations throughout all rCFD simulations: two probe volumes mimicking the experimental conductivity probe positions (probe 1 and probe 2, each comprising approximately 400 computational cells at $r = 0.112$ m near the free surface), and the full computational domain ($\approx 1.8 \cdot 10^6$ cells). The 95% mixing time, $\tau_{m,95}$, is defined as the elapsed time from tracer injection until the concentration at all monitored locations remains within $\pm 5\%$ of the final equilibrium value. The reported $\tau_{m,95}$ is taken as the maximum across all three monitors, ensuring it reflects the last region to reach homogeneity.

Journal Pre-proof

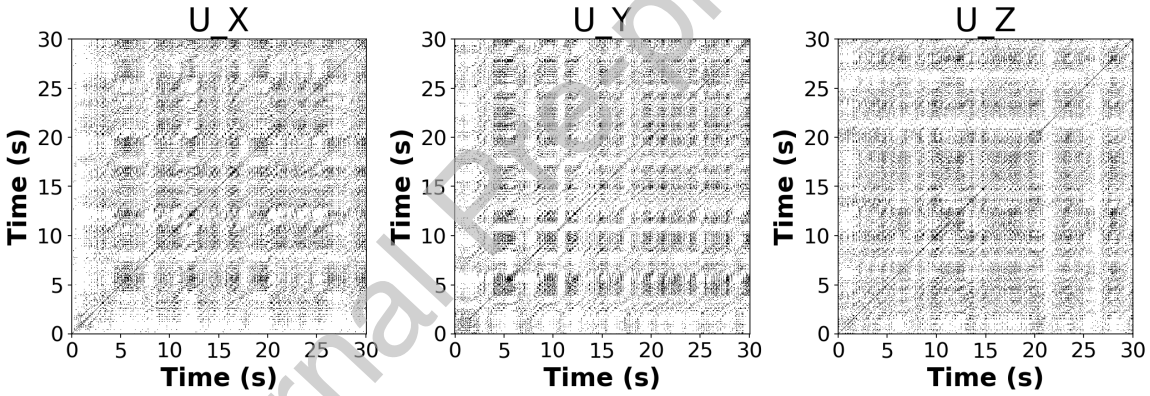
Journal Pre-proof

3. Results and Discussion

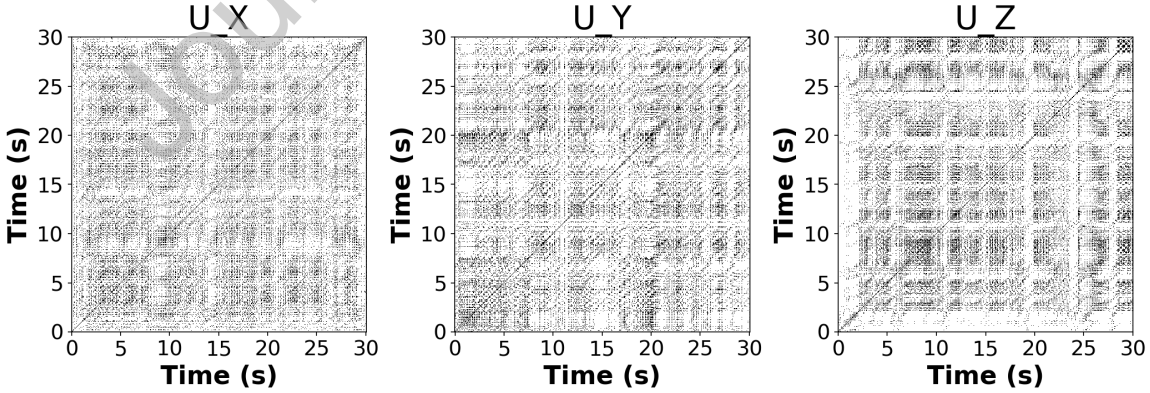
3.1. Nonlinear Dynamics and Flow Recurrence Analysis

Establishing the pseudo-recurrent character of the flow field is a prerequisite for rCFD, yet the instantaneous velocity signal at any probe location in a turbulence STR exhibits irregular amplitude variations that can obscure this underlying structure. To verify that the selected 4.5 s database window ($\approx 2.65\,\tau_{\mathrm{rec}}$, capturing 15 complete impeller revolutions at 200 RPM) adequately represents the pseudo-recurrent attractor, recurrence quantification analysis (RQA) was applied to a 30 s velocity record at two lattice probe locations sampled at 200 Hz.

As shown in Figure 7, the recurrence structure repeats on timescales of approximately 4 to 5 s throughout the full 30 s record. The self-similar block pattern confirms that the 4.5 s window captures a statistically representative portion of the pseudo-recurrent flow cycle and that extending the database beyond this duration would not alter the recurrence statistics governing the extrapolation. Full RQA methodology, cell-level velocity and acceleration analysis, and recurrence plots at representative cell locations are provided in Appendix Nonlinear Dynamics and Flow Recurrence Analysis.



(a) Lattice point 1: M-Star lattice node near the Rushton turbine at $(x, y, z) = (0.06,\ 0.035,\ 0)$ m.



(b) Lattice point 2: M-Star lattice node near the Rushton turbine at $(x, y, z) = (0.05,\ 0.035,\ -0.04)$ m.

Figure 7: Recurrence plots from a 30 s M-Star dataset (recurrence threshold 15%, sampling frequency 200 Hz), showing that the pseudo-recurrent structure repeats within windows of approximately 4–5 s and validating the choice of a 4.5 s rCFD database window.

3.2. Mixing Time Sensitivity to Temporal Injection

The fidelity of the lattice-to-cell data conversion is assessed through velocity profile comparisons and statistical metrics in appendix Conversion Fidelity Assessment; the conversion preserves the essential flow field structure with an overall Pearson correlation of $R_{pearson} = 0.986$.

The rCFD methodology is validated by comparing predicted mixing times against those obtained from M-Star ($4.7{\cdot}10^6$ lattice nodes, LB-LES, Smagorinsky–Lilly), Ansys Fluent ($1.2{\cdot}10^6$ cells, FV-LES), and experimental conductivity measurements at a fixed injection location in the 30 L stirred tank reactor, following the configuration of Rautenbach et al. [45]. The tracer was introduced as a pulse within a cubic volume ($\Delta x = \Delta y = \Delta z = 10$ mm) centered at (x, y, z) = (0, 0.41, 0.116) m, corresponding to $y/H = 0.95$ and $r/R = 0.80$, a near-surface position in the outer radial region of the vessel.

To evaluate both the temporal sensitivity and the run-to-run reproducibility of the method, tracer injection was initiated at ten discrete time instances $t_{\text{inj}} = 1, 2, \ldots, 10$ s, sampling different instantaneous flow field states within the transient stirred-tank flow. Each temporal case was repeated five times under otherwise identical conditions, yielding 50 simulations in total. All 50 rCFD runs shared a single recurrence database constructed from $\tau_{\text{rec}} = 4.5$ s of M-Star velocity field data, confirming that database generation is performed only once and all subsequent predictions are obtained through time extrapolation alone.

Figure 8 compares the mean $\tau_{m,95}$ values across the ten injection slots for rCFD, M-Star, and Ansys Fluent prior to any diffusivity calibration (uncalibrated case f_{swap}=0.0125). All three methods operate at comparable spatial resolutions: M-Star employs a lattice of $4.7 \cdot 10^6$ nodes with the Smagorinsky–Lilly subgrid-scale model (LB-LES), and Ansys Fluent employs a finite-volume mesh of $1.2 \cdot 10^6$ cells with an LES turbulence closure. The rCFD scalar transport is driven by the precomputed velocity displacement field extracted from the same $4.7 \cdot 10^6$ node M-Star simulation; no additional flow solve is performed during the rCFD extrapolation phase. This equal-resolution basis ensures that observed differences in predicted mixing time reflect methodological distinctions rather than resolution artifacts.

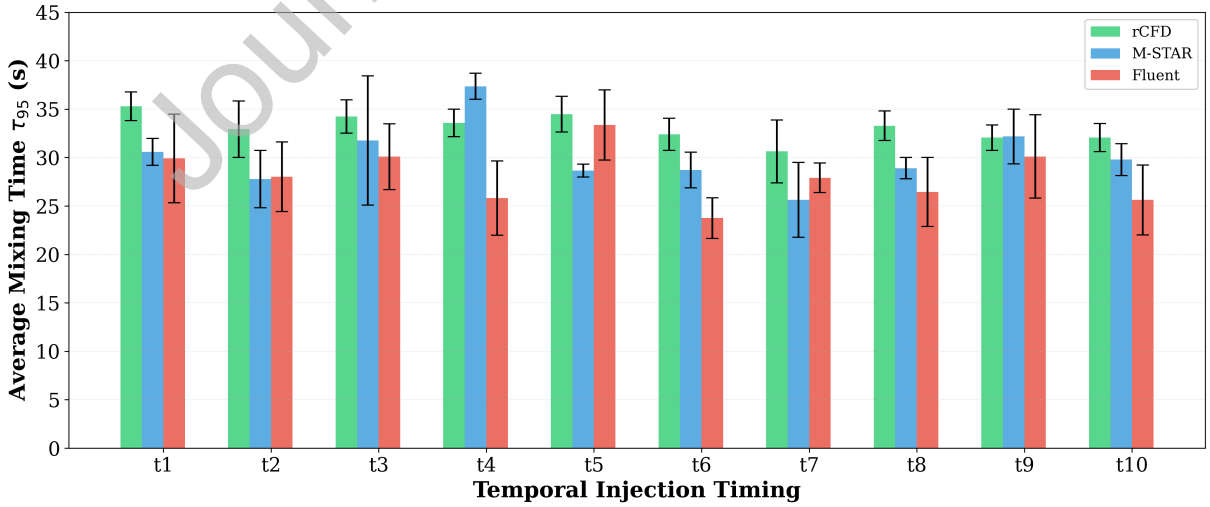
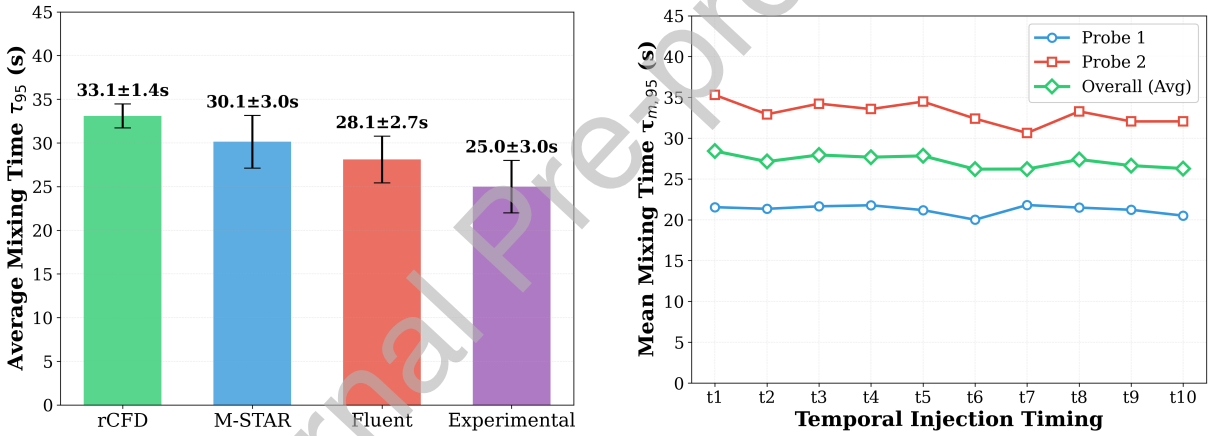


Figure 8: Effect of tracer injection timing on predicted $\tau_{m,95}$ for uncalibrated rCFD, M-Star, and Ansys Fluent. Each data point represents the mean over five independent repetitions.

Journal Pre-proof

Consistent with the findings of Rautenbach et al. [45], all three methods exhibit slot-to-slot variation in predicted mixing time, attributable to differences in the instantaneous flow field state at the moment of tracer injection, a direct manifestation of the macro-instabilities inherent to stirred-tank hydrodynamics. This temporal sensitivity is method-independent and reflects a physical property of the flow rather than a numerical artifact.

In addition to slot-to-slot variation, the rCFD results exhibit a secondary source of variability: run-to-run differences within nominally identical injection slots. This spread arises from the stochastic nature of the frame-sequencing algorithm in the rCFD time-extrapolation step, in which the order of recurrence frame assignments is not fixed between runs, introducing small but non-negligible differences in the reconstructed scalar transport history. Notably, as evident from the error bars in Figure 8, the within-slot standard deviation for rCFD remains small and consistent across all ten injection points, exhibiting less variability than the reproducibility spread observed for M-Star and Ansys Fluent. This indicates that stochastic frame sequencing, while a source of run-to-run variation, produces a more uniform and bounded uncertainty than the numerical irreproducibility accumulated over the course of full transient simulations in conventional solvers.



(a) Overall mean $\tau_{m,95}$ pooled across 50 realizations data points.

(b) Per-slot mean $\tau_{m,95}$ at probe 1, probe 2, and domain average across ten injection slots.

Figure 9: Uncalibrated rCFD $\tau_{m,95}$ predictions: (a) pooled mean; (b) per-slot breakdown by monitor.

When pooled across all 50 realizations, the uncalibrated rCFD predicts a mean $\tau_{m,95}$ of 33.1 ± 1.4 s, compared with 30.1 ± 3.0 s for M-Star, 28.1 ± 2.7 s for Ansys Fluent, and an experimental reference of 25.0 ± 3.0 s, as shown in Figure 9a. All three computational methods over-predict the experimental mixing time: M-Star and Ansys Fluent deviate by approximately 20 % and 12 %, respectively, while uncalibrated rCFD deviates by approximately 32 %. The larger over-prediction in rCFD reflects the uncalibrated diffusivity setting $f_{swap} = 0.0125$, the lowest value of the calibration range (minimum diffusive exchange), which does not compensate for the sub-grid mixing deficit at this resolution; this is addressed through experimental calibration of f_{swap} in the following subsection.

Before presenting the calibration results, it is instructive to examine why the probe-based and domain-averaged mixing times differ systematically as shown in Figure 9b. In the simulations, $\tau_{m,95}$ is monitored simultaneously at probe 1, probe 2, and over the full

Journal Pre-proof

computational domain, with the reported mixing time taken as the maximum of the three. The domain-averaged indicator reaches $\tau_{m,95}$ earlier than either probe because it reflects the bulk homogenization of the tank, whereas the probes are positioned near the free surface at a radius of $r = 0.112$ m, a region characterized by weaker axial exchange and, in the case of probe 2, a recirculation shadow immediately downstream of a baffle plate. As a result, probe 2 consistently records the largest mixing time and governs the reported $\tau_{m,95}$ across all simulation conditions, a behavior that mirrors the probe-position sensitivity documented experimentally in Rautenbach et al. [45]

As described in section Numerical Diffusion via Face-Based Concentration Exchange, the face-swap diffusion parameter f_{swap} compensates for two physical gaps inherent to the present configuration: sub-grid scalar dispersion unresolved at the employed lattice resolution, and the buoyancy contrast of the experimental tracer not captured by the passive-scalar assumption. To identify the appropriate value, rCFD simulations were performed across fourteen values of f_{swap} spanning 0.0125 to 0.1425, with tracer injection at $t_{\text{inj}} = 1$ s and five independent repetitions at each value, yielding 70 simulations in total. The resulting probe-based $\tau_{m,95}$ values are shown in Figure 10 together with the experimental uncertainty band (25.0 ± 3.0 s). A monotonically decreasing trend is observed for both probes and the domain average as f_{swap} increases, consistent with progressively enhanced inter-cell mixing. Probe 2 remains the governing monitor throughout, entering the experimental band at $f_{\text{swap}} = 0.0825$, the value at which all monitored indicators fall within the experimental uncertainty range. Since all simulation methods consistently predict lower $\tau_{m,95}$ at later injection times relative to $t_{\text{inj}} = 1$ s, the selected f_{swap} value is the one whose prediction at the earliest injection time falls within the experimental band, providing a conservative and reproducible selection criterion. It is acknowledged that this calibration relies on conductivity data at a single probe location; validation against an independent measurement at a different vessel height would strengthen confidence in the selected value and is recommended for future work.

Journal Pre-proof

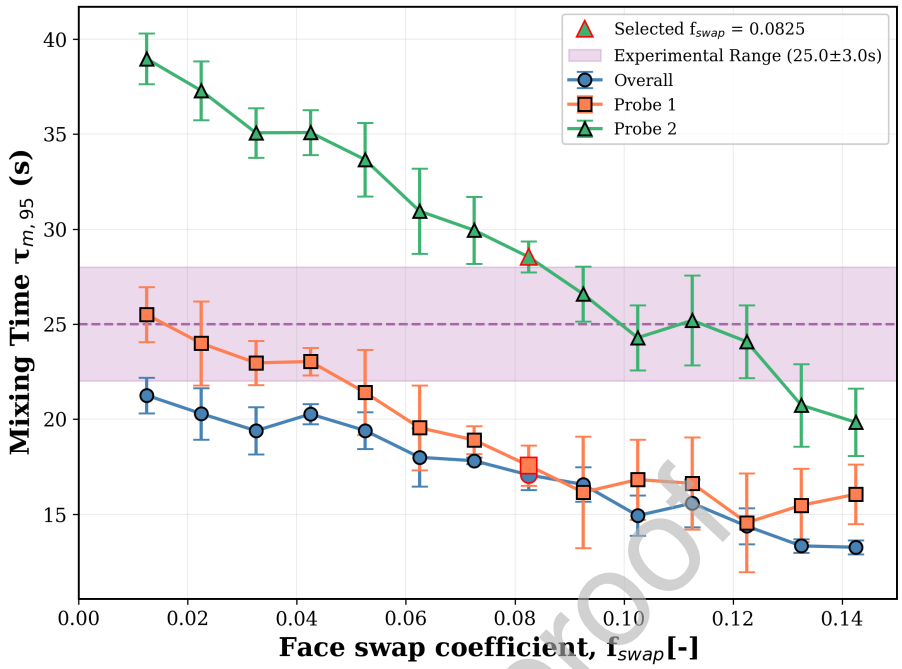


Figure 10: Sensitivity of $\tau_{m,95}$ to f_{swap} at probe 1, probe 2, and domain average (five repetitions each). The shaded band marks the experimental range (25.0 ± 3.0 s); $f_{\text{swap}} = 0.0825$ is selected where all indicators enter the band.

The value $f_{\text{swap}} = 0.0825$ is retained for all subsequent simulations: it is the minimum diffusivity that brings all monitored indicators within the experimental band, avoiding the introduction of artificial diffusion beyond what is physically warranted.

With $f_{\text{swap}} = 0.0825$, the calibrated rCFD predictions are shown in Figures 11 and 12. Across all injection slots, calibrated rCFD consistently predicts $\tau_{m,95}$ values within the experimental uncertainty band, with a pooled mean of 24.5 ± 1.0 s, in close agreement with the experimental reference of 25.0 ± 3.0 s. M-Star and Ansys Fluent, operating at the same spatial resolution with LES turbulence closures, predict pooled means of 30.1 ± 3.0 s and 28.1 ± 2.7 s, respectively both outside the experimental band. The narrower standard deviation of the calibrated rCFD predictions reflects the inherent smoothing of stochastic variability through repeated sampling of the recurrence database, rather than an artificial suppression of mixing dynamics; the underlying slot-to-slot trend is preserved, as confirmed by Figure 11.

The per-slot breakdown in Figure 12b shows that probe 2 remains the governing monitor across all injection points following calibration, consistent with its baffle-adjacent position and weaker axial exchange near the free surface. Probe 1 and the domain-averaged indicator reach $\tau_{m,95}$ earlier at every slot, confirming that the reported mixing times characterize the most poorly mixed region of the vessel, a physically appropriate and conservative definition of process mixing time. The slot-to-slot variation pattern observed in the uncalibrated case is qualitatively preserved after calibration, confirming that the adjusted diffusivity shifts the absolute mixing time into agreement with experiment without altering the underlying sensitivity of the method to the instantaneous flow state at injection.

Journal Pre-proof

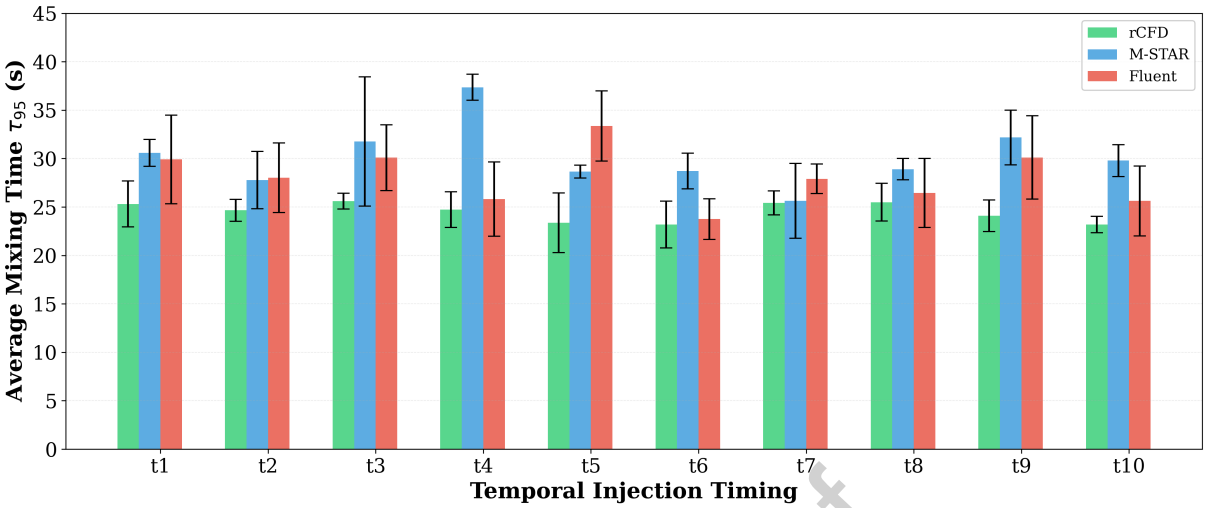
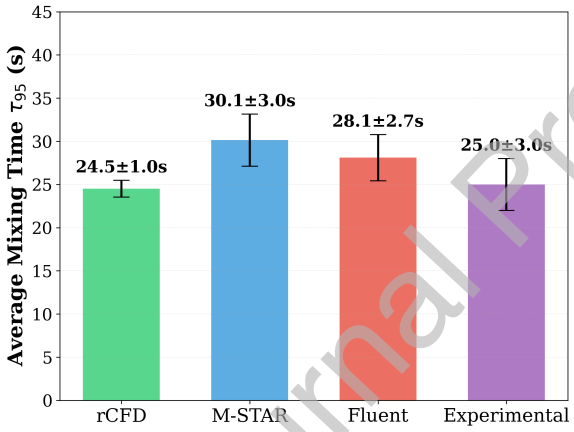
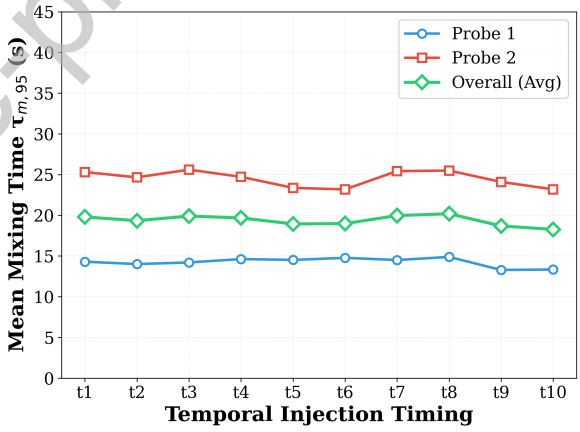


Figure 11: Per-slot mean $\tau_{m,95}$ (s) for calibrated rCFD, M-Star, and Ansys Fluent across ten injection slots; reported value is the maximum among probe 1, probe 2, and domain average.



(a) Overall mean $\tau_{m,95}$ pooled across 50 realizations with $f_{\text{swap}} = 0.0825$).



(b) Per-slot mean $\tau_{m,95}$ at probe 1, probe 2, and domain average following calibration.

Figure 12: Calibrated rCFD $\tau_{m,95}$ predictions ($f_{\text{swap}} = 0.0825$): (a) pooled mean; (b) per-slot breakdown by monitor.

The individual slot-mean values show that rCFD reproduces the experimental mixing time to within the experimental uncertainty at nine of the ten injection slots, with the largest individual deviation occurring at slot t5. M-Star and Ansys Fluent over-predict at the majority of slots, with slot-to-slot spread reflecting the macro-instabilities of the flow at the moment of injection.

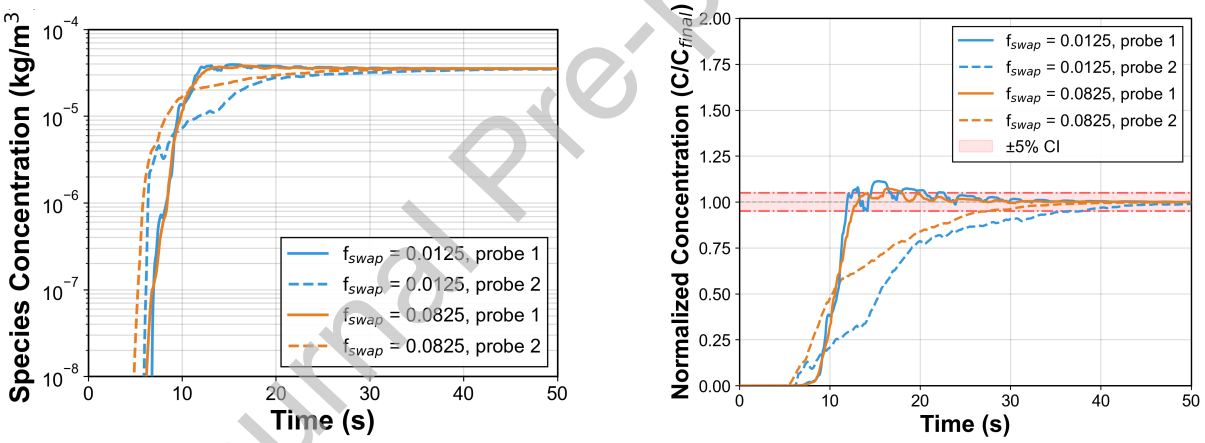
These results demonstrate that transport-based rCFD, when supplied with a pre-computed conventional flow field and calibrated against a single experimental probe measurement, reproduces the 95 % mixing time in close agreement with experiment. It should be noted that rCFD does not replace conventional simulation: it relies on it as input. The advantage it offers lies in the subsequent transport phase once the recurrence database is established,

Journal Pre-proof

passive tracer studies across varied injection conditions, diffusivity parameters, or timing can be conducted at a fraction of the cost of re-running a full transient solver for each case.

3.3. Dynamic Tracer Concentration

This section examines the temporal evolution of tracer concentration at the two probe locations and in the full three-dimensional field, providing insight into the mixing pathway and the sensitivity of the transport predictions to the face-swap diffusion parameter. Figure 13a shows the absolute tracer concentration on a logarithmic scale for both the calibrated ($f_{\text{swap}} = 0.0825$) and uncalibrated ($f_{\text{swap}} = 0.0125$) cases, while Figure 13b presents the corresponding normalized concentrations. Probe 1 exhibits closely similar concentration histories for both diffusivity values, in both magnitude and temporal profile, indicating that scalar transport at this location is governed primarily by convective flow structures rather than diffusion. Probe 2, by contrast, reaches equilibrium noticeably earlier under the higher diffusivity, consistent with its baffle-adjacent position where diffusive exchange plays a proportionally larger role in tracer homogenization. The qualitative trend obtained from rCFD at both probes is broadly consistent with the Ansys Fluent predictions reported by Rautenbach et al. [45], though a quantitative comparison of the transient concentration signals is reserved for future work.

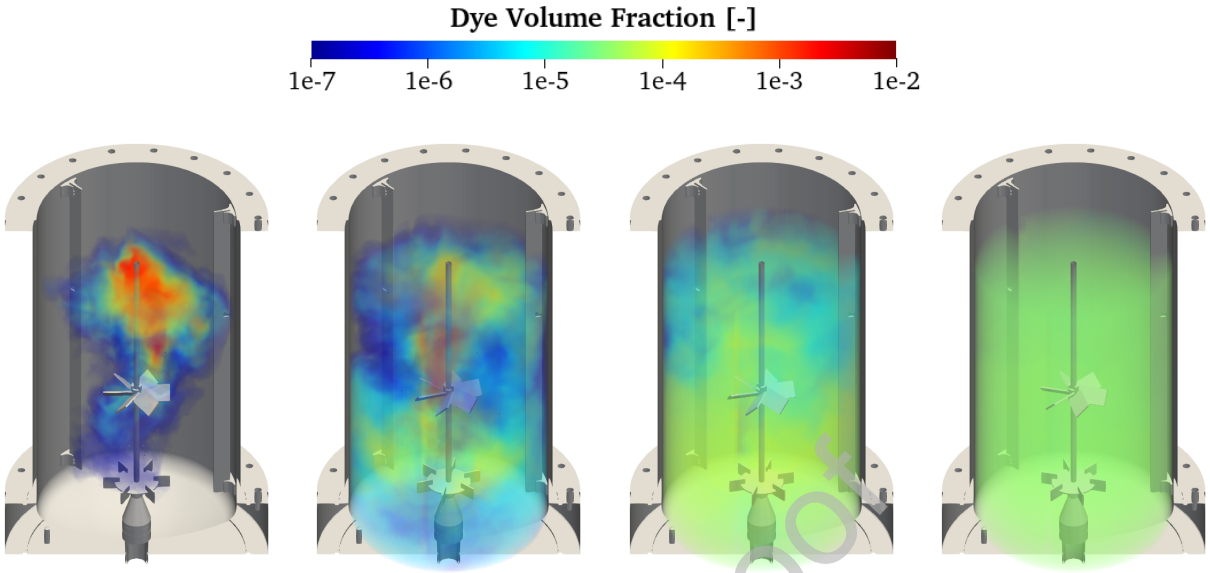


(a) Absolute tracer concentration

(b) Normalized concentration

Figure 13: Tracer concentration histories at probe 1 and probe 2 for calibrated ($f_{\text{swap}} = 0.0825$) and uncalibrated rCFD; probe 2 equilibrates earlier under higher diffusivity.

Journal Pre-proof



(a) 5 s (b) 7 s (c) 10 s (d) 25 s

Figure 14: Three-dimensional tracer concentration at $t = 5$, 7, 10, and 25 s following pulse injection at $y/H = 0.95$, $r/R = 0.80$ (calibrated rCFD).

Journal Pre-proof

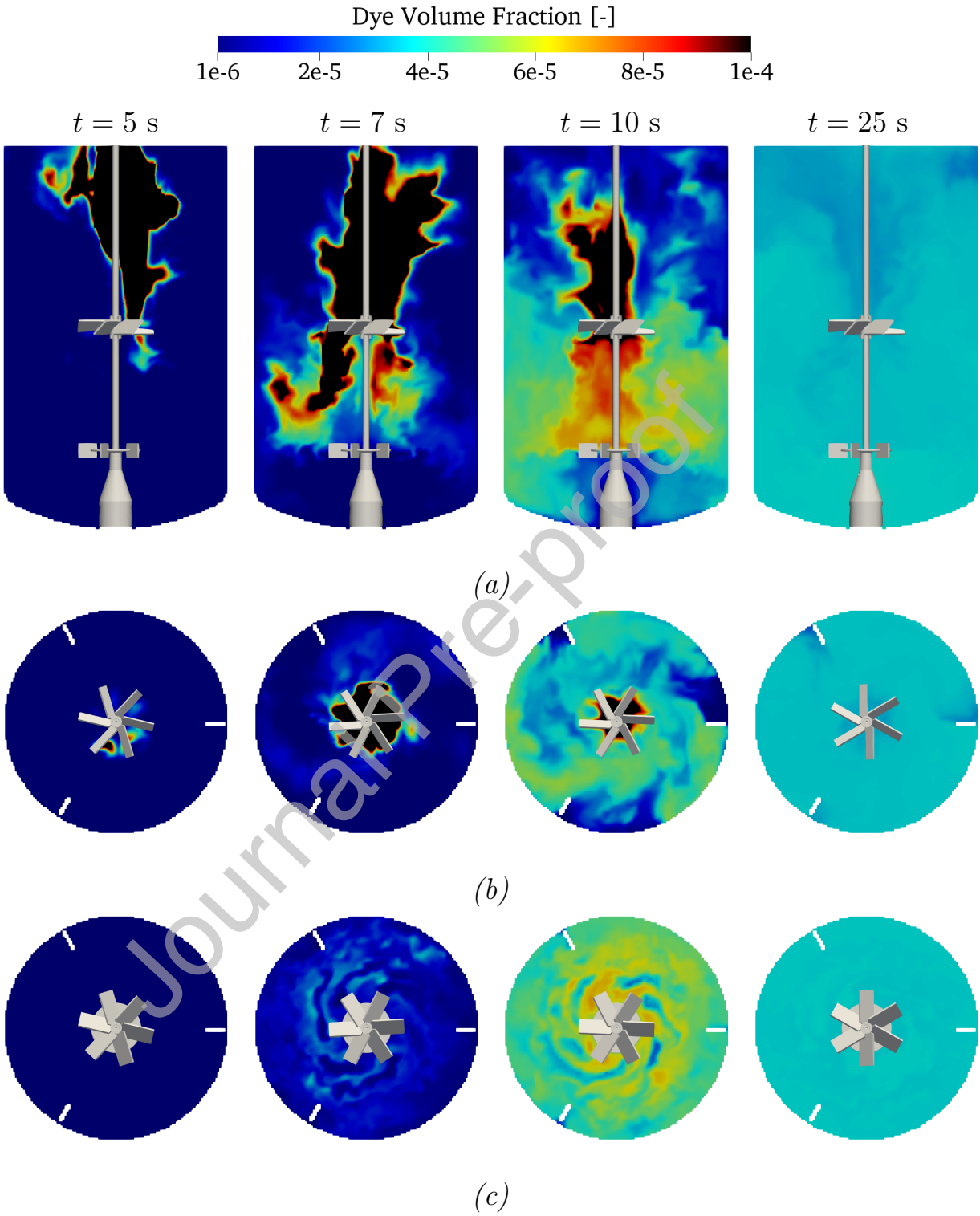


Figure 15: Cross-sectional tracer concentration at $t = 5$, 7, 10, and 25 s: (a) vertical centreline slice; (b) horizontal slice at the pitched blade turbine plane ($y/H = 0.45$); (c) horizontal slice at the Rushton turbine plane ($y/H = 0.09$).

Journal Pre-proof

Journal Pre-proof

The three-dimensional distribution of tracer following pulse injection is illustrated in Figures 14 and 15 at four representative time instants. Immediately after injection at the near-surface outer radial position ($y/H = 0.95$), the tracer accumulates as a concentrated hot spot near the pitched blade turbine swept volume, visible at $t = 5$ s in both the 3D rendering and the vertical cross-section. By $t = 7$ s, a strong downward plume has developed, transporting tracer toward the Rushton turbine plane, where it is subsequently dispersed radially outward. The horizontal slices at $t = 10$ s capture the characteristic swirling dispersion pattern driven by each impeller, with the Rushton turbine producing the more pronounced tangential structure at the lower elevation. By $t = 25$ s, the concentration field is near-uniform throughout the bulk of the vessel. Persistent weakly-mixed regions are visible immediately downstream of the baffle plates in the vertical slice from the earliest time instants onward, consistent with the local deflection of tangential flow and suppression of axial exchange at these locations, a structural feature of the baffled geometry that is independent of the injection configuration.

3.4. Computational Performance

Increasing Δt_{rCFD} from 0.005 s to 0.025 s achieves a 4.9× reduction in wall-clock time while keeping $\tau_{m,95}$ within 5% of the calibration-setting value; beyond $\Delta t_{\mathrm{rCFD}} = 0.025$ s the accuracy penalty becomes more pronounced. The full sensitivity analysis across four time step widths is presented in Appendix Effect of Time Step Width on Mixing Time and Computational Cost. All subsequent simulations employ $\Delta t_{\mathrm{rCFD}} = 0.025$ s.

The computational cost of rCFD relative to conventional simulation methods is summarized in Tables 1 and 2. Per second of simulated process time, rCFD requires approximately 3 s of wall-clock time, compared with 1,268 s for M-Star and 3,600 s for Ansys Fluent corresponding to speedup factors of 423× and 1200×, respectively. These figures should be interpreted in light of the hardware configurations listed in Table 1: M-Star was executed on an NVIDIA A100-PCIE GPUs and Ansys Fluent on NVIDIA A100-SXM4 GPUs, while rCFD ran on 32 CPU cores of an AMD EPYC 9354 node. Since the GPU hardware used for M-Star and Ansys Fluent is considerably more capable than the CPU cluster used for rCFD, the reported speedup factors represent conservative lower bounds on the true computational advantage of the method.

Table 1: Wall-clock time per second of simulated process time for each computational method, with speedup factors relative to Ansys Fluent and M-Star.

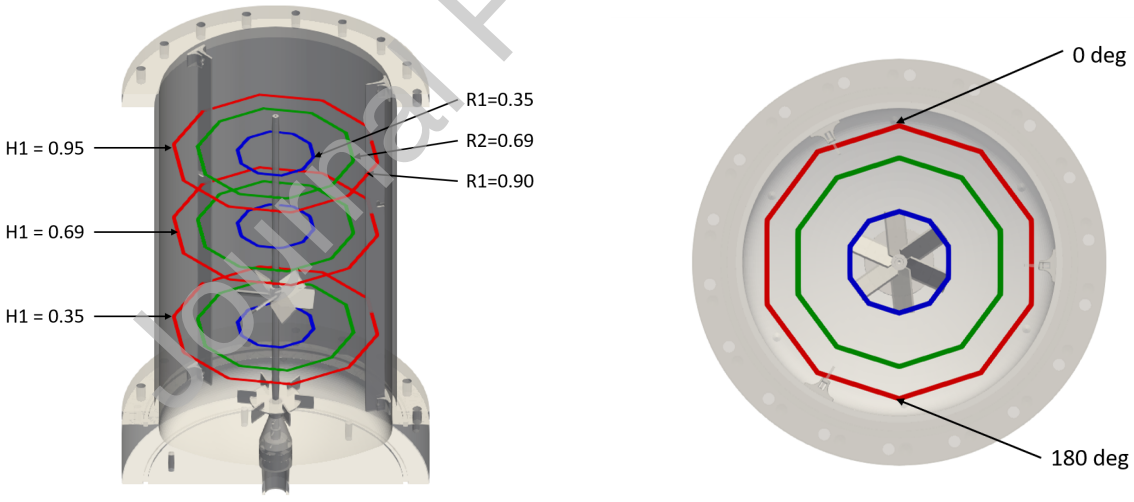
Method	Clock Time (s)	Speedup vs Ansys Fluent (-)	Speedup vs M-Star (-)	Hardware
Ansys Fluent (FVM)	3600	1×	–	1 × NVIDIA A100-SXM4-40GB GPUs
M-Star (LBM)	1268	3×	1×	1 × NVIDIA A100-PCIE-40GB GPUs
rCFD	3	1200×	423×	AMD EPYC 9354, 32-Cores CPU

Table 2: Estimated wall-clock cost for the 90-case injection location parametric study, assuming a mean mixing time of 30 s per case.

Method	Time per Case	Total (90 Cases)
Ansys Fluent	≈30 hours	≈113 days
M-Star	≈11 hours	≈40 days
rCFD	≈1.5 minutes	≈2.25 hours

Long-timescale mixing is a standing bottleneck for conventional CFD. For the 90-configuration sweep considered here, each case covering approximately 30 s of mixing, Table 2 estimates the total cost at roughly 113 days with Ansys Fluent and 40 days with M-Star on the present CPU hardware figures that effectively preclude iterative use in design or optimization. The same sweep completes in approximately 2.25 hours under rCFD, making the systematic injection-location study in Section Injection Location Sensitivity tractable within this work. This speed-up is not unconditional: rCFD replays a previously computed flow field rather than resolving the momentum equations in each run, so its predictive accuracy is inherited from the underlying CFD database and from the transport parameters calibrated to it. An upfront full-CFD simulation and case-specific validation against a trusted reference therefore, remain prerequisites for any parametric deployment of the method, and the quoted wall-clock saving should be interpreted with that scope in mind.

3.5. Injection Location Sensitivity



(a) Normalized radial and axial injection positions.

(b) Azimuthal distribution of ten injection angles.

Figure 16: Fastest and slowest $\tau_{m,95}$ configurations from the 90-case parametric study. Locations at $r/R = 0.35$, $y/H \leq 0.69$ are physically inaccessible and excluded from practical consideration.

Building on the computational efficiency characterized in section Computational Performance, a systematic parametric study was conducted to map $\tau_{m,95}$ across the geometric design space of the vessel. The tracer was introduced as a $10 \times 10 \times 10$ mm^3 pulse at locations defined by three normalized radii ($r/R = 0.35$, 0.69, 0.90), three normalized heights

(y/H = 0.35, 0.69, 0.95), and ten azimuthal angles (θ = 0° to 324° in 36° increments), yielding 90 simulations in total. The geometric parameters are summarized in Table 3 and the injection grid is illustrated in Figure 16.

Table 3: Dimensionless geometric parameters for parametric injection study.

Parameter	Actual (m)	Reference (m)	Normalized (-)	Approximation (-)
Radial Position				
R_1	0.05		0.346	0.35
R_2	0.10	0.1445	0.692	0.69
R_3	0.13		0.899	0.90
Vertical Height				
H_1	0.15		0.346	0.35
H_2	0.30	0.433	0.693	0.69
H_3	0.41		0.947	0.95

Figure 17 shows the mean $\tau_{m,95}$ as a function of normalized radius and height, reported separately for probe 1, probe 2, and the domain average. Probe 2 consistently records the longest mixing time across all configurations, attributable to its baffle-adjacent position in the weakly exchanged near-surface region, and therefore governs the reported $\tau_{m,95}$ throughout the parametric study. A clear trend is observed with injection height: as the injection location moves from y/H = 0.35 toward the free surface at y/H = 0.95, the mean $\tau_{m,95}$ increases monotonically at r/R = 0.69 and r/R = 0.90, while the effect is less pronounced at the inner radius r/R = 0.35. This height dependence reflects the weakening of axial convective transport with distance from the impeller swept volumes.

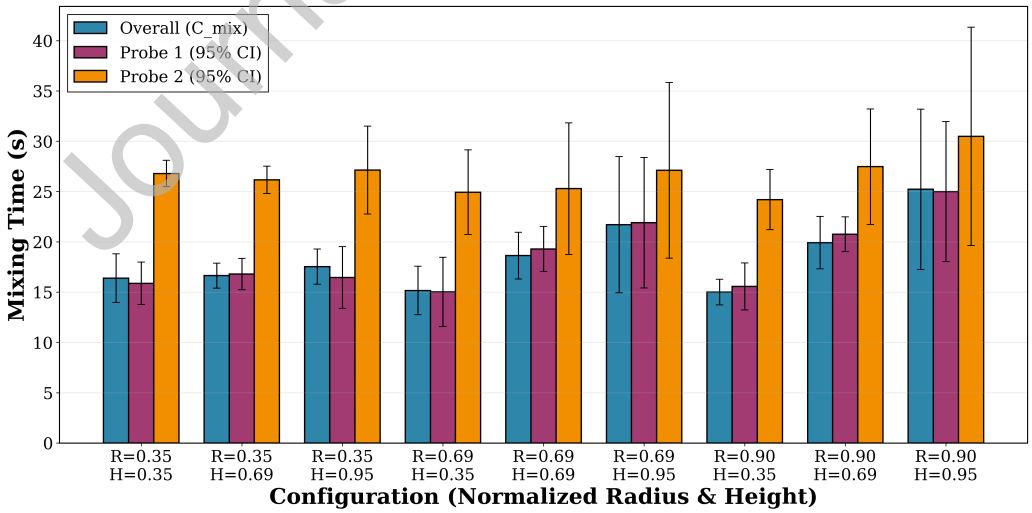


Figure 17: Mean $\tau_{m,95}$ as a function of normalized injection radius and height, reported for probe 1, probe 2, and the full domain average. Values represent the mean over ten azimuthal angles at each (r/R, y/H) combination.

Journal Pre-proof

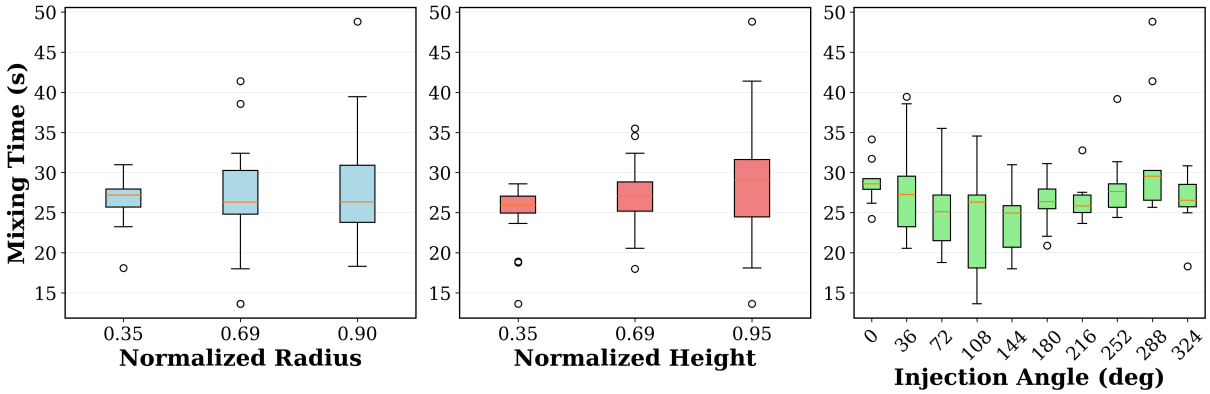


Figure 18: Box plot distributions of $\tau_{m,95}$ aggregated by normalized radius r/R (left), normalized height y/H (center), and injection angle θ (right) for calibrated rCFD. Each box represents 30 simulations (radius and height panels) or 9 simulations (angle panel). Whiskers extend to 1.5× the inter-quartile range; outliers shown individually.

The distribution of $\tau_{m,95}$ across the three parametric dimensions is presented as box plots in Figure 18. The median mixing time is relatively insensitive to normalized radius, remaining near 26-27 s across all three r/R values, though the spread widens progressively at the outer positions. Height shows a more pronounced effect: the median increases and the inter-quartile range broadens substantially from $y/H = 0.35$ to $y/H = 0.95$, with the upper whisker extending beyond 40 s at the uppermost level. The injection angle introduces the most asymmetric response: $\theta = 108°$ consistently yields lower median values (≈24–25 s) and wider downside spread, while $\theta = 0°$, 288°, and 324° cluster near 27–29 s with narrower distributions, reflecting preferential mixing pathways aligned with the impeller rotational direction.

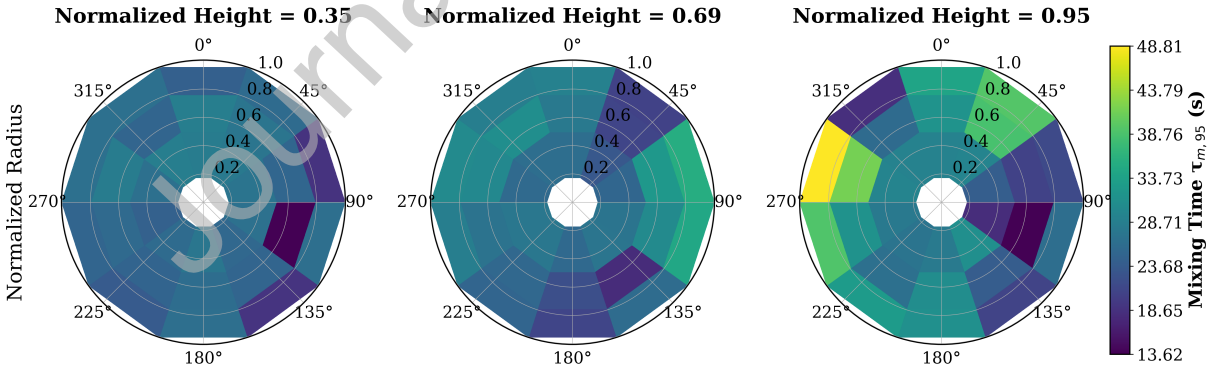


Figure 19: Polar heatmaps of $\tau_{m,95}$ as a function of normalized radius and injection angle at three normalized heights: $y/H = 0.35$ (left), $y/H = 0.69$ (center), and $y/H = 0.95$ (right) for calibrated rCFD. A shared colormap spanning 13.62–48.81 s is used across all panels.

Journal Pre-proof

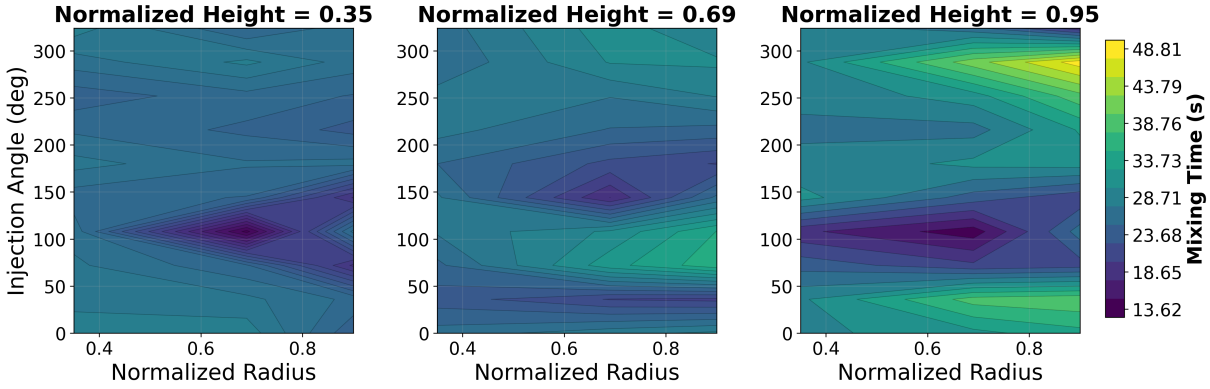


Figure 20: Cartesian contour maps of $\tau_{m,95}$ as a function of injection angle and normalized radius at three normalized heights: $y/H = 0.35$ (left), $y/H = 0.69$ (center), and $y/H = 0.95$ (right) for calibrated rCFD. A shared colormap spanning 13.62–48.81 s is used across all panels.

Table 4: Optimal and suboptimal injection locations identified from the parametric study. Configurations at $r/R = 0.35$ near PBT and RT ($r = 0.05$ m) are physically inaccessible due to proximity to the impeller shaft and are excluded from practical consideration regardless of their predicted $\tau_{m,95}$.

Rank	Height (y/H)	Radius (r/R)	Angle (°)	t_{mix} Domain (s)	t_{mix} Probes (s)
Fastest mixing (top 3)					
1	0.95	0.69	108	13.75	13.62
2	0.35	0.69	108	12.59	13.63
3	0.69	0.69	144	17.16	17.98
Slowest mixing (bottom 3)					
88	0.95	0.90	36	27.50	39.45
89	0.95	0.69	288	33.90	41.40
90	0.95	0.90	288	41.83	48.81

The full spatial distribution of $\tau_{m,95}$ across all 90 configurations is shown as polar heatmaps in Figure 19 and as Cartesian contour maps in Figure 20, both using a consistent colormap spanning the global range of 13.62–48.81 s. At $y/H = 0.35$, the mixing time is largely uniform across the r/R–θ plane, confined predominantly to the teal-cyan range of 24–28 s, with the notable exception of a sharply favorable configuration at $r/R = 0.69$, $\theta = 108°$ ($\tau_{95} = 13.63$ s). At $y/H = 0.69$, variability increases moderately, with elevated mixing times appearing at the outer radius near $\theta = 72°$ and 108°. The most pronounced spatial heterogeneity is observed at $y/H = 0.95$, where the mixing time range widens to 13.62–48.81 s. A large unfavorable zone is visible in both the polar and Cartesian representations between $\theta = 270°$ and 315° at $r/R = 0.90$, corresponding to the worst-performing configuration in the entire study ($\tau_{95} = 48.81$ s). Conversely, the favorable $\theta = 108°$ sector persists at this height, yielding $\tau_{95} = 13.62$ s, confirming that certain angular positions produce short mixing times across all height levels. The configurations with the lowest and with the highest mixing time are compiled in Table 4.

Journal Pre-proof

Journal Pre-proof

Not all configurations in the parametric space are physically accessible. At the innermost radial position ($r/R = 0.35$, $r = 0.05$ m), the injection point falls in the immediate vicinity of the impeller shaft at mid and lower vessel heights ($y/H = 0.35$ and $y/H = 0.69$), where placement of an injection device would obstruct impeller motion or disturb the flow structures governing tracer dispersal. The 20 configurations at these locations are therefore not representative of achievable operating conditions and should be excluded from practical consideration. The actionable design space therefore, comprises 70 of the 90 simulated configurations.

Finally, as highlighted by Rautenbach et al. [45], the flow field in stirred tank reactors exhibits sensitivity to small perturbations, such that nominally identical operating conditions can produce measurable differences in mixing behavior. The design space presented here should therefore be treated as a probabilistic guide rather than a deterministic lookup table, to be interpreted alongside physical accessibility constraints and flow stability considerations when selecting an injection configuration for a given process.

4. Conclusions

This study presents the first experimental validation of transport-based recurrence CFD (rCFD) for mixing time prediction in a stirred tank reactor. Using a coarse M-Star Lattice Boltzmann flow field ($4.7 \cdot 10^6$ lattice nodes, LB-LES) as sole input, rCFD reproduces the 95 % mixing time in a 30 L dual-impeller STR following calibration of the face-swap diffusion parameter f_{swap}. The calibrated model predicts $\tau_{m,95} = 24.5 \pm 1.0$ s, in agreement with the experimental reference of 25.0 ± 3.0 s. M-Star at the same lattice resolution ($4.7 \cdot 10^6$ nodes, LB-LES) predicts 30.1 ± 3.0 s, and Ansys Fluent ($1.2 \cdot 10^6$ cells, FV-LES) predicts 28.1 ± 2.7 s, both outside the experimental uncertainty band. All three methods operate at comparable spatial resolutions, making the comparison methodologically consistent. Run-to-run reproducibility, governed by stochastic frame sequencing in rCFD, is comparable to or better than that of conventional solvers across all temporal injection slots.

A time step sensitivity analysis across $\Delta t_{\text{rCFD}} = 0.005$ to 0.050 s reveals that $\tau_{m,95}$ increases progressively with step width when f_{swap} is held constant, owing to unresolved intermediate diffusion at larger extrapolation intervals. Up to $\Delta t_{\text{rCFD}} = 0.025$ s, the domain $\tau_{m,95}$ remains within 5% of the baseline, while wall-clock time reduces by 4.9× to 3 s per second of simulated process time yielding speedup factors of 423× relative to M-Star and 1,200× relative to Ansys Fluent on equivalent or superior GPU hardware. This efficiency enables the 90-case parametric injection study to be completed in approximately 2.25 hours, compared with an estimated 40 days for M-Star and 113 days for Ansys Fluent.

The parametric study reveals that injection height exerts the strongest influence on $\tau_{m,95}$, while azimuthal angle introduces asymmetry attributable to the fixed baffle geometry. Injection at $r/R = 0.69$, $\theta = 108°$ is identified as consistently favorable across all accessible height levels ($\tau_{m,95}$ as low as 13.62 s), while near-surface outer radial positions ($r/R = 0.90$, $y/H = 0.95$, $\theta = 288°$) yield the slowest mixing ($\tau_{m,95} = 48.81$ s) due to weak axial exchange and baffle-dominated recirculation. Configurations at $r/R = 0.35$ and $y/H \leq 0.69$ are physically inaccessible and excluded from practical consideration.

Three limitations of the current implementation merit acknowledgment. First, f_{swap} was calibrated against a single probe location; validation against an independent spatial mea-

surement would strengthen confidence in the selected value. Second, the rCFD database is constructed from a single statistically stationary flow field, restricting the present application to fixed operating conditions; extension to variable impeller speed or fed-batch operation would require multi-point databases and an interpolation strategy. Third, the intermediate node-to-cell conversion between M-Star and the Ansys Fluent-based rCFD solver introduces computational overhead and a potential interpolation error source; native rCFD integration within M-Star would eliminate this step. Collectively, these results establish transport-based rCFD as a computationally viable and experimentally consistent methodology for mixing time prediction and injection location optimization in stirred tank reactors of this configuration; extension to geometrically distinct vessels, different impeller types, or variable operating conditions will require independent validation.

Nomenclature

Latin symbols

c	Tracer concentration	$\mathrm{kg\,m^{-3}}$
c_0, c_1	Tracer concentration in adjacent cells	$\mathrm{kg\,m^{-3}}$
C_S	Smagorinsky model constant	–
f_{imp}	Impeller rotation frequency	Hz
f_{max}	Maximum characteristic flow frequency	Hz
f_{min}	Minimum characteristic flow frequency	Hz
f_{out}	Velocity field snapshot frequency	Hz
f_{swap}	Face-swap diffusion parameter	–
H	Vessel working height	m
n	Impeller rotational speed	rpm
N	Number of data points	–
$N^{(\phi)}$	Recurrence norm normalization factor	–
r	Radial position	m / –
R	Vessel inner radius	m
$R^{(\phi)}(t,t')$	Recurrence norm between flow states at t and t'	–
$\mathcal{R}(t,t')$	Recurrence matrix	–
R^2	Coefficient of determination	–
t	Time	s
t_{inj}	Tracer injection time	s
$\mathbf{u}$	Velocity vector	$\mathrm{m\,s^{-1}}$
u_{max}	Maximum velocity magnitude	$\mathrm{m\,s^{-1}}$
u_{RMS}	Root-mean-square velocity	$\mathrm{m\,s^{-1}}$
V_{flip}	Effective face-exchange volume, $\min(V_{c_0}, V_{c_1})$	$\mathrm{m^3}$
V_{reac}	Reactor working volume	$\mathrm{m^3}$
y	Vertical (axial) coordinate	m

Journal Pre-proof

Journal Pre-proof

Greek symbols

α	Local phase volume fraction	–
Δc	Concentration difference across cell face	$\mathrm{kg\,m^{-3}}$
Δm_{diff}	Diffusive mass exchanged per rCFD time step	kg
Δt_{CFD}	CFD snapshot time interval	s
Δt_{rCFD}	rCFD extrapolation time step	s
Δx	LBM lattice spacing	m
θ	Azimuthal injection angle	$^\circ$
ϕ	Generic passive scalar field	–
ϕ_x, ϕ_y, ϕ_z	Position-tracking passive scalars	m
ρ_{f}	Fluid density	$\mathrm{kg\,m^{-3}}$
$\tau_{m,95}$	95 % mixing time	s
τ_{rec}	Recurrence window duration	s

Abbreviations

BGK	Bhatnagar–Gross–Krook
C2C	Cell-to-cell
CFD	Computational fluid dynamics
Co	Courant number
FFT	Fast Fourier transform
FVM	Finite volume method
GPU	Graphics processing unit
LBM	Lattice Boltzmann method
LES	Large-eddy simulation
MAE	Mean absolute error
MRF	Multiple reference frame
NRMSE	Normalised root-mean-square error
PBT	Pitched blade turbine
RANS	Reynolds-averaged Navier–Stokes
rCFD	Recurrence-based computational fluid dynamics
RMSE	Root-mean-square error
RQA	Recurrence quantification analysis
RT	Rushton turbine
SGS	Subgrid-scale
STR	Stirred tank reactor
UDF	User-defined function
UDS	User-defined scalar

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the authors used Claude (Anthropic) to improve the clarity, grammar, and structure of the manuscript text. Specifically, AI tools were employed for: (1) language editing and grammar correction, (2) restructuring sentences and paragraphs

Journal Pre-proof

for improved readability, (3) ensuring consistency in technical terminology and notation, and (4) formatting LaTeX code. All scientific content, methodology, results interpretation, and conclusions were developed independently by the authors. AI tools were not used for data analysis, figure generation, experimental design, or scientific decision-making. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Asif Shaik reports financial support was provided by German Research Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1615 – 503850735.

Data Availability Statement

The data that support the findings of this study, including M-Star and rCFD simulation results, optimization study, and post-processing scripts, will be openly available in TORE (TUHH Open Research Repository) at `https://hdl.handle.net/11420/61857` upon publication.

The persistent identifier (DOI) is `https://doi.org/10.15480/882.16811`

References

[1] A. W. Nienow, M. F. Edwards, N. Harnby, Mixing in the process industries, Butterworth-Heinemann, 1997.

[2] A. L. Bowler, S. Bakalis, N. J. Watson, A review of in-line and on-line measurement techniques to monitor industrial mixing processes, Chemical Engineering Research and Design 153 (2020) 463–495.

[3] S. M. Kresta, A. W. Etchells III, D. S. Dickey, V. A. Atiemo-Obeng, Advances in industrial mixing: a companion to the handbook of industrial mixing, John Wiley & Sons, 2015.

[4] I. Shevchenko, E. Aliiev, Improving the efficiency of the process of continuous flow mixing of bulk components, Eastern-European Journal of Enterprise Technologies 6 (1) (2020) 108.

Journal Pre-proof

[5] D. S. Dickey, Mixing and blending, Kirk-Othmer Encyclopedia of Chemical Technology (2000) 1–71.

[6] F. X. McConville, S. B. Kessler, Scale-up of mixing processes: A primer, Chemical Engineering in the Pharmaceutical Industry: Active Pharmaceutical Ingredients (2019) 241–259.

[7] V. Hessel, H. Löwe, F. Schönfeld, Micromixers—a review on passive and active mixing principles, Chemical engineering science 60 (8-9) (2005) 2479–2501.

[8] J. P. Valdés, L. Kahouadji, O. K. Matar, Current advances in liquid–liquid mixing in static mixers: A review, Chemical Engineering Research and Design 177 (2022) 694–731.

[9] M. Schrimpf, J. Esteban, T. Rösler, A. J. Vorholt, W. Leitner, Intensified reactors for gas-liquid-liquid multiphase catalysis: From chemistry to engineering, Chemical Engineering Journal 372 (2019) 917–939.

[10] M. Jafarian, Y. Chisti, G. J. Nathan, Gas-lift circulation of a liquid between two interconnected bubble columns, Chemical Engineering Science 218 (2020) 115574.

[11] K. Wadaugsorn, S. Limtrakul, T. Vatanatham, P. A. Ramachandran, Mixing characteristics of gas and liquid phases in bubble column reactors from virtual tracer simulation, Industrial & Engineering Chemistry Research 57 (42) (2018) 14064–14079.

[12] N. A. Ibrahim, A. Khalid, I. Zaman, A. Sapit, B. Manshoor, Flow visualization of bubble structure in bubble column reactor for fluid mixing, in: AIP Conference Proceedings, Vol. 1831, AIP Publishing LLC, 2017, p. 020024.

[13] G. Forte, F. Alberini, M. Simmons, H. E. Stitt, Use of acoustic emission in combination with machine learning: monitoring of gas–liquid mixing in stirred tanks, Journal of Intelligent Manufacturing 32 (2) (2021) 633–647.

[14] F. Maluta, F. Alberini, A. Paglianti, G. Montante, Hydrodynamics, power consumption and bubble size distribution in gas-liquid stirred tanks, Chemical Engineering Research and Design 194 (2023) 582–596.

[15] X. Hu, A. D. Ilgun, A. Passalacqua, R. O. Fox, F. Bertola, M. Milosevic, F. Visscher, Cfd simulations of stirred-tank reactors for gas-liquid and gas-liquid-solid systems using openfoam®, International Journal of Chemical Reactor Engineering 19 (2) (2021) 193–207.

[16] D. Huang, H. Wang, F. Wang, H. Xu, X. Duan, X. Feng, G. Wang, W. Zhang, C. Yang, Micromixing characterization of gas–liquid fast reactive flow: Test reaction scheme and experimental investigation, AIChE Journal (2025) e70058.

[17] G. N. Sajida, H. K. Sari, S. Saripudin, N. Nurcahyo, Sustainable revitalization and hydrodynamic characterization of a lab-scale cstr for process engineering education, in: E3S Web of Conferences, Vol. 664, EDP Sciences, 2025, p. 09011.

Journal Pre-proof

[18] C.-Y. Yang, Z. Lin, Y.-S. Chien, The nonideal mixing effect on the selectivity dynamic of consecutive-parallel reactions in an isothermal continuous stirred tank reactor based on cholette's model, International Journal of Chemical Reactor Engineering (2025).

[19] C.-Y. Yang, Z. Lin, Y.-S. Chien, Nonideal mixing effect based on cholette's model on van de vusse reaction selectivity dynamics in an isothermal continuous stirred-tank reactor, International Journal of Chemical Reactor Engineering 22 (9) (2024) 1079–1089.

[20] A. Cugini, D. Smith, D. Cillo, Liquid mixing in a four phase continuous stirred tank reactor, Tech. rep., New York, NY; American Institute of Chemical Engineers (1987).

[21] P. Satjaritanun, Y. Khunatorn, N. Vorayos, S. Shimpalee, E. Bringley, Numerical analysis of the mixing characteristic for napier grass in the continuous stirring tank reactor for biogas production, Biomass and Bioenergy 86 (2016) 53–64.

[22] J. Tamminen, T. Koiranen, Mixing performance comparison of milliscale continuous high-shear mixers, The Canadian Journal of Chemical Engineering 93 (12) (2015) 2245–2252.

[23] Q. Zhang, Y. Yong, Z.-S. Mao, C. Yang, C. Zhao, Experimental determination and numerical simulation of mixing time in a gas–liquid stirred tank, Chemical Engineering Science 64 (12) (2009) 2926–2933.

[24] J. Fitschen, S. Hofmann, J. Wutz, A. Kameke, M. Hoffmann, T. Wucherpfennig, M. Schlüter, Novel evaluation method to determine the local mixing time distribution in stirred tank reactors, Chemical engineering science: X 10 (2021) 100098.

[25] G. Ascanio, Mixing time in stirred vessels: A review of experimental techniques, Chinese Journal of Chemical Engineering 23 (7) (2015) 1065–1076.

[26] D. Rajavathsavai, A. Khapre, B. Munshi, Study of mixing behavior of cstr using cfd, Brazilian Journal of Chemical Engineering 31 (2014) 119–129.

[27] A. Ramírez-López, Analysis of mixing efficiency in a stirred reactor using computational fluid dynamics, Symmetry 16 (2) (2024) 237.

[28] M. Wu, N. Jurtz, A. Walle, M. Kraume, Evaluation and application of efficient cfd-based methods for the multi-objective optimization of stirred tanks, Chemical Engineering Science 263 (2022) 118109.

[29] M. Jahoda, M. Moštěk, A. Kukuková, V. Machoň, Cfd modelling of liquid homogenization in stirred tanks with one and two impellers using large eddy simulation, Chemical Engineering Research and Design 85 (5) (2007) 616–625.

[30] A. Rosseburg, J. Fitschen, J. Wutz, T. Wucherpfennig, M. Schlüter, Hydrodynamic inhomogeneities in large scale stirred tanks–influence on mixing time, Chemical Engineering Science 188 (2018) 208–220.

Journal Pre-proof

[31] T. Płusa, J. Talaga, A. Duda, P. Duda, Modeling mixing dynamics in uncovered baffled and unbaffled stirred tanks, AIChE Journal 67 (9) (2021) e17322.

[32] G. L. Lane, Improving the accuracy of cfd predictions of turbulence in a tank stirred by a hydrofoil impeller, Chemical Engineering Science 169 (2017) 188–211.

[33] T. Lichtenegger, S. Pirker, Recurrence cfd–a novel approach to simulate multiphase flows with strongly separated time scales, Chemical Engineering Science 153 (2016) 394–410.

[34] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, et al., Recurrence plots of dynamical systems, World Scientific Series on Nonlinear Science Series A 16 (1995) 441–446.

[35] T. Lichtenegger, E. Peters, J. Kuipers, S. Pirker, A recurrence cfd study of heat transfer in a fluidized bed, Chemical Engineering Science 172 (2017) 310–322.

[36] P. Kieckhefen, T. Lichtenegger, S. Pietsch, S. Pirker, S. Heinrich, Simulation of spray coating in a spouted bed using recurrence cfd, Particuology 42 (2019) 92–103.

[37] T. Lichtenegger, P. Kieckhefen, S. Heinrich, S. Pirker, Dynamics and long-time behavior of gas–solid flows on recurrent-transient backgrounds, Chemical engineering journal 364 (2019) 562–577.

[38] S. Pirker, T. Lichtenegger, Efficient time-extrapolation of single-and multiphase simulations by transport based recurrence cfd (rcfd), Chemical Engineering Science 188 (2018) 65–83.

[39] S. Pirker, M. Atzori, S. Heinrich, T. Lichtenegger, Particle size segregation in bi and penta-disperse gas–solid fluidized beds: Cfd-dem and recurrence cfd simulations, Chemical Engineering Science 309 (2025) 121469.

[40] A. Shaik, S. Pietsch-Braune, S. Pirker, M. Özdemir, M. Adrian, A. Penn, S. Heinrich, MRI-validated CFD–DEM simulation and recurrence-based time extrapolation (rCFD) of a bubbling fluidized bed: Drag model selection and computational speed-up, Chemical Engineering Research and Design 227 (2026) 879–893.

[41] H. Lumetzberger, S. Pirker, T. Lichtenegger, Propagator-moments approximation for recurrence cfd: Application to species transport in turbulent flows, Chemical Engineering Science 311 (2025) 121624.

[42] T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, Model order reduction in fluid dynamics: challenges and perspectives, Reduced Order Methods for modeling and computational reduction (2014) 235–273.

[43] B. N. Hanna, N. T. Dinh, R. W. Youngblood, I. A. Bolotnov, Machine-learning based error prediction approach for coarse-grid computational fluid dynamics (cg-cfd), Progress in Nuclear Energy 118 (2020) 103140.

Journal Pre-proof

Journal Pre-proof

[44] F. Dabbagh, S. Pirker, S. Schneiderbauer, On the fast modeling of species transport in fluidized beds using recurrence computational fluid dynamics, AIChE Journal 66 (5) (2020) e16931.

[45] R. Rautenbach, H. M. de León, P. Brorens, M. Schlüter, C. Haringa, Reproducibility of gpu-based large eddy simulations for mixing in stirred tank reactors, Computers and Chemical Engineering 210 (2026) 109615.

[46] C. Sullivan, A. Kaszynski, Pyvista: 3d plotting and mesh analysis through a streamlined interface for the visualization toolkit (vtk), Journal of Open Source Software 4 (37) (2019) 1450.

[47] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, L.-S. Young (Eds.), Dynamical Systems and Turbulence, Warwick 1980, Springer Berlin Heidelberg, Berlin, Heidelberg, 1981, pp. 366–381.

Data Processing Pipeline and Conversion Fidelity

LBM-to-FVM Conversion Pipeline

The C2C mapping database required by rCFD is constructed from cell-centered velocity fields. Because M-Star stores velocity data at lattice node positions, the following four-step pipeline was developed to produce compatible input for the rCFD framework.

1. **Node-to-cell conversion:** The lattice-node velocity data exported from M-Star are converted to a cell-centered representation using the `PointToCellData` filter in PyVista[46], an open-source Python (version 3.9) library for VTK-based mesh processing. Cell values are obtained by averaging the nodal values of all lattice points surrounding each cell centroid.
2. **IP file generation:** The converted cell-centered velocity fields are written to Ansys Fluent Interpolation (IP) file format, one of the native data import formats supported by Ansys Fluent, for each stored time step.
3. **rCFD database construction:** Ansys Fluent reads the IP files sequentially, and the rCFD UDF initializes the position tracking scalar fields (ϕ_x, ϕ_y, ϕ_z) from the imported velocity field at each time step to construct the C2C mapping database, as described in Section 2.3.3.
4. **Time extrapolation:** Using the completed C2C database, rCFD performs the time extrapolation as described in Section Time Extrapolation of Scalar Transport.

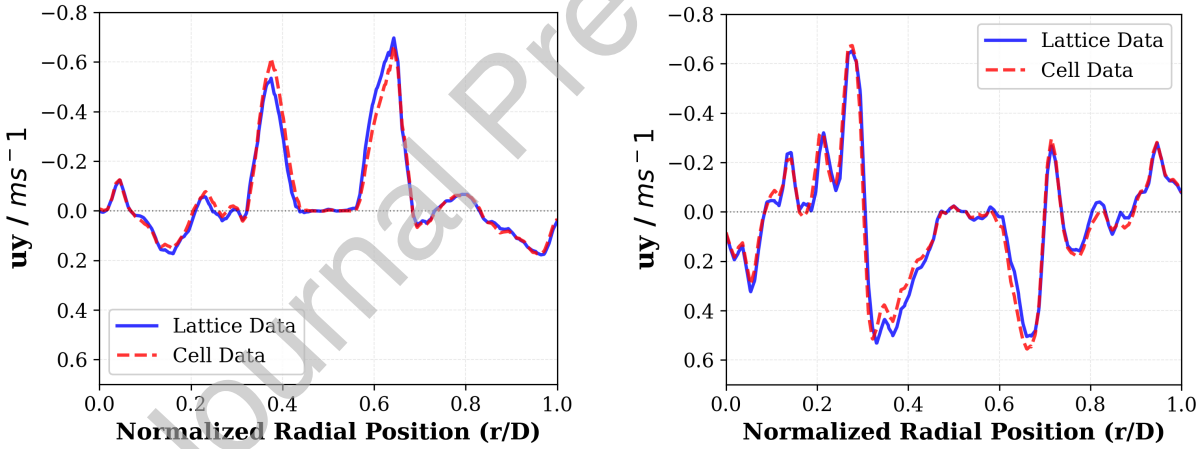
A schematic overview of the pipeline is shown in Figure 21. A comparable approach was employed by Kieckhefen et al. [36], who constructed rCFD input files from CFDEMcoupling data for granulation modeling in a gas-solid fluidized bed.



Figure 21: Node-to-cell data conversion pipeline: lattice data from M-Star are interpolated to cell centers via PyVista and written to Ansys Fluent IP format for rCFD input.

.1. *Conversion Fidelity Assessment*

The `PointToCellData` interpolation does not guarantee strict conservation of all flow quantities, as the averaging of nodal values introduces interpolation error at cell boundaries. To assess the fidelity of the converted velocity field, profiles along horizontal transects were extracted at two characteristic elevations: $y/H = 0.38$, coinciding with the pitched blade turbine plane, and $y/H = 0.18$, coinciding with the Rushton turbine plane. These locations span the two dynamically distinct flow regimes within the dual-impeller system and expose the conversion to high-velocity-gradient conditions near the impeller swept volumes, where interpolation error is largest.



(a) Pitched blade turbine plane ($y/H = 0.38$).

(b) Rushton turbine plane ($y/H = 0.18$).

Figure 22: Comparison of vertical velocity component (at $t = 25$ s) along horizontal transects at two elevations. Solid lines: original LBM (M-Star) data; dashed lines: converted cell-centered data.

Journal Pre-proof

Journal Pre-proof

Table 6: Statistical validation metrics for LBM-to-FVM data conversion.

Metric	Mid-plane ($y/H = 0.38$)	Bottom plane ($y/H = 0.18$)	Overall
Data points (N)	1001	1001	2002
MAE (m/s)	0.0197	0.0302	0.0249
RMSE (m/s)	0.0325	0.0412	0.0371
NRMSE (%)[a]	16.1	17.1	16.7
Mean bias (m/s)	−0.0016	−0.00001	−0.0008
Pearson r	0.986	0.985	0.986
r^2	0.972	0.970	0.972
$u_{\max}$ (m/s)[b]	0.697	0.652	0.697
u_{RMS} (m/s)[b]	0.202	0.241	0.222

[a] $\mathrm{NRMSE} = 100 \cdot \mathrm{RMSE}/u_{\mathrm{RMS}}$

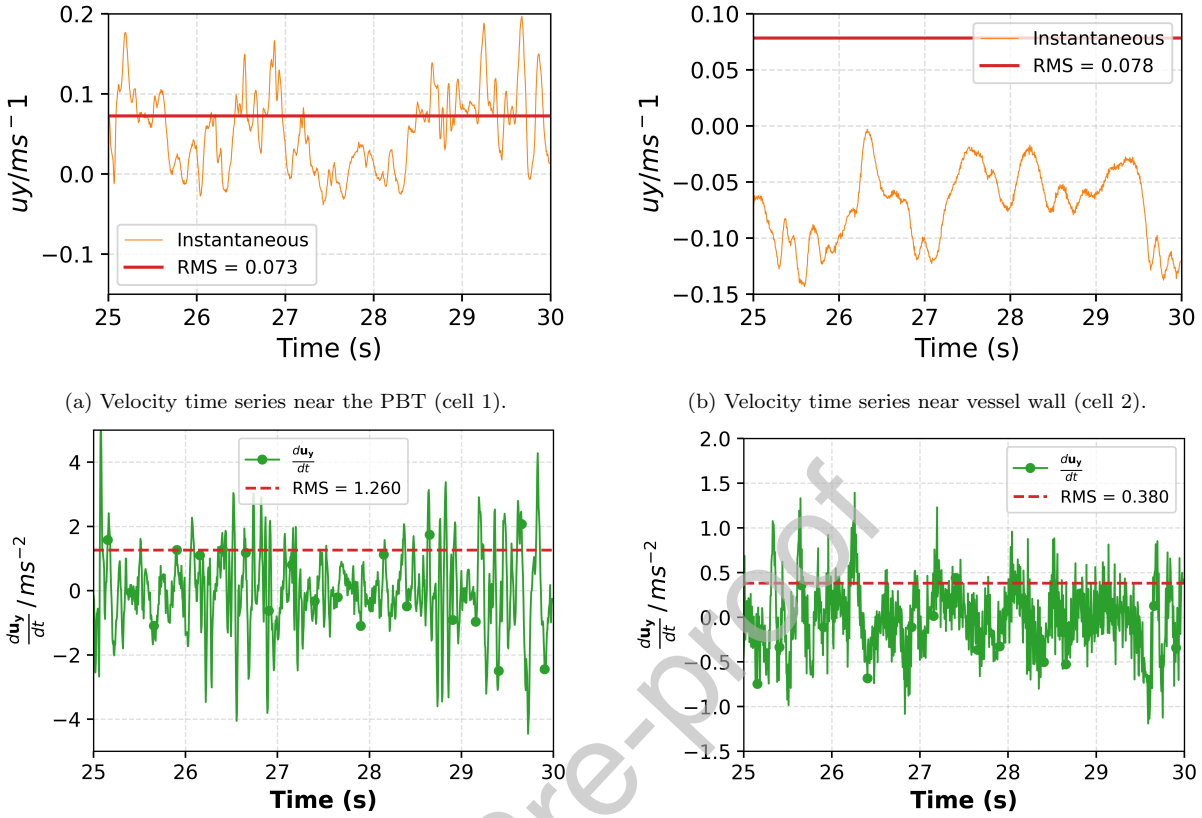
[b] Reference values from original LBM data.

As shown in Figure 22, the converted data closely tracks the original lattice-based velocity field at both elevations. The overall Pearson correlation is $r = 0.986$, $r^2 = 0.972$, and the NRMSE remains below 17% of the local RMS velocity. The mean bias of -0.0008 m/s confirms the absence of systematic offset. Residual discrepancies are largest near the impeller tips ($y/H = 0.18$, NRMSE = 17.1%), attributable to steep velocity gradients that amplify discretization error during node-to-cell mapping. These results confirm that the conversion pipeline preserves the essential flow field structure with sufficient fidelity for use as rCFD input.

Nonlinear Dynamics and Flow Recurrence Analysis

The pseudo-recurrent character of STR flow is not immediately apparent from raw velocity time series. Frequency-domain and phase-space analyses were performed at approximately 500 cells distributed across critical regions of the flow domain; trends were qualitatively consistent across all locations, with differences confined to signal magnitude near the impeller and wall regions. Results are presented for two representative cells: cell 1, located near the pitched blade turbine swept volume, and cell 2, located near the vessel wall.

At cell 1, the instantaneous velocity profile shows a pseudo-periodic character that is not immediately apparent from the raw signal (Figure 23a) but becomes clearly resolved when the acceleration is examined at the same location (Figure 23c). At cell 2, the instantaneous velocity profile (Figure 23b) appears irregular, with a predominantly negative axial component consistent with recirculating flow returning toward the lower impeller. The acceleration signal at cell 2 (Figure 23d) nonetheless reveals a repetitive fluctuation pattern, confirming that pseudo-recurrent dynamics persist even in regions dominated by large-scale recirculation.



(a) Velocity time series near the PBT (cell 1).

(b) Velocity time series near vessel wall (cell 2).

(c) Acceleration time series at cell 1.

(d) Acceleration time series at cell 2.

Figure 23: Instantaneous velocity and acceleration at two representative cells, illustrating the quasi-periodic character of the stirred-tank flow used as rCFD input.

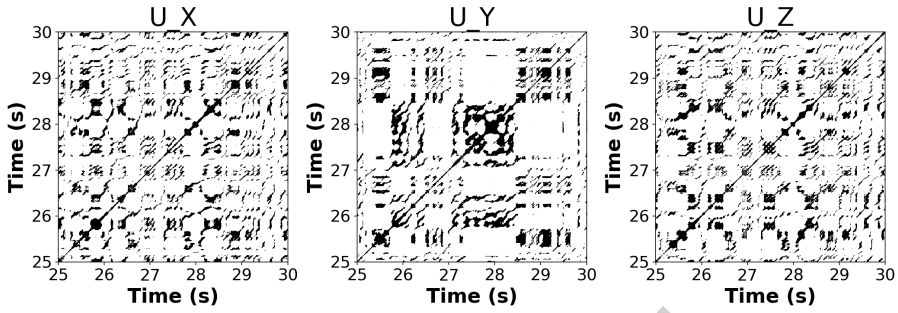
To characterize the underlying flow structure, RQA was applied to the velocity time series at both cells. In the present STR, operated at $n = 200$ rpm ($f_{\mathrm{imp}} = 3.33$ Hz), the 4.5 s database window captures 15 complete impeller revolutions at a sampling frequency of 200 Hz (60 data points per revolution), providing sufficient temporal density to reconstruct the system trajectory in phase space via Takens' embedding theorem [47]. Because high-energy dissipation and turbulence mixing prevent exact return to previous states, the flow is pseudo-recurrent rather than strictly periodic.

Recurrence Plots (RPs) serve as a topological fingerprint of the phase-space trajectory, enabling distinction between three qualitative flow regimes: fully periodic dynamics (long continuous diagonal lines at uniform spacing); purely stochastic fluctuations (uniform scatter of isolated points); and the intermediate pseudo-recurrent regime expected for turbulent impeller-driven flows. The RPs at both cells (Figure 24) exhibit short, parallel diagonal segments frequently interrupted by checkerboard and rectangular block patterns. The diagonal segments confirm deterministic impeller-driven forcing; their truncated length reflects the dissipative, divergent nature of high-Reynolds-number turbulence. This topology is the defining signature of a chaotic pseudo-recurrent attractor: impeller-driven macro-scale memory persists, but small-scale turbulence prevents exact recurrence. The stationarity and

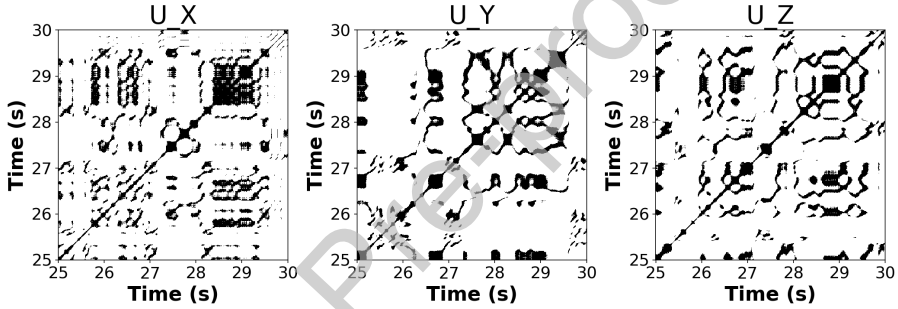
Journal Pre-proof

Journal Pre-proof

repetition of these topological structures across all three velocity components (U_x, U_y, U_z) confirm that the 4.5 s window captures 15 complete and representative cycles.



(a) Cell 1 near the pitched blade turbine.



(b) Cell 2 near the vessel wall.

Figure 24: Recurrence plots of velocity components at two distinct cell locations (recurrence threshold 15%), confirming that the 4.5 s M-Star database captures the pseudo-recurrent behavior of the 30 L STR at 200 RPM.

A. Effect of Time Step Width on Mixing Time and Computational Cost

To investigate the sensitivity of $\tau_{m,95}$ predictions to the rCFD extrapolation time step, simulations were conducted at four values of Δt_{rCFD}: 0.005, 0.010, 0.025, and 0.050 s. For each setting, a dedicated C2C transport database was constructed from the M-Star flow field using the corresponding time interval, tracer was injected at $t_{\mathrm{inj}} = 1$ s, and five independent repetitions were performed, yielding 20 simulations in total. The face-swap diffusion parameter was held fixed at $f_{\mathrm{swap}} = 0.0825$ across all settings, the value calibrated for $\Delta t_{\mathrm{rCFD}} = 0.005$ s in Section Mixing Time Sensitivity to Temporal Injection.

The physical basis for expecting $\tau_{m,95}$ to vary with Δt_{rCFD} is as follows. When the extrapolation step increases, the C2C transport mapping is constructed over a proportionally larger time interval, capturing the net convective displacement of species between two endpoints. However, the intermediate diffusive exchange that would occur at sub-step timescales is not resolved; only the final scalar state at the endpoint of each Δt_{rCFD} interval is carried forward. Since f_{swap} compensates for sub-grid diffusion at each rCFD step, a value calibrated

at $\Delta t_{\mathrm{rCFD}} = 0.005$ s underestimates the effective diffusion contribution at larger step widths, where more intermediate exchange is bypassed. Consequently, $\tau_{m,95}$ is expected to increase with Δt_{rCFD} when f_{swap} is held constant.

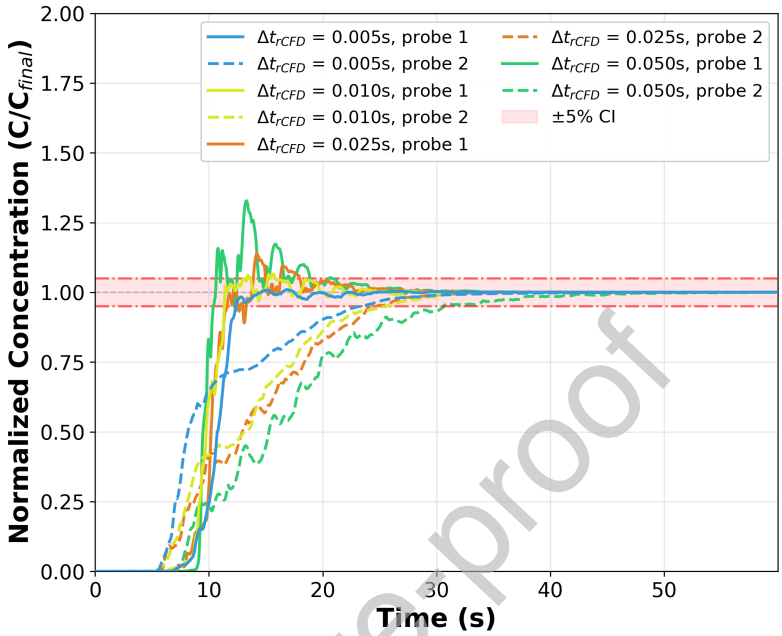


Figure 25: Normalized tracer concentration C/C_{final} at probe 1 (solid) and probe 2 (dashed) for $\Delta t_{\mathrm{rCFD}} = 0.005$, 0.010, 0.025, and 0.050 s ($f_{\mathrm{swap}} = 0.0825$, $t_{\mathrm{inj}} = 1$ s). The ±5% band marks the equilibrium criterion.

Figure 25 shows the normalized tracer concentration at probe 1 and probe 2 for all four time step settings. Probe 1 exhibits a qualitatively and quantitatively similar concentration history across all four values, indicating that scalar transport at this location is predominantly convection-driven and therefore relatively insensitive to Δt_{rCFD}. Probe 2, by contrast, shows a progressive delay in reaching equilibrium as Δt_{rCFD} increases, consistent with its baffle-adjacent position where diffusive exchange contributes more substantially to tracer homogenization. At $\Delta t_{\mathrm{rCFD}} = 0.050$ s, probe 1 additionally exhibits greater concentration fluctuation over time than at smaller step sizes, attributable to the coarser temporal resolution of the C2C shift patterns at this location.

Journal Pre-proof

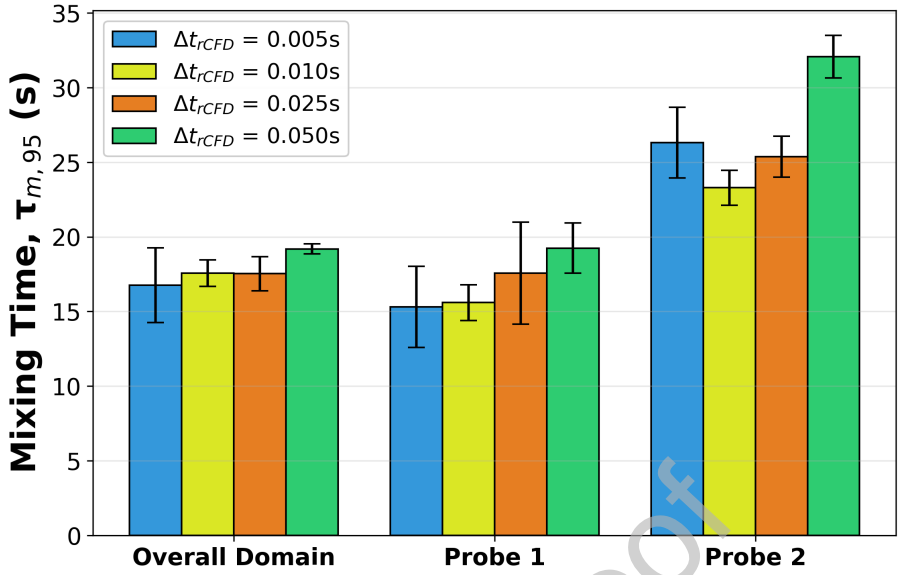


Figure 26: Mean $\tau_{m,95}$ at the overall domain, probe 1, and probe 2 for four Δt_{rCFD} settings. Probe 2 shows the strongest sensitivity, consistent with its diffusion-influenced transport regime.

The effect on $\tau_{m,95}$ is summarized in Figure 26. For the overall domain, $\tau_{m,95}$ increases from 16.75±2.50 s at $\Delta t_{\mathrm{rCFD}} = 0.005$ s to 17.58±0.89 s at 0.010 s, 17.53±1.14 s at 0.025 s, and 19.19 ± 0.34 s at 0.050 s. The values at 0.010 s and 0.025 s are statistically indistinguishable, while the step to 0.050 s produces a more discernible increase of approximately 15% relative to the baseline. Probe 2 shows the largest absolute variation across settings, consistent with the diffusion-sensitive nature of tracer transport at that location, while probe 1 and the domain average remain comparatively stable up to $\Delta t_{\mathrm{rCFD}} = 0.025$ s.

Table 7: Computational efficiency comparison for four rCFD time step widths. Wall-clock time per simulated second is reported as the mean over five independent repetitions.

Δt_{rCFD} (s)	N	Mean (s/s)	Std (s/s)	Min (s/s)	Max (s/s)	Speedup factor
0.005	5	15.19	0.49	14.46	15.61	1.00×
0.010	5	7.56	0.30	7.22	8.03	2.01×
0.025	5	3.08	0.02	3.07	3.11	4.92×
0.050	5	1.73	0.05	1.69	1.83	8.76×

These results indicate that Δt_{rCFD} sensitivity is governed by the relative contribution of diffusion to the local transport regime. Convection-dominated transport is less sensitive to step width, since the dominant displacement is captured by the C2C shift regardless of step size. Liquid- and gas-phase tracers, where diffusive exchange is non-negligible, are inherently more sensitive, as observed here for probe 2. For the present passive-scalar liquid-phase application, $\Delta t_{\mathrm{rCFD}} = 0.025$ s represents a practical upper bound: it achieves a 4.9× reduction in wall-clock time relative to the baseline (see Table 7) while keeping the domain $\tau_{m,95}$ within 5% of the calibration-setting value.

Journal Pre-proof

Journal Pre-proof

The time step dependency identified here stems from a specific limitation of the current implementation: f_{swap} compensates for sub-grid diffusion at the scalar level but is calibrated for a fixed Δt_{rCFD} and does not adapt when the step size changes. A more fundamental resolution is offered by the propagator-based framework introduced by Lumetzberger et al. [41], in which diffusion is incorporated directly into the C2C shift by tracking the zeroth, first, and second moments of the scalar distribution originating from each cell. The resulting propagator kernel carries both convective and diffusive contributions within the mapping itself, making scalar transport predictions inherently independent of Δt_{rCFD}. Transitioning the present STR implementation to this propagator framework is identified as the natural next step toward a time-step-robust rCFD methodology for liquid-phase mixing applications.

Journal Pre-proof