Predicting complex multicomponent particle–liquid flow in a mechanically agitated vessel via machine learning

Li, Kun; Savari, Chiya; Barigou, Mostafa

DOI:
10.1063/5.0142198

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Li, K, Savari, C & Barigou, M 2023, 'Predicting complex multicomponent particle–liquid flow in a mechanically agitated vessel via machine learning', *Physics of Fluids*, vol. 35, no. 5, 053301.
https://doi.org/10.1063/5.0142198

Physics of Fluids

RESEARCH ARTICLE | MAY 01 2023

Predicting complex multicomponent particle–liquid flow in a mechanically agitated vessel via machine learning

Li Kun (李坤) ; Chiya Savari ; Mostafa Barigou

Physics of Fluids 35, 053301 (2023)
https://doi.org/10.1063/5.0142198

Physics of Fluids ARTICLE scitation.org/journal/phf

Predicting complex multicomponent particle–liquid flow in a mechanically agitated vessel via machine learning

Cite as: Phys. Fluids **35**, 053301 (2023); doi: 10.1063/5.0142198
Submitted: 12 January 2023 · Accepted: 8 April 2023 ·
Published Online: 1 May 2023

View Online Export Citation CrossMark

Kun Li (李坤), Chiya Savari, and Mostafa Barigou[a)]

AFFILIATIONS

School of Chemical Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

Note: This paper is part of the special topic, Multiphase flow in energy studies and applications: A special issue for MTCUE-2022.
[a)]**Author to whom correspondence should be addressed:** m.barigou@bham.ac.uk

ABSTRACT

Machine learning (ML) is used to build a new computationally efficient data-driven dynamical model for single-phase and complex multicomponent particle–liquid turbulent flows in a stirred vessel. By feeding short-term trajectories of flow phases or components acquired experimentally for a given flow condition via a positron emission particle tracking (PEPT) technique, the ML model learns primary flow dynamics from the input driver data and predicts new long-term trajectories pertaining to new flow conditions. The model performance is evaluated over a wide range of flow conditions by comparing ML-predicted flow fields with extensive long-term experimental PEPT data. The ML model predicts the local velocities and spatial distribution of each flow phase and component to a high degree of accuracy, including conditions of impeller speeds, particle loadings and sizes within and without the range of the input driver datasets. A new flow analysis and modeling strategy is thus developed, whereby only short-term experiments (or alternatively high-fidelity simulations) covering a few typical flow situations are sufficient to enable the prediction of complex multiphase flows, significantly reducing experimental and/or simulation costs.

I. INTRODUCTION

Machine learning (ML) is an extremely powerful data-driven technique, which automatically builds a certain mathematic model using supplied sample data to make decisions and predictions without being explicitly programmed. It has achieved advances in a diverse range of applications, such as climatology, fluid turbulence, finance, robotics, and neuroscience.[1–3] ML is especially useful for complex dynamical systems whose characteristics of nonlinearity, multiscale, high-dimensionality, and dynamics often limit the use of conventional methods to understand, predict, design, and control them. ML strategies offer an agile and modular modeling framework that can solve specific issues based purely on input data.[4] For example, ML recently played an important role in better understanding how COVID-19 spread, thus helping to inform the fight against the pandemic worldwide.[5,6] Although data-driven ML has achieved significant success in a number of fields, this is still a growing science and there is a need to explore and develop new strategies to widen and improve its applicability to science and engineering.

Mixing flows in mechanically agitated vessels are a typical example of a dynamical system in the study of multiphase fluid mechanics, where the blending of different phases and/or phase components produces a complex dynamic flow behavior. The main objective is to rapidly reduce the inhomogeneity of phase, temperature, and concentration, thus, speeding up mixture production and ensuring good product physical mixing and/or enhancing chemical reaction.[7,8] The selection of a suitable system for a specific mixing application should consider a number of factors including the vessel geometry, impeller type, fluid properties and operating conditions since the resulting mixing flow pattern is a complex function of these parameters.[9] Thus, understanding the complex flow dynamics involved is key to the successful design, operation and optimization of these devices and processes. Conventionally, the prediction of complex flow dynamics is achieved by establishing mechanistic models based on the underlying knowledge of physico-chemical phenomena or simplistic empirical models.[10] However, the high complexity of such flow systems

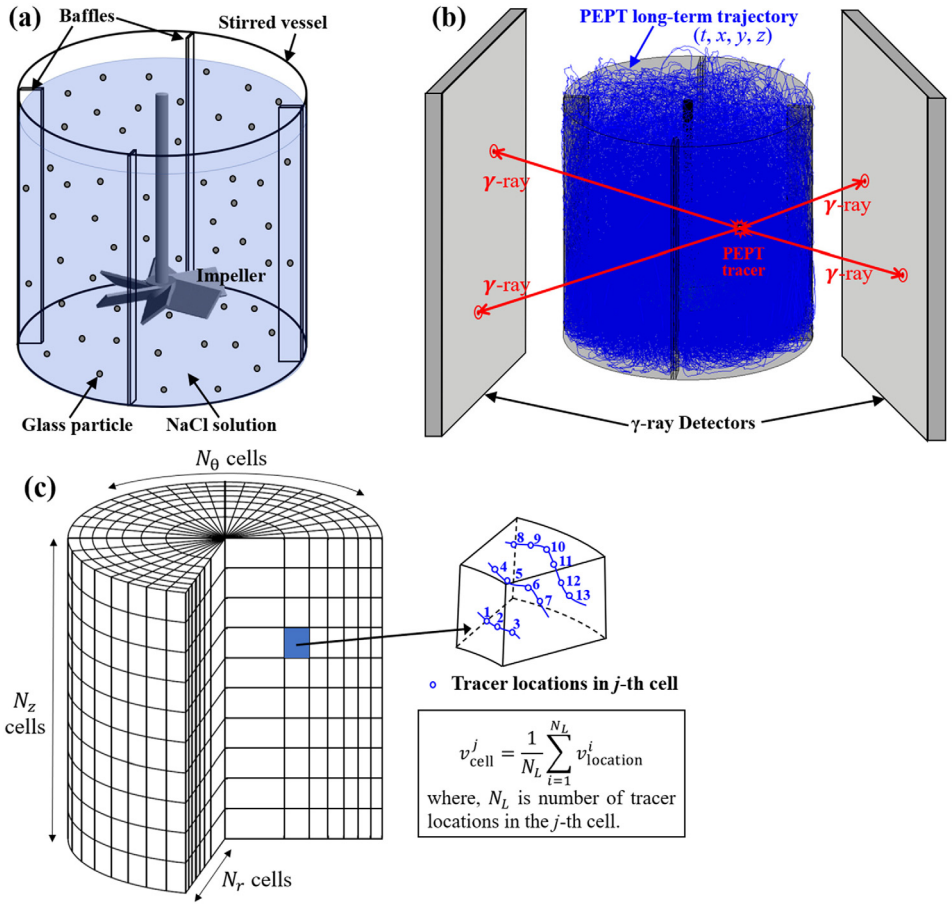

FIG. 1. PEPT experimental setup and computational grid for Lagrangian data analysis: (a) stirred vessel; (b) illustration of PEPT measurement and determined long-term trajectory; and (c) grid of equal-volume cells used for Lagrangian data analysis and method of calculation of local velocity and phase concentration.

often poses challenges for creating accurate models, and numerical solutions may not always be practicable due to extensive computation costs.

Fortunately, the ever increasing availability of high-fidelity experimental and numerical data has opened up new routes for modeling complex flows and, in this respect, data-driven ML methods have attracted significant attention in recent years. One of the main topics of interest is using ML to analyze fluid dynamics phenomena with reduced-order modeling, which projects dynamical methods in reduced form and improves computational efficiency on high-dimensional data. Hasegawa *et al.*[11] used a convolutional neural network autoencoder to extract the evolution of laminar bluff body wakes, which took advantage of the low-dimensional feature of the latent space and captured hidden wakes around a body of an arbitrary shape. Another promising topic is closure modeling or establishing physically informed models that improve the speed or accuracy of conventional computational fluid dynamics (CFD) models. For example, Ling *et al.*[12] applied ML to identify and model the Reynolds stress tensor discrepancies between the Reynolds averaged Navier–Stokes (RANS) model and high-fidelity simulations. Hou *et al.*[13] used neural networks to predict flow properties and detect flow disturbances in dynamical systems, and Zhai *et al.*[14] developed a semi-physics-informed neural network to predict the micro-bubble dynamics in bubbly flows. Moreover, ML was applied to flow pattern identification, sensor placement, and flow control,[15] and an adaptive sensing and control strategy was developed to arrange the sensor placement for obtaining maximal information.[16] Rabault *et al.*[17] used artificial neural networks to discover control strategies for active flow control.

More recently, a number of other data-driven theoretical methodologies based on high-fidelity driver data have also been reported, but have not been widely used to model and analyze engineering fluid flows, including Lagrangian stochastic modeling (LSM), Lagrangian recurrence tracking and Lagrangian coherent structure detection.[18–20] Sheikh *et al.*[20,21] developed a data-driven Lagrangian stochastic model (LSM) to simulate single-phase and particle–liquid flows in stirred vessels, which was driven by experimental local velocity measurements in conjunction with decorrelation statistics. LSM has also been used to model geophysical ocean flows for underwater vehicles,[22] to predict the mixing and transport of pollutants in water or atmosphere,[23,24] and to study solar wind turbulence and turbulent combustion.[25,26]

Recently, we successfully developed a ML modeling framework to reconstruct the turbulent flow fields of single-phase and two-phase flows in a stirred vessel. The strategy relied on feeding a very short-term Lagrangian trajectory of the phase concerned experimentally determined by a positron emission particle tracking (PEPT)

technique.[27] The method proved very efficient in producing the corresponding long-term Lagrangian trajectory of the particular phase, having the same statistical features learned from the short driver dataset input, which greatly reduces the experimental costs and/or numerical costs needed for complex flow simulations. This work represents another branch of ML applications in fluid dynamics, i.e., flow field and parameter estimation via machine learning from limited data input.

In this study, we extend the capability of the ML framework to predict the flow developed under conditions different to those under which the limited sets of driver data are experimentally acquired. In addition to different conditions of single-phase flow, the strategy is also extended to complex two-phase multicomponent particle–liquid flows. The framework strategy is borrowed from the Reynolds averaging concept, where the instantaneous flow field is approximated by a mean flow field coupled with a Gaussian distributed fluctuation. The mean flow field is predicted by a k-nearest neighbors (KNN) regressor, which is trained by the driver Lagrangian trajectories data, and the fluctuation is produced by a Gaussian noise generator with the same statistical pattern corresponding to the driver flow conditions. Extensive experimental PEPT data are used to validate the ML framework including the local velocity field and spatial distribution of each flow phase and component involved.

II. EXPERIMENTAL AND TURBULENT DYNAMIC ANALYSIS

A. Mixing apparatus

Single-phase and complex particle–liquid mixing experiments were conducted in a stirred vessel of standard configuration with diameter $T = 288$ mm (radius $R = 0.5T$), fitted with four wall baffles of width $0.1T$ and filled to a height $H = T$. The vessel was agitated by a down-pumping six-blade 45° pitched-turbine (PBTD) of diameter $D = 0.5T$, blade height $0.1T$, and off-bottom clearance $0.25T$, as depicted in Fig. 1(a). In single-phase flow tests, the vessel was filled with NaCl solution (density 1150 kg/m^3). Agitation speeds ranged from 60 to 540 rpm, corresponding to a fully turbulent regime, i.e., impeller Reynolds number $Re_{imp} \geq 24 \times 10^3$. In particle–liquid flow experiments, two different types of ballotini suspensions (density 2485 kg/m^3) were used, i.e., monodisperse and polydisperse. The suspending medium was an aqueous NaCl solution. The monodisperse suspensions were studied at mean solid mass loadings of $C_m = 5$ and 20 wt. %, and agitated at impeller speeds ranging from the minimum speed for particle suspension N_{js} up to $2N_{js}$. The polydisperse suspensions consisted of five particle sizes of equal mass fraction and overall mean solid mass loadings of $C_m = 5$, 10, 20, and 40 wt. % and were agitated at an impeller speed corresponding to N_{js}. In each case, the N_{js} speed was experimentally determined based on the well-known Zwietering criterion.[28] The conditions of the experiments conducted are summarized in Table I.

B. PEPT measurements

PEPT allows noninvasive imaging of opaque flows in opaque devices by using a representative positron-emitting particle tracer to track the three-dimensional (3D) motion of each phase.[29] In a typical PEPT experiment, a radiolabelled tracer is introduced in the flow and its movement is recorded, providing a long-term Lagrangian trajectory, as shown in Fig. 1(b). The liquid phase was tracked using a ~600 μm neutrally buoyant resin particle tracer (note that NaCl was added to water to match the density of the resin tracer). Each component of the particle phase was individually tracked using a representative glass bead tracer taken from the particle fraction considered. Being able to visualize opaque flows in 3D with a comparable accuracy to leading optical techniques, such as particle image velocimetry

TABLE I. Experimental flow conditions.

Flow system	C_m (wt. %)	C_v (vol. %)	d (mm)	T (mm)	N (rpm)	Re_{imp} ($\times 10^5$)
Single-phase	...	...	...	288	60	0.24
				288	100	0.40
				288	150	0.60
				288	260	1.03
				288	330	1.31
				288	400	1.59
				288	500	1.98
				288	540	2.13
Monodisperse particle–liquid	5	2.5	3	288	360 (N_{js})	1.43
	5	2.5		288	540 ($1.5N_{js}$)	2.13
	5	2.5		288	720 ($2N_{js}$)	2.84
	20	10.4		288	490 (N_{js})	1.94
	20	10.4		288	613 ($1.25N_{js}$)	2.42
	20	10.4		288	735 ($1.5N_{js}$)	2.91
Polydisperse particle–liquid	5	2.5	1.1, 1.7, 2.1, 2.7, 3.1	288	380 (N_{js})	1.50
	10	5.2	(equal mass fraction = $C_m/5$)	288	450 (N_{js})	1.78
	20	10.4		288	510 (N_{js})	2.01
	40	23.6		288	610 (N_{js})	2.41

(PIV)/laser Doppler velocimetry (LDV),[30,31] gives the PEPT technique a unique advantage. PEPT has been extensively used to study single and multiphase flows in pipes and stirred vessels and more details can be found in our earlier papers.[18,19,21,29,32–37] Here, the Lagrangian trajectories of the carrier fluid and the suspended particles were acquired over a period of ~40 min, providing ample data to reliably study their flow behaviour.[33]

C. Dynamic analysis of turbulent flow field

Three-dimensional (3D) trajectories were acquired for each individual flow phase/component in the mechanically agitated vessel described above, in the form of data arrays that consist of one temporal and three spatial coordinates (t, x, y, and z). After converting the Cartesian coordinates to cylindrical coordinates, the local instantaneous velocity ($\boldsymbol{v}$) of a given flow phase/component was estimated at each location from the discrete-time derivation of its Lagrangian trajectory. By applying a pre-defined cylindrical mesh with $N_r \times N_z \times N_\theta$ equal-volume cells, as illustrated in Fig. 1(c), the Eulerian (ensemble-averaged) velocity field ($\overline{\boldsymbol{v}}$) was obtained by

$$\overline{\boldsymbol{v}}^j = \frac{1}{N_L}\sum_{i=1}^{N_L} \boldsymbol{v}^i, \tag{1}$$

where $\boldsymbol{v}^i$ and $\overline{\boldsymbol{v}}^j$ are, respectively, the local instantaneous velocity at the ith detected location of the trajectory and the mean velocity in the jth cell of the mesh. N_L is the total number of detected tracer locations in the jth cell [Fig. 1(c)]. Then, the local fluctuating velocity component ($\boldsymbol{v}'$) at each location is

$$\boldsymbol{v}' = \boldsymbol{v} - \overline{\boldsymbol{v}} = [v_r,\ v_z, v_\theta] - [\overline{v}_r,\ \overline{v}_z, \overline{v}_\theta], \tag{2}$$

where $v_r,\ v_z, v_\theta$ and $\overline{v}_r,\ \overline{v}_z, \overline{v}_\theta$ are the instantaneous and mean velocity components in the radial, axial and tangential directions, respectively. Thus, a new Lagrangian data array containing the time,

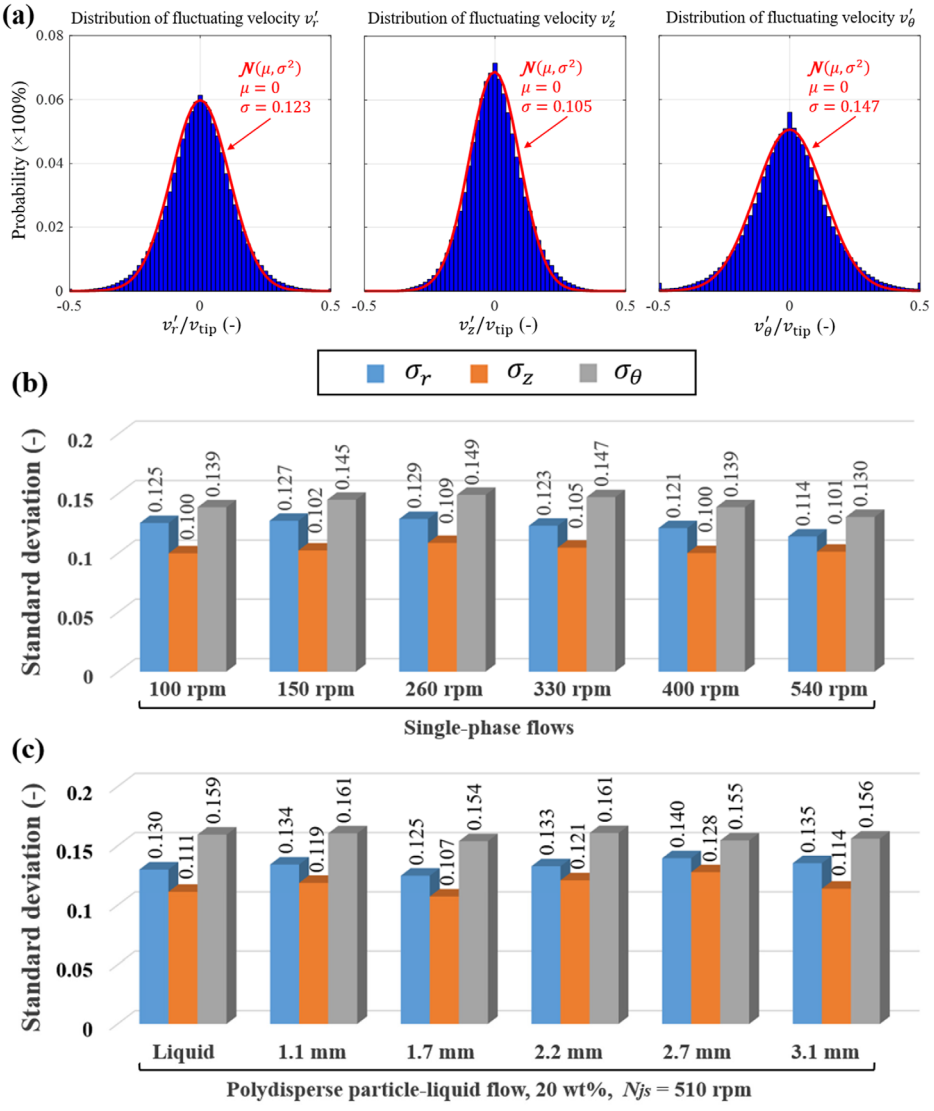

FIG. 2. Global fluctuating velocity components in single and two-phase flows obtained from PEPT measurements over a wide range of conditions fitted by a Gaussian distribution: (a) typical distribution of radial, axial and tangential fluctuating velocity components (single-phase flow, $N = 330$ rpm); (b) standard deviation of fluctuating velocities in single-phase flow; and (c) standard deviation of liquid and solid fluctuating velocities in a typical polydisperse particle–liquid flow.

cylindrical coordinates, instantaneous and fluctuating velocity components of each phase in the flow was generated, i.e., $[t, r, z, \theta, v_r, v_z, v_\theta, v'_r, v'_z, v'_\theta]$.

In order to extract the fluctuation characteristics of a turbulent flow for feeding to the ML framework, the fluctuating velocity components of each phase/component in the single and two-phase flows considered were statistically analyzed for all experimental conditions, as depicted in Fig. 2(a), and their local distributions were all well represented by a Gaussian distribution with a zero mean value since most of local instantaneous velocities are very close to the local mean velocity. As reported in previous works, the Gaussian nature of these fluctuating velocity components is approximately invariant throughout the whole flow field.[38–40] Consequently, the global fluctuating velocity components in the vessel were normalized and fitted by

$$v'_i/v_{\text{tip}} \sim \boldsymbol{N}\left(\mu_i, \sigma_i^2\right), \tag{3}$$

where v_{tip} is the impeller tip velocity, and μ_i and σ_i are, respectively, the mean and standard deviation of the normalized fluctuating velocity component in the i-direction.

Comparing the standard deviation of fluctuating velocity components in Fig. 2, an ascending order $\sigma_\theta > \sigma_r > \sigma_z$ is observed in both

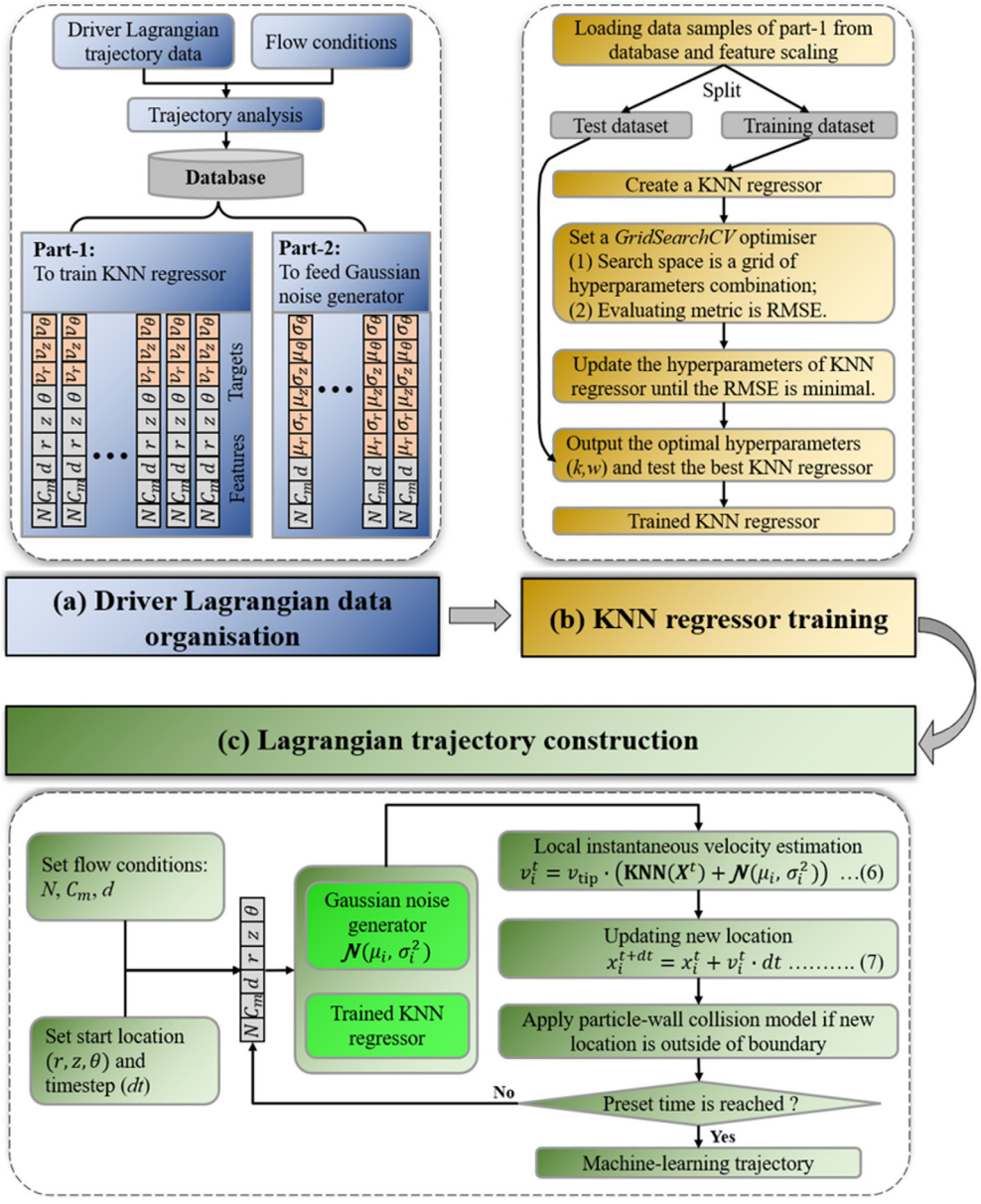

FIG. 3. Flowchart of ML model for Lagrangian trajectory construction in turbulent flow in a stirred vessel.

single and two-phase flows. The fact that the largest standard deviation belongs to the tangential fluctuations is mainly because of the periodic impeller motion and the vortex breaking effect of the wall-mounted baffles. In single-phase flows, the standard deviation is not significantly affected by the agitation speed over the range considered, as depicted in Fig. 2(b), which indicates that the flows studied belong to the same regime (i.e., fully turbulent). In a particle–liquid flow at a given condition, there is no significant difference in the fluctuation characteristics between each individual phase, as shown in Fig. 2(c). Comparing the standard deviation between different flow systems in Figs. 2(b) and 2(c), the fluctuations in two-phase flow are slightly greater than those in single-phase flow, which is consistent with reports that adding large solid particles results in an increase in the turbulence intensity of the carrier phase.[41–43] Hence, the statistical features of fluctuating velocity components in similar flow regimes are close to each other, i.e., flows of similar regimes possess a similar fluctuation pattern.

III. MACHINE-LEARNING MODELING FRAMEWORK

The proposed ML framework consists of three main modules, as illustrated in Fig. 3: (a) driver Lagrangian data organization; (b) KNN regressor training; and (c) Lagrangian trajectory construction. First, the driver data which consists of short experimental Lagrangian trajectories are analyzed and the extracted flow characteristics (i.e., instantaneous velocities and distribution of fluctuating velocities) are stored in a database. Then, a KNN regressor is trained to learn the primary flow pattern governing the flow field. By setting new flow conditions similar to those of the driver data, a corresponding new long-term Lagrangian trajectory is constructed by advancing a seed tracer throughout the instantaneous velocity field that is approximated by the KNN-predicted mean velocity coupled with a Gaussian fluctuation. The key equations and parameters are presented and discussed in Secs. III A and III B.

A. Key equations

As described in Sec. II C, the dynamic analysis of the input driver trajectories provides the local instantaneous velocities and the global fluctuation distributions in each direction, which are combined with the flow operating conditions used namely, the impeller rotational speed (N), particle mass concentration (C_m) and particle size (d), and organized into two parts in the database. Part-1 data are encoded in the format $[N, C_m, d, r, z, \theta, v_r, v_z, v_\theta]$ and used to train the KNN regressor, while Part-2 data formatted as $[N, C_m, d, \mu_r, \sigma_r, \mu_z, \sigma_z, \mu_\theta, \sigma_\theta]$ are used to feed the Gaussian noise generator.

In the KNN training process, normalization is first applied to rescale the data into a unit scale to weaken the multiple features spanning effect caused by different units and magnitudes of features, whereby all features $[N, C_m, d, r, z, \theta]$ are normalized using their

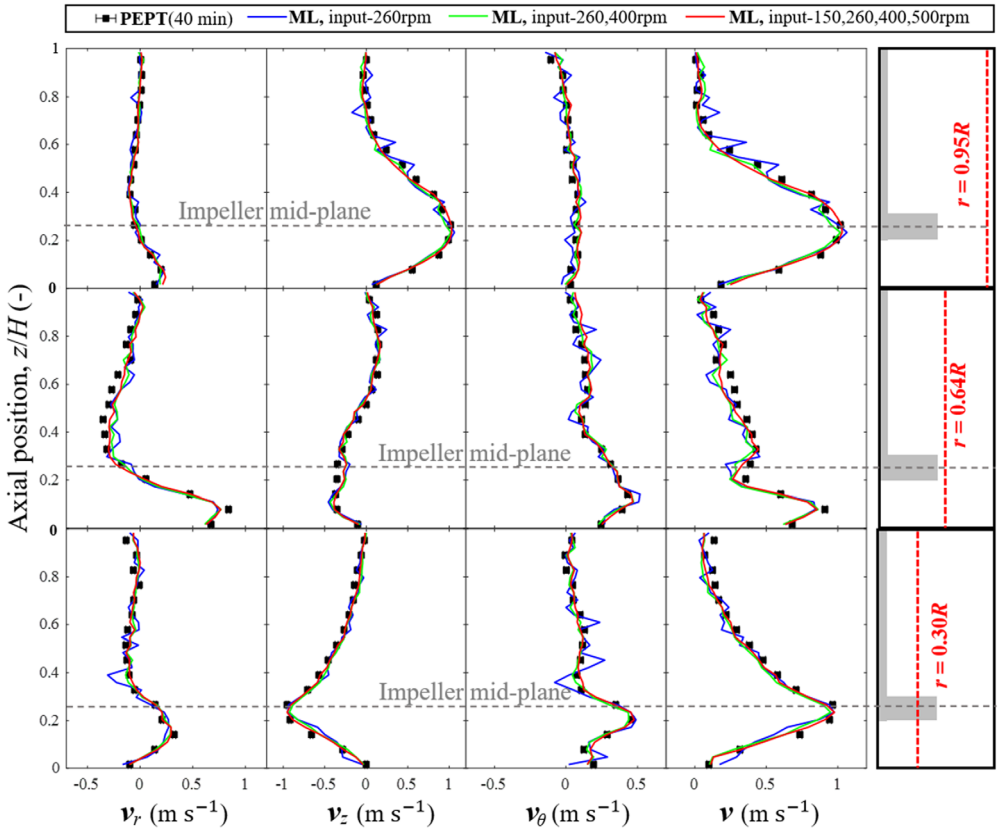

FIG. 4. Sample case showing the effects of using one or more 5 min driver datasets corresponding to different agitation speeds, to predict the long-term 40 min trajectory corresponding to $N = 330$ rpm: ML model predictions and PEPT measurement compared in single-phase flow.

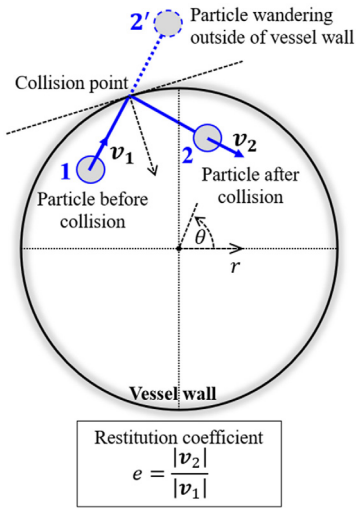

FIG. 5. Illustration of particle–wall collision model.

maximum value, i.e., $N/N_{\max}$, $C_m/C_{m-\max}$, $d/d_{\max}$, r/R, z/H, θ/π, and targets $[v_r, v_z, v_\theta]$ are normalized by the impeller tip velocity, v_{tip}. The KNN training process is illustrated in Fig. 3(b). First, the data in Part-1 are randomly split into a training dataset (85%) and a test dataset (15%). A pipeline is then created to search for the best hyperparameters for the KNN regressor using a *GridSearchCV* algorithm.[44] During each step of the search, the performance of the trained KNN regressor is independently validated by the tenfold cross-validation technique,[45] and the best KNN regressor is determined by searching for the minimum of the root mean square error (RMSE). The test dataset is then used to recheck the model performance. The trained KNN regressor remembers the correspondence between normalized features $\boldsymbol{X}$ and targets $\boldsymbol{V}$, and stores the training data in a fast-indexing structure. Given a new query instance $\boldsymbol{X}_i$, the KNN regressor rapidly indexes a number (k) of neighboring instances in the training data and returns their weighted average as the target $\boldsymbol{V}_i^{\text{pred}}$ of the new instance, as follows:

$$\boldsymbol{V}_i^{\text{pred}} = \mathbf{KNN}(\boldsymbol{X}_i) = \sum_{j=1}^{k} (w_{ij} \times \boldsymbol{V}_j), \tag{4}$$

where $\boldsymbol{V}_j$ is the target of the jth instance in the training data; w_{ij} is the distance weight function that is the inverse of the feature distance L_p between the query instance $\boldsymbol{X}_i$ and training instance $\boldsymbol{X}_j$, thus

$$w_{ij} = \frac{1}{L_p(\boldsymbol{X}_i, \boldsymbol{X}_j)} = \left(\sum_{l=1}^{n} \left| X_i^{(l)} - X_j^{(l)} \right|^p \right)^{-\frac{1}{p}}, \tag{5}$$

where p is a constant for different types of distance function and $X_j^{(l)}$ is the lth dimensional feature of the jth instance. In this study, the value of l is 6 which is the number of feature dimensions; the value of p is set to 2 to correspond to the normalized Euclidean distance to calculate the feature distance (L_p).[44]

When the ML framework is applied to construct a trajectory for a given phase/component in a new multiphase system operated under a similar flow regime, the flow operating conditions (N, C_m, d) and the start location $\boldsymbol{x} = (r, z, \theta)$ of a virtual seed tracer should be set to initialize the simulations. These initial data are then combined into a

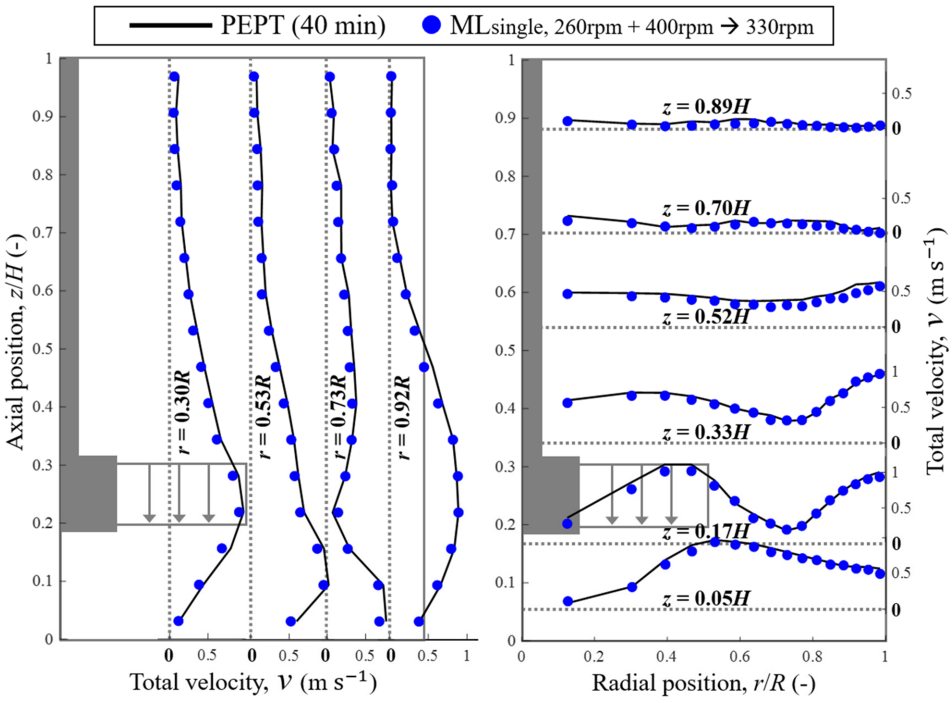

FIG. 6. Azimuthally averaged axial and radial profiles of total flow velocity: ML model and PEPT compared in single-phase flow; $N = 330$ rpm.

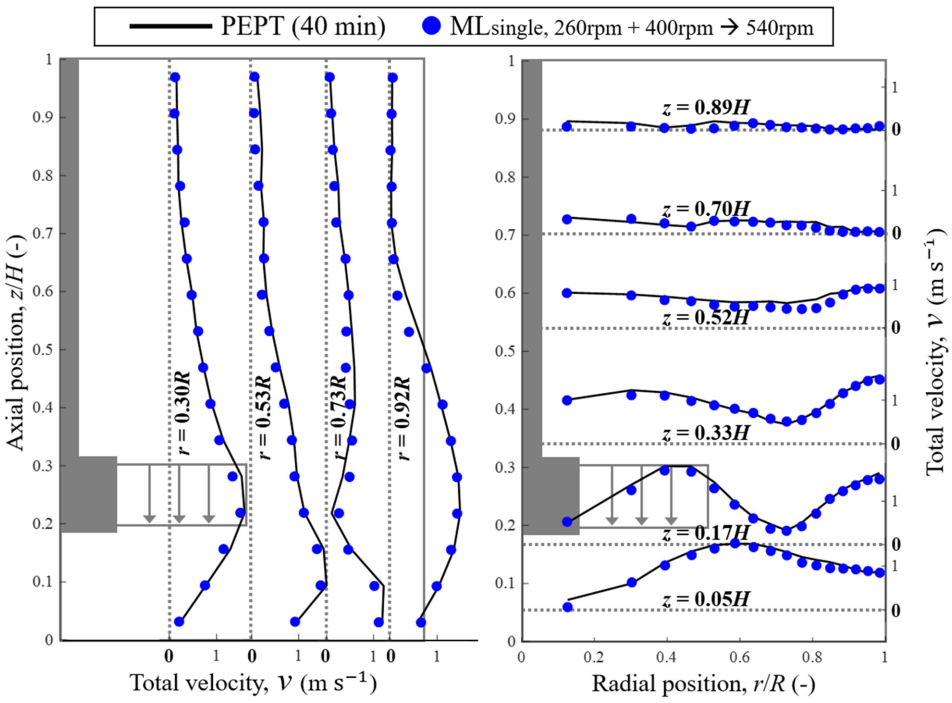

FIG. 7. Azimuthally averaged axial and radial profiles of total flow velocity: ML model and PEPT compared in single-phase flow; $N = 540$ rpm.

new instance $\boldsymbol{X}$ that requests a mean velocity from the trained KNN regressor and a fluctuation from the Gaussian noise generator. Thus, the ML framework approximates the local instantaneous velocity field by

$$v_i^t = \overline{v_i}^t + v_i' t = v_{\text{tip}} \times \left(\mathbf{KNN}(\boldsymbol{X}^t) + \boldsymbol{N}\left(\mu_i,\ \sigma_i^2\right)\right), \tag{6}$$

where v_i^t, $\overline{v_i}^t$, and $v_i' t$ are, respectively, the instantaneous, mean, and fluctuating velocities of the tracer in the ith direction at time t.

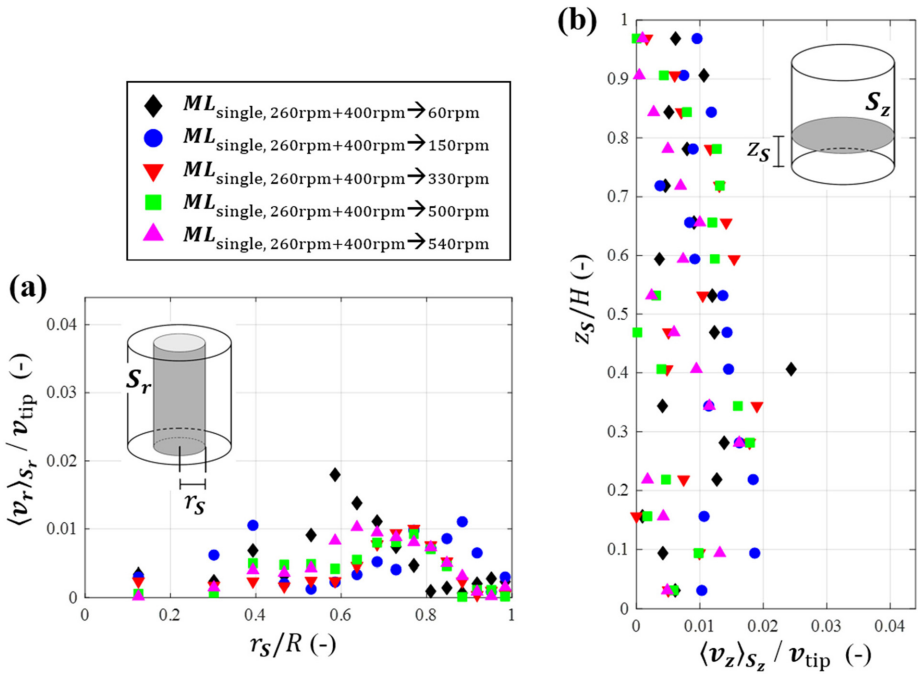

FIG. 8. Mass continuity verification for ML-modelled single-phase flows: (a) normalized mean radial velocity averaged over cylindrical envelope $\boldsymbol{S}_r$; (b) normalized mean axial velocity averaged over horizontal surface $\boldsymbol{S}_z$.

Subsequently, the new location $\boldsymbol{x}^{t+dt}$, after time step dt, is updated by advancing the seed tracer through the local velocity field in the whole vessel volume, thus

$$x_i^{t+dt} = x_i^t + v_i^t \times dt, \tag{7}$$

where x_i^t is the tracer location in the ith direction at time t. Equations (6) and (7) are iterated until the pre-set time length is reached. Thus, the ML framework constructs a long-term flow trajectory for the selected phase/component having characteristics closely similar to those learned from the input driver data.

B. Key parameters

To implement the ML framework, several key parameters should be determined and used to simulate the new phase/component trajectory of flow systems in stirred vessels. First and foremost is the minimum amount of input Lagrangian driver data, which should include tracer visits to most of the grid cells in the vessel domain, and which serves as a benchmark for the flow pattern. Trial and error tests showed that using a 5 min trajectory was sufficient to achieve an excellent agreement between ML-predicted flow fields and PEPT measurements for the same flow condition.

The aim here is to use ML to predict long-term Lagrangian trajectories pertaining to new flow conditions; hence, several short-term trajectories from different flow conditions should be utilized to drive the ML framework. For example, to predict the single-phase flow field in a vessel agitated for any given value of N, different input driver datasets obtained at various agitation speeds should be used, and similarly for other flow variables. A sample case is illustrated in Fig. 4. The three velocity components and total velocity are computed from 40 min trajectories predicted by the ML approach for $N = 330$ rpm, using driver datasets consisting of individual trajectories (5 min each) corresponding to: (i) one agitation speed (260 rpm); (ii) two agitation speeds (260, 400 rpm); and (iii) four agitation speeds (150, 260, 400, and 500 rpm). Comparison of results with PEPT measurements shows that using a single 5 min driver dataset corresponding to 260 rpm, overall yields a good prediction of the velocity field at 330 rpm despite some minor discrepancies throughout the vessel. However, feeding more driver data corresponding to more agitation speeds clearly enhances predictability, with no significant improvements being obtained from four datasets compared to two, as depicted in Fig. 4. Similar tests showed that this also applies to other flow variables, e.g., C_m or d.

Moreover, the quality of the driver datasets is also crucial since it serves as a benchmark, i.e., a poor driver dataset (inaccurate experimental data) would mislead the ML framework. Henceforth, for optimum accuracy and computational efficiency, two datasets of short-term Lagrangian trajectories will be used to drive the framework. Furthermore, intensive tests showed that the prediction accuracy of the three velocity components is totally reflected in the total velocity and to avoid duplication of the information presented, therefore, only the total velocity profiles will be presented in Sec. IV. It should be noted, however, that the flow to be predicted should not pertain to a hydrodynamic regime in terms of flow (e.g., turbulent) and particle suspension (e.g., just-suspended) which is too dissimilar to the one from which the driver data are obtained. In other words, extrapolation from one hydrodynamic regime to a completely different one is not currently possible.

The next key parameter is k in Eq. (4), which is the number of instances in the training data used to predict a new instance,

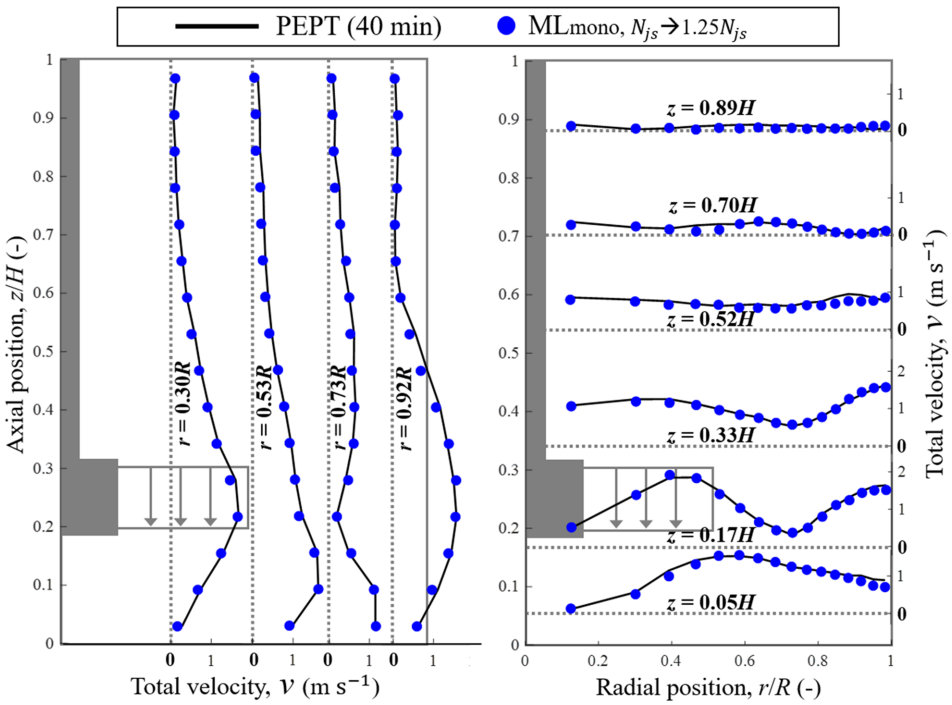

FIG. 9. Azimuthally averaged axial and radial profiles of total liquid phase velocity: ML model and PEPT compared in monodisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $N = 1.25N_{js}$ (613 rpm).

determining the performance of the KNN regressor. Too large or too small a k value will produce serious errors by under-fitting or over-fitting the training data. In other words, as k increases, the ML framework performance gradually improves to an optimum point beyond which it starts deteriorating. In general, this optimum k value is automatically determined by minimizing the KNN prediction error during the training process [Fig. 3(b)].

During construction of the predicted Lagrangian trajectory, a particle–wall collision model is also needed to avoid the particle tracer wandering outside of the flow boundary, as depicted in Fig. 5. The ratio of the tracer velocities before and after collision is defined as the restitution coefficient, e

$$e = \frac{|\boldsymbol{v}_2|}{|\boldsymbol{v}_1|}, \tag{8}$$

where $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the velocity vectors before and after the particle–wall collision, respectively. The restitution coefficient normally ranges from 0 to 1, where 0 means a perfectly inelastic collision and 1 means a perfectly elastic collision. A sensitivity test for the restitution coefficient showed that the ML framework predictions are not very sensitive to the selected value,[27] provided it is not zero (perfectly inelastic collision) to avoid particles clustering at the wall. Hence, a perfectly elastic collision ($e = 1$) was assumed in this work.

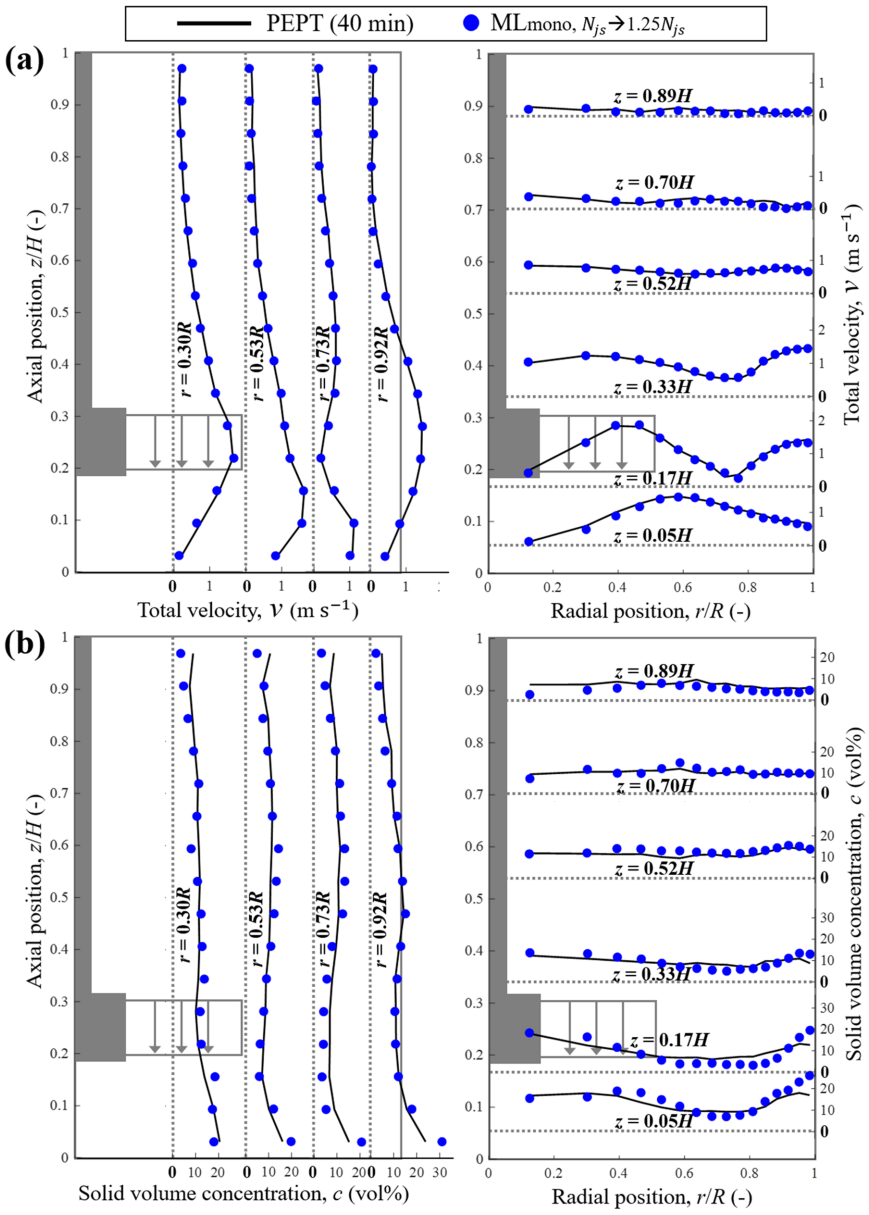

FIG. 10. Azimuthally averaged axial and radial profiles of (a) total particle velocity and (b) particle volume concentration: ML model and PEPT compared in monodisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $N = 1.25N_{js}$ (613 rpm).

IV. RESULTS AND DISCUSSION

The ML model was implemented using the Python language, to predict long-term Lagrangian trajectories in single-phase and complex multicomponent particle–liquid turbulent flows under a range of conditions ($N = 60$–540 rpm and $C_m = 5$–40 wt. %). The typical computation time for predicting and constructing a 40 min ML trajectory with a typical 5 ms time step was around 2 h, which is orders of magnitude less than most conventional numerical simulations of turbulent flows. Note that the size of the time step does not affect the accuracy of the results provided it is on the order of milliseconds; 5 ms was selected to match the time step of the acquisition of the experimental PEPT data. For each application, the ML framework was fed with short-term Lagrangian trajectory driver datasets, as described above. To evaluate the framework, the azimuthally averaged profiles of local total velocity and solid phase concentration predicted by ML were compared with the PEPT measured long-term profiles. Comparison was performed axially over four cylindrical envelopes spanning the vessel radius ($r = 0.30R$–$0.92R$) and radially over six horizontal planes spanning the height of the vessel ($z = 0.05H$–$0.89H$). Further validation was conducted via verification of mass continuity for all ML modeling cases.

A. Evaluation of ML model in single-phase flows

To evaluate the capability of the dynamical model to predict the velocity field in single-phase flows, two short-term (5 min) trajectories measured by PEPT at $N = 260$ and 400 rpm were used as input driver data. The flow fields corresponding to agitation speeds of 330 rpm (interpolation within the range of input driver data, $\boldsymbol{ML}_{\text{single, 260rpm+400rpm}\rightarrow\text{330rpm}}$) and 540 rpm (extrapolation outside the range of input driver data, $\boldsymbol{ML}_{\text{single, 260rpm+400rpm}\rightarrow\text{540rpm}}$) were then computed from the long-term (40 min) trajectories predicted by the ML model. Comparison of the ML-predicted and PEPT-measured velocity fields is presented in Figs. 6 and 7, showing an excellent agreement for both 330 and 540 rpm. Hence, the ML model is capable of predicting flow both within and without the range of impeller speeds pertaining to the supplied driver data. More extensive validation was performed using PEPT measurements in fully turbulent single-phase flows across a wide range of agitation speeds (see Table I: $N = 60$–540 rpm). More results are depicted in Fig. S1 in the supplementary material.

Mass continuity calculations were performed to verify the reliability and accuracy of the ML-predicted flow data, i.e., the net mass flux through any given closed surface $\boldsymbol{S}$ should be zero, thus

$$\sum_{S} \rho \boldsymbol{v} \cdot \Delta \boldsymbol{S} \cong 0, \tag{9}$$

where ρ is the phase density. Note that the term $\boldsymbol{v} \cdot \Delta \boldsymbol{S}$ is always zero over the base and wall of the vessel. Therefore, the closed surface $\boldsymbol{S}$ can be reduced to a cylindrical envelope $\boldsymbol{S}_r$ or a horizontal plane $\boldsymbol{S}_z$, as shown in Fig. 8, and the surface element $\Delta \boldsymbol{S}$ is then replaced by $\Delta \boldsymbol{S}_r$ or $\Delta \boldsymbol{S}_z$. Using equal-volume cells [Fig. 1(c)], $\Delta \boldsymbol{S}_r$ or $\Delta \boldsymbol{S}_z$ for each cell is constant, and the phase density is also constant. Hence, the left term in Eq. (9) reduces to the average axial velocity over a horizontal plane (i.e., $\langle \boldsymbol{v}_z \rangle_{S_z}$) or the average radial velocity over a cylindrical envelope (i.e., $\langle \boldsymbol{v}_r \rangle_{S_r}$). Extensive tests of ML-predicted single-phase flows show a very good verification of mass continuity, i.e., the average velocities $\langle \boldsymbol{v}_z \rangle_{S_z}$ and $\langle \boldsymbol{v}_r \rangle_{S_r}$ are close to zero throughout the vessel, generally less than 2% of the impeller tip velocity v_{tip}, as depicted in Fig. 8.

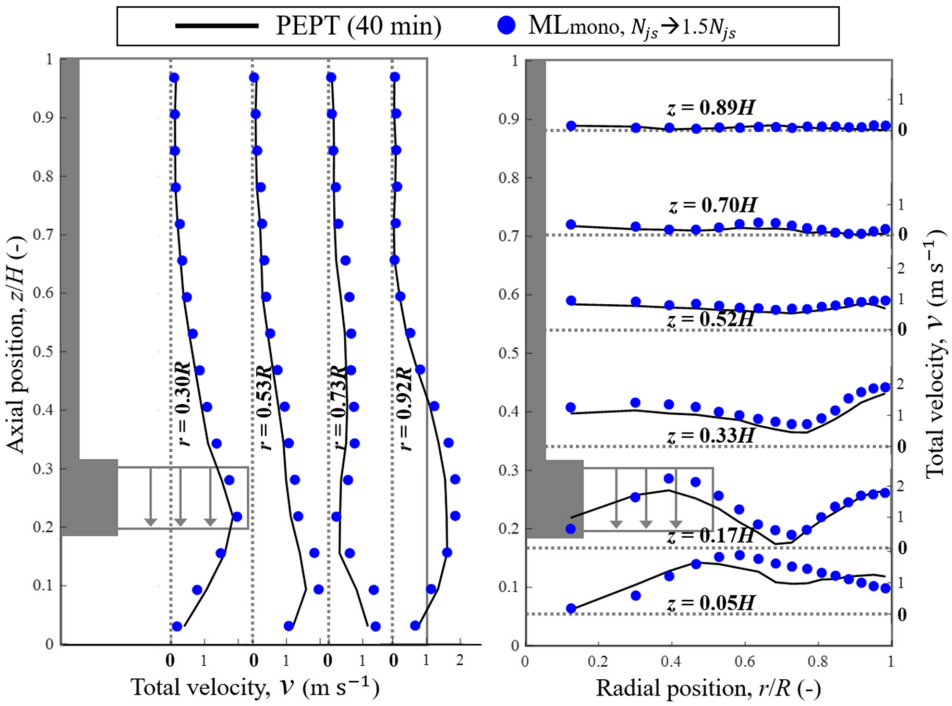

FIG. 11. Azimuthally averaged axial and radial profiles of total liquid phase velocity: ML model and PEPT compared in monodisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $N = 1.5N_{js}$ (735 rpm).

B. ML model evaluation in monodisperse particle–liquid suspensions

The capability of the model to predict the flow field of monodisperse particle–liquid suspensions was assessed under different agitation regimes (i.e., at the just-suspension speed N_{js} and at speeds above it up to $2N_{js}$). Thus, 5-min-long liquid and solid trajectories determined by PEPT at a certain agitation speed were used as input driver data to simulate the flow field developed at higher speeds. For example, predicting flow at $1.25N_{js}$ was implemented using the liquid and solid PEPT trajectories developed at N_{js} as input driver data, which is denoted by $\boldsymbol{ML}_{\text{mono},\,N_{js}\rightarrow 1.25N_{js}}$. The azimuthally averaged profiles of local total liquid and particle velocities predicted by ML for $N = 1.25N_{js}$ and $C_m = 20$ wt. % are, respectively, compared in Figs. 9 and 10(a) with long-term PEPT measurements. Results show that the ML-predicted velocities for both liquid and solid phases are generally in very good agreement with the PEPT measurements.

The local solid concentration can be inferred from a long-term Lagrangian trajectory of a trackable solid particle in a particle–liquid flow using our previously established occupancy approach.[33,35] The local occupancy is defined as the ratio of the cumulative time Δt spent

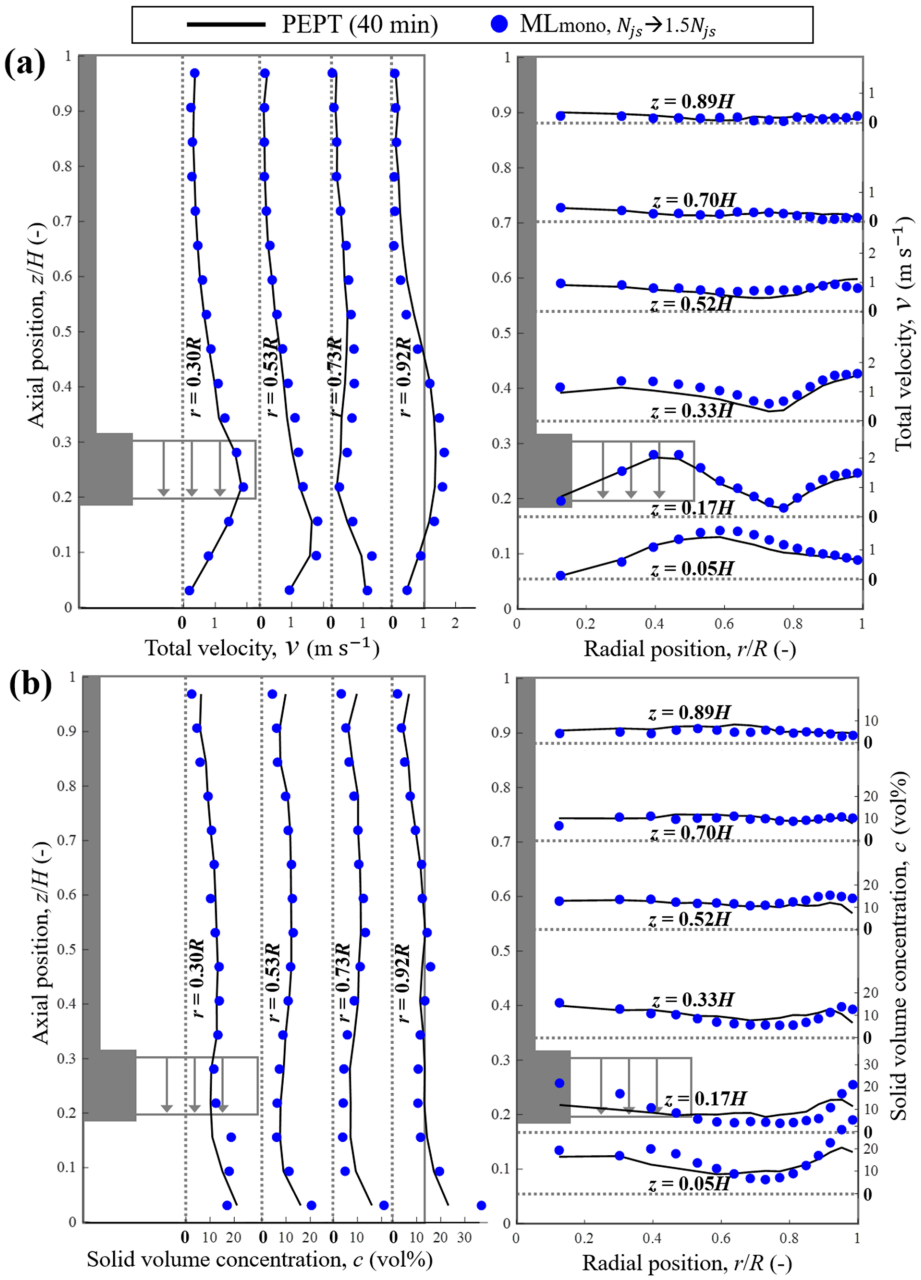

FIG. 12. Azimuthally averaged axial and radial profiles of (a) total particle velocity and (b) particle volume concentration: ML model and PEPT compared in monodisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $N = 1.5N_{js}$ (735 rpm).

by particle tracer in a grid cell to the ergodic time t_E, i.e., the experimental runtime t_∞ divided by the number N_c of equal-volume cells in the grid ($t_E = t_\infty/N_c$). Thus, the local occupancy (O_E) can be defined as follows:

$$O_E = \frac{\Delta t}{t_E}. \tag{10}$$

The local occupancy has been shown to be equal to the ratio of the local volume particle concentration (c) and mean volume particle concentration in the whole vessel (C_v), that is

$$O_E = \frac{c}{C_v}. \tag{11}$$

Using the above approach, for a given suspension flow, the ML-predicted and PEPT-measured local particle concentration fields were inferred from their respective long-term Lagrangian trajectories. The theoretical and experimental axial and radial particle volume concentration profiles corresponding to the above case of 20 wt. % solids at $1.25N_{js}$ are displayed in Fig. 10(b). Overall, the ML predictions are close to the experimental values except for some discrepancies near the bottom corner. This is most likely due to frequent particle–wall collisions resulting in more stochastic and larger velocity fluctuations in this region, which are less well approximated by the global Gaussian noise generator during the ML trajectory simulation.

Similarly, the flow field of the 20 wt. % case was predicted by the ML model at a higher speed of $1.5N_{js}$ using the liquid and solid PEPT trajectories developed at N_{js} as input driver data, and is denoted by $\boldsymbol{ML}_{\text{mono},\,N_{js}\rightarrow 1.5N_{js}}$. Comparison of the obtained local total liquid velocity, particle velocity, and particle volume concentration distribution with experiments is depicted in Figs. 11 and 12. Results show that the ML predictions are overall in good agreement with PEPT measurements, although the deviations of particle volume concentration near the bottom corner observed for the case of $\boldsymbol{ML}_{\text{mono},\,N_{js}\rightarrow 1.25N_{js}}$ are now slightly larger.

Attempts to predict the flow field at the even higher speed of $2N_{js}$ based on driver data obtained at N_{js} failed for 20 wt. % as well as other concentrations, providing a particle volume concentration with a significant overestimation in the bottom half of the vessel and a significant underestimation in the upper half. This may be attributed to the large difference in the hydrodynamic regime approaching complete suspension homogeneity at $2N_{js}$ compared to the just-suspended regime at N_{js}. In other words, the ML model is unable to predict flow in a regime that is too remote from the regime of the input driver data. However, using input driver data based on particle–liquid flow developed at N_{js}, the ML model is able to predict with good accuracy flows developed at agitation speeds up to $1.5N_{js}$, which corresponds to a large difference (>threefold) in power input.

Further tests were conducted using driver data that are more diverse, i.e., acquired at two different impeller speeds, for example, at N_{js} and $2N_{js}$. For example, comparing the results in $\boldsymbol{ML}_{\text{mono},\,N_{js}+2N_{js}\rightarrow 1.5N_{js}}$ (Fig. S2 in the supplementary material) and $\boldsymbol{ML}_{\text{mono},\,N_{js}\rightarrow 1.5N_{js}}$ (Fig. 12), shows that the former case provides improved predictions especially near the bottom of the vessel. It further demonstrates that accuracy and reliability of the ML model are improved with a larger and more diverse amount of input driver data, provided the particle–liquid flow regime is not too dissimilar to that of the input driver data. Mass continuity was also verified for all the above cases, as shown in Fig. S3.

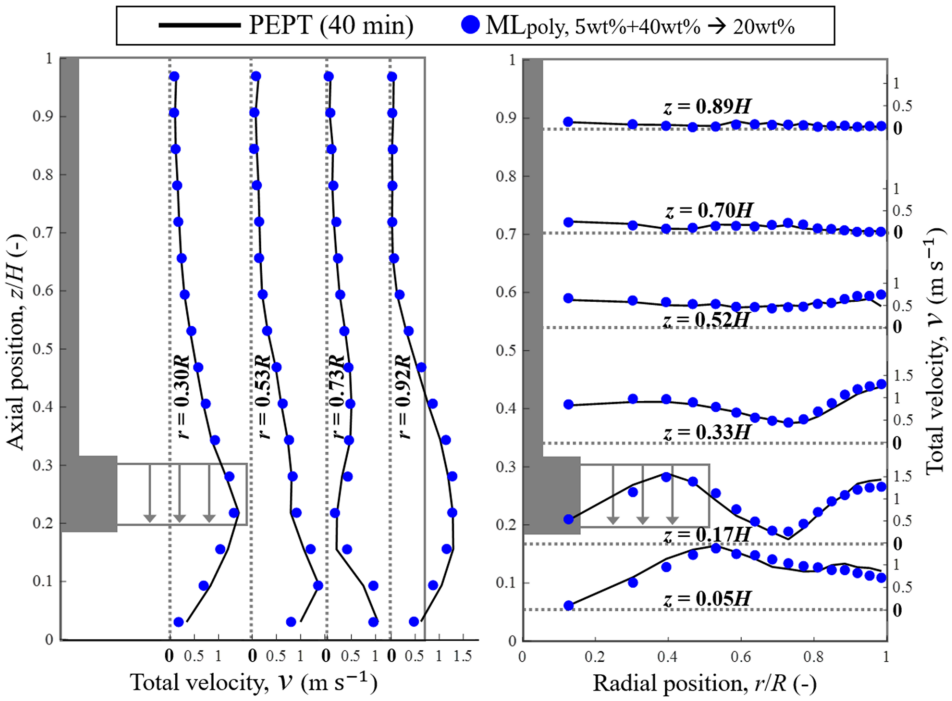

FIG. 13. Azimuthally averaged axial and radial profiles of total liquid phase velocity: ML model and PEPT compared in polydisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $N = N_{js}$ (510 rpm).

C. Evaluation of ML model evaluation in polydisperse particle–liquid suspensions

The capability of the proposed ML approach to predict the multi-component particle–liquid flow field developed at the just-suspension regime ($N = N_{js}$) was assessed under varying conditions of solid loading. Thus, 5-min PEPT-determined trajectories of all individual components corresponding to two different particle concentrations were used as input driver data to simulate the flow developed at a different particle concentration. For example, the interpolation case of predicting flow at 20 wt. % solid loading, which is denoted by $ML_{\text{poly, 5wt.\%+40wt.\%} \rightarrow \text{20wt.\%}}$, was implemented using PEPT trajectories of all components from 5 and 40 wt. % suspensions as input driver data. Comparison of the ML-predicted and PEPT-measured liquid velocity field is presented in Fig. 13, showing excellent agreement. Sample validation results of the velocity field and particle distribution are exhibited in Figs. 14 and 15, for two particle size fractions, 1.1 and 3.1 mm. Results for the other particle size fractions (1.7, 2.2, 2.7 mm) are presented in Figs. S4–S6 in the supplementary material. Again, the predictions are overall very good. Mass continuity was verified by all the flow components, and the results are summarized in Fig. S7,

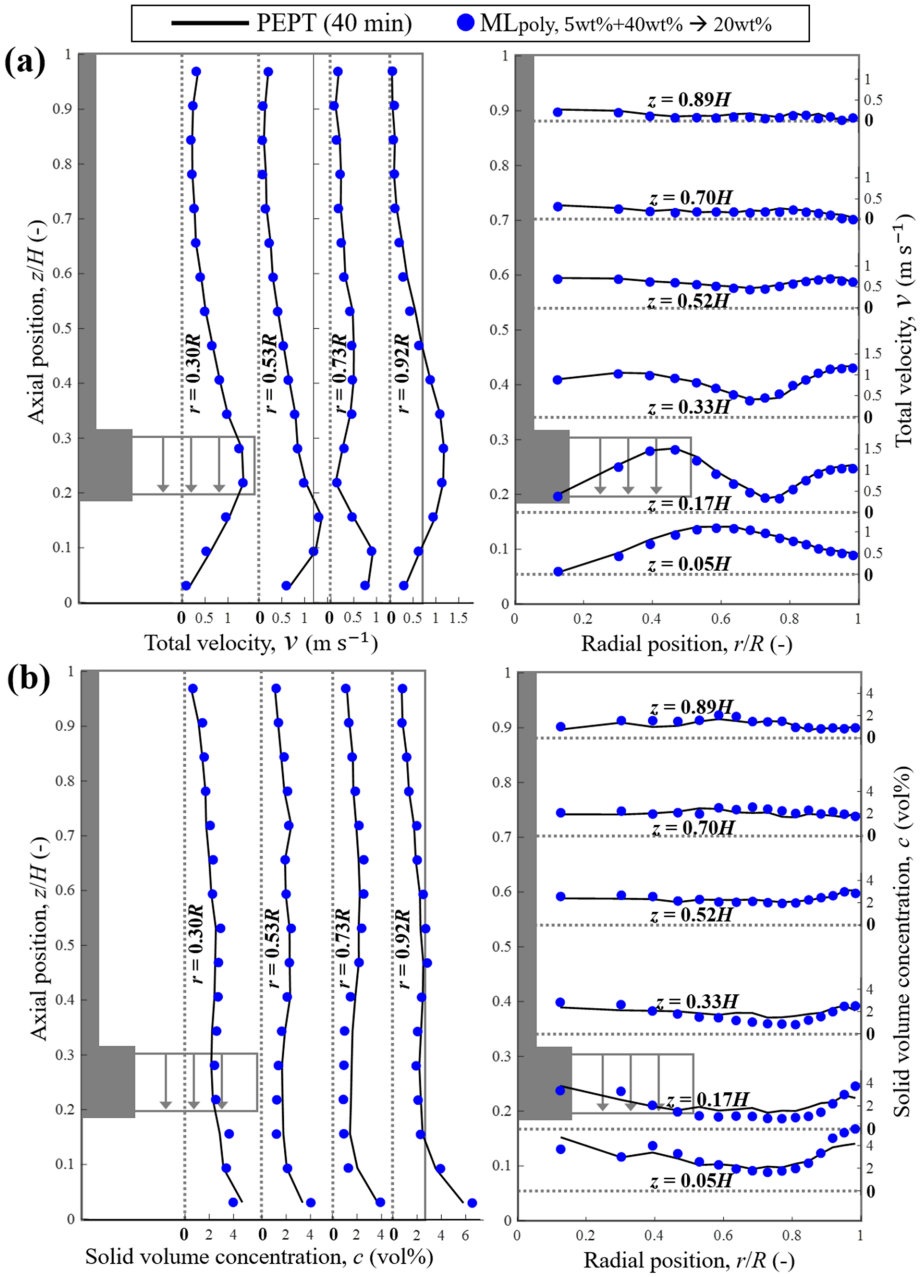

FIG. 14. Azimuthally averaged axial and radial profiles of (a) total particle velocity and (b) particle volume concentration: ML model and PEPT compared in polydisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $d = 1.1$ mm; $N = N_{js}$ (510 rpm).

showing close-to-zero average velocities $\langle v_z \rangle_{S_z}$ and $\langle v_r \rangle_{S_r}$, generally less than $0.02 v_{tip}$.

Other tests including interpolation as well as extrapolation cases were conducted using polydisperse suspensions of different solid loadings, 5, 10, and 40 wt. %, denoted by $ML_{\text{poly, 10wt.\%+20wt.\%} \rightarrow \text{5wt.\%}}$, $ML_{\text{poly, 5wt.\%+40wt.\%} \rightarrow \text{10wt.\%}}$, and $ML_{\text{poly, 10wt.\%+20wt.\%} \rightarrow \text{40wt.\%}}$. Sample results of an extrapolation case at 40 wt. % are depicted in Fig. S8 in the supplementary material. Overall, a high degree of model predictability was achieved. In conclusion, the ML framework is capable of very good predictions within and without the range of solid loading concentrations used to provide driver data for the model.

Furthermore, the ability to predict the flow field of a given particle size fraction using input driver data pertaining to other size fractions was investigated in the just-suspension regime ($N = N_{js}$). For example, the interpolation case of predicting the flow field of 2.2 mm particles at 40 wt. % suspensions, which is denoted by $ML_{\text{poly, 1.1mm+3.1mm} \rightarrow \text{2.2mm}}$, was implemented using 5 min PEPT trajectories of 1.1 and 3.1 mm particles as input driver data, as depicted

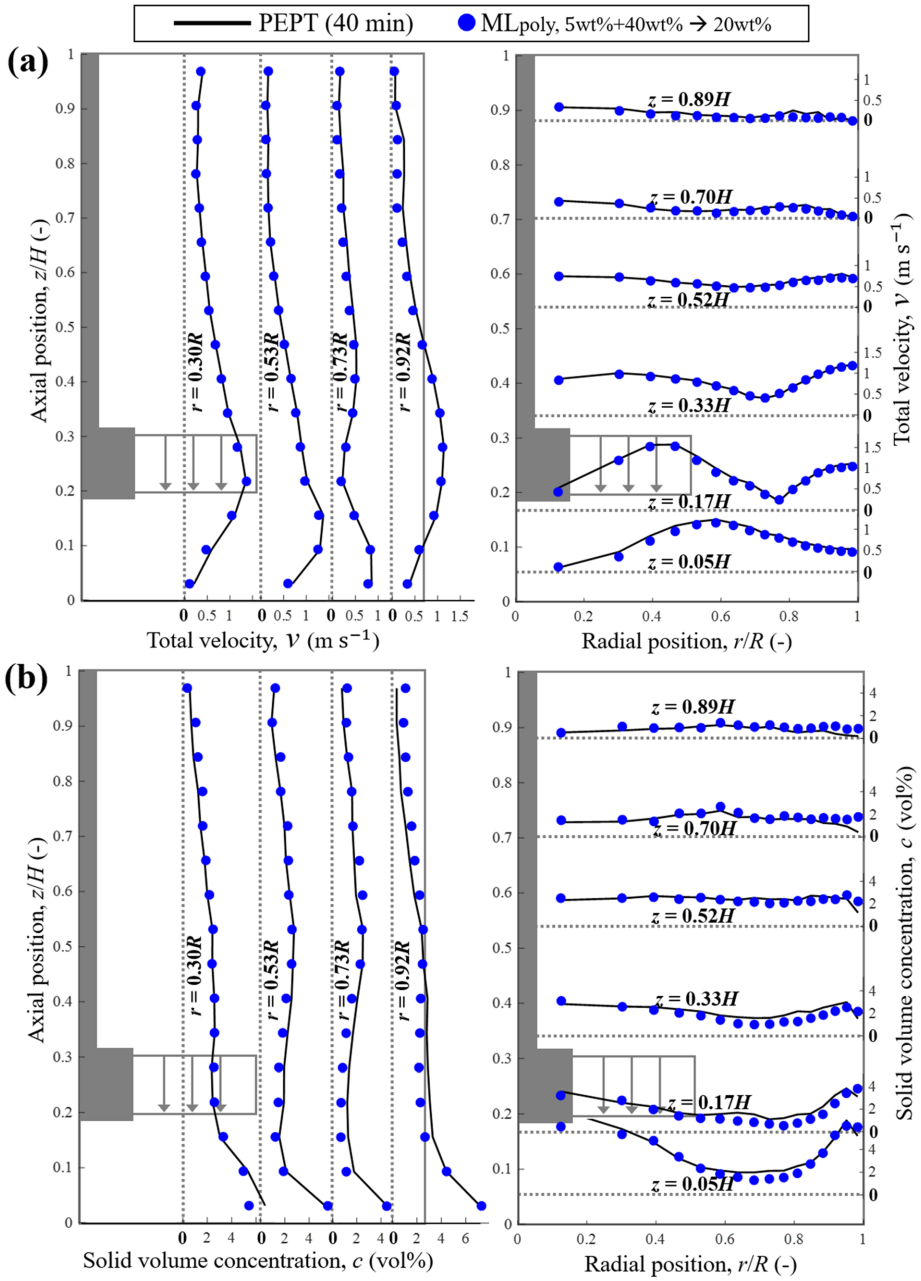

FIG. 15. Azimuthally averaged axial and radial profiles of (a) total particle velocity and (b) particle volume concentration: ML model and PEPT compared in polydisperse particle–liquid flow; $C_m = 20$ wt. % ($C_v = 10.4$ vol. %); $d = 3.1$ mm; $N = N_{js}$ (510 rpm).

in Fig. 16. Other extrapolation cases, $ML_{\text{poly, 1.7mm+2.2mm} \rightarrow \text{1.1mm}}$ and $ML_{\text{poly, 1.7mm+2.2mm} \rightarrow \text{3.1mm}}$, and sample results for 1.1 mm are presented in Fig. S9 in the supplementary material. In all cases, results show that the ML predictions are overall very good. The mass continuity was also verified for all investigated cases, as shown in Fig. S10.

V. CONCLUSIONS

We have successfully developed and validated a computationally efficient machine learning framework for predicting single-phase and complex two-phase multicomponent turbulent flows in stirred vessels. Using a very small amount of Lagrangian driver trajectories, the proposed ML model learns the primary flow dynamics from the driver data and efficiently produces new long-term trajectories of the flow components corresponding to new flow conditions. Evaluation was conducted by comparing the long-term ML-predicted trajectories and long-term PEPT-measured trajectories using the local flow characteristics inferred, namely, local flow component velocities, and concentration distribution. The overall excellent agreement obtained confirmed the accuracy and reliability of the presented ML model for predicting, such complex flows under a wide range of conditions. Thus, accurate

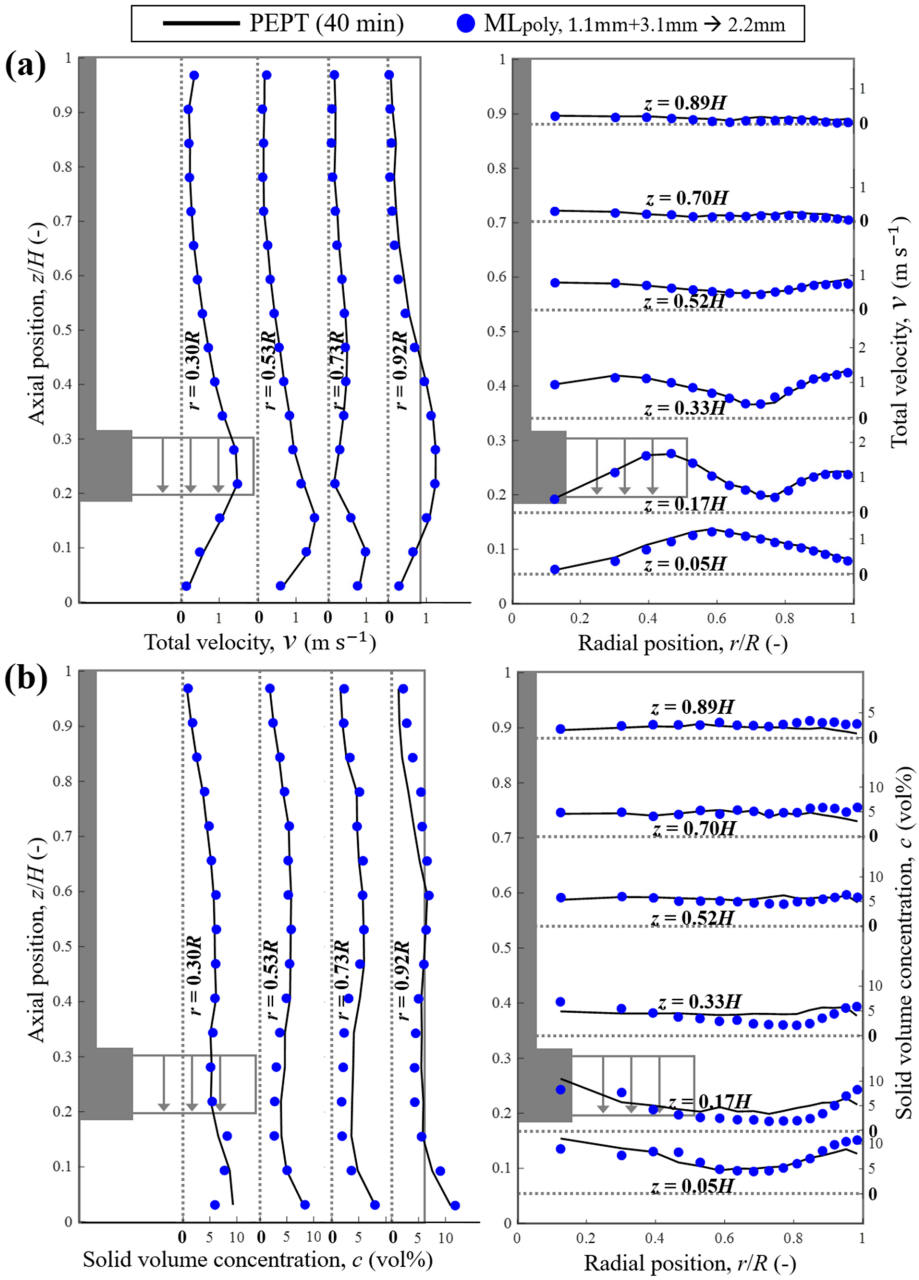

FIG. 16. Azimuthally averaged axial and radial profiles of (a) total particle velocity and (b) particle volume concentration: ML model and PEPT compared in polydisperse particle–liquid flow; $C_m = 40$ wt. % ($C_v = 23.6$ vol. %); $d = 2.2$ mm; $N = N_{js}$ (610 rpm).

ML predictions can be achieved within and without the respective range of the supplied driver data for (i) different impeller speeds, provided the hydrodynamic flow regime is not dissimilar to the one corresponding to the driver data; (ii) for different solid loading concentrations; and (iii) for different particle size fractions.

The proposed ML framework provides a new flow analysis and modeling strategy, whereby only short-term experiments (or alternatively high-fidelity simulations) covering a few typical flow situations are sufficient to enable the prediction of complex multiphase flows, significantly reducing experimental and/or simulation costs. The ML technique is applicable to other impeller configurations provided that driver data are available for the particular impeller configuration considered. However, using driver data from one impeller configuration to predict the performance of another impeller configuration is generally not possible because of the large differences in flow patterns. For example, our investigations showed that data pertaining to a down-pumping PBT could not be used to drive the ML model to predict the performance of an up-pumping PBT.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional validation results of the proposed machine-learning framework.

ACKNOWLEDGMENTS

This work was supported by EPSRC Programme Grant No. EP/R045046/1: Probing Multiscale Complex Multiphase Flows with Positrons for Engineering and Biomedical Applications (PI: Professor M. Barigou, University of Birmingham).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Kun Li: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Chiya Savari:** Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Writing – review & editing (equal). **Mostafa Barigou:** Conceptualization (lead); Funding acquisition (lead); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

NOMENCLATURE

Symbols

c	Local volume solid concentration (%)
C_m	Mean solid mass concentration (%)
C_v	Mean solid volume concentration (%)
D	Impeller diameter (m)
d	Solid particle diameter (mm)
dt	Time step of machine learning trajectory (s)
e	Restitution coefficients (-)
H	Height of the fluid in vessel (m)
k	Number of neighbors used in KNN regressor (-)
L_p	Distance between two samples in feature space (-)
N	Impeller rotational speed (rpm)
N_c	Number of equal-volume cells of cylindrical mesh system (-)
N_{js}	Minimal impeller rotational speeds for just suspension (rpm)
N_L	Number of detected locations in a cell (-)
N_r, N_z, N_θ	Number of cells in three directions of cylindrical mesh system (-)
$\boldsymbol{N}(\mu, \sigma^2)$	Gaussian distribution with mean value μ and standard derivation σ (-)
O_E	Local occupancy (-)
R	Vessel radius (m)
r	Radial cylindrical coordinate (m)
Re_{imp}	Impeller Reynolds number (-)
T	Vessel diameter (m)
t	Time (s)
t_E	Ergodic time (s)
t_∞	Total runtime of Lagrangian trajectory (s)
$\boldsymbol{V}$	Targets of training sample (-)
$\boldsymbol{v}$	Lagrangian velocity (m/s)
$\boldsymbol{v}'$	Fluctuating velocity (m/s)
$\overline{\boldsymbol{v}}$	Mean velocity (m/s)
v_{tip}	Impeller tip velocity (m/s)
$\langle \boldsymbol{v}_r \rangle_{S_r}$	Averaged radial velocity over a cylindrical envelope (m/s)
$\langle \boldsymbol{v}_z \rangle_{S_z}$	Averaged axial velocity over a horizontal plane (m/s)
v_θ, v_r, v_z	Lagrangian velocity components (m/s)
v'_θ, v'_r, v'_z	Fluctuating velocity components (m/s)
$\boldsymbol{v}_1, \boldsymbol{v}_2$	Velocities before and after collision (m/s)
$\boldsymbol{V}^{pred}$	New predicted targets (-)
w	Weight function of KNN regressor (-)
$\boldsymbol{x}$	3D space location of tracer (m)
x, y, z	Cartesian coordinates (m)
$\boldsymbol{X}_i, \boldsymbol{X}_j$	ith query instance and jth training instance (-)
$X_i^{(l)}, X_j^{(l)}$	lth dimensional feature of ith instance and jth instance (-)
Δt	Cumulative time of tracer spent in each cell (s)

Greek symbols

θ	Azimuthal coordinates of cylindrical system (rad)
μ	Mean normalized fluctuating velocity (-)
ρ	Phase density (kg/m^3)
σ	Standard deviation of normalized fluctuating velocity (-)

Abbreviations

CFD	Computational fluid dynamics
KNN	k-nearest neighbors
LDV	Laser Doppler velocimetry

ML	Machine learning
PBT	Six-blade 45° pitched-turbine
PBTD	Down-pumping PBT
PEPT	Positron emission particle tracking
PIV	Particle image velocimetry
RANS	Reynolds-averaged Navier–Stokes
3D	Three-dimensional

REFERENCES

[1]H. Ghoddusi, G. G. Creamer, and N. Rafizadeh, "Machine learning in energy economics and finance: A review," Energy Econ. **81**, 709 (2019).
[2]X. D. Chen and V. Yang, "Direct numerical simulation of multiscale flow physics of binary droplet collision," . Fluids **32**, 062103 (2020).
[3]P. Rajendra and V. Brahmajirao, "Modeling of dynamical systems through deep learning," Biophys. Rev. **12**, 1311 (2020).
[4]S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," Annu. Rev. Fluid Mech. **52**, 477 (2020).
[5]R. M. Elavarasan and R. Pugazhendhi, "Restructured society and environment: A review on potential technological strategies to control the COVID-19 pandemic," Sci. Total Environ. **725**, 138858 (2020).
[6]Q. V. Pham, D. C. Nguyen, H. T. Thien, W. J. Hwang, and P. N. Pathirana, "Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: A survey on the state-of-the-arts," IEEE Access **8**, 130820 (2020).
[7]N. Harnby, M. F. Edwards, and A. W. Nienow, *Mixing in the Process Industries* (Butterworth-Heinemann, 1997).
[8]E. L. Paul, V. A. Atiemo-Obeng, and S. M. Kresta, *Handbook of Industrial Mixing* (John Wiley & Sons, Inc., 2003).
[9]M. Jaszczur and A. Mlynarczykowska, "A general review of the current development of mechanically agitated vessels," Processes **8**, 982 (2020).
[10]L. B. Chen, Y. Hontoir, D. Huang, J. Zhang, and A. J. Morris, "Combining first principles with black-box techniques for reaction systems," Control Eng. Pract. **12**, 819 (2004).
[11]K. Hasegawa, K. Fukami, T. Murata, and K. Fukagata, "CNN-LSTM based reduced order modeling of two-dimensional unsteady flows around a circular cylinder at different Reynolds numbers," Fluid Dyn. Res. **52**, 065501 (2020).
[12]J. Ling, A. Kurzawski, and J. Templeton, "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance," J. Fluid Mech. **807**, 155 (2016).
[13]W. Hou, D. Darakananda, and J. D. Eldredge, "Machine-learning-based detection of aerodynamic disturbances using surface pressure measurements," AIAA J. **57**, 5079 (2019).
[14]H. F. Zhai, Q. Zhou, and G. H. Hu, "Predicting micro-bubble dynamics with semi-physics-informed deep learning," AIP Adv. **12**, 035153 (2022).
[15]B. R. Noack, *Closed-Loop Turbulence Control-From Human to Machine Learning (and Retour)* (Springer, Singapore, 2017).
[16]B. Colvert, G. Liu, H. B. Dong, and E. Kanso, "Flowtaxis in the wakes of oscillating airfoils," Theor. Comput. Fluid Dyn. **34**, 545 (2020).
[17]J. Rabault, M. Kuchta, A. Jensen, U. Reglade, and N. Cerardi, "Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control," J. Fluid Mech. **865**, 281 (2019).
[18]C. Savari, H. A. Sheikh, and M. Barigou, "Lagrangian recurrence tracking: A novel approach for description of mixing in liquid and particle-liquid flows," Ind. Eng. Chem. Res. **60**, 18501 (2021).
[19]K. Li, C. Savari, and M. Barigou, "Computation of Lagrangian coherent structures from experimental fluid trajectory measurements in a mechanically agitated vessel," Chem. Eng. Sci. **254**, 117598 (2022).
[20]H. A. Sheikh, C. Savari, and M. Barigou, "Lagrangian stochastic modelling of liquid flow in a mechanically agitated vessel," Chem. Eng. Sci. **249**, 117318 (2022).
[21]H. A. Sheikh, C. Savari, and M. Barigou, "A data-driven stochastic model for velocity field and phase distribution in stirred particle-liquid suspensions," Powder Technol. **411**, 117940 (2022).
[22]D. Chang, C. R. Edwards, F. M. Zhang, and J. Sun, "A data assimilation framework for data-driven flow models enabled by motion tomography," Int. J. Intell. Rob. Appl. **3**, 158 (2019).
[23]G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics* (Cambridge University Press, 2017).
[24]E. Ferrero, S. Alessandrini, S. Meech, and C. Rozoff, "A 3D Lagrangian stochastic particle model for the concentration variance dispersion," Bull. Atmos. Sci Technol. **3**, 2 (2022).
[25]M. Strumik and W. M. Macek, "Testing for Markovian character and modeling of intermittency in solar wind turbulence," Phys. Rev. E **78**, 026414 (2008).
[26]C. Dopazo, L. Cifuentes, J. Martin, and C. Jimenez, "Strain rates normal to approaching iso-scalar surfaces in a turbulent premixed flame," Combust. Flame **162**, 1729 (2015).
[27]K. Li, C. Savari, H. A. Sheikh, and M. Barigou, "A data-driven machine learning framework for modelling of turbulent mixing flows," Phys. Fluids **35**, 015150 (2023).
[28]T. N. Zwietering, "Suspending of solid particles in liquid by agitators," Chem. Eng. Sci. **8**, 244 (1958).
[29]M. Barigou, "Particle tracking in opaque mixing systems: An overview of the capabilities of PET and PEPT," Chem. Eng. Res. Des. **82**, 1258 (2004).
[30]P. Pianko-Oprych, A. W. Nienow, and M. Barigou, "Positron emission particle tracking (PEPT) compared to particle image velocimetry (PIV) for studying the flow generated by a pitched-blade turbine in single phase and multi-phase systems," Chem. Eng. Sci. **64**, 4955 (2009).
[31]C. Tropea, "Laser-Doppler anemometry—Recent developments and future challenges," Meas. Sci. Technol. **6**, 605 (1995).
[32]M. Eesa and M. Barigou, "Horizontal laminar flow of coarse nearly-neutrally buoyant particles in non-Newtonian conveying fluids: CFD and PEPT experiments compared," Int. J. Multiphase Flow **34**, 997 (2008).
[33]A. Guida, X. Fan, D. J. Parker, A. W. Nienow, and M. Barigou, "Positron emission particle tracking in a mechanically agitated solid-liquid suspension of coarse particles," Chem. Eng. Res. Des. **87**, 421 (2009).
[34]A. Guida, A. W. Nienow, and M. Barigou, "PEPT measurements of solid-liquid flow field and spatial phase distribution in concentrated monodisperse stirred suspensions," Chem. Eng. Sci. **65**, 1905 (2010).
[35]A. Guida, A. W. Nienow, and M. Barigou, "Lagrangian tools for the analysis of mixing in single-phase and multiphase flow systems," AIChE J. **58**, 31–45 (2012).
[36]C. Savari, K. Li, and M. Barigou, "Multiscale wavelet analysis of 3D Lagrangian trajectories in a mechanically agitated vessel," Chem. Eng. Sci. **260**, 117844 (2022).
[37]Z. J. Yang, C. Savari, and M. Barigou, "Numerical and experimental investigations of horizontal turbulent particle-liquid pipe flow," Ind. Eng. Chem. Res. **61**, 12040 (2022).
[38]G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, 1953).
[39]P. A. Krogstad and P. A. Davidson, "Freely decaying, homogeneous turbulence generated by multi-scale grids," J. Fluid Mech. **680**, 417 (2011).
[40]M. E. Morsy and J. Yang, "Numerical and experimental study on turbulence statistics in a large fan-stirred combustion vessel," Exp. Fluids **62**, 116 (2021).
[41]T. Tanaka and J. K. Eaton, "Sub-Kolmogorov resolution particle image velocimetry measurements of particle-laden forced turbulence," J. Fluid Mech. **643**, 177 (2010).
[42]O. Semeraro, G. Bellani, and F. Lundell, "Analysis of time-resolved PIV measurements of a confined turbulent jet using POD and Koopman modes," Exp. Fluids **53**, 1203 (2012).
[43]C. Savari and M. Barigou, "Lagrangian wavelet analysis of turbulence modulation in particle–liquid mixing flows," Phys. Fluids **34**, 115121 (2022).
[44]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," J. Mach. Learn. Res. **12**, 2825 (2011).
[45]T.-S. Wong and P. Y. Yeh, "Reliable accuracy estimates from *k*-fold cross validation," IEEE Trans. Knowl. Data Eng. **32**, 1586–1594 (2020).