Development of machine learning framework for interface force closures based on bubble tracking data

Tai, C.-K.; Bolotnov, I.; Evdokimov, I.; Schlegel, F.; Lucas, D.;

Originally published:

November 2022

Nuclear Engineering and Design 399(2022), 112032

DOI: https://doi.org/10.1016/j.nucengdes.2022.112032

Perma-Link to Publication Repository of HZDR:

https://www.hzdr.de/publications/Publ-32961

Nuclear Engineering and Design

DEVELOPMENT OF MACHINE LEARNING FRAMEWORK FOR INTERFACE FORCE CLOSURES BASED ON BUBBLE TRACKING DATA

--Manuscript Draft--

Manuscript Number:	
Article Type:	VSI: NURETH19
Section/Category:	Thermal Hydraulics
Corresponding Author:	Cheng-Kai Tai NC State University UNITED STATES
First Author:	Cheng-Kai Tai
Order of Authors:	Cheng-Kai Tai
	Ilya Evdokimov, Ph.D.
	Fabian Schlegel, Ph.D.
	Igor A. Bolotnov, Ph.D.
	Dirk Lucas, Ph.D.
Abstract:	Interfacial force closures in the two-fluid model play a critical role for the predictive capabilities of void fraction distribution. However, the practices of interfacial force modeling have long been challenged by the inherent physical complexity of the two-phase flows. The rapidly expanding computational capabilities in the recent years have made high-fidelity data from the interface-captured direct numerical simulation become more available, and hence potential for data-driven interfacial force modeling has prevailed. In this work, we established a data-driven modeling framework integrated to the HZDR multiphase Eulerian-Eulerian framework for computational fluid dynamics simulations. The data-driven framework is verified in a benchmark problem, where a feedforward neural network managed to capture the non-linear mapping between bubble Reynolds number and drag coefficient and reproduce the void distribution resulting from the baseline model in the test case. The second focus is on utilizing the bubble tracking data set to form a closure for the bubble drag in the turbulent bubbly flow, in which the drag coefficient is set to be correlated with the bubble Reynolds number and the Eötvös number. Pseudo-steady state filtering in the Frenet Frame was carried out to obtain the drag coefficient from the turbulent bubbly flow data. The performance of the data-driven drag model is also examined through a case study, where improvement of model's prediction near-wall is regarded necessary. Discussion and further plans of investigation are provided.
Suggested Reviewers:	Gretar Tryggvason, Ph.D. Department Head and Charles A. Miller, Jr. Distinguished Professor, Johns Hopkins University gtryggv1@jhu.edu
	Xiaodong Sun, Ph.D. Professor, University of Michigan xdsun@umich.edu
	Mamoru Ishii, Ph.D. Associate Professor, Purdue University ishii@purdue.edu

DEVELOPMENT OF MACHINE LEARNING FRAMEWORK FOR INTERFACE FORCE CLOSURES BASED ON BUBBLE TRACKING DATA

Highlights

- A data-driven modeling (DDM) framework is established for applications in the HZDR multiphase fluid simulation framework based on the Eulerian-Eulerian approach. This work further focuses on the application of the DDM framework in data-driven drag model for bubbly flows.
- The implementation of DDM framework is verified via an artificial problem, in which the simulation results based on feedforward neural network drag model emulating correlation by Tomiyama et al. (1992) showed good agreement with the reference.
- Pseudo-steady state filtering in the Frenet moving frame is applied to the bubble tracking data set (Fang et al., 2017) in order to obtain drag coefficients of bubbles from the data of interface-captured direct numerical simulation of turbulent bubbly flow.
- The interface-captured DNS data-driven drag model was applied to the case based on the setup by Shawkat et al., (2008). The performance of the data-driven drag model is discussed.

DEVELOPMENT OF MACHINE LEARNING FRAMEWORK FOR INTERFACE FORCE CLOSURES BASED ON BUBBLE TRACKING DATA

Cheng-Kai Tai[1,*], Ilya Evdokimov[2], Fabian Schlegel[2], Igor A. Bolotnov[1] and Dirk Lucas[2]

[1]Department of Nuclear Engineering, North Carolina State University
[2]Department of FWDC, Helmholtz-Zentrum Dresden Rossendorf
[*]Corresponding author: ctai2@ncsu.edu

ABSTRACT

Interfacial force closures in the two-fluid model play a critical role for the predictive capabilities of void fraction distribution. However, the practices of interfacial force modeling have long been challenged by the inherent physical complexity of the two-phase flows. The rapidly expanding computational capabilities in the recent years have made high-fidelity data from the interface-captured direct numerical simulation become more available, and hence potential for data-driven interfacial force modeling has prevailed.

In this work, we established a data-driven modeling framework integrated to the HZDR multiphase Eulerian-Eulerian framework for computational fluid dynamics simulations. The data-driven framework is verified in a benchmark problem, where a feedforward neural network managed to capture the non-linear mapping between bubble Reynolds number and drag coefficient and reproduce the void distribution resulting from the baseline model in the test case.

The second focus is on utilizing the bubble tracking data set to form a closure for the bubble drag in the turbulent bubbly flow, in which the drag coefficient is set to be correlated with the bubble Reynolds number and the Eötvös number. Pseudo-steady state filtering in the Frenet Frame was carried out to obtain the drag coefficient from the turbulent bubbly flow data. The performance of the data-driven drag model is also examined through a case study, where improvement of model's prediction near-wall is regarded necessary. Discussion and further plans of investigation are provided.

Keyword: direct numerical simulation, data-driven modeling, interfacial force modeling, machine learning, Eulerian-Eulerian approach

DEVELOPMENT OF MACHINE LEARNING FRAMEWORK FOR INTERFACE FORCE CLOSURES BASED ON BUBBLE TRACKING DATA

Cheng-Kai Tai[1,*], Ilya Evdokimov[2], Fabian Schlegel[2], Igor A. Bolotnov[1] and Dirk Lucas[2]

[1]Department of Nuclear Engineering, North Carolina State University
[2]Department of FWDC, Helmholtz-Zentrum Dresden Rossendorf
[*]Corresponding author: ctai2@ncsu.edu

ABSTRACT

Interfacial force closures in the two-fluid model play a critical role for the predictive capabilities of void fraction distribution. However, the practices of interfacial force modeling have long been challenged by the inherent physical complexity of the two-phase flows. The rapidly expanding computational capabilities in the recent years have made high-fidelity data from the interface-captured direct numerical simulation become more available, and hence potential for data-driven interfacial force modeling has prevailed.

In this work, we established a data-driven modeling framework integrated to the HZDR multiphase Eulerian-Eulerian framework for computational fluid dynamics simulations. The data-driven framework is verified in a benchmark problem, where a feedforward neural network managed to capture the non-linear mapping between bubble Reynolds number and drag coefficient and reproduce the void distribution resulting from the baseline model in the test case.

The second focus is on utilizing the bubble tracking data set to form a closure for the bubble drag in the turbulent bubbly flow, in which the drag coefficient is set to be correlated with the bubble Reynolds number and the Eötvös number. Pseudo-steady state filtering in the Frenet Frame was carried out to obtain the drag coefficient from the turbulent bubbly flow data. The performance of the data-driven drag model is also examined through a case study, where improvement of model's prediction near-wall is regarded necessary. Discussion and further plans of investigation are provided.

Keyword: direct numerical simulation, data-driven modeling, interfacial force modeling, machine learning, Eulerian-Eulerian approach

1. Introduction

Two-phase flow phenomenon has a wide range of applications in engineering practices. In the context of water-cooled reactors, understanding of gas-liquid bubbly flow serves as the foundation to the operational safety of the reactor systems. However, the inherent complexity of the multi-phase flow phenomenon rendered modeling of bubbly flow a challenging task and hence a still popular topic among two-phase flow researchers.

One of the two-phase flow modeling frameworks commonly adopted in engineering practices is the Eulerian-Eulerian (E-E) approach [1, 2, 3], as shown in eq. (1) to (3):

$$\frac{\partial \alpha_k \rho_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k u_k) = \Gamma_k, k = l, g \quad (1)$$

$$\frac{\partial \alpha_k \rho_k u_k}{\partial t} + \nabla \cdot \alpha_k \rho_k u_k u_k = -\nabla \alpha_k p_k + \nabla \cdot (\alpha_k \tau_k) + \alpha_k \rho_k g + M_{ik} \quad (2)$$

$$M_{ik} = F_D + F_L + F_{VM} + F_{TD} + F_W \quad (3)$$

where $k = l, g$ stands for the liquid and gas phase, respectively. Γ_k is the mass source for phase k. Under the scope of this work, $\Gamma_k = 0$ since phase change is excluded. M_{ik} accounts for the momentum source for the phase k during interfacial momentum exchange, which is modeled into the separate effects by drag (F_D), lift (F_L), virtual mass (F_{VM}), turbulent dispersion (F_{TB}), and wall force (F_W). Compared to the direct numerical simulation (DNS), in which the smallest time and spatial scales of turbulence and bubble dynamics are resolved, the E-E approach accounts for the ensemble of the instantaneous conservation equations for mass and momentum. And the interface between liquid and gas is not explicitly captured. This brings tremendous advantage in computational cost in performing E-E approach-based computational fluid dynamics (CFD) simulations over DNS, and hence CFD simulation based on two-fluid model remains the workhorse for engineering applications. On the other hand, however, the predictive capabilities on void distribution and phasic velocity hinges heavily upon the robustness of individual interfacial force closure as well as the interaction between closures for different interfacial force terms. However, due to the inherent complexity of interfacial momentum exchange, a set of interfacial force closures generalizable to different engineering applications remains a goal pursued by researchers nowadays [4].

In the recent decade, the ever-growing computational capability of high-performance computing (HPC) systems has made DNS an affordable source of high-fidelity and high-resolution data. Specifically for the two-phase flows, the interface captured DNS can provide further insight to the complicated two-phase flow phenomena. Fang et al. [5] performed interface captured DNS with PHASTA to study the bubbly flow in pressurized water reactor (PWR)-relevant subchannel geometry. In the simulation, 17 bubbles are explicitly tracked with the interfaces between gas and liquid phases captured with the level-set method [6]. The bubble tracking capabilities allows the exportation of flow data in the vicinity of the interfaces and hence suitable for the study of interfacial momentum export. Cambareri et al. [7] studied bubble dynamics of adiabatic bubbly flow in a PWR subchannel with the spacer grid and mixing vanes with interface captured DNS. Five cases with void fraction $1\%, 3\%, 5\%, 10\%$ and 15% were studied. The presence of the spacer grid and mixing vanes resulted in promoted liquid cross flow and mixing between subchannels. On the other hand, the swirling liquid caused the bubbles swarm towards the centerline of the subchannel and drastically altered liquid statistics there. Note that local liquid velocity data and flow statistics are also available thanks to the bubble tracking algorithm. Saini and Bolotnov [8] investigated the interaction of droplets with PWR subchannel structures in the dispersed flow film regime with interface captured DNS. This research topic is of importance to understand the droplet dynamics in post-loss of coolant accident (LOCA) scenarios. The explicitly captured interfaces allow detailed understanding of droplet morphology along the flow, which paves the path toward improvement of heat transfer correlations for system thermal hydraulics analysis. Fan et al. [9] developed proportional-

integral-derivative (PID) bubble controller for the interface captured DNS with PHASTA. The PID controller anchors the bubble in a relatively small computational domain by compensating the interfacial forces experienced by the bubble. The main advantage of the PID controller is allowing lower computational cost for investigation of bubble deformation and interfacial force terms. With the PID controller, a bubble deformation map is established, and the bubble topology was validated by experimental works.

On the other hand, the growth in data supply and computational capabilities a has made data-driven approach an attractive measure to tackle challenges in closure development for thermal fluid simulations. Sandberg et al. [10] employed large eddy simulation (LES) data and gene expression programming to develop the closures for non-linear shear stress and turbulent heat flux for trailing-edge cooling slot cases. Zhu et al. [11] developed an adaptive turbulence modeling approach for Reynolds-averaged Navier Stokes equations (RANS) framework based on ensemble learning. The proposed multi-model data driven turbulence modeling (MDTM) approach modulates the balance between knowledge base and high-fidelity data base by introducing multiple baseline models for low-fidelity simulation. The flexibility in baseline model allows the high-to-low (Hi2Lo) turbulence modeling scheme to choose the best one and hence improves the generalization capability of data-driven model. Bao et al. [12] established feature similarity measurement (FSM) for the coarse-grid CFD (CG-CFD) simulation based on the data-driven method. The proposed framework allows to bridge the gap in global simulation parameters of data sets by evaluating distribution of local (cell or grid point level) flow features. The FSM combined with deep learning technique was deployed to study the connection between error in CG-CFD simulation to the local physical coverage conditions. It was shown that the FSM has good predictive capability when the physical coverage condition is in the interpolation range. On the ML applications in two-phase flows, the high-fidelity data by first principle-based simulations, such as DNS, contains valuable information to support/inform the development of closures for sub-grid phenomena to the simulations in the engineering practices. Ma et al. [13, 14] applied neural network and interface tracked DNS data of to form closure for in the two-fluid model. The considered problem is the evolving bubbly flow of (nearly) spherical bubbles of the same size in a rectangular box. The data-driven model showed decent performance in predicting evolution of void distribution in the cases generating training data and those with different initial condition and void fractions. Albeit the simplicity of the modeled problem, this work has shown the potential of DNS data informing turbulence models in the lower-order simulations. Liu et al. [15] leveraged the data from large eddy simulation (LES) of pool boiling with interface explicitly tracked and employed deep feedforward neural networks (FNN) to inform the pool boiling heat transfer behaviors for the two fluid model-based simulation. The quantities of interest (QoI) predicted by the deep FNN showed good agreement between the reference data from LES in both extrapolation and interpolation test cases. The authors also pointed out that the generalization capabilities of the FNN model, in terms of scale invariances and flow regime, should be further examined.

From the review above, it is obvious that the ML methods' potential has prevailed in different aspects of thermal fluid simulation. the advantage of machine learning (ML) methods in tackling two-phase flow problems are: (1) efficiency handling large data sets, (2) flexibility in model form and complexity, and (3) capability representing/approximating complicated mathematical relations and physical processes. These traits make ML methods an attractive

measure to approach two-phase flow closures.

This work aims to establish a data-driven modeling (DDM) framework upon the HZDR E-E framework and baseline models [16] that can extract useful information from the high-fidelity data using ML algorithms. Then, given the significance of drag force in predictive capability of velocity profile in the problem of interest, this work also presents the development of data-driven model for drag force in bubbly flow using the DNS data [5]. The resulting data-driven drag model is employed in place of baseline drag to demonstrate the DDM framework. However, the DDM practice with the bubble tracking data set is challenged by the following: (1) the turbulent bubbly flow is inherently transient, (2) bubble interface is constantly being created/eliminated due to break up/coalescence of bubbles in the source DNS case. (3) momentum exchange between gas and liquid phases is further complicated by the topology evolution in turbulent liquid flow. The quality of the DDM framework will be compromised if without proper feature identification and data pre-processing. Therefore, a benchmark problem (case study I) utilizing artificial data set to model bubble drag coefficient is firstly studied to verify the proposed DDM framework. Also, sensitivity study on model size, amount of data learned are carried out to help the development of data utilization strategies. The artificial data is generated based on known correlation or analytical solution of known relation. Such characteristics make artificial data set an ideal candidate for model and framework verification. Then, based on the experience on case study I, we show the preliminary results on using DNS bubble tracking data set [5] for data-driven drag model for bubbly flow. Details of problem formulation and results will be discussed in the following sections.

The rest of the paper is organized as following: section 2 will introduce the methodologies adopted, including the DDM framework, the interface-captured DNS data set, and the ML algorithm adopted. In section 3, we discuss the results of ML-based drag closure for E-E formulation with two case studies. The paper is concluded in section 4.

2. Methodology

2.1. The data-driven modeling framework

Figure 1 shows the overall schematic of the DDM framework established for the OpenFOAM-based Eulerian-Eulerian multiphase solver developed at HZDR. Data-driven and Machine Learning (ML) applications involve close interaction of raw data sources, top-level algorithms, and software. Among other challenges, the reproducibility of obtained results and models demands special approach when ML models are developed and deployed. The DDM framework proposed here aims to achieve reproducibility and extensibility by modularizing the data adaptation, pre-processing, integration to solver and post-processing of the results.

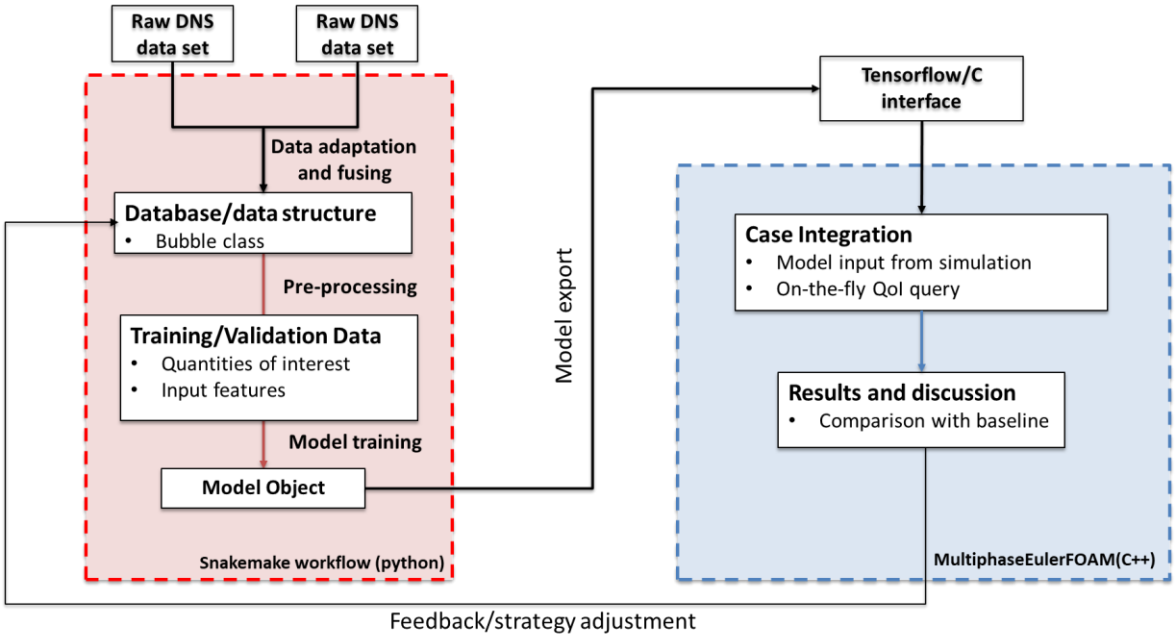


Figure 1 Schematic of data-driven modeling framework for HZDR Euleraian-Eulerian multiphase solver. On the left hand side of the schematic (red box), the actions between each intermediate stage are automated using Snakemake [17].

The proposed framework pipelines the data processing, model training and implementation to final deployment into HZDR maintained OpenFOAM multiphase solver. The main functional component of the framework is the Snakemake [17] workflow. The workflow is composed of various "jobs" representing logical steps in reading, pre-processing datasets, and training models (actions listed besides the arrows in the red box of Figure 1). Snakemake oversees correct execution and checks for proper file-structure in the input and output of each job as well. Under the scope of this work, the workflow is driven by DNS bubble tracking data set [5]. The bubble data structure is established by adapting and merging raw data sets. Then, to meet the modeling and application need, the quantities of interest and input features are extracted, and form the training/validation sets that are consumed in the model training stage. Finally, a model object along with the trained parameters are encapsulated into and binary file for later usage in prediction phase (right hand side flow in Figure 1). The only possible successful job result in this case would be the presence of all three files (2 input and 1 output) in the respective directory. It makes top-level ML pipeline well defined and deterministic. Note that the DDM framework presented here can also combine various ML algorithms and models. For example, in additions to the ML interfacial force models implemented in this work, it currently also contains PyTorch model used for prediction of phenomena unrelated to the main topic of present paper.

On the lower level, data pre-processing and model training are performed in Python. As for main ML framework, Tensorflow [18] and Keras [19] ML libraries are used in model setup, training and performance evaluation. The integration of model object into OpenFOAM solver environment (C++-based) is fulfilled using TensorFlow C API [20] (will be denoted as interface onward). In the case setup phase, the ML object acts as a plug-in that is passed into the interfacial force slot in the HZDR E-E multiphase solver. During the simulation, the model object is evoked by the interface in OpenFOAM solver [16] per step for on-the-fly query of

quantity of interest. For the drag force case in this work, the ML model receives cell input values (e.g. local Re and Eo) from the solver, predicts quantities of interest then returns it to the solver (e.g. $CdRe$ for particular phase pair for which child class of dragModel is initialized). Modular structure of the OpenFOAM CFD library allows flexibly change solver sub-models. Lastly, the data-driven drag model is tested using the cases selected from the repository [21] dedicated for broad multiphase model development and testing. The description of the chosen case is provided in the next section.

On the Figure 2, the structural component of the DDM framework called "Models" contains defined drag prediction Keras model called from top-level workflow and exported on final step. It can be easily extended with models from other frameworks, and they may be loaded separately in various environments. "Data Structures" and "Data Adapters" components incorporate data loading, parsing, and filtering. "Data Adapters" allow to work with different raw data. They bridge particular file formats and allow to combine them into tables with "native" DDM framework formats implemented in "Data Structures" of Figure 3.

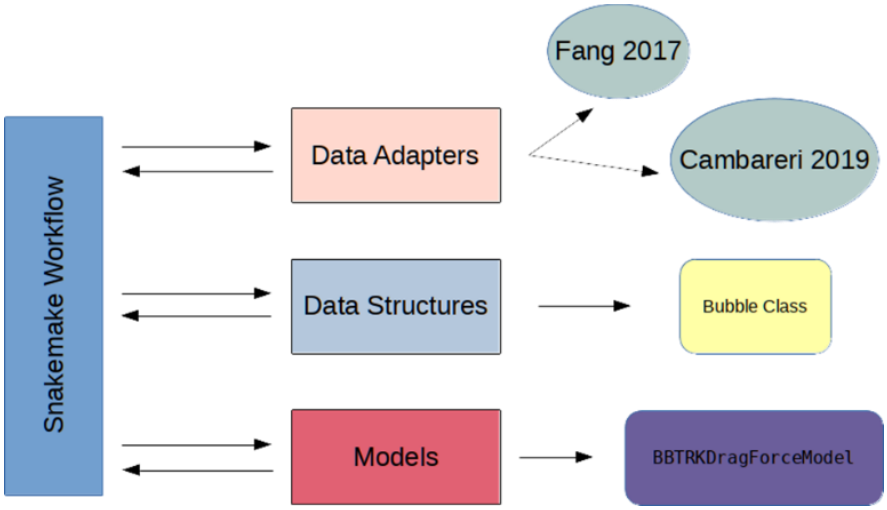


Figure 2 Top-level Workflow and low-level components of the HZDR ML Framework

Embedding of the Tensorflow binary model (for example, exported from *BBTRKDragForceModel* object on the Figure 2) for the purposes of inference demands further development of the new interfacial model. This new "ML" drag model implementing Tensorflow API must be deployed thereafter via an independent shared library. CFD solver reads the "*controlDict*" and loads corresponding *.so library which should provide a new interfacial model if it is called for drag inference in the simulation.

2.2. The bubble tracking data set

The high-fidelity bubble tracking data set is utilized for ML method-based interfacial force modeling tasks in this study. The data set is obtained from the state-of-the-art interface capturing DNS study on PWR-relevant subchannel domain by Fang et al. [5]. In this work, the level-set method [6] is deployed to explicitly track the advection of liquid-gas interface. The bubble tracking algorithm implemented in this work enables the access of the turbulence data in the vicinity of the interface. The data set adopted in this work is from the case "RE01", whose Reynolds number based on the subchannel hydraulic diameter is 29,079 and wherein

17 bubbles are tracked. Such information is conducive to research on interfacial force of bubbles in the sense that physical quantities local to the bubble interfaces would be influential to, or as a result of gas-liquid interaction. Note that the local liquid velocity and rate-of-strain are obtained from the 6-element thick "liquid shell" surrounding the interface. Summary of data source DNS case is listed in Table 1. The data content of the bubble tracking data set is summarized in Table 2.

Table 1 Summary of bubble tracking data set source DNS case.

Rod radius (mm)	4.57
Pitch-to-diameter (P/D) ratio	1.38
Domain size (mm^3)	$40.5 \times 12.6 \times 12.6$
Rod diameter (mm)	4.57
Pitch-to-diameter ratio	1.38
Number of bubbles	17
Time steps size (s)	$6.3e-5$
Total time steps	18816

Table 2 Data contents of bubble tracking data set

Bubble centroid position	(x_b, y_b, z_b)
Bubble/liquid velocity (m/s)	$(u_b, v_b, w_b)/(u_l, v_l, w_l)$
Local liquid shear ($1/s$)	$\dot{\gamma}_l$
Bubble equivalent radius (m)	$r_{b,eqv}$
Bubble mass (kg)	m_b
Bubble deformation f actor	γ_{def}

2.3. The feedforward neural network

In this work, the feedforward neural network (FNN) is adopted for the regression tasks. The FNNs are often considered as one of the simplest yet powerful neural network (NN) among NN families. The "feedforward" indicates the NN has no recurrent structure between neurons (often called nodes as well), and the information passing occurs in a uniform direction. Besides its simplicity, the Universal Approximation Theorem states that multi-layer FNN with at least one hidden layer, and appropriate smooth activation functions can approximate arbitrary functions and their derivatives [22]. This makes FNN an attractive option as a starting point of DDM framework implementation.

Figure 3 shows the schematic of a densely connected FNN (neurons in a certain layer are fully connected to those in the vicinity layers). The FNN can be roughly divided into three parts: an input layer, hidden layers, and an output layer. Each layer consists of neurons, the basic functional component that performs functional transformation to the incoming information (weighted input from the previous layer plus the bias). The weights associated with the connection between neurons ($\underline{\underline{w_j}}$) and the biases ($\underline{b_j}$) are the trainable parameters of an FNN. In the forward pass of the information, the input data ($\underline{x_i}$) received by the input layer undergoes a series of non-linear transformations towards the output layer, where the prediction on the quantities of interest ($\underline{y_{out}}$) is made. Mathematically, the forward propagation for a FNN with depth (number of hidden layers plus 1) N_L can be expressed as:

$$\underline{x}_1 = A_f(\underline{\underline{w}}_{1,i}\underline{x}_i + \underline{b}_1) \text{ for the input layer} \quad (4)$$

$$\underline{x}_j = A_f(\underline{\underline{w}}_j\underline{x}_{j-1} + \underline{b}_j) \text{ for } j \geq 2 \text{ and } j < N_L \quad (5)$$

$$y_{out} = A_f(w_{N_L}x_{N_L-1} + b_{N_L}) \text{ for the output layer} \quad (6)$$

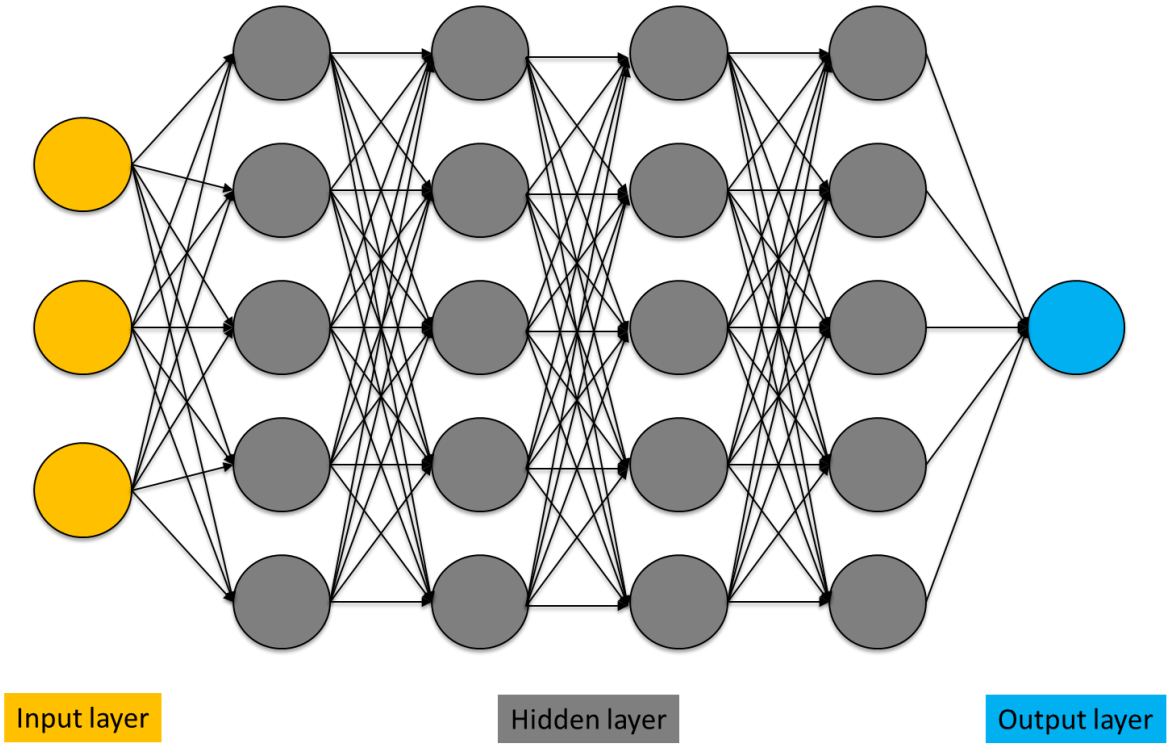
$$E = L(\underline{y}_{label}, \underline{y}_{out}) \quad (7)$$

$$w'_{kj} = w_{kj} - \epsilon_l(\frac{\partial E}{\partial w_{kj}}) \text{ for } k^{th} \text{ neuron in the } j^{th} \text{ layer} \quad (8)$$

$$b'_j = b_j - \epsilon_l(\frac{\partial E}{\partial b_j}) \quad (9)$$

A loss function (L) is used to evaluate the performance of the FNN by the discrepancies (or E_{out}) between prediction and the training label (y_{label}). Then, in the back propagation, the weights and biases of the network are adjusted according to the partial derivatives of the loss function with respect to each of the adjustable parameters, which are obtained by the chain rule.

Note the selection of hyperparameters, such as number of hidden layers (depth), number of neurons per layer (width) and learning rate are subject to the optimization process in each problem. Details of the FNN in the case studies will be listed.



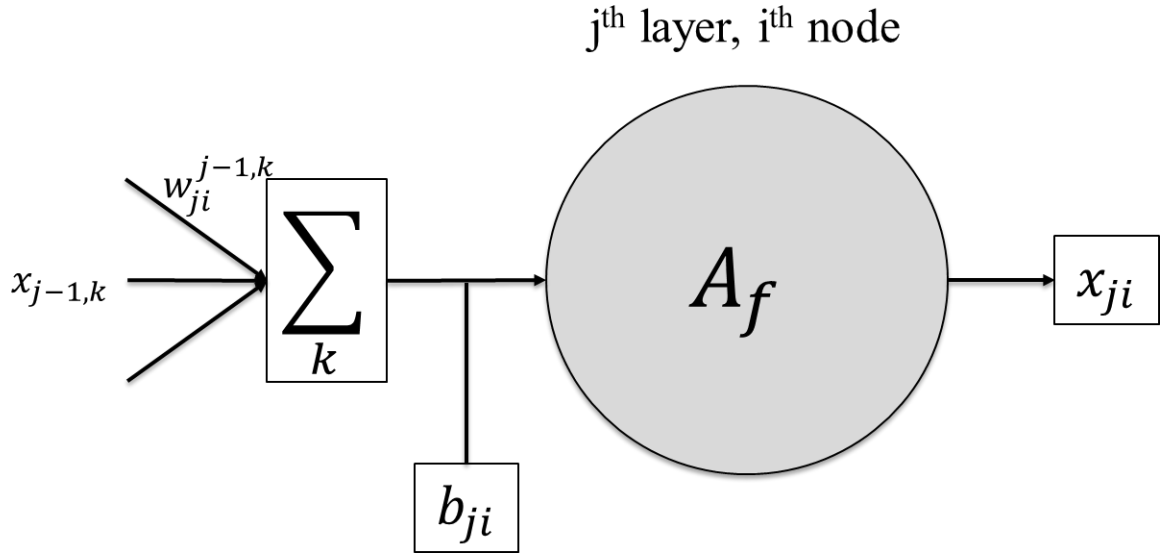


Figure 3 Schematic of (A) a densely connected FNN. This example of FNN has 3 input nodes, 4 hidden layers, 5 hidden nodes per layer, and 1 output node, and (B)a neuron.

3. Case Study and Discussion

In this section, the DDM framework is demonstrated via two case studies focusing on data-driven modeling of drag force in the bubbly flow. In the first problem, we verify our DDM methodology by training FNN using a designed problem and artificial data sets. Then, in the second problem study, we present the progress of development of a drag model for bubbly flow using the DNS bubble tracking data set. Details of problem formation and results will be discussed in the corresponding subsections.

3.1. Benchmark Case: Bubble Drag Coefficient Model with Artificial Data

In this case study, we demonstrate and verify the implementation of FNN in the DDM framework via a simple regression problem, where we aim to use FNN to emulate the drag correlation proposed by [23], as shown in eq.(10), in which the drag coefficient of a spherical bubble is correlated as the function of bubble Reynolds number:

$$C_d = \frac{24}{Re_{bub}}(1 + 0.15Re_{bub}^{0.687}) \approx fn_{FNN}(Re_{bub}) \tag{10}$$

where the fn_{FNN} denotes the approximation by FNN. The trained FNN will be used in place of the drag model in the baseline to examine the performance of FNN in the HZDR multiphase E-E framework against the results from the "parent" drag model in the same case setup.

Artificial data set is adopted to make the FNN emulate the drag correlation listed above. The artificial data set is generated by using known analytical mathematical relations or sampled from (complicated) correlations. Hence, there's a clear expectation that the FNN models mimic the underlying data generator (correlations adopted). Due to this reason, artificial data combined with simplified problem setup is suitable to diagnose implementation of the DDM.

Secondly, the cost and human effort involved in generating the artificial data is low. Compared to the DNS data, whose amount is often limited, the low generation cost of artificial data ensures a sufficient supply of data required by FNN to achieve assigned tasks.

3.1.1 The artificial data set

In this case study, the artificial data set will contain (Re_b, C_d) sampled from the correlation by [23] to fulfill the specified task. An imaginary laminar pipe flow is devised for the calculation of the drag coefficient of spherical bubbles. The fluid and geometric properties are listed in Table 3. Assuming the liquid flow is upward and fully-developed, the velocity profile in the pipe can be expressed by:

$$u_{liq}(r) = \frac{\nabla P + \rho_{liq} g}{4\mu_{liq}} R(1 - \frac{r^2}{R^2}) \tag{11}$$

Then, bubbles with random size uniformly sampled from range $r_{bub} \in [1.25\text{e}-4, 1.0\text{e}-2]$ are initialized at random radial position ($r \in [0, R]$) in the channel. Further, we assume the bubbles are under steady state, and influence from other neighboring bubbles are excluded. With the local liquid velocity at radial position r is known, the terminal velocity for a bubble rising in the laminar flow can be found by iteratively searching the bubble velocity (or essentially relative velocity) that would result in drag force on a par with the buoyancy experienced by the bubbles. Note that the search of terminal velocity started from static bubble state (i.e., $u_{bub} = 0$). Lastly, the bubble Reynolds number and the drag coefficient can be obtained using the following expressions:

$$Re_{bub} = \frac{\rho_{liq} u_{rel} D_{bub}}{\mu_{liq}} \tag{12}$$

$$F_d = \frac{1}{2} \rho_{liq} u^2{}_{rel} C_d (\pi r_{bub}^2) \tag{13}$$

The workflow and algorithm of artificial data generation are summarized in Figure 4.

Table 3 Summary of the pipe flow case for artificial data generation

radius of the pipe (m)	0.5
liquid/gas density ρ_{liq}/ρ_{gas} (kg/m^3)	$1000.0/1.0$
liquid/gas dynamic viscosity μ_{liq}/μ_{gas} $(Pa\ s)$	$1.0\text{e}-3/1.0\text{e}-5$
pressure gradient applied (N/m^3)	-9799.984

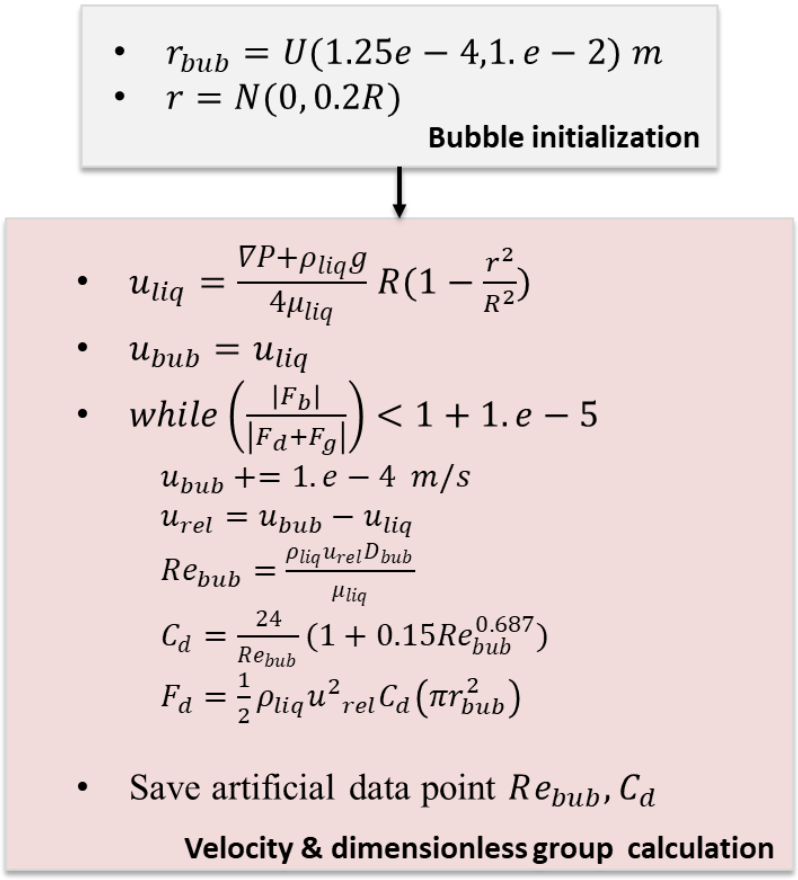
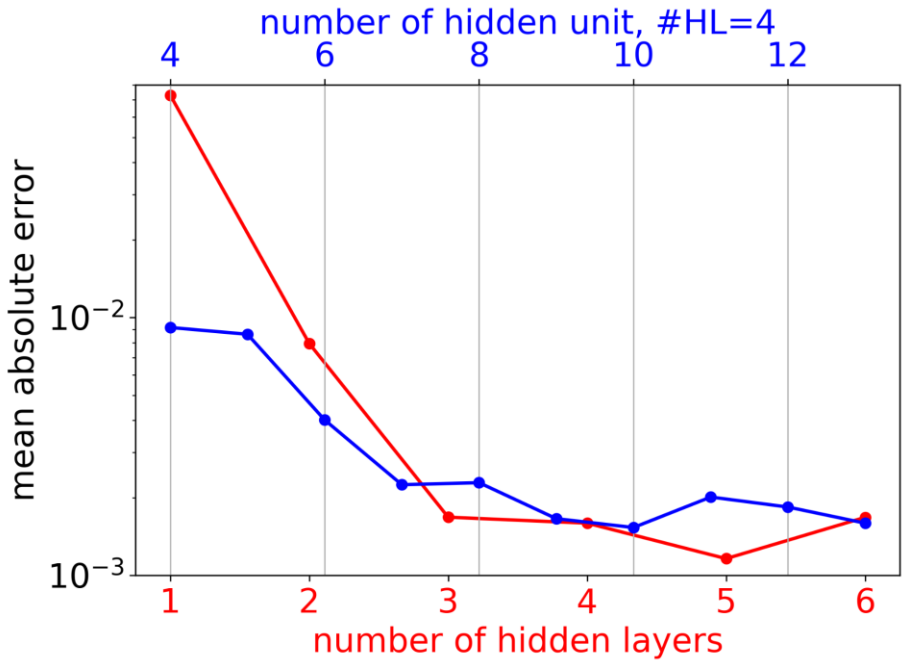


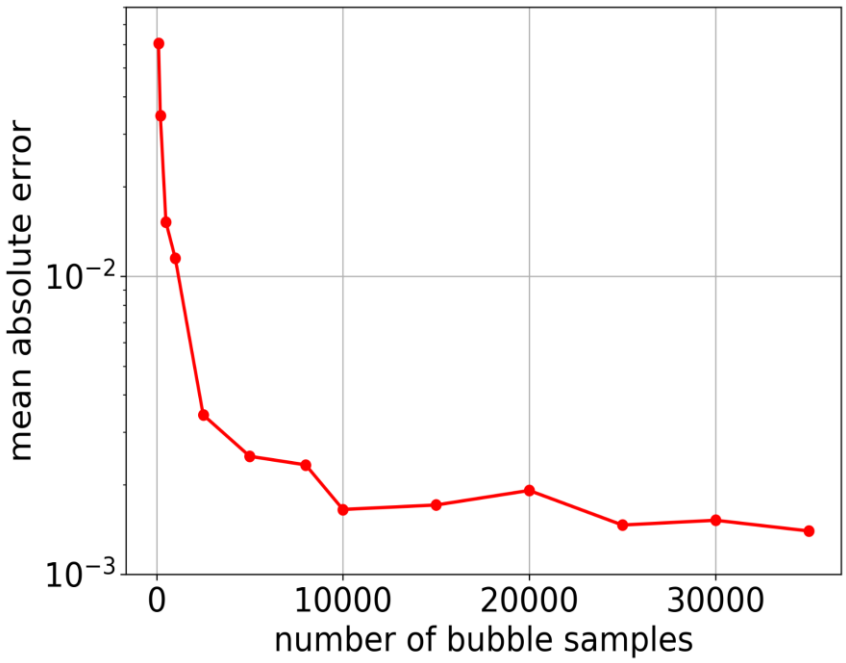
Figure 4 Workflow and algorithm for artificial data generation

3.1.2 FNN setup and training

The FNN is trained by the artificial data set to capture the mathematical mapping between Re_{bub} and C_d. Here, the L_1 lost function (mean absolute error, MAE), is adopted to characterize the model performance against the label (C_d). In order to achieve the balance between the variance and bias of the FNN model, the width and depth of the FNN is determined by observing the trend of MAE on testing set along the variation of the FNN size. Results are shown in Figure 5A. The convergence of the FNN performance can be observed with more than 4 hidden layers and 10 neurons per hidden layer. With the FNN's structure established, we further investigated the performance of the FNN with different amounts of artificial data learned. Figure 5B shows the trend of model MAE on the test set after training by different amounts of data. It can be identified that the models' MAE on the test set stabilized with more than 20,000 samples in the training set. The FNN model for the benchmark case is hence established. The summary of the FNN hyperparameters are listed in Table 4.



(A)



(B)

Figure 5 Variation of MAE on the test set along with (A) FNN width and depth, and (B) fixed FNN size (4 hidden layers, 10 hidden neurons) but varying number of samples learned

Table 4 Summary of FNN hyperparameters for the benchmark case

Activation function	relu
number of hidden layers	4
number of neurons per hidden layer	10
weight and bias initialization	random uniform over $[-0.05,0.05]$ different seeds for weights and biases
optimization algorithm	ADAM
Training/validation ratio	80%/20%
number of test data points	30,000

3.1.3 Results and discussion

The DDM framework along with the FNN model trained by artificial data is deployed to the bubbly flow in a pipe case by [24] from OpenFOAM Foundation release tutorials set. The specification of the test case is given in Table 5. After the training phase, the FNN along with the weights and biases are exported and used in place of the drag model in the baseline. The FNN model is deployed for on-the-fly prediction of local (cell-wise) C_d by receiving cell Re from the solver on per time step basis. Results of void fraction distribution were averaged over 10 seconds after convergence in residue was observed.

Table 5 Summary of the benchmark bubbly flow case

Geometric specifications	
pipe radius (m)	$1.905\text{e}-2$
pipe length (m)	$2\text{e}-2$
Flow setup and fluid properties	
water/air density ρ_{liq}/ρ_{gas} (kg/m^3)	$995.7/1.17$
water/air dynamic viscosity μ_{liq}/μ_{gas} $(Pa\ s)$	$7.99\text{e}-4/1.868\text{e}-5$
bubble diameter $D_{bub}(m)$	$3.4\text{e}-3$
Initial velocity and profiles (m/s)	0.916, uniform
time step Δt (s)	$3.0\text{e}-3$

Figure 6 shows the comparison of the resulting radial void fraction distribution based on FNN and the reference drag model [23]. Despite the slight discrepancies in peaking value and the corresponding radial position for the max void fraction, the FNN model trained by artificial data gave the void distribution in accordance with the reference drag model. This shows the FNN's competence capturing the non-linear mapping between C_d and Re_{bub} of the artificial data, and the implementation of the DDM framework to the HZDR multiphase E-E framework has been verified [16]. To summarize, this benchmark case has shown the potential of applying ML models for sophisticated two-phase flow modeling problems. On the other hand, for ML modeling tasks based on high fidelity numerical and experimental data, the performance of ML models lies in the quality of training data (e.g., proper physical representation of input to quantity of interest), which will be discussed in the next section.

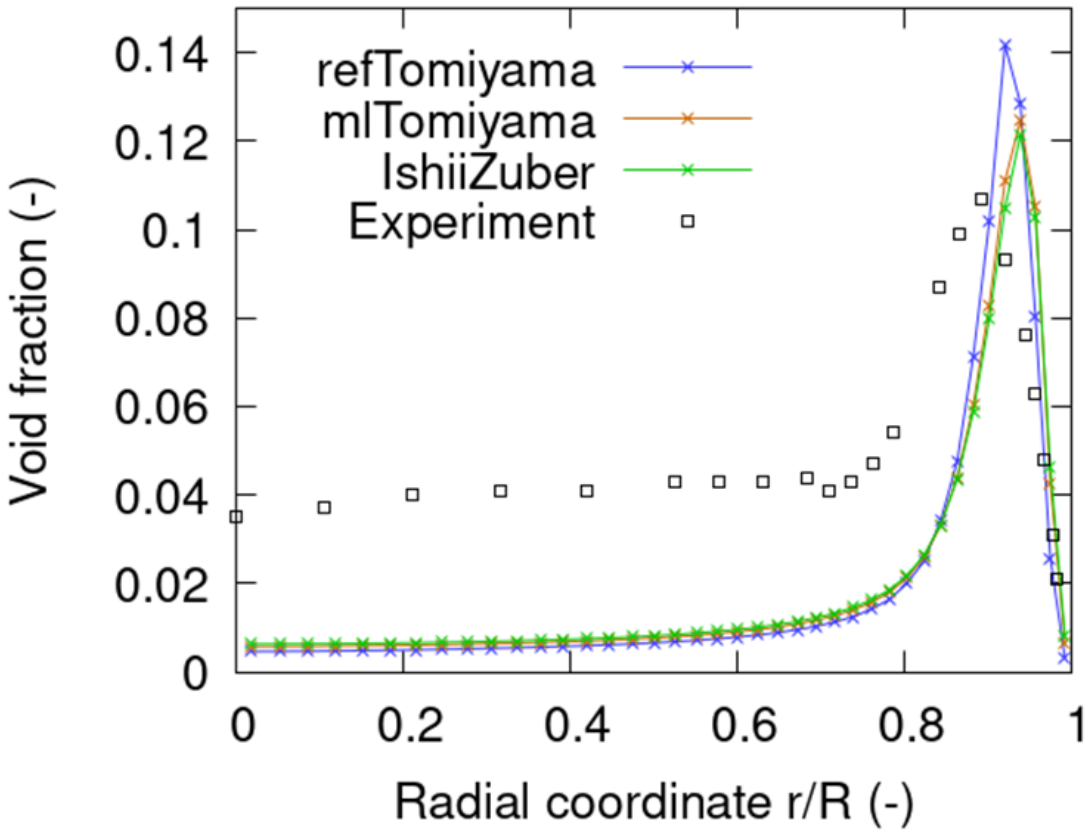


Figure 6 Comparison of void fraction distribution between reference drag force model [23], FNN drag model based on artificial data (this work), and the experimental data [24]

3.2. Bubble Drag Coefficient Model Utilizing Bubble Tracking Data Set

With the DDM framework verified in the benchmark problem, this part of the case study will utilize the bubble tracking data set [5] to formulate a ML drag model and discuss its performance. In order to extend the model's generalization ability across different geometric and flow setups, the non-dimensional groups derived from the raw data will be adopted as input physical features for the ML drag model. Besides the well-known dependence of drag coefficient on relative velocity (and thus Re_{bub}), the bubble Eötvös number (Eo) is also considered to account for the effect of bubble shape on the drag force experienced. The Eo in this case is derived from the bubble deformation factor, which is the ratio of the length of the semi-minor axis of deformed bubble to its equivalent radius. Note that the drag coefficient's dependence on the Eo can also be seen in the data pre-processing, which will be discussed in the later subsections. The equation form of the proposed problem can be written as:

$$C_d \approx fn_{FNN}\left(Re_{bub}, Eo\right) \tag{14}$$

While the plan for deriving input features is rather clear, the bubble drag coefficient is not directly available from the transient data set. In the interface-captured DNS, the turbulent bubbly flow is always transient. This poses challenges on obtaining the bubbles' drag coefficient using conventional definition under steady state. To tackle this, the following section will shed light on the derivation of bubble drag coefficient from the turbulent bubbly flow in order to train FNN via supervised learning.

3.2.1 Pseudo-steady state data filtering from Frenet frame

Three independent groups employed moving reference frame attached to the bubble and bubble-like bodies for modeling forces acting on a moving body in a fluid. Shew et al. [25] focused research efforts on small millimeter-sized bubbles, Fernandes et al. [26] investigated thick, slightly buoyant disks and Veldhuis et al. [27] built a model for light solid freely rising spheres. All models are based on Kirchhoff's equations and address general complexity of the researched zigzag and the spiral motions. In present research co-moving (Frenet) reference frame allowed unified approach for both simpler bubble motions and more complex ones.

Conventionally, the drag coefficient of a particle or bubble is computed from the force balance between drag and buoyancy after the object of interest reaches a steady state in a static fluid. However, such a definition is inapplicable in the case of turbulent bubbly flow, where bubbles and fluid dynamics show strong time-dependent characteristics when observed from fine time scales such as DNS time steps. In order to obtain the drag information from a transient flow phenomenon, data filtering based on pseudo-steady state assumption is proposed.

In this work, the pseudo-steady state filtering is carried out in the Frenet frame, which is a coordinate anchored on a moving particle, as shown in Figure 7. The construction basis of the Frenet frame can be expressed as:

$$\vec{t} = \frac{d\vec{r}}{ds} \tag{15}$$

$$\vec{n} = \frac{\frac{d\vec{t}}{ds}}{\left|\frac{d\vec{t}}{ds}\right|} \tag{16}$$

$$\vec{b} = \vec{t} \times \vec{n} \tag{17}$$

where $\vec{r}$ is the position vector with respect to the origin of the fixed reference frame. $\vec{t}$ is the unit vector parallel to the direction of bubble motion and tangential to the trajectory s, $\vec{n}$ accounts for the rate of change of tangential vector along trajectory s, and $\vec{b}$ is the unit vector orthogonal to both $\vec{t}$ and $\vec{n}$.

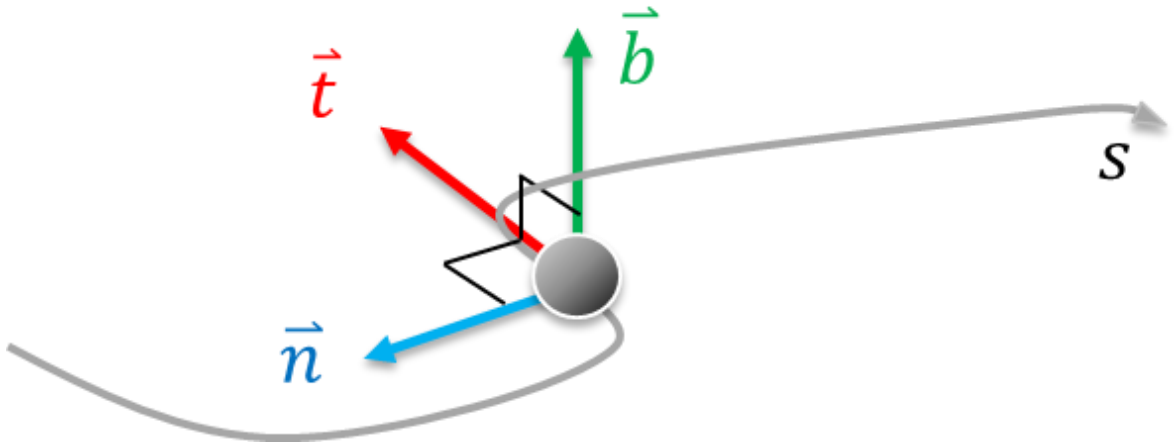


Figure 7 Schematic of the Frenet Frame attached to a spherical bubble with arbitrary motion

The momentum equation along the direction of the construction basis can be derived from the Kirchhoff equations [27]:

$$\{1+\frac{1}{2}(\frac{\rho_{liq}}{\rho_{gas}})\}m_{bub}\frac{d^2s}{dt^2}-\{1-(\frac{\rho_{liq}}{\rho_{gas}})\}m_{bub}g_t=F_t \tag{18}$$

$$\{1+\frac{1}{2}(\frac{\rho_{liq}}{\rho_{gas}})\}m_{bub}\kappa(\frac{ds}{dt})^2-\{1-(\frac{\rho_{liq}}{\rho_{gas}})\}m_{bub}g_n=F_n \tag{19}$$

$$-\{1-(\frac{\rho_{liq}}{\rho_{gas}})\}m_{bub}g_b=F_b \tag{20}$$

The first term on the left-hand side of eqs. (18) to (20) stands for the virtual mass force exerted on the direction of motion when the bubbles accelerate/decelerate. The second term in eqs. (18) and (19) represents the gravitational force experienced by the bubble. The term on the right-hand side of the equations stands for the summation of external forces exerted on the bubble. Here, we further focus on eq. due to our interest in the drag force. By definition, drag is the resistance along the direction of relative velocity of particles or bubbles in motion, or naturally $F_{drag}\,\vec{t}=-F_t$ under the Frenet Frame point of view. Therefore, to calculate the drag force in the turbulent bubbly flow, we applied time-averaging with a moving window of 2,000 time steps on the raw data. Then, we filter the raw bubble tracking data set by the pseudo-steady state, in which the bubbles are undergoing minimal acceleration. The metric adopted for pseudo-steady state filtering is the acceleration number [28, 29], as shown in the eq. (21).

$$A_c=\frac{|u_{liq}-u_{bub}|^2}{D_{bub}\frac{d|u_{liq}-u_{bub}|}{dt}} \tag{21}$$

Note that the bulk liquid velocity $u_{liq} = (0.27,0.0,0.0)\ m/s$ is adopted for the relative velocity calculation in the formulation above. This is owing to the fact that the flow has reached a statistically steady state in the subchannel domain. Also, using mean liquid velocity instead of the one local to the interface excludes the influence of turbulence fluctuation. The raw data is filtered with the following criteria: (1) $A_c < 0.1$, (2) $C_d < 3.0$ and (3) $0.5 < Eo < 3.25$.

Figure 8 shows the distribution of the filtered data points (dots) on the $C_d - Re_{bub}$ plane along with the drag correlation by [28](curves). A cluster appeared around $C_d \sim 0.75$, $1000 < Re_{bub} < 3000$ followed by a tail extending into the high C_d region. Also, it can be seen that the majority of filtered data points, in general, C_d close to the correlation of [28]. However, the pattern with regard to Eo appears to be different. As we further grouped the data points by its Eo, a clear dependence of $C_d - Re_{bub}$ on Eo showed up. We observed that as the Eo of the bubbles grows, the drag coefficient is larger than those with lower Eo. This further confirms the need to add Eo as one of the input features for the proposed regression problem. The distribution of the C_d of the filtered data points is shown in Figure 9.

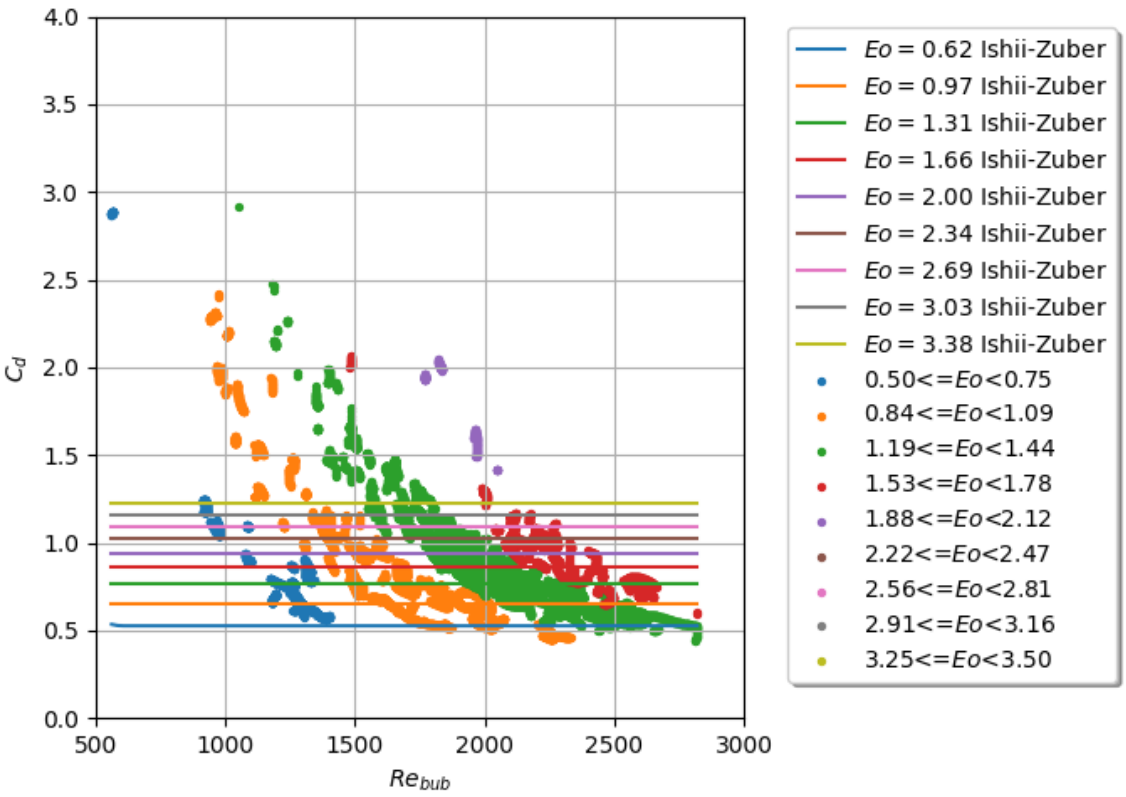


Figure 8 Scatter plot of $C_d - Re_{bub}$ of filtered data set (colored dots). Filtering criteria: $A_c < 0.1, C_d < 3.0, 0.5 < Eo < 3.25$. The filtered data set is further colored by the Eo and compared to the correlation by [28].

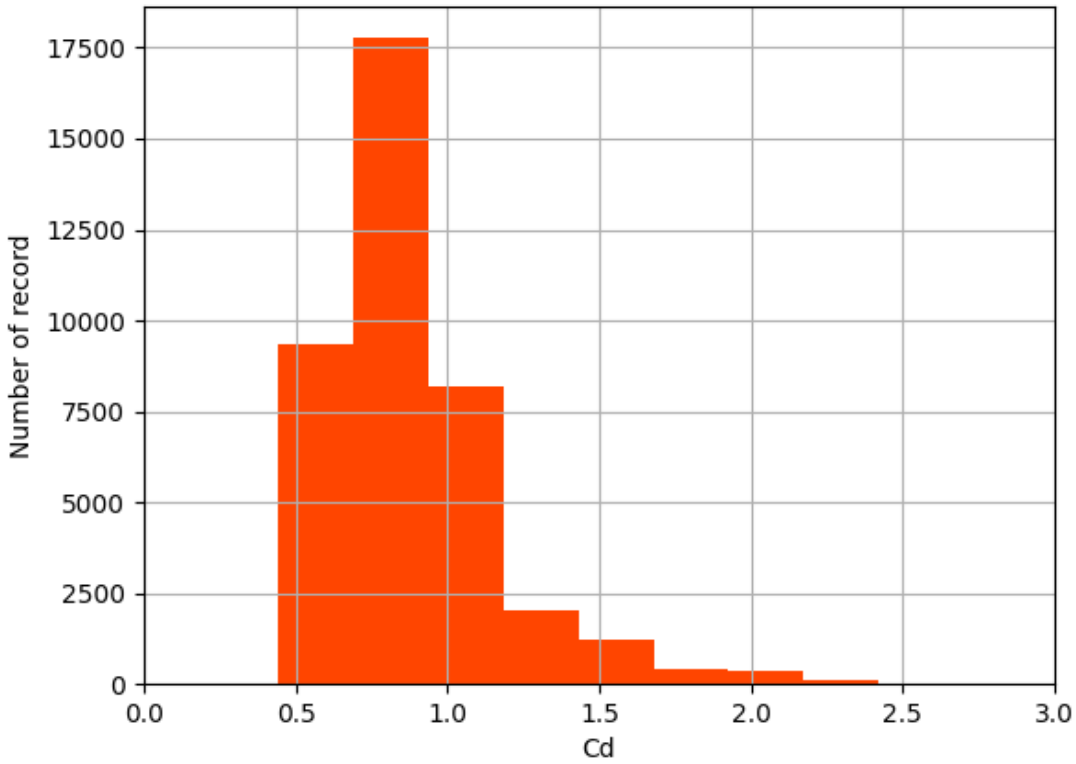


Figure 9 Distribution of the C_d of the filtered data points. Criteria: $A_c < 0.1$, $C_d < 3.0$, $0.5 < Eo < 3.25$.

3.2.2 FNN model setup and training

To capture the mapping between (Re_{bub}, Eo) and C_d, a FNN with specifications summarized in Table 6 is trained by the filtered data points. Result of FNN regression is shown in Figure 10. The FNN showed reasonable performance capturing mapping between input and label in the range $C_d < 1.5$. For the high drag coefficient range, the model tends to underpredict the drag coefficient. However, we should note that the model should be excluded from working in that range due to the insufficient data coverage of that range, which could result in high variance in the FNN's prediction on the drag coefficient. This FNN model will be referred to as the "ML drag model" onward.

Table 6 Summary of the FNN deployed for ML drag modeling

Activation function	relu
number of hidden layers	2
number of neurons per hidden layer	30
weight and bias initialization	random uniform over $[-0.05, 0.05]$ different seeds for weights and biases
optimization algorithm	ADAM
Training/validation ratio	80%/20%
number of test data points	7889

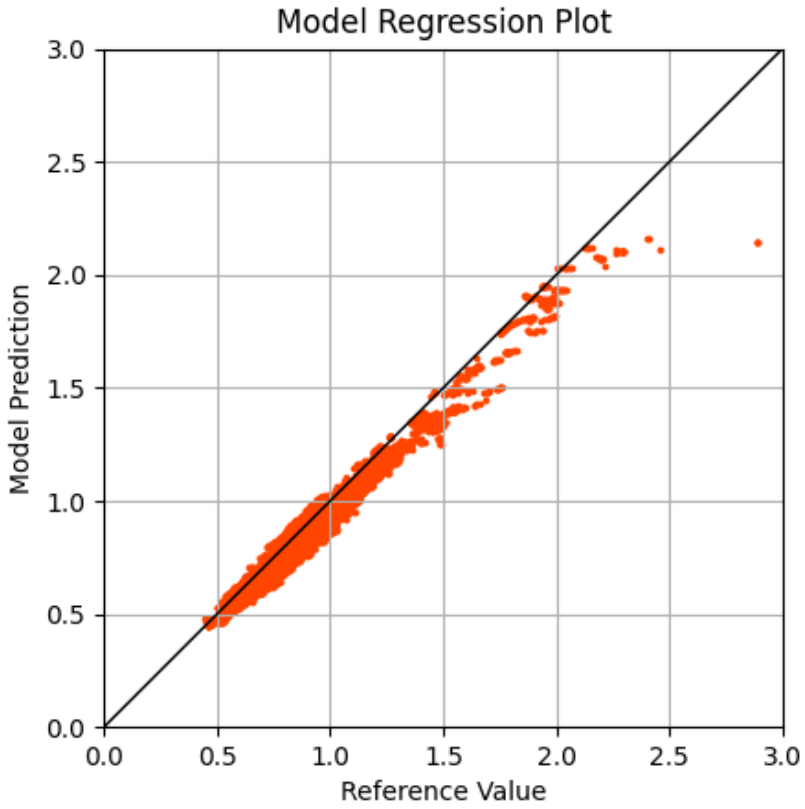


Figure 10 Results of model regression on the filtered DNS bubble tracking data set.

3.2.3 Results and discussion

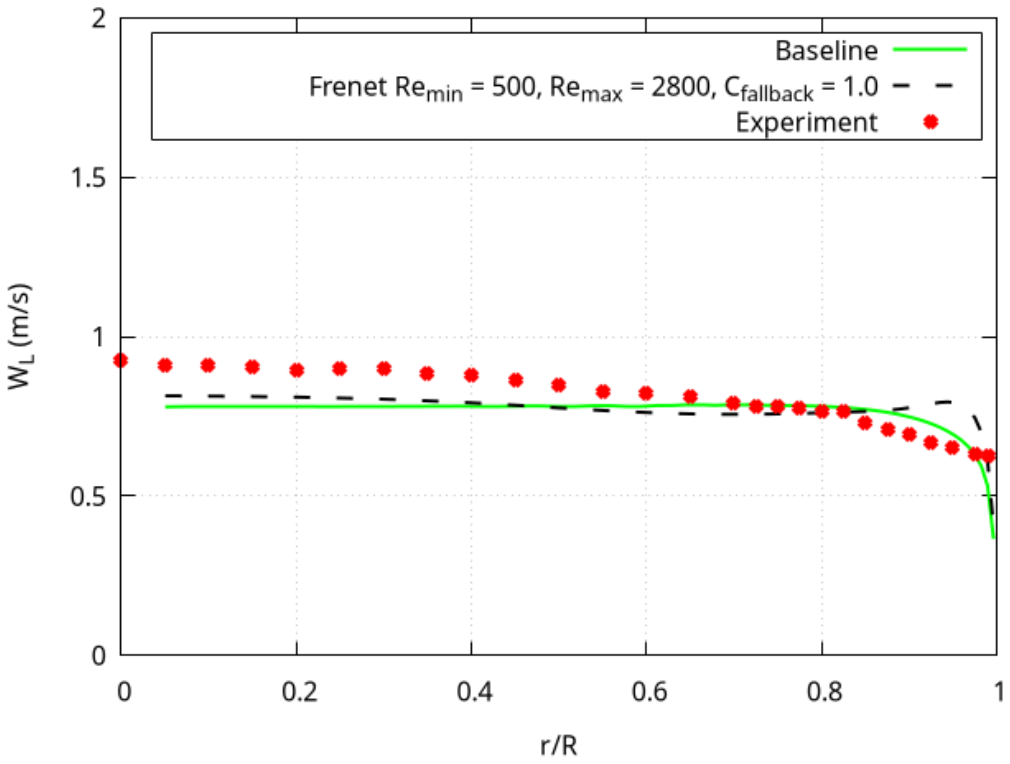
The ML drag model is deployed in place of the baseline drag model [23, 16, 31] as in the benchmark problem. The test case selected from the repository is the bubbly flow in a pipe emulating the experimental setup by Shawkat et al. [31] (referred to as the "S33" case onward). The S33 case is a monodisperse high Reynolds number vertical upward pipe flow and equipped with validation data for the gas phase fraction, liquid phase velocity and liquid phase turbulent fluctuations. The flow is assumed to be axi-symmetric, and the passage of the flow is represented by a wedge with 2.5° opening angle. Both phases enter the bottom of the domain and flow upward. Summary of the S33 case setup is listed in Table 7.

Table 7 Summary of the "S33" case based on Shawkat [31]

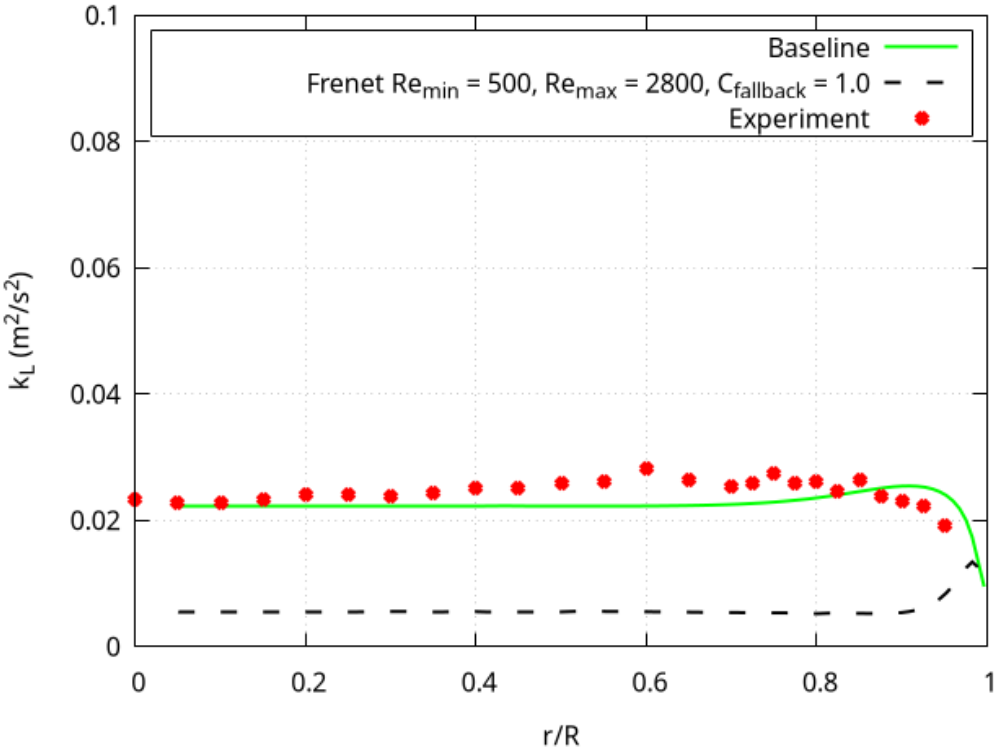
Geometric specifications	
pipe radius (m)	$1e-2$
pipe length (m)	9.56
Flow setup and fluid properties	
water/air density ρ_{liq}/ρ_{air} (kg/m^3)	$997.0/1.205$ (Standard ambient temperature and pressure, SATP)
water/air dynamic viscosity μ_{liq}/μ_{air} $(Pa\ s)$	$8.9e-4/1.82e-5$ (SATP)

bubble diameter $D_{bub}(m)$	4.4e − 3
Initial velocity and profiles (m/s)	0.78, uniform
time step Δt (s)	1.0e − 3
prescribed turbulence intensity	5%

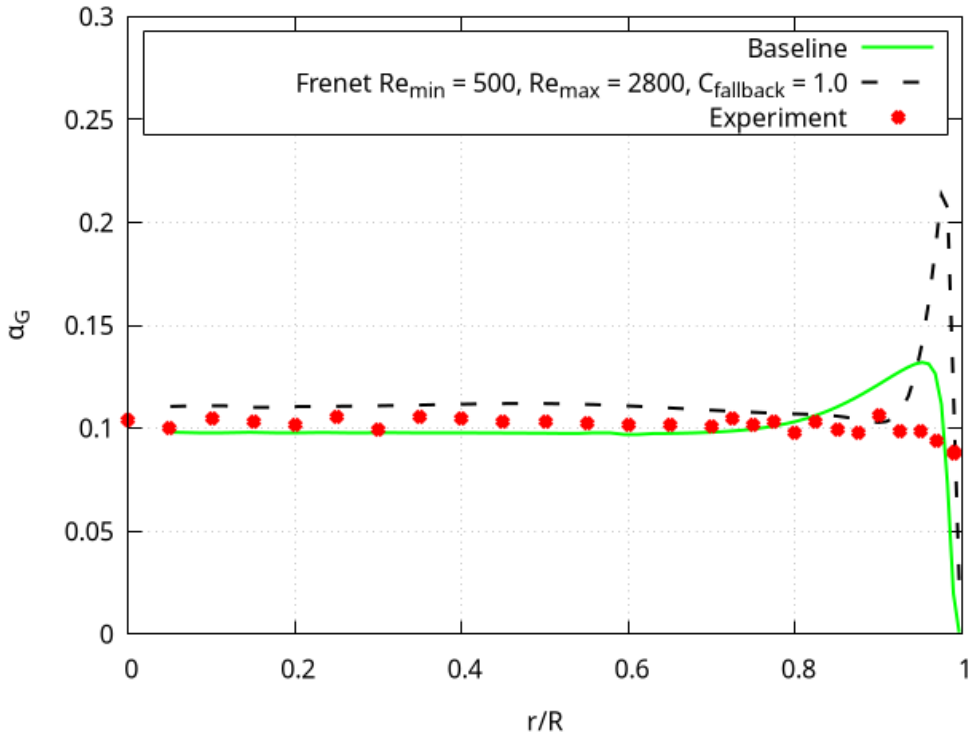
Figure 11 shows the comparison of streamwise velocity, turbulent kinetic energy (TKE) of liquid phase, and void fraction distribution from the experiment, ML drag model and the baseline, respectively. Note that hard limits were applied to the input bubble Reynolds number and output drag coefficient to prevent the ML drag model working in the extrapolation range and divergence of simulation due to unreasonable drag coefficient predictions. The resulting liquid velocity distribution was shown to be in close agreement with the baseline and the experimental measurement. On the other hand, overall underprediction of TKE is accompanied by the overprediction of the void fraction near wall. We currently attribute the observed discrepancies to two possible reasons. The first is the complicated interaction between the ML drag model to the other interfacial force closures. For instance, there's possibility that the output of the ML drag model made the lift force model work out of the applicable range. Secondly, the large discrepancy in void fraction prediction implies necessity to improve model's capability capturing near-wall flow behavior. One possible way to further investigate this is to re-examine necessity to expand input feature space of the FNN. For instance, distance to the wall and the local derivatives of liquid velocity may also be important features to characterize the drag force near wall.



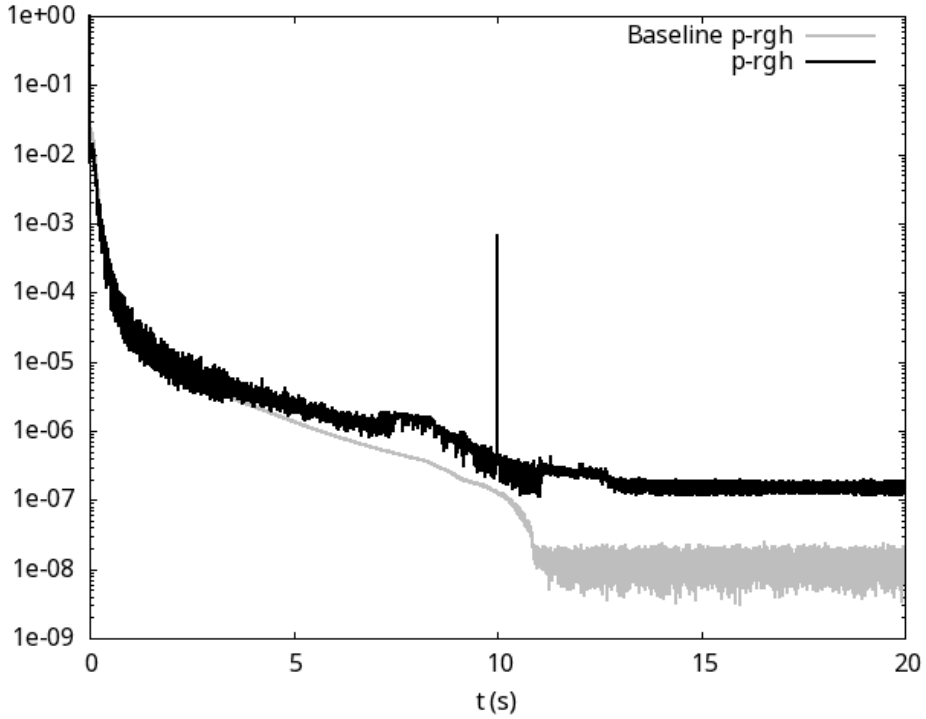
(A)


Baseline
Frenet Re$_{min}$ = 500, Re$_{max}$ = 2800, C$_{fallback}$ = 1.0
Experiment
k$_L$ (m^2/s^2)
r/R
0.1
0.08
0.06
0.04
0.02
0
0
0.2
0.4
0.6
0.8
1

(B)


Baseline
Frenet Re$_{min}$ = 500, Re$_{max}$ = 2800, C$_{fallback}$ = 1.0
Experiment
α$_G$
r/R
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.2
0.4
0.6
0.8
1

(C)



(D)

Figure 11 Comparison of resulting (A) streamwise velocity, (B) liquid phase TKE, (C) void fraction, and (D) residuals plot from case S33, ML drag model and the baseline.

4. Conclusions

This work established a data-driven modeling (DDM) framework and explored the topic of leveraging high-fidelity data sets from the interface-captured direct numerical simulation of bubbly flow for the closures of the bubble drag. The implementation of the DDM framework on the HZDR multiphase Eulerian-Eulerian framework is verified via a benchmark problem, in which a simple feedforward neural network trained by artificial data set managed to emulate the non-linear drag correlation and delivered results of bubbly flow in accordance with the reference case.

Secondly, we focused on the development of a data-driven drag force model utilizing the bubble tracking data set. To obtain drag coefficients from the unsteady turbulent bubbly flow, pseudo-steady state filtering with transformation to the Frenet Frame is applied. The ML drag model is examined in the bubbly flow case based on Shawkat et al. [31]. Resulting distribution of liquid TKE and void fraction urges further investigation on the interaction between ML drag model and other interfacial force surrogates.

Here, the authors would like to highlight the necessity of application-oriented high-fidelity simulation and data generation. Despite rapid growth in the capabilities of the computational systems, the DNS remains an expensive simulation that eludes from broad engineering applications. Hence, close collaboration between the simulation design and modelers is required to maximize the value of the DNS and the data sets. This work, as a

demonstration of leveraging data from DNS to inform the closure of engineering scale CFD simulation, would serve as an example for future works relevant to the topic.

Acknowledgement

The presented research was supported by Alexander von Humboldt Fellowship for experienced researchers (Bolotnov), Provost's Doctoral Fellowship by The Gradual School of NCSU, DOE-NEUP, and internship support by HZDR (Tai). The mesh of the data source simulation was generated using SimModeler by Simmetrix Inc, and the bubble tracking data was obtained using Acusim by Altair Engineering Inc.

Nomenclature

Nomenclature

A_f: activation function
$\underline{b_j}$: biases associated with the neurons in the jth layer of the FNN
$\vec{b}$: unit vector orthogonal to both $\vec{t}$ and $\vec{n}$ (Frenet Frame basis)
D_{bub}: bubble diameter
E : Loss/error between label and FNN prediction
F_b: buoyancy
F_d: drag force
F_g: body force due to gravity
L: loss function for the FNN
m_{bub}: mass of the bubble
M_{ik}: momentum source for phase k due to interfacial momentum exchange
$\vec{n}$: unit vector normal to the direction of particle/bubble motion accounting for rate of change of $\vec{t}$ along the trajectory (Frenet Frame basis)
p_k: pressure of phase k
r: radial position in the pipe
r_{bub}: radius of the bubble
$\vec{r}$: position vector with respect to the origin of the coordinate system
R: radius of the pipe
s : trajectory of particle/bubble (Frenet Frame)
t: time
$\vec{t}$: unit vector tangential to the direction of particle/bubble motion (Frenet Frame basis)
u_k: velocity of phase k
$\underline{\underline{w_j}}$: weight matrix associated with neurons in the jth layer in the FNN

Greek alphabet

α_k: volumetric fraction occupied by phase k
ϵ_l: learning rate
Γ_k: mass source term for phase k
μ: dynamic viscosity
ν: kinematic viscosity
ρ_k: density of phase k
σ: surface tension coefficient
τ_k: stress tensor of phase k

$\underline{x_j}$: output from the neurons in the jth layer of the FNN

Dimensionless groups

$Re_{bub} = \frac{\rho_{liq} u_{rel} D_b}{\mu_{liq}}$: Bubble Reynolds number

$Eo = \frac{(\rho_{liq} - \rho_{gas}) g D_{bub}}{\sigma}$: Eötvös number

$A_c = \frac{|u_{liq} - u_{air}|^2}{D_{bub} \frac{d|u_{liq} - u_{air}|}{dt}}$: Acceleration number

sub-, and superscripts

X_{liq}: quantity or properties of liquid phase

X_{gas}: quantity or properties of gas phase

X_{rel}: relative quantities

X_{bub}: quantities associated with bubbles

References

[1] M. Z. Podowski, "Multidimensional modeling of two-phase flow and heat transfer," *International Journal of Numerical Methods for Heat & Fluid Flow*, vol. 18, no. 3/4, pp. 491-513, 2007.

[2] D. A. Drew and S. L. Passman, Theory of Multicomponent Fluids, Springer, 1999.

[3] M. Ishii and T. Hibiki, Thermo-Fluid Dynamics of Two-Phase Flow, Springer, 2011.

[4] I. Khan, M. Wang, Y. Zhang, W. Tian, G. Su and S. Qiu, "Two-phase bubbly flow simulation using CFD method: A review of models for interfacial forces," *Progress in Nuclear Energy*, vol. 125, p. 103360, 2020.

[5] J. Fang, M. Rasquin and I. A. Bolotnov, "Interface tracking simulations of bubbly flows in PWR relevant geometries," *Nuclear Engineering and Design 312*, pp. 205-213, 2017.

[6] M. Sussman, E. Fatemi, P. Smereka and S. Osher, "An adaptive level set approach for incompressible two-phase flows," *J. Comput. Phys.*, vol. 27, pp. 663-680, 1998.

[7] J. J.Cambareri, J. Fang and I. A. Bolotnov, "Interface capturing simulations of bubble population effects in PWR subchannels," *Nuclear Engineering and Design*, vol. 365, no. 15, 2020.

[8] N. Saini and I. A. Bolotnov, "Interface capturing simulations of droplet interaction with spacer grids under DFFB conditions," *Nuclear Engineering and Design*, vol. 364, no. 1, 2020.

[9] Y. Fan, J. Fang and I. Bolotnov, "Complex bubble deformation and break-up dynamics studies using interface capturing approach," *Experimental and Computational Multiphase Flow*, vol. 3, no. 3, pp. 139-151, 2020.

[10] R. D. Sandberg, R. Tan, J. Weatheritt, A. Ooi, A. Haghiri, V. Michelassi and G. Laskowski, "Applying Machine Learnt Explicit Algebraic Stress and Scalar Flux Models to a Fundamental Trailing Edge Slot," *J. Turbomach.*, vol. 140, no. 10, p. 101008, 2018.

[11] Y. Zhu, N. T. Dinh, N. Saini and I. A. Bolotnov, "An adaptive knowledge-based data-driven approach for turbulence modeling using ensemble learning technique under complex flow configuration: 3D PWR sub-channel with DNS data," *Nuclear Engineering and Design*, vol. 393, p. 111814, 2022.

[12] H. Bao, N. Dinh, L. Lin, R. Youngblood, J. Lane and H. Zhang, "Using deep learning to explore local physical similarity for global-scale bridging in thermal-hydraulic simulation," *Annals of Nuclear Energy*, vol. 147, p. 107684, 2020.

[13] M. Ma, J. Lu and G. Tryggvason, "Using statistical learning to close two-fluid multiphase flow equations for a simple bubbly system," *Physics of Fluids,* vol. 27, p. 092101, 2015.

[14] M. Ma, J. Lu and G. Tryggvason, "Using statistical learning to close two-fluid multiphase flow equations for bubbly flows in vertical channels," *International Journal of Multiphase Flow,* vol. 85, pp. 336-347, 2016.

[15] Y. Liu, N. Dinh, Y. Sato and B. Niceno, "Data-driven modeling for boiling heat transfer: Using deep neural networks and high-fidelity simulation results," *Applied Thermal Engineering,* vol. 144, pp. 305-320, 2018.

[16] F. Schlegel, K. G. Bilde, M. Draw, I. Evdokimov, S. Hänsch, V. V. Kamble, H. Khan, B. Krull, R. Lehnigk, J. Li, H. Lyu, R. Meller, G. Petelin and M. Tekavčič, "HZDR Multiphase Addon for OpenFOAM. Rodare," 21 June 2021. [Online]. Available: https://rodare.hzdr.de/record/1742.

[17] F. Mölder, K. P. Jablonski, B. Letcher, M. B. Hall, C. H. Tomkins-Tinch, V. Sochat, J. Forster, S. Lee, S. O. Twardziok, A. Kanitz, A. Wilm, M. Holtgrewe, S. Rahmann, S. Nahnsen and J. Köster, "Sustainable data analysis with Snakemake [version 1; peer review: 1 approved, 1 approved with reservations]," *F1000Research,* 2021.

[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, C. C. Zhifeng Chen, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz and a. others, "TensorFlow: Large-scale machine learning on heterogeneous systems, Software available from tensorflow.org.," 2015.

[19] F. Chollet and others, "Keras," 2015. [Online]. Available: https://keras.io.

[20] "tensorflow/c_api.h at master - GitHub," Tensorflow.org, [Online]. Available: https://github.com/tensorflow/tensorflow/tree/master/tensorflow/c.

[21] S. Hänsch, M. Draw, I. Evdokimov, H. Khan, B. Krull, R. Lehnigk, Y. Liao, H. Lyu, R. Meller, F. Schlegel and M. Tekavčič, "HZDR Multiphase Case Collection for OpenFOAM," 15 Feburary 2021. [Online]. Available: https://rodare.hzdr.de/record/812.

[22] K. Hornik, M. Stinchcombe and H. White, "Universal Approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neurul Networks 3,* pp. 551-560, 1990.

[23] A. Tomiyama, I. Kataoka, I. Zun and T. Sakaguchi, "Drag coefficients of single bubbles under normal and micro gravity conditions," *JSME International Journal 41,* pp. 472-479, 1998.

[24] C. Grossetete, "Experimental investigation and numerical simulations of void profile development in a vertical cylindrical pipe," (EDF--96-NB-00120), France, 1995.

[25] W. L. Shew and J.-F. Pinton, "Dynamical Model of Bubble Path Instability," *Phys. Rev. Lett,* vol. 97, p. 144508, 2006.

[26] P. C. Fernandes, E. Patricia, R. Frédéric and J. Magnaudet, "Dynamics of axisymmetric bodies rising along a zigzag path," *Journal of Fluid Mechanics,* vol. 606, pp. 209-223, 2008.

[27] C. Veldhuis, A. Biesheuvel and D. Lohse, "Freely rising light solid spheres," *International Journal of Multiphase Flow,* vol. 2009, pp. 312-322, 2009.

[28] I. Kim, S. Elghobashi and W. Sirignano, "On the equation for spherical-particle motion: Effect of Reynolds and acceleration numbers," *Journal of Fluid Mechanics,* vol. 367, pp. 221-253, 1998.

[29] C. W. M. v. d. Geld, H. v. Wingaarden and B. A. Brand, "Experiments on the effect of

acceleration on the drag of tapwater bubbles," *Experiments in Fluids,* vol. 31, pp. 708-722, 2001.

[30] M. Ishii and N. Zuber, "Drag coefficient and relative velocity in bubbly, droplet or particulate flows," *AlChE Journal 25,* pp. 843-855, 1979.

[31] M. Shawkat, C. Ching and M. Shoukri, "Bubble and liquid turbulence characteristics of bubbly flow in a large diameter vertical pipe," *International Journal of Multiphase Flow,,* vol. 34, no. 8, pp. 767-785, 2008.

[32] Y. Liao, K. Upadhyay and F. Schlegel, "Eulerian-Eulerian two-fluid model for laminar bubbly pipe flows: Validation of the baseline model," *Computers & Fluids 202,* p. 104496, 2020.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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

References

[1] M. Z. Podowski, "Multidimensional modeling of two-phase flow and heat transfer," *International Journal of Numerical Methods for Heat & Fluid Flow,* vol. 18, no. 3/4, pp. 491-513, 2007.

[2] D. A. Drew and S. L. Passman, Theory of Multicomponent Fluids, Springer, 1999.

[3] M. Ishii and T. Hibiki, Thermo-Fluid Dynamics of Two-Phase Flow, Springer, 2011.

[4] I. Khan, M. Wang, Y. Zhang, W. Tian, G. Su and S. Qiu, "Two-phase bubbly flow simulation using CFD method: A review of models for interfacial forces," *Progress in Nuclear Energy,* vol. 125, p. 103360, 2020.

[5] J. Fang, M. Rasquin and I. A. Bolotnov, "Interface tracking simulations of bubbly flows in PWR relevant geometries," *Nuclear Engineering and Design 312,* pp. 205-213, 2017.

[6] M. Sussman, E. Fatemi, P. Smereka and S. Osher, "An adaptive level set approach for incompressible two-phase flows," *J. Comput. Phys.*, vol. 27, pp. 663-680, 1998.

[7] J. J.Cambareri, J. Fang and I. A. Bolotnov, "Interface capturing simulations of bubble population effects in PWR subchannels," *Nuclear Engineering and Design,* vol. 365, no. 15, 2020.

[8] N. Saini and I. A. Bolotnov, "Interface capturing simulations of droplet interaction with spacer grids under DFFB conditions," *Nuclear Engineering and Design,* vol. 364, no. 1, 2020.

[9] Y. Fan, J. Fang and I. Bolotnov, "Complex bubble deformation and break-up dynamics studies using interface capturing approach," *Experimental and Computational Multiphase Flow,* vol. 3, no. 3, pp. 139-151, 2020.

[10] R. D. Sandberg, R. Tan, J. Weatheritt, A. Ooi, A. Haghiri, V. Michelassi and G. Laskowski, "Applying Machine Learnt Explicit Algebraic Stress and Scalar Flux Models to a Fundamental Trailing Edge Slot," *J. Turbomach.,* vol. 140, no. 10, p. 101008, 2018.

[11] Y. Zhu, N. T. Dinh, N. Saini and I. A. Bolotnov, "An adaptive knowledge-based data-driven approach for turbulence modeling using ensemble learning technique under complex flow configuration: 3D PWR sub-channel with DNS data," *Nuclear Engineering and Design,* vol. 393, p. 111814, 2022.

[12] H. Bao, N. Dinh, L. Lin, R. Youngblood, J. Lane and H. Zhang, "Using deep learning to explore local physical similarity for global-scale bridging in thermal-hydraulic simulation," *Annals of Nuclear Energy,* vol. 147, p. 107684, 2020.

[13] M. Ma, J. Lu and G. Tryggvason, "Using statistical learning to close two-fluid multiphase flow equations for a simple bubbly system," *Physics of Fluids,* vol. 27, p. 092101, 2015.

[14] M. Ma, J. Lu and G. Tryggvason, "Using statistical learning to close two-fluid multiphase flow equations for bubbly flows in vertical channels," *International Journal of Multiphase Flow,* vol. 85, pp. 336-347, 2016.

[15] Y. Liu, N. Dinh, Y. Sato and B. Niceno, "Data-driven modeling for boiling heat transfer: Using deep neural networks and high-fidelity simulation results," *Applied Thermal Engineering,* vol. 144, pp. 305-320, 2018.

[16] F. Schlegel, K. G. Bilde, M. Draw, I. Evdokimov, S. Hänsch, V. V. Kamble, H. Khan, B. Krull, R. Lehnigk, J. Li, H. Lyu, R. Meller, G. Petelin and M. Tekavčič, "HZDR

Multiphase Addon for OpenFOAM. Rodare," 21 June 2021. [Online]. Available: https://rodare.hzdr.de/record/1742.

[17] F. Mölder, K. P. Jablonski, B. Letcher, M. B. Hall, C. H. Tomkins-Tinch, V. Sochat, J. Forster, S. Lee, S. O. Twardziok, A. Kanitz, A. Wilm, M. Holtgrewe, S. Rahmann, S. Nahnsen and J. Köster, "Sustainable data analysis with Snakemake [version 1; peer review: 1 approved, 1 approved with reservations]," *F1000Research,* 2021.

[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, C. C. Zhifeng Chen, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz and a. others, "TensorFlow: Large-scale machine learning on heterogeneous systems, Software available from tensorflow.org.," 2015.

[19] F. Chollet and others, "Keras," 2015. [Online]. Available: https://keras.io.

[20] "tensorflow/c_api.h at master - GitHub," Tensorflow.org, [Online]. Available: https://github.com/tensorflow/tensorflow/tree/master/tensorflow/c.

[21] S. Hänsch, M. Draw, I. Evdokimov, H. Khan, B. Krull, R. Lehnigk, Y. Liao, H. Lyu, R. Meller, F. Schlegel and M. Tekavčič, "HZDR Multiphase Case Collection for OpenFOAM," 15 Feburary 2021. [Online]. Available: https://rodare.hzdr.de/record/812.

[22] K. Hornik, M. Stinchcombe and H. White, "Universal Approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neurul Networks 3,* pp. 551-560, 1990.

[23] A. Tomiyama, I. Kataoka, I. Zun and T. Sakaguchi, "Drag coefficients of single bubbles under normal and micro gravity conditions," *JSME International Journal 41,* pp. 472-479, 1998.

[24] C. Grossetete, "Experimental investigation and numerical simulations of void profile development in a vertical cylindrical pipe," (EDF--96-NB-00120), France, 1995.

[25] W. L. Shew and J.-F. Pinton, "Dynamical Model of Bubble Path Instability," *Phys. Rev. Lett,* vol. 97, p. 144508, 2006.

[26] P. C. Fernandes, E. Patricia, R. Frédéric and J. Magnaudet, "Dynamics of axisymmetric bodies rising along a zigzag path," *Journal of Fluid Mechanics,* vol. 606, pp. 209-223, 2008.

[27] C. Veldhuis, A. Biesheuvel and D. Lohse, "Freely rising light solid spheres," *International Journal of Multiphase Flow,* vol. 2009, pp. 312-322, 2009.

[28] I. Kim, S. Elghobashi and W. Sirignano, "On the equation for spherical-particle motion: Effect of Reynolds and acceleration numbers," *Journal of Fluid Mechanics,* vol. 367, pp. 221-253, 1998.

[29] C. W. M. v. d. Geld, H. v. Wingaarden and B. A. Brand, "Experiments on the effect of acceleration on the drag of tapwater bubbles," *Experiments in Fluids,* vol. 31, pp. 708-722, 2001.

[30] M. Ishii and N. Zuber, "Drag coefficient and relative velocity in bubbly, droplet or particulate flows," *AlChE Journal 25,* pp. 843-855, 1979.

[31] M. Shawkat, C. Ching and M. Shoukri, "Bubble and liquid turbulence characteristics of bubbly flow in a large diameter vertical pipe," *International Journal of Multiphase Flow,,* vol. 34, no. 8, pp. 767-785, 2008.

[32] Y. Liao, K. Upadhyay and F. Schlegel, "Eulerian-Eulerian two-fluid model for laminar bubbly pipe flows: Validation of the baseline model," *Computers & Fluids 202,* p. 104496, 2020.