Article

Machine Learning Augmented Two-Fluid Model for Segregated Flow

Ayush Rastogi [1] and Yilin Fan [2],*

[1] BP America Inc., Denver, CO 80202, USA; ayush.rastogi@outlook.com
[2] Petroleum Engineering Department, Colorado School of Mines, Golden, CO 80401, USA
* Correspondence: yilinfan@mines.edu

Abstract: Segregated flow, including stratified and annular flows, is commonly encountered in several practical applications such as chemical, nuclear, refrigeration, and oil and gas industries. Accurate prediction of liquid holdup and the pressure gradient is of great importance in terms of system design and optimization. The current most widely accepted model for segregated flow is a physics-based two-fluid model that treats gas and liquid phases separately by incorporating mass and momentum conservation equations. It requires empirically derived closure relationships that have the limitation of being applicable only under a narrow range of input parameters under which they were developed. In this paper, we proposed a more generalized machine learning augmented two-fluid model, using a database that spans the range of various flowing conditions and fluid properties. Machine learning algorithms such as random forest, neural networks, and gradient boosting were tested for the best performing data-driven predictive model. The new model proposed in this work successfully captures the complex, dynamic, and non-linear relationships between the friction factor and flowing conditions. A comprehensive model evaluation against nineteen existing correlations shows the best results from the proposed model.

Keywords: segregated flow; multiphase flow; machine learning; two-fluid model; hydraulic models; data-driven approach

1. Introduction

Segregated flow, which includes stratified smooth, stratified wavy, and annular flows, is commonly encountered in several practical applications such as chemical, nuclear, refrigeration, and oil and gas industries. Accurate prediction of liquid holdup and the pressure gradient is of great importance in terms of system design and optimization. This section first reviews the two-fluid hydraulic model that is widely used in the industry for segregated flow pressure gradient and liquid holdup predictions, followed by a review of the common application of machine learning (ML) in multiphase flow modeling.

1.1. Review of Two-Fluid Hydraulic Models for Segregated Flow

The current most widely accepted two-fluid hydraulic models for steady-state gas-liquid two-phase segregated flow are primarily based on the mass and momentum conservation equations for gas and liquid phases separately [1,2], which is also referred to as a two-fluid model in the current study. Separate governing equations are derived by treating each phase or component as a separate entity. Figure 1 depicts a two-phase flow system, as well as the forces acting on the gas and liquid phase. Equations (1) and (2) provide the integral form of conservation of momentum equations for gas and liquid under steady-state conditions.

$$-A_L\left(\frac{dp}{dL}\right)_L - \tau_{WL}S_L + \tau_I S_I - A_L\rho_L g\sin\theta = 0, \tag{1}$$

$$-A_C\left(\frac{dp}{dL}\right)_C - \tau_{WC}S_C - \tau_I S_I - A_C\rho_C g\sin\theta = 0, \tag{2}$$

where A_L is the area occupied by the liquid phase, A_C is the area occupied by the gas core, τ_{WL} is the liquid wall shear stress, τ_{WC} is the wall shear stress for the gas core, τ_I is the interfacial shear stress, S_L is the liquid wetted perimeter, S_C is the gas core wetted perimeter, and S_I is the interfacial perimeter. dp/dL is the pressure gradient, and ρ_L and ρ_C are liquid and gas core densities respectively. The subscript L represents the liquid phase, C refers to the gas core with the entrained liquid droplet. θ is the inclination angle from the horizontal. v_L and v_C are the actual in-situ velocities for the liquid and gas core, respectively.
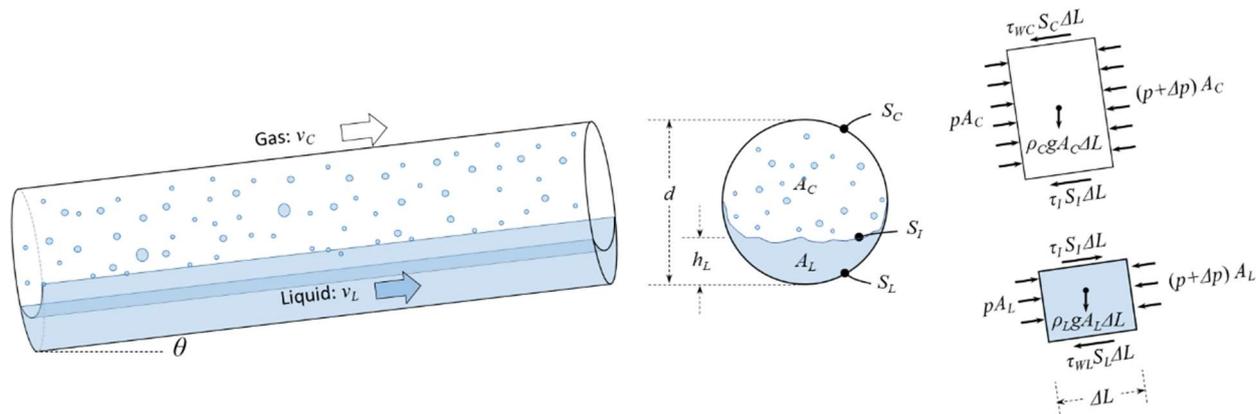
Figure 1. Parameters and forces in two-fluid model for segregated flow.

Equation (3) is the combined momentum equation of Equations (1) and (2) by canceling out the pressure gradient term, which can be used to solve for the liquid holdup, while Equations (4)–(6) provide the formulae to calculate shear stresses [1,3].

$$\tau_{WL}\frac{S_L}{A_L} - \tau_{WC}\frac{S_C}{A_C} - \tau_I\left(\frac{S_I}{A_L} + \frac{S_I}{A_C}\right) + (\rho_L - \rho_C)g\sin\theta = 0, \tag{3}$$

$$\tau_{WL} = f_L\frac{\rho_L v_L |v_L|}{2}, \tag{4}$$

$$\tau_{WC} = f_C\frac{\rho_C v_C |v_C|}{2}, \tag{5}$$

$$\tau_I = f_I\frac{\rho_C(v_C - v_L)|v_C - v_L|}{2}, \tag{6}$$

To solve Equation (3), several closure relationships are needed, including liquid and gas wall friction factors (f_L and f_C), interfacial friction factor (f_I), geometrical parameters such as S_L, S_C, S_I, A_C, and A_L, and liquid entrainment (F_E) defined by the entrained liquid mass flow rate in the gas core over the total liquid mass flow rate.

There are several different models for the geometrical parameters, such as the flat interface assumption that is widely used by various studies thereafter [3–7], apparent rough surface that assumes a liquid film with a constant thickness [8,9], and unified "double-circle" model and its simplified version applied to all inclination angles [10,11]. The geometrical parameters can be easily extended to annular flow by assuming a uniform film thickness [12,13].

Previous studies have shown that the interfacial friction factor plays the most important role in model performance. There are dozens of different interfacial friction factor correlations available in the literature, which were developed for specific experimental or field conditions. A previous study discussed the importance of the interfacial friction factor in gas–liquid stratified flow modeling and also emphasized the basic pitfalls in the commonly used single-phase closure relationships [1]. The primary reason behind the inaccuracy was attributed to the empirical nature through which the correlations were developed and the narrow range considered in the correlation development. Using the cor-

relations outside the conditions originally used to develop the model can result in dramatic errors. Another problem is that the model prediction is discontinuous when switching from one correlation to another, which may potentially cause problems in production design. Due to the absence of a correlation built for a wide range of flow conditions, there is a desire for the industry to look for generalized correlations or approaches that can capture all the parameters that affect the pressure gradient and liquid holdup prediction, such as gas and liquid flow rates, inclination angle, fluid properties, pipe diameter, etc.

The wall friction factors in Equations (4) and (5) are commonly calculated using correlations for single-phase flow as functions of the Reynolds number, given in Equations (7) and (8), using hydraulic diameters. The formulae for the liquid and gas hydraulic diameters are given in Equations (9) and (10) by assuming open-channel flow for the liquid phase and closed-duct flow for the gas phase, respectively [3,14].

$$Re_L = \frac{\rho_L v_L d_L}{\mu_L}, \tag{7}$$

$$Re_C = \frac{\rho_C v_C d_C}{\mu_C}, \tag{8}$$

$$d_L = \frac{4A_L}{S_L}, \tag{9}$$

$$d_C = \frac{4A_C}{S_I + S_C}, \tag{10}$$

Experimental data show that the use of the hydraulic diameter and single-phase correlation for the wall friction factor may also result in inaccuracy [15]. This conclusion can be obtained by comparing the experimental measured frictional pressure gradient and the one calculated from correlations, while assuming negligible errors from geometrical and entrainment correlations. The experimental measured frictional pressure gradient can be calculated from Equation (12), which is derived from Equations (1) and (2) by canceling out the interfacial shear stress terms, shown in Equation (11). The frictional pressure gradient from correlations can be obtained from Equation (13), in which the liquid and gas core wall shear stresses are calculated from Equations (4), (5) and (7)–(10). In this study, we used the liquid holdup obtained from experiments to calculate the geometrical parameters using a unified geometrical prediction model developed for all inclination angles [11], with a zero liquid entrainment assumption. The comparison of the two calculated frictional pressure gradients is plotted in Figure 2a for various datasets from [16–19] that cover all inclination angles from 0° to 90° and wide pressure and liquid flow rate ranges, showing unsatisfactory results. The sensitivity study on the single-phase wall friction factor correlation from [20–24] also shows that it does not have a significant impact on the conclusion (Figure 2b).

$$\left.\frac{dp}{dL}\right)_{total} = \left.\frac{dp}{dL}\right)_f + \left.\frac{dp}{dL}\right)_g = \underbrace{\tau_{WL}\frac{S_L}{A_p} + \tau_{WC}\frac{S_C}{A_p}}_{frictional\ term} + \underbrace{[\rho_L H_L + \rho_G(1 - H_L)]g\sin\theta}_{gravitational\ term}, \tag{11}$$

$$\left.\frac{dp}{dL}\right)_f = \left.\frac{dp}{dL}\right)_{total} - \underbrace{[\rho_L H_L + \rho_G(1 - H_L)]g\sin\theta}_{gravitational\ term}, \tag{12}$$

$$\left.\frac{dp}{dL}\right)_f = \tau_{WL}\frac{S_L}{A_p} + \tau_{WC}\frac{S_C}{A_p}, \tag{13}$$

(a)　　　　　　　　　　　　　　　　　　　　　**(b)**
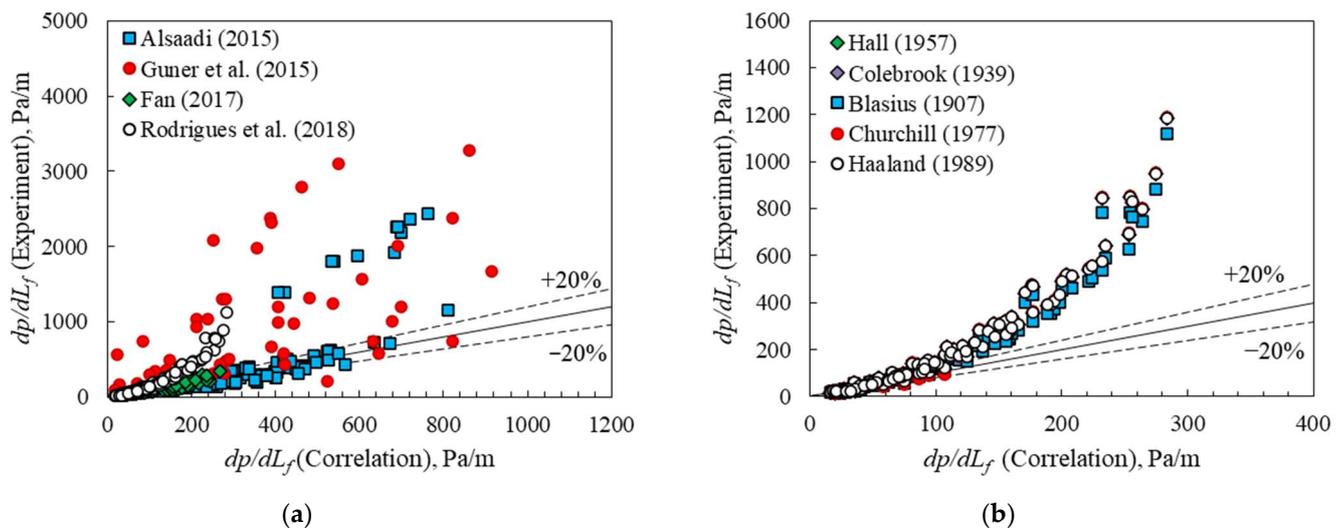
Figure 2. Comparison of experimentally measured frictional pressure gradient with calculated values from correlations for (**a**) different datasets with various flowing conditions and (**b**) various wall friction correlations.

1.2. Application of Machine Learning in Multiphase Flow Modeling

As discussed in the previous section, closure relationships are used to estimate wall shear stresses and interfacial shear stress in a two-fluid model by empirically derived friction factor correlations [25]. Modeling of interfacial shear stress is one of the most critical issues in gas–liquid stratified flow models [1]. This is primarily due to the drag forces acting between the gas and liquid phases, which cause a considerable increase in the pressure gradient. Although mechanistic models are based on the fundamental laws of natural sciences, the closure relationships used in the process have a limitation of being applicable only under the conditions used for model development for acceptable accuracy [26,27].

In the past few years, multiphase flow solutions based on artificial intelligence were also developed and have proven to be a promising solution to model flow behavior. The advantage of using data-driven models lies in the reduced computational cost and a more simplistic approach without any user-induced bias in the form of model assumptions. A primary focus in the application of machine learning solutions in the past has been the prediction of flow patterns using pressure drop signal data as discussed by Mi et al. (2001), Shaban and Tavoularis (2014) [28,29]. Multiple other studies are based on fluid property experimental data using support vector machines, neural networks, or deep learning as a machine learning tool. Some of the related work was conducted by Osman (2004), Li et al. (2014), Ezzatabadipour et al. (2017), Rondon-Guillen et al. (2018), Mask et al. (2019), Hernandez et al. (2019), and Seong et al. (2020) [30–36]. The studies prove machine learning is an effective tool for predicting flow patterns based on pipe characteristics, superficial velocities, and other fluid properties. Mohammadi (2020) proposed an unsupervised learning model to determine clusters for closure relationships and used a genetic algorithm to find a near-optimal set of closure relationships and validate the results using statistics [37]. Lin (2020) incorporated physics by introducing a partial differential equation (PDE) aware deep learning model but was focused on solving the physical phenomenon of the saturation front [38].

Most of the previous studies are predominantly data-driven and did not capture the physics of the multiphase flow behavior, which limits their application range to the condition where the models were trained. In order to address the limitations of previous studies, this paper focuses on developing a machine learning model for friction factors and combining them with a physics-based two-fluid model to predict pressure drop and liquid holdup for segregated flow for a wide range of flow conditions and fluid properties.

2. Methods

In the following section, a hybrid physics-data-driven algorithm is proposed to model segregated flow using machine learning. Specifically, the new approach adopts the framework of a two-fluid model, incorporating the physics of gas and liquid segregated flow, and develops new generalized data-driven correlations for the liquid wall friction factor and interfacial friction factor using machine learning.

2.1. Model Development

Considering the inaccuracy using wall friction factor correlations developed for single-phase pipe flow as discussed previously (Figure 2), we proposed a new liquid wall friction factor (f_L) data-driven model for two-phase segregated flow, which is a function of the liquid wall friction factor for single-phase pipe flow ($f_{L\text{-}SP}$), given in Equation (14). The reason for choosing the liquid wall friction factor instead of the gas wall friction factor is the consideration of annular flow in which gas has zero contact with the pipe wall. The new data-driven model for f_L is applicable to both stratified and annular flows.

$$f_L = \phi f_{L-SP}, \tag{14}$$

With Equations (4)–(6) and (14), the gas and liquid conservation of momentum equations at steady-state becomes:

$$-A_L \left(\frac{dp}{dL}\right)_L - \phi f_{L-SP}\frac{\rho_L v_L |v_L|}{2} S_L + f_I \frac{\rho_C (v_C - v_L)|v_C - v_L|}{2} S_I - A_L \rho_L g \sin\theta = 0 \tag{15}$$

$$-A_G \left(\frac{dp}{dL}\right)_G - f_{G-SP}\frac{\rho_C v_C |v_C|}{2} S_C - f_I \frac{\rho_C (v_C - v_L)|v_C - v_L|}{2} S_I - A_C \rho_C g \sin\theta = 0 \tag{16}$$

where $f_{L\text{-}SP}$ is the liquid wall friction factor that can be calculated from correlations developed for single-phase pipe flow such as [20–24], using hydraulic diameter given in Equation (9); $f_{G\text{-}SP}$ is the gas wall friction calculated from correlations [20–24], using hydraulic diameter given in Equation (10).

In addition to ϕ, we also proposed a new data-driven machine learning model for f_I, that is supposed to capture the complex behaviors under various flowing conditions and fluid properties. ϕ and f_I can be directly calculated from experimental data for training purposes if both the pressure gradient and liquid holdup are known. The formulas to determine ϕ and f_I from experimental data sets are given in Equations (17) and (18), which are derived from Equations (15) and (16).

$$\phi = 2\frac{-A_p \left(\frac{dp}{dL}\right)_L - f_{G-SP}\frac{\rho_C v_C |v_C|}{2} S_C - [\rho_L H_L + \rho_G (1 - H_L)]A_p g \sin\theta}{f_{L-SP}\rho_L v_L |v_L| S_L}, \tag{17}$$

$$f_I = 2A_p \frac{H_L \left(\frac{dp}{dL}\right)_L + \phi f_{L-SP}\frac{\rho_L v_L |v_L|}{2}\frac{S_L}{A_p} + \rho_L H_L g \sin\theta}{\rho_C (v_C - v_L)|v_C - v_L| S_I}, \tag{18}$$

In the current study, a data-driven methodology is proposed to model ϕ and f_I to capture their non-linear relationships with flowing conditions and fluid properties for determining pressure drops and liquid holdups. The data source used in the model development covers a wide range of flowing conditions and fluid properties, discussed in the next section. The back-calculated interfacial friction factor, f_I, and the coefficient, ϕ, from experimental measurement, become the target/response variable, and machine learning can be used to generate a correlation for accurately predicting this response variable while using fluid properties and flowing parameters as input variables. Figure 3 shows the stepwise workflow of the model development using machine-learning (Steps 1–5) and the implementation of the new model to predict pressure gradient and liquid holdup for segregated flow (Step 6).

Step 1. Input Data

Flowing Conditions:
- Pipe inclination angle, θ
- Pipe diameter, d
- Superficial gas velocity, v_{Sg}
- Superficial liquid velocity, v_{SL}

Fluid Properties:
- Density, ρ_L and ρ_G
- Viscosity, μ_L and μ_G
- Surface Tension, σ

Measurement:
- Pressure gradient, dp/dL
- Liquid holdup, H_L

Step 2: Calculate geometrical parameters, S_L, S_C, S_I, using Zhang and Sarica (2011) wetted wall fraction and geometrical correlation, assuming zero liquid entrainment.

Step 3: Estimate liquid wall fraction factor coefficient, ϕ, from experimental data using Equation (17). f_{L-SP} and f_{G-SP} are calculated from Churchill (1977), friction factor correlation developed for single phase pipe flow, as given by:

$$f = 2\left[\left(\frac{8}{Re}\right)^{12} + \frac{1}{(A+B)^{1.5}}\right]^{1/12} \qquad A = \left(2.457\ln\left(\frac{1}{\left(\frac{7}{Re}\right)^{0.9} + 0.27\frac{\varepsilon}{d}}\right)\right)^{16}$$

where Re is determined from Equations (7-10), and ε is the pipe roughness.

$$B = \left(\frac{37530}{Re}\right)^{16}$$

Step 5. Build a Machine Learning (ML) predictive model. The following three algorithms were investigated in this study:
- Random Forest
- Neural Network
- XGBoost

The input variables are the flowing conditions and fluid properties listed in Step 1. The target/Responses are ϕ and f_I.

Step 4. Calculate f_I using Equation (18) from experimental data, where the calculations of f_{L-SP}, S_L, S_I are the same as in Steps 2-3.

Step 6. Implementation of new model for pressure gradient and liquid holdup predictions from best performing model:
- ϕ and f_I are calculated from ML models as developed in Step 5.
- Geometrical parameters, S_L, S_C, S_I, A_L, A_C, are calculated from Zhang and Sarica (2011), assuming zero liquid entrainment (same as Step 2).
- Calculate liquid holdup from the combined momentum equation of gas and liquid phases, given by:

$$\phi f_{L-SP}\frac{\rho_L v_L |v_L|}{2}\frac{S_L}{A_L} - f_{G-SP}\frac{\rho_C v_C |v_C|}{2}\frac{S_C}{A_C} - f_I\frac{\rho_C (v_C - v_L)|v_C - v_L|}{2}\left(\frac{S_I}{A_L} + \frac{S_I}{A_C}\right) + (\rho_L - \rho_C)g\sin\theta = 0$$

- Pressure gradient can be calculated from either gas or liquid momentum equations as in Equations (15) or (16).

Figure 3. Workflow of model development (Steps 1–5) and model implementation (Step 5).

The biggest advantage of the proposed method is that it preserves the physics of two-phase segregated flow while generating more robust and generalized interfacial and liquid wall friction factor correlations.

2.2. Dataset Description

The dataset was extracted from the open literature, including multiple studies where pressure gradient and liquid holdup were both presented, such as [16,18,19,39–48]. The two response variables, the coefficient for liquid wall friction factor, ϕ, and the interfacial friction factor, f_I, were estimated from the experimental data first. These two parameters were correlated with the input variables using three machine learning algorithms—random forest, gradient boosting, and an artificial neural network. The input parameters include liquid and gas superficial velocities (v_{SL} and v_{Sg}), liquid and gas densities and viscosities

(ρ_L, ρ_G, μ_L, and μ_G), surface tension (σ), pipe diameter (d), and inclination angle (θ). The total number of filtered data points is approximately 1500.

Exploratory Data Analysis (EDA) is a critical step for understanding quantitative variables in an experimental dataset and is often used as a visual tool to understand the high-dimensional dataset. This step plays a critical role in learning from the data during the scientific process of model building and testing and discovering patterns in the dataset. In this section, the dataset is analyzed in a univariate and multivariate graphical sense by looking at statistical distributions in the form of histograms, scatter plots, and box plots. Data pre-processing includes the process of data wrangling by identifying the irregularities, cleaning the data, identifying, and removing anomalies and outliers as well as data transformation and feature selection.

There are four basic questions about the data that need to be asked before starting with data analysis. The first question relates to whether the data are discrete or continuous. The second step involves looking at the symmetry using distributions to identify if there is skewness present in the dataset. The third step is related to the upper and lower bounds of the data, and the final question determines the likelihood of observing extreme values in the distribution. The data used in the study are primarily of continuous form with values in a finite interval. Figure 4 shows the univariate distribution of all the variables that form the dataset. This figure also helps to determine the wide range of each variable, which is essential to create a model with high generalization capability. As an example, it can be observed from the inclination distribution plot that a majority of studies have been conducted on horizontal pipes.
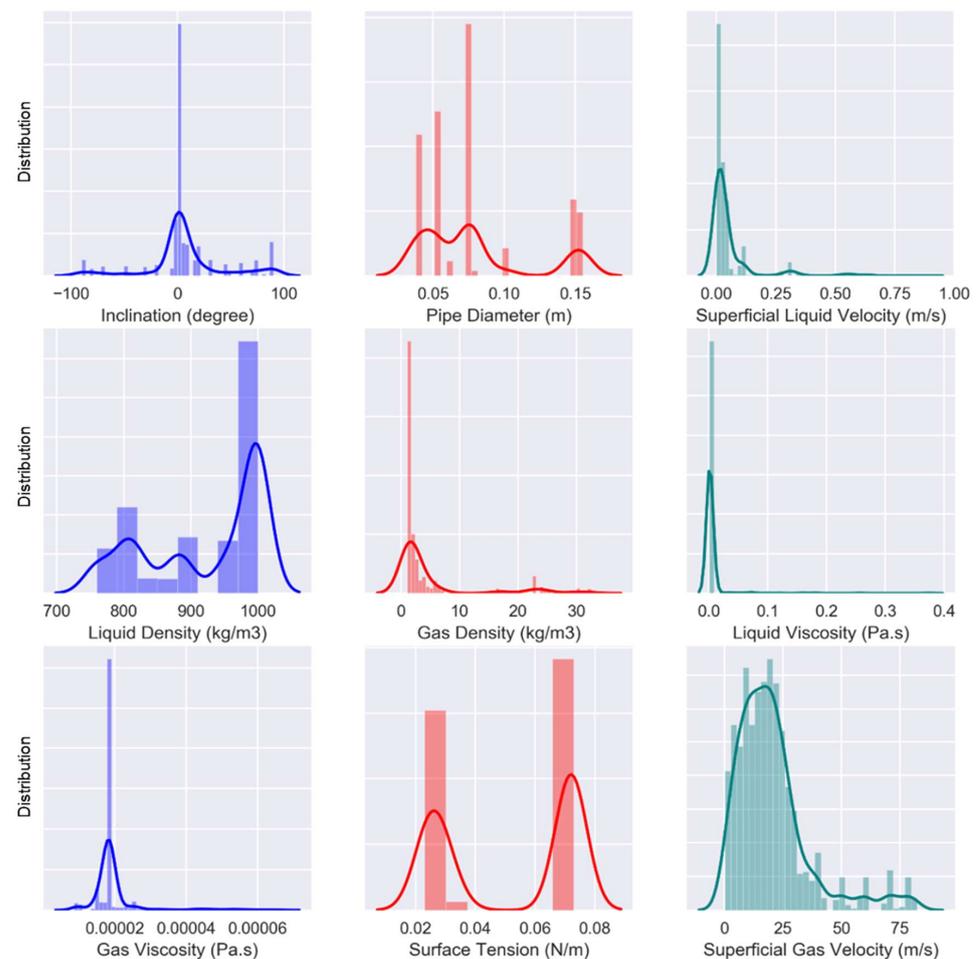
Figure 4. Univariate distribution plot for the entire dataset.

A multivariable bivariate distribution leads to an identification of significant skewness in liquid density, gas density, liquid viscosity, and gas viscosity, but it is not necessarily an outlier, rather another subset of data points with lower frequency.

Box plots are often used as a graphical tool to display information for continuous univariate data and to identify highly influential points or possible outliers. The classical method from Tukey (1977) [49] was used to remove outliers by using a threshold of 1.5×interquartile range (IQR), which is defined as the difference between the first and third quartile [50]. Predictions for ϕ and fI using the two-fluid model were analyzed to identify and remove outliers. The results for the case of ϕ and fI before and after the outlier removal are shown in Figure 5.
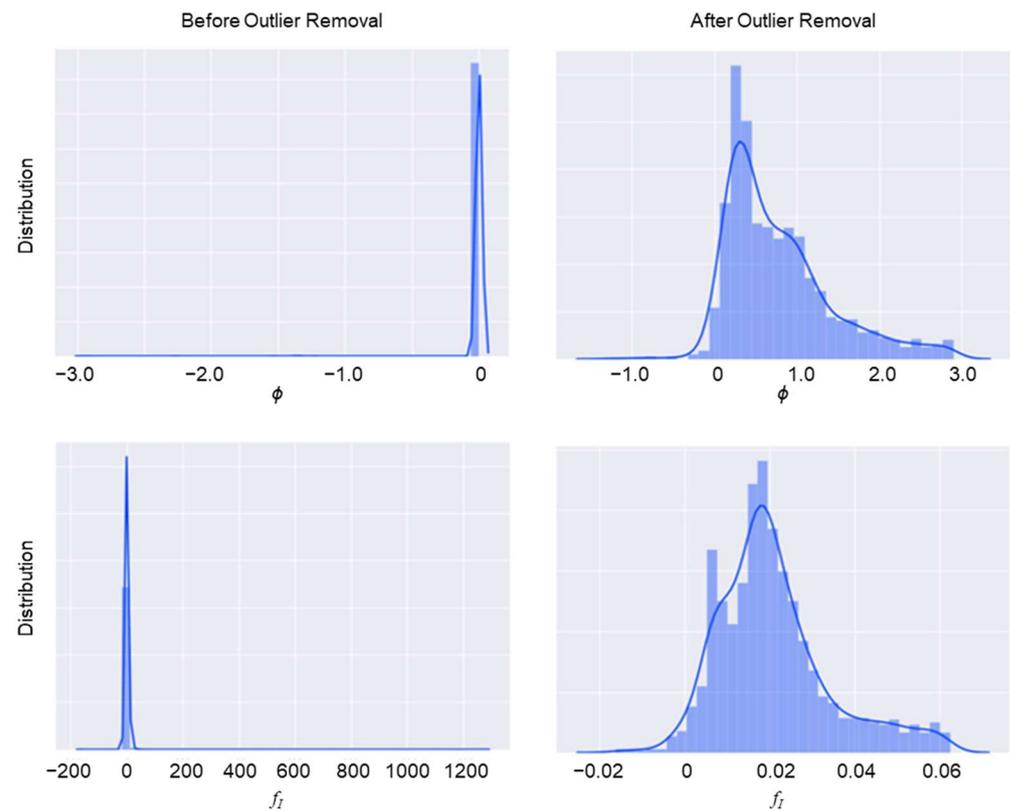
Figure 5. Comparison of the distribution for ϕ (top) and f_I (bottom) before (left column) and after (right column) the outliers were removed.

2.3. Machine Learning Algorithms

Three different machine learning algorithms are implemented for the dataset—Random Forest [51], eXtreme Gradient Boosting—XGBoost [52], and Artificial Neural Networks [53]. The objective is to generate a predictive model for ϕ and f_I with the highest accuracy. Although these three algorithms could be used for both regression and classification problems, in this work the focus will be supervised regression since the dataset has continuous variables, with no categorical information. A primary reason for the selection of these algorithms is their high performance to account for the non-linearity associated with the predictor variables or features [54]. A detailed explanation of each algorithm is readily available in the literature and is omitted from this paper.

2.4. Model Evaluation—Accuracy Metrics

The model performance is comprehensively evaluated with other existing methods available in the literature to show improvement while predicting pressure gradient and liquid holdup. To quantify the performance of different regression models, accuracy metrics such as the Root Mean Squared Error (RMSE), Absolute Average Error (AAE), and R^2 are

commonly used. The mathematical explanation of some of these error metrics is provided below. The error for each prediction is defined as the difference between the measured and predicted values.

Root Mean Squared Error (RMSE)—This metric is defined as the square root of the average square distance between the measured value and the prediction. It represents the sample standard deviation of the residuals, which is a measure of how concentrated the data are around the line of best fit. It can be calculated by Equation (19).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}, \tag{19}$$

Mean Absolute Error (MAE)—The metric is the absolute of the difference between the predicted value and observed value and provides a linear score where all the individual differences are weighted equally in the average. The metric can be calculated by Equation (20).

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|, \tag{20}$$

R^2—Defined as the coefficient of determination, R^2 provides the goodness of fit of a model in the form of an intuitive scale varying between 0 and 1. Another way to define the metric is with respect to the proportion of the explainable variance in the dependent variable with the independent variable(s). A lower R^2 indicates a low level of correlation while a score of 1 indicates a perfect correlation between the two variables with no variance. Equation (21) can be used to determine the metric.

$$R^2 = \frac{\sum(\hat{y}_j - \bar{y})^2}{\sum(y_j - \bar{y})^2}, \tag{21}$$

Average Absolute Percentage Relative Error (AAPRE): This metric can be derived from the absolute error and is an indication of how good a prediction is with respect to the observed value. Mathematically the metric can be calculated from Equation (22) below.

$$\text{AAPRE} = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{y_j - \hat{y}_j}{y_j} \times 100\right|, \tag{22}$$

In the equations above, y_j represents the experimentally measured value for the jth data point; $\hat{y}_j$ represents the predicted value from the machine learning model; and $\bar{y}$ denotes the mean of the dataset, with n being the total number of data points.

3. Results

Predictive modeling results for different regression models along with different accuracy metrics are discussed in this section. Table 1 shows the results in a tabulated format, broken down into the base model and random search cross-validated results, for the three machine learning algorithms analyzed in this work.

Based on the results for different machine learning models, XGBoost performed better when average performance was evaluated between predictions for both ϕ and f_I, in terms of RMSE for the test dataset, and was selected as the final model. The learning curve after hyper-parameter tuning for interfacial friction factor is shown in Figure 6. An 'early stopping' approach is used to prevent overfitting. The process enables the input of the number of iterations (epochs) after which the algorithm stops if the validation score does not increase. Cross plots between the known value and its corresponding prediction are also shown in Figure 7. Hyper-parameter tuning was carried out for each machine learning algorithm, and predictions for ϕ and f_I were made.

Table 1. Comparison of the accuracy metric for different machine learning models.

Model and Response Variables		Data	MAE	R^2	RMSE	Data	R^2	RMSE
RF (ϕ)	Base Model	Train	0.076	0.949	0.129	Entire Data	0.868	0.207
		Test	0.215	0.620	0.348			
	Random Search CV	Train	0.052	0.972	0.094	Entire Data	0.898	0.182
		Test	0.192	0.669	0.325			
RF (f_I)	Base Model	Train	0.0015	0.949	0.0028	Entire Data	0.885	0.0043
		Test	0.0037	0.704	0.0069			
	Random Search CV	Train	0.0010	0.972	0.0020	Entire Data	0.901	0.0039
		Test	0.0037	0.703	0.0070			
XGB (ϕ)	Base Model	Train	0.207	0.773	0.304	Entire Data	0.741	0.3270
		Test	0.258	0.651	0.386			
	Random Search CV	Train	0.072	0.974	0.101	Entire Data	0.910	0.1920
		Test	0.201	0.729	0.340			
XGB (f_I)	Base Model	Train	0.0039	0.800	0.0057	Entire Data	0.767	0.0063
		Test	0.0050	0.686	0.0079			
	Random Search CV	Train	0.0013	0.979	0.0018	Entire Data	0.919	0.0037
		Test	0.0038	0.776	0.0067			
ANN (ϕ)	Base Model	Train	0.302	0.425	0.436	Entire Data	0.416	0.436
		Test	0.295	0.380	0.439			
	Grid Search CV	Train	0.280	0.448	0.427	Entire Data	0.439	0.428
		Test	0.284	0.400	0.432			
ANN (f_I)	Base Model	Train	0.0048	0.665	0.0073	Entire Data	0.648	0.0074
		Test	0.0052	0.593	0.0076			
	Grid Search CV	Train	0.0044	0.698	0.0069	Entire Data	0.683	0.007
		Test	0.0047	0.631	0.0073			
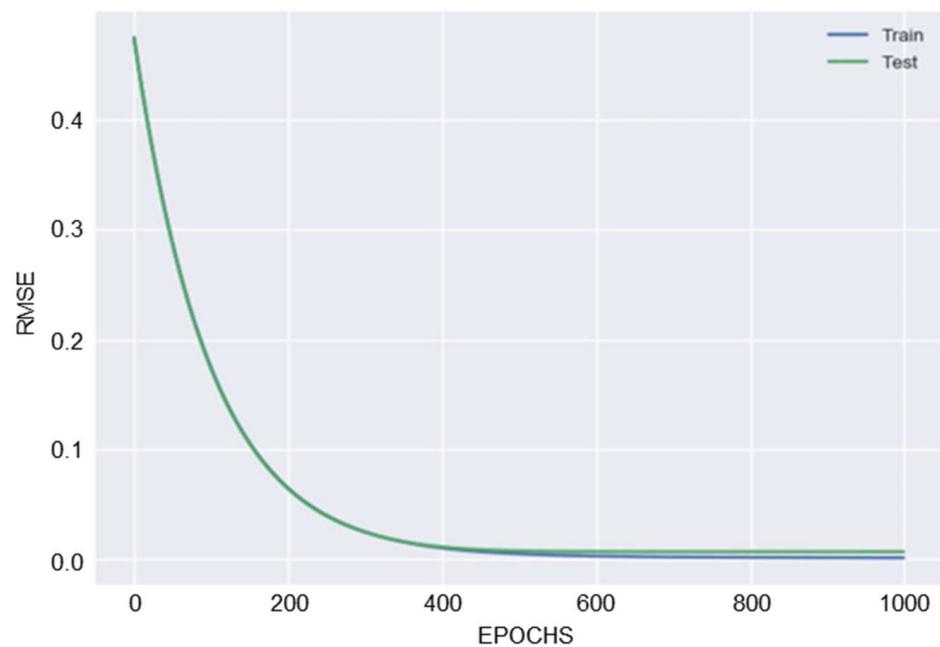
Figure 6. Learning curve for the XGBoost model with respect to RMSE for 1000 epochs using Randomized Search CV from scikit-learn.
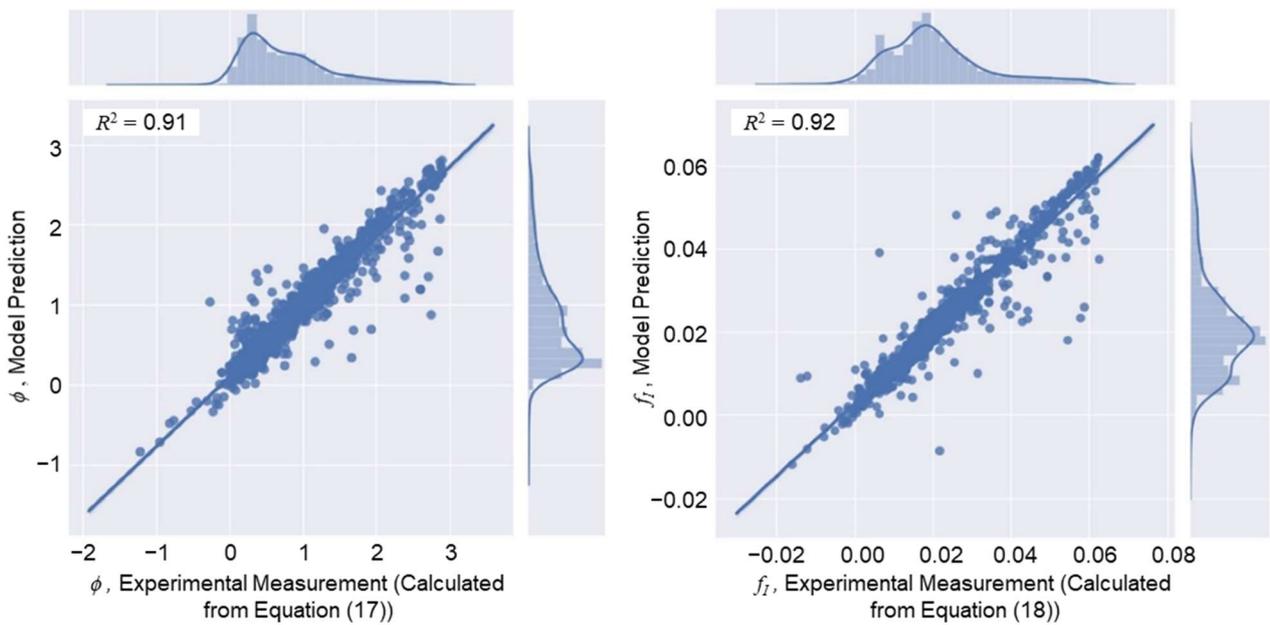
Figure 7. Prediction performance using XGBoost as an algorithm for ϕ and f_I.

Once the predictions are made for the dataset, the error can be calculated between the experimental and predicted values.

XGBoost also provides the capability to determine the feature importance, which plays a critical role to identify predictors with the highest impact on the model. Figure 8 shows the results of feature importance for the current dataset. The metric shown in the figure is the 'F-score'—weight, which is determined simply by how many times each feature is used to split the data across all trees.
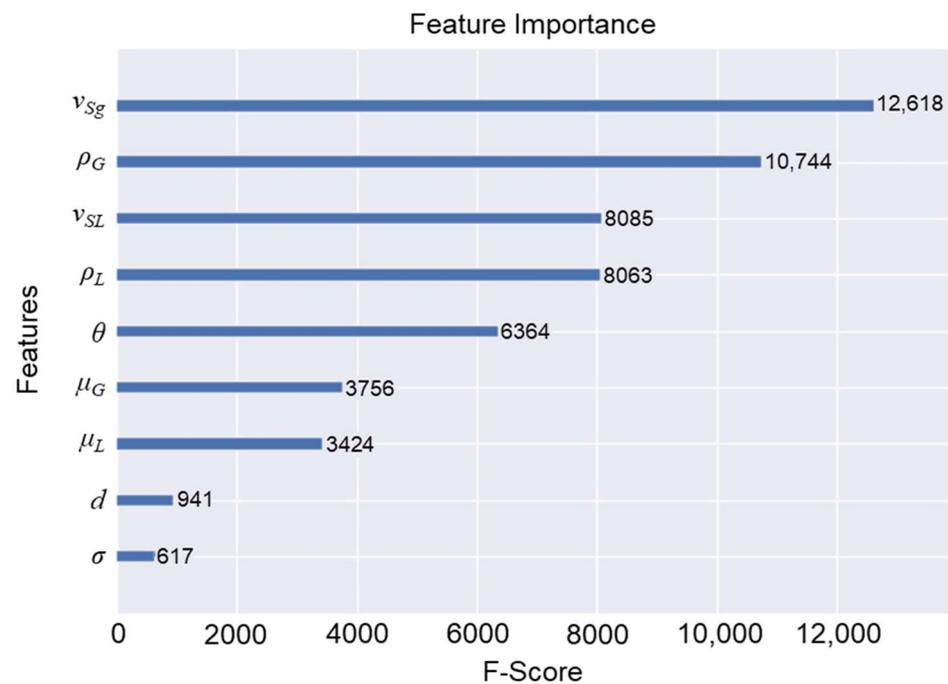
Figure 8. Feature importance of predictors determined using XGBoost.

The next step is to evaluate the model in terms of the pressure gradient and liquid holdup predictions, using the physics-based two-fluid model, as shown in the workflow in Step 6 of Figure 3. Figures 9 and 10 show the comparison between experimentally

determined and predictions from the proposed model for pressure drop and liquid holdup, respectively, for the entire dataset on the left and each individual data source on the right. In general, the model gives good predictions for all data sets.
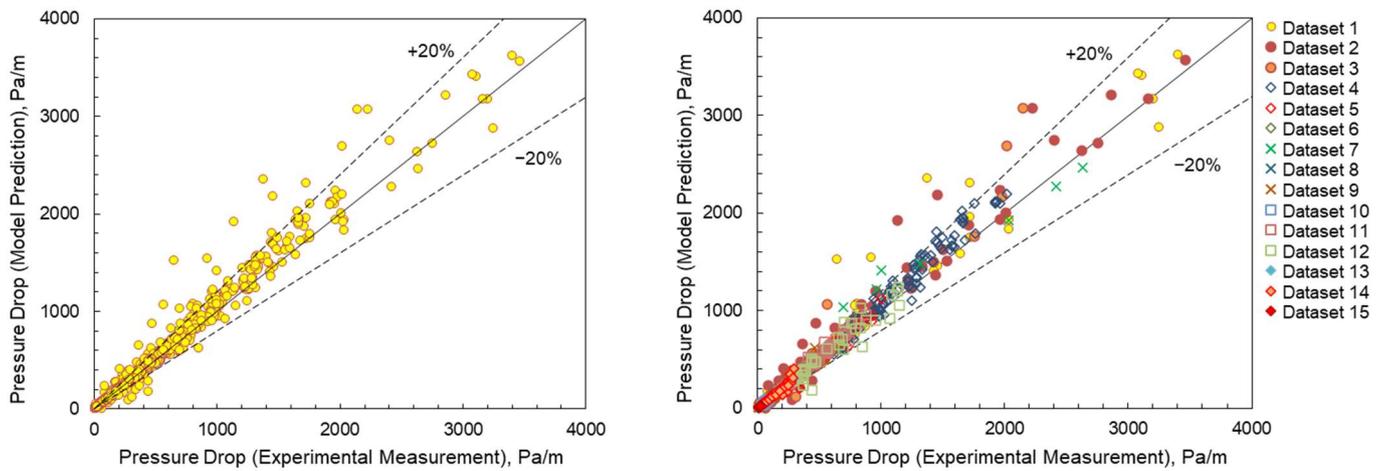
Figure 9. Comparison of experimental measurement and predictions for the pressure gradient.
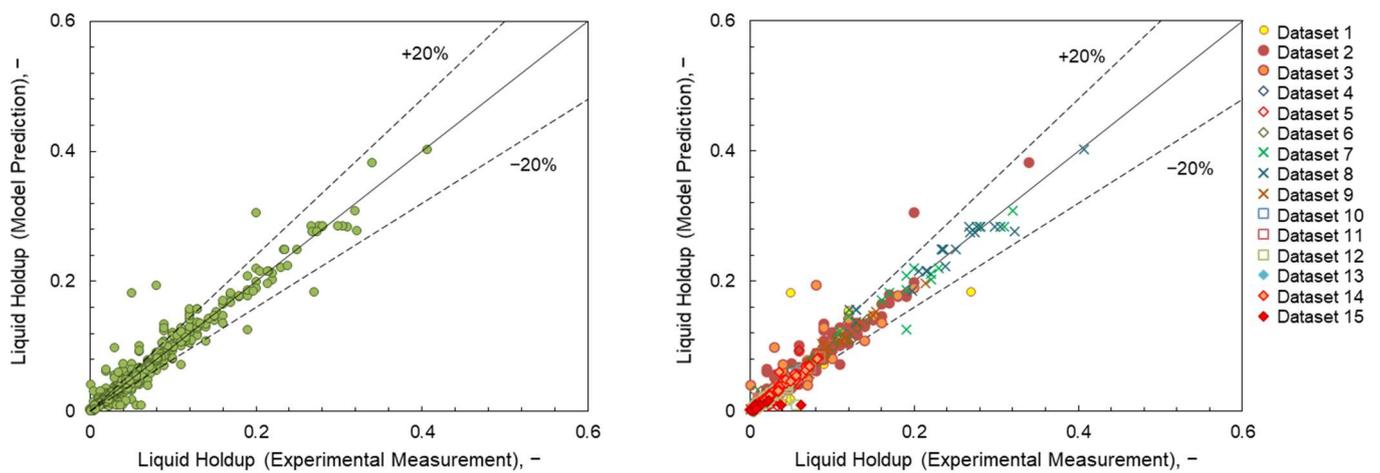
Figure 10. Comparison of experimental measurement and predictions for liquid holdup.

The results were also compared with other existing models in the literature. The primary equations of these models are based on the two-fluid model but incorporated with different correlations for geometrical parameters and interfacial friction factor (f_I). Two wetted wall fraction correlations tested were Zhang and Sarica (2011) [11], and Taitel and Dukler (1976) [3]. Churchill (1977) was used to determine wall friction factors [21]. In total, 19 different interfacial friction factor correlations from the literature were tested, including [3,5,7,8,12,55–64], which led to a total of 38 different combinations of wetted wall friction and interfacial friction factor correlations. The list was divided into different cases, referred to in Figures 11 and 12, and is included in the Appendix A.

Model performance is gauged through metrics such as the average absolute relative error (AAPRE) and root mean squared error (RMSE). Figure 11 shows the performance in the form of bar charts for both metrics during the prediction of the pressure gradient. The comparison shows that the new model has the lowest average absolute relative error and RMSE.

Predictions are also made for liquid holdup, and comparisons of results using the same metric are shown in Figure 12. Based on the error metrics, the new model proposed in this work provides the lowest error in comparison to two-fluid models that use empirically derived interfacial friction factor correlations.
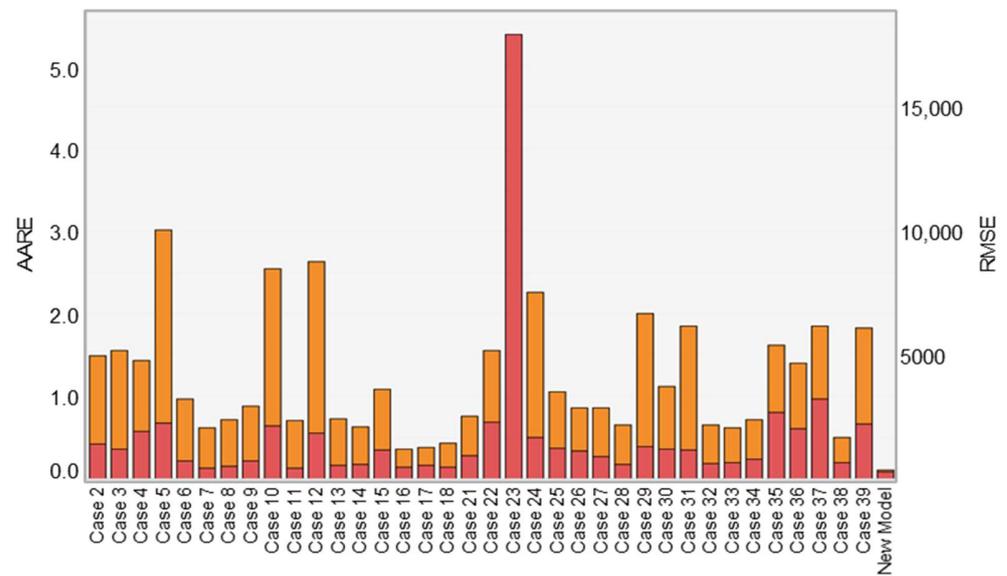
Figure 11. Model evaluation for pressure gradient predictions in terms of average absolute relative error and RMSE.
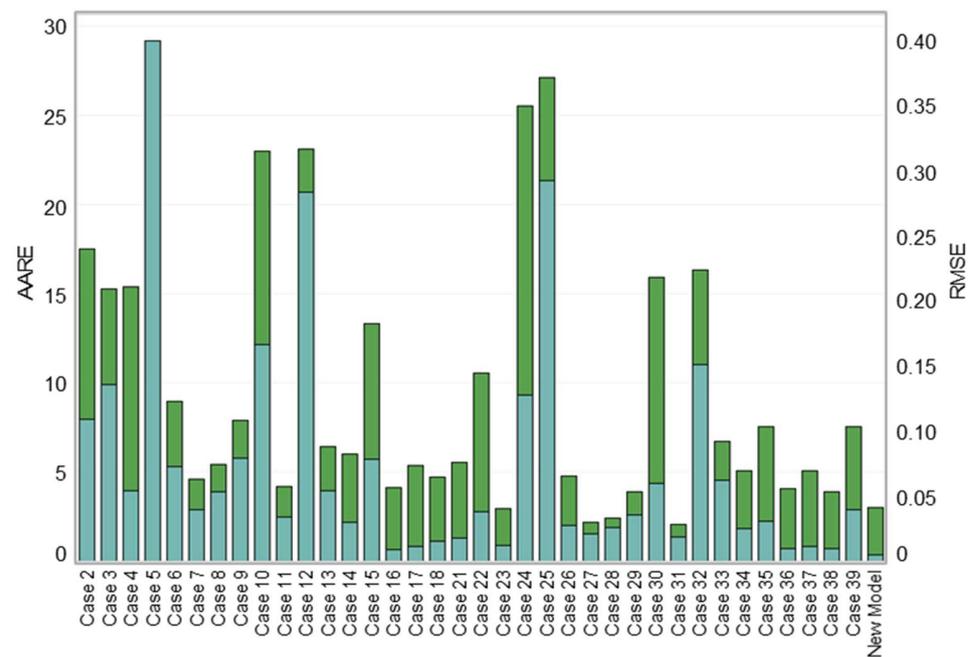
Figure 12. Model evaluation for liquid holdup predictions in terms of average absolute relative error and RMSE.

In addition to liquid holdup and pressure gradient predictions, the new model developed from this study was incorporated into a state-of-art mechanistic model for onset of liquid accumulation prediction proposed in [65]. That model is based on liquid film reversal of segregated flow and requires pressure gradient and liquid holdup prediction for segregated flow. The critical gas velocity determined from the coupled mechanistic model and the proposed ML algorithm from this study was further evaluated against existing droplet and film reversal models from [25,66–74], as listed in Table 2. It is worth mentioning that the critical gas velocity refers to the minimum gas superficial velocity that maintains segregated flow in upward inclined pipes. The experimental data used in model

evaluation were extracted from [16,17,19,48,64,65,74–76]. Details of the mechanistic model and experimental datasets are well described in [65] and were omitted from this paper.

Table 2. Start-of-the-art droplet and film reversal-based modeling studies for critical gas velocity estimation.

No.	Models Based on Liquid Droplet Removal	No.	Models Based on Liquid Film Dynamics/Reversal
1	Turner et al. (1969)	7	Barnea (1986)
2	Coleman et al. (1991)	8	Luo et al. (2014)
3	Li et al. (2002)	9	Shekhar et al. (2017)
4	Wang and Liu (2007)	10	Fan et al. (2018)
5	Belfroid et al. (2008)	11	New Model (Rastogi and Fan (2019) with new ML algorithm from this study)
6	Zhou and Yuan (2010)		

As expected, the liquid film models provide a better prediction than the liquid droplet models, which is consistent with the findings from previous studies [65,74]. Overall, the new model gives better predictions for critical gas velocity compared to other models evaluated in this study. This is illustrated with the lowest AARE score of 12.78 units, shown in Figure 13 between the existing models and the proposed model in this paper.
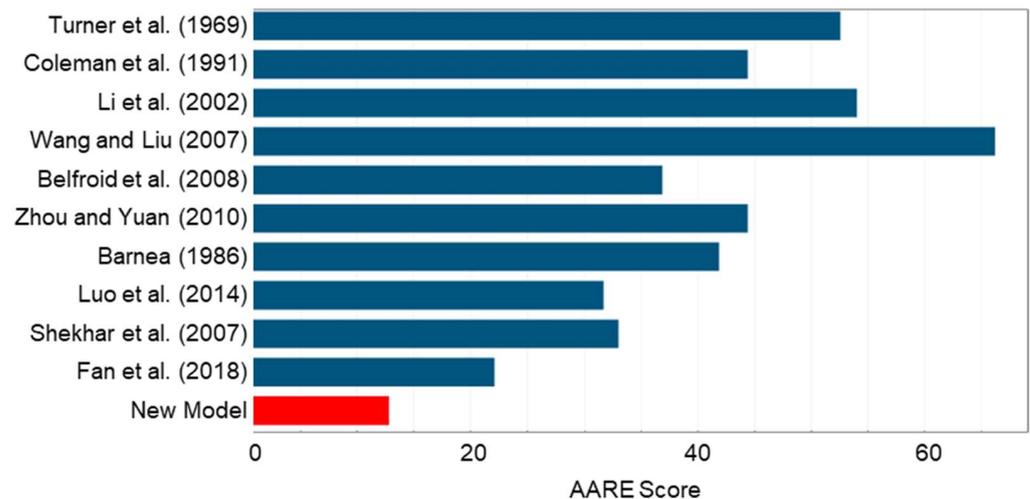
Figure 13. Average absolute relative error for different model predictions of critical gas velocity.

4. Summary

Segregated flow is one of the most commonly encountered flow patterns in several practical applications such as chemical, nuclear, refrigeration, and oil and gas industries. Accurate prediction of its pressure gradient and the liquid holdup is an important consideration for facility design and operations. It is also crucial for predicting the onset of liquid accumulation in the petroleum industry, which needs to be prevented and carefully managed during oil and gas production. It has been found from previous studies that the interfacial friction factor plays a crucial role in two-fluid model performance. The previous approach is to use empirically derived correlations that are only applicable to the small subset of working conditions under which those solutions are derived.

Using predictions from a mechanistic model augmented with data-driven analysis, a novel and more generalized approach is proposed that combines a physics-based two-fluid model and machine learning algorithms to accurately predict pressure drop and liquid holdup. The new model is comprehensively evaluated with existing models with various closure relationships and shows a significant improvement in terms of prediction accuracy. A comparison of the average absolute percentage error was also made for critical gas

velocity between the experimentally determined values and model predictions, which include both droplet and film-reversal based models. The hybrid model gives the best prediction for the critical gas velocity compared with other existing modeling approaches.

The proposed modeling workflow helps to reduce the dependence on empirically derived correlations and the need for interfacial friction factor correlation selection. Considering the gradual transition to the digital oil field, more field data will be available in the future, which can be used to augment the physics-based mechanistic multiphase flow modeling. The proposed approach in this study will add significant value in segregated flow modeling in terms of combining physics-based and data-driven models, and therefore the optimal design of production systems.

5. Discussion

A hybrid machine learning and mechanistic model to determine the pressure gradient and liquid holdup, as well as the critical gas velocity, presented in this work, provides better accuracy in comparison to studies derived from literature, which are primarily mechanistic or empirical-based. This has been made possible by leveraging the predictive capabilities of a machine learning algorithm along with the causal understanding of fluid flow mechanisms through a mechanistic model. However, this conclusion is accompanied by a caveat that significant, good-quality data are available before the implementation of any machine learning model. It is important to understand that machine learning and mechanistic modeling are two different paradigms and should never be considered to compete with or replace one other.

According to Baker (2018) [77], if large-scale datasets are available, machine learning algorithms can be an efficient and scalable modeling approach, as well as provide the ability to avoid the need to understand complex mechanisms. Advanced algorithms in the field of machine learning have been proven to identify hidden correlations between parameters that are not easy to identify through conventional approaches. One of the biggest drawbacks of a pure data-driven model is the applicability and generalization of using the model outside the boundary conditions in which the model was trained. This is where a mechanistic model has an advantage since the predictions from a mechanistic model, based on fundamental laws of nature, can be applied under conditions where experiments are either costly or difficult to perform. As discussed earlier, a mechanistic model relies on the generation of novel hypotheses and is usually built to mimic real-life events with certain assumptions. In some cases, these oversimplified assumptions and their extreme specific nature, commonly found in empirical models, prevent the model predictions to be applicable on a wider scale. This paper shows how a symbiotic relationship can be created between machine learning and mechanistic modeling approaches by harnessing the positive aspects of both approaches.

Nomenclature

A_C	Gas core occupied cross-sectional area, m^2
A_L	Liquid film occupied cross-sectional area, m^2
A_p	Pipe cross-sectional area, m^2
d	Pipe inner diameter, m
d_L	Liquid hydraulic diameter, m
d_C	Gas core hydraulic diameter, m
$(dp/dL)_{total}$	Total pressure gradient, Pa/m
$(dp/dL)_L$	Liquid phase pressure gradient, Pa/m
$(dp/dL)_C$	Gas core pressure gradient, Pa/m
$(dp/dL)_f$	Frictional component of pressure gradient, Pa/m
$(dp/dL)_g$	Gravitational component of pressure gradient, Pa/m
f_C	Gas wall friction factor, -
$f_{G\text{-}SP}$	Gas wall friction factor for single phase pipe flow, -
f_I	Interfacial friction factor, -
f_L	Liquid wall friction factor, -
$f_{L\text{-}SP}$	Liquid wall friction factor for single phase pipe flow, -
F_E	Entrainment fraction, -
g	Gravitational acceleration, m/s^2
h_L	Liquid film thickness, m
H_L	Total liquid holdup, -
Re	Reynolds number, -
S_C	Gas wetted perimeter, m
S_I	Interfacial perimeter, m
S_L	Liquid wetted perimeter, m
v_L	Average liquid velocity, m/s
v_C	Average gas core velocity, m/s
v_{Sg}	Gas superficial velocity, m/s
v_{SL}	Liquid superficial velocity, m/s
ε	Pipe roughness, m
θ	Inclination angle from horizontal, ° (degree)
μ_G	Gas viscosity, Pa·s
μ_L	Liquid viscosity, Pa·s
ρ_C	Gas core density, kg/m^3
ρ_g	Gas density, kg/m^3
ρ_L	Liquid density, kg/m^3
σ	Surface tension, N/m
τ_{WC}	Gas core wall shear stress, Pa
τ_I	Interfacial shear stress, Pa
τ_{WL}	Average liquid wall shear stress, Pa
ϕ	Liquid wall friction factor coefficient, -
AAE	Absolute Average Error
AAPRE	Average Absolute Percent Relative Error
ANN	Artificial Neural Network
CV	Cross-Validation
EDA	Exploratory Data Analysis
IQR	Inter-Quartile Range
MAE	Mean Absolute Error
ML	Machine Learning
PDE	Partial Differential Equation
RF	Random Forest
RMSE	Root Mean Squared Error
XGBoost	eXtreme Gradient Boosting

Appendix A

Table indicating case number and the corresponding combination of geometrical parameters, wall friction factor, and interfacial friction factor used for model validation.

Table A1. Case number and the corresponding correlations used for model validation in Figures 11 and 12.

Case #	Geometrical Parameter	Wall Friction Factor, f_L	Interfacial Friction Factor, f_I
New Model	Zhang and Sarica (2011)	New ML Model from Current Study	New ML Model from Current Study
Case 02	Taitel and Dukler (1976)	Churchill (1977)	Cohen and Hanratty (1968)
Case 03	Taitel and Dukler (1976)	Churchill (1977)	Hart et al. (1989)
Case 04	Taitel and Dukler (1976)	Churchill (1977)	Kowalski (1987)
Case 05	Taitel and Dukler (1976)	Churchill (1977)	Taitel and Dukler (1976)
Case 06	Taitel and Dukler (1976)	Churchill (1977)	Vlachos et al. (1997)
Case 07	Taitel and Dukler (1976)	Churchill (1977)	Wallis (1969)
Case 08	Taitel and Dukler (1976)	Churchill (1977)	Wallis Modified (1969)
Case 09	Taitel and Dukler (1976)	Churchill (1977)	Whalley and Hewitt (1978)
Case 10	Taitel and Dukler (1976)	Churchill (1977)	Oliemans et al. (1986)
Case 11	Taitel and Dukler (1976)	Churchill (1977)	Fore et al. (2000)
Case 12	Taitel and Dukler (1976)	Churchill (1977)	Dallman et al. (1979)
Case 13	Taitel and Dukler (1976)	Churchill (1977)	Ambrosini et al. (1991)
Case 14	Taitel and Dukler (1976)	Churchill (1977)	Hamersma and Hart (1987)
Case 15	Taitel and Dukler (1976)	Churchill (1977)	Chen et al. (1997)
Case 16	Taitel and Dukler (1976)	Churchill (1977)	Andritsos and Hanratty (1987)
Case 17	Taitel and Dukler (1976)	Churchill (1977)	Andritsos et al. (2008)
Case 18	Taitel and Dukler (1976)	Churchill (1977)	Zhang et al. (2003)
Case 19	Taitel and Dukler (1976)	Churchill (1977)	Brito (2015)
Case 20	Taitel and Dukler (1976)	Churchill (1977)	Grolman and Fortuin (1997)
Case 21	Zhang and Sarica (2011)	Churchill (1977)	Cohen and Hanratty (1968)
Case 22	Zhang and Sarica (2011)	Churchill (1977)	Hart et al. (1989)
Case 23	Zhang and Sarica (2011)	Churchill (1977)	Kowalski (1985)
Case 24	Zhang and Sarica (2011)	Churchill (1977)	Taitel and Dukler (1976)
Case 25	Zhang and Sarica (2011)	Churchill (1977)	Vlachos et al. (1997)
Case 26	Zhang and Sarica (2011)	Churchill (1977)	Wallis (1997)
Case 27	Zhang and Sarica (2011)	Churchill (1977)	Wallis Modified (1969)
Case 28	Zhang and Sarica (2011)	Churchill (1977)	Whalley and Hewitt (1978)
Case 29	Zhang and Sarica (2011)	Churchill (1977)	Oliemans et al. (1986)
Case 30	Zhang and Sarica (2011)	Churchill (1977)	Fore et al. (2000)
Case 31	Zhang and Sarica (2011)	Churchill (1977)	Dallman et al. (1979)
Case 32	Zhang and Sarica (2011)	Churchill (1977)	Ambrosini et al. (1991)
Case 33	Zhang and Sarica (2011)	Churchill (1977)	Hamersma and Hart (1987)
Case 34	Zhang and Sarica (2011)	Churchill (1977)	Chen et al. (1997)
Case 35	Zhang and Sarica (2011)	Churchill (1977)	Andritsos and Haratty (1987)
Case 36	Zhang and Sarica (2011)	Churchill (1977)	Andritsos et al. (2008)
Case 37	Zhang and Sarica (2011)	Churchill (1977)	Zhang et al. (2003)
Case 38	Zhang and Sarica (2011)	Churchill (1977)	Brito (2015)
Case 39	Zhang and Sarica (2011)	Churchill (1977)	Grolman and Fortuin (1997)

References

1. Ullmann, A.; Brauner, N. Closure Relations for Two-fluid Models for Two-phase Stratified Smooth and Stratified Wavy Flows. *Int. J. Multiph. Flow* **2006**, *32*, 82–105. [CrossRef]
2. Shoham, O. *Mechanistic Modeling of Gas-Liquid Two-Phase Flow in Pipes*; Society of Petroleum Engineers: Houston, TX, USA, 2006.
3. Taitel, Y.; Dukler, A.E. A Model for Predicting Flow Regime Transitions in Horizontal and Near Horizontal Gas-liquid Flow. *AIChE J.* **1976**, *22*, 47–55. [CrossRef]
4. Shoham, O.; Taitel, Y. Stratified Turbulent-turbulent Gas-liquid Flow in Horizontal and Inclined Pipes. *AIChE J.* **1984**, *30*, 377–385. [CrossRef]
5. Kowalski, J.E. Wall and Interfacial Shear Stress in Stratified Flow in a Horizontal Pipe. *AIChE J.* **1987**, *33*, 274–281. [CrossRef]
6. Andritsos, N.; Hanratty, T. Interfacial Instabilities for Horizontal Gas-liquid Flows in Pipelines. *Int. J. Multiph. Flow* **1987**, *13*, 583–603. [CrossRef]
7. Andritsos, N.; Hanratty, T.J. Influence of Interfacial Waves in Stratified Gas-liquid Flows. *AIChE J.* **1987**, *33*, 444–454. [CrossRef]

8. Hart, J.; Hamersma, P.; Fortuin, J. Correlations Predicting Frictional Pressure Drop and Liquid Holdup during Horizontal Gas-liquid Pipe Flow with a Small Liquid Holdup. *Int. J. Multiph. Flow* **1989**, *15*, 947–964. [CrossRef]
9. Hamersma, P.; Hart, J. A Pressure Drop Correlation for Gas/liquid Pipe Flow with a Small Liquid Holdup. *Chem. Eng. Sci.* **1987**, *42*, 1187–1196. [CrossRef]
10. Chen, X.T.; Cal, X.D.; Brill, J.P. Gas-Liquid Stratified-Wavy Flow in Horizontal Pipelines. *J. Energy Resour. Technol.* **1997**, *119*, 209–216. [CrossRef]
11. Zhang, H.-Q.; Sarica, C. A Model for Wetted-Wall Fraction and Gravity Center of Liquid Film in Gas/Liquid Pipe Flow. *SPE J.* **2011**, *16*, 692–697. [CrossRef]
12. Oliemans, R.; Pots, B.; Trompé, N. Modelling of Annular Dispersed Two-phase Flow in Vertical Pipes. *Int. J. Multiph. Flow* **1986**, *12*, 711–732. [CrossRef]
13. Alves, I.; Caetano, E.; Minami, K.; Shoham, O. Modeling Annular Flow Behavior for Gas Wells. *SPE Prod. Eng.* **1991**, *6*, 435–440. [CrossRef]
14. Gazley, C. Interfacial Shear and Stability in Two-Phase Flow. Ph.D. Thesis, University of Delaware, Newark, DE, USA, 1948.
15. Rastogi, A. Onset of Liquid Loading in Large Diameter Inclined Pipes. Ph.D. Thesis, Colorado School of Mines, Golden, CO, USA, 2020.
16. Alsaadi, Y.; Pereyra, E.; Torres, C.; Sarica, C. Liquid loading of highly deviated gas wells from 60° to 88°. In Proceedings of the SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 28–30 September 2015. [CrossRef]
17. Guner, M.; Pereyra, E.; Sarica, C.; Torres, C.F. An Experimental Study of Low Liquid Loading in Inclined Pipes from 90° to 45°. In Proceedings of the SPE Production and Operations Symposium, Oklahoma City, OK, USA, 1–5 March 2015. [CrossRef]
18. Fan, Y. A Study of the Onset of Liquid Accumulation and Pseudo-Slug Flow Chracterization. Ph.D. Thesis, University of Tulsa, Tulsa, OK, USA, 2017.
19. Rodrigues, H.T.; Pereyra, E.; Sarica, C. Pressure Effects on Low-Liquid-Loading Oil/Gas Flow in Slightly Upward Inclined Pipes: Flow Pattern, Pressure Gradient, and Liquid Holdup. *SPE J.* **2019**, *24*, 2221–2238. [CrossRef]
20. Blasius, H. Grenzschichten in Flüssigkeiten mit kleiner Reibung. *Zeits. Math. Phys.* **1907**, *56*, 1–37.
21. Churchill, S.W. Friction-Factor Equation Spans All Fluid-Flow Regimes. *Chem. Eng. J.* **1977**, *84*, 91–92.
22. Hall, N.A. *Thermodynamics of Fluid Flow*; Longmans Green: New York, NY, USA, 1957.
23. Massey, B.S. *Mechanics of Fluids*; Chapman and Hall: Boston, MA, USA, 1989.
24. Colebrook, C.F. Turbulent Flow in Pipes, with Particular Reference to the Transition Region between the Smooth and Rough Pipe Laws. *J. Inst. Civ. Eng.* **1939**, *11*, 133–156. [CrossRef]
25. Barnea, D. Transition from Annular Flow and from Dispersed Bubble Flow—Unified Models for the Whole Range of Pipe Inclinations. *Int. J. Multiph. Flow* **1986**, *12*, 733–744. [CrossRef]
26. Kanin, E.; Osiptsov, A.; Vainshtein, A.; Burnaev, E. A Predictive Model for Steady-state Multiphase Pipe Flow: Machine Learning on Lab Data. *J. Pet. Sci. Eng.* **2019**, *180*, 727–746. [CrossRef]
27. Buist, J.; Sanderse, B.; Van Halder, Y.; Koren, B.; Van Heijst, G. Machine Learning for Closure Models in Multiphase Flow Applications. In Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP 2019), Crete, Greece, 24–26 June 2019. [CrossRef]
28. Mi, Y.; Ishii, M.; Tsoukalas, L. Flow Regime Ddentification Methodology with Neural Networks and Two-phase Flow Models. *Nucl. Eng. Des.* **2001**, *204*, 87–100. [CrossRef]
29. Shaban, H.; Tavoularis, S. Measurement of Gas and Liquid Flow Rates in Two-phase Pipe Flows by the Application of Machine Learning Techniques to Differential Pressure Signals. *Int. J. Multiph. Flow* **2014**, *67*, 106–117. [CrossRef]
30. Osman, E.-S.A. Artificial Neural Network Models for Identifying Flow Regimes and Predicting Liquid Holdup in Horizontal Multiphase Flow. *SPE Prod. Facil.* **2004**, *19*, 33–40. [CrossRef]
31. Li, X.; Miskimins, J.L.; Sutton, R.P.; Hoffman, B.T. Multiphase Flow Pattern Recognition in Horizontal and Upward Gas-Liquid Flow Using Support Vector Machine Models. In Proceedings of the SPE Annual Technical Conference and Exhibition, Amsterdam, The Netherlands, 27–29 October 2014. [CrossRef]
32. Ezzatabadipour, M.; Singh, P.; Robinson, M.D.; Guillen-Rondon, P.; Torres, C. Deep Learning as a Tool to Predict Flow Patterns in Two-Phase Flow. *arXiv* **2017**, arXiv:1705.07117. Available online: Https://arxiv.org/pdf/1705.07117 (accessed on 1 August 2021).
33. Guillen-Rondon, P.; Robinson, M.D.; Torres, C.; Pereya, E. Support Vector Machine Application for Multiphase Flow Pattern Prediction. *arXiv* **2018**, arXiv:1806.05054. Available online: Http://arxiv.org/abs/1806.05054 (accessed on 1 August 2021).
34. Mask, G.; Wu, X.; Ling, K. An Improved Model for Gas-liquid Flow Pattern Prediction Based on Machine Learning. *J. Pet. Sci. Eng.* **2019**, *183*, 106370. [CrossRef]
35. Hernandez, J.S.; Valencia, C.; Ratkovich, N.; Torres, C.F.; Muñoz, F. Data Driven Methodology for Model Selection in Flow Pattern Prediction. *Heliyon* **2019**, *5*, e02718. [CrossRef]
36. Seong, Y.; Park, C.; Choi, J.; Jang, I. Surrogate Model with a Deep Neural Network to Evaluate Gas–Liquid Flow in a Horizontal Pipe. *Energies* **2020**, *13*, 968. [CrossRef]
37. Mohammadi, S. A Machine Learning Model to Find a Near-Optimal Set of Closure Relationships for the Multiphase Unified Model. Ph.D. Thesis, The University of Tulsa, Tulsa, OK, USA, 2020.
38. Lin, Z.; Liu, X.; Lao, L.; Liu, H. Prediction of Two-phase Flow Patterns in Upward Inclined Pipes via Deep Learning. *Energy* **2020**, *210*, 118541. [CrossRef]

39. Mukherjee, H. An Experimental Study of Inclined Two-Phase Flow. Master's Thesis, The University of Tulsa, Tulsa, OK, USA, 1979.
40. Magrini, K.L. Liquid Entrainment in Annular Gas-liquid Flow in Inclined Pipes. Master's Thesis, The University of Tulsa, Tulsa, OK, USA, 2009.
41. Yuan, G. Liquid Loading of Gas Wells. Master's Thesis, The University of Tulsa, Tulsa, OK, USA, 2011.
42. Fan, Y. An Investigation of Low Liquid Loading Gas-Liquid Stratified Flow in Near-Horizontal Pipes. Ph.D. Thesis, The University of Tulsa, Tulsa, OK, USA, 2005.
43. Gokcal, B. An Experimental and Theoretical Investigation of Slug Flow for High Oil Viscosity in Horizontal Pipes. Ph.D. Thesis, The University of Tulsa, Tulsa, OK, USA, 2008.
44. Brito, R. Effect of Medium Oil Viscosity on Two-Phase Oil-Gas Flow Behavior in Horizontal Pipes. Master's Thesis, The University of Tulsa, Tulsa, OK, USA, 2012.
45. Brill, J.P.; Chen, X.T.; Flores, J.G.; Marcano, R. *Transportation of Liquids in Multiphase Pipelines under Low Liquid Loading Conditions*; Final Report; The Pennsylvania State University: University Park, PA, USA, 1995.
46. Guner, M. Liquid Loading of Gas Wells with Deviations from $0°$ to $45°$. Master's Thesis, The University of Tulsa, Tulsa, OK, USA, 2012.
47. Meng, W. Low Liquid Loading Gas-Liquid Two-Phase Flow in Near-Horizontal Pipes. Ph.D. Thesis, The University of Tulsa, Tulsa, OK, USA, 1999.
48. Langsholt, M.; Holm, H. Liquid Accumulation in Gas-Condensate Pipelines—An Experimental Study. In Proceedings of the 13th International Conference on Multiphase Production Technology, Edinburgh, UK, 13–15 June 2007.
49. Exploratory Data Analysis. *The Concise Encyclopedia of Statistics*; Springer: New York, NY, USA, 2008; pp. 192–194.
50. Seo, S. A Review and Comparison of Methods for Detecting Outliers in Univariate Datasets. Master's Thesis, University of Pittsburgh, Pittsburgh, PA, USA, 2006.
51. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
52. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA, 2016; Volume 10, pp. 785–794.
53. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media: Newton, MA, USA, 2019.
54. Pandey, Y.N.; Rastogi, A.; Kainkaryam, S.; Bhattacharya, S.; Saputelli, L. *Machine Learning in the Oil and Gas Industry: Including Geosciences, Reservoir Engineering, and Production Engineering with Python*; Apress: New York, NY, USA, 2020.
55. Cohen, L.S.; Hanratty, T.J. Effect of Waves at a Gas—Liquid Interface on a Turbulent Air Flow. *J. Fluid Mech.* **1968**, *31*, 467–479. [CrossRef]
56. Vlachos, N.; Paras, S.; Karabelas, A. Liquid-to-wall Shear Stress Distribution in Stratified/atomization Flow. *Int. J. Multiph. Flow* **1997**, *23*, 845–863. [CrossRef]
57. Wallis, G.B. *One-Dimensional Two-Phase Flow*; McGraw-Hill: New York, NY, USA, 1969.
58. Whalley, P.B.; Hewitt, G.F. *The Correlation of Liquid Entrainment Fraction and Entrainment Rate in Annular Two-Phase Flow*; UKAEA Atomic Energy Research Establishment: Harwell, UK, 1978.
59. Fore, L.; Beus, S.; Bauer, R. Interfacial Friction in Gas–liquid Annular Flow: Analogies to Full and Transition Roughness. *Int. J. Multiph. Flow* **2000**, *26*, 1755–1769. [CrossRef]
60. Dallman, J.C.; Jones, B.G.; Hanratty, T.J. Interpretation of Entrainment Measurements in Annular Gas-Liquid Flows. In *Two-phase Momentum, Heat and Mass Transfer in Chemical, Process, and Energy Engineering Systems*; Hemisphere Publishing Corporation: London, UK, 1979; Volume 2, pp. 681–693.
61. Ambrosini, W.; Andreussi, P.; Azzopardi, B. A Physically Based Correlation for Drop Size in Annular Flow. *Int. J. Multiph. Flow* **1991**, *17*, 497–507. [CrossRef]
62. Andritsos, N.; Tzotzi, C.; Hanratty, T.J. Interfacial Shear Stress in Wavy Stratified Gas-liquid Two-phase Flow. In Proceedings of the 5th European Thermal-Sciences Conference, Eindhoven, The Netherlands, 18–22 May 2008.
63. Zhang, H.-Q.; Wang, Q.; Sarica, C.; Brill, J.P. Unified Model for Gas-Liquid Pipe Flow via Slug Dynamics—Part 1: Model Development. *J. Energy Resour. Technol.* **2003**, *125*, 266–273. [CrossRef]
64. Brito, R.M. Effect of Horizontal Well Trajectory on Two-Phase Gas-Liquid Flow Behavior. Ph.D. Theis, The University of Tulsa, Tulsa, OK, USA, 2015.
65. Rastogi, A.; Fan, Y. Experimental and Modeling Study of Onset of Liquid Accumulation. *J. Nat. Gas Sci. Eng.* **2020**, *73*, 103064. [CrossRef]
66. Turner, R.G.; Hubbard, M.G.; Dukler, A.E. Analysis and Prediction of Minimum Flow Rate for the Continuous Removal of Liquids from Gas Wells. *J. Pet. Technol.* **1969**, *21*, 1475–1482. [CrossRef]
67. Coleman, S.B.; Clay, H.B.; McCurdy, D.G.; Norris, L.H.I. A New Look at Predicting Gas-Well Load-Up. *J. Pet. Technol.* **1991**, *43*, 329–333. [CrossRef]
68. Li, M.; Lei, S.; Li, S. New View on Continuous-Removal Liquids from Gas Wells. In Proceedings of the SPE Permian Basin Oil and Gas Recovery Conference, Midland, TX, USA, 15–17 May 2001.
69. Wang, Y.Z.; Liu, Q.W. A New Method to Calculate the Minimum Critical Liquid Carrying Flow Rate for Gas Wells. *Pet. Geol. Oilfield Dev. Daqing* **2007**, *6*, 82–85.

70. Belfroid, S.; Schiferli, W.; Alberts, G.; Veeken, C.A.M.; Biezen, E. Predicting Onset and Dynamic Behavior of Liquid Loading Gas Wells. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 21–24 September 2018. [CrossRef]

71. Zhou, D.; Yuan, H. A New Model for Predicting Gas-Well Liquid Loading. *SPE Prod. Oper.* **2010**, *25*, 172–181. [CrossRef]

72. Luo, S.; Kelkar, M.; Pereyra, E.; Sarica, C. A New Comprehensive Model for Predicting Liquid Loading in Gas Wells. *SPE Prod. Oper.* **2014**, *29*, 337–349. [CrossRef]

73. Shekhar, S.; Kelkar, M.; Hearn, W.J.; Hain, L.L. Improved Prediction of Liquid Loading in Gas Wells. *SPE Prod. Oper.* **2017**, *32*, 539–550. [CrossRef]

74. Fan, Y.; Pereyra, E.; Sarica, C. Onset of Liquid-Film Reversal in Upward-Inclined Pipes. *SPE J.* **2018**, *23*, 1630–1647. [CrossRef]

75. Nair, J.V. Gas Lift Application and Severe Slugging in Toe-Down Horizontal Wells. Ph.D. Thesis, The University of Tulsa, Tulsa, OK, USA, 2017.

76. Espedal, M. An Experimental Investigation of Stratified Two-Phase Pipe Flow at Small Inclinations. Ph.D. Thesis, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 1998.

77. Baker, R.E.; Peña, J.-M.; Jayamohan, J.; Jérusalem, A. Mechanistic Models versus Machine Learning, a Fight worth Fighting for the Biological Community? *Biol. Lett.* **2018**, *14*, 20170660. [CrossRef] [PubMed]