ARTICLE

Application of Supervised Machine Learning Techniques for the Prediction of Pressure Drop and Void Fraction in Gas–Liquid Two-Phase Flow in Horizontal Pipelines

Marco Cruz[1] **and Janelle Santos**[2]

[1] Polytechnic University of the Philippines, Anonas Street, Sta. Mesa, Manila 1016, Philippines
[2] University of Southeastern Philippines, Iñigo Street, Obrero, Davao City 8000, Philippines

Abstract

Gas–liquid two-phase flow in horizontal pipelines is encountered in a variety of industrial systems, including natural gas gathering networks, chemical reactors, and district heating lines. In such systems the prediction of frictional pressure drop and void fraction is relevant for equipment sizing, control strategies, and safety assessment. Classical approaches are based on mechanistic or empirical correlations that depend on assumptions about flow pattern and interfacial momentum transfer. These approaches can exhibit limited robustness when applied beyond the conditions for which they were developed, especially under wide variations in mass flux, fluid properties, or pipe diameter. Supervised machine learning offers an alternative route that can accommodate nonlinear relations between operating variables and hydraulic responses by learning from combined experimental and numerically generated data. This work examines the application of several supervised learning techniques to the prediction of pressure drop and cross-section-averaged void fraction in horizontal gas–liquid flow. The study considers models ranging from linear regressors to nonlinear kernel methods, ensemble tree algorithms, and feed-forward neural networks, and compares them under a common data-processing workflow. Particular attention is given to the use of non-dimensional inputs, to the incorporation of physically motivated constraints, and to the analysis of model generalization across operating ranges. The results illustrate how data-driven models can complement mechanistic formulations and how error structure, uncertainty representation, and sensitivity analysis can be used to interpret the learned relations without relying on flow-regime classification schemes.

1 Introduction

Gas–liquid two-phase flow in pipelines is of interest in many engineering applications, including upstream oil and gas transport, wet-gas transmission, refrigeration systems, and certain chemical processing lines [1]. In horizontal or nearly horizontal configurations, the coexisting phases may form a variety of topologies such as stratified, wavy, intermittent, annular, and dispersed patterns. For design and operation, key output quantities include the overall pressure drop and the void fraction, which together determine the required pumping or compression power, the liquid holdup, and the capacity limits of the pipeline. These quantities depend on phase mass fluxes, physical properties, pipe geometry, and the details of interfacial shear and phase distribution [2]. The complexity of the associated transport phenomena makes prediction a nontrivial task.

Traditional predictive tools generally fall into three broad classes: homogeneous models, drift-flux formulations, and two-fluid or mechanistic models. Homogeneous models assume equal phase velocities and treat the mixture as a single effective fluid, which is convenient but often inaccurate when slip between phases is significant [3]. Drift-flux and related formulations introduce a slip relation between phase velocities using empirical parameters, allowing

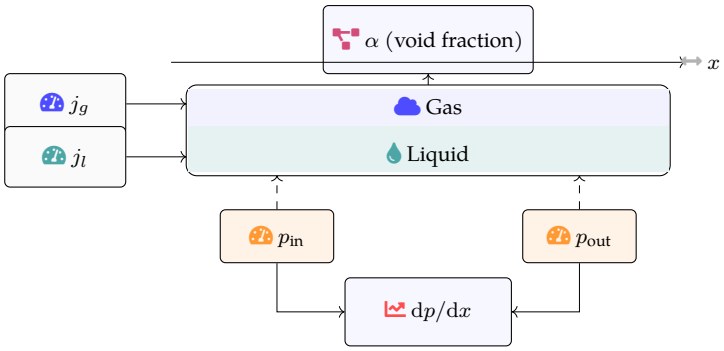

Figure 1. Simplified schematic of horizontal gas–liquid two-phase flow. Gas and liquid enter with superficial velocities j_g and j_l; the void fraction α characterizes the cross-section. Pressure sensors at inlet and outlet measure p_{in} and p_{out}, yielding the frictional gradient $\mathrm{d}p/\mathrm{d}x$ along the pipe.

Table 1. Representative engineering contexts for horizontal gas–liquid flow

Application	Typical fluids	Key purpose
Oil and gas transport	Crude oil, gas, water	Flow assurance, pressure loss
Gas networks	Wet gas	Compression and line sizing
Refrigeration loops	Refrigerant–oil	Thermal performance
District heating	Water–air	Pumping and cavitation control
Chemical pipelines	Reactants–gases	Safety and residence time

improved representation of void fraction while preserving a relatively simple structure. Two-fluid models attempt to write separate momentum balances for each phase and require closure relations for wall and interfacial shear, interfacial pressure terms, and, in some cases, mass transfer. To connect these models to practical calculations, many correlations have been developed for friction factors, flow regime boundaries, and slip or distribution parameters [4].

Although mechanistic and correlation-based methods are widely used, their predictive capability is closely tied to the data and flow conditions used in their derivation [5]. Extrapolation to new fluids, extended ranges of mass flux, or different pipe diameters can lead to increased uncertainty. In addition, many correlations require explicit classification of the flow pattern, which itself may depend on empirical maps. The resulting workflows can be sensitive to classification errors and discontinuities near regime boundaries [6]. As measurements and high-fidelity simulations of two-phase flow have become more available, there has been growing interest in complementing traditional models with data-driven approaches that can exploit larger datasets and flexible functional representations.

Supervised machine learning is one such approach, in which a model is trained to map input variables, such as mass fluxes, superficial velocities, and fluid properties, to target outputs, such as pressure gradient and void fraction. Various algorithms, including kernel-based regression, ensemble decision trees, and neural networks, can approximate nonlinear mappings without the need to specify explicit mechanistic forms. Nevertheless, applying these techniques to two-phase flow requires careful consideration of data quality, feature construction, and physical constraints [7]. Void fraction must remain between zero and one, pressure drop must be consistent with energy dissipation, and extrapolation beyond the training domain needs to be treated cautiously.

The present work examines the use of supervised learning techniques for predicting pressure drop and void fraction in gas–liquid flow through horizontal pipes. The focus is on steady, adiabatic conditions with no phase change, so that the mass flow rates of each phase are constant along the pipeline [8]. A combination of experimental and numerically generated data is considered to span a range of gas and liquid superficial velocities, viscosities, densities, and pipe diameters. Methodologically, the study emphasizes the interplay between physical modeling and data-driven regression, including the use of non-dimensional groups, physics-based constraints in the loss function, and hybrid configurations in which mechanistic estimates are refined by machine learning models. The following sections describe the governing equations and physical modeling assumptions, the structure and preparation of the data, the supervised

Table 2. Common two-phase modeling approaches

Model	Main idea	Limitation
Homogeneous	Equal phase velocity	Fails with strong slip
Drift-flux	Slip via C_0, V_{gj}	Regime-dependent
Two-fluid	Separate momentum	Many closures required
Empirical maps	Regime-based fits	Sensitive to classification
CFD	Field-resolved	High computational cost

Table 3. Key governing relations

Relation	Equation	Role
Mass conservation	$\frac{\partial}{\partial x}(\alpha_k \rho_k u_k) = 0$	Constant flux per phase
Pressure split	$\frac{\mathrm{d}p}{\mathrm{d}x} = (\frac{\mathrm{d}p}{\mathrm{d}x})_f + (\frac{\mathrm{d}p}{\mathrm{d}x})_a$	Friction + acceleration
Frictional loss	$(\frac{\mathrm{d}p}{\mathrm{d}x})_f = -2 f_{tp} \rho_m u_m^2 / D$	Main loss term
Void fraction	$\alpha = \frac{j_g}{C_0 j + V_{gj}}$	Drift-flux form

learning framework, the integration of physical priors, and the resulting predictive behavior across operating conditions [9].

2 Governing Equations and Physical Modeling

The two-phase flow considered here is assumed to be steady on average, one-dimensional along the pipe axis, and adiabatic, with no mass transfer between phases. The pipeline is horizontal, so gravitational effects enter primarily through hydrostatic pressure distribution associated with nonuniform phase holdup rather than elevation changes along the axis. For modeling purposes, the gas and liquid phases are treated as interpenetrating continua with volume fractions that sum to unity. The flow is represented using a one-dimensional two-fluid formulation under the assumption of cross-sectionally averaged quantities [10].

Let the subscripts g and l denote gas and liquid, respectively. The cross-sectionally averaged void fraction is denoted by α, representing the volume fraction of the gas phase, and the liquid holdup is $1-\alpha$. The densities are ρ_g and ρ_l, while the average axial velocities are u_g and u_l. The pipe has internal diameter D and cross-sectional area $A = \pi D^2/4$. Under the one-dimensional two-fluid approach, the mass conservation equation for each phase in differential form can be written as [11]

$$\frac{\partial}{\partial t}\left(\alpha_k \rho_k\right) + \frac{\partial}{\partial x}\left(\alpha_k \rho_k u_k\right) = 0, \tag{1}$$

where the index k stands for g or l. For steady flow with constant mass flow rates of each phase, the time derivative in equation (1) vanishes and the product $\alpha_k \rho_k u_k$ is constant along the pipe. In the absence of phase change, densities can be taken as constant for many applications, so that mass conservation reduces to conservation of volumetric flux.

The momentum conservation equations for each phase in the axial direction can be expressed in one-dimensional form as

$$\frac{\partial}{\partial t}\left(\alpha_k \rho_k u_k\right) + \frac{\partial}{\partial x}\left(\alpha_k \rho_k u_k^2\right) = -\alpha_k \frac{\partial p}{\partial x} + M_k + S_k, \tag{2}$$

where p is the common pressure field shared by the phases, M_k represents interfacial momentum exchange, and S_k contains source terms associated with wall shear and gravity. For steady, fully developed horizontal flow with negligible acceleration of the average velocities, the convective and storage terms on the left-hand side of equation (2) often play a reduced role, and the dominant balance involves pressure gradient and wall and interfacial shear.

A practical decomposition of the axial pressure gradient in two-phase flow separates frictional, accelerational, and hydrostatic contributions [12]. For horizontal flow, the elevation change term is zero, but the nonuniform phase distribution can still contribute through effective density terms in stratified configurations. The overall pressure gradient is often written as

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \left(\frac{\mathrm{d}p}{\mathrm{d}x}\right)_f + \left(\frac{\mathrm{d}p}{\mathrm{d}x}\right)_a, \tag{3}$$

where the subscript f denotes frictional and a denotes accelerational contributions [13]. Under conditions of roughly constant cross-section and modest density change with pressure, the accelerational term is often small compared to frictional losses, particularly in many liquid-dominated flows. Nevertheless, for

Table 4. Dimensionless input groups

Symbol	Definition	Effect
Re_k	$\rho_k j_k D/\mu_k$	Flow regime
Fr	$u_m/\sqrt{gD}$	Gravity vs inertia
We_k	$\rho_k j_k^2 D/\sigma$	Surface tension effects
r_ρ	ρ_g/ρ_l	Slip and buoyancy

Table 5. Learning models compared

Model	Advantage	Limitation
Linear	Simple, interpretable	Misses nonlinearities
Kernel	Smooth nonlinear fit	Scales poorly with data
Tree ensemble	Robust on tabular data	Less transparent
Neural net	Flexible mapping	Needs tuning

completeness, it can be included in a mechanistic estimate that is later combined with machine learning corrections.

To calculate the frictional component of the pressure gradient, many approaches introduce an effective two-phase friction factor and mixture density [14]. A common form for the frictional pressure gradient is

$$\left(\frac{\mathrm{d}p}{\mathrm{d}x}\right)_f = -\frac{2 f_{tp} \rho_m u_m^2}{D}, \tag{4}$$

where f_{tp} is the two-phase friction factor, ρ_m is an effective mixture density, and u_m is a mixture velocity defined as the total volumetric flow rate divided by the cross-sectional area. The effective density can be written as

$$\rho_m = \alpha \rho_g + (1-\alpha)\rho_l, \tag{5}$$

which depends on the void fraction. For the mixture velocity one may write [15]

$$u_m = j_g + j_l, \tag{6}$$

where the superficial velocities j_g and j_l are given by

$$j_k = \frac{\dot{m}_k}{\rho_k A}. \tag{7}$$

Here $\dot{m}_k$ is the mass flow rate of phase k. The two-phase friction factor f_{tp} can be expressed in terms of single-phase friction factors and a multiplier function of operating conditions, or it can be correlated directly as a function of the non-dimensional groups characterizing the flow.

A key quantity for interpreting gas–liquid flow is the void fraction α, which determines the volumetric proportions of gas and liquid and influences both interfacial and wall shear stresses. In many engineering correlations, the drift-flux concept is used to relate α to the phase superficial velocities. In the drift-flux framework, the volumetric flux of gas phase is written as [16]

$$j_g = C_0 j \alpha + V_{gj} \alpha, \tag{8}$$

where $j = j_g + j_l$ is the total volumetric flux, C_0 is the distribution parameter, and V_{gj} is the drift velocity that accounts for relative motion between the gas and the mixture. Rearranging equation (8) yields an expression for the void fraction as

$$\alpha = \frac{j_g}{C_0 j + V_{gj}}. \tag{9}$$

The parameters C_0 and V_{gj} typically depend on flow pattern, pipe orientation, and fluid properties, and are commonly described by empirical correlations. In a mechanistic context, these parameters can be linked to the balance of forces acting on bubbles or slugs; in this work they are used as example physical priors that can be incorporated into machine learning feature sets.

The non-dimensionalization of governing equations is useful both for physical interpretation and for constructing scalable input features for supervised learning models. Important dimensionless groups include the phase Reynolds numbers

$$Re_k = \frac{\rho_k j_k D}{\mu_k}, \tag{10}$$

where μ_k is the dynamic viscosity of phase k, the Froude number

$$Fr = \frac{u_m}{\sqrt{gD}}, \tag{11}$$

which compares inertial and gravitational effects, and the Weber number [17]

$$We_k = \frac{\rho_k j_k^2 D}{\sigma}, \tag{12}$$

Table 6. Hybrid modeling schemes

Type	Form	Purpose
Additive correction	$y = y_{\text{mech}} + \delta_\theta$	Learn residuals
Bounded output	Sigmoid/softplus layer	Enforce physical range
Penalty term	Dual loss	Stay close to physics
Monotone net	Sign-limited weights	Keep trends

Table 7. Error and uncertainty metrics

Metric	Expression	Meaning
MAE	$\frac{1}{N}\sum \lvert\hat{y} - y\rvert$	Average absolute error
RMSE	$\sqrt{\frac{1}{N}\sum(\hat{y} - y)^2}$	Penalizes large errors
R^2	$1 - \frac{\sum(\hat{y}-y)^2}{\sum(y-\bar{y})^2}$	Explained variance
Var_{ens}	Ensemble spread	Predictive uncertainty

where σ is the surface tension. These groups enter into many flow regime maps and correlation forms and also serve as candidate features for machine learning models, either directly or after further transformation.

When mechanistic predictions are used as inputs or baselines for learning, the conservation equations must be closed by specifying wall and interfacial shear stresses [18]. In a stratified or annular flow representation, one may write a wall shear stress for phase k as

$$\tau_{w,k} = \frac{1}{8} f_k \rho_k u_k^2, \tag{13}$$

where f_k is a Darcy friction factor based on the hydraulic diameter of the wetted region associated with that phase. The interfacial shear stress can be expressed similarly as

$$\tau_i = \frac{1}{8} f_i \rho_g \left(u_g - u_l\right)^2, \tag{14}$$

with an interfacial friction factor f_i. Correlations for f_k and f_i usually depend on Reynolds numbers constructed with appropriate characteristic velocities and length scales.

For the purposes of this work, the mechanistic model is used as a source of physically consistent baseline values for pressure gradient and void fraction. These values can be obtained by discretizing the momentum balance along the pipe and solving for the pressure distribution under prescribed inlet conditions [19]. If the flow is nearly fully developed, the pressure gradient can be taken as approximately uniform over a sufficient length, and a one-dimensional algebraic relation such as equation (4) can be solved directly for $\mathrm{d}p/\mathrm{d}x$ once friction factors and void fraction are specified. Numerical solution schemes can involve finite volume or finite difference approximations in the axial direction; for instance, one may discretize the pipe into segments of length Δx and write a discrete pressure update

$$p^{n+1} = p^n + \left(\frac{\mathrm{d}p}{\mathrm{d}x}\right)_f^n \Delta x, \tag{15}$$

with the frictional gradient at cell n evaluated using local estimates of void fraction, mixture density, and friction factor. Although in the present context the focus lies on steady, averaged behavior, such discretizations illustrate the link between partial differential equations and algebraic forms that enter machine learning feature construction [20].

3 Data Sources, Feature Construction, and Preprocessing

The application of supervised learning methods for predicting pressure drop and void fraction in gas–liquid horizontal flow requires a dataset that captures relevant variations of operating and geometric parameters. Such a dataset may combine experimental measurements from test loops with synthetic values generated by mechanistic models under different conditions. The target variables of interest are the frictional pressure gradient and the cross-section-averaged void fraction, which are treated as continuous quantities [21]. The input variables describe the flow rates, fluid properties, and pipe characteristics.

Consider a dataset consisting of N samples, each corresponding to a combination of gas and liquid mass flow rates, densities, viscosities, surface tension, pipe diameter, and possibly pressure and temperature. For

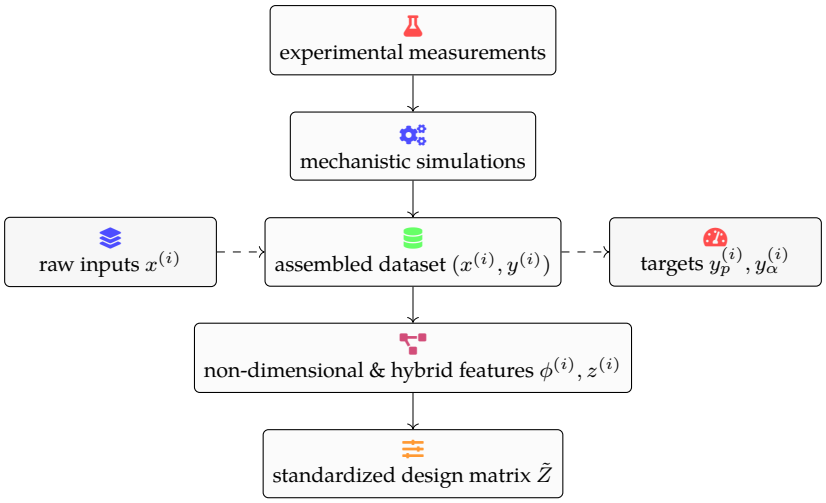

Figure 2. Data and feature-processing pipeline for supervised learning of pressure drop and void fraction. Experimental measurements and mechanistic simulations are merged into an assembled dataset of raw inputs $x^{(i)}$ and targets $y^{(i)}$, which are transformed into non-dimensional and hybrid feature vectors $\phi^{(i)}$ and $z^{(i)}$ and standardized into the design matrix $\tilde{Z}$ used for model training and evaluation.

sample index i, one can define the raw input vector

$$x^{(i)} = \left[j_g^{(i)}, j_l^{(i)}, \rho_g^{(i)}, \rho_l^{(i)}, \mu_g^{(i)}, \mu_l^{(i)}, \sigma^{(i)}, D^{(i)} \right]^\top \in \mathbb{R}^d, \tag{16}$$

with d equal to the number of scalar descriptors selected [22]. The corresponding targets are pressure gradient and void fraction,

$$y^{(i)} = \left[y_p^{(i)}, y_\alpha^{(i)} \right]^\top, \tag{17}$$

where $y_p^{(i)}$ represents the measured or computed frictional pressure gradient and $y_\alpha^{(i)}$ represents the measured or computed void fraction. Depending on the chosen learning strategy, one may train separate models for each target or a multi-output model.

Raw physical variables often span several orders of magnitude and may be correlated [23]. To improve conditioning and facilitate learning, it is common to construct derived features based on non-dimensional groups. For each sample, one can compute phase Reynolds numbers, Froude and Weber numbers, mixture velocity, and density and viscosity ratios. A feature vector consisting of non-dimensional quantities can be written as

$$\phi^{(i)} = \left[Re_g^{(i)}, Re_l^{(i)}, Fr^{(i)}, We_g^{(i)}, We_l^{(i)}, r_\rho^{(i)}, r_\mu^{(i)} \right]^\top, \tag{18}$$

where r_ρ and r_μ denote density and viscosity ratios. The mapping from $x^{(i)}$ to $\phi^{(i)}$ is deterministic and based on the relations presented earlier.

In addition to non-dimensional groups, mechanistic estimates can be used to augment the feature set [24]. For instance, a drift-flux estimate of void fraction $\alpha_{\mathrm{df}}^{(i)}$ and a mechanistic estimate of the frictional pressure gradient $(\mathrm{d}p/\mathrm{d}x)_{\mathrm{mech}}^{(i)}$ can be computed from equations similar to (9) and (4). These quantities can be treated as additional inputs to the learning model, providing it with physically motivated baselines. One can write an extended feature vector

$$z^{(i)} = \left[\phi^{(i)}, \alpha_{\mathrm{df}}^{(i)}, \left(\frac{\mathrm{d}p}{\mathrm{d}x} \right)_{\mathrm{mech}}^{(i)} \right]^\top, \tag{19}$$

which combines non-dimensional descriptors and mechanistic outputs in a single representation [25]. The dimension of $z^{(i)}$ remains moderate, which is advantageous for avoiding severe curse-of-dimensionality effects.

To prepare the dataset for supervised learning, the features are often standardized or normalized. A common approach is to subtract the empirical mean and divide by the empirical standard deviation for each feature dimension. If the design matrix for the extended features is denoted by [26]

$$Z = \begin{bmatrix} z^{(1)\top} \\ \vdots \\ z^{(N)\top} \end{bmatrix} \in \mathbb{R}^{N \times d_z}, \tag{20}$$

then for each column j one computes

$$\tilde{Z}_{ij} = \frac{Z_{ij} - \mu_j}{\sigma_j}, \tag{21}$$

where μ_j and σ_j are the sample mean and standard deviation of column j. This transformation produces a

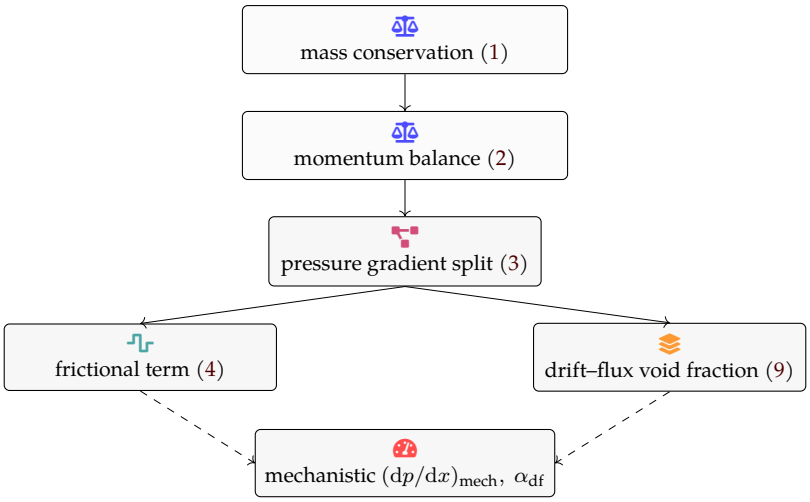

Figure 3. Connection between governing equations and mechanistic baseline predictions. Cross-sectionally averaged mass and momentum balances lead to a decomposition of the axial pressure gradient into frictional and accelerational components. Frictional models and drift–flux approximations for void fraction are combined to yield mechanistic estimates $(\mathrm{d}p/\mathrm{d}x)_{\mathrm{mech}}$ and α_{df} that can be used directly or as priors and additional features in the supervised learning framework.

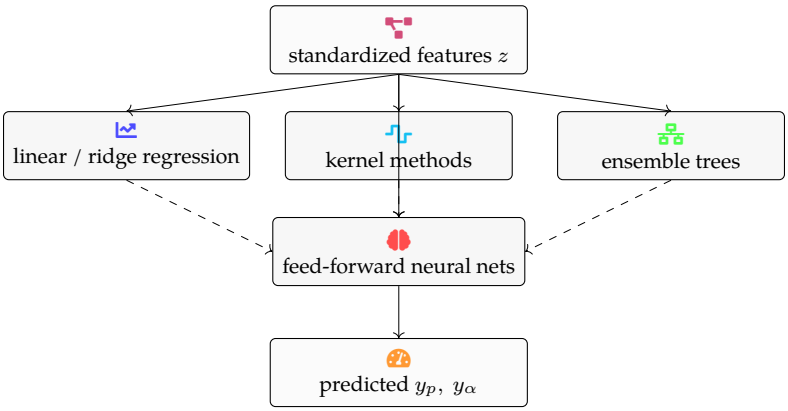

Figure 4. Supervised regression models mapping standardized feature vectors z to pressure-gradient and void-fraction outputs. Linear and kernel-based methods, ensemble decision trees, and feed-forward neural networks are trained under a unified empirical risk minimization framework, and their predictions for y_p and y_α are compared using consistent data splits and preprocessing.

standardized design matrix $\tilde{Z}$ with zero mean and unit variance in each feature direction, which can improve the conditioning of optimization problems and the convergence behavior of some learning algorithms.

The target variables may also be transformed. For example, the frictional pressure gradient may span several orders of magnitude between low and high flow rates, motivating a logarithmic transform. One may define [27]

$$\tilde{y}_p^{(i)} = \log\left(\varepsilon + y_p^{(i)}\right), \tag{22}$$

with a small positive offset ε to avoid taking the logarithm of very small values, while keeping void fraction in its original linear scale. The use of transformed targets can yield more uniform residual distributions and influence the bias between relative and absolute errors.

Splitting the dataset into training, validation, and test subsets is another part of preprocessing [28]. In many cases the data are partitioned randomly, but for assessing generalization across operating ranges it can be informative to stratify the split based on pipe diameter, fluid pair, or flow regime indicator. For instance, one may allocate all samples from a particular pipe diameter to the test set to assess extrapolation performance with respect to geometry. Alternatively, cross-validation procedures can be used, in which the dataset is divided into folds and models are trained on subsets and evaluated on the remaining data. Such strategies can reduce variance in performance estimates and help in hyperparameter selection [29].

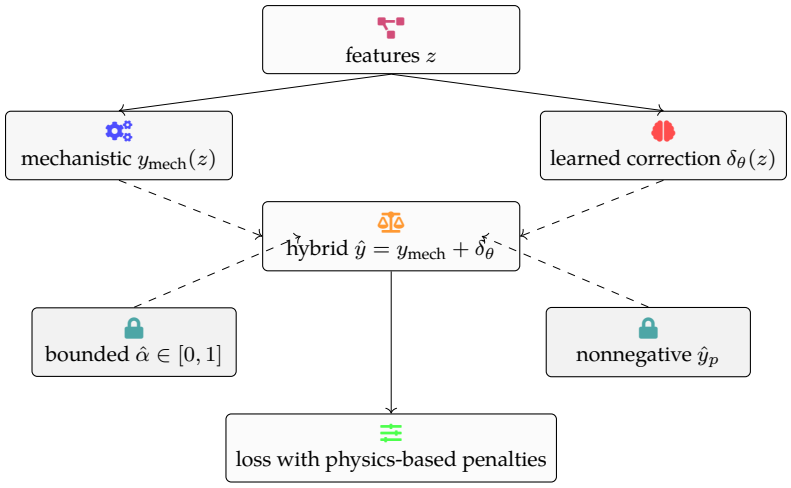

Figure 5. Hybrid and physics-constrained regression strategy for pressure drop and void fraction. Mechanistic predictions $y_{\text{mech}}(z)$ are refined by a residual model $\delta_\theta(z)$ to obtain hybrid outputs $\hat{y}$, while constraints such as bounded void fraction and nonnegative frictional pressure gradients are enforced through output parameterizations and additional penalty terms in the training objective.

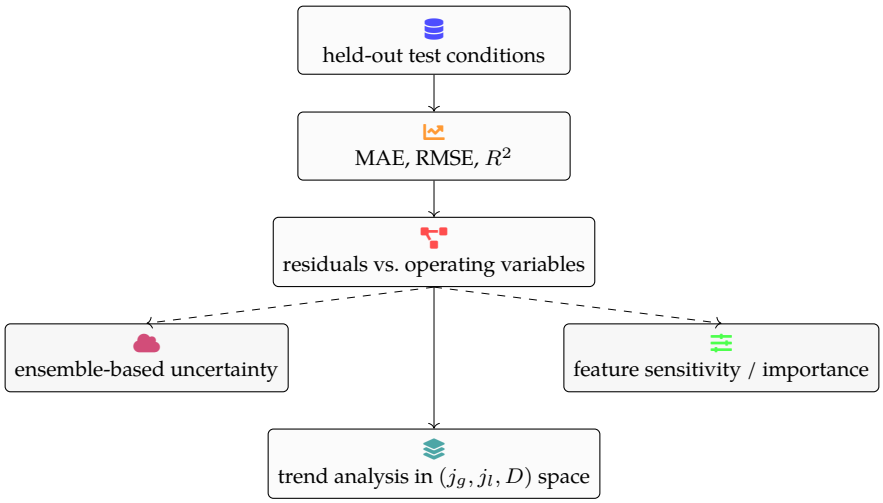

Figure 6. Evaluation and interpretation workflow for learned models. Predictions on held-out test conditions are summarized using error metrics such as MAE, RMSE, and R^2, followed by residual analysis across operating variables. Ensemble predictions provide uncertainty estimates, while feature importance and sensitivity diagnostics support interpretation of trends in pressure gradient and void fraction over the (j_g, j_l, D) space.

Outliers and measurement noise must also be considered. In experimental datasets, some points may exhibit inconsistent combinations of inputs and outputs due to sensor errors, transient conditions, or unsteady flow patterns. Simple screening based on physical constraints, such as void fraction between zero and one and nonnegative pressure drop, can remove obviously inconsistent points [30]. More subtle inconsistencies may be analyzed through residual diagnostics after initial model training. In the context of this work, the goal is not to remove all irregularities but to ensure that the learning algorithms operate on data that satisfy basic conservation and positivity constraints, so that they do not need to learn to correct gross violations.

4 Supervised Learning Framework for Pressure Drop and Void Fraction

The problem of predicting pressure drop and void fraction from operating conditions and physical properties can be formulated as a supervised regression task. Given a dataset of feature–target pairs $\{(z^{(i)}, y^{(i)})\}_{i=1}^{N}$, the goal is to find a function f such that $f(z^{(i)})$ approximates $y^{(i)}$ for the training data and generalizes to new inputs. The mapping f may be scalar-valued, predicting either pressure gradient or void fraction separately, or vector-valued, predicting both simultaneously [31]. Many supervised learning methods can be expressed within the framework of empirical risk minimization, in which one minimizes a regularized loss function.

Let the design matrix of standardized extended features be denoted by $\tilde{Z} \in \mathbb{R}^{N \times d_z}$, as in equation (20), and let $y \in \mathbb{R}^N$ denote the vector of target values for a single output variable, such as $\tilde{y}_p$ or y_α. A general supervised learning algorithm seeks parameters θ that minimize a cost function of the form

$$J(\theta) = \frac{1}{N} \sum_{i=1}^{N} \ell \left(f_\theta(z^{(i)}), y^{(i)} \right) + \lambda R(\theta), \tag{23}$$

where ℓ is a loss function that measures discrepancy between predictions and targets, $R(\theta)$ is a regularization term, and λ is a nonnegative regularization coefficient [32]. A typical choice for regression is the squared loss,

$$\ell\left(\hat{y}, y\right) = \left(\hat{y} - y\right)^2, \tag{24}$$

though absolute loss or Huber loss functions can be used for robustness to outliers. Regularization can take the form of an ℓ_2-norm, an ℓ_1-norm, or more structured penalties depending on the model.

Linear regression offers a baseline model in which the mapping from features to target is affine [33]. In matrix notation, the prediction for all samples can be written as

$$\hat{y} = \tilde{Z} w + b\mathbf{1}, \tag{25}$$

where $w \in \mathbb{R}^{d_z}$ is a weight vector, b is a scalar bias term, and $\mathbf{1}$ is the vector of ones in $\mathbb{R}^N$. The empirical risk with squared loss and ℓ_2-regularization is

$$J(w, b) = \frac{1}{N} \|\tilde{Z} w + b\mathbf{1} - y\|_2^2 + \lambda \|w\|_2^2, \tag{26}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Minimizing equation (26) leads to the normal equations associated with ridge regression. Linear models provide interpretable coefficients and can capture some trends when non-dimensional features are well chosen, but they may not be sufficient to represent strong nonlinear interactions among flow parameters.

To capture nonlinear dependencies, kernel methods such as kernel ridge regression and support vector regression can be employed. In kernel ridge regression, one introduces a positive definite kernel function $K(z, z')$ that measures similarity between feature vectors, and expresses the predictor as a linear combination of kernel evaluations at training points [34]. The prediction for a new input z is

$$\hat{y}(z) = \sum_{i=1}^{N} \alpha_i K(z, z^{(i)}), \tag{27}$$

where the coefficients α_i are obtained by solving a linear system that arises from minimizing a regularized squared loss. A common choice is the radial basis function kernel

$$K(z, z') = \exp\left(-\frac{\|z - z'\|_2^2}{2\sigma_k^2} \right), \tag{28}$$

with kernel width σ_k. Kernel methods can approximate a wide class of smooth functions, but their computational cost scales with the number of training samples, which can be a limitation for very large datasets [35].

Ensemble methods based on decision trees, such as random forests and gradient boosting machines, offer an alternative nonlinear modeling approach. Although the internal structure of trees is not easily expressed in compact analytical form, their learning process can still be described within a predictive framework. For example, gradient boosting constructs a model as a sum of weak learners,

$$\hat{y}(z) = \sum_{m=1}^{M} \gamma_m h_m(z), \tag{29}$$

where each h_m is a regression tree and γ_m is a step size determined during training. At each iteration, a new tree is fitted to the negative gradient of the loss function with respect to current predictions [36]. This stagewise additive modeling can approximate complex functional forms and is often competitive on tabular data.

Neural networks provide yet another class of flexible function approximators. In a feed-forward neural network with one hidden layer, the prediction can be written as [37]

$$\hat{y}(z) = v^\top \sigma\left(W z + c\right) + b, \tag{30}$$

where $W \in \mathbb{R}^{h \times d_z}$ is a weight matrix connecting inputs to h hidden units, $c \in \mathbb{R}^h$ is a hidden bias vector, $v \in \mathbb{R}^h$ is a vector connecting hidden units to the output, b is an output bias, and $\sigma(\cdot)$ denotes an activation function applied componentwise, such as a rectified linear or hyperbolic tangent function. The parameters $\theta = (W, c, v, b)$ are typically learned by minimizing a loss function via stochastic gradient descent or its variants. Deeper networks with multiple hidden layers can represent increasingly complicated functions, though they require careful regularization and sufficient data.

For multivariate outputs, such as simultaneous prediction of pressure gradient and void fraction, one

can extend the neural network to produce a vector output [38]. If the output dimension is q, the final linear layer may have a weight matrix $V \in \mathbb{R}^{q\times h}$ and bias vector $b \in \mathbb{R}^q$, yielding

$$\hat{y}(z) = V\sigma\left(Wz + c\right) + b. \tag{31}$$

The loss function can be defined as the sum of squared errors across outputs, possibly with different weights reflecting the relative importance or scaling of each variable. In the context of two-phase flow, one may assign weights to pressure and void fraction errors to balance their contributions to the overall cost function.

Probabilistic regression models, such as Gaussian process regression, provide predictive distributions rather than point estimates. In a Gaussian process model, the prior over functions is specified by a mean function and a covariance kernel, often chosen to be zero mean with a covariance similar to equation (28). Observations are assumed to be noisy, and the posterior over functions given the data yields a Gaussian predictive distribution at new points [39]. The predictive mean provides a point prediction, while the predictive variance quantifies uncertainty. Although Gaussian process regression may be computationally expensive for large datasets, it can be informative for quantifying uncertainty in regions of sparse data and for guiding the design of new experiments.

The empirical performance of a supervised learning model is commonly evaluated using metrics such as mean absolute error, root mean square error, and the coefficient of determination [40]. For a test set of size N_{test} with predictions $\hat{y}^{(i)}$ and targets $y^{(i)}$, the mean absolute error is

$$\text{MAE} = \frac{1}{N_{\text{test}}}\sum_{i=1}^{N_{\text{test}}}\left|\hat{y}^{(i)} - y^{(i)}\right|, \tag{32}$$

and the root mean square error is

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{test}}}\sum_{i=1}^{N_{\text{test}}}\left(\hat{y}^{(i)} - y^{(i)}\right)^2}. \tag{33}$$

The coefficient of determination can be expressed as

$$R^2 = 1 - \frac{\sum_{i=1}^{N_{\text{test}}}\left(\hat{y}^{(i)} - y^{(i)}\right)^2}{\sum_{i=1}^{N_{\text{test}}}\left(y^{(i)} - \bar{y}\right)^2}, \tag{34}$$

where $\bar{y}$ is the empirical mean of the target values in the test set. These metrics, along with error distributions and residual plots, provide insight into how well the models capture trends in pressure drop and void fraction under varying flow conditions.

5 Physics-Guided Hybrid Modeling and Constraints

While flexible supervised learning models can approximate complex nonlinear mappings, their predictions may violate physical constraints if these constraints are not encoded explicitly [41]. In gas–liquid two-phase flow, basic requirements include nonnegativity of pressure drop for dissipative flows, void fraction constrained between zero and one, monotonic relations between certain variables in given regimes, and consistency with conservation laws. Hybrid modeling strategies aim to combine mechanistic and data-driven components to respect such constraints while still exploiting data-driven flexibility.

One class of hybrid models treats machine learning as a correction to mechanistic predictions [42]. Let $y_{\text{mech}}(z)$ denote a mechanistic estimate of a target variable, such as the pressure gradient obtained from a friction factor correlation combined with a drift-flux void fraction estimate. A hybrid predictor can then be written as

$$\hat{y}(z) = y_{\text{mech}}(z) + \delta_\theta(z), \tag{35}$$

where $\delta_\theta(z)$ is a learned correction term parameterized by θ. The correction function can be represented by any of the supervised learning models discussed earlier, such as a neural network or gradient-boosted trees. Training proceeds by minimizing the discrepancy between the hybrid prediction $\hat{y}(z)$ and observed targets, effectively fitting the residuals of the mechanistic model. This structure can preserve the qualitative behavior of mechanistic predictions in regions where data are sparse, while improving accuracy in data-rich regions.

Another hybrid strategy incorporates mechanistic quantities into the loss function as soft constraints [43]. For example, consider the prediction of void fraction $\hat{\alpha}(z)$ by a neural network. To enforce that $\hat{\alpha}$ remains between zero and one, one can apply a bounded activation function at the output layer, such as a logistic sigmoid, yielding

$$\hat{\alpha}(z) = \frac{1}{1 + \exp\left(-s(z)\right)}, \tag{36}$$

where $s(z)$ is the unbounded network output before the final transformation. This mapping ensures that predicted void fractions satisfy $0 < \hat{\alpha} < 1$, thereby enforcing a basic physical constraint as a hard boundary. Similarly, pressure gradients can be constrained to be nonnegative by using functions

such as exponential or softplus transformations at the output.

Mechanistic priors can also be incorporated via penalty terms in the loss function [44]. Suppose that for each sample one has a mechanistic estimate $y_{\text{mech}}^{(i)}$ and wishes to prevent the learned model from deviating excessively from this estimate in regions where mechanistic behavior is considered reliable. A composite loss function can be defined as

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left[\left(f_\theta(z^{(i)}) - y^{(i)}\right)^2 + \beta\left(f_\theta(z^{(i)}) - y_{\text{mech}}^{(i)}\right)^2\right] + \lambda R(\theta), \tag{37}$$

where β is a nonnegative coefficient controlling the strength of the penalty. The first term in the summation measures data misfit, while the second term penalizes deviation from mechanistic predictions [45]. The balance between these components can be adjusted to reflect confidence in the mechanistic model relative to the measurements.

Monotonicity constraints represent another form of physical prior. For many two-phase flow regimes and fluids, the pressure gradient increases monotonically with increasing mass flux at fixed void fraction or fixed phase ratios. To encode such behavior in a neural network model, one can restrict certain weights to be nonnegative, yielding monotone dependence on chosen input features [46]. For example, consider a network in which the first layer applies a linear transformation with weight matrix W. If particular columns of W are constrained to be nonnegative, then the output of the network becomes nondecreasing with respect to the corresponding input components under common activation functions. Enforcing these sign constraints can be expressed as a projection step during training, in which negative values in selected weights are set to zero after each gradient update [47].

From a probabilistic perspective, hybrid models can be interpreted within a Bayesian framework as specifying priors centered around mechanistic functions. Let f_{mech} denote a mechanistic mapping from features to outputs, and consider a stochastic model in which the true function f is modeled as

$$f(z) = f_{\text{mech}}(z) + \epsilon(z), \tag{38}$$

where $\epsilon(z)$ is a random function representing deviations from mechanistic behavior. Placing a Gaussian process prior on $\epsilon(z)$ with zero mean and a covariance kernel $k(z, z')$ yields a prior over functions $f(z)$ centered at $f_{\text{mech}}(z)$. Conditioning on observed data produces a posterior that balances mechanistic expectations and empirical evidence. While a full Bayesian treatment may be computationally demanding, this perspective motivates regularization strategies that penalize large deviations from mechanistic baselines [48].

Hybrid modeling also interacts with numerical analysis considerations. Mechanistic predictions derived from discretized conservation equations may contain numerical errors due to grid resolution, iterative tolerances, or simplified representations of interfacial geometry. When such predictions are used as inputs or targets, the learning model implicitly learns to compensate for numerical artifacts [49]. To reduce this effect, one can assess the convergence of mechanistic computations with respect to discretization parameters. For example, if a finite volume scheme is used to approximate pressure gradients, one may examine the convergence of the discrete solution as the axial grid spacing Δx decreases, by computing differences between solutions on successive grids and verifying that they decrease according to the expected discretization order. The more converged the mechanistic solution, the less likely it is that the learning model will have to correct numerical artifacts [50].

The integration of constraints can extend beyond individual predictions to correlations between outputs. For instance, the coupling between pressure drop and void fraction can be reflected by constructing joint loss terms. Under quasi-homogeneous assumptions, the frictional pressure gradient can be linked to mixture properties via relations similar to equation (4), in which void fraction influences mixture density and viscosity. If separate models are trained for pressure gradient and void fraction, one can require that their outputs satisfy approximate consistency with such relations within a tolerance. This can be expressed as an additional penalty term in the loss function, [51]

$$J_{\text{cons}}(\theta_p, \theta_\alpha) = \frac{1}{N}\sum_{i=1}^{N}\left[\hat{y}_p^{(i)} - g\left(\hat{\alpha}^{(i)}, z^{(i)}\right)\right]^2, \tag{39}$$

where $\hat{y}_p^{(i)}$ and $\hat{\alpha}^{(i)}$ are the model predictions for pressure gradient and void fraction, respectively, g is a mechanistic relation, and θ_p, θ_α are the

corresponding parameter sets. Including J_{cons} in the overall objective encourages the two models to produce mutually consistent outputs, reducing the likelihood of contradictory predictions.

6 Results, Error Analysis, and Interpretation

The performance of supervised learning models for predicting pressure drop and void fraction in gas–liquid two-phase flow can be examined by analyzing test-set errors, residual distributions, and response trends as a function of input variables. While detailed numerical values depend on the specific dataset, some qualitative patterns are typically observed when comparing linear models, kernel methods, ensemble trees, and neural networks in this context.

For pressure gradient prediction, simple linear models built on non-dimensional features often capture the main dependence on mixture velocity and density ratios but may underfit regions where flow regime transitions introduce sharp changes in frictional behavior [52]. Residual plots against mixture velocity and Froude number can reveal systematic biases, for example underprediction in high-velocity, gas-dominated conditions and overprediction in liquid-dominated slug regimes. These patterns suggest that linear models have limited capacity to represent the nonlinear influence of parameters such as surface tension and viscosity ratio on interfacial waves and flow structure.

Nonlinear models such as gradient boosting and neural networks tend to reduce these systematic biases. For instance, when trained with extended features that include mechanistic predictions, a gradient boosting model may achieve lower mean absolute error and root mean square error across the test set [53]. Error distributions can be examined by constructing histograms of relative error in predicted pressure gradient, defined as

$$e_p^{(i)} = \frac{\hat{y}_p^{(i)} - y_p^{(i)}}{y_p^{(i)} + \varepsilon}, \tag{40}$$

where ε is a small offset introduced to regularize division by small targets. A model that captures nonlinearities more effectively typically yields a distribution of $e_p^{(i)}$ that is more tightly centered around zero and exhibits fewer large deviations.

For void fraction prediction, models that incorporate bounded output transformations often provide physically consistent results [54]. When a neural network with a sigmoid output layer is trained on a dataset with diverse flow conditions, predicted void fractions remain within the admissible interval without requiring post-processing. The error analysis can be performed similarly to pressure prediction, using absolute and relative error measures. Void fraction tends to be more sensitive to local flow structure and may exhibit larger scatter due to measurement uncertainties, especially in regimes with intermittent or wavy interfaces. Consequently, the achievable accuracy may be limited by data quality as much as by model expressiveness [55].

An important aspect of model evaluation is the assessment of extrapolation behavior. For instance, one may train models on data from a subset of pipe diameters and fluids, and then test them on conditions with different diameters or fluid pairs. In a case where training data cover smaller diameters and a test set involves a larger diameter, linear models based on non-dimensional numbers often preserve reasonably consistent trends, as non-dimensionalization reduces explicit dependence on geometry [56]. Nonlinear models can perform similarly if they rely on the same non-dimensional features, but models that heavily depend on raw dimensional inputs may experience degraded performance under extrapolation. Cross-validation schemes that leave out entire groups of conditions can help quantify such behavior.

Uncertainty estimation provides additional insight, particularly in safety-critical applications [57]. Ensemble-based methods can be used to approximate predictive uncertainty by training multiple models with different initializations or bootstrap resampling and analyzing the spread in their outputs. If $\hat{y}^{(i,1)}, \ldots, \hat{y}^{(i,M)}$ denote predictions from an ensemble of M models for sample i, an empirical predictive variance can be computed as

$$s^2(z^{(i)}) = \frac{1}{M-1} \sum_{m=1}^{M} \left[\hat{y}^{(i,m)} - \bar{y}^{(i)}\right]^2, \tag{41}$$

where $\bar{y}^{(i)}$ is the ensemble mean. Elevated predictive variance may indicate regions of input space that are underrepresented in the training data or where the mapping is intrinsically noisy. Comparing predictive variance with actual residuals helps gauge the calibration of the uncertainty estimates.

Interpretation of learned models can be facilitated by sensitivity analysis and feature importance measures [58]. For tree-based ensembles, one can compute

impurity-based or permutation-based importance scores for each feature, indicating their relative contribution to reducing prediction error. In pressure drop prediction, features such as mixture velocity, gas superficial velocity, and density ratio often rank highly, while for void fraction, drift-flux estimates and superficial velocities may dominate. To complement global importance, partial dependence plots can be constructed [59]. For a chosen feature, say mixture velocity u_m, the partial dependence function represents the average model prediction as u_m varies over a specified range while other features are marginalized. Such plots can reveal approximate monotonic trends, saturation behavior, or interactions between features.

The relation between two-phase pressure drop and void fraction offers an opportunity to perform joint interpretation. By evaluating the models at sequences of operating points where, for example, gas superficial velocity increases at fixed liquid superficial velocity, one can trace corresponding changes in predicted pressure gradient and void fraction. These trajectories can be compared to mechanistic expectations, such as increased void fraction and, depending on regime, either increased or decreased pressure gradient with growing gas velocity [60]. If the learned models exhibit trends consistent with known physics, this provides qualitative support for their reliability in regions lacking direct data.

From the perspective of numerical analysis, it is also useful to examine the smoothness and regularity of the learned function. For differentiable models such as neural networks and kernel regressors, one can compute gradients of the predictions with respect to input features, obtaining sensitivity information [61]. For example, the gradient of predicted pressure gradient with respect to mixture velocity, $\partial \hat{y}_p / \partial u_m$, can be evaluated analytically via automatic differentiation or backpropagation. Excessive oscillations or sign changes in these derivatives might indicate overfitting or lack of regularization. In contrast, smooth monotone gradients suggest that the model has captured stable trends.

Finally, comparisons between purely data-driven models and hybrid models that include mechanistic inputs or constraints can highlight the effect of physics guidance [62]. In many test cases, hybrid models show similar or slightly improved average accuracy compared to unconstrained models, while exhibiting fewer unphysical predictions in extrapolation. For example, a hybrid void fraction model that uses both a bounded output and a penalty for deviation from drift-flux estimates may produce predictions that stay in a narrow band around mechanistic values where data are scarce, while adjusting more substantially where dense data indicate systematic mechanistic errors. This behavior is consistent with a regularization effect that anchors the model to known physics.

7 Conclusion

The application of supervised machine learning techniques to the prediction of pressure drop and void fraction in gas–liquid two-phase flow in horizontal pipelines has been examined from the viewpoint of physical modeling, data preparation, and hybrid learning strategies [63]. Starting from the one-dimensional two-fluid formulation, basic conservation equations for mass and momentum were discussed, together with typical decompositions of the pressure gradient into frictional and accelerational components. The role of void fraction in determining mixture properties and interfacial shear was highlighted, and drift-flux concepts were introduced as a bridge between phase superficial velocities and volumetric phase distribution. Dimensionless groups such as Reynolds, Froude, and Weber numbers were identified as useful ingredients for constructing features that reduce explicit dependence on dimensional scales [64].

The supervised learning framework was formulated in terms of empirical risk minimization for regression, with linear, kernel-based, tree-based, and neural network models considered as possible function approximators. The importance of data preprocessing, including non-dimensionalization, standardization, target transformations, and careful train–test splitting, was emphasized. Mechanistic estimates were incorporated both as additional features and as baselines in hybrid modeling strategies, enabling models to learn corrections to established correlations rather than entirely new mappings. Various forms of regularization and constraint enforcement were discussed, including bounded output transformations for void fraction, nonnegativity constraints for pressure gradients, monotonicity constraints through sign restrictions on network weights, and consistency penalties tying pressure and void predictions to approximate mechanistic relationships [65].

Error analysis and interpretation methods were outlined, including the use of absolute and relative error metrics, residual diagnostics, and uncertainty

quantification through ensemble predictions. Strategies for assessing extrapolation performance across pipe diameters and fluid pairs were described, along with the use of feature importance and partial dependence analyses to gain insight into learned relations between flow parameters, pressure drop, and void fraction. The interplay between physical understanding and data-driven learning emerged as a central aspect of building models that are both accurate and consistent with basic hydraulic principles [66]. The considerations presented suggest that supervised machine learning can complement mechanistic and correlation-based methods for predicting pressure drop and void fraction in horizontal gas–liquid two-phase flow, particularly when combined with physics-guided feature construction and constraint-aware training. At the same time, the reliability of such models depends on the coverage and quality of underlying datasets, the careful design of input features, and attention to extrapolation and uncertainty. Extensions could include the incorporation of additional phenomena such as phase change, inclination effects, or transient behavior, as well as the integration of machine learning models into system-level simulators for networks of pipelines. Further developments in physics-informed learning and uncertainty quantification may provide additional tools for balancing empirical flexibility with adherence to conservation laws and thermodynamic consistency in complex multiphase flows [67].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgement

This work was supported without any funding.

References

[1] D. Janiga, J. Kwaśnik, and P. Wojnarowski, "Utilization of discrete fracture network (dfn) in modelling and simulation of a horizontal well-doublet enhanced geothermal system (egs) with sensitivity analysis of key production parameters," *Energies*, vol. 15, no. 23, pp. 9020–9020, Nov. 29, 2022. DOI: 10.3390/en15239020

[2] M. Yamamoto and M. Murai, "A numerical analysis of the internal flow effect on the riser's mechanical behavior," in *29th International Conference on Ocean, Offshore and Arctic Engineering: Volume 5, Parts A and B*, ASMEDC, Jan. 1, 2010, pp. 345–353. DOI: 10.1115/omae2010-20316

[3] J. Cromb, A. H. Summers, and A. Jefferies, "Overcoming the challenges of completion and installation operations associated with the gemini system from the ocean star," in *Offshore Technology Conference*, OTC, May 1, 2000. DOI: 10.4043/11865-ms

[4] D. Stephens and W. Clements, *Llnl underground-coal-gasification project. quarterly progress report, july-september 1981*, Nov. 9, 1981. DOI: 10.2172/5415197

[5] K. Manikonda et al., "Application of machine learning classification algorithms for two-phase gas-liquid flow regime identification," in *Abu Dhabi International Petroleum Exhibition and Conference*, SPE, 2021, D041S121R004.

[6] T. Ma, J. Liu, J. Fu, and B. Wu, "Drilling and completion technologies of coalbed methane exploitation: An overview," *International Journal of Coal Science & Technology*, vol. 9, no. 1, Sep. 22, 2022. DOI: 10.1007/s40789-022-00540-x

[7] S. Blackburn, R. Sanders, C. Boyer, E. Lasseter, J. Stevenson, and R. Mills, *Demonstration of the enrichment of medium quality gas from gob wells through interactive well operating practices. final report, june–december, 1995*, Dec. 1, 1995. DOI: 10.2172/244559

[8] P. J. L. Webster, B. Y. C. Leung, J. X. Z. Yu, M. D. Anderson, T. Hoult, and J. M. Fraser, "Coaxial real-time metrology and gas assisted laser micromachining: Process development, stochastic behavior, and feedback control," in *SPIE Proceedings*, vol. 7590, SPIE, Feb. 11, 2010, pp. 759 003–. DOI: 10.1117/12.842409

[9] *Geothermal energy research and development program; project summaries*, Mar. 1, 1994. DOI: 10.2172/860862

[10] V. N. Zakharov, V. A. Trofimov, Y. A. Filippov, and A. V. Shlyapin, "Abnormal gas phenomena in coal seams," *Gornyi Zhurnal*, pp. 80–87, Dec. 30, 2021. DOI: 10.17580/gzh.2021.12.15

[11] J. Sinner, *Improved miscible nitrogen flood performance utilizing advanced reservoir characterization and horizontal laterals in a class i reservoir - east binger (marchand) unit*, Jun. 30, 2004. DOI: 10.2172/923026

[12] B. Dong, "Drilling and completion challenges and countermeasures in kenkiyak complex sub-salt oil field development," in *Middle East Drilling Technology Conference & Exhibition*, SPE, Oct. 26, 2009. DOI: 10.2118/125482-ms

[13] F. Andritsos and H. Cozijn, "An innovative oil pollution containment method for ship wrecks proposed for offshore well blow-outs," in *Volume 2: Structures, Safety and Reliability*, ASMEDC, Jan. 1, 2011, pp. 73–81. DOI: 10.1115/omae2011-49110

[14] L. Peng, W. Kai, L. Bo, J. Yifeng, and G. Jianqiang, "Research on the effective influence radius of hydraulic reaming in mining seam," *The Open Fuels & Energy Science Journal*, vol. 8, no. 1, pp. 161–167, Jul. 31, 2015. DOI: 10.2174/1876973x01508010161

[15] J. S. S. Toralde, "Technology update: Retrofitting mpd systems to deepwater rigs aids drilling, efficiency, and process safety," *Journal of Petroleum Technology*,

vol. 69, no. 02, pp. 20–21, Feb. 1, 2017. DOI: 10.2118/0217-0020-jpt

[16] R. Matalucci, C. Esparza-Baca, and R. Jimenez, *Characterization, monitoring, and sensor technology catalogue*, Dec. 1, 1995. DOI: 10.2172/197816

[17] G. Karlsen, "A fully air deployable well capping stack and rov tooling system for worldwide support," in *SPE Asia Pacific Oil and Gas Conference and Exhibition*, SPE, Oct. 22, 2012. DOI: 10.2118/160409-ms

[18] H. Slater, D. Stiles, and W. Chmilowski, "Successful sealing of vent flows with ultra-low-rate cement squeeze technique," *SPE/IADC Drilling Conference*, Feb. 27, 2001. DOI: 10.2118/67775-ms

[19] F. Xiao, Z. Zhao, and L. Yang, "Innovative measures for thermal performance enhancement of single well-based deep geothermal systems: Existing solutions and some viable options," *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, vol. 8, no. 5, Jul. 28, 2022. DOI: 10.1007/s40948-022-00433-y

[20] C. Yang et al., "Experimental and numerical analysis of flow behavior for reverse circulation drill bit with inserted swirl vanes," *Geofluids*, vol. 2022, pp. 1–13, Jan. 10, 2022. DOI: 10.1155/2022/2472280

[21] D. Cohen, J. Dallas, F. Knight, and E. Woerheide, "Development of a gas handling hydraulic submersible pump and planning a field trial, captain field," in *Offshore Technology Conference*, OTC, May 5, 1997. DOI: 10.4043/8511-ms

[22] D. V. Eliseev et al., "New vision: Ic taml5 wells on caspian offshore. reasons, implementation and results," in *SPE Russian Petroleum Technology Conference and Exhibition*, SPE, Oct. 24, 2016. DOI: 10.2118/181901-ms

[23] A. L. Neverov*, L. S. Batalina, E. I. Starostina, and F. A. Buryukin, "Interaction between polymer-based drilling fluids containing ordinary salts and clay rocks," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 2, pp. 789–795, Dec. 30, 2019. DOI: 10.35940/ijitee.b6891.129219

[24] Q. Nie, S. Zhang, Y. Huang, X. Yi, and J. Wu, "Numerical and experimental investigation on safety of downhole solid–liquid separator for natural gas hydrate exploitation," *Energies*, vol. 15, no. 15, pp. 5649–5649, Aug. 4, 2022. DOI: 10.3390/en15155649

[25] Y. Tang, B. Wang, F. Zeng, and J. Wang, "The key and the countermeasures research of shale gas fracturingtechnology," *The Open Petroleum Engineering Journal*, vol. 8, no. 1, pp. 325–332, Aug. 19, 2015. DOI: 10.2174/1874834101508010325

[26] A. T. Bozdana and N. K. Al-Karkhi, "Comparative experimental investigation and gap flow simulation in electrical discharge drilling using new electrode geometry," *Mechanical Sciences*, vol. 8, no. 2, pp. 289–298, Sep. 20, 2017. DOI: 10.5194/ms-8-289-2017

[27] T. N. Ofei, S. Irawan, and W. Pao, "Cuttings-liquid frictional pressure loss model for horizontal narrow annular flow with rotating drillpipe," *IOP Conference Series: Materials Science and Engineering*, vol. 78, no. 1, pp. 012 009–, Apr. 2, 2015. DOI: 10.1088/1757-899x/78/1/012009

[28] S. Won, "Investigation of mud filtrate invasion using computational fluid dynamics," in *SPE Eastern Regional/AAPG Eastern Section Joint Meeting*, SPE, Oct. 11, 2008. DOI: 10.2118/117769-ms

[29] G. Liland, L. J. Prebeau-Menezes, and S. J. de Castro Mjolhus, "World's first taml level 5 multi-lateral well with individual remote inflow control of three branches on troll oil field," in *OTC Brasil*, OTC, Oct. 29, 2013. DOI: 10.4043/24427-ms

[30] S. Phillips, M. Tavana, K. Leung, and S. Schwartz, *Geothermal energy databook for the western united states (draft copy)*, Jun. 1, 1979. DOI: 10.2172/860864

[31] M. Crocker, M. Madden, and R. Porter, *Identification of environmental rd d needs*, Sep. 1, 1991. DOI: 10.2172/6248347

[32] R. Maglione and G. Robotti, "Field rheological parameters improve stand pipe pressure prediction while drilling," in *SPE Latin America/Caribbean Petroleum Engineering Conference*, SPE, Apr. 23, 1996. DOI: 10.2118/36099-ms

[33] H. A. Dogan, I. H. Gucuyener, and M. E. Ozbayoglu, "Investigation of bit hydraulics for gasified drilling fluids," in *SPE/ICoTA Coiled Tubing Conference & Exhibition*, SPE, Apr. 4, 2006. DOI: 10.2118/99596-ms

[34] K. Manikonda et al., "Horizontal two-phase flow regime identification with machine learning classification models," in *International Petroleum Technology Conference*, IPTC, 2022, D011S021R002.

[35] L. C. Sevillano and S. Sangesland, "Assessment of power requirements for alternative vertical transportation system for deepsea mining," in *Volume 4: Ocean Space Utilization*, American Society of Mechanical Engineers, Jun. 5, 2022. DOI: 10.1115/omae2022-80149

[36] S. O. R, C. Iturrios, A. Mahmood, and A. Kadadha, "Historical performance analysis and upcoming development of rotating control devices in the kingdom of saudi arabia," in *International Petroleum Technology Conference*, IPTC, Jan. 13, 2020. DOI: 10.2523/iptc-20238-ms

[37] A. L. Martins, A. Lourenço, P. H. Andrade, V. Silva, P. R. Silva, and E. Nakagawa, "Evaluation of solids carrying capacity in aerated fluid drilling: Real scale tests and modeling," *Revista de Engenharia Térmica*, vol. 2, no. 1, pp. 03–10, Jun. 30, 2003. DOI: 10.5380/ret.v2i1.3513

[38] *The objectives for deep scientific drilling in yellowstone national park*, Jan. 1, 1987. DOI: 10.2172/6956439

[39] R. Mann, *Slant hole completion test. final report*, Jul. 1, 1993. DOI: 10.2172/10108370

[40] Z. Zhang, S. Xiang, S. Liu, M. Luo, and J. Wu, "Study on the influence of tripping operation on annular transient surge-swab pressure of herschel–bulkley fluid," *Journal of Energy Resources Technology*, vol. 145, no. 5, Dec. 27, 2022. DOI: 10.1115/1.4056330

[41] R. P. Holst, "Preparing for lng: A montney gas field optimization workflow," in *SPE/CSUR*

Unconventional Resources Conference, SPE, Oct. 20, 2015. DOI: 10.2118/175939-ms

[42] A. Islamov, D. Kolupaev, R. Faskhutdinov, and S. Vereschagin, "On the mechanisms of the formation of zones with abnormally high rock pressure and methods for predicting them in undeveloped rock systems," in *SPE Russian Petroleum Technology Conference*, SPE, Oct. 15, 2018. DOI: 10.2118/191501-18rptc-ms

[43] M. J. Kieba and C. J. Ziolkowski, *Differential soil impedance obstacle detection*, Jun. 30, 2004. DOI: 10.2172/833263

[44] T. Huszar, G. Wittenberger, and E. Skvarekova, "Warning signs of high-pressure formations of abnormal contour pressures when drilling for oil and natural gas," *Processes*, vol. 10, no. 6, pp. 1106–1106, Jun. 1, 2022. DOI: 10.3390/pr10061106

[45] S. Z. Luan, X. Sun, K. L. Wang, and A. X. Geng, "Summary in research on kick detection technology," *Advanced Materials Research*, vol. 821-822, pp. 1422–1425, Sep. 18, 2013. DOI: 10.4028/www.scientific.net/amr.821-822.1422

[46] G. Buslaev, P. Tsvetkov, A. Lavrik, A. Kunshin, E. Loseva, and D. Sidorov, "Ensuring the sustainability of arctic industrial facilities under conditions of global climate change," *Resources*, vol. 10, no. 12, pp. 128–128, Dec. 15, 2021. DOI: 10.3390/resources10120128

[47] A. C. Vargas and R. Rodrigano, "Controlled flow-drilling induced by gas lift - an economic approach to the improvement of productivity in offshore scenarios," in *Offshore Technology Conference*, OTC, May 5, 2003. DOI: 10.4043/15207-ms

[48] P. Pournazari, P. Ashok, E. van Oort, S. Unrau, and S. Lai, "Enhanced kick detection with low-cost rig sensors through automated pattern recognition and real-time sensor calibration," in *SPE Middle East Intelligent Oil and Gas Conference and Exhibition*, SPE, Sep. 15, 2015. DOI: 10.2118/176790-ms

[49] D. Luo and F.-S. Li, "Discussion on the sustainable development of environment in the context of shale gas development," *DEStech Transactions on Environment, Energy and Earth Science*, no. sses/icfse, Jun. 1, 2017. DOI: 10.12783/dteees/sses/icfse2016/10648

[50] T. R. Salakhov, V. U. Yamaliev, and V. Dubinsky, "An expert system approach to real-time diagnosis of drill bit condition and preventing its damage," in *SPE Oil and Gas India Conference and Exhibition*, SPE, Jan. 20, 2010. DOI: 10.2118/127371-ms

[51] T. McVey, *Final report: Technoeconomic evaluation of undergroundcoal gasification (ucg) for power generationand synthetic natural gas*, Jun. 15, 2011. DOI: 10.2172/1118020

[52] S. Veenhuizen, T. O'Hanlon, D. Kelley, J. Duda, and J. Aslakson, "Ultra-high pressure down hole pump for jet-assisted drilling," *IADC/SPE Drilling Conference*, Mar. 12, 1996. DOI: 10.2118/35111-ms

[53] Amanullah, "A novel method of assessment of spurt and filtrate related formation damage potential of drilling and drilling-in fluids," in *SPE Asia Pacific Oil and Gas Conference and Exhibition*, SPE, Sep. 9, 2003. DOI: 10.2118/80484-ms

[54] Y. Guimin, J. Hao, and K. Qingwen, "Study on hydrate risk in the water drainage pipeline for offshore natural gas hydrate pilot production," *Frontiers in Earth Science*, vol. 9, Feb. 1, 2022. DOI: 10.3389/feart.2021.816873

[55] P. Du, W. Xia, and C. Yu, "Extension limit of a straight-swirling mixed jet bit and its influential factors in radial jet drilling," *Geofluids*, vol. 2022, pp. 1–10, May 23, 2022. DOI: 10.1155/2022/3078233

[56] Z. Khademian, K. M. Ajayi, S. J. Schatzel, G. S. Esterhuizen, and B. H. Kim, "Rockmass permeability induced by longwall mining under deep cover: Potential gas inflow from a sheared gas well," *Mining, Metallurgy & Exploration*, vol. 39, no. 4, pp. 1465–1473, Jun. 19, 2022. DOI: 10.1007/s42461-022-00635-8

[57] R. Lee, R. B. Grigg, and B. McPherson, *Validation and comparison of carbon sequestration project cost models with project cost data obtained from the southwest partnership*, Apr. 15, 2011. DOI: 10.2172/1025540

[58] R. Zhang and C. Hao, "Research on the development of hydraulic flushing caverning technology and equipment for gas extraction in soft and low permeability tectonic coal seams in china.," *ACS omega*, vol. 7, no. 25, pp. 21 615–21 623, Jun. 14, 2022. DOI: 10.1021/acsomega.2c01465

[59] W. J. McDonald and G. T. Pittard, *Development of a near-bit mwd system*, May 1, 1995. DOI: 10.2172/132678

[60] W. Wolfram and T. Law, "Santa barbara salm - a prototype deepwater production riser and floating production system," *Journal of Petroleum Technology*, vol. 32, no. 02, pp. 311–318, Feb. 1, 1980. DOI: 10.2118/7317-pa

[61] B. Wang, L. Wang, X. Liu, and F. Ren, "Bifurcation diagram and dynamic response of a drill string applied in ngh drilling," *Processes*, vol. 10, no. 6, pp. 1111–1111, Jun. 2, 2022. DOI: 10.3390/pr10061111

[62] E. D. Mattson and L. Hull, *Documentation of inl's in situ oil shale retorting water usage system dynamics model*, Dec. 1, 2012. DOI: 10.2172/1070124

[63] O. Kaldirim and J. Schubert, "An experimental study on riser gas behavior for dual gradient drilling," in *IADC/SPE Managed Pressure Drilling & Underbalanced Operations Conference & Exhibition*, SPE, Mar. 28, 2017. DOI: 10.2118/185297-ms

[64] K. Manikonda, C. Obi, A. A. Brahmane, M. A. Rahman, and A. R. Hasan, "Vertical two-phase flow regimes in an annulus image dataset-texas a&m university," *Data in Brief*, vol. 58, p. 111 245, 2025.

[65] L. M. LaFreniere and E. Yan, *Final work plan: Phase i investigation at bladen, nebraska*, Jul. 1, 2014. DOI: 10.2172/1172001

[66] A. Matei and N. Ianc, "Classification of underground mining works within the tg. ocna salt mine from the point of view of gas emissions," *MATEC Web of Conferences*, vol. 354, pp. 28–00 028, Jan. 6, 2022. DOI: 10.1051/matecconf/202235400028

[67] J. B. Lewis, "The use of the computer and other special tools for monitoring a gas well blowout during the kill operation - offshore louisiana," *SPE Annual Fall Technical Conference and Exhibition*, Oct. 9, 1977. DOI: 10.2118/6836-ms