Evaluation of Deep Learning Architectures for Local Flow Parameter Estimation in Gas–Liquid Two-Phase Flow Under Variable Conditions

Li Wei[1], Zhou Ming[2], and Han Rui[3]

[1] Liaoning University of Technology, Xueyuan Street, Jinzhou 121001, Liaoning, China
[2] Henan University of Science and Technology, Kaiyuan Avenue, Luolong District, Luoyang 471023, Henan, China
[3] Jiangxi University of Science and Technology, Ganzhou Avenue, Zhanggong District, Ganzhou 341000, Jiangxi, China

Abstract

Gas–liquid two-phase flows arise in pipelines, power plants, and process equipment where reliable prediction of local flow parameters is required for design, control, and safety. Local quantities such as phase volume fraction, phase velocity, and interfacial area are strongly affected by flow regime transitions and operating conditions, which makes empirical correlations difficult to generalize. High-fidelity computational fluid dynamics can resolve these fields, but the computational cost prohibits their use in real-time applications or large parametric studies. Recent progress in deep learning offers alternatives for learning low-cost surrogates directly from data collected by numerical simulation and experimental sensors. This work evaluates the ability of different deep learning architectures to infer local flow parameters from available measurements under variable gas and liquid flow rates, fluid properties, and pipe inclinations. Convolutional, recurrent, and attention-based models are trained on synthetically generated and experimentally inspired data that span stratified, slug, and annular flow conditions. The study compares prediction accuracy, robustness to measurement noise, and generalization to unseen operating conditions, and examines how architectural choices affect the reconstruction of small-scale interfacial features. Furthermore, the analysis explores the benefits of incorporating approximate physical constraints during training. The results highlight consistent trends in the relative performance of competing architectures and map regions in the operating space where data-driven estimators remain reliable, providing guidance for selecting models in practical monitoring and control systems.

1 Introduction

Gas–liquid two-phase flows appear in a wide range of industrial systems, including oil and gas production pipelines, nuclear reactor coolant circuits, chemical reactors, and refrigeration loops [1]. In these systems, design margins, operational safety, and efficiency depend on the ability to characterize local flow parameters such as phase volume fraction, phase velocity, turbulence intensity, interfacial area concentration, and pressure fluctuations. These local quantities control interfacial transfer of mass, momentum, and energy, and they are strongly influenced by flow regime, geometry, and operating conditions. Even in simple geometries like horizontal and inclined pipes, the complex interaction of buoyancy, inertia, surface tension, and wall friction yields a rich set of regimes, ranging from stratified and intermittent flows to dispersed bubbles and annular films. Each regime exhibits different spatial and temporal structures, which complicates the development of universal empirical correlations for local parameters.

Traditional modeling of gas–liquid two-phase flow relies on averaged conservation equations for each phase, together with closure relations for interfacial momentum transfer, turbulence, and phase distribution [2]. Drift-flux models and two-fluid models are widely used in one-dimensional system analysis, where the pipe cross-section is represented by area-averaged quantities. For local analysis, methods such as direct numerical simulation of interface-resolving equations and large-eddy simulations of Euler–Euler or Euler–Lagrange formulations have been employed to resolve detailed fields. While these approaches offer high fidelity, their computational requirements remain prohibitive for real-time estimation and online control across broad operating envelopes. The resulting gap between high-fidelity modeling and practical monitoring motivates the development of reduced-order models and data-driven surrogates that can approximate local fields from a limited set of accessible measurements.

Symbol	Description	Units
$k \in \{g, l\}$	Phase index for gas (g) and liquid (l)	–
α_k	Local volume fraction of phase k	–
ρ_k	Density of phase k	$\mathrm{kg\,m^{-3}}$
$\boldsymbol{u}_k$	Velocity field of phase k	$\mathrm{m\,s^{-1}}$
p	Pressure field	Pa
$\boldsymbol{M}_k$	Interfacial momentum exchange term	$\mathrm{N\,m^{-3}}$

Table 1: Key physical quantities in the two-phase flow formulation.

Dimensionless number	Definition	Physical role
Re_k	$\mathrm{Re}_k = \dfrac{\rho_k U L}{\mu_k}$	Inertial vs. viscous forces in phase k
We_k	$\mathrm{We}_k = \dfrac{\rho_k U^2 L}{\sigma}$	Inertia vs. surface tension
Fr	$\mathrm{Fr} = \dfrac{U}{\sqrt{gL}}$	Inertia vs. gravity
$\bar{\alpha}_g$	Cross-sectional average gas volume fraction	Regime classification and closure input
Inclination angle	Geometric angle of pipe relative to horizontal	Controls stratification and secondary flows

Table 2: Non-dimensional parameters controlling regime transitions and interfacial structures.

On the experimental side, techniques such as wire-mesh sensors, electrical capacitance tomography, electrical resistance tomography, gamma-ray tomography, and optical probes have enabled high-resolution measurements of phase distribution and interface dynamics [3]. However, the mapping from the available sensor signals to the underlying local flow parameters is often nonlinear, high-dimensional, and regime-dependent. Conventional reconstruction algorithms may require strong prior assumptions about the spatial distribution or may rely on regularization strategies that do not adapt well to changing flow conditions. At the same time, operators typically measure only a small subset of possible signals, such as pressure gradients, differential pressures, and global void fraction indicators, which further complicates the inversion from sparse measurements to fully resolved fields.

Deep learning offers a flexible framework for learning nonlinear mappings from measurements to flow parameters directly from data, without explicit specification of functional forms. Convolutional architectures can exploit spatial correlations in tomographic images, while recurrent and sequence models can exploit temporal correlations in sensor signals. Attention-based architectures can learn to focus on informative subsets of input features, potentially improving robustness under variable operating conditions [4]. Nevertheless, applying deep learning to gas–liquid two-phase flow presents several challenges. The available data sets may be limited in size and biased toward certain regimes. Changes in flow configuration and operating conditions can induce distribution shifts between training and deployment. In addition, purely data-driven models may violate basic physical constraints such as mass conservation, leading to inconsistent reconstructions and reduced extrapolation capability.

The present work investigates the performance of several deep learning architectures for estimating local flow parameters in gas–liquid two-phase flow under variable conditions [5]. The focus is on the reconstruction of cross-sectional fields of gas volume fraction, phase velocities, and derived quantities from synthetic sensor data generated by high-fidelity simulations and simplified virtual measurement models. The study considers a range of superficial velocities, fluid properties, and pipe inclinations that span multiple classical flow regimes. Convolutional encoder–decoder networks, recurrent models with temporal convolutions or gated units, and attention-based transformers are evaluated with respect to predictive accuracy, robustness to measurement noise, and generalization to unseen operating conditions. In addition, the study investigates the impact of incorporating approximate physical constraints into the learning objective, such as penalties on residuals of averaged conservation laws. By comparing these approaches on a common benchmark, the analysis aims to clarify the strengths and limitations of modern deep learning architectures in this application and to provide guidance for their integration with model-based approaches in two-phase flow analysis and monitoring.

2 Governing Equations and Physical Modeling

The underlying physics of gas–liquid two-phase flow in pipes can be described at the continuum level by separate conservation equations for each phase [6]. For a phase index $k \in \{g, l\}$, representing gas and liquid, the volume

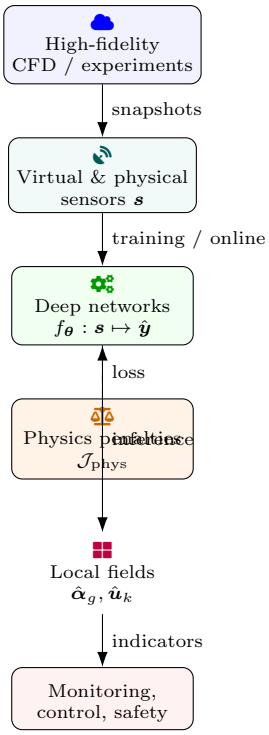

Figure 1: Overall workflow: high-fidelity simulations and experiments generate local flow fields that are mapped to synthetic and physical sensor signals. Deep networks trained with optional physics-based penalties infer local quantities from measurements for online monitoring and control.

fraction α_k, density ρ_k, and velocity field $\boldsymbol{u}_k$ satisfy a mass conservation equation of the form

$$\frac{\partial}{\partial t}\left(\alpha_k \rho_k\right) + \nabla \cdot \left([7]\alpha_k \rho_k \boldsymbol{u}_k\right) = \Gamma_k, \tag{1}$$

where Γ_k denotes volumetric mass transfer between phases due to processes such as phase change or dissolution. In many isothermal adiabatic gas–liquid flows without phase change, Γ_k is negligible and the equations reduce to conservative forms with constant densities. The momentum balance can be written as

$$\frac{\partial}{\partial t}\left(\alpha_k \rho_k \boldsymbol{u}_k\right) + \nabla \cdot \left([8]\alpha_k \rho_k \boldsymbol{u}_k \otimes \boldsymbol{u}_k\right) = -\alpha_k \nabla p + \nabla \cdot \left([9]\alpha_k \boldsymbol{\tau}_k\right) \tag{2}$$

$$+ \alpha_k \rho_k \boldsymbol{g} + \boldsymbol{M}_k + \boldsymbol{S}_k, \tag{3}$$

where p is the pressure field, $\boldsymbol{\tau}_k$ is the viscous stress tensor in phase k, $\boldsymbol{g}$ is the gravitational acceleration, $\boldsymbol{M}_k$ represents interfacial momentum exchange with the other phase, and $\boldsymbol{S}_k$ includes additional source terms such as wall friction and turbulence modeling contributions. The coupling between the phases arises through the common pressure field and the interfacial exchange terms, which depend on local interface geometry, relative velocity, and turbulence.

In interface-resolving formulations such as the volume-of-fluid method, a single momentum equation is solved with a variable phase indicator field, and surface tension forces are distributed over a finite thickness interface region. The phase indicator satisfies an advection equation for the volume fraction field,

$$\frac{\partial \alpha_g}{\partial t} + \nabla \cdot [10]\left(\alpha_g \boldsymbol{u}\right) = 0, \tag{4}$$

with $\alpha_l = 1 - \alpha_g$. In this setting, the local flow parameters of interest, such as gas volume fraction and phase velocities, can be derived from the resolved fields. However, the spatial and temporal resolution required to capture interfacial structures such as ligaments, droplets, and waves leads to fine computational grids and small time steps governed by capillary and inertial constraints. Direct use of such simulations for repeated evaluation in design or control remains computationally intensive, which motivates the development of surrogate models that approximate these local fields [11].

For many system-level analyses, cross-sectional average quantities are used. One-dimensional two-fluid models introduce area-averaged void fraction $\bar{\alpha}_g$ and phase velocities $\bar{u}_g$ and $\bar{u}_l$, along the streamwise coordinate z. Under

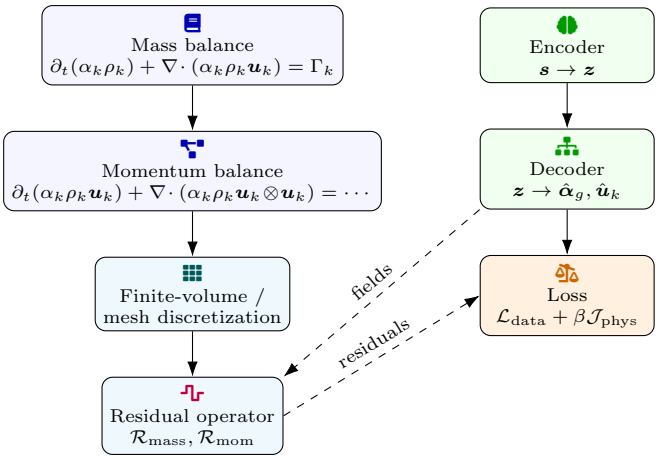

Figure 2: Relationship between continuum two-phase flow equations and the data-driven surrogate. Governing mass and momentum balances are discretized to define residual operators that act on network predictions. These residuals enter physics-based loss terms, encouraging consistency with conservation laws while the encoder–decoder network maps sensor data to local fields.

Architecture	Primary input type	Main characteristics	Typical advantage
Conv. encoder–decoder	Cross-sectional images / low-res tomography	Strided convs, skip connections, multiscale features	Accurate spatial reconstructions
Temporal CNN	Sensor time series	1D convs over time, parallel processing	Efficient capture of local temporal patterns
GRU / RNN	Sequential probes, pressures	Gated recurrence, hidden states	Long-term temporal correlations
Transformer	Multi-sensor sequences	Self-attention, learned focus on informative steps	Robust to variable temporal patterns
Graph decoder	Mesh-based fields	Message passing on unstructured meshes	Flexible cross-section representation
Multi-branch hybrid	Scalars + fields / sequences	Fuses global and local streams in latent space	Combines global constraints with local detail

Table 3: Deep learning architectures considered for local flow parameter estimation.

stationarity and suitable closure relations, the gas mass balance can be written as

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(A\bar{\alpha}_g\rho_g\bar{u}_g\right) = 0, \tag{5}$$

where A is the pipe cross-sectional area. In these averaged models, closure correlations relate wall and interfacial shear stresses to the mean velocities, and additional correlations approximate interfacial area concentration [12]. While these models are widely used, they do not provide direct access to local distributions within the cross-section and they require calibration of regime-dependent closure coefficients. The uncertainties in these closures often dominate the prediction error when operating away from calibrated conditions.

Local flow parameters can be defined as functions of space and time. For instance, the local gas volume fraction field in a pipe of circular cross-section can be written as

$$\alpha_g = \alpha_g\left([13]r, \theta, z, t\right), \tag{6}$$

where (r, θ) are polar coordinates in the cross-section. For data-driven estimation, it is convenient to discretize this field on a computational grid or an unstructured mesh, yielding discrete vectors that approximate the continuous fields. Let $\boldsymbol{\alpha}_g \in R^n$ denote the discrete void fraction at n spatial points, and correspondingly $\boldsymbol{u}_g$ and $\boldsymbol{u}_l$ for the velocities. The goal of local estimation in the present context is to reconstruct these high-dimensional vectors from a smaller set of sensor measurements $\boldsymbol{s} \in R^m$, where typically $m \ll n$. The mapping from $\boldsymbol{s}$ to $\boldsymbol{\alpha}_g$ and $\boldsymbol{u}_k$ is governed implicitly by the underlying PDE system, the geometry, and the sensor placement.

Non-dimensional analysis helps organize the influence of operating conditions on flow structure. Consider a characteristic length L such as the pipe diameter, a characteristic velocity U, and characteristic density and

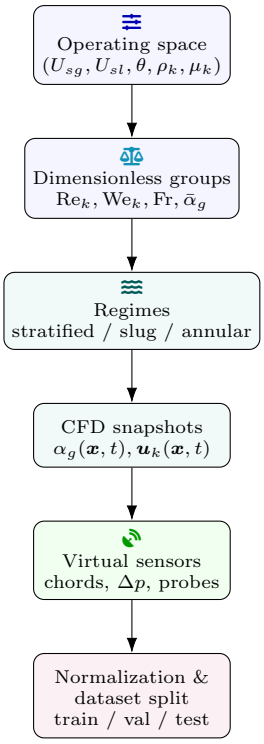

Figure 3: Numerical experiment pipeline: operating conditions are sampled in terms of superficial velocities, inclination, and fluid properties, mapped to dimensionless groups that organize classical flow regimes. High-fidelity simulations generate snapshots, which are processed by virtual sensors to emulate realistic measurements and then normalized and split into training, validation, and test sets.

Scenario	Split strategy	Evaluation focus
Random split	Mixed operating points in all subsets	Interpolation performance
Velocity hold-out	Reserve high U_g or U_l for testing	Extrapolation in superficial velocity
Inclination hold-out	Exclude extreme inclinations from training	Sensitivity to gravity orientation
Regime-based split	Train on e.g. stratified, test on slug	Cross-regime generalization
Noise-level split	Train on low noise, test on high noise	Robustness to sensor degradation

Table 4: Data partitioning scenarios used to probe generalization behaviour.

viscosity scales [14]. The Reynolds number for phase k is

$$\mathrm{Re}_k = \frac{\rho_k U L}{\mu_k}, \tag{7}$$

and the Weber number is

$$\mathrm{We}_k = \frac{\rho_k U^2 L}{\sigma[15]}, \tag{8}$$

where μ_k is the dynamic viscosity and σ is the surface tension. The Froude number characterizes the relative importance of inertia to gravity and can be written as

$$\mathrm{Fr} = \frac{U}{\sqrt{gL}}. \tag{9}$$

Combinations of these dimensionless numbers, together with void fraction and pipe inclination, influence regime transitions and interfacial structures. Local parameter estimation must remain reliable across variations in these quantities, which generally modify the mapping from global measurements to local fields [16].

An alternative viewpoint emphasizes the statistical nature of turbulent two-phase flow. The local void fraction field and velocities can be interpreted as random fields driven by stochastic fluctuations in interfacial position and

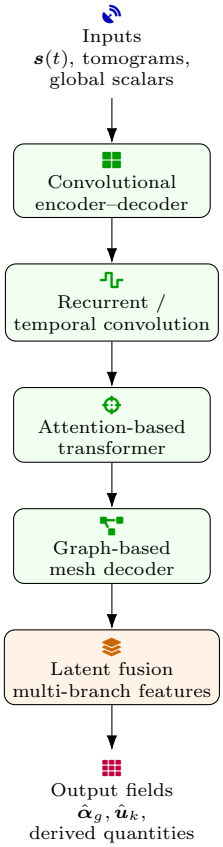

Figure 4: Architectural variants for local flow estimation. Input measurements feed a hierarchy of models that exploit spatial structure (convolutional encoder–decoders), temporal correlations (recurrent and temporal convolutional networks), adaptive feature weighting (attention mechanisms), and unstructured meshes (graph-based decoders). A fused latent representation is decoded to cross-sectional fields and derived quantities.

turbulence. Under ergodicity assumptions, ensemble averages over realizations can be replaced by time averages. The spatial covariance of the void fraction field,

$$C_\alpha\left(\boldsymbol{x}_1, \boldsymbol{x}_2\right)[17] = E\left[\alpha_g\left(\boldsymbol{x}_1\right)\alpha_g\left(\boldsymbol{x}_2\right)[18]\right] - \bar{\alpha}_g\left(\boldsymbol{x}_1\right)\bar{\alpha}_g\left(\boldsymbol{x}_2\right), \tag{10}$$

describes spatial correlations that can be exploited by learning architectures [19]. Convolutional layers in deep networks implicitly assume some form of local translation invariance in these correlations, which may be approximately valid in homogeneous directions of the flow. However, boundaries and gravity break full invariance, so model design must account for inhomogeneous structures such as stratified interfaces near the bottom of a horizontal pipe.

To link PDE-based modeling with data-driven estimation, one can consider physics-informed loss terms or constraints that penalize violations of conservation laws. For example, if $\hat{\boldsymbol{\alpha}}_g$ and $\hat{\boldsymbol{u}}_k$ are predicted fields from a network, discrete residuals of the averaged continuity equation can be constructed on a numerical grid. Let $\mathcal{R}_{\text{mass}}$ denote a vector of residuals obtained by applying a finite-volume discretization operator to the predicted fields. A physics-based penalty can then be defined as

$$\mathcal{J}_{\text{phys}} = \|[20]\mathcal{R}_{\text{mass}}\|_2^2 + \lambda\,\|\mathcal{R}_{\text{mom}}\|_2^2, \tag{11}$$

where $\mathcal{R}_{\text{mom}}$ are momentum residuals and λ is a weighting factor. This term can be combined with data misfit terms during training in order to encourage predictions that respect global conservation constraints, potentially improving generalization to conditions not represented in the training set.

The physical modeling perspective thus provides the context and constraints under which deep learning surrogates operate. While fully resolving the governing equations is often too costly for real-time applications, insights from these equations about invariants, symmetries, scale separation, and parameter dependencies can be embedded into network architectures, normalization strategies, and loss functions. The following sections build on these foundations to formulate deep learning models for local parameter estimation and to design numerical experiments that evaluate their performance under varying operating conditions [21].

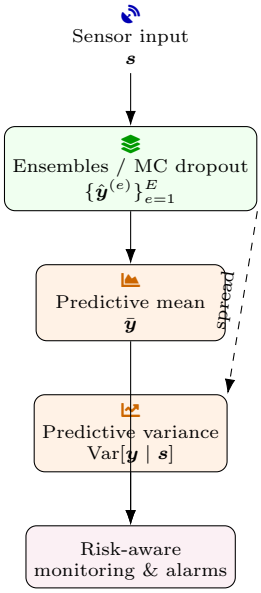

Figure 5: Uncertainty quantification concept: multiple stochastic realizations of the network, obtained via ensembles or Monte Carlo dropout, produce a distribution of local field estimates. From these samples, predictive means and variances are constructed to support calibrated, risk-aware decisions in monitoring and control.

Loss component	Representative form	Role in training	Tuning knob
Data fidelity	$\mathcal{L}_{\text{data}} = \frac{1}{N}\sum w_p \|\hat{\boldsymbol{y}}^{(p)} - \boldsymbol{y}^{(p)}\|_2^2$	Match predicted and reference fields	Weights w_p
Mass residual	$\|\mathcal{R}_{\text{mass}}\|_2^2$	Encourage discrete continuity	Physics weight β
Momentum residual	$\lambda\|\mathcal{R}_{\text{mom}}\|_2^2$	Promote momentum balance	Multiplier λ
Regularization	$\Omega(\boldsymbol{\theta})$ (e.g. weight decay)	Control overfitting	Decay factor
Bound penalties	Soft constraints on $\alpha_g \in [0,1]$, $k \geq 0$	Enforce physical ranges	Penalty strength

Table 5: Main loss components and their roles in the optimization problem.

3 Deep Learning Architectures for Local Flow Estimation

The task of local flow parameter estimation can be formulated as a supervised learning problem. Let $\boldsymbol{s} \in R^m$ denote the vector of sensor measurements, which may include pressure drops, differential pressures across segments, void fraction indicators, and potentially low-resolution tomographic images or time series from local probes. Let $\boldsymbol{y} \in R^d$ denote the target vector containing discretized local flow parameters, such as the gas volume fraction field $\boldsymbol{\alpha}_g$, phase velocities $\boldsymbol{u}_g$ and $\boldsymbol{u}_l$, and interfacial area concentration. The data set consists of pairs $(\boldsymbol{s}_i, \boldsymbol{y}_i)$ for samples indexed by $i = 1, \ldots, N$, generated from numerical simulations or experiments under different operating conditions. A deep learning model approximates a nonlinear mapping

$$f_{\boldsymbol{\theta}} : R^m \to R^d, \tag{12}$$

parametrized by network weights and biases collected in $\boldsymbol{\theta}$. The objective is to find $\boldsymbol{\theta}$ that minimizes a suitable empirical risk over the data set, while providing robust generalization to new inputs.

A standard choice for regression tasks is the mean-squared error loss. Defining the predicted output for sample i as [22] $\hat{\boldsymbol{y}}_i = f_{\boldsymbol{\theta}}(\boldsymbol{s}_i)$, the data fidelity loss can be expressed as

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i\|_2^2 . \tag{13}$$

In practice, the vector $\boldsymbol{y}_i$ may be composed of multiple physical fields with different scales and physical units, so normalization and weighting strategies are applied. Let P denote the number of distinct target components, for instance void fraction and phase velocities, and decompose $\boldsymbol{y}_i$ into subvectors $\boldsymbol{y}_i^{(p)}$ with corresponding predictions

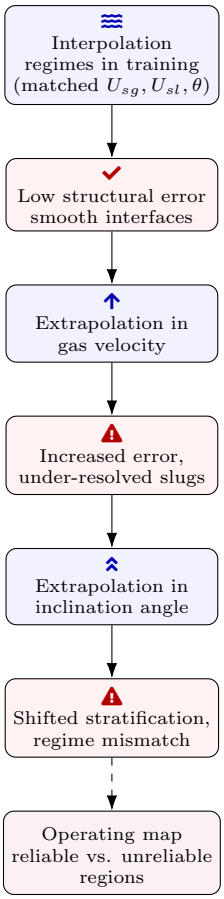

Figure 6: Schematic view of model performance across operating conditions. Errors remain small for interpolation within the training envelope but grow when gas velocity or inclination deviate significantly, altering flow regimes and interfacial structures. Aggregating error statistics across such scenarios defines an operating map that delineates regions where data-driven estimators provide reliable local reconstructions.

$\hat{\boldsymbol{y}}_i^{(p)}$. A weighted loss can be written as

$$\mathcal{L}_{\text{data}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{p=1}^{P} w_p \left\| [23] \hat{\boldsymbol{y}}_i^{(p)} - \boldsymbol{y}_i^{(p)} \right\|_2^2, \tag{14}$$

where the weights w_p control the relative importance of each component in training.

Convolutional neural networks are well suited to reconstructing spatial fields from image-like inputs. When the sensors provide cross-sectional tomographic images, such as low-resolution void fraction maps, a convolutional encoder–decoder architecture can be used to infer higher-resolution fields. Let $\boldsymbol{X} \in R^{H \times W \times C}$ represent an input tensor with spatial dimensions $H \times W$ and C channels, and let $\boldsymbol{Y} \in R^{H' \times W' \times C'}$ represent the output tensor of local parameters on a potentially finer grid. A convolutional layer with kernel weights $K^{(l)}$ and bias $\boldsymbol{b}^{(l)}$ computes

$$\boldsymbol{Z}^{(l)} = \phi \left(\text{conv} \left(\boldsymbol{X}^{(l)}, K^{(l)} \right) [24] + \boldsymbol{b}^{(l)} \right), \tag{15}$$

where ϕ is a nonlinear activation function, such as a rectified linear unit, and $\boldsymbol{X}^{(l)}$ is the input feature map at layer l. In an encoder–decoder structure, downsampling operations reduce spatial resolution while increasing channel depth, and upsampling operations reconstruct the original or higher resolution. Skip connections between corresponding encoder and decoder levels preserve multiscale information and help reconstruct fine interfacial structures such as thin films and small bubbles.

When sensor data are primarily temporal, for example pressure signals and void fraction measurements collected at fixed locations along the pipe, recurrent architectures can exploit temporal correlations. Let $\boldsymbol{s}_{i,t}$ denote the feature vector at time index t for sample i. A gated recurrent unit layer with hidden state $\boldsymbol{h}_t$ can be written in

Case	Noise model	Relative standard deviation	Purpose
Clean	$\tilde{s}_j = s_j^{\text{true}}$	0%	Baseline accuracy
Low noise	$\tilde{s}_j = s_j^{\text{true}} + \epsilon_j$	1–2% of s_j^{true}	Mild measurement uncertainty
Moderate noise	Additive Gaussian + small multiplicative term	5% range	Typical experimental levels
High noise	Strong additive noise on all channels	$> 10\%$ range	Stress-test robustness
Mismatched	Train on clean, test on noisy	Varies	Diagnose overfitting to ideal data

Table 6: Sensor noise configurations used in robustness experiments.

Metric	Symbol	Type	Interpretation
Mean-squared error	MSE	Pointwise regression	Penalizes large local deviations
Mean-absolute error	MAE	Pointwise regression	Robust to outliers
Normalized error	–	Scale-adjusted	Compares across fields / units
Spatial correlation	$\rho(x)$	Structural	Agreement of spatial patterns
Integral quantity error	–	Scalar summary	Error in $\bar{\alpha}_g$, Δp, etc.
Conservation residual	–	Physics consistency	Mass / momentum balance satisfaction

Table 7: Performance metrics used for quantitative comparison of models.

terms of update and reset gates,

$$\boldsymbol{z}_t = \sigma\left([25]W_z \boldsymbol{s}_t + U_z \boldsymbol{h}_{t-1} + \boldsymbol{b}_z\right), \tag{16}$$

$$\boldsymbol{r}_t = \sigma\left(W_r \boldsymbol{s}_t + U_r \boldsymbol{h}_{t-1} + \boldsymbol{b}_r\right), \tag{17}$$

$$\tilde{\boldsymbol{h}}_t = \tanh\left([26]W_h \boldsymbol{s}_t + U_h\left(\boldsymbol{r}_t \odot \boldsymbol{h}_{t-1}\right) + \boldsymbol{b}_h\right), \tag{18}$$

$$\boldsymbol{h}_t = (1 - \boldsymbol{z}_t) \odot \boldsymbol{h}_{t-1} + \boldsymbol{z}_t \odot \tilde{\boldsymbol{h}}_t, \tag{19}$$

where σ denotes the logistic sigmoid, $\odot$ is the elementwise product, and W and U matrices are trainable parameters [27]. The final hidden state, or an aggregation over time, can be used as a latent representation from which local flow fields are decoded using fully connected or convolutional layers. Temporal convolutions provide an alternative with parallelizable operations by applying one-dimensional convolutions across the time axis, capturing local temporal patterns in the signals.

Attention-based architectures extend this idea by enabling the model to learn pairwise interactions between time steps or spatial locations. For a sequence of embedded inputs $\boldsymbol{e}_t$, the self-attention mechanism constructs queries, keys, and values,

$$\boldsymbol{q}_t = W_Q \boldsymbol{e}_t, \tag{20}$$

$$\boldsymbol{k}_t = W_K \boldsymbol{e}_t, \tag{21}$$

$$\boldsymbol{v}_t = W_V \boldsymbol{e}_t, \tag{22}$$

and computes attention weights

$$a_{tt'} = \frac{[28]\exp\left(\boldsymbol{q}_t^\top \boldsymbol{k}_{t'}/\sqrt{d_k}\right)}{\sum_j \exp\left(\boldsymbol{q}_t^\top \boldsymbol{k}_j/\sqrt{d_k}\right)}, \tag{23}$$

with output representations [29]

$$\boldsymbol{o}_t = \sum_{t'} a_{tt'} \boldsymbol{v}_{t'}. \tag{24}$$

This mechanism allows the network to focus adaptively on informative time steps or sensor channels, which may be important in flows where intermittent structures such as slugs and large waves dominate the signals.

Method	Uncertainty type	Output	Extra cost
Ensemble of networks	Epistemic	Sample mean and variance fields	High (multiple models)
MC dropout	Epistemic + aleatoric	Stochastic forward-pass statistics	Moderate
Gaussian likelihood	Aleatoric	Per-pixel variance parameter	Low
Temperature scaling	Calibration	Rescaled predictive spread	Negligible
Reliability analysis	Diagnostic	Coverage vs. nominal plots	Post-processing only

Table 8: Approaches for quantifying and calibrating predictive uncertainty.

Architecture	Main strengths	Main limitations
Conv. encoder–decoder	Accurate cross-sectional reconstructions in seen regimes; efficient inference	Interface smoothing; reduced extrapolation fidelity
Temporal CNN / GRU	Captures slug signatures and intermittent dynamics from time series	Sensitive to sequence length and unseen frequencies
Transformer	Learns to focus on informative events; good with multi-sensor signals	Higher computational cost; attention degradation under heavy noise
Physics-informed variant	Better global conservation and integral quantities in extrapolation	May trade local detail for constraint satisfaction
Small-capacity model	More stable in low-data regimes	Higher bias; limited representation of complex structures

Table 9: Qualitative comparison of architectural behaviour across operating conditions.

Graph-based architectures are also relevant because computational meshes for local fields are often unstructured. A mesh can be represented as a graph with nodes corresponding to spatial locations and edges connecting neighboring cells. A graph convolution layer can update node features $\boldsymbol{h}_i$ according to

$$\boldsymbol{h}_i^{(l+1)} = \sigma\left(W_0\boldsymbol{h}_i^{(l)} + \sum_{j\in\mathcal{N}(i)} W_1\boldsymbol{h}_j^{(l)} + \boldsymbol{b}\right), [30] \tag{25}$$

where $\mathcal{N}(i)$ denotes the set of neighbors of node i. Such layers can approximate discrete diffusion and propagation of information along the mesh, aligning well with the physical couplings in discretized PDE systems. In the present study, graph convolutions are used to post-process latent representations into fields defined on irregular cross-sectional meshes for some configurations.

Regardless of architecture, the optimization problem solved during training can be expressed as

$$\min_{\boldsymbol{\theta}} \ \mathcal{L}_{\text{data}}\left(\boldsymbol{\theta}\right)[31] + \beta\mathcal{J}_{\text{phys}}\left(\boldsymbol{\theta}\right) + \Omega\left(\boldsymbol{\theta}\right), \tag{26}$$

where β controls the contribution of physics-based penalties and Ω denotes regularization terms such as weight decay [32]. Stochastic gradient-based methods are employed to approximate solutions to this optimization problem. At each iteration, a mini-batch of samples is drawn and gradients of the loss with respect to $\boldsymbol{\theta}$ are estimated using backpropagation. The parameters are updated according to

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \eta_k\widehat{\nabla}_{\boldsymbol{\theta}}\mathcal{L}\left(\boldsymbol{\theta}^{(k)}\right), \tag{27}$$

with learning rate η_k and stochastic gradient estimate $\widehat{\nabla}_{\boldsymbol{\theta}}\mathcal{L}$. Learning rate schedules and normalization layers are selected to ensure stable training across the range of operating conditions [33].

From a linear algebra perspective, the forward pass of a deep network can be viewed as a composition of affine transformations and elementwise nonlinearities. For a fully connected layer with weight matrix $W^{(l)}$ and bias vector $\boldsymbol{b}^{(l)}$, the operation on an input vector $\boldsymbol{h}^{(l)}$ is

$$\boldsymbol{h}^{(l+1)} = \sigma\left(W^{(l)}\boldsymbol{h}^{(l)} + \boldsymbol{b}^{(l)}\right). \tag{28}$$

When stacking many such layers, the overall mapping from inputs to outputs can, in principle, approximate high-dimensional nonlinear functions with controlled complexity. Convolutional layers replace dense matrices with structured sparse operators that share weights spatially, which reduces the number of parameters and encodes a form of translational invariance [34]. The expressivity of the architectures considered here is governed by depth, width, and kernel supports, balanced against the amount of available data and the risk of overfitting.

In the specific context of gas–liquid two-phase flow, the relevant architectures must capture both local spatial patterns and global context. Stratified flows feature large-scale separation between phases with comparatively smooth interfaces, whereas slug and churn flows exhibit three-dimensional structures that can be challenging to reconstruct from sparse measurements. To address this, multi-branch architectures are considered, where one branch processes global features such as overall pressure drop and bulk flow rates through fully connected layers, and another branch processes spatial or temporal sensor fields through convolutions or recurrent units. The outputs of these branches are concatenated in a latent space and decoded into local parameter fields, allowing the model to combine local and global information [35]. The following section describes how data sets are generated and how these architectures are instantiated and trained for quantitative evaluation.

4 Numerical Experiments and Data Generation

A central component of the evaluation is the construction of data sets that span a representative range of flow conditions and regimes. High-fidelity numerical simulations are used as a primary source of local flow fields, supplemented by simplified virtual measurement models that emulate plausible sensor configurations. The simulations are based on a two-fluid or volume-of-fluid formulation of incompressible gas–liquid flow in pipes with circular cross-section. For each configuration, gas and liquid superficial velocities, pipe inclination angles, and fluid properties are sampled over predefined ranges chosen to cover stratified, slug, and annular flow conditions. Steady and statistically steady unsteady simulations are conducted until the flow reaches a regime that exhibits stationary statistical properties, after which snapshots of the local fields are recorded at specified time intervals [36].

The computational domain is discretized using either structured or unstructured meshes with sufficient resolution to capture key interfacial features. Time integration employs schemes that satisfy stability conditions associated with capillary and inertial time scales. For example, the time step Δt is constrained by a Courant number condition of the form

$$\mathrm{CFL} = \max\left(\frac{[37]\,\|\boldsymbol{u}\|\,\Delta t}{\Delta x}\right) \leq \mathrm{CFL}_{\max}, \tag{29}$$

and, in interface-resolving simulations, by a capillary time step condition [38]

$$\Delta t \leq c_\sigma \sqrt{\frac{\rho \Delta x^3}{\sigma}}, \tag{30}$$

with typical constants determined from stability analysis [39]. The resulting simulation data provide sequences of void fraction, velocity, and pressure fields over time.

To generate training inputs that resemble practical measurements, virtual sensors are placed in the computational domain. For instance, cross-sectional void fraction distributions can be integrated along chords to emulate gamma-ray or electrical tomography signals. The integrated signal along chord index j can be written as

$$s_j = \frac{1}{L_j} \int_{C_j} \alpha_g\,(\boldsymbol{x})\,[40]\,\mathrm{d}\ell, \tag{31}$$

where C_j denotes the path of the chord and L_j its length. Pressure sensors located along the pipe measure local or averaged pressure values, which are derived from the simulated pressure field and phase distributions. Differential pressures between specified locations approximate common instrumentation used in industrial pipelines. These virtual sensors are arranged in configurations that vary in density and coverage, allowing assessment of how sensor layout influences the identifiability of local parameters.

Measurement noise is modeled to reflect realistic uncertainties. Additive Gaussian noise with specified standard deviation is applied to each sensor channel, and, for some experiments, multiplicative noise components are added to reflect calibration uncertainties [41]. If s_j^{true} denotes a noise-free simulated signal, the noisy measurement is defined as

$$\tilde{s}_j = s_j^{\mathrm{true}} + \epsilon_j, \tag{32}$$

where ϵ_j is drawn from a normal distribution with zero mean and variance σ_j^2 proportional to the magnitude of the signal. In experiments involving robustness analysis, noise levels are varied up to levels representing several percent of the signal magnitude, and networks are trained and evaluated at multiple noise intensities.

For data-driven models, normalization of inputs and outputs is essential. Sensor signals are standardized to zero mean and unit variance based on statistics computed from the training set. Local field quantities such as void fraction and velocities are scaled to fixed ranges, for example by mapping void fraction to the interval $[0, 1]$ and velocities to nondimensional forms using characteristic velocity scales. Let $\boldsymbol{\mu}_s$ and $\boldsymbol{\sigma}_s$ denote the mean and standard deviation vectors of the sensor data in the training set. Normalized inputs are computed as

$$\boldsymbol{s}^{\mathrm{norm}} = \frac{[42]\boldsymbol{s} - \boldsymbol{\mu}_s}{\boldsymbol{\sigma}_s}, \tag{33}$$

with analogous transformations for outputs. These transformations are inverted when interpreting predictions in physical units.

The data set is partitioned into three subsets for training, validation, and testing. The partitioning strategy plays an important role in assessing generalization to unseen conditions. In one scenario, random sampling is used so that all operating conditions appear across subsets, and the test set primarily reflects interpolation performance [43]. In another scenario, specific ranges of superficial velocities, inclination angles, or fluid properties are withheld entirely from training and used only for testing, thereby probing extrapolation across operating conditions. Flow-regime-based splits are also considered, in which networks are trained on selected regimes and evaluated on others. These splitting strategies probe whether the architectures learn regime-specific or more generalizable representations.

The network architectures introduced previously are instantiated with hyperparameters such as depth, width, kernel size, and latent dimension. For convolutional encoder–decoder models, multiple depth levels are used, with downsampling via strided convolutions [44]. The number of filters per layer increases with depth in the encoder and decreases symmetrically in the decoder. Activation functions, normalization layers, and dropout rates are chosen based on preliminary experiments that check training stability and avoid overfitting. Recurrent and attention-based models are configured with sequence lengths corresponding to windows of sensor data, selected to capture several characteristic time scales of the flow, such as slug passage periods or wave propagation times.

Training is performed using mini-batch stochastic optimization. The batch size is chosen to balance computational efficiency and gradient estimation variance. Learning rates and schedules, including constant, stepwise, and cosine annealing strategies, are investigated [45]. Gradient clipping is applied in some configurations to mitigate occasional instabilities arising from rare high-gradient samples. Early stopping based on validation loss is employed to determine the number of training epochs and to avoid overfitting, particularly in parameter ranges with relatively sparse data. Checkpoints of network parameters at different epochs are stored for later analysis of convergence behavior and for ensemble-based uncertainty quantification.

As a final ingredient, physics-based penalties and constraints are incorporated for selected models. For example, the predicted void fraction and velocity fields are inserted into discrete mass and momentum balance equations to compute residuals [46]. These residuals are included in the loss function with tunable weights, as described earlier. In some configurations, constraints enforcing bounds on void fraction in the range $[0, 1]$ and enforcing non-negativity of predicted turbulent kinetic energy are implemented via appropriate activation functions or penalty terms. The comparative evaluation across models with and without such physics-based components allows assessment of their impact on accuracy and robustness across varying flow conditions.

5 Results and Discussion

The performance of the deep learning architectures is assessed using several quantitative metrics computed on the test data sets. For each predicted local field, the mean-squared error and mean-absolute error relative to the reference high-fidelity solution are evaluated. Let $\hat{y}_{i,j}$ denote the predicted value of component j of sample i and $y_{i,j}$ the corresponding reference value. The mean-squared error for a given field is defined as [47]

$$\mathrm{MSE} = \frac{1}{Nd} \sum_{i=1}^{N} \sum_{j=1}^{d} \left(\hat{y}_{i,j} - y_{i,j} \right)^2, \tag{34}$$

and the mean-absolute error is defined analogously. To interpret performance across fields with different scaling, normalized errors are computed by dividing by the variance of the reference field. Spatial correlation coefficients between predicted and reference fields are also calculated to capture structural agreement. For scalar summary quantities derived from the fields, such as cross-sectional averaged void fraction or pressure drop, relative errors are reported [48].

In scenarios where the training and test sets share the same ranges of operating conditions, the architectures demonstrate varying levels of interpolation performance. Convolutional encoder–decoder models tend to achieve low errors in reconstructing cross-sectional void fraction distributions when supplied with sufficient tomographic input information. They capture primary features such as stratified interfaces, core–annular structures, and gross slug shapes. However, they sometimes smooth sharp interfacial regions and underpredict small-scale structures

such as droplets and ligaments, particularly when trained with strong regularization or when the training data lack high-resolution interfacial detail. Multi-branch architectures that combine global scalar inputs with image-based branches often yield improved performance on integral quantities while maintaining comparable local field accuracy.

Recurrent and temporal convolutional models provide advantages when temporal sensor sequences are available [49]. For flows with pronounced intermittent structures, such as slug flow, these models learn temporal signatures that correlate with phase distribution patterns. For example, the passage of a gas slug produces characteristic fluctuations in differential pressure and void fraction signals that the network can associate with particular cross-sectional distributions. Attention-based models further refine this capability by assigning higher weights to time steps containing informative fluctuations and lower weights to relatively quiescent periods. Quantitative results show that attention-based sequence models reduce errors in predicting the spatial extent and position of slugs along the pipe under conditions that match the training data distribution.

The evaluation across architectures also considers sensitivity to measurement noise [50]. For each noise level configuration, models are trained and tested on data with the same noise statistics, and, in additional experiments, models trained on noiseless data are applied to noisy test data. As noise increases up to several tens of % of the signal standard deviation, models without explicit noise handling exhibit a gradual increase in error and a tendency to overfit noise patterns when trained on small data sets. Architectures that include explicit regularization, such as dropout and weight decay, and those trained with augmentations that randomly perturb sensor signals, display improved robustness. In attention-based models, noise can degrade the reliability of attention weights, causing the model to focus on spurious fluctuations. This effect is mitigated by input normalization, smoothing of attention distributions, and training strategies that expose the model to a range of noise levels.

An important aspect of the evaluation is generalization to operating conditions not represented in the training set [51]. When models trained at lower gas superficial velocities are tested at higher velocities that produce more vigorous interfacial structures, prediction errors for all architectures increase. Convolutional encoder–decoder models trained on regimes with relatively smooth interfaces may underpredict void fraction fluctuations and over-smooth structures in more turbulent regimes. Sequence-based models that rely on temporal patterns may misinterpret high-frequency fluctuations if such patterns did not appear in the training data. Architectures augmented with physics-based penalties show somewhat improved extrapolation performance in these scenarios. By penalizing violations of global mass conservation, these models tend to maintain more accurate overall void fraction even when local structures differ from training examples, although detailed interfacial patterns remain difficult to reconstruct [52].

The effects of pipe inclination and gravity orientation provide another test of generalization. Networks trained primarily on horizontal configurations are evaluated on inclined flows where stratification patterns shift and buoyancy-induced secondary flows change phase distribution in the cross-section. Architectures that rely heavily on spatial symmetries learned from horizontal configurations may not generalize directly. Introducing inclination angles as additional input features and including a subset of inclined cases in training improves performance but does not fully eliminate errors at extreme inclinations. The sensitivity of intermediate network representations to inclination suggests that embedding appropriate rotational or reflectional invariances into the architecture, or using coordinate transformations aligned with gravity, could further improve generalization.

From a physics perspective, an informative diagnostic is the evaluation of conservation law residuals on the predicted fields [53]. For each predicted void fraction and velocity field, discrete mass and momentum balance residuals are computed using finite-volume operators. Aggregate measures of residual magnitude, such as volume-integrated squared residuals, are used to quantify the degree to which predictions satisfy the governing equations. Models trained with physics-based penalties exhibit systematically lower residual magnitudes on both training and test data, indicating that they have internalized some aspects of the conservation structure. However, reductions in residuals do not always directly translate into lower pointwise prediction errors, particularly in strongly nonlinear regimes where multiple local configurations can satisfy approximate conservation. Balancing data fidelity and residual minimization remains a design choice that influences both accuracy and physical consistency [54].

The computational cost of each architecture is also characterized, including training time and inference time per sample. Convolutional encoder–decoder models with moderate depth generally offer favorable trade-offs, achieving reliable accuracy with inference times suitable for near real-time applications on modern hardware. Sequence models and transformers introduce additional computational cost that scales with sequence length and model width. Graph-based decoders incur overhead from sparse matrix operations on unstructured meshes, but benefit from mesh flexibility. For applications where latency is critical, such as online monitoring or control, architectures with fewer parameters or compressed versions of larger models may be preferred, even at the expense of some accuracy. Techniques such as model pruning, quantization, and knowledge distillation are potential strategies for reducing computational load while retaining predictive performance [55].

The spatial structure of prediction errors provides further insight into model behavior. Error fields often show concentration near interfaces, especially in regions with steep gradients or rapidly changing topology. In stratified regimes, errors may be largest near the contact line where the liquid film meets the pipe wall and gas

core, reflecting sensitivity to small geometric changes. In slug flow, errors can be localized near the front and tail of slugs, where velocity and void fraction gradients are strong. In annular flow, thin liquid films on the wall and entrained droplets are challenging to reconstruct, leading to underprediction of film thickness and misrepresentation of droplet concentration [56]. These patterns highlight the need for architectures and loss functions that pay particular attention to interfacial regions, possibly through adaptive weighting or specialized sub-networks focused on interface refinement.

Finally, the interaction between data set size and model capacity is examined. As training data are reduced, larger models tend to overfit and show increased variance in test performance across different initializations. Smaller models with fewer parameters may generalize better in low-data regimes, albeit with higher bias. Physics-based penalties and regularization techniques can partially compensate for limited data, effectively injecting prior knowledge into the learning process [57]. However, there remains a trade-off between flexible architectures that can capture complex flow patterns and the risk of overfitting in regimes with sparse data. These observations suggest that practical deployment should consider both the anticipated range of operating conditions and the availability of representative training data.

6 Uncertainty Quantification and Statistical Evaluation

Beyond point estimates of local flow parameters, quantifying the uncertainty of deep learning predictions is important for risk-aware decision making in monitoring and control. Several approaches are employed to characterize predictive uncertainty. One approach uses ensembles of networks trained with different random initializations and, in some cases, different subsets of the training data. For a given input $\boldsymbol{s}$, the ensemble provides a collection of predictions $\hat{\boldsymbol{y}}^{(e)}$ indexed by ensemble member e. The empirical mean and variance across the ensemble can be used to approximate epistemic uncertainty associated with model parameters and limited data [58]. The predictive mean is computed as

$$\bar{\boldsymbol{y}} = \frac{1}{E} \sum_{e=1}^{E} \hat{\boldsymbol{y}}^{(e)}, \tag{35}$$

and the sample covariance provides a measure of dispersion. Spatial maps of predictive variance highlight regions where the model is less confident, often correlating with interfacial zones and extrapolation in operating conditions.

Another approach employs approximate Bayesian techniques, such as Monte Carlo dropout, where dropout layers are kept active during inference. For each stochastic forward pass, a different subset of network connections is dropped according to specified probabilities, yielding a distribution of outputs for the same input [59]. This approach approximates Bayesian model averaging under certain assumptions. The predicted variance can be decomposed into components arising from observation noise and model uncertainty. For a scalar quantity of interest at location index j, the total predictive variance under a Gaussian likelihood assumption can be approximated as

$$\mathrm{Var}\left[y_j \mid \boldsymbol{s}\right] \approx [60] \frac{1}{E} \sum_{e=1}^{E} \left(\hat{y}_j^{(e)} - \bar{y}_j\right)^2 + \sigma_{\mathrm{noise}}^2, \tag{36}$$

where $\sigma_{\mathrm{noise}}^2$ represents an estimate of observation noise variance. In regions of the operating space where training data are dense, the epistemic component tends to be small, and uncertainty is dominated by measurement noise. In extrapolation regions, epistemic uncertainty increases, reflecting reduced confidence in predictions.

Calibration of uncertainty estimates is crucial. A prediction interval is considered well calibrated if the empirical frequency with which true values fall inside the interval matches the nominal coverage probability [61]. Calibration is assessed by constructing prediction intervals at specified nominal coverage levels and computing the fraction of test samples whose true values lie within them. Reliability diagrams plotting empirical coverage versus nominal coverage provide a visual diagnostic. Overconfident models produce intervals that are too narrow, leading to empirical coverages below nominal levels, while underconfident models produce excessively wide intervals. Temperature scaling and other post-processing techniques can be applied to adjust uncertainty estimates without changing mean predictions.

Statistical hypothesis testing is used to assess whether observed differences in performance metrics across architectures are significant. For each test case, metrics such as mean-squared error or structural similarity index are computed [62]. Paired tests comparing two architectures use the fact that predictions are generated on the same set of test samples. Let d_k denote the difference in metric value for sample k between two models. The sample mean and variance of d_k are used to construct test statistics under assumptions of approximate normality of differences. When distributional assumptions are doubtful, nonparametric tests based on ranks, such as Wilcoxon-type statistics, can be employed. These tests indicate whether one architecture consistently outperforms another across the operating space, beyond fluctuations due to finite sample sizes.

Error decomposition into bias and variance components provides further insight. For a scalar predicted quantity y at a given location and operating condition, the expected squared error can be decomposed as [63]

$$E\left[(\hat{y}-y)^2\right] = \mathrm{Bias}^2 + \mathrm{Var} + \sigma^2_{\mathrm{noise}}, \tag{37}$$

where Bias measures the systematic deviation of the mean prediction from the true value, Var is the variance of predictions due to model randomness, and $\sigma^2_{\mathrm{noise}}$ is the irreducible noise variance. Estimation of these quantities requires multiple independent training runs or approximations via ensembles. In practice, high-capacity models trained with ample data tend to exhibit low bias but higher variance, while constrained models may show higher bias but lower variance. Physics-informed regularization can reduce bias in directions constrained by conservation laws, but may also increase variance if the penalty introduces additional sensitivity to modeling errors in the physics [64] [65].

Spatially resolved uncertainty maps reveal where predictions are most uncertain. These maps often align with regions of high physical variability or limited training coverage. In stratified flows, uncertainty may be highest near the interface, reflecting sensitivity to small changes in level that strongly affect local void fraction. In slug flows, uncertainty may be concentrated near transient fronts where local flow topology changes quickly. Understanding these patterns can inform sensor placement strategies; for example, placing additional sensors in high-uncertainty regions may reduce overall uncertainty more efficiently than uniform sensor distributions [66].

In addition to pointwise metrics, distributional comparisons between predicted and reference fields are performed. Empirical cumulative distribution functions of void fraction and velocity at selected locations are constructed, and divergence measures between predicted and reference distributions are computed. Probability integral transform diagnostics evaluate whether predictive distributions, when transformed by the reference cumulative distribution function, approximate uniform distributions. Deviations indicate miscalibration or mismatch in distributional shape. For flow regimes where the statistical behavior of local parameters has operational importance, such as where maxima or exceedance probabilities matter, such distributional assessments complement mean-square-based metrics.

Overall, the uncertainty quantification and statistical evaluation highlight both the capabilities and limitations of deep learning architectures in this application [67]. While point estimates can achieve useful levels of accuracy across a range of conditions, uncertainties may be substantial in extrapolation regimes and in regions with complex interfacial dynamics. Ensemble and approximate Bayesian methods provide avenues for representing this uncertainty, though they increase computational cost. Calibration and statistical testing ensure that reported uncertainties and performance differences are interpreted appropriately, supporting cautious use of these models in practical settings.

7 Conclusion

This work has examined the use of deep learning architectures for local flow parameter estimation in gas–liquid two-phase flow under variable operating conditions. Starting from the governing equations of two-phase flow, the study has outlined how local quantities such as gas volume fraction and phase velocities relate to the underlying conservation laws and how they can be represented on discrete computational meshes. Deep learning models are used to approximate the nonlinear mapping from sparse or aggregated sensor measurements to high-dimensional local fields, leveraging architectures that exploit spatial, temporal, and graph-structured correlations in the data. The models are trained and evaluated on data sets generated from high-fidelity simulations with virtual sensors, covering a range of superficial velocities, pipe inclinations, and flow regimes.

The results indicate that convolutional encoder–decoder networks can reconstruct many salient features of cross-sectional phase distributions when provided with sufficiently informative inputs, particularly in regimes similar to those represented in the training data. Sequence-based and attention-based models offer advantages in capturing temporal signatures associated with intermittent structures, improving predictions in slug and churn flows when temporal sensor data are available. Architectures that incorporate physics-based penalties on conservation law residuals show improvements in preserving global consistency, especially when extrapolating to operating conditions beyond the training set, although detailed reconstruction of interfacial structures remains challenging in strongly nonlinear regimes.

The analysis of sensitivity to measurement noise and of generalization across operating conditions underscores the importance of both data diversity and model design [68]. Networks trained on narrow ranges of velocities or inclinations may struggle when confronted with substantially different flow conditions, and models without explicit regularization may overfit noise when data are limited. Physics-informed training, careful normalization, and augmentation strategies can mitigate some of these effects, but they do not eliminate the fundamental dependence on representative training data. The spatial and temporal patterns of prediction errors, often concentrated near interfaces and rapidly evolving structures, point toward potential benefits of specialized treatments of interfacial regions in network architectures and loss functions.

Uncertainty quantification and statistical evaluation provide a complementary view of model performance. Ensemble methods and approximate Bayesian techniques yield spatially resolved estimates of predictive uncertainty, which tend to be larger in extrapolation regions and in zones with complex interfacial dynamics [69]. Calibration analyses and hypothesis tests help distinguish genuine differences between architectures from variations due to finite sample sizes and stochastic training. These tools are useful for assessing the reliability of predictions and for guiding the cautious deployment of deep learning surrogates in operational environments.

Future developments may focus on closer integration between physics-based and data-driven approaches. Strategies include embedding reduced-order models derived from the governing equations into network architectures, using simulation data to pre-train models before fine-tuning on experimental measurements, and applying domain adaptation techniques to handle shifts in flow conditions and sensor characteristics. Extensions to three-dimensional geometries, non-Newtonian fluids, and flows with phase change are also of interest. While challenges remain, particularly in extrapolation and uncertainty characterization, the findings obtained here provide a basis for further exploration of deep learning as a component of multi-fidelity frameworks for analysis and monitoring of gas–liquid two-phase flows under variable conditions [70].

References

[1] M. Yamamoto, M. Murai, K. Maeda, and S. Uto, "An experimental study of the interaction between pipe structure and internal flow," in *Volume 3: Pipeline and Riser Technology*, ASMEDC, Jan. 1, 2009, pp. 213–220. DOI: 10.1115/omae2009-79312

[2] D. Chandrasekharam and A. Baba, "Carbon dioxide emissions mitigation strategy through enhanced geothermal systems: Western anatolia, turkey.," *Environmental earth sciences*, vol. 81, no. 8, pp. 235–, Apr. 7, 2022. DOI: 10.1007/s12665-022-10345-5

[3] K. Witwer, *Test report for core drilling ignitability testing*, Aug. 8, 1996. DOI: 10.2172/657933

[4] 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.

[5] J. Zhang and D. Li, "Research on optimization of large flow hydraulic punching process in zhaozhuang mine," *Geofluids*, vol. 2022, pp. 1–14, Dec. 29, 2022. DOI: 10.1155/2022/7535215

[6] H. N. Patel, J. Bruton, E. Dietrich, and R. Prince-Wright, "Lessons learned and safety considerations for installation and operation of a managed pressure drilling system on classed floating drilling rigs," in *Offshore Technology Conference*, OTC, May 4, 2015. DOI: 10.4043/25946-ms

[7] K. Y, "Statistical and numerical investigation of the effect of wellbore trajectories in williston basin horizontal wells and their effects on production performance," *Petroleum & Petrochemical Engineering Journal*, vol. 6, no. 4, pp. 1–15, Oct. 20, 2022. DOI: 10.23880/ppej-16000320

[8] E. Laraque, "Smart laser hole drilling for gas turbine combustors," in *SAE Technical Paper Series*, vol. 1, United States: SAE International, Apr. 1, 1991. DOI: 10.4271/911140

[9] G. G. Karkashadze, E. V. Mazanik, and G. P. Ermak, "Reduction in gas-dynamic energy of outburst-hazardous coal bed under degassing," *Gornyi Zhurnal*, pp. 37–40, Sep. 30, 2015. DOI: 10.17580/gzh.2015.09.07

[10] L. Wang, X. Zhang, Z. Zhang, Z. Q. Li, and W. Ren, "Analysis of development status of high performance water-based drilling fluid," *Journal of Innovation and Social Science Research*, vol. 8, no. 10, Oct. 31, 2021. DOI: 10.53469/jissr.2021.08(10).23

[11] M. Tham et al., *Research needs to maximize economic producibility of the domestic oil resource*, Oct. 1, 1991. DOI: 10.2172/6221899

[12] G. Muktadir, M. Amro, and A. Schramm, "Review and applicability of downhole separation technology," in *SPE Middle East Artificial Lift Conference and Exhibition*, SPE, Nov. 30, 2016. DOI: 10.2118/184201-ms

[13] F. Seidel and T. Greene, "Use of expanding cement improves bonding and aids in eliminating annular gas migration in hobbs grayburg-san andres wells," *SPE Annual Technical Conference and Exhibition*, Sep. 22, 1985. DOI: 10.2118/14434-ms

[14] 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

[15] J. Xia and R. Taghipour, "Feasibility of tlp with tender assisted drilling for northwest australian waters - a case study," in *Offshore Technology Conference*, OTC, Apr. 30, 2012. DOI: 10.4043/23247-ms

[16] S. Zhironkin and D. Szurgacz, *Mining Technologies Innovative Development*. MDPI, Feb. 10, 2022. DOI: 10.3390/books978-3-0365-3224-0

[17] N. Zhang, F. Xue, N. Zhang, and X. Feng, "Patterns and security technologies for co-extraction of coal and gas in deep mines without entry pillars," *International Journal of Coal Science & Technology*, vol. 2, no. 1, pp. 66–75, May 9, 2015. DOI: 10.1007/s40789-015-0058-1

[18] "Research of the lost circulation curing," *Acta Montanistica Slovaca*, no. 26, pp. 582–592, Dec. 23, 2021. DOI: 10.46544/ams.v26i3.15

[19] X. Ding, C. Zhai, J. Xu, X. Yu, and Y. Sun, "Study on coal seepage characteristics and secondary enhanced gas extraction technology under dual stress disturbance," *Sustainability*, vol. 14, no. 22, pp. 15 118–15 118, Nov. 15, 2022. DOI: 10.3390/su142215118

[20] K. Kinik, F. Gumus, and N. Osayande, "A case study: First field application of fully automated kick detection and control by mpd system in western canada," in *SPE/IADC Managed Pressure Drilling & Underbalanced Operations Conference & Exhibition*, SPE, Apr. 8, 2014. DOI: 10.2118/168948-ms

[21] M. Parker, J. Wielenga, V. Bochkarev, I. Poletzky, M. Juskiw, and S. Saeed, "Planning managed pressure drilling with two-phase fluid in a depleted reservoir," in *SPE/IADC Managed Pressure Drilling & Underbalanced Operations Conference & Exhibition*, SPE, Apr. 8, 2014. DOI: 10.2118/168946-ms

[22] X. J. Zhang, S. Taoutaou, Y. Guo, Y. L. An, S. M. Zhong, and Y. Wang, "Engineering cementing solution for hutubi underground-gas-storage project," *SPE Drilling & Completion*, vol. 29, no. 01, pp. 88–97, Feb. 11, 2014. DOI: 10.2118/163974-pa

[23] A. D. Black and A. Judzis, *Smaller footprint drilling system for deep and hard rock environments; feasibility of ultra-high-speed diamond drilling*, Oct. 1, 2004. DOI: 10.2172/835941

[24] M. Murphy, *Advanced oil recovery technologies for improved recovery from slope basin clastic reservoirs, nash draw brushy canyon pool, eddy county, nm*, Feb. 1, 1999. DOI: 10.2172/3257

[25] J. Han, J. Sun, K. Lv, J. Yang, and Y. Li, "Polymer gels used in oil-gas drilling and production engineering.," *Gels (Basel, Switzerland)*, vol. 8, no. 10, pp. 637–637, Oct. 7, 2022. DOI: 10.3390/gels8100637

[26] C. Winter and A. Edgington, "Computer analysis of gas field deliverability and development economics," *Journal of Canadian Petroleum Technology*, vol. 3, no. 04, pp. 180–186, Dec. 1, 1964. DOI: 10.2118/64-04-10

[27] K. D. Lyons, S. Honeygan, and T. Mroz, "Netl extreme drilling laboratory studies high pressure high temperature drilling phenomena," in *Volume 2: Structures, Safety and Reliability; Petroleum Technology Symposium*, ASMEDC, Jan. 1, 2007, pp. 791–796. DOI: 10.1115/omae2007-29478

[28] N. Hedzyk and L. Haidarova, "The research of the gas wells bottomhole formation zone treatment efficiency," *Prospecting and Development of Oil and Gas Fields*, no. 1(78), pp. 51–60, Mar. 29, 2021. DOI: 10.31471/1993-9973-2021-1(78)-51-60

[29] K. C. Manger, *Assessing geothermal energy potential in upstate new york. final report, tasks 1, 3, and 4*, Jul. 25, 1996. DOI: 10.2172/501499

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

[31] W. Parks, W. Going, and D. Beebe, "The unitized well cluster: A low-cost alternative to large multi-well templates and clustered satellite wells," in *Offshore Technology Conference*, OTC, May 1, 1995. DOI: 10.4043/7906-ms

[32] M. J. Watson, N. Hawkes, P. F. Pickering, and L. D. Brown, "Efficient conceptual design of an offshore gas gathering network," in *SPE Asia Pacific Oil and Gas Conference and Exhibition*, SPE, Oct. 20, 2008. DOI: 10.2118/116593-ms

[33] L. Yang, A. Li, X. Cheng, and Y. Meng, "Under-balance drilling technology of slim directional well in region xiaoguai of karamay oilfield," in *International Oil and Gas Conference and Exhibition in China*, SPE, Nov. 7, 2000. DOI: 10.2118/64700-ms

[34] A. Wilson, "Hydraulic percussion drilling system boosts rate of penetration, lowers cost," *Journal of Petroleum Technology*, vol. 67, no. 12, pp. 79–80, Dec. 1, 2015. DOI: 10.2118/1215-0079-jpt

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

[36] T. C. Chidsey, *Production analysis: Cherokee and bug fields, san juan county, utah*, Dec. 1, 2003. DOI: 10.2172/835836

[37] L. Payne, "The effect of circulating media and nozzle design on rock bit performance," *Journal of Petroleum Technology*, vol. 4, no. 01, pp. 9–13, Jan. 1, 1952. DOI: 10.2118/144-g

[38] P. H. Lampén and A. Nikula, "Groundwater investigations in the lavia deep drilling hole-sampling techniques and results," *Water Science and Technology*, vol. 20, no. 3, pp. 223–224, Mar. 1, 1988. DOI: 10.2166/wst.1988.0103

[39] *Geothermal energy r&d program annual progress report for fiscal year 1992*, Jul. 1, 1993. DOI: 10.2172/860708

[40] A. A. Almashwali, B. Lal, A. S. Maulud, and K. S. Foo, "Kinetic inhibition effect of valine on methane hydrate nucleation time in oil system," *Key Engineering Materials*, vol. 938, pp. 133–138, Dec. 26, 2022. DOI: 10.4028/p-d7t27z

[41] K. Witwer, *Ignitability testing of the no-flow push bit*, Sep. 20, 1996. DOI: 10.2172/290551

[42] *Final report on uncertainties in the detection, measurement, and analysis of selected features pertinent to deep geologic repositories*, Jul. 10, 1978. DOI: 10.2172/6552380

[43] T. Chrulski, "Statistical review of the italian gas transmission system operator under conditions of the covid-19 pandemic and the supply restriction from the russian federation," *Gases*, vol. 2, no. 4, pp. 134–145, Oct. 9, 2022. DOI: 10.3390/gases2040008

[44] J. Sperry, *Development and field testing of a process for recovering heavy crude oil in the carlyle pool-allen county, kansas using the vapor therm generator. final report*, Sep. 1, 1980. DOI: 10.2172/5036541

[45] V. O. Yablonskii, "Modeling of degassing of viscoplastic liquids in a cylindrical hydrocyclone," *Russian Journal of Applied Chemistry*, vol. 95, no. 2, pp. 270–276, Jun. 15, 2022. DOI: 10.1134/s1070427222020069

[46] *Underground gasification for steeply dipping coal beds: Phase iii. quarterly progress report, april 1-june 30, 1981. [rawlins test 2]*, Dec. 1, 1981. DOI: 10.2172/5220905

[47] N. Nakaten and T. Kempka, "Techno-economic comparison of onshore and offshore underground coal gasification end-product competitiveness," *Energies*, vol. 12, no. 17, pp. 3252–3252, Aug. 23, 2019. DOI: 10.3390/en12173252

[48] *Oil shale, tar sand, coal research, advanced exploratory process technology, jointly sponsored research. quarterly technical progress report, july–september 1992*, Dec. 31, 1992. DOI: 10.2172/10124838

[49] S. C. Soni, A. Bhargava, A. B. Desai, R. Kumar, and H. G. Suryadi, "Continuous improvement in drilling with rotary steerable system in mumbai high field.," in *SPE Oil and Gas India Conference and Exhibition*, SPE, Mar. 28, 2012. DOI: 10.2118/154823-ms

[50] A. A. Sharkh et al., "The challenge of improving water quality prior to discharge to the sea during well clean up and testing operations by the rig/barge," in *Abu Dhabi International Petroleum Exhibition and Conference*, SPE, Oct. 13, 2002. DOI: 10.2118/78553-ms

[51] B. Qin, Z. Shi, J. Hao, D. Ye, B. Liang, and W. Sun, "Analysis of the space-time synergy of coal and gas co-mining.," *ACS omega*, vol. 7, no. 16, pp. 13 737–13 749, Apr. 12, 2022. DOI: 10.1021/acsomega.2c00034

[52] C. Jee, J. White, and S. Bhatia, *Study of true in situ oil shale technology*, May 1, 1977. DOI: 10.2172/5255026

[53] *Preliminary performance assessment for the waste isolation pilot plant, december 1992. volume 2, technical basis*, Dec. 1, 1992. DOI: 10.2172/10152565

[54] R. Gehle and R. Thoms, *Tests of us rock salt for long-term stability of caes reservoirs*, Jan. 1, 1986. DOI: 10.2172/5903334

[55] A. L. D'Almeida, N. C. R. Bergiante, G. de Souza Ferreira, F. R. Leta, C. B. de Campos Lima, and G. B. A. Lima, "Digital transformation: A review on artificial intelligence techniques in drilling and production applications.," *The International journal, advanced manufacturing technology*, vol. 119, no. 9-10, pp. 5553–5582, Jan. 22, 2022. DOI: 10.1007/s00170-021-08631-w

[56] J. Labaj and D. Barta, "Unsteady flow simulation and combustion of ethanol in diesel engines," *Communications - Scientific letters of the University of Zilina*, vol. 8, no. 2, pp. 27–37, Jun. 30, 2006. DOI: 10.26552/com.c.2006.2.27-37

[57] L. Lou, M.-j. Chen, W.-y. Qin, W.-r. Wu, and H.-y. Rui, "Research on the synchronization and shock characteristics of an air adjustment mechanism for cluster-type dth hammers under partial loads," *Shock and Vibration*, vol. 2022, pp. 1–17, Mar. 31, 2022. DOI: 10.1155/2022/9794391

[58] Y. Wang, T. Luan, D. Gao, and J. Wang, "Research progress on recoil analysis and control technology of deepwater drilling risers," *Energies*, vol. 15, no. 19, pp. 6897–6897, Sep. 21, 2022. DOI: 10.3390/en15196897

[59] N. G. Author, *Testing geopressured geothermal reservoirs in existing wells. saldana well no. 2, zapata county, texas. volume i. completion and testing. final report*, Oct. 7, 1981. DOI: 10.2172/6054041

[60] W. Overbey, R. Carden, C. Locke, and S. Salamy, *Drilling, completion, stimulation, and testing of hardy hw[number sign]1 well, putnam county, west virginia*, Mar. 1, 1992. DOI: 10.2172/6910867

[61] L. J. Spadaccini and E. J. Szetela, "Approaches to the prevaporized-premixed combustor concept for gas turbines," in *Volume 1B: General*, American Society of Mechanical Engineers, Mar. 2, 1975. DOI: 10.1115/75-gt-85

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

[63] A. Hosseini, M. Najafi, and A. Morshedy, "Determination of suitable distance between methane drainage stations in tabas mechanized coal mine (iran) based on theoretical calculations and field investigation," *Journal of Mining Institute*, vol. 258, pp. 1050–1060, Dec. 30, 2022. DOI: 10.31897/pmi.2022.106

[64] W. Lee, L. Chaturvedi, M. K. Silva, R. Weiner, and R. —. Neill, *An appraisal of the 1992 preliminary performance assessment for the waste isolation pilot plant*, Sep. 1, 1994. DOI: 10.2172/10186287

[65] 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.

[66] A. Nandan and S. Imtiaz, "Nonlinear model predictive control of managed pressure drilling.," *ISA transactions*, vol. 69, pp. 307–314, Apr. 8, 2017. DOI: 10.1016/j.isatra.2017.03.013

[67] 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

[68] L. Brown and A. Vadie, *The utilization of the microflora indignous to and present in oil-bearing formations to selectively plug the more porous zones thereby increasing oil recovery during waterflooding*, Apr. 20, 1997. DOI: 10.2172/644597

[69] G. Thoma et al., *Probabilistic risk based decision support for oil and gas exploration and production facilities in sensitive ecosystems*, May 31, 2009. DOI: 10.2172/975095

[70] W. G. Deskins, W. J. McDonald, R. G. Knoll, and S. J. Springer, *Horizontal oil well applications and oil recovery assessment. volume 2: Applications overview, final report*, Mar. 1, 1995. DOI: 10.2172/28416