ASME Journal of Fluids Engineering
Online journal at:
https://asmedigitalcollection.asme.org/fluidsengineering

Enhanced Liquid Detection in Wet Gas Metering Via Microwave Sensing and Random Forest Regression

Ishigita Lucas Shunashu
Department of Mechanical and Industrial Engineering,
College of Engineering and Technology,
Mbeya University of Science and Technology (MUST),
Mbeya, Tanzania
e-mail: ishigitalc@gmail.com

Osmund Kaunde
College of Engineering and Technology,
Mbeya University of Science and Technology (MUST),
Mbeya, Tanzania
e-mail: kaunde.ok@gmail.com

Duncan Mwakipesile
Department of Engineering,
St. Joseph University in Tanzania (SJUIT),
Dar es Salaam, Tanzania
e-mail: dmwakipesile63@gmail.com

This study explored the integration of machine learning regression models with a microwave transmission line sensor for estimating liquid volume fraction and liquid flowrate in wet gas flows. Under low liquid loading conditions (gas volume fraction 95–99.9%), four models: Bruggeman, support vector regression, Gaussian process regression, and random forest regression were evaluated. Random forest regression delivered the best tradeoff between accuracy, robustness, and computational efficiency, achieving a relative absolute error of 2.23% for liquid volume fraction and approximately 5% for liquid flowrate, with a Durbin–Watson statistic of 2.02 indicating minimal residual autocorrelation. Feature importance analysis identified the mixture dielectric constant as the dominant predictor (approximately 97% contribution), while other dimensionless parameters had a limited impact. Support vector regression failed to generalize, and although Gaussian process regression showed slightly higher accuracy, its computational cost limited real-time applicability. Overall, random forest regression combined with microwave sensing offers a scalable, nonintrusive solution for wet gas metering, with future validation needed under industrial hydrocarbon–water conditions and liquid loading flow regimes.
[DOI: 10.1115/1.4070244]

Keywords: wet gas measurement, microwave sensor, machine learning regression, liquid fraction detection

1 Introduction

Accurate measurement of liquid fraction in wet gas flows is critical in oil and gas, petrochemical, and related industries, where even small errors can significantly affect production optimization, flow assurance, and resource allocation. Wet gas refers to a two-phase flow in which the gas volume fraction (GVF) typically exceeds 95%, which is common in production and transportation systems. Its liquid volume fraction (LVF) is a key parameter for process control, allocation metering, and corrosion monitoring. In industrial practice, wet gas is often characterized using the Lockhart–Martinelli parameter X_{LM} <0.3 or by GVF exceeding 95%, as adopted in ISO/TR 12748 and widely accepted in the oil and gas sector.

Despite its importance, wet gas measurement remains challenging due to the complex interaction of phases, the wide range of possible flow regimes, and the influence of environmental and operational conditions such as pressure, temperature, salinity, and fluid composition. Traditional measurement techniques often face tradeoffs between accuracy, cost, safety, and ease of deployment. Microwave sensing has emerged as a promising nonintrusive approach for LVF detection, leveraging the sensitivity of electromagnetic waves to changes in the dielectric properties of multiphase mixtures. By monitoring phase shift and amplitude attenuation, microwave sensors can infer the dielectric constant of the mixture, which correlates with phase fractions. However, despite their potential, microwave-based systems still face limitations such as noise sensitivity, environmental variability, and calibration complexity.

Recent advances in data-driven modeling, particularly machine learning (ML), offer new opportunities to address these challenges. ML algorithms can learn complex, nonlinear relationships between sensor outputs and flow parameters, enabling improved calibration, compensation for environmental effects, and enhanced robustness under dynamic conditions. This study investigates the integration of microwave sensing with three ML models: Support vector regression (SVR), Gaussian process regression (GPR), and random forest regression (RFR) for accurate, real-time LVF detection in wet gas flows. The objectives are to evaluate the predictive performance of these ML models using microwave-derived dielectric properties and dimensionless flow parameters, identify the most effective model for improving LVF measurement accuracy and robustness, and demonstrate the potential of the integrated approach for scalable, field-deployable wet gas measurement.

2 Literature Review

2.1 Liquid Content Measurement Techniques. A variety of techniques have been developed for wet gas measurement, each with distinct strengths and limitations. Electrical capacitance/conductance impedance (ECI) methods offer strong temporal resolution but

Contributed by the Fluids Engineering Division of ASME for publication in the Journal of Fluids Engineering. Manuscript received July 3, 2025; final manuscript received October 14, 2025; published online November 27, 2025. Assoc. Editor: Jinxiang Xi.

suffer from low spatial resolution, noise susceptibility, and frequent calibration requirements [1]. Gamma-ray attenuation systems can achieve ±5% accuracy but are costly and require strict radiation safety measures. Optical methods, such as laser-based and fiber-optic sensors, provide high precision and nonintrusiveness but are sensitive to contamination and require careful alignment [2].

Microwave-based techniques have emerged as promising alternatives for phase fraction detection. The transmission–reflection method examines signal changes across the flow but requires precise calibration and frequency-specific optimization [3]. Doppler-based systems measure velocity changes due to phase motion with accuracies of ±10%, but are limited by noise and environmental sensitivities such as salinity, condensate, and turbulence [4]. Microwave tomography reconstructs cross-sectional flow images with ±5% accuracy but demands complex setups and robust sensor arrangements [5]. Two principal microwave measurement approaches are widely applied: the resonant cavity method and the transmission line method. The resonant cavity approach detects shifts in resonant frequency due to dielectric property changes; however, its accuracy can degrade at high liquid content, leading to measurement uncertainty [3,6]. In contrast, the transmission line method monitors phase and amplitude variations along a transmission line in contact with the flowing mixture, offering enhanced reliability under high liquid content conditions and supporting continuous, real-time monitoring across a wide range of flow regimes [7,8]. Table 1 below summarizes the main wet gas measurement techniques, highlighting their operating principles and typical accuracy.

Several studies have demonstrated the potential of microwave sensing for LVF measurement. Tayyab et al. [9] developed a noninvasive microwave sensor for dielectric constant estimation in multiphase oil flows, achieving ±10% accuracy but with sensitivity to phase composition and dynamic conditions. Sabzevari et al. [10] applied ultrawideband synthetic aperture radar (UWB-SAR) for multiphase flow monitoring, achieving a maximum error of 3.8% but still facing environmental interference issues. Other work has targeted specific applications, such as offshore natural gas fields [11,12] gas bubbles in liquids [13], and water-liquid ratio sensing [14], but limitations persist in handling varying compositions, pressures, and temperatures. Broader applications of microwave sensors have been explored for environmental monitoring and material diagnostics [7], yet challenges such as sensitivity to salt content, temperature, liquid density, and multipath effects remain.

2.2 Machine Learning Applications in Wet Gas and Two-Phase Flow Measurement. Traditional models for LVF estimation often struggle with dynamic flow regimes, environmental variability, and sensor noise. To address these limitations, researchers have increasingly turned to ML techniques, particularly RFR, SVR, and GPR, to enhance predictive accuracy and robustness.

Random forest regression has shown strong capability in handling nonlinear relationships and high-dimensional datasets. Hosseini et al. [15] applied RFR to wet gas flow modeling, achieving superior performance over empirical models in predicting LVF under varying gas compositions and flow rates. SVM has been used for both classification and regression in multiphase systems; Zhang and Xia [16] combined RS-SVM classifiers with genetic algorithm-optimized neural networks to estimate water content in oil–water mixtures, achieving high accuracy even in noisy conditions. GPR offers a probabilistic regression framework ideal for uncertainty quantification; Lin et al. [5] integrated GPR with microwave and Venturi sensor data to estimate oil–water flow rates with high precision across varying salinity and temperature. Table 2 summarizes key ML applications in wet gas and multiphase flow contexts.

Collectively, these studies underscore the versatility of RFR, SVR, and GPR in modeling fluid dynamics and validate their integration with microwave sensor data for enhanced wet gas metering.

2.3 Measurement Research Gap and Objectives. While conventional wet gas measurement techniques and microwave-based sensing methods have advanced significantly, they still face persistent challenges related to calibration complexity, environmental sensitivity, and reduced accuracy under varying flow regimes. ML models have demonstrated strong potential to address these issues, yet their integration with microwave sensing for LVF detection in wet gas remains limited. Few studies have systematically compared RFR, SVR, and GPR using identical experimental datasets, and even fewer have assessed their robustness across a wide range of flow regimes and environmental conditions.

This study addresses the identified gaps by developing and evaluating an integrated microwave–machine learning (ML) framework for LVF detection in wet gas flows, combining advanced sensing techniques with data-driven modeling to improve measurement accuracy, enhance robustness under low liquid loading flow conditions, and reduce environmental sensitivity. The work specifically aims to develop a microwave sensing framework for LVF detection in wet gas; train and test RFR, SVR, and GPR models using dielectric and dimensionless flow parameters; compare the models in terms of accuracy, robustness, and suitability for real-time application; identify the most effective model for integration with microwave sensing; and demonstrate the scalability and non-intrusive applicability of the integrated approach for industrial wet gas metering.

3 Materials and Methods

This section presents the integrated methodology for estimating liquid volume fraction (LVF) in wet gas flows using a microwave transmission line sensor and machine learning regression. The sensor system, illustrated in Fig. 1, operates at 2.7 GHz and captures dielectric variations through phase shift and amplitude attenuation. Both Bruggeman and Maxwell models were jointly applied to replicate and interpret experimental data from wet gas flows with GVF ranging from 95% to 99.9%. The reconstructed dielectric constants were used to calculate LVF via the Bruggeman model. This performance was further enhanced through the integration of machine learning regression models.

3.1 Theoretical Models and Sensor Architecture. The two theoretical models are essential for interpreting microwave sensor data: the Bruggeman model, which estimates the effective dielectric constant of mixtures, and the Maxwell framework, which describes the propagation of electromagnetic waves through lossy media. Together, these models facilitate the reconstruction of the dielectric constant of wet gas and its conversion into Liquid Volume Fraction (LVF) under conditions of low liquid loading flow [7].

3.1.1 Maxwell Electromagnetic Wave Propagation Framework. In this study, Maxwell's equations were applied as the theoretical basis for processing the microwave sensor measurements and quantifying the propagation of electromagnetic waves through the wet gas mixtures tested. In the context of wet gas flows, microwave sensors exploit these principles to detect variations in dielectric properties by measuring phase shifts and amplitude attenuation of the transmitted signal. These variations are directly related to the complex dielectric constant of the gas–liquid mixture [3]. For a sinusoidal plane wave propagating through a lossy dielectric medium, the electric field can be expressed as

$$E(x) = E_0 e^{-jkx} = E_0 e^{-jk'x} e^{-k''x} \tag{1}$$

where $k = k' - jk''$ is the complex propagation constant, ω is the angular frequency, μ is the magnetic permeability, and $\varepsilon^* = \varepsilon' - j\varepsilon''$ is the complex dielectric constant of the medium. From Maxwell's wave equation, the propagation constant is related to the electromagnetic properties of the medium by

$$k^2 = \omega^2 \mu \varepsilon^* \tag{2}$$

Microwave sensors measure two key quantities: the phase shift ($\Delta\theta$) and the attenuation (ΔA), which are linked to the real k' and imaginary k'' components of the propagation constant, respectively

$$\Delta\theta = -8.68\alpha x = -8.68k''x \tag{3}$$

$$\Delta A = \beta x = k'x \tag{4}$$

From the measured phase shift $\Delta\theta$ (radians) and attenuation $\Delta\theta$ (dB), the real β and imaginary α components of the propagation constant are extracted, allowing the effective dielectric constant of the mixture to be reconstructed as

$$\varepsilon_{mix} = \frac{c^2}{\omega^2}\left(\beta^2 - \alpha^2\right) \tag{5}$$

Table 1 Summary of wet gas measurement techniques and their characteristics

Method	Principle/application focus	Accuracy	References
Electrical capacitance/conductance impedance (ECI)	Measures changes in capacitance/conductance due to dielectric property variations in gas–liquid mixtures	±5–10%	[1]
Gamma-ray attenuation	Uses gamma radiation attenuation to infer phase fractions	±5%	[12]
Optical (laser/fiber-optic)	Detects light scattering/reflection from dispersed phases	±2–5%	[2]
Microwave transmission–reflection	Measures phase shift and attenuation of microwaves passing through the flow.	±5–10%	[8]
Microwave Doppler shift	Detects frequency shift from moving phases.	±10%	[3]
Microwave tomography	Reconstructs cross-sectional images from multiple microwave paths.	±5%	[4]
Microwave resonant cavity	Monitors resonant frequency shifts due to dielectric changes in the cavity.	±5–10%	[5,6]
Microwave transmission line	Tracks phase/amplitude changes along a transmission line in contact with the flow	±5%	[7,8]

Table 2 Summary of ML models in wet gas/multiphase flow measurement

ML model	Application focus	Key metrics	References
RFR, SVR, GPR	Wet gas flow rate prediction using dielectric and flow-derived features	RMSE = 0.011 (RFR)	[15]
SVM, GPR	Multiphase flow rate estimation from time-series ultrasonic/microwave data	RMSE < 0.02 (GPR)	[20]
SVR	Predicting liquid/gas flow rates from pressure, temperature, and velocity	MSE = 0.0169 (liquid)	[19]
RFR, GPR	Blockage prediction in multiphase flow.	Accuracy > 95%	[22]
SVM + GA-NN	Water content estimation in oil–water mixtures	Accuracy > 90%	[16]

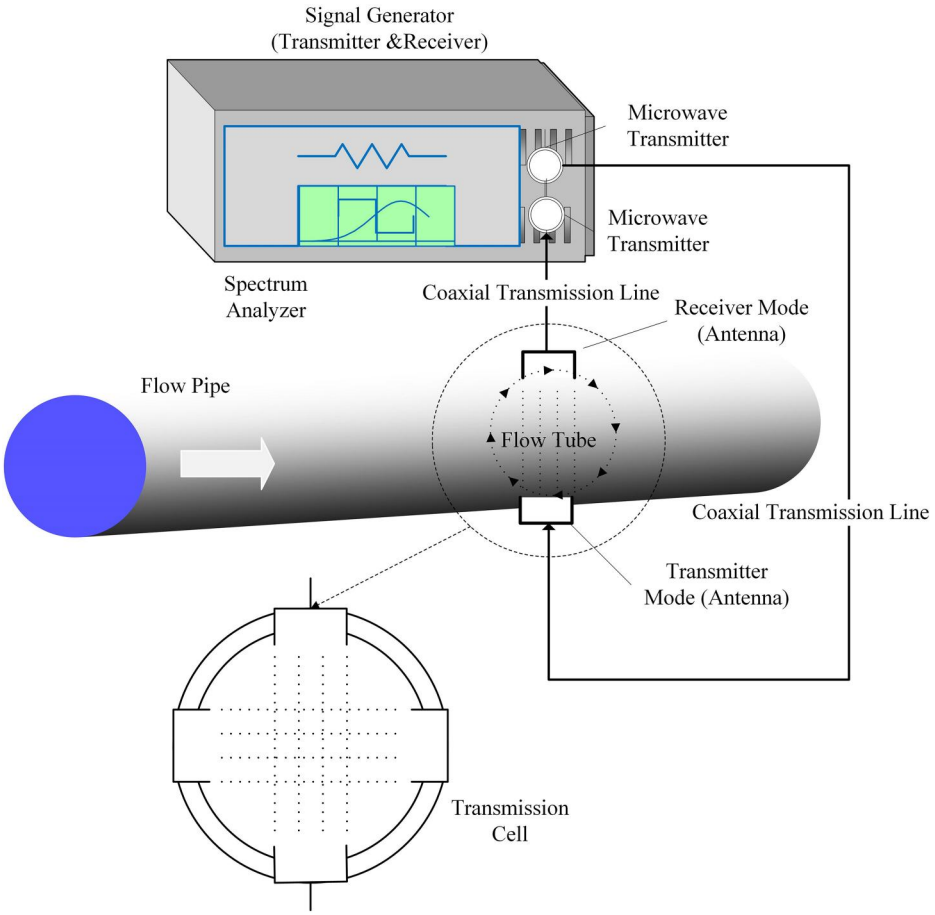

Fig. 1 Structure of the microwave transmission line sensor

where $\beta = k', \alpha = k''$.The reconstructed ε_{mix} is subsequently used as an input to the Bruggeman model for estimating the liquid volume fraction (LVF) in wet gas flows.

3.1.2 Bruggeman Model for Liquid Volume Fraction Estimation. In this study, the Bruggeman effective medium theory was applied to estimate the dielectric constant of wet gas mixtures from measured dielectric properties and phase volume fractions. The method was suited to the mist-flow conditions tested, where low liquid loading allowed the mixture to be treated as statistically homogeneous [6]. Originally developed for composite and porous media [17], it was here adapted to link microwave-measured dielectric constants to liquid volume fraction in wet gas metering [7,18] as expressed in Eq. (1).

In this work, the Bruggeman effective medium model was formulated and implemented to determine the liquid volume fraction (*LVF*) in wet gas flows from microwave sensor measurements. The gas phase (ε_g) and liquid phase (ε_l) were treated as dispersed within a statistically homogeneous mixture, enabling the measured effective dielectric constant (ε_{mix}) to be related directly to LVF and the gas volume fraction (GVF = 1 − LVF). The formulation assumed that each phase was embedded in the effective medium and contributed to the overall polarization. For the two-phase gas–liquid system tested, the general Bruggeman equation was applied as

$$(1-\text{LVF})\cdot\frac{\varepsilon_g-\varepsilon_{\text{mix}}}{\varepsilon_g+L(\varepsilon_{\text{mix}}-\varepsilon_g)}+LVF\cdot\frac{\varepsilon_l-\varepsilon_{\text{mix}}}{\varepsilon_l+L(\varepsilon_{\text{mix}}-\varepsilon_l)}=0 \quad (6)$$

where L is the geometric (depolarization) factor, taken as $1/3$ for the spherical droplet assumption. The value of ε_{mix} was obtained experimentally from the microwave sensor, and Eq. (1) was inverted to estimate LVF. Under the wet gas flow condition of this study, which is characterized by a statistically homogeneous distribution of spherical droplets [7,16], the LVF was computed using the closed-form expression

$$\text{LVF} = 1-\left(\frac{\varepsilon_l-\varepsilon_{mix}}{\varepsilon_l-\varepsilon_g}\cdot\frac{\varepsilon_g}{\varepsilon_{mix}}\right)^{1/3} \quad (7)$$

This closed form was implemented in the data-processing workflow for real-time LVF estimation in low liquid loading conditions. For tests involving saline or conductive liquids, the dielectric constant was corrected for conductivity effects using

$$\varepsilon^* = \varepsilon' - j(\sigma/\omega) \quad (8)$$

where σ is the electrical conductivity and $\omega = 2\pi f$ is the angular frequency. Pressure effects were also incorporated by adjusting ε_g for pressure-dependent gas density and dielectric behavior, following the pressure-compensated Bruggeman formulation [7]

$$\text{LVF}(P) = \left(\frac{\varepsilon_l-\varepsilon_{\text{mix}}}{\varepsilon_l-\varepsilon_g(P)}\cdot\frac{\varepsilon_g(P)}{\varepsilon_{\text{mix}}}\right)^{1/3} \quad (9)$$

Once LVF was determined, the corresponding liquid flowrate Q_l was calculated from the measured total two-phase flowrate Q_{tp} as

$$Q_l = \text{LVF}\cdot Q_{tp} = \left[1-\left(\frac{\varepsilon_l-\varepsilon_{\text{mix}}}{\varepsilon_l-\varepsilon_g}\cdot\frac{\varepsilon_g}{\varepsilon_{\text{mix}}}\right)^{1/3}\right]\cdot Q_{tp} \quad (10)$$

This implementation assumed isotropic phase distribution and neglected interfacial effects, which may reduce accuracy in stratified or slug flows. Nevertheless, the model's symmetry, closed-form solution, and low computational cost made it well-suited for the real-time sensing requirements of the wet gas pipeline tests conducted in this study [9].

3.2 Machine Learning Regression Model. To extend liquid volume fraction (LVF) estimation beyond the analytical framework outlined in Sec. 3.1, this study integrated microwave sensor measurements with three advanced regression models: RFR, SFR, and GPR. Each model was independently trained and tested to assess its predictive performance and suitability for wet gas applications. This integration was designed to exploit the complementary strengths of the algorithms, RFR's robustness to noise and ability to model complex feature interactions [15], SVR's suitability for precise fitting on limited datasets [8], and GPR's inherent capability for uncertainty quantification, enabling the model to capture complex, nonlinear relationships between sensor-derived features and LVF [19]. The following subsections describe the physics-informed feature set, model training procedures, and performance evaluation metrics.

3.2.1 Input Feature Selection and Engineering. The machine learning models were trained on six physics-informed features that collectively represent both electromagnetic and hydrodynamic characteristics of wet gas flow

$$f\left(\varepsilon_g/\varepsilon_l, \varepsilon_{\text{mix}}, \text{Re}_l, \text{Re}_g, \text{We}, S\right) \quad (11)$$

The wet gas dielectric constant ε_{mix}, was obtained from phase shift and amplitude attenuation measurements obtained via a microwave transmission line sensor. The sensor output was calibrated against the known dielectric constants of the pure liquid ε_l and gas ε_g phases. The ratio $\varepsilon_g/\varepsilon_l$ quantifies the dielectric contrast between the carrier gas and dispersed liquid. Hydrodynamic conditions are expressed by the liquid and gas Reynolds numbers, (Re_l, Re_g), calculated from phase-specific volumetric flow rates, densities and viscosities to reflect inertial viscous force balances. The Weber number, We is derived from phase velocity, fluid density, characteristic droplet diameter and surface tension to characterize interfacial dynamics. Finally, the slip ratio, S estimated the dispersed-phase velocity divided by the carrier-phase velocity, accounts for phase-slip effects. These features formed the independent variable set for all three regression frameworks. The target LVF was also expressed in a Bruggeman-based form

$$\text{LVF} = f\left(\varepsilon_g/\varepsilon_l, \varepsilon_{\text{mix}}, \text{Re}_l, \text{Re}_g, \text{We}, S\right) \quad (12)$$

where A_o= 1/3 for spherical droplets, consistent with mist-flow observations (GVF 95–99%).

3.2.2 Model Functions, Training, and Hyperparameter Tuning. All models were trained and tested using the same dataset and split ratio. The dataset was randomly partitioned into 70% training, 20% validation, and 10% testing subsets to ensure robust model development and unbiased performance evaluation. Hyperparameters were optimized using a grid search with fivefold cross-validation, minimizing root mean squared error (RMSE) on the validation set while monitoring R^2, mean absolute percentage error (MAPE), and the Durbin–Watson statistic for residual independence. Final models were retrained on the combined training with validation data before evaluation on the held-out test set.

Support vector regression implemented with an RBF kernel to capture nonlinear relationships. The ε-insensitive loss function-controlled tolerance for prediction errors. Hyperparameters C, ε, and γ were tuned over defined search ranges, with the optimal set selected based on the lowest cross-validated RMSE. The SVR prediction function is

$$\text{LVF}_{\text{SVR}}(x) = \sum_{i=1}^{N}\left(\alpha_i-\alpha_i^*\right)K(x,x_i)+b \quad (13)$$

where $K(x,x_i)$ is the RBF kernel function, α_i, α_i^* are support vector coefficients, b is the bias term, and N refers to the number of support vectors. However, RFR modeled LVF as an ensemble of decision trees trained on bootstrap samples with random feature selection at each split. Hyperparameters tuned included the number of trees ($n_{\text{estimators}}$), maximum depth, minimum samples per split/leaf, and

maximum features per split. Selection criteria included out-of-bag error and validation RMSE. The RFR prediction function is

$$\mathrm{LVF}_{\mathrm{RF}}(x) = \frac{1}{M}\sum_{m=1}^{M} T_m(x) \tag{14}$$

where M is the number of trees and $T_m(x)$ is the prediction from the mth tree. Consequently, by using the same dataset and split ratio, GPR was applied with a probabilistic regression model with an automatic relevance determination (ARD) Matérn-5/2 kernel for smooth nonlinear regression and uncertainty estimation. Hyperparameters (length-scales l_i, signal variance σ_f^2, noise variance σ_n^2) were tuned by maximizing the log-marginal likelihood with multiple starters. The GPR predictive mean is

$$\mathrm{LVF}_{\mathrm{GPR}}(x_*) = k_*^{\top}\left[K + \sigma_n^2 I\right]^{-1} y \tag{15}$$

where K is the covariance matrix of training inputs, k_* is the covariance vector between x_* and the training set, and σ_n^2 is the noise variance.

3.3 Data Collection and Feature Computation. All experiments were conducted at the multiphase process engineering (PSE) Laboratory, Cranfield University, using an industrial-scale horizontal multiphase flow rig designed for controlled wet gas testing. The facility comprises independent gas and liquid supply loops, enabling precise adjustment of phase flow rates, pressures, and temperatures. Compressed air from the gas loop and water from the liquid loop were combined at the test section inlet to generate wet gas mixtures with GVF ranging from 95% to 99.9%, covering stratified to mist flow regimes. The test section was constructed using a stainless-steel flow pipe with an internal diameter of 50.8 mm, ensuring representative flow dynamics for field-scale applications. Figure 2 illustrates the setup of the test rig in the PSE lab.

A 2.7 GHz microwave phase fraction sensor (Analog Devices Inc., Wilmington, MA) was installed in the test section to provide nonintrusive, real-time measurements of the mixture's electromagnetic properties. The sensor was integrated with an RF signal processor to record the phase shift $(\Delta\theta)$ and amplitude attenuation (ΔA) at a sampling rate of 60 samples/min. These raw measurements were synchronized with auxiliary instrumentation, including Coriolis mass flow meters for each phase, differential pressure transducers, and temperature sensors, ensuring complete characterization of the flow conditions.

Prior to multiphase testing, baseline calibration was performed by circulating each pure phase separately through the test section to determine the dielectric constants of the pure gas (ε_g) and pure liquid (ε_l) phases. These reference values were later used to normalize dielectric measurements and compute the dielectric contrast ratio $(\varepsilon_g/\varepsilon_l)$.

For each test point, the gas and liquid flow rates were set to achieve the target superficial velocities (U_g, U_l), along with specified pressures and temperatures, and the resulting flow regime was confirmed visually via the sight glass section and supported by pressure drop signatures. The microwave sensor outputs ($\Delta\theta$, ΔA) were converted to the mixture dielectric constant $(\varepsilon_{\mathrm{mix}})$ using the Maxwell-based reconstruction described in Sec. 2.1.

Fluid properties such as densities (ρ_g, ρ_l), viscosities (μ_g, μ_l), and surface tension (σ) were determined at operating conditions using standard thermophysical correlations. From the measured and derived quantities, dimensionless parameters were calculated for each run to characterize flow dynamics and support model generalization. These included the liquid Reynolds number ($\mathrm{Re}_l = \frac{\rho_l U_l D}{\mu_l}$), gas Reynolds number ($Re_g = \frac{\rho_g U_g D}{\mu_g}$), gas Weber

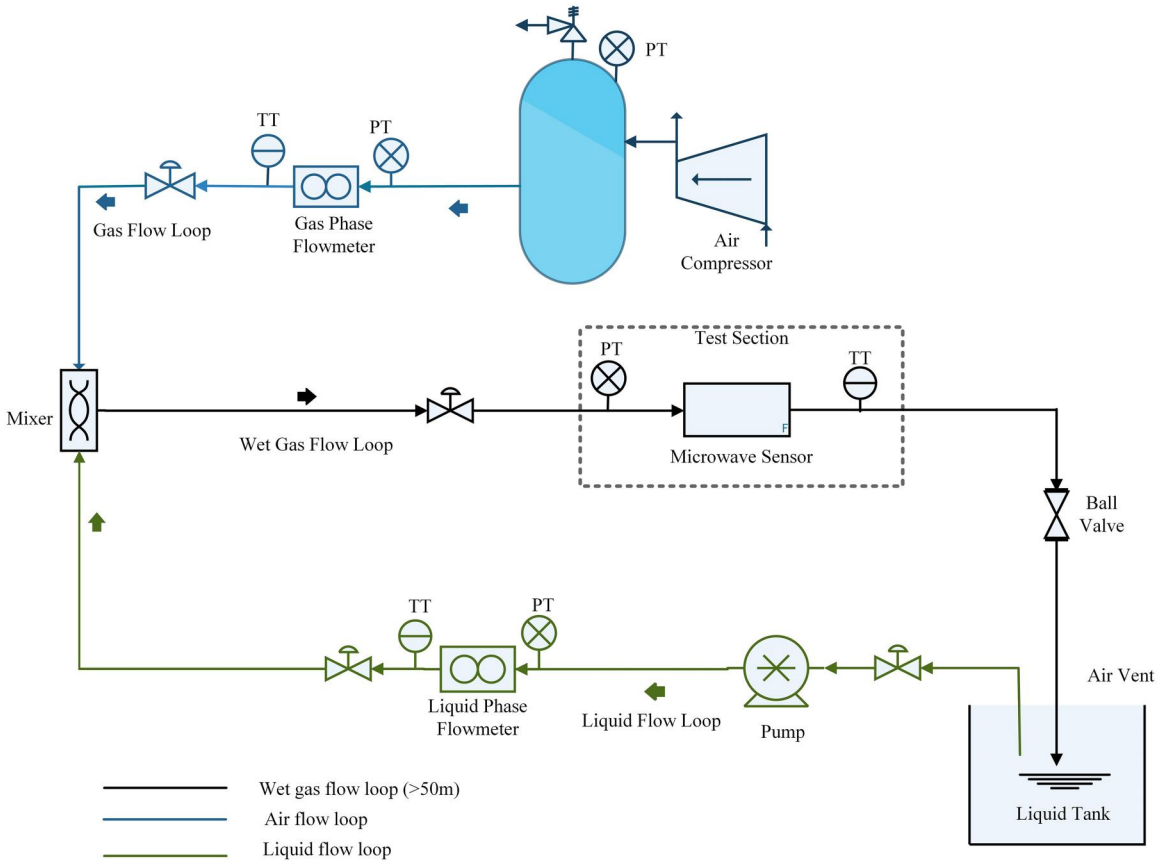

Fig. 2 Experimental setup of microwave sensor for liquid fraction detection in the multiphase flow lab

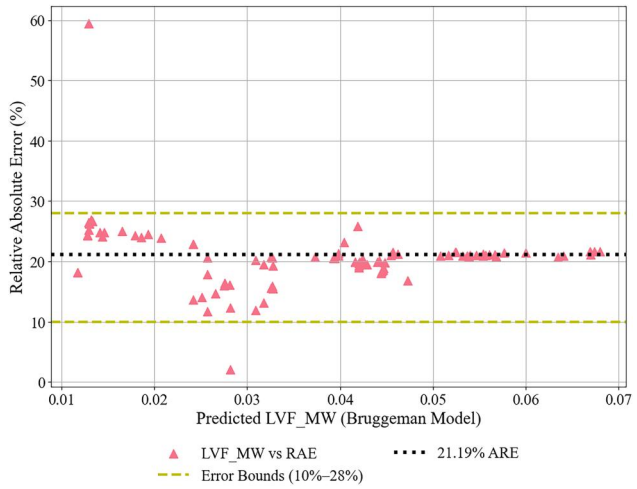

Fig. 3 Relative absolute error (RAE) for the microwave sensor liquid fraction detection based on the Bruggeman model (LVF_MW)

number ($We_g = \frac{\rho_g U_g^2 D}{\sigma}$), and slip ratio ($S = \frac{U_g}{U_l}$). Here, D is internal pipe diameter.

The experimental setting in Sec. 3.3 was conducted under controlled laboratory conditions, with pressures of 1.2–25 bar(a) and temperatures of 20–40 °C, using air–water mixtures. While these conditions aimed to replicate stratified and mist flow regimes, they do not encompass the full range of industrial wet gas scenarios. High-pressure/high-temperature environments, saline liquids, crude oil–water mixtures, and varying gas compositions were not included, all of which can influence dielectric behavior and flow stability.

4 Results

This section presents the results of liquid fraction detection in wet gas flows using microwave sensors and machine learning models. Experimental data collected under a stratified and mist flow regime were analyzed using both analytical and data-driven approaches. The Bruggeman model provided baseline LVF estimates from dielectric measurements, while machine learning models, SVR, GPR, and RFR, were trained to enhance prediction accuracy.

4.1 Liquid Fraction Based on the Bruggeman Model. The Bruggeman model was applied to estimate the LVF from microwave sensor data, using phase shift ($\Delta\theta$) and amplitude attenuation (ΔA) to derive the mixture dielectric constant (ε_{mix}). This analytical approach provided a baseline for evaluating the sensor's capability in detecting liquid content in wet gas flows.

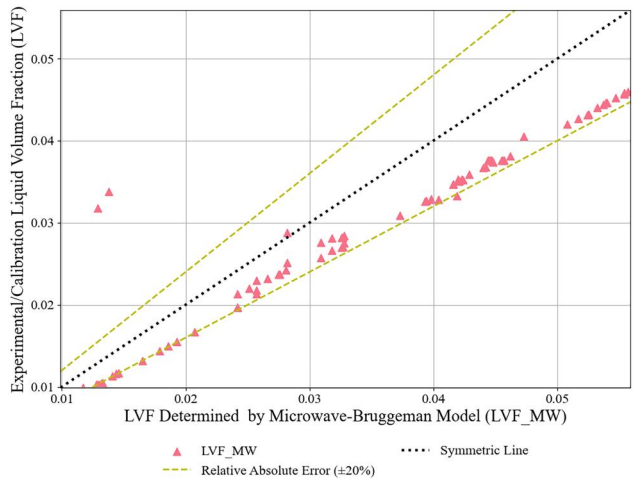

Fig. 4 The correlation of LVF based on the Bruggeman model (LVF_MW) with the calibration liquid volume fraction (LVF)

As shown in Fig. 3, the relative absolute error (RAE) for LVF detection using the Bruggeman model averaged 21.23% across all test conditions. The LVF values derived from controlled gas–liquid flow rates were compared with those calculated from microwave data (LVF-MW), revealing a consistent linear correlation between LVF_MW and actual LVF and validating the model's applicability under stratified and mist flow regimes.

Figure 4 illustrates the correlation between the actual LVF, determined from reference flowmeter measurements, and the LVF estimated from microwave sensor data using the Bruggeman model (LVF_MW). The scatter plot reveals a consistent linear trend across the mist-flow regime, confirming the sensor's ability to detect liquid content in low-liquid-loading (wet gas) without prior calibration. In addition to LVF estimation, the Bruggeman model was used to calculate liquid flowrate, with results presented in Fig. 5. A strong correlation was observed between the flowrate measured by reference instrumentation and that derived from microwave data, achieving an average RAE of approximately 10%. This confirms the model's utility in quantifying liquid flow under varying gas volume fractions (GVF), ranging from 95% to 99.9%.

While the Bruggeman model establishes a reliable baseline for microwave-based LVF detection, its performance is limited under dynamic and noisy flow conditions. The results in Figs. 3–5 clearly demonstrate the model's strengths and limitations, creating a compelling case for the integration of machine learning models. Compared to conventional techniques such as electrical capacitance tomography (ECT), gamma-ray densitometry, and differential pressure meters, the microwave-based approach demonstrates comparable or superior correlation performance. While ECT methods are prone to noise and require frequent recalibration, and gamma-ray systems pose safety concerns, the microwave sensor offers a nonintrusive and adaptable alternative.

4.2 Modeling Results and Generalization Performance. As shown in Table 3 below, the performance of the models was evaluated using RMSE, R^2, MAPE, and the Durbin–Watson statistic. The GPR model achieved the lowest RMSE of 0.00167 and the highest R^2 of 0.99955. However, the RFR model was chosen for integration with the microwave sensing framework due to its superior residual independence (Durbin–Watson statistic ≈ 2.02), enhanced robustness to noise, and greater computational efficiency for real-time applications. GPR's dependence on covariance matrix inversion results in higher computational costs and a greater risk of overfitting in high-dimensional, noisy environments. Additionally, support vector regression (SVR) exhibited significantly higher errors, with an RMSE of 0.08264 and MAPE exceeding 450%, making it unsuitable for this application.

Random forest regression's ensemble-based approach yielded stable predictions across the entire mist flow range, achieving a low mean absolute percentage error (MAPE) of 2.01%. This model also enabled a thorough feature importance analysis, revealing that the dielectric constant of the wet gas mixture accounted for approximately 97% of the predictive capability. In contrast, other input features had minimal contributions, such as the gas Reynolds number (0.72%), liquid Reynolds number (0.92%), gas Weber number (0.46%), slip ratio (0.76%), and the dielectric constant ratio (0.01%).

Generalization testing showed that both RFR and GPR maintained high accuracy when trained on as little as 50% of the available data, with only modest increases in RMSE and MAPE. Below this threshold, performance degradation became more pronounced, particularly for SVR. Noise-robustness evaluation using Gaussian perturbations of 1%, 3%, and 5% of feature ranges indicated that RFR was more resilient than GPR at higher noise levels ($\geq 3\%$), consistent with its noise-averaging ensemble structure. GPR retained an edge in low-noise scenarios but declined more sharply as noise increased.

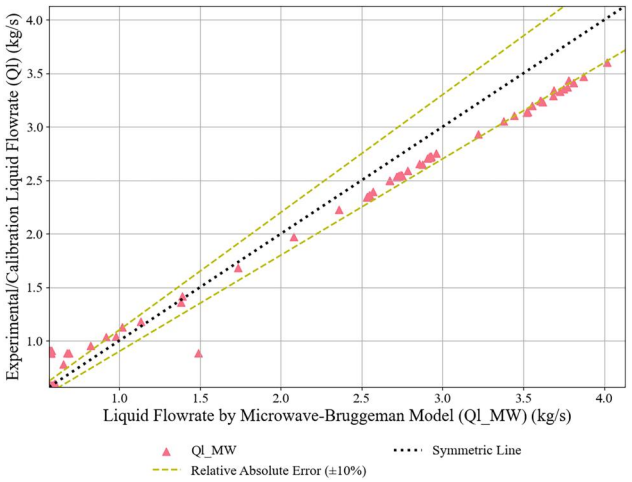

Fig. 5 The correlation of liquid flow rate based on the Bruggeman model and the calibration liquid flow rate

4.3 Liquid Fraction Estimation Via Random Forest Regression. The integration of random forest regression (RFR) with microwave sensing significantly enhanced the accuracy and reliability of liquid volume fraction (LVF) detection in wet gas flows. By leveraging ensemble learning, RFR effectively captured nonlinear relationships between microwave-derived dielectric properties and flow dynamics, outperforming both Gaussian process regression (GPR) and support vector regression (SVR) in predictive performance and operational robustness. The dielectric constant of the wet gas mixture (ε_{mix}), measured in real-time by the microwave sensor, served as the primary input feature, offering a direct and responsive indicator of phase fraction (low liquid loading).

Random forest regression (RFR) demonstrated robust generalization across mist-flow regimes, significantly enhancing the accuracy of liquid volume fraction (LVF) prediction. As shown in Fig. 6, the predicted LVF values (LVF_RF) exhibit a strong positive correlation with experimentally measured liquid fractions, confirming the model's reliability and consistency under varying flow conditions. Building on this, Fig. 7 illustrates the improvement in liquid flowrate estimation. The RFR-enhanced model reduced the Relative Absolute Error (RAE) from 10%, as observed with the Bruggeman baseline model, to below 5%, marking a substantial gain in predictive fidelity.

The most notable advancement is captured in Fig. 8, where RFR achieved an average RAE of just 2.23% for LVF prediction. This represents an order-of-magnitude improvement over the Bruggeman model, which yielded an RAE of 21.19%. The overall relative uncertainty of the RFR-integrated microwave sensing system reflects a sixfold reduction compared to the analytical baseline, underscoring its superior performance in complex multiphase environments

5 Discussions

The primary objective of this study was to determine whether integrating microwave sensing with advanced ML regression models could improve LVF estimation in wet gas flows beyond the accuracy of established physics-based approaches. The results confirm that physics-informed, data-driven models can significantly outperform the Bruggeman dielectric mixing model in mist-flow regimes, where droplet dispersion, slip effects, and sensor noise reduce the reliability of purely analytical formulations [9]. The Bruggeman model remains attractive for its physical interpretability, closed-form computation, and independence from training data, and in this study, it achieved an average LVF relative absolute error (RAE) of approximately 21.2% and a liquid flowrate RAE near 10%. However, its simplifying assumptions: spherical droplets, homogeneous dispersion, and negligible slip limit its applicability under dynamic, high-GVF conditions, consistent with previous findings that analytical dielectric models lose accuracy when phase distributions are unstable [6,12].

Table 3 Performance metrics of machine learning models for LVF prediction

Performance metric	SVR	RFR	GPR
RMSE	0.08264	0.00509	0.00167
R^2	−0.16628	0.99559	0.99955
Durbin–Watson	1.06850	2.02286	1.90533
MAPE	450.85%	2.0143%	1.0392%

By contrast, the ML models, particularly RFR and GPR, learned complex nonlinear relationships between microwave-derived dielectric properties and hydrodynamic parameters without imposing restrictive geometric or homogeneity assumptions. This flexibility allowed them to capture subtle interactions between flow dynamics and electromagnetic response that the Bruggeman model could not. In this work, RFR reduced LVF RAE to 2.23% and liquid flowrate RAE to below 5%, representing an order-of-magnitude improvement over the physics-based baseline. These gains are in line with other studies where ML integration with multiphase sensing has narrowed error margins compared to empirical or analytical models, especially at high GVF and under environmental perturbations, as reported in studies such as Hosseini et al. [15] and Wang et al. [20].

Although GPR achieved the lowest RMSE (0.00167) and highest R^2 (0.99955), RFR was ultimately selected for integration with the microwave sensing framework due to its superior residual independence (Durbin–Watson ≈ 2.02), robustness to noise, and computational efficiency in real-time applications. GPR's reliance on covariance matrix inversion increases computational cost and susceptibility to overfitting in high-dimensional, noisy environments, while SVR exhibited markedly higher errors (RMSE = 0.08264, MAPE > 450%), confirming its limited suitability for stochastic dielectric signals in mist-flow regimes. Feature importance analysis further reinforced the dominance of the reconstructed mixture dielectric constant (ε_{mix}), which accounted for over 97% of RFR's predictive capability. This finding aligns with electromagnetic theory and is consistent with prior microwave sensing studies in dilute liquid–gas mixtures, as reported by Liu et al. [18], Ma et al. [8] and Xu et al. [7], who demonstrated that dielectric permittivity is the most sensitive and reliable indicator of phase composition in low liquid loading regimes. The relatively minor contribution of Reynolds numbers, Weber number, slip ratio, and dielectric constant ratio reflects the limited variability of these

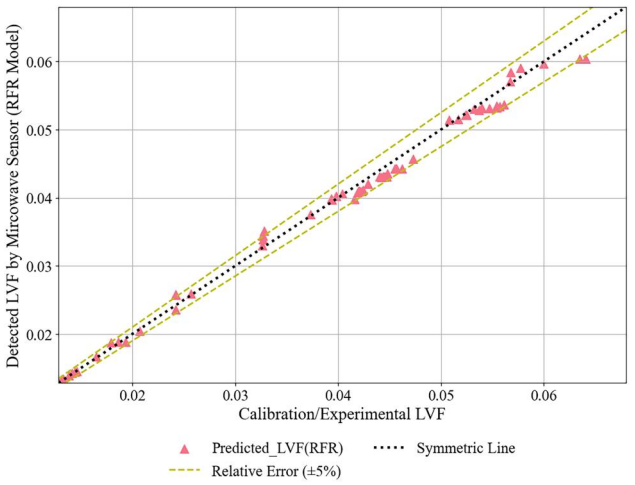

Fig. 6 Correlation of microwave detected liquid volume fraction based on random forest regression (LVF-RFR) with calibration liquid volume fraction (LVF)

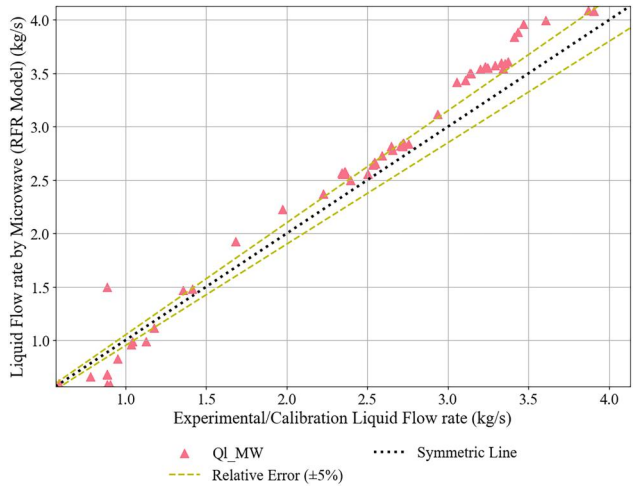

Fig. 7 Correlation of detected liquid flow rate by microwave (Q_r RFR) with calibration liquid flow rate determined by reference flowmeter (Q_l)

parameters under the controlled test conditions, though literature suggests they may become more influential in field environments with variable fluid properties and regime transitions [12,21]. This pronounced influence can be attributed to several factors: the dielectric constant ratio remained constant due to the fixed air–water composition; mist flow conditions led to significant liquid dispersion, which diminished hydrodynamic sensitivity; and overlapping physical definitions resulted in multicollinearity.

Generalization analysis showed that both RFR and GPR maintained high predictive accuracy when trained on as little as 50% of the available data, with only modest increases in RMSE and MAPE, while SVR degraded more rapidly. Noise-robustness testing, involving Gaussian perturbations of 1–5% to ε_{mix} and dimensionless features, revealed that RFR was more resilient than GPR at higher noise levels ($\geq 3\%$), consistent with the ensemble model's inherent noise-averaging properties. GPR retained a slight advantage in low-noise scenarios but declined more sharply as noise increased. These findings align with prior reports that ensemble methods offer superior robustness to measurement variability and environmental disturbances in multiphase metering [15,22].

Overall, the results demonstrate that while the Bruggeman model remains a valuable, fast, and interpretable baseline, its structural assumptions constrain accuracy in complex, high-GVF wet gas

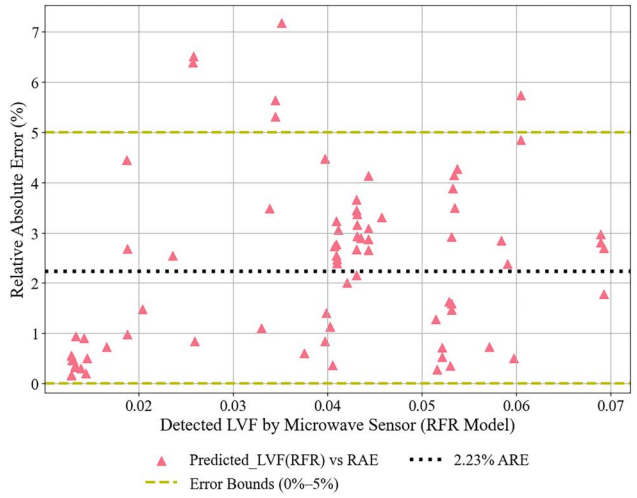

Fig. 8 Relative absolute error (RAE) for the microwave sensor in liquid detection based on random forest regression model (LVF_RFR)

flows. Physics-informed RFR not only meets but exceeds the study's objective by delivering accurate, stable, and interpretable LVF estimates across the full mist-flow range, scaling efficiently for real-time processing, and maintaining performance under reduced training data and noisy measurement conditions. These attributes, combined with scalability and operational resilience, make RFR a strong candidate for integration into industrial wet gas metering systems, with the potential to reduce calibration requirements and improve reliability in challenging field environments.

6 Limitations

While the results demonstrate the strong potential of integrating Random Forest Regression with microwave sensing for accurate LVF and liquid flowrate estimation, several constraints remain. The use of air–water mixtures, though replicating key dielectric contrasts, does not capture the full thermodynamic, compositional, and phase-behavior complexities of hydrocarbon–water systems, including effects of gas compressibility, condensates, salinity, and temperature gradients. Tests were limited to a fixed pipe diameter, moderate pressures (1.2–25 bar(a)), and temperatures (20–40 ° C), which may affect scalability, so scaling to larger diameters, higher pressures, or more variable thermal environments may introduce additional uncertainties. Sensor performance was assessed mainly in stratified and mist-flow regimes, leaving slug, churn, and annular flows unexamined. Finally, model generalization to industrial conditions depends on training data representativeness, calibration stability, and robustness under higher noise and more complex flow compositions. Addressing these gaps through field trials, broader flow regime coverage, and adaptive or hybrid modeling will be essential for reliable deployment in diverse wet gas metering applications.

7 Conclusions

This study demonstrated that integrating microwave sensing with machine learning regression models significantly improves the estimation of liquid volume fraction (LVF) and liquid flowrate in wet gas flows. Comparative evaluation of the Bruggeman dielectric mixing model with Support Vector Regression (SVR), Gaussian Process Regression (GPR), and Random Forest Regression (RFR) under mist-flow conditions (GVF 95–99.9%) revealed that RFR offers the best balance of accuracy, robustness, and computational efficiency. RFR achieved a Relative Absolute Error (RAE) of 2.23% for LVF and approximately 5% for liquid flowrate, with minimal residual autocorrelation (Durbin–Watson ≈ 2.02). Feature importance analysis confirmed the reconstructed mixture dielectric constant (ε_{mix}) as the dominant predictor, contributing over 97% to model accuracy, while other dimensionless parameters had negligible influence. SVR failed to generalize effectively, and although GPR exhibited slightly higher accuracy, its computational demands and sensitivity to noise limited its suitability for real-time deployment.

Integrating Random Forest Regression (RFR) with microwave sensing offers a scalable, nonintrusive solution for accurate wet gas metering. By reducing LVF estimation error from $\sim$21% (Bruggeman model) to 2.23% and liquid flowrate error from $\sim$10% to $< 5\%$, the framework meets industrial standards for allocation metering, corrosion monitoring, and flow assurance. Its robustness to limited training data and sensor noise makes it ideal for subsea and remote deployments. The results underscore the effectiveness of physics-informed ML models for multiphase flow diagnostics. Future research should validate the approach under industrial conditions, including hydrocarbon–water mixtures, elevated pressures, and broader temperature ranges and extend flow regime coverage to slug, churn, and annular patterns. Field trials and adaptive calibration strategies, including hybrid physical ML frameworks, may further enhance generalization across high-GVF, low-liquid-loading scenarios.

Acknowledgment

I would like to express my sincere gratitude to the Multiphase Process Engineering (PSE) Laboratory at Cranfield University for

providing access to its wet gas flow facilities. Special thanks to Dr. Liyun Lao for his expert supervision and commitment, and to Mr Stan Collins for his exceptional support in laboratory setup and operations, which were crucial for the success of the Multiphase Flow Rig tests.

Nomenclature

D = internal diameter of flow pipe, m
f = frequency of microwave signal, Hz
k = complex propagation constant $k' - jk''$, 1/m
$K(x, x_i)$ = Kernel function in SVR
k' = real part of propagation constant (attenuation), 1/m
k'' = imaginary part of propagation constant (phase shift), 1/m
L = geometric factor (assumed 1/3 for spherical droplets), dimensionless
M = number of trees in RFR ensemble
Q_l = liquid flow rate, m^3/s
Q_{tp} = total two-phase flow rate, m^3/s
R^2 = coefficient of determination
Re_l = liquid Reynolds number, dimensionless
Re_g = gas Reynolds number, dimensionless
S = slip ratio (gas velocity/liquid velocity), dimensionless
$T_m(x)$ = prediction from m^{th} tree in RFR
We = Weber number, Dimensionless
x^* = test input vector in GPR prediction
GVF = gas volume fraction $(1 - LVF)$, dimensionless
LVF = liquid volume fraction, dimensionless
MAPE = mean absolute percentage error, %
MSE = root mean squared error
σ = electrical conductivity of the liquid, S/m
μ = magnetic permeability, H/m
ω = angular frequency $(2\pi f)$, rad/s
ε_g = dielectric constant of the gas phase, dimensionless
ε_l = dielectric constant of the liquid phase, dimensionless
$\varepsilon_{\mathrm{mix}}$ = effective dielectric constant of the gas–liquid mixture, dimensionless
ε^* = complex dielectric constant $\varepsilon' - j\varepsilon''$, dimensionless
$\Delta\theta$ = phase shift of microwave signal, radians
ΔA = amplitude attenuation, dB
ρ_g = gas density, kg/m^3
ρ_l = liquid density, kg/m^3
μ_g = gas dynamic viscosity, Pa·s
μ_l = liquid dynamic viscosity, Pa·s
U_g = superficial gas velocity, m/s
U_l = superficial liquid velocity, m/s
α_i, α_i^* = support vector coefficients in SVR

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Li, Y., Yang, W., Xie, C-G., Huang, S., Wu, Z., Tsamakis, D., and Lenn, C., 2013, "Gas/Oil/Water Flow Measurement by Electrical Capacitance Tomography," Meas. Sci. Technol., **24**(7), p. 074001.
[2] Wang, J., Huang, Z., Xu, Y., and Xie, D., 2024, "Gas–Liquid Two-Phase Flow Measurement Based on Optical Flow Method With Machine Learning Optimization Model," Appl. Sci., **14**(9), p. 3717.
[3] Al-Kizwini, M. A., Wylie, S. R., Al-Khafaji, D. A., and Al-Shamma'a, A. I., 2013, "Monitoring of the Two-Phase Annular Flow Regime Using Microwave Sensor Technique," Measurement, **46**(1), pp. 45–51.
[4] Xie, C. G., and Wu, Z., 2012, "Microwave Doppler System for Multiphase Flow Measurement," AIP Conf. Proc., **1428**, pp. 319–326.
[5] El-Aziz, D. A. A., Attiya, A. M., Elhennawy, H. M., and Ali, R. S. M., 2025, "Comparative Study of Microwave Relative Phase Measurement Methods," Mapan - J. Metrol. Soc. India, **40**, pp. 1–14.
[6] Lin, X., Wang, H., Chen, Z., Zhang, H., and Li, Y., Oct. 2020, "Measurement of the Flow Rate of Oil and Water Using Microwave and Venturi Sensors With End-to-End Dual Convolutional Neural Network," Meas. Sens., **10-12**, p. 100018.
[7] Liu, C., Liao, C., Peng, Y., Zhang, W., Wu, B., and Yang, P., 2024, "Microwave Sensors and Their Applications in Permittivity Measurement," Sensors, **24**(23), p. 7696.
[8] Ma, H., Xu, Y., Yuan, C., Yang, Y., Zuo, R., Liu, J., and Li, T., 2023, "A Water Fraction Measurement Method of Gas-Water Flow in a Wide Conductivity Range," Measurement , **216**, p. 112895.
[9] Tayyab, M., Sharawi, M. S., and Al-Sarkhi, A., 2017, "A Radio Frequency Sensor Array for Dielectric Constant Estimation of Multiphase Oil Flow in Pipelines," IEEE Sens. J., **17**(18), pp. 5900–5907.
[10] Sabzevari, F. M., Winter, R. S. C., Oloumi, D., and Rambabu, K., 2020, "A Microwave Sensing and Imaging Method for Multiphase Flow Metering of Crude Oil Pipes," IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, **13**, pp. 1286–1297.
[11] Oon, C. S., Ateeq, M., Shaw, A., Wylie, S., Al-Shamma'a, A., and Kazi, S. N., 2016, "Detection of the Gas–Liquid Two-Phase Flow Regimes Using Non-Intrusive Microwave Cylindrical Cavity Sensor," J. Electromagnetic Waves Appl., **30**(17), pp. 2241–2255.
[12] van Maanen, H. R. E., 2007, "Measurement of the Liquid Water Flow Rate Using Microwave Sensors in Wet-Gas Meters: Not as Simple as You Might Think," Energy Institute – 26th International North Sea Flow Measurement Workshop, St. Andrews, Scotland, UK, Oct. 21–24, pp. 345–354.
[13] Pochet, S., Teyssedou, A., and Akyel, C., 2014, "Development and Assessment of a Microwave Void-Fraction Measurement System," Rev. Sci. Instrum., **85**(1), p. 015103.
[14] Sharma, P., Lao, L., and Falcone, G., 2018, "A Microwave Cavity Resonator Sensor for Water-in-Oil Measurements," Sens. Actuators B: Chem., **262**, pp. 200–210.
[15] Hosseini, S., Chinello, G., Lindsay, G., Smith, S., and McGlinchey, D., 2025, "Multiphase Flow Measurement of Wet Gas Flow Using Machine Learning Modelling Algorithms," Meas. Sens., **38**, p. 101556.
[16] Zhang, D., and Xia, B., 2014, "Soft Measurement of Water Content in Oil-Water Two-Phase Flow Based on RS-SVM Classifier and GA-NN Predictor," Meas. Sci. Rev., **14**(4), pp. 219–226.
[17] Lund Bø, Ø., and Nyfors, E., 2002, "Application of Microwave Spectroscopy for the Detection of Water Fraction and Water Salinity in Water/Oil/Gas Pipe Flow," J. Non-Cryst. Solids, **305**, pp. 345–353.
[18] Liu, W., Sun, H., and Xu, L., 2018, "A Microwave Method for Dielectric Characterization Measurement of Small Liquids Using a Metamaterial-Based Sensor," Sensors, **18**(5), p. 1438.
[19] Bao, M., Wang, M., Li, K., and Jia, X., 2024, "Integrating Machine Learning With Sensor Technology for Multiphase Flow Measurement," IEEE Sens. J., **24**(19), pp. 29603–29618.
[20] Wang, H., Zhang, M., and Yang, Y., 2020, "Machine Learning for Multiphase Flowrate Estimation With Time Series Sensing Data," Meas. Sens., **10-12**, p. 100025.
[21] Makeev, Y. V., Lifanov, A. P., and Sovlukov, A. S., 2013, "Microwave Measurement of Water Content in Flowing Crude Oil," Autom. Remote Control, **74**(1), pp. 157–169.
[22] Saparbayeva, N., Balakin, B. V., Struchalin, P. G., Rahman, T., and Alyaev, S., 2024, "Application of Machine Learning to Predict Blockage in Multiphase Flow," Computation, **12**(4), p. 67.