Global Flow Measurement Workshop
22 - 24 October 2024

Technical Paper

Enhancing Wet Gas Metering Capabilities Using Machine Learning: Implications of Data-Driven Models and Integration with Physical Principles

SeyedAhmad Hosseini, Glasgow Caledonian University
Gabriele Chinello, TÜV SÜD National Engineering Laboratory
Gordon Lindsay, Glasgow Caledonian University
Don McGlinchey, Glasgow Caledonian University

---

1 INTRODUCTION

In industrial energy and process design applications, accurate measurement of multiphase flow is of significant importance. Wet gas, which consists primarily of gas with a small amount of liquid (either condensed water and/or liquid hydrocarbon), forms because of changes in the thermophysical properties as the gas moves from the reservoir to the surface or because of formation water presence from the reservoir. Accurate measurement of wet gas is essential for various purposes, including reservoir management, production optimisation, fiscal allocation, process control, and compliance with environmental regulations. Furthermore, inaccurate wet gas measurements can have significant financial implications because they may lead to considerable discrepancies in revenue allocation and production accounting [1–4].

Wet gas is commonly defined as a gas-liquid mixture with a Gas Volume Fraction (GVF) higher than 95% or when the Lockhart-Martinelli (X_{LM}) is below 0.3. The Lockhart-Martinelli (X_{LM}) parameter is defined as follows:

$$X_{LM} = \frac{\dot{m}_l}{\dot{m}_g}\sqrt{\frac{\rho_g}{\rho_l}} = \frac{1-\mathrm{GVF}}{\mathrm{GVF}}\sqrt{\frac{\rho_l}{\rho_g}} \tag{1}$$

For decades, the Oil and Gas industry has employed various flow metering technologies for wet gas flow measurement. The Venturi tube has emerged as a particularly favoured option because of its durability and capacity to generate less permanent pressure loss, which is often beneficial in practical settings. The primary hurdle for DP flow meters is the quantification of the liquid content (wetness) within the stream. Numerous methods have been devised to determine liquid flow rate, including test separators, sampling, tracers, microwave technology (specific to water), overall pressure loss across differential pressure meters, infrared spectroscopy, and sophisticated signal processing. However, each approach has its own constraints in terms of time, cost efficiency, reliability, application scope, and other operational issues. Moreover, wet gas metering is further complicated by factors such as phase slip, varying velocities of liquid and gas phases, and the intricacy of flow regimes that fluctuate between annular, mist, and slug flows, contingent on the pipe orientation, gas and liquid flow rates, and fluid characteristics. Precise measurements are hindered by nonequilibrium conditions, liquid loading, and instrumentation limitations. Additionally, flow regimes and diverse gas-liquid ratios make the calibration more complex, whereas harsh operating environments can affect the longevity and efficiency of meters [5–7].

It is worth noting that further assessment of wet gas metering is beneficial in moving towards energy transition and using clean fuel sources, particularly in the metering of carbon capture and storage (CCS) and hydrogen infrastructure. The carbon capture, utilisation, and storage (CCUS) transport chain can experience wet gas flow depending on the temperature and pressure conditions and impurity levels in the CO_2 stream [8]. It is well known that wet-gas flow has a substantial negative impact on the accuracy of single-phase flow meters performance. Wet-gas conditions may arise in CCS transportation pipelines because of events such as pressure drop through equipment (e.g. a choke valve), local temperature drop, transient operations (e.g. pipeline pressurisation and depressurisation), suboptimal design of equipment, or equipment failure [9].

When employing differential pressure meters to measure wet-gas flow, the liquid amount (X_{LM}) flowing into the pipe should be known to correct for meter over-reading. The development of the Pressure Loss Ratio (PLR) method, such as ISO-11583 (2012) [10], to determine X_{LM} represents a notable advancement. However, this method is currently limited to low X_{LM} values, typically below 0.06, leading to diminished applicability. Advancements in data analytics and machine learning (ML) could offer solutions to address this limitation.

Using extensive datasets and sophisticated algorithms, ML methods can discern intricate nonlinear correlations between flow conditions, fluid characteristics, and metering outputs [11–16]. Over time, these data-driven models can enhance their accuracy and dependability by adjusting to shifting conditions and by assimilating new data. Moreover, ML approaches are particularly adept at handling incomplete or noisy data, which is a frequent issue in practical metering applications where sensor malfunctions or communication interruptions may arise as well as signal noise.

2 Objective

This study presents a novel method for addressing the complexity of accurate wet-gas measurements over an extended X_{LM} range. It utilises sophisticated machine learning (ML) techniques through a purely data-driven approach. The primary objective is to create and assess these ML models for predicting liquid and gas flow rates and their efficacy under various operating conditions by relying solely on sensor data. This study aims to extend the applicability of wet gas metering with Venturi tubes at higher X_{LM} values by employing advanced data-driven methods that may be able to outperform conventional techniques. This study has the potential to considerably enhance the creation of more efficient and economical metering solutions for the oil and gas sector. Furthermore, it could broaden applications in fields such as carbon capture and storage (CCS) initiatives and the development of hydrogen infrastructure. These outcomes seek to boost the operational efficiency, economic performance, and environmental sustainability of the energy industry. The study concludes by examining the implications and limitations of data-based models, and proposing potential solutions to overcome these challenges.

3 Methodology

3.1 Experimental Setup and Data Collection

This study collected data at the NEL's multiphase flow facility, which is one of the world's leading flow measurement facilities. The experimental setup featured a 6-inch pipe with a Venturi tube having a 0.6 beta ratio. Nitrogen was used as the gas phase, and water and Exssol D140 were used as the liquid phases in the fluid system, allowing different wet gas flow scenarios to be simulated. Experiments were conducted under a wide range of operating conditions.

- Pressure (P): 20-60 bar
- Temperature (T): 19-28°C
- Water cut: 0 – 100%
- Flow regimes: From stratified wavy to annular flow
- Lockhart-Martinelli parameter (X_{LM}): Up to 0.5, extending the model's applicability to higher liquid loading conditions

The accompanying illustration depicts the allocation of the experimental conditions on the Shell flow regime map. The diagram indicates that the test data span from stratified-wavy flow to annular flow, with some data points situated near the transition boundary between the annular and slug flows.

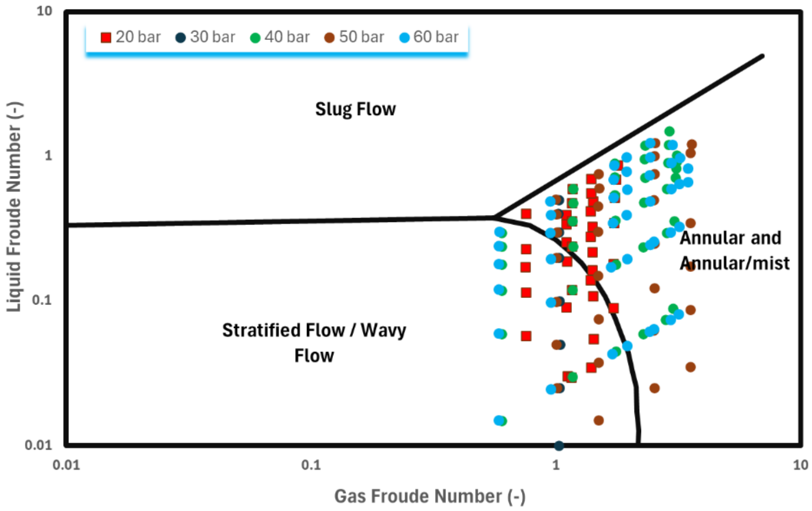
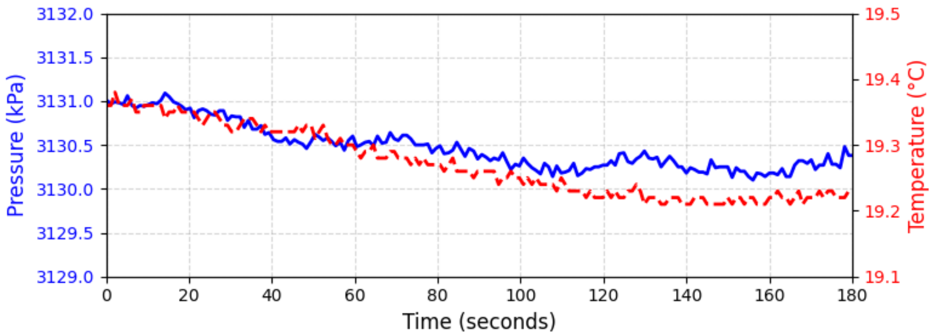


Figure 1. Experimental scenarios depicted on the Shell flow pattern diagram

3.2 Experimental Setup and Data Collection

An experimental arrangement utilising a Venturi tube configuration was employed to compile a comprehensive wet-gas flow analysis dataset. Four crucial parameters were monitored using the sensors: temperature (T), pressure (P), pressure drop across the venturi (PPL), and differential pressure at the throat (DP). To ensure quality, consistency, and optimal performance for subsequent modelling, the initial data were preprocessed. The Z-score method was applied to standardise all the input features, adjusting them to have a mean of 0 and a standard deviation of 1. This normalisation process aids in aligning the scale of various features and enhancing model convergence. A standard deviation threshold of three was used to identify and eliminate outliers, thereby minimising the impact of anomalous measurements. The refined dataset was subsequently divided into training (70%), validation (15%), and test (15%) sets to ensure separate data for model development and evaluation. This meticulous preprocessing approach established a solid foundation for subsequent ML modelling, thereby enhancing the reliability and accuracy of the wet gas flow predictions. Figure 2 illustrates a sample of raw signals for P, T, DPt, and PPL under representative experimental conditions, demonstrating the variation over a 180-second testing period.



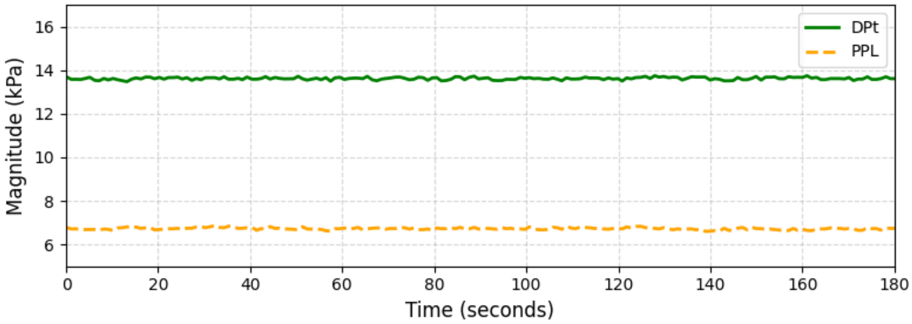


Figure 2. An illustration of unprocessed sensor readings obtained from a Venturi tube under conditions of P: 31 barg, mg: 3.5 kg/s, and ml: 8.83 kg/s

3.2 Machine Learning Models

This study utilised a variety of neural network models based on the following:

1. Deep Neural Network (DNN): Learning complex nonlinear relationships through feedforward neural networks.
2. Convolutional neural networks (CNN): Analyse the input data for spatial correlations.
3. Long Short-Term Memory (LSTM): Represents long-term dependencies on sequential data.
4. Gated Recurrent Unit (GRU): Sequential data processing using a simpler version of LSTM.

3.4 Model Training and Optimization

The training and optimisation of the model were meticulously designed to ensure peak performance and broad applicability across various ML models. The process began with K-fold cross-validation, chosen for its comprehensive assessment of model efficacy and ability to reduce overfitting by evaluating the model on different data subsets. Following this, hyperparameter optimisation was conducted using grid search and cross-validation, allowing for systematic exploration of each model's parameter space. This includes refining the number of layers and nodes in the neural network structures and adjusting the model-specific variables for other algorithms. To prevent overfitting during training, early stopping was implemented, which halted the process when the model's performance on the validation set ceased to improve. The Adam optimiser was selected for its adaptive learning rate capabilities, which can lead to faster convergence and enhanced performance in deep learning models. The initial learning rate (0.001) was carefully chosen to strike a balance between the convergence speed and optimisation stability. For the training process, a batch size of 64 was utilised for DNN and CNN models, whilst a reduced batch size of 32 was employed for LSTM and GRU models. This adjustment was made to strike a balance between computational efficiency and the ability to capture data variability, taking into account the higher memory demands and computational complexity of recurrent models. The smaller batch size for LSTM and GRU is a standard approach that helps address memory limitations whilst still enabling effective learning of temporal dependencies. This reduction also aligns with the sequential processing nature of these models, potentially facilitating more frequent weight updates and improved convergence on time-series data. To enhance the model's ability to generalise, l2 regularisation was implemented, aiming to achieve an optimal balance between the algorithm's bias and variance. The framework for flow rate prediction using neural network algorithms is depicted in the flowchart presented in Figure 3.

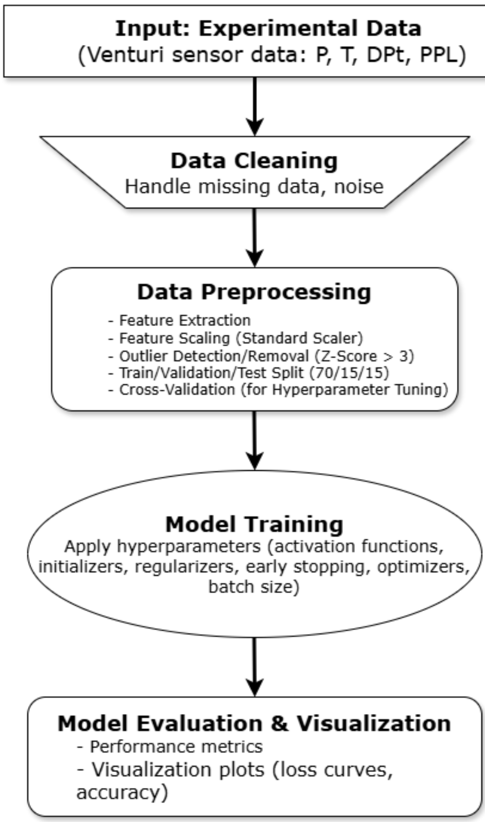


Figure 3. Schematic representation of the Neural Network algorithms' training process

3.5 Validation Methodology

The assessment methodology employed a comprehensive array of performance indicators to meticulously assess the precision and dependability of the wet gas metering models. These indicators included the root mean squared error (RMSE), mean absolute percentage error (MAPE), and standard deviation (STD). The RMSE offers an interpretation of the units of the target variable. MAPE was selected as the reference metric because of its intuitive interpretation as a percentage error, making it particularly valuable for comparing model performance across diverse flow rates and conditions. MAPE expresses the accuracy as an error percentage, delivering a clear and easily comprehensible measure of model performance. Finally, STD provides an idea of how much variability is present in prediction errors. The RMSE, MAPE, and STD formulations are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{2}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3}$$

$$STD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left(y_i - \hat{y}_i\right) - \mu\right)^2} \tag{4}$$

where:

- μ is the mean error, calculated as $\mu = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)$.
- n is the number of test samples.

- $\hat{y}_i$ is the predicted flowrate for the i-th sample.
- y_i is the reference (true) flowrate for the i-th sample.

4. Result and discussion

Figure 4 illustrates the loss curves for training and validation across the four models (DNN, CNN, LSTM, and GRU). These graphs plot the loss (Mean Square Error) against the number of iterations, providing insights into how each model learns and generalises over time. In machine learning, an iteration represents one complete pass through the training data, whereas a batch is a subset of the training data used in a single update of the model's parameters. The relationship between batches and iterations (*I*) can be expressed as:

$$I = \frac{N}{B} \times E \qquad (5)$$

where, *N* is the total number of samples in the dataset, *B* is the batch size, and *E* is the number of epochs.

The loss error plots for the four neural network models (DNN, CNN, LSTM, and GRU) illustrated the performance of each model during training and validation. Because of the early stopping during training, each model halts at specific iterations when the improvements plateau. Both LSTM and GRU showed the fastest and most effective convergence, with minimal gaps between the training and validation losses, indicating excellent generalisation and low final loss values.

The DNN model, while stable, converges more slowly and ends with a higher final loss, suggesting that it is less capable of capturing the complexity of the data than sequential models. Finally, despite a rapid decrease in training loss, the CNN model showed a significant gap between the training and validation losses, indicating overfitting and poor generalisation. In summary, the LSTM and GRU consistently outperformed the other models by achieving low loss values and maintaining minimal overfitting, making them the most suitable models for accurate gas and liquid flow rate prediction.

The results are the mean squared errors (MSE) for the training and validation sets over multiple iterations. The loss curves demonstrate the capacity of the models to generalise and prevent overfitting, while achieving high accuracy. Each graph illustrates a unique training behaviour, highlighting the specific characteristics of the model architectures and reactions to the input data. This study's four ML models, DNN, CNN, LSTM, and GRU, showed distinct abilities for managing wet gas flow data complexities. Despite the occasional instability in the validation loss, the DNN effectively captured nonlinear relationships. CNN can detect spatial correlations well, but overfitting is an issue. The LSTM and GRU models, tailored to handle sequential data, delivered consistent and stable performance across epochs. GRU is more computationally efficient. Accordingly, the wet gas flow is best predicted using LSTM and GRU, particularly when time-dependent patterns are essential.

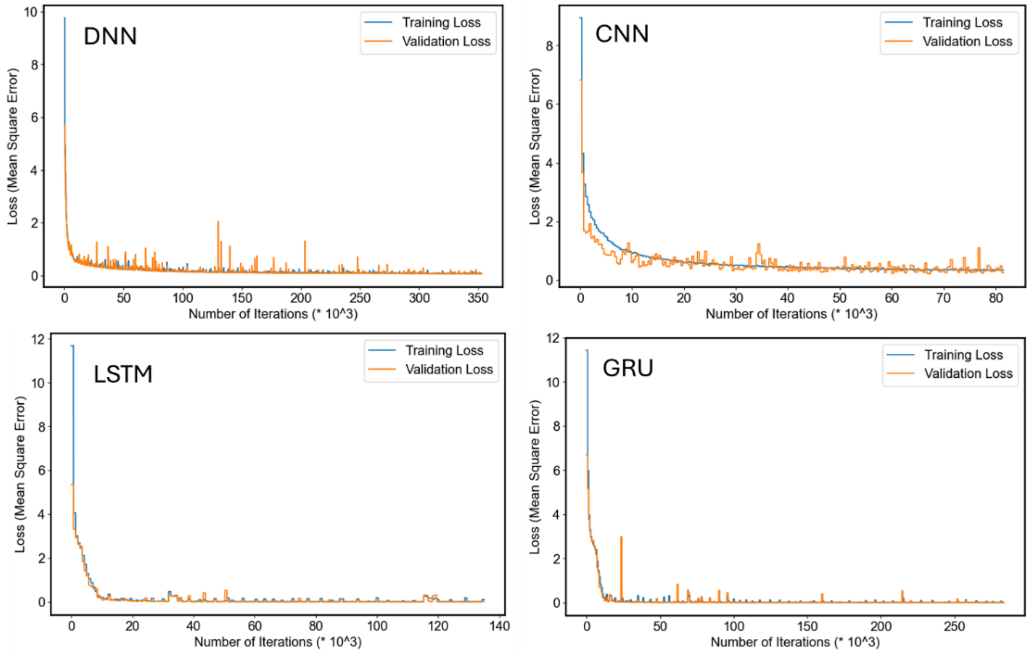


Figure 4. Loss curves for training and validation across the four models

Figure 5 illustrates the comparison between the predicted and reference flow rates for both the gas and liquid across the four models. The plots reveal distinct variations in accuracy and performance among the models. For gas flow rate predictions, the LSTM and GRU models demonstrated superior performance, with predictions closely aligned with a perfect fit line and minimal scatter. This is corroborated by the metrics in Table 1, where LSTM and GRU exhibit the lowest RMSE values (0.0349 and 0.0252, respectively) and MAPE values (0.52% and 0.37%, respectively) for the test set. The DNN model shows acceptable accuracy with slightly higher RMSE (0.0496) and MAPE (0.69%) values, although there is more visible scatter in Figure 5. However, the CNN model exhibits significant scatter and deviation from the perfect fit line, reflected in its high RMSE (0.2338) and MAPE (2.43%) values.

A similar trend was observed for the liquid flow rate predictions. The LSTM and GRU models again outperform the others, with tight clustering around the perfect fit line in Figure 5 and low RMSE (0.059 and 0.053, respectively) and MAPE (2.0% and 2.1%, respectively) values, as shown in Table 1. The DNN model showed moderate performance with higher error rates (RMSE: 0.125, MAPE: 3.1%), whereas the CNN model struggles considerably with the highest RMSE (0.363) and MAPE (7.0%) among all models. Figure 6 provides further insight into model performance by plotting the relative error against the Lockhart-Martinelli (X_{LM}) parameter. This figure reveals that the LSTM and GRU models maintain consistent performance across different X_{LM} values, with most predictions falling within the ±3% relative error range for both the gas and liquid flow rates. The DNN model showed increased variability, particularly for liquid flow rates at lower X_{LM} values. The CNN model exhibited the highest variability and largest errors across the entire X_{LM} range for both gas and liquid predictions. The superior performance of LSTM and GRU can be attributed to their ability to capture temporal dependencies in the data, which is crucial for accurate flow rate prediction. These models utilise memory cells and gating mechanisms to retain relevant information over time, allowing them to better model the complex dynamics of gas-liquid flow.

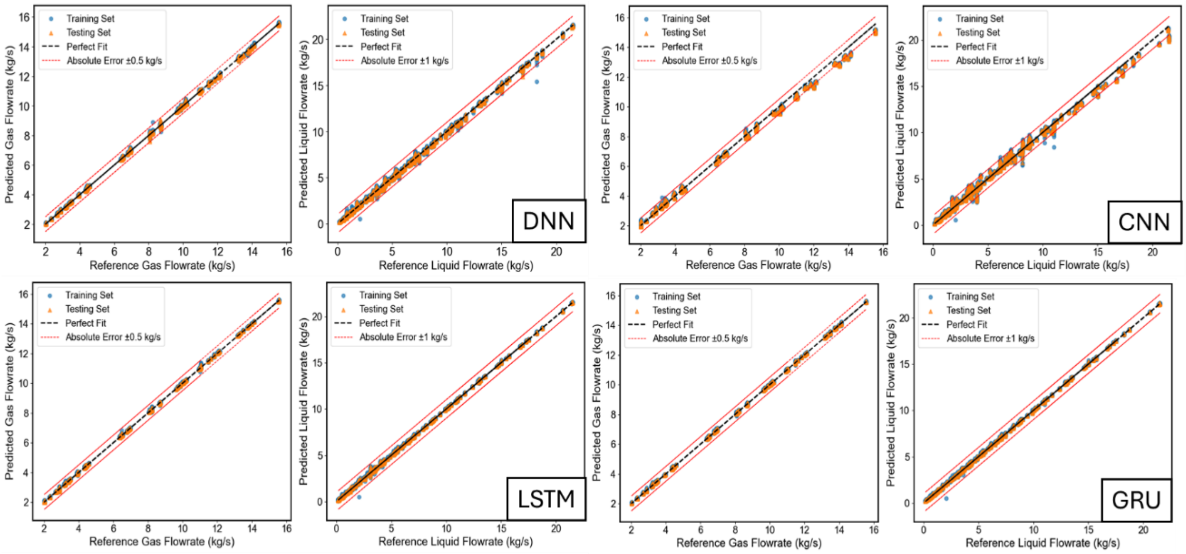


Figure 5. Comparison of actual and estimated gas and liquid flow rates for training and testing datasets across multiple models.

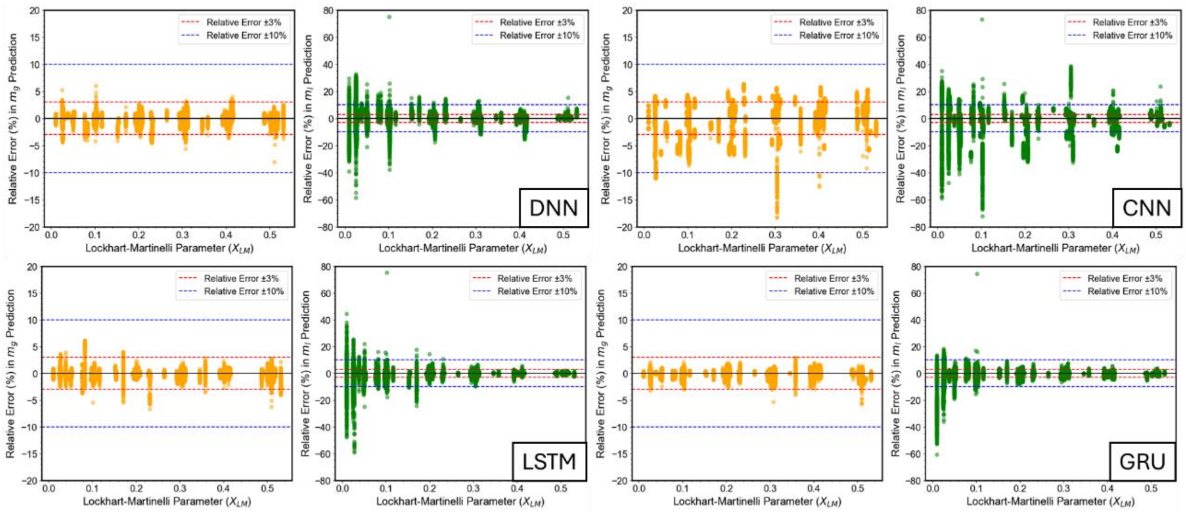


Figure 6. Comparative analysis of gas and liquid flow rate prediction relative error (%) versus the Lockhart-Martinelli (X_{LM})

The DNN, which is simpler in architecture, still manages to capture some of the underlying patterns, particularly for gas flow rate prediction. However, its performance lags behind that of recurrent models, especially for liquid flow rates, where the dynamics may be more complex. The poor performance of the CNN across both gas and liquid predictions suggests that its architecture, typically suited for spatial data and image processing, struggles to capture the temporal and multidimensional aspects of the flow rate data effectively.

Based on the analysis of the plots and metrics presented in Table 1, the LSTM and GRU models consistently outperform DNN and CNN in both gas and liquid flow rate predictions. These recurrent models demonstrate superior accuracy, lower error rates, and better generalisation across different flow regimes, making them the most suitable choice for wet gas flow metering applications where precise predictions are critical.

Table 1. Evaluation metrics of investigated models based on train and test data

Models	Metrics	Gas Flowrate Train	Gas Flowrate Test	Liquid Flowrate Train	Liquid Flowrate Test
DNN	RMSE	0.0477	0.0496	0.1265	0.1259
	MAPE	0.0069	0.0069	0.0325	0.0314
	STD	0.0477	0.0496	0.1202	0.1195
CNN	RMSE	0.2303	0.2338	0.3512	0.3635
	MAPE	0.0242	0.0243	0.0779	0.0766
	STD	0.2159	0.2183	0.3498	0.3625
LSTM	RMSE	0.034	0.0349	0.0591	0.0598
	MAPE	0.0051	0.0052	0.0203	0.0202
	STD	0.0334	0.0345	0.0587	0.0594
GRU	RMSE	0.0256	0.0252	0.0547	0.0536
	MAPE	0.0037	0.0037	0.0220	0.0216
	STD	0.0253	0.0250	0.0539	0.0527

5. Implications of Pure Data-Driven Modelling and Limitations

Similar to the neural networks examined in this study, purely data-driven modelling offers significant advantages in predicting complex systems such as wet gas metering. These approaches can directly extract nonlinear relationships from data without depending on explicit physical principles, thus rendering them flexible and adaptable to diverse operational scenarios. Despite their strengths, they have inherent constraints, particularly when extrapolating to unfamiliar situations or to test sets. In many instances, these models fail to produce accurate predictions because of inadequate, incomplete, or inconsistent data collection. Modelling real-world flow regimes and fluid compositions presents challenges, particularly under field conditions where substantial variations occur. Data-driven models often require extensive representative datasets to prevent overfitting. As they do not incorporate the underlying physics of the system, such as conservation laws, these models can be difficult to interpret and may lack consistency. Consequently, predictions based on such models can be unreliable under certain circumstances, potentially lacking physical significance, or violating fundamental principles when extrapolated.

6. Hybrid Approaches and Integration of Physical Principles

To address the shortcomings of purely data-driven models, a combined approach integrating neural networks with physical principles has been proposed for more comprehensive and resilient wet gas metering solutions [17]. By incorporating physical models (such as the Bernoulli principle) into machine learning, systems can be developed to comply with fundamental fluid mechanics principles, ensuring precise and consistent predictions, even in situations where data are limited or unavailable. This integration can be achieved by incorporating physical constraints into the loss function, or by preprocessing the data and generating features for the neural network. The fusion of machine learning with physical principles enables hybrid systems to adapt better to various flow regimes and conditions, thereby reducing the need for extensive datasets. This allows the models to manage complex flow regimes, including annular or slug flows, at high gas-to-liquid ratios. The aim was to enhance the reliability and generalisability of the model.

7. Conclusions

In summary, the data-driven models analysed in this study, notably LSTM and GRU, exhibited robust predictive abilities for gas and liquid flow rates during wet gas metering. These models demonstrate superior performance across various flow regimes, particularly in their consistent accuracy across different Lockhart-Martinelli parameter values. However, limitations inherent to purely data-driven approaches persist in their generalisability, as they may struggle to extrapolate to unfamiliar conditions or lack consistency when data are scarce. To address these challenges, future research should focus on integrating feature engineering and incorporating physical principles such as conservation laws into neural network models. Implementing physics-informed constraints or developing hybrid models can enhance the robustness and reliability of predictions, enabling better generalisation across various flow regimes and operational conditions. This strategy diminishes reliance on extensive datasets and ensures that the models maintain physical validity, resulting in more consistent and interpretable outcomes in practical applications. By combining the strengths of data-driven approaches with physics-based constraints, future wet-gas metering systems can achieve higher accuracy, improved generalisability, and greater reliability across a broader range of operating conditions. Further investigations are currently underway to study the prediction capability of other machine learning models and focus on the generalisation of AI-based models, aiming to advance the field of wet gas metering and expand the applicability of these techniques in industrial settings.

8. NOTATION

X_{LM}	Lockhart-Martinelli
$\dot{m}_g$	Gas mass flowrate
m_l	Liquid mass flowrate
GVF	Gas volume fraction
ρ_g	Gas density
ρ_l	Liquid density
DPt	Throat differential pressure
PPL	Pressure loss across Venturi
DNN	Deep Neural Network
CNN	Convolutional neural networks
GRU	Gated Recurrent Unit
$LSTM$	Long Short-Term Memory

9. REFERENCES

[1] S.M. Salehi, L. Lao, L. Xing, N. Simms, W. Drahm, Devices and methods for wet gas flow metering: A comprehensive review, Flow Measurement and Instrumentation (2023) 102518.

[2] Y. Yan, L. Wang, T. Wang, X. Wang, Y. Hu, Q. Duan, Application of soft computing techniques to multiphase flow measurement: A review, Flow Measurement and Instrumentation 60 (2018) 30–43.

[3] M. Mowbray, M. Vallerio, C. Perez-Galvan, D. Zhang, A.D.R. Chanona, F.J. Navarro-Brull, Industrial data science–a review of machine learning applications for chemical and process industries, React Chem Eng 7 (2022) 1471–1509.

[4] G. Chinello, Wet gas flow measurement with Venturi tubes, PhD Thesis (2019).

[5] R. Steven, Wet gas metering, Ph.D. Dissertation, Dept. Mech. Eng. Univ. Strathclyde, Scotland U. K. (2001).

[6] R. Steven, V-cone wet gas metering, in: CITESEER. North Sea Flow Measurement Workshop, Citeseer, 2007: p. 24.

[7] A. Gryzlov, E. Nyfors, L. Jordaan, E. Undheim, N.A. Braaten, S.E. Gregersen, E. Steinsland, Fluid Mechanical Aspects of wet gas metering, in: SPE Russian Oil and Gas Exploration and Production Technical Conference and Exhibition, OnePetro, 2012.

[8] C. Mills, G. Chinello, M. Henry, Flow measurement challenges for carbon capture, Utilisation and storage, Flow Measurement and Instrumentation (2022) 102261.

[9] D.D. Erickson, H. Haghighi, C. Phillips, The Importance of CO2 Composition Specification in the CCUS Chain, in: Offshore Technology Conference, OTC, 2022: p. D021S025R008.

[10] ISO/TR 11583, Measurement of wet gas flow by means of pressure differential devices inserted in circular cross-section conduits, International Standards Organisation (2012).

[11] M. Farsi, H.S. Barjouei, D.A. Wood, H. Ghorbani, N. Mohamadian, S. Davoodi, H.R. Nasriani, M.A. Alvar, Prediction of oil flow rate through orifice flow meters: Optimized machine-learning techniques, Measurement 174 (2021) 108943.

[12] T. Bikmukhametov, J. Jäschke, Oil production monitoring using gradient boosting machine learning algorithm, Ifac-Papersonline 52 (2019) 514–519.

[13] Z. Jiang, H. Wang, Y. Yang, Y. Li, Comparison of machine learning methods for multiphase flowrate prediction, in: 2019 IEEE International Conference on Imaging Systems and Techniques (IST), IEEE, 2019: pp. 1–6.

[14] T. Barbariol, E. Feltresi, G.A. Susto, Machine learning approaches for anomaly detection in multiphase flow meters, IFAC-PapersOnLine 52 (2019) 212–217.

[15] H. Wang, M. Zhang, Y. Yang, Machine learning for multiphase flowrate estimation with time series sensing data, Measurement: Sensors 10 (2020) 100025.

[16] S. Hosseini, G. Chinello, G. Lindsay, S. Smith, D. McGlinchey, Gas-Liquid Multiphase Flow Measurement in Venturi Tube Through Data-Driven Modelling, in: 2024 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), IEEE, 2024: pp. 1–6.

[17] S. Li, B. Bai, Gas-liquid two-phase flow rates measurement using physics-guided deep learning, International Journal of Multiphase Flow 162 (2023) 104421.