Evaluation of different architectures to determine gas fraction in internal two-phase flows with an artificial neural network model across orifice plate

Tiago F. B Camargo
Engenharia Elétrica
CCET-UFAC
Rio Branco, Brasil
tiago.camargo@ufac.br

José H. A. Monteiro
Engenharia Elétrica
CCET-UFAC
Rio Branco, Brasil
humberto.monteiro@ufac.br

Emilio E. Paladino
EMC
CTC-UFSC
Florianópolis, Brazil
emilio.paladino@ufsc.br

***Abstract*—In industry, accurate monitoring of multiphase flow—simultaneous flows of oil, gas, water, and other combinations—is crucial for optimizing production processes. This study investigates the behavior of phase flows, focusing specifically on the relationship between statistical parameters derived from PDF and gas flow. By training models using pressure signals upstream of an orifice plate, the research explores the capability of artificial neural networks (ANNs) to predict the void fraction using minimal parameters. Data was acquired from a 25.4 mm gas-liquid flow circuit over a wide range of flow rates.**

***Index Terms*—two-phase flow, artificial neural network, orifice plate, vapor-liquid mixture measurement, energy production, model generalization**

I. Introduction

The measurement of flow parameters in multiphase systems, particularly the mass flow rate of the transported phases, is both a critical and challenging aspect. In energy transformation systems, for example, inadequate measurement can result in low efficiency or catastrophic failures. In transport processes involving the transfer of ownership, measurement accuracy is crucial, especially in terms of calculating transfer taxes.

One of the most important applications is in the oil and gas industry, where multiphase flows are predominant in extraction, transport, and processing. The streams transported in production systems generally consist of oil, water, gas, and suspended solids.

In the nuclear energy sector, parameters related to coolant flow, including the mass flow rate of the vapor-liquid mixture, are vital for the safety analysis of nuclear power plants. Monitoring properties such as gas fraction and phase flow is important for efficiency control in power generation. A measurement system capable of collecting such data at multiple points within the system is crucial for this purpose but can be complex. An effective approach is to use intelligent systems to measure multiphase flows (e.g., vapor-liquid flows in a steam power system) using minimally invasive techniques and additional parameters, such as wall pressure near singularities.

Multiphase flows exhibit significant complexity, partly due to the various phase arrangements (patterns) that emerge. These patterns depend on the flow rates of the phases and the properties of the constituent fluids. The arrangements of phases in two-phase flows present an opportunity to extract attributes from transient signals that are suitable for training a neural network model for gas fraction measurement.

In multiphase flow, measuring the gas fraction is essential for applying analytical models and serves as a critical parameter for production or system monitoring [1]. The model parameters are trained using inputs derived from a probability density function (PDF) distribution [2], focusing on distribution characteristics such as mean, skewness, kurtosis, and mode. This study aims to determine the most effective activation function and the minimum number of neurons required for an application that detects the gas fraction. Each activation function may demand a varying number of neurons and computational efforts for further processing. These models can initially be integrated in parallel with existing monitoring systems and retrained with subsequent real-world phase flow data, or with more precise data from a gas fraction sensor at the same point in the line, given the complexity of methods for measuring multiphase parameters [1].

II. Methodology

The raw data, collected over 2 minutes using the setup illustrated in Fig. 1, undergoes pre-processing, which includes filtering out noise above 13 Hz and normalizing the data to ensure consistent scaling across all parameters. The data, obtained from a 25.4 mm gas-liquid flow circuit over a wide range of flow rates, is used to train and test the models. The range of flow rates is unspecified. The characteristics extracted from the pressure signal include mean, standard deviation, skewness, kurtosis, and mode. These features are essential for capturing the dynamics of multiphase flow and will be used as inputs for the artificial neural networks (ANNs) in the first case study.

The second set of features includes the values of the PDF function curve at 30 positions, limited by z-scores between -6 and 6. Normalizing the signal is crucial to aligning different flow conditions and allowing for closer comparison of their shapes. The objective of this second set of features is to

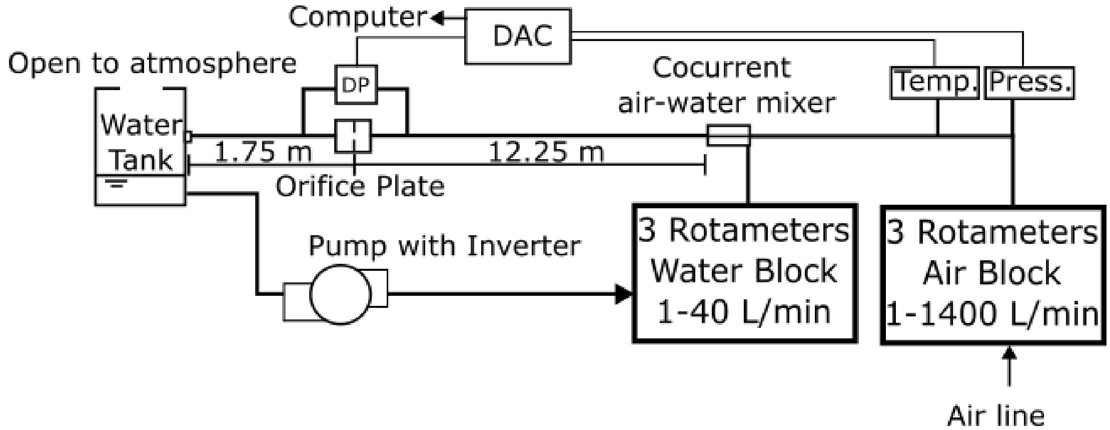

Fig. 1. Experimental circuit with Differential Pressure (DP), Temperature (Temp.), and Absolute Pressure (Press.) Transmitter positions.

gain insights into the model's performance using additional information directly from the PDF, alongside the PDF-derived parameters used in the other set. The analysis focuses on the complexity of the trained model in extrapolating orifice geometry ratios, as well as the model's precision in interpolation (using the same geometry) and extrapolation (using different geometries) scenarios.

This work also studies the impact of the number of neurons used in the architecture, with results evaluated using Mean Absolute Percentage Error (MAPE). To train the ANNs, different neuron configurations were tested. Training was conducted using the Python library TensorFlow, focusing on the objective function errors MAPE and Mean Squared Error (MSE). The minimum number of neurons required to achieve acceptable performance was investigated, starting with a single neuron and gradually increasing the number using a power of two approach, specifically 2^n where n = 1, 2, 3, ..., 10. The data was divided into training, validation, and test sets. In the two scenarios, the features were either whitened or standardized. The data distribution was 40% for training, 10% for validation to assist in the training process, and the remaining 50% for testing. Training was repeated 25 times to calculate the mean and 90% confidence interval for each model with varying neuron counts.

Dimensional reduction of the input variables can be evaluated to achieve a smaller set of input parameters [3], thereby reducing the number of network parameters. The data used in this study was collected from a 25.4 mm gas-liquid flow circuit over a wide range of flow rates. A correlation based on measured phase flow rate data from the literature was used to determine the labels/targets for parameter training [4], [5]. Specifically, the correlation for gas fraction proposed by Lockhart-Martinelli, which was obtained using a 25.4 mm horizontal gas-liquid flow circuit similar to the one used in this study, was employed among other options available in the literature [5]. The study by Lockhart and Martinelli [6] presents a correlation for horizontal flow in pipes with diameters ranging from 1.5 mm to 25.4 mm, expressed as,

$$\alpha = \left[1 + 0.28\left(\frac{1-\mathtt{x}}{\mathtt{x}}\right)^{0.64}\left(\frac{\rho_g}{\rho_l}\right)^{0.36}\left(\frac{\mu_l}{\mu_g}\right)^{0.07}\right]^{-1}. \quad (1)$$

A greater number of neurons has the potential to create a more accurate model by capturing more complex functions in the data. However, if there is a variation in the geometry of the orifice, this increased complexity could lead to poorer model performance. One of the challenges of training complex models is that they often struggle to generalize well to new or slightly altered conditions. To address this, the work proposes evaluating the performance of models by varying the architecture, specifically by adjusting the number of neurons in the hidden layer, to assess how these changes impact the model's ability to generalize.

Two types of changes in the scenario used to develop the model are highlighted: altering the set of features used and varying the number of neurons. Additionally, tests were conducted using two objective functions, MAPE and MSE. Although these functions sometimes yielded similar performance, MAPE consistently demonstrated superior results in all cases. Therefore, the remaining models were trained using MAPE as the sole loss function.

III. Results

In the initial study, the following statistical parameters were used as features: mean, standard deviation, skewness, kurtosis, and mode of the differential pressure signal. This set of features, referred to as Feature Set 1, encompasses five entries. Training was conducted using a dataset with an orifice geometry of 0.5 diameter ratio (β), but the testing procedures were applied to both 0.5 and 0.74 geometry ratios under conditions (pair liquid and gas flow rates) not seen during training. The performance results of the models are shown in Fig. 2.

The bottom plot illustrates the results for interpolation test points, which indicate better performance with 32 neurons in the hidden layer, achieving an error rate of 15.88%. This is a strong prediction performance under conditions not seen during training. A range of neuron counts from 8 to 256 also demonstrated good performance. Notably, the architectures with the two highest neuron counts encountered issues during training, likely due to the difficulty in creating gradients to train with so many parameters and few entries.

In the case of geometry extrapolation performance, the optimal number of neurons using Feature Set 1 was lower than in the interpolation scenario. Beyond this optimal number, the precision decreased. This trend may be attributed to the fact that the global function becomes more complex when using fewer neurons. The results for the 0.74 geometry, combined with the interpolation performance for the 0.5 geometry, indicate that increasing the number of neurons does not improve performance. This suggests that for this feature set, a better and less complex model is achieved with 8 to 32 neurons.

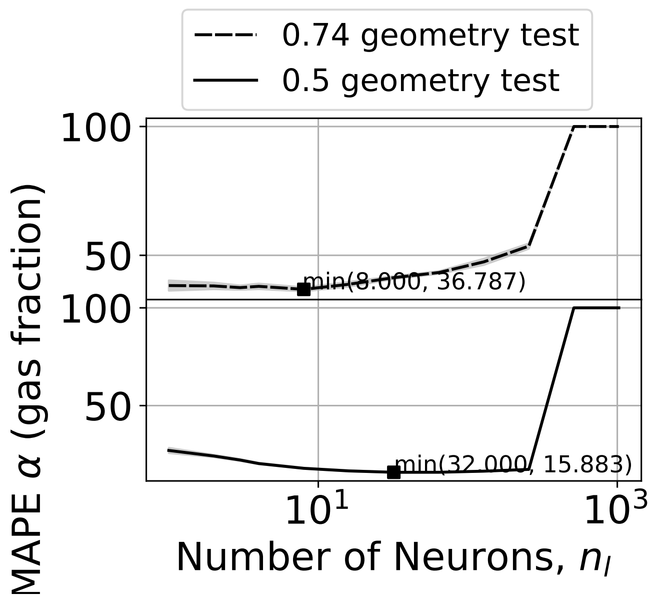

Fig. 2. Variation of Neurons used on ANN architecture using Feature Set 1.

Feature Set 1, when combined with values from the PDF distribution itself—specifically, 30 function points—forms Feature Set 2 with 35 entries. While Feature Set 1 is derived from the PDF, this second scenario increases the amount of information available, but also makes the mapping more complex.

The second set of features was also used to develop ANN models by varying the number of neurons, following the same procedure as the previous scenario. The results of the obtained models are shown in Fig. 3. Testing on the same trained geometry revealed that the results were less accurate compared to those obtained using Feature Set 1. In this case, performance achieved a minimum MAPE value of 30.50% with 16 neurons, and then the MAPE increased slightly as the number of neurons continued to rise, indicating a drop in performance. Furthermore, in the case of extrapolation to a geometry with a diameter ratio of 0.74, the predictions were worse than those obtained using the first set of features. This indicates that adding features directly from the PDF distribution increased the difficulty of extrapolation, possibly due to the increased complexity of the global function. However, in this scenario, all architectures were able to produce plausible models without the training issues encountered with Feature Set 1 when using neuron counts of 512 and 1024.

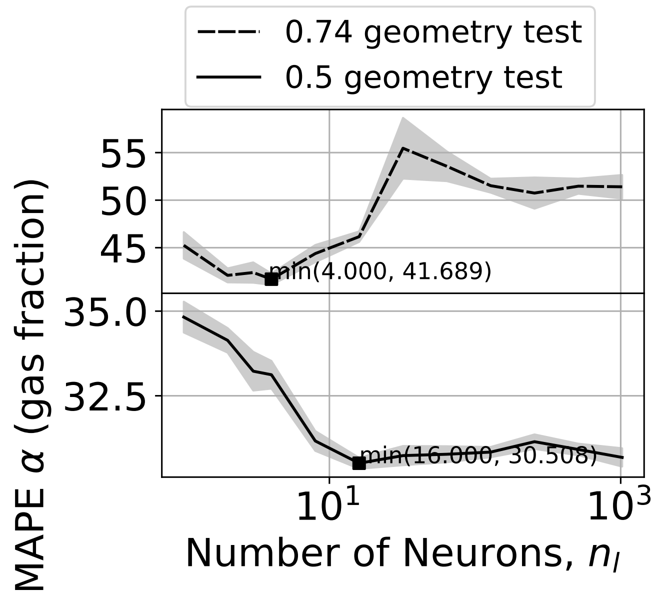

Fig. 3. Variation of Neurons used on ANN architecture.

IV. Conclusion

This work investigated the impact of the number of neurons on the performance of ANN models. The obtained models demonstrated good precision, with a lower bound error of 15.88%. The analysis also examined the complexity of the models in relation to their ability to generalize to different orifice geometries. The results indicate that a lower number of neurons improves the estimation of gas fraction when the model is applied to a different orifice geometry. This may be due to the lower complexity of the model, which results in a function that is more suitable for a wider range of extrapolation. A reflection of this study can be applied orifice plates that can suffer of erosion or corrosion during life time. The findings highlight the importance of carefully selecting input features and neuron counts in ANN models, particularly in applications where the model must generalize across different physical scenarios, such as varying orifice geometries. In practice, this means that simpler models with fewer neurons may be more reliable in scenarios where generalization is critical, such as in the monitoring of multiphase flows in energy generation.

Acknowledgment

No acknowledgment

References

[1] G. FalconE, G. F. Hewitt, and C. Alimonti, *MULTIPHASE FLOW METERING*, 8 2009. [Online]. Available: https://www.sciencedirect.com/bookseries/developments-in-petroleum-science/vol/54/suppl/C

[2] S. Li, F. Zhao, X. Zheng, D. He, and B. Bai, “Wet gas metering by cone throttle device with machine learning,” *Measurement*, vol. 164, p. 108080, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224120306187

[3] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[4] D. Chisholm, “Flow of incompressible two-phase mixtures through sharp-edged orifices,” *Journal of Mechanical Engineering Science*, vol. 9, no. 1, pp. 72–78, 1967.

[5] A. R. Gardenghi, E. d. S. Filho, D. G. Chagas, G. Scagnolatto, R. M. Oliveira, and C. B. Tibiriçá, “Overview of void fraction measurement techniques, databases and correlations for two-phase flow in small diameter channels,” *Fluids*, vol. 5, no. 4, 2020. [Online]. Available: https://www.mdpi.com/2311-5521/5/4/216

[6] R. W. Lockhart and R. C. Martinelli, “Proposed correlation of data for isothermal two-phase, two-component flow in pipes,” *Chemical Engineering Progress*, vol. 45, no. 1, pp. 39–48, 1949.