Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Application of K-means for Identification of Multiphase Flows Based on Computational Fluid Dynamics

Patrick S. Lima[a*], Leonardo S. Souza[b], Leizer Schnitman[b], Idelfonso B. R. Nogueira[a]

[a] Norwegian University of Science and Technology, Department of Chemical Engineering, Trondheim, Trøndelag, Norway
[b] Universidade Federal da Bahia, Salvador, Bahia, Brazil
* Corresponding Author: Patrick.lima@ntnu.no.

ABSTRACT

This study explores multiphase flow dynamics with a focus on the annular flow regime using Computational Fluid Dynamics (CFD) simulations. The methodology included defining the physical model, generating the computational mesh, and analyzing flow patterns. The Volume of Fluid (VOF) model captured fluid interactions, while the k-ω SST turbulence model ensured accurate flow predictions. Simulations examined mixture density behavior and identified optimal configurations. A dataset was generated and analyzed using k-means clustering to classify flow patterns effectively. The results demonstrate the reliability of this approach for improving multiphase flow systems, with applications in oil-water processes.

Keywords: Multiphase Flow, Computational Fluid Dynamics, k-Means Clustering, Flow Pattern Classification

INTRODUCTION

Multiphase flows play an essential role in various industrial sectors, being characterized by the interaction of two or more phases, such as liquid, gas, or solid, moving together.

This phenomenon is present in applications ranging from oil and gas extraction and transportation processes to refrigeration systems and sediment transport in rivers. A detailed understanding of these flows is fundamental for optimizing processes, reducing operational costs, and ensuring safety in industrial operations.

In the oil and gas industry, for instance, multiphase flows are frequently associated with the transportation of mixtures of oil, gas, and water in pipelines. The efficiency of these operations directly depends on understanding the interactions between phases, as well as the ability to predict behaviors such as bubble formation, flow patterns, and pressure gradients [1,2].

Among the various regimes of multiphase flow, core annular flow (CAF) stands out as a highly relevant configuration. One example of this regime can be seen in Figure 1, a central phase, core in red, is surrounded by another phase, annulus in blue, offering significant advantages in applications such as the transportation of high-viscosity fluids. In the oil and gas industry, core-annular flow is widely employed for transporting ultra-viscous oil through pipelines. By surrounding the viscous oil with a layer of less viscous water, this regime reduces wall friction, allowing the fluid to be transported more efficiently and economically [3,4].

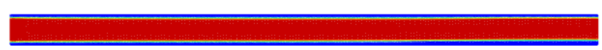

Figure 1. Example of CAF.

The identification of multiphase flow patterns in real systems has been a critical area of research, given its impact on industrial process optimization and safety.

Traditional experimental methods, such as high-speed imaging, electrical capacitance tomography (ECT), and pressure drop analysis, have been widely employed for flow visualization and pattern classification. For instance, Hewitt (1978) provided a foundational approach to visualizing flow regimes in horizontal and vertical pipes, emphasizing the role of flow characteristics such as velocity, phase distribution, and interfacial dynamics in defining patterns [5]. Another notable contribution was made by Soleimani et al. (2006), who demonstrated the effectiveness of ECT in reconstructing real-time images of multiphase flows, offering a non-intrusive method to map flow patterns accurately [6].

While these methods provide valuable insights, they often require sophisticated equipment and are limited by operational constraints in high-pressure or high-temperature environments.

Another approach to address these challenges, advanced tools like Computational Fluid Dynamics (CFD) have emerged as powerful approaches. CFD enables the precise simulation of complex flows, providing detailed insights into critical variables such as flow patterns and pressure gradients. However, the detailed analysis of the data generated by these simulations requires methods for identifying and classifying flow regimes.

This article explores the application of the K-means algorithm to identify flow regimes applied data generated by CFD in the context of core-annular flow systems for transporting ultra-viscous oil.

METHODOLOGY

The methodology was divided into three main stages. In the first stage, the physical model used for the simulation was defined. Subsequently, the process of mesh generation for the system was developed, where the control volumes were defined. These volumes are the elements where the mass and momentum balance equations of the system are solved. Furthermore, in this stage, the models of viscosity, turbulence, and interface governing the interactions between fluids with different properties were established.

In the second stage, individual simulations of different flow patterns were performed, based on the previously constructed geometry and mesh. During these simulations, the behavior of the mixture's density over time was observed, with the objective of understanding and identifying the most suitable approaches for identifying the system.

Finally, in the third stage, 24 simulations were conducted to generate a dataset, which was analyzed using the k-means clustering method. This procedure evaluated the efficiency of the technique in identifying flow patterns, enabling conclusions about its applicability. The resume methodology of this work can be seen in Figure 2.

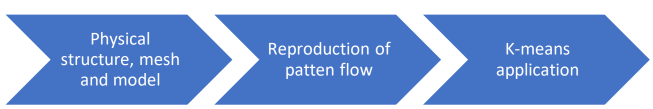

Figure 2. Resume of methodology.

PHYSICAL STRUCTURE MESH AND MODEL

To induce annular flow, it is common to use a nozzle that allows the simultaneous injection of oil and water through separate orifices. An example of such a setup can be seen in Figure 3.

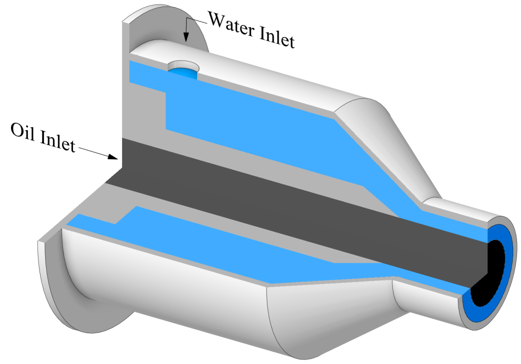

Figure 3. Example of nozzle.

To simplify the system's geometry and make the computational cost of the simulation more reasonable, the structure was adapted to a two-dimensional system, as shown in Figure 4.

Figure 4. Simplification of nozzle inlet.

This geometry features a centralized inlet of 40 mm for oil injection, along with two additional inlets of 5 mm for water injection. Connected to this inlet structure is a straight horizontal tube with a length of 1000 mm.

In general, any simplification that reduces the dimensionality of a problem results in lower computational effort at the cost of potentially minimizing certain three-dimensional effects. However, the two-dimensional simulations conducted in this study proved to be sufficiently representative in describing the classic multiphase flow patterns reported in the literature.

After preparing the geometry, a mesh was created to solve the partial differential equations for mass balance and the Navier-Stokes equations. The mesh consists of around 10000 elements and can be visualized in Figure 5.a and Figure 5.b with zoom.

Figure 5.a. Mesh of system.

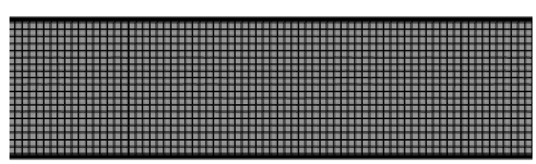

Figure 5.b. Mesh of system with zoom.

The mesh scale was determined based on a heuristic approach, and it was considered reasonable when it successfully represented flow patterns closely resembling those already established in the literature.

Since the primary objective of this work was to explore a strategy for identifying flow patterns rather than performing a numerical validation of simulations against real experiments, the choice of a cost-effective mesh was prioritized.

From a rigorous perspective, creating a dedicated numerical mesh for each type of flow pattern would be necessary, but this was not within the scope of the study. Instead, a balanced approach was adopted to ensure computational efficiency while still achieving meaningful results.

The model that was used to simulate the system was Volume of Fluid (VOF). The VOF model was employed to simulate the interfaces between immiscible fluids in the system. This model is well-suited for resolving complex interfacial dynamics, as it uses a single set of momentum equations for the domain and tracks the volume fraction of each fluid. This approach allows for an accurate representation of the interactions between fluids with different properties and is widely used in multiphase flow simulations [7].

During the simulation, an isothermal system was considered, where the physicochemical properties of the fluids water and oil remained constant.

The water density was assumed to be 998.2 kg/m³, while the oil density was set to 854 kg/m³. The dynamic viscosity was defined as 0.001 Pa·s for water and 0.62 Pa·s for oil, with an interfacial tension of 0.032 N/m between the fluids [8].

Another fundamental part of one CFD simulation is the turbulence model. For this application, we use k-ω SST turbulence model due capability to handle flows with adverse pressure gradients and boundary layer separation. It combines the advantages of the k-ω model near walls with the k-ε model in free-stream regions, providing reliable performance across a wide range of flow regimes [9].

REPRODUCTION OF PATTERN FLOW

In this section, classic flow patterns reported in the literature were simulated [3]. To enhance visualization, oil was represented in red and water in blue, with the transition between colors indicating the interface between the two fluids.

Additionally, the average fluid density at the pipe outlet was monitored over time. It is important to note that the pipe was initially filled entirely with water. As a result, during the early stages of the simulation, the time signal often closely resembled the density of water, as some time was required for the oil to travel from the pipe inlet to its outlet.

The first flow pattern simulated was the oil in water dispersion. This pattern is characterized by small oil droplets dispersed within the water. As shown in Figure 6, the fluid density over time remains close to the density of water, with minor oscillations caused by droplets of varying sizes passing through the pipe outlet.

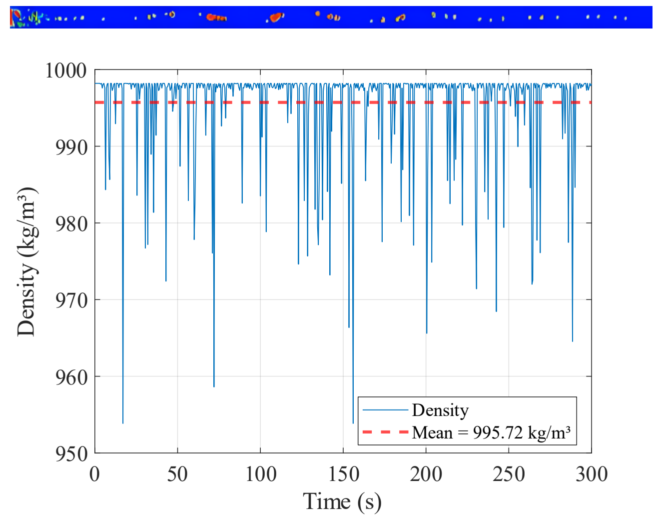

Figure 6. Oil in water dispersion signal in the time.

The second simulated flow pattern was the bubbly flow, where the droplets observed in the oil in water dispersion evolve into larger bubbles. This transition leads to a slightly lower average density compared to the previous pattern as shown in Figure 7.

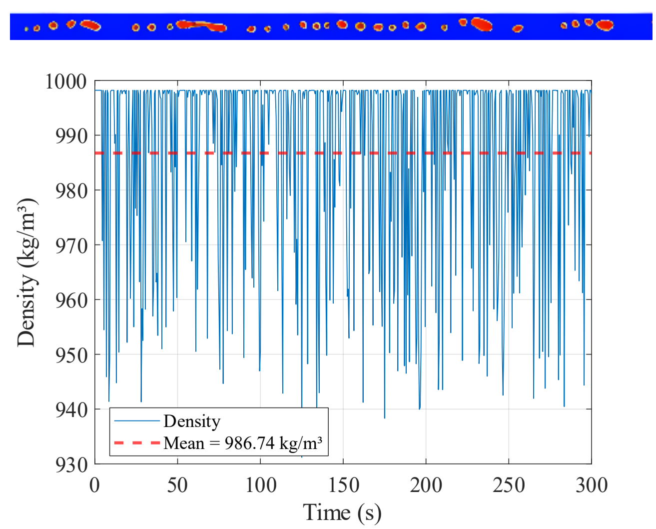

Figure 7. Bubble signal in the time.

The third simulated flow pattern was slug flow, which follows a similar progression to bubbly flow. In this regime, the increased volume of oil promotes the coalescence of bubbles, forming elongated oil slugs. These slugs occupy significant portions of the pipe cross-section, resulting in a further reduction of the average fluid density over time. As depicted in Figure 8, the density signal exhibits larger and more pronounced oscillations, reflecting the alternating passage of oil slugs and water segments at the pipe outlet. This pattern highlights the impact of bubble coalescence on flow dynamics and density distribution.

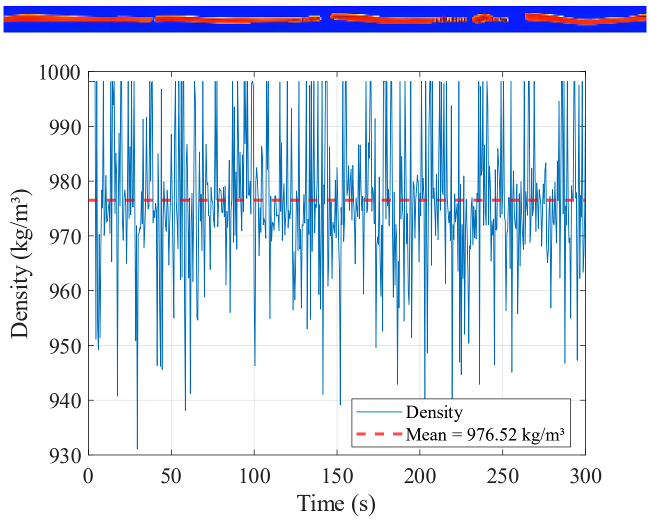

Figure 8. Slug signal in the time.

The fourth simulated flow pattern was stratified flow, characterized by oil and water flowing in distinct layers along the pipe. In this regime, the oil does not achieve sufficient velocity to traverse the pipe effectively. As a result, buoyancy forces become significantly relevant, allowing the oil mass to accumulate at the top of the pipe. This behavior increases the likelihood of oil deposition on the pipe walls, leading to fouling over time. As shown in Figure 9, the density signal stabilizes at a lower average value due to the layered separation of the fluids and the reduced mixing between phases. This pattern emphasizes the importance of managing flow conditions to minimize fouling risks in pipeline systems.

The fifth simulated flow pattern was annular flow, that can be seen in Figure 10, which exhibited the lowest average density among all patterns. This observation aligns with expectations, as annular flow contains the highest volume of oil within its cross-section. A notable characteristic of this pattern is the smooth and consistent behavior of the density signal, which contrasts with the noisier signals observed in other patterns. This is attributed to the well-defined interfaces between oil and water, resulting in reduced fluctuations at the pipe outlet. The stability and distinct separation of phases in annular flow highlight its unique dynamics and emphasize the importance of precise modeling for accurate system analysis.

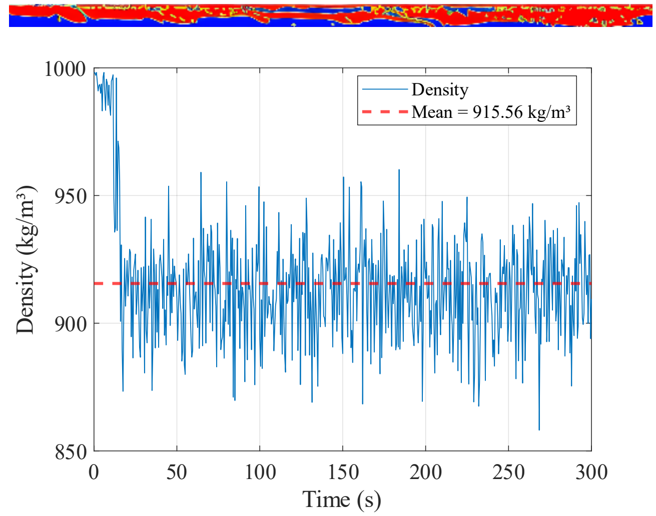

Figure 9. Stratified flow signal in the time.

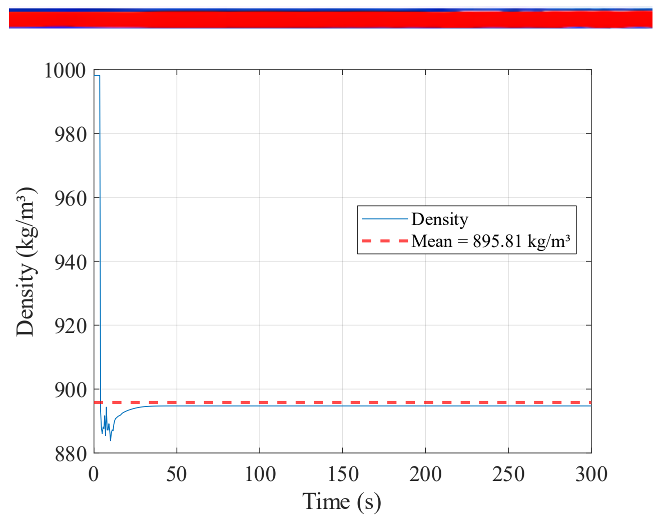

Figure 9. Annular Flow sign al in the time

CLUSTERING AND IDENTIFICATION

To analyze the flow patterns, a one-dimensional k-means clustering was performed, considering the average density of the time signal as the primary variable. To minimize the impact of transient effects in the system, the first 50 seconds of each simulation were excluded from the analysis. This adjustment ensured that the calculated average density more accurately reflected the steady-state behavior of each flow pattern.

A total of 25 simulations were conducted, each running for 300 seconds under varying boundary conditions. This approach allowed for the evaluation of a wide range of scenarios, capturing the diversity in flow dynamics. The use of average density as a clustering metric provided a straightforward yet effective means of distinguishing between patterns, particularly in cases where differences were more pronounced, such as with core-annular flow.

The Figure 10 shows the results of clustering performed using the k-means method with five clusters based on the average density of the time signal for 25 simulations conducted under different boundary conditions. Each point in the graph represents a specific simulation, while the dashed lines indicate the cluster centers identified by the algorithm.

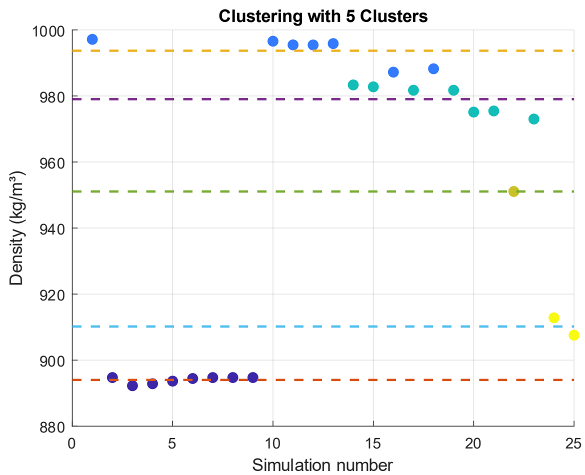

Figure 10. Results of K-means

The clusters are well defined and distributed along the density axis, reflecting the differences in the simulated flow patterns. Most points are closely grouped around their respective cluster centers, suggesting that average density was an effective parameter for distinguishing between the different patterns. Additionally, the low variation within each cluster indicates that this parameter reliably captures the general behavior of the flow patterns.

The cluster with the lowest density, located below 900 kg/m³, corresponds to the annular flow pattern, characterized by the highest volume of oil in the pipe's cross-section. This significantly reduces the average density of the signal and shows a clear separation from the other clusters.

On the other hand, the intermediate clusters, such as those around 920 kg/m³ and 940 kg/m³, represent patterns like slug flow and stratified flow, where there is greater interaction between phases. In these cases, some overlap in average density values can be observed, indicating that the boundaries between these patterns may not be perfectly distinct depending on the boundary conditions.

The cluster with the highest density, near 1000 kg/m³, is associated with patterns such as oil in water dispersion, where the amount of oil in the mixture is smaller, and the behavior is dominated by the characteristics of water. This cluster is well-separated from the others, emphasizing the uniqueness of this pattern in terms of average density.

Overall, the results presented in the image demonstrate that the k-means method was effective in identifying flow patterns based on average density. However, the overlap of values observed between some intermediate clusters highlights the need to explore more complex approaches in future studies. These approaches could include the use of multiple variables or the consideration of temporal relationships to improve the discrimination of patterns with very similar characteristics.

CONCLUSIONS

The strategy of identifying multiphase flow patterns using the k-means clustering method has proven to be promising, as it allowed for a reasonable clustering of results. This approach was particularly effective for core annular flow, which exhibits a highly distinct behavior compared to other patterns due to its low density oscillations over time. However, for other flow patterns, some challenges may arise. Depending on the boundary conditions, there can be an overlap of values between patterns that are very similar, making differentiation more difficult.

Despite these limitations, this study successfully demonstrated a functional method, aligning with its premise of developing a simple approach for flow pattern identification. At the same time, it opens the door for further research into more advanced clustering techniques or methodologies. Future approaches could explore grouping data across multiple dimensions and incorporating temporal relationships, moving beyond reliance on a single variable to achieve more robust and comprehensive results.

REFERENCES

1. Ishii M, Hibiki T. Thermo-fluid dynamics of two-phase flow. Springer Science & Business Media (2011).
2. Brennen CE. Fundamentals of multiphase flow. Cambridge University Press (2005).
3. Joseph DD, Bai R, Chen K, Renardy Y. Core-annular flows. Annu Rev Fluid Mech 29:65-90 (1997). https://doi.org/10.1146/annurev.fl.29.010197.00043
4. Bannwart AC. Modeling aspects of core-annular flows. J Pet Sci Eng 32:127-143 (2001). https://doi.org/10.1016/S0920-4105(01)00155-5
5. Hewitt GF. Flow regimes. Int J Multiphase Flow 4:89-114 (1978). https://doi.org/10.1016/0301-9322(78)90002-3
6. Soleimani M, Caunce J, Peyton AJ. Imaging of two-phase flows using electrical capacitance tomography. Meas Sci Technol 17:1926-1936 (2006). https://doi.org/10.1088/0957-0233/17/8/037

7. Hirt, C. W., & Nichols, B. D. (1981). Volume of fluid (VOF) method for the dynamics of free boundaries. Journal of Computational Physics, 39(1), 201–225.
8. Lima, P., Nogueira, I., Schnitman, L., Mendes, T. P. G., Costa, E., & Skogestad, S. (2023). Harnessing Feedback Control Strategy for Improved Core Annular Flow Stability in Heavy Oil Transportation. AIChE Annual Meeting, Computing and Systems Technology Division. Orlando, FL.
9. Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. AIAA Journal, 32(8), 1598–1605.