

Research Article

Enhancing Fractional Flow Curve Modeling with Advanced Data-driven Techniques: A Comparative Evaluation of Machine Learning Frameworks

Caleb John^{*} , Oluwatoyin Akinsete , Abosede A. Fadayomi ,
Samuel B. Aderemi 

Department of Petroleum Engineering, University of Ibadan, Ibadan, Nigeria

Abstract

Modeling fractional flow curves accurately is essential for optimizing reservoir performance and improving hydrocarbon recovery. This study introduces a robust analytical framework utilizing advanced computational techniques to predict fractional flow behavior. The model leverages Gradient Boosted Decision Trees (GBDT) and integrates key physical parameters such as water saturation, viscosity ratios, and relative permeability. The performance of the proposed framework was evaluated using data from reservoir simulations and experiments. The model demonstrated high predictive accuracy, achieving a Root Mean Square Error (RMSE) of 0.005, a Coefficient of Determination (R^2) of 0.99, and a Mean Absolute Percentage Error (MAPE) of 1%. Compared to conventional fractional flow models based on Buckley-Leverett theory, which yielded an RMSE of 0.16 and a MAPE of 12.8%, the new approach showed significant improvement. Additionally, it outperformed other computational approaches, including Random Forest (RMSE: 0.02, MAPE: 10.4%) and Artificial Neural Networks (RMSE: 0.016, MAPE: 6.0%), providing both enhanced accuracy and consistency. A sensitivity analysis confirmed the robustness of the model across a range of viscosity ratios, showing strong alignment with physical principles, such as shock front behavior and saturation constraints. The practical utility of this model lies in its ability to accurately predict fractional flow under varying conditions, bridging gaps between analytical methods and data-driven techniques, while remaining computationally efficient. This development enhances the tools available for reservoir engineers, offering new insights for waterflooding strategies, enhanced oil recovery (EOR), and other multi-phase flow applications, with direct relevance to field operations.

Keywords

Buckley-Leverett Theory, Enhanced Oil Recovery, Fractional Flow Model, Machine Learning, Reservoir Engineering

1. Introduction

Fractional flow curve description is the essence of multi-phase fluid displacement in porous media, directly influencing reservoir management practices like waterflooding and enhanced oil recovery (EOR). [19] Fractional flow curves, based

on Buckley–Leverett theory, describe fluid front behavior, breakthrough times, and sweep efficiency. Analytical solutions, however, are usually not able to describe the intricacies of actual reservoir conditions like heterogeneous permeability,

^{*}Correspondence: Caleb John (calebjohn628@gmail.com)

Received: 1 October 2025; **Accepted:** 14 October 2025; **Published:** 2 February 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

capillary pressure effects, and location-dependent mobility ratios.

Over the past decade, data-driven methods, i.e., machine learning (ML) models, have proven to be strong tools for enhancing the accuracy of prediction in complicated reservoir systems. By combining vast amounts of data and sophisticated algorithms, such models are capable of identifying patterns and nonlinear correlations that classical approaches may not recognize. Of these methodologies, Gradient Boosted Decision Trees (GBDT) have proven to be popular options because they are capable of dealing with high-dimensional data, identifying intricate interactions, and producing solid predictions.

The focus of this work is to recommend an analytical framework with GBDT to obtain high-precision fractional flow curves. The devised model utilizes standout reservoir properties such as water saturation, viscosity ratio, and relative permeability as an attempt at being physically compliant. Furthermore, the model's performance is compared with conventional models and other ML-based methods of Random Forest and ANNs as well for making its predictability superiority as well as usability credentials.

By obtaining smaller error measures and signs of stability with sensitivity analyses, this research contributes to data-driven methods in reservoir engineering. The results of this research can be used to give smart suggestions for optimizing waterflooding operations, enhancing EOR efficiency, and facilitating data-driven decision-making in complicated reservoir operations.

1.1. Theory of Study

Fractional flow theory is a fundamental concept in multiphase flow dynamics, particularly in the context of waterflooding and enhanced oil recovery (EOR) processes. [11] It describes the relationship between water saturation and the proportion of water in the total fluid flowing through a porous medium. [15] The concept was first introduced by Buckley and Leverett [1], who developed a mathematical model for one-dimensional, immiscible displacement in porous media. Their work laid the foundation for understanding fluid front behavior during displacement processes.

The fractional flow of water (f_w) is defined as the ratio of the water flow rate to the total flow rate:

$$f_w = \frac{q_w}{q_w + q_o} = \frac{1}{1 + \frac{\kappa_{ro}\mu_w}{\kappa_{rw}\mu_o}} \quad (1)$$

where:

f_w = fractional flow of water

q_w = water flow rate (m³/s)

q_o = oil flow rate (m³/s)

κ_{ro}, κ_{rw} = relative permeabilities of oil and water (unitless)

μ_o, μ_w = viscosities of oil and water (cP)

The equation highlights the impact of relative permeabilities and fluid viscosities on the fractional flow behavior im-

part of relative permeabilities and fluid viscosities on the fractional flow behavior. [7] According to Dake, relative permeability functions are critical in determining how effectively water displaces oil in the reservoir. [2].

1.2. Buckley-Leverett Theory and Waterflooding Front

The Buckley–Leverett theory describes the propagation of a saturation front in a linear reservoir during waterflooding. The solution to the Buckley–Leverett equation yields the fractional flow curve and the saturation profile, which are essential for understanding breakthrough behavior and sweep efficiency essential for understanding breakthrough behavior and sweep efficiency. [4] The water saturation profile is obtained from the fractional flow derivative, known as the fractional flow gradient:

$$\frac{df_w}{ds_w} = \frac{\frac{\kappa_{rw}}{\mu_w}}{\left(\frac{\kappa_{rw}}{\mu_w} + \frac{\kappa_{ro}}{\mu_o}\right)^2} \quad (2)$$

where s_w is the water saturation to support pressure-fractional flow relationship. [15].

Welge introduced a graphical technique to determine water breakthrough and recoverable oil based on the fractional flow curve [3]. This method involves plotting the tangent to the fractional flow curve to estimate the average water saturation at breakthrough.

2. Literature Review

Recent advancements in fractional flow theory modeling have focused on addressing traditional limitations through data-driven and physics-informed approaches enhancing the accuracy of prediction in complicated reservoir systems. [12] Buckley and Leverett provided foundational insights into fluid displacement in porous media. However, their assumptions of homogeneous media and linear flow limit applicability in real reservoirs [1]. Fadairo et al. introduced a novel analytical model, integrating production data to enhance fractional flow predictions [5], integrating production data to enhance fractional flow predictions. [9]. Additionally, Zhao et al. employed physics-informed neural networks (PINNs) to improve accuracy in multiphase flow simulations with sparse data. [20, 21].

Recent research on reduced-order modeling (ROM) for reservoir simulation has emphasized the development of surrogate models to address computational challenges. [22] Traditional models often lack integration with physical knowledge, limiting their generalization capabilities. Nagao et al. proposed a physics-informed spatial-temporal neural network (PI-STNN) that incorporates flow theory into its loss function and combines a deep convolutional encoder-decoder (DCED) with a convolutional long short-term memory (ConvLSTM) network [6]. This model demonstrated superior accuracy and robustness compared to purely data-driven approaches and the

Fourier neural operator (FNO), particularly in heterogeneous and fractured reservoirs. The PI-STNN also enabled efficient uncertainty quantification, supporting faster and more informed decision-making in oil and gas development. [24] This work highlights the potential of integrating physical principles with machine learning to enhance reservoir simulation. [13].

Zhou et al. reviewed machine learning applications in reservoir engineering, emphasizing XGBoost for key parameter prediction and flooding optimization [20]. XGBoost's ability to handle missing data and control overfitting made it highly effective in CO₂ flooding design, achieving superior performance over traditional models [20]. Gao et al. combined XGBoost with Particle Swarm Optimization (PSO) for CO₂ flooding parameter design, demonstrating its advantage in predicting minimum miscibility pressure and optimizing injection schemes [28]. These studies highlight XGBoost's efficiency and adaptability in petroleum engineering.

Researchers have gone into advanced fractional flow modeling using data-driven and physics-informed approaches. [27] These advancements reflect a shift toward hybrid models that integrate physical knowledge with machine learning, offering more accurate and interpretable solutions for reservoir engineering challenges. [10].

3. Methodology

The aim of this study is to develop a robust data-driven approach for predicting fractional flow curves, utilizing machine

learning models, including GBDT (XGBoost), Random Forest, and ANN. [13] A systematic workflow was implemented, covering data preprocessing, model training, and performance evaluation. This section provides an overview of the machine learning techniques employed, highlighting their mechanisms and contributions to the results analyzed later.

3.1. Acquisition and Description of Dataset

The dataset for this study was gotten from published sources, covering key reservoir engineering parameters relevant to fractional flow modeling. [23] Experimental data for water saturation, relative permeability, viscosity ratio and fractional flow were obtained based on the findings of Ren et al [8]. The results from Ren et al were simulated to create a 3569 data points with 6 quantitative variables, setting up a supervised regression task to enhance fractional flow curve modeling. Input features include water saturation, viscosity ratio, relative permeabilities for water and oil, permeability ratio. The target variable, fractional flow, is modeled as a function of these input parameters to capture varying wettability effects, capillary-driven flow, and mobility transitions. The dataset's structure makes it well-suited for training Gradient Boosted Decision Trees (GBDT) to improve fractional flow predictions.

3.2. Exploratory Data Analysis

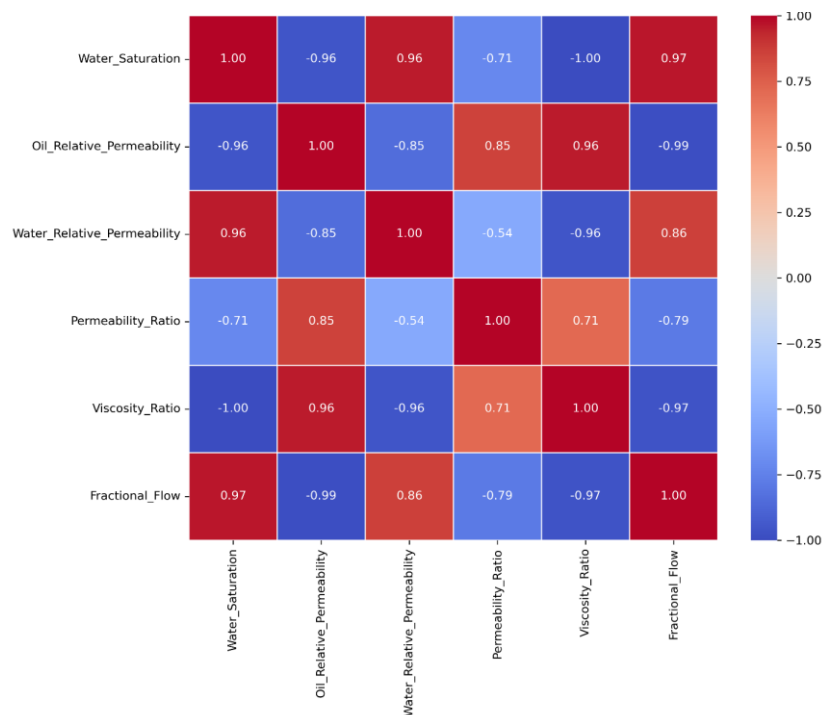


Figure 1. Heatmap illustrating correlation of features.

The heatmap (Figure 1) displays the correlation between key parameters affecting fractional flow in two-phase flow within porous media. A strong positive correlation (close to +1) is observed between water saturation and water relative permeability (0.96), as well as oil relative permeability and viscosity ratio (0.96), indicating that as water saturation increases, water mobility also increases, while oil mobility decreases. Conversely, there is a strong negative correlation between water saturation and oil relative permeability (-0.96) and between viscosity ratio and water relative permeability (-0.96), suggesting that increased water saturation reduces oil flow capacity. Fractional flow is highly correlated with water saturation (0.97) and negatively correlated with oil relative permeability (-0.99), which aligns with Buckley-Leverett theory predictions. The permeability ratio shows moderate influence on other variables, indicating its role in controlling flow behavior but not being the primary driver. The heatmap confirms expected fluid flow relationships in immiscible displacement processes.

Table 1 provides statistical insights into key parameters influencing fractional flow in two-phase flow systems. Water saturation has a mean of 0.50 with a standard deviation of 0.17, ranging from 0.2 to 0.8, indicating moderate variability. The viscosity ratio varies significantly, with a mean of 3.0 and a high standard deviation of 1.8, suggesting diverse fluid viscosity contrasts. Relative permeability to water and oil both have identical means of 0.17 but exhibit large variations, with maximum values of 0.51 and minimum values as low as 0.008, reflecting strong dependence on saturation. The permeability ratio has a notably high maximum value of 64.1 but a small standard deviation (0.016), implying that most data points cluster near the mean (6.72). Lastly, fractional flow averages 0.65 with a low standard deviation of 0.06, ranging from 0.31 to 1.00, indicating that flow conditions are generally stable. This substantial variation emphasizes the complexity of multiphase flow behavior, making the dataset well-suited for advanced fractional flow curve modeling.

Table 1. Descriptive summary of dataset.

	Water Saturation	Viscosity Ratio	Relative Permeability- Water	Relative Permeability- Oil	Permeability Ratio	Fractional flow
Mean	0.50	3.0	0.17	0.17	6.72	0.65
Max	0.8	4.2	0.51	0.51	64.1	1.00
Min	0.2	0.69	0.008	0.008	12.46	0.31
Std	0.17	1.8	0.15	0.15	0.016	0.06

3.2.1. Data Processing

Outliers were identified based on extreme values in water saturation, relative permeabilities for water and oil, and fractional flow. These outliers were handled using a combination of capping at the 99th percentile and log transformation for highly skewed values. The dataset contained a total of 3569 data points, which were split into 80% for training and 20% for testing to ensure a balanced evaluation. Additionally, missing values were imputed using mean or median values where necessary, and data normalization was applied to scale the features between 0 and 1. These pre-processing steps ensured a clean and well-structured dataset for accurate fractional flow modeling using machine learning techniques. These pre-processing steps ensured a clean and well-structured dataset for accurate fractional flow modeling using machine learning techniques.

3.2.2. Model Development

In fractional flow curve modeling, machine learning models such as Gradient Boosting Decision Trees (GBDT), Random

Forest (RF), and Artificial Neural Networks (ANN) are increasingly utilized due to their ability to handle complex, non-linear relationships in reservoir data. GBDT, an ensemble learning technique, builds models iteratively by correcting errors from prior iterations, making it highly effective for capturing intricate patterns in fluid displacement data. Its robustness against overfitting and capability to work with limited datasets make it a valuable tool for predicting fractional flow curves and analyzing the impact of reservoir parameters on production performance [16].

Random Forest (RF), another ensemble-based approach, constructs multiple decision trees and aggregates their outputs to improve prediction accuracy and stability. RF is particularly well-suited for fractional flow modeling as it can handle high-dimensional data and identify the relative importance of various reservoir features influencing fluid saturation and recovery. This model's versatility allows researchers to assess the interplay of multiple factors affecting fluid flow while minimizing the risk of overfitting [17]. On the other hand, Artificial Neural Networks (ANN) excel at capturing nonlinear relationships by mimicking the structure of biological neural

networks. ANN is particularly effective in predicting fractional flow behavior under varying conditions due to its ability to learn from historical data and generalize for future scenarios [25], making it indispensable for optimizing reservoir management strategies [18, 26].

The integration of GBDT, RF, and ANN into fractional flow curve modeling enhances the accuracy and reliability of predictions, enabling better decision-making in reservoir engineering. These models provide complementary strengths: GBDT excels at iterative learning from limited datasets, RF offers interpretability and stability in high-dimensional spaces, and ANN provides unmatched flexibility in capturing nonlinear trends. Together, they form a robust framework for developing representative fractional flow curves that account for complex reservoir dynamics, ultimately improving oil recovery predictions and informing effective reservoir management practices.

3.2.3. Model Evaluation Metrics

In our regression analysis, we employed three key evaluation metrics to assess the performance of our supervised models: *Root Mean Square Error (RMSE)*, *Coefficient of Determination (R^2)*, and *Mean Absolute Percentage Error (MAPE)*. Each metric offers distinct insights into model accuracy and predictive reliability.

(i). Root Mean Square Error (RMSE)

RMSE quantifies the standard deviation of prediction errors, measuring how much predicted values deviate from actual values. It is mathematically defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

This metric is particularly useful in rock-fluid applications, where large errors must be minimized due to their significant impact on results. RMSE is sensitive to large errors, making it an essential criterion for evaluating model robustness.

(ii). Coefficient of Determination (R^2)

The R^2 metric, or coefficient of determination, indicates the proportion of variance in the dependent variable that can be explained by the independent variables. It serves as a measure of the model's goodness of fit, ranging from 0 to 1, where 1 signifies a perfect fit. It is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

A higher R^2 value signifies a better-fitting model, enhancing confidence in its predictive capability.

(iii). Mean Absolute Percentage Error (MAPE)

MAPE evaluates prediction accuracy by measuring the average absolute percentage error between actual and predicted values. It is a scale-independent metric, making it particularly useful for comparing models across different datasets. The formula is:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (5)$$

Since MAPE expresses errors as a percentage, it provides an intuitive understanding of model performance relative to the scale of the data.

4. Result and Discussion

4.1. Traditional Model Performance

The result of the traditional Buckley-Leverett equation showed a coefficient of correlation value of 0.73, mean average percentage error of 12.8% and root mean square error value of 0.16 on the experimental data. Figure 2 shows an actual vs predicted value plot to demonstrate the result.

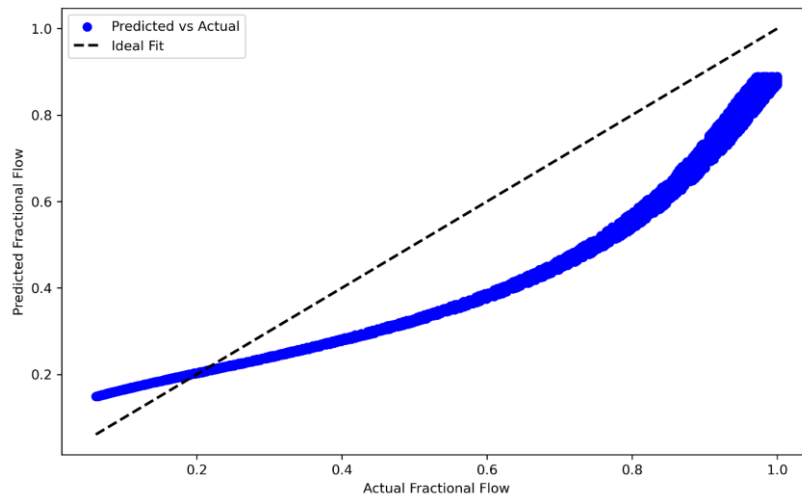


Figure 2. Predicted Vs Actual Plot for Buckley-Leverett Model.

Advanced Models Results and Performance Assessment

Table 2 displays the cross-validation results for all models used in this study, along with their performance on the train-test split. The performance of the three machine learning models—XGBoost, Random Forests, and Artificial Neural Networks (ANN)—was evaluated using R^2 , RMSE, and MAPE across the training, 5-fold cross-validation, and test sets. The model performance evaluation confirms that XGBoost delivers superior predictive accuracy and generalization capability compared to Random Forest and Artificial Neural Network (ANN) models. XGBoost achieves the highest R^2 scores (0.99

for training, 0.99 for cross-validation, and 0.98 for testing), indicating an excellent fit with minimal errors. Its RMSE and MAPE values remain consistently low across all datasets, highlighting its robustness. While Random Forests and ANN also perform well, Random Forests exhibit slightly lower R^2 values (0.96 for training and testing, 0.95 for cross-validation) and higher RMSE and MAPE, suggesting sensitivity to training data. ANN shows reasonable performance, with an improved cross-validation R^2 of 0.97, but slightly higher errors than XGBoost. XGBoost emerges as the most reliable and accurate model for predicting fractional flow.

Table 2. Model Evaluation Result.

Model	Train set			5-fold cross validation			Test set		
	R^2	RMSE	MAPE	R^2	RMSE	MAPE	R^2	RMSE	MAPE
XGboost	0.99	0.002	0.005	0.99	0.005	0.010	0.98	0.005	0.010
Random Forests	0.96	0.019	0.097	0.95	0.02	0.104	0.96	0.02	0.100
Artificial Neural Network	0.95	0.012	0.04	0.97	0.016	0.06	0.96	0.011	0.03

The comparative analysis (Figure 3) of predictive models for fractional flow reveals that XGBoost outperforms both Random Forest and Artificial Neural Networks (ANN) in accuracy and consistency. The XGBoost predictions exhibit a near-perfect alignment with the actual fractional flow values, demonstrating minimal deviation and high precision. In contrast, the Random Forest model shows a staircase-like pattern, indicating overfitting to discrete training data points, which

may limit its generalization ability. The ANN predictions, while closely following the actual values, display slightly more scatter compared to XGBoost, suggesting minor inconsistencies in capturing complex nonlinear relationships. XGBoost proves to be the most robust and reliable model for predicting fractional flow with high accuracy and minimal error.

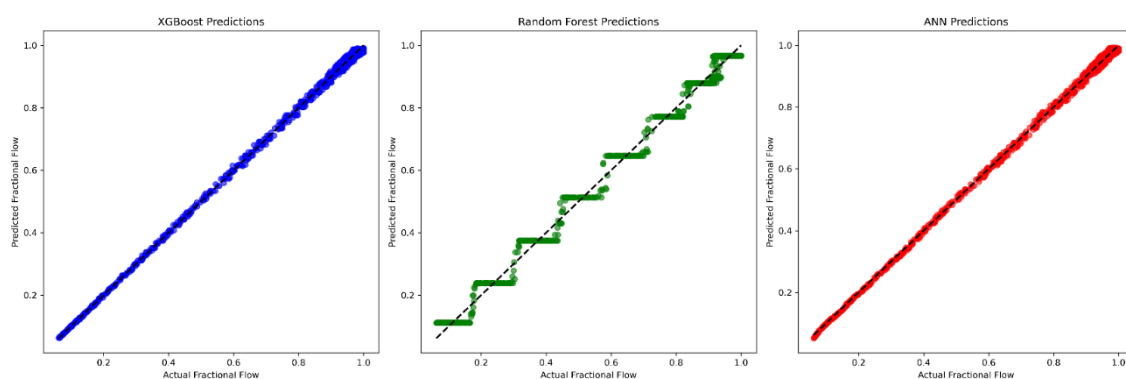


Figure 3. Predicted Vs Actual Fractional Flow Values.

4.2. Sensitivity Analysis

The XGBoost model demonstrates strong robustness in predicting fractional flow under varying reservoir conditions, as

evidenced by the sensitivity analysis charts. The model exhibits smooth and expected trends in response to perturbations in oil relative permeability, water relative permeability, and viscosity ratio, indicating that it accurately captures reservoir

flow behavior. The logical sensitivity responses—such as decreasing fractional flow with increasing oil relative permeability and increasing fractional flow with higher water relative permeability—align with fundamental reservoir engineering principles. [14] Additionally, the model remains stable around the baseline, with no drastic deviations, showing resilience to minor changes in input parameters and mitigating risks of

overfitting. The controlled variability in predictions further reinforces the model's reliability, ensuring that it provides realistic outputs across different perturbation levels. These characteristics highlight the XGBoost model's capability to generalize well and maintain predictive accuracy in fractional flow modeling.

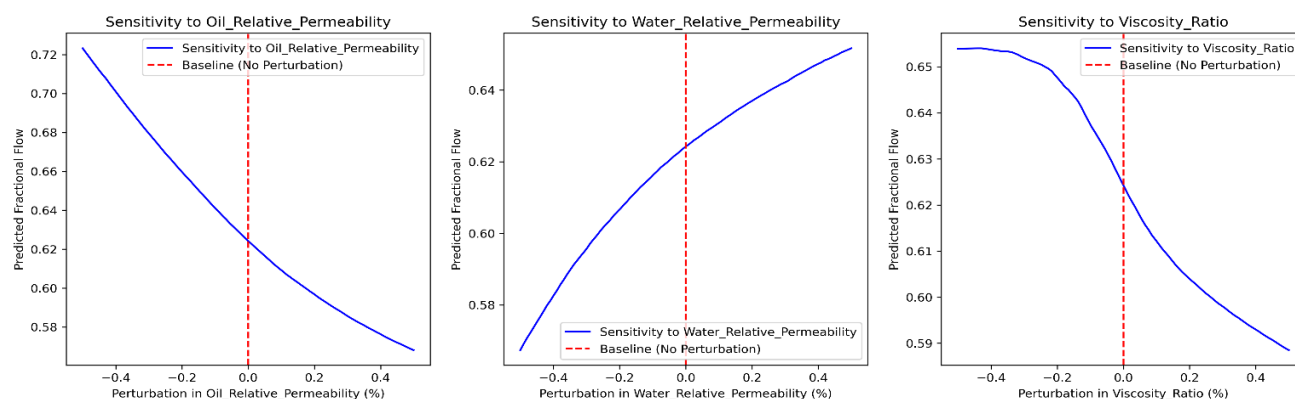


Figure 4. Sensitivity analysis plot.

5. Conclusion and Recommendation

The result of this study underscores the transformative potential of data-driven techniques, particularly Gradient Boosted Decision Trees (GBDT), in advancing fractional flow modeling for reservoir engineering applications. By integrating key reservoir parameters such as water saturation, viscosity ratio, and relative permeability, the proposed GBDT-based framework demonstrated superior predictive accuracy compared to traditional Buckley-Leverett models and other machine learning approaches, including Random Forest and Artificial Neural Networks (ANN). The model achieved a Root Mean Square Error (RMSE) of 0.005, a Coefficient of Determination (R^2) of 0.99, and a Mean Absolute Percentage Error (MAPE) of 1%, significantly outperforming the traditional method and other computational approaches. By leveraging the strengths of GBDT, this work contributes to the growing body of research on physics-informed machine learning in reservoir engineering, paving the way for more accurate and efficient reservoir management practices.

Conclusively, to facilitate widespread adoption and enhance practical utility, user-friendly software tools or platforms should be developed, allowing reservoir engineers to easily implement the GBDT-based fractional flow model, complete with visualization features for fractional flow curves and sensitivity analysis. Given its superior performance in handling high-dimensional data, capturing nonlinear relationships, and providing accurate predictions, Gradient Boosted Decision Trees (GBDT) should be adopted as a standard tool for fractional flow curve modeling in reservoir engineering.

Additionally, the model's computational efficiency makes it a strong candidate for real-time reservoir monitoring and management, and future studies should focus on integrating real-time production data to enable dynamic updates of fractional flow predictions, thereby facilitating proactive decision-making and optimizing reservoir operations.

Abbreviations

ANN	Artificial Neural Network
cP	Centipoise (Unit of Viscosity)
DCED	Deep Convolutional Encoder-Decoder
EOR	Enhanced Oil Recovery
FNO	Fourier Neural Operator
GBDT	Gradient Boosted Decision Trees
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
PI-STNN	Physics-Informed Spatial-Temporal Neural Network
PINN	Physics-Informed Neural Network
PSO	Particle Swarm Optimization
R^2	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Square Error
ROM	Reduced-Order Modeling
XGBoost	Extreme Gradient Boosting

Acknowledgments

I will like to acknowledge the invaluable contribution of

Prince Ikabi, Alex Ifenaike, and Mubarak Olasunkanmi to this work as well as Society of Petroleum Engineers (SPE) Nigeria Council for the opportunity to write this paper.

Author Contributions

Caleb John: Conceptualization, Data curation, Methodology, Visualization, Writing – original draft.

Oluwatoyin Akinsete: Writing – review & editing, validation.

Samuel B. Aderemi: Funding acquisition.

Conflicts of Interest

The authors declare no conflicts of interests.

References

- [1] Buckley, S. E., and Leverett, M. C. (1942). Mechanism of fluid displacement in sands. *Transactions of the AIME*, 146(1), 107–116. <https://doi.org/10.2118/942107-G>
- [2] Dake, L. P. (1983). *Fundamentals of reservoir engineering*. Elsevier.
- [3] Welge, H. J. (1952). A simplified method for computing oil recovery by gas or water drive. *Journal of Petroleum Technology*, 4(4), 91–98. <https://doi.org/10.2118/124-G>
- [4] Tarek, A. (2010). *Reservoir engineering handbook* (4th ed.). Gulf Professional Publishing.
- [5] Fadairo, A., Adeyemi, A., & Falode, O. (2014). A new analytical model for developing fractional flow curve using production data. *Journal of Chemical and Petroleum Engineering*, 55(1), 45–56. <https://doi.org/10.22059/JCHPE.2014.7559>
- [6] Nagao, M., Datta-Gupta, A., Onishi, T., & Sankaran, S. (2024). Physics-informed machine learning for reservoir connectivity identification and robust production forecasting. *SPE Journal*, 29(9), 4527–4541. <https://doi.org/10.2118/219773-PA>
- [7] Kantzas, A., Bryan, J., & Taheri, S. (2018). *Fundamentals of fluid flow in porous media*. Perm Inc. Retrieved from <https://perminc.com/resources/fundamentals-of-fluid-flow-in-porous-media/>
- [8] Ren, X., Li, A., Fu, S., & Wang, S. (2018). Experimental study on the oil-water relative permeability relationship for tight sandstone considering the nonlinear seepage characteristics. *Journal of Petroleum Science and Engineering*, 163, 655–664. <https://doi.org/10.1016/j.petrol.2017.12.003>
- [9] Sitorus, J., Sofyan, A., & Abdulfatah, M. Y. (2006). Developing a fractional flow curve from historic production to predict performance of new horizontal wells—Bekasap Field, Indonesia. *Society of Petroleum Engineers*. <https://doi.org/10.2118/101144-MS>
- [10] Cao, F., Luo, H., & Lake, L. (2015). Oil rate forecast by inferring fractional flow models from field data. *Society of Petroleum Engineers*. <https://doi.org/10.2118/173315-MS>
- [11] Nagoo, A. S. (2018). Insights on pipe fractional flow theory and associated relationships. Retrieved from <https://media.licdn.com/dms/document/media/v2/D561FA-QGvEGyzCdFNFW/feedshare-document-pdf-analyzed/feedshare-document-pdf-analyzed/0/1721826593406?e=1741219200&v=beta&t=SXxGSsiSpUaK1eNi0sBzEuPUJTFuADIBcLXCXXEITKw>
- [12] Sankaran, S., Matringe, S., Sidahmed, M., Saputelli, L., Wen, X.-H., Popa, A., & Dursun, S. (2020). *Data analytics in reservoir engineering*. Society of Petroleum Engineers. <https://doi.org/10.2118/9781613998205>
- [13] Ramalho, L. A. M., Tueros, J. A. R., & Horowitz, B. (2023). Machine learning applied to evaluation of reservoir connectivity. *Neural Computing and Applications*, 36(2). <https://doi.org/10.1007/s00521-023-09056-0>
- [14] Gjennestad, M. A., Winkler, M., & Hansen, A. (2019). Pore network modeling of the effects of viscosity ratio and pressure gradient on steady-state incompressible two-phase flow in porous media. <https://doi.org/10.1007/s11242-020-01395-z>
- [15] Knudsen, H. A., & Hansen, A. (2001). A new relation between pressure and fractional flow in two-phase flow in porous media. <https://doi.org/10.1103/PhysRevE.65.056310>
- [16] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. <https://doi.org/10.1145/2939672.2939785>
- [17] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [18] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- [19] Bassir, S. M., Zargar, G., Moghadasi, J., & Roghanian, R. (2016). A comprehensive review of low salinity water-flooding in sandstone/carbonate reservoirs: From theory to practice. *The Fourth International Conference on Oil, Gas, Refining & Petrochemical*. Retrieved from https://www.researchgate.net/publication/329058576_A_Comprehensive_Review_of_Low_Salinity_Waterflooding_in_SandstoneCarbonate_Reservoirs_From_Theory_to_Practice
- [20] Zhou, W., Liu, C., Liu, Y., Zhang, Z., Chen, P., & Jiang, L. (2024). Machine learning in reservoir engineering: A review. *Processes*, 12(6), 1219. <https://doi.org/10.3390/pr12061219>
- [21] Zhang, J., Braga-Neto, U. M., & Gildin, E. (2024). Physics-informed neural networks for multiphase flow in porous media considering dual shocks and interphase solubility. *Energy & Fuels*, 38(18), 17781–17795. <https://doi.org/10.1021/acs.energyfuels.4c02888>
- [22] Cao, F., Luo, H., & Lake, L. W. (2015). Oil-rate forecast by inferring fractional-flow models from field data with Koval method combined with the capacitance resistance model. *SPE Reservoir Evaluation & Engineering*. Society of Petroleum Engineers. <https://doi.org/10.2118/173315-PA>

- [23] Ertekin, T., Abou-Kassem, J. H., & King, G. R. (2001). *Basic applied reservoir simulation* (Vol. 7). SPE Textbook Series. <https://doi.org/10.2118/9781555630898>
- [24] Yousef, A. A., Gentil, P. H., & Jensen, J. L. (2005, October 9–12). A capacitance model to infer interwell connectivity from production and injection rate fluctuations. *SPE Annual Technical Conference and Exhibition*, Dallas, Texas. Society of Petroleum Engineers. <https://doi.org/10.2118/95322-PA>
- [25] Temirchev, P., Simonov, M., Kostoev, R., Burnaev, E., Oseledets, I., Akhmetov, A., Margarit, A., Sitnikov, A., & Koroteev, D. (2019). Deep neural networks predicting oil movement in a development unit. *arXiv*. <https://doi.org/10.48550/arXiv.1901.02549>
- [26] Ertekin, T., & Sun, Q. (2019). Artificial intelligence applications in reservoir engineering: A status check. *Energies*, 12(15), 2897. <https://doi.org/10.3390/en12152897>
- [27] Zhao, C., Zhang, F., Lou, W., Wang, X., & Yang, J. (2024). A comprehensive review of advances in physics-informed neural networks and their applications in complex fluid dynamics. *Physics of Fluids*, 36(101301). <https://doi.org/10.1063/5.0226562>
- [28] Gao, M., Liu, Z., Qian, S., Liu, W., Li, W., Yin, H., & Cao, J. (2023). Machine-learning-based approach to optimize CO₂-WAG flooding in low permeability oil reservoirs. *Energies*, 16(17), 6149. <https://doi.org/10.3390/en16176149>

Biography



Caleb John is a final year undergraduate student at University of Ibadan, Department of Petroleum Engineering. He is Research Assistant to Dr Akinsete Oluwatoyin at the Department of Petroleum Engineering, University of Ibadan.



Oluwatoyin Akinsete is an Associate Professor at University of Ibadan, Department of Petroleum Engineering. He completed his PhD in Petroleum Engineering from University of Ibadan in 2015, and his Master of Science in Petroleum Engineering from the same institution in 2001. Recognized for his exceptional contributions, Dr. Akinsete has been honoured with the Professional Engineer designation by the esteemed Council for the Regulation of Engineering in Nigeria (COREN). He has participated in multiple international research collaboration projects in recent years.



Abosede A. Fadayomi is a Senior Director of Petroleum Engineering and Executive Projects with over 19 years of cross industry leadership experience. She had her Bachelor of Science in Petroleum Engineering from University of Ibadan, Master of Science Petroleum Engineering from the same institution, Master's Degree in Business Administration and Data Analytics from Obafemi Awolowo University, Nigeria and India, respectively. She is currently pursuing a Ph.D. in Petroleum Engineering in Petroleum Engineering at the University of Ibadan.



Samuel B. Aderemi is a Regional Digital Strategy Manager for SLB ACE based in Namibia. He has worked almost in all continents in roles from Reservoir Engineer to Global Training Manager, Digital Subsurface Manager and Digital Strategy Manager. He is also driving digital transformation in energy. He holds a Diploma in Electrical/Electronics, B.Sc. and M.Sc. in Petroleum Engineering, an MBA in Energy & Sustainability. He is currently pursuing his Ph.D. in Petroleum Engineering at University of Ibadan, Nigeria.

Research Field

Caleb John: Reservoir Engineering, Reservoir Simulation, Machine Learning, Modeling.

Oluwatoyin Akinsete: Reservoir Engineering, Gas Engineering, Mathematical Modeling, Formation Evaluation, Flow Assurance, Data Analytics.

Abosede A. Fadayomi: Reservoir Engineering, Enhanced oil Recovery, Reservoir Simulation, Project Management.

Samuel B. Aderemi: Reservoir Engineering, Modeling, Data Science, Digital Transformation.