GEOMECHANICS AND GEOENERGY JOURNAL
(GGJ)

Extended Abstract

Estimating Oil-Water Relative Permeability Using Machine Learning: A Case Study from Southwest Iran

Ali Ranjbar[1]*, Mohammad Rasool Dehghani[1], Mahdi Maleki[1]

1- Faculty of Petroleum, Gas and Petrochemical Engineering, Persian Gulf University, Bushehr, Iran

Received: 12 July 2025; Accepted: 22 November 2025

DOI: 10.22107/ggj.2025.559197.1260

Keywords

Relative permeability, CatBoost, Core analysis, Extra trees, SHAP analysis

Abstract

Relative permeability is a key petrophysical parameter controlling multiphase flow behavior in porous media and plays a vital role in reservoir simulation, recovery forecasting, and enhanced oil recovery design. Conventional laboratory methods for determining relative permeability, although accurate, are time-consuming and costly. This study investigates the application of machine learning techniques to estimate oil and water relative permeabilities using core data obtained from a reservoir located in southwest Iran. A total of sixteen core samples were analyzed, and a dataset consisting of 428 data points was constructed using seven input parameters, including absolute permeability, porosity, irreducible water saturation, oil permeability at Swi, viscosities of oil and water, and pressure differential. Four machine learning models—Extra Trees, K-Nearest Neighbors (KNN), CatBoost, and XGBoost—were developed and optimized using Bayesian hyperparameter tuning and validated using five-fold cross-validation.

For water relative permeability estimation, the Extra Trees model demonstrated superior performance with an R^2 of 0.9974, RMSE of 0.0045, and MAE of 0.0007, indicating excellent accuracy and generalizability. In contrast, for oil relative permeability, the KNN model provided the highest predictive accuracy, achieving an R^2 of 0.9973 and the lowest RMSE (0.0113) and MAE (0.0024). Sensitivity analysis using SHAP revealed that water saturation is the most influential parameter in predicting both oil and water relative permeabilities, while oil viscosity and oil permeability at Swi also play significant roles.

Overall, the results demonstrate that machine learning offers a powerful and efficient alternative to experimental methods for relative permeability estimation. The findings contribute to improved reservoir modeling, reduced laboratory workload, and enhanced decision-making in reservoir management.

1. Introduction

Relative permeability is an essential concept for describing multiphase fluid flow in porous media, influencing reservoir simulation outcomes, waterflood performance, and hydrocarbon recovery forecasting[1-3]. Traditional experimental approaches—such as steady-state and unsteady-state coreflood tests—are accurate but operationally intensive, costly, and limited in data availability[4, 5]. Recent advances in machine learning provide an opportunity to model complex nonlinear relationships governing multiphase flow using routinely measured core and fluid properties.

This study aims to develop robust machine learning models capable of predicting oil–water relative permeability using an extensive dataset obtained from southwest Iran. The research addresses the need for faster, scalable, and cost-effective predictive tools and explores the potential of four advanced algorithms—Extra Trees, KNN, CatBoost, and XGBoost. Additionally, the future perspective highlights integrating machine learning predictions into

* **Corresponding Author:** ali.ranjbar@pgu.ac.ir

dynamic reservoir simulation workflows and extending the approach to heterogeneous or low-data environments such as carbonates and shales.

2. Methodology

This study utilized 16 core samples covering various zones of a southwestern Iranian reservoir. Seven key input features—absolute permeability, porosity, irreducible water saturation, oil permeability at Swi, oil viscosity, water viscosity, and differential pressure—were selected to construct a dataset of 428 points. The Spearman correlation analysis indicated the absence of strong linear correlations, supporting the use of machine learning for modeling.

Four machine learning techniques (Extra Trees[6], KNN[7], CatBoost[8], XGBoost[9]) were trained using 70% of data, with 30% retained for testing. Hyperparameters were optimized using Bayesian optimization, and 5-fold cross-validation ensured robust training. Model performance was evaluated using R^2, RMSE, and MAE metrics. Additionally, SHAP analysis[6] was performed to quantify the contribution of each input feature to model outputs.

3. Results and Conclusions

Results indicate that the Extra Trees model achieved the highest accuracy for water relative permeability ($R^2 = 0.9974$; RMSE = 0.0045), displaying excellent generalization. For oil relative permeability, the KNN model outperformed all others ($R^2 = 0.9973$; RMSE = 0.0113). Extra Trees performed poorly for oil permeability estimation, especially at higher permeability values.

SHAP sensitivity analysis revealed water saturation as the dominant factor in both permeability types. Oil permeability at Swi significantly influenced water relative permeability, whereas oil viscosity was the second most important feature for oil permeability estimation.

Overall, machine learning proved to be efficient and reliable for predicting relative permeability, reducing reliance on laboratory measurements and improving the quality of reservoir simulation inputs. The study recommends Extra Trees for modeling water relative permeability and KNN for oil relative permeability.

4. Acknowledgment

The authors acknowledge the support of the Faculty of Petroleum, Gas, and Petrochemical Engineering at Persian Gulf University for providing laboratory data and research facilities. No external funding was received for this work.

5. References

1. *Chapter 7 Relative permeability*, in *Developments in Petroleum Science*, J.R. Fanchi, Editor. 2000, Elsevier. p. 85-101.
2. Viera, M.A.D., et al., *Mathematical and numerical modeling in porous media: Applications in geosciences*. 2012: CRC Press.
3. Muccino, J.C., W.G. Gray, and L.A.J.R.o.G. Ferrand, *Toward an improved understanding of multiphase flow in porous media.* 1998. **36**(3): p. 401-422.
4. Ecay, L., et al., *On the prediction of permeability and relative permeability from pore size distributions.* 2020. **133**: p. 106074.
5. Hossain, Z. *Relative permeability prediction from Image Analysis of thin sections*. in *SPE Europec featured at EAGE Conference and Exhibition?* 2011. SPE.
6. Cao, L., et al., *Interpretable Soft Sensors using Extremely Randomized Trees and SHAP.* 2023. **56**(2): p. 8000-8005.
7. Khamis, Y.E., et al., *Rate of penetration prediction in drilling operation in oil and gas wells by k-nearest neighbors and multi-layer perceptron algorithms.* 2023. **14**(3): p. 755-770.
8. Pu, B., et al., *Machine Learning Evaluation Method of Inter-well Sand Body Connectivity by CatBoost Model: A Field Example from Tight Oil Reservoir, Ordos Basin, China.*
9. Xie, Z., et al., *Prediction of conformance control performance for cyclic-steam-stimulated horizontal well using the XGBoost: a case study in the Chunfeng heavy oil reservoir.* 2021. **14**(23): p. 8161.