

Research Article

Optimizing Well Selection in Hydraulic Fracturing Using Advanced Machine Learning Approaches

Hai T. Nguyen ^{1,2}, Tarek Al-Arbi Ganat,³ and Tu V. Truong⁴

¹Laboratory of Applied Physics, Science and Technology Advanced Institute, Van Lang University, Ho Chi Minh City, Vietnam

²Faculty of Applied Technology, School of Technology, Van Lang University, Ho Chi Minh City, Vietnam

³Petroleum and Chemical Engineering Department, Sultan Qaboos University, Muscat, Oman

⁴Drilling and Production Department, Faculty of Petroleum and Energy, Hanoi University of Mining and Geology, Hanoi, Vietnam

Correspondence should be addressed to Hai T. Nguyen; nguyenthilai@vlu.edu.vn

Received 27 December 2024; Accepted 8 May 2025

Academic Editor: Pramita Mishra

Copyright © 2025 Hai T. Nguyen et al. Geofluids published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

This research evaluates the performance of artificial neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFIS) in selecting well candidates for hydraulic fracturing (HF) in the Bach Ho oilfield, Vietnam. Traditional well selection often depends on expert judgment and deterministic criteria, which may be limited in uncertain and data-constrained reservoir environments. To address this limitation, machine learning models are applied to improve decision-making accuracy. A dataset of 41 wells was analyzed using permeability, porosity, skin factor, reservoir pressure, water cut, and reservoir thickness to predict post-HF daily production rates. Both models were trained and evaluated using RMSE, MSE, MAE, and R^2 . The ANFIS model demonstrated superior accuracy, achieving an RMSE of 4.24, R^2 of 0.93, and MAE of 4.24 on the training set. On the testing set, ANFIS achieved an RMSE of 40.44, R^2 of 0.81, and MAE of 30.33, outperforming the ANN model, which recorded an RMSE of 40.43, R^2 of 0.59, and MAE of 31.86. These results suggest that ANFIS is more effective in capturing nonlinear relationships and handling input uncertainties. The study presents a practical, interpretable tool for supporting petroleum engineers in prioritizing HF candidates, ultimately enhancing oil recovery and resource allocation in complex reservoir settings.

Keywords: ANFIS; ANN; fuzzy logic; hydraulic fracturing; well selection

1. Introduction

Well selection plays a pivotal role in the success of hydraulic fracturing (HF) for enhancing hydrocarbon recovery. This initial and critical step traditionally depends on a comprehensive array of input data, including well information, reservoir characteristics, and geomechanical properties. Conventional well selection methods are often intricate and lack precision, relying heavily on key parameters such as permeability, porosity, viscosity, skin factor, reservoir pressure, wellbore condition, formation thickness, and water cut [1]. However, the integration of artificial intelligence (AI) and machine learning (ML) offers a promising avenue for refining this process. ML techniques, particularly artificial neural networks (ANNs) and adaptive neuro-fuzzy

inference systems (ANFIS) have shown potential in efficiently identifying suitable HF targets, even in the presence of incomplete data. This capability not only reduces costs but also enhances the accuracy of well selection [2]. Consequently, embracing ML in well selection is essential for addressing data limitations and optimizing HF outcomes, underscoring the urgent need for further research in this domain.

AI and ML have significantly transformed petroleum engineering, enhancing operational efficiency, resource optimization, and production outcomes [3]. For instance, Li et al. [4] demonstrated how ML algorithms in drilling and completion engineering improve efficiency and reduce costs through techniques such as ANN, support vector machines, and reinforcement learning. Candra et al. [5] assessed

supervised ML and regression methods to enhance matrix acidizing success in hydraulically fractured sandstone formations.

In production forecasting, Han and Kwon [6] developed a deep neural network that accurately predicted shale gas production by analyzing data from 1239 Montney shale wells. Other notable advancements include Alakbari et al. [7], who introduced a fuzzy logic (FL) model for predicting total drawdown, and Van Pham et al. [8], who enhanced shale well predictions by integrating tree-based pipeline optimization with Bayesian methods. Additionally, Xiao et al. [9] employed a deep convolutional autoencoder for shale gas forecasting without traditional history matching, while Wang et al. [10] utilized physics-informed ANN with Caputo fractional derivatives for long-term forecasts.

The application of ML in HF design has evolved, integrating innovative methodologies and expanding predictive capabilities [11]. Wang and Chen [12] pioneered data mining techniques, including random forest (RF) and ANN, to evaluate well production in unconventional reservoirs. Morozov et al. [13] developed a digital database from thousands of HF jobs using CatBoost for production forecasting with a high R^2 of 0.815. Liu et al. [14] addressed the challenge of predicting production from shale oil wells using particle swarm optimization (PSO) and gradient boosting decision trees. Furthermore, Li et al. [15] combined RF with PSO to optimize fracturing parameters in horizontal wells, enhancing production rate predictions and economic returns.

In contrast to general applications of ML in petroleum engineering, well selection for HF presents unique challenges. This paper focuses on addressing these challenges through ML approaches. Early work by Mohaghegh et al. [16] introduced a three-layer backpropagation ANN that predicted poststimulation well performance with $R^2 = 0.95$ accuracy, eliminating the need for detailed reservoir data. This was further expanded in 1998 by integrating ANN with genetic algorithms, optimizing stimulation design and providing a ranked list of candidate wells with recommended restimulation parameters [17]. Later, Mohaghegh et al. [18] developed an intelligent systems approach, combining ANN, genetic algorithms, and FL to refine restimulation well selection in tight gas sands. Yanfang and Salehi [19] advanced this work by employing a hybrid method that combined FL with interval Type 2 fuzzy sets and ANN for HF well selection, addressing the complexity of nonlinear datasets. Aryanto et al. [20] also emphasized the challenges of well selection, noting that advanced methods such as ANN and ANFIS help reduce uncertainties caused by the nonlinearities of geological and reservoir characteristics. More recently, Asala et al. [21] introduced an integrated ML approach that employed feed-forward ANN and RF models for shale gas refrac well identification, optimizing decision-making and enhancing profitability in shale projects.

The application of FL, first introduced by Zadeh [22], in well selection has advanced considerably. Erlong [23] used fuzzy mathematics to score and optimize target wells, leading to notable improvements in oil production. Xiang-jun and Ting [24] developed an FL-based protocol using a

Mamdani-type fuzzy inference system, which utilizes expert knowledge to translate complex input data into fuzzy outputs, facilitating decisions on HF quality. Zoveidavianpoor et al. [25] reviewed various FL applications, highlighting how this approach reduces uncertainty in well selection. They further introduced ANFIS in 2013 [26], enhancing accuracy in complex carbonate reservoirs by combining FL with ANN. In 2015, they proposed an interval Type 2 fuzzy logic system (FLS), which better manages data uncertainty through fuzzy sets defined by imprecise membership functions [27]. This innovation is particularly valuable in HF well selection, where data ambiguity is common. Guo et al. [28] expanded upon these advancements by integrating FL with grey system theory, effectively managing uncertainties in reservoir evaluation. Their hybrid approach employed the analytic hierarchy process and grey relational analysis for more reliable well selection. Similarly, Zeng et al. [29] applied a Type 1 FLS to evaluate tight gas formations, successfully addressing the complexities of these reservoirs. Most recently, Artun and Kulga [30] implemented a Mamdani-type fuzzy inference system for assessing HF quality, providing a practical method for improving recovery in tight gas reservoirs. Gou et al. [31] developed an advanced hybrid intelligence model integrating a Type 2 Takagi–Sugeno–Kang FLS with grey clustering analysis, which adapts to the nonlinear and uncertain nature of well selection, significantly enhancing prediction accuracy for post-fracturing production outcomes. This review underscores the critical role of FL in managing uncertainties and improving decision-making in well selection. Given the inherent uncertainties and nonlinearities in reservoir characterization and well performance prediction, FL offers a robust framework for addressing imprecision and incomplete data, making it particularly relevant in HF contexts.

Despite significant advancements in ML for production optimization, the challenge of well selection in HF remains unresolved, especially in complex reservoirs with incomplete data. Existing methods often overlook the inherent uncertainties and nonlinear relationships in reservoir data, resulting in inefficient well choices. This paper addresses this critical gap by evaluating two innovative models: ANN and ANFIS, which effectively capture the complexities of well candidate selection. By utilizing the FL capabilities of ANFIS to manage uncertainties, our approach yields more accurate predictions of well performance, even in data-sparse scenarios. This methodology not only enhances well selection accuracy but also streamlines resource allocation in HF operations. Specifically applied to the Bach Ho oilfield, it presents a robust, cost-effective solution for identifying high-potential wells, ultimately optimizing production outcomes and reducing operational costs.

Despite significant advancements in ML for production optimization, the challenge of well selection in HF remains unresolved, especially in complex reservoirs with incomplete data. Existing methods often overlook the inherent uncertainties and nonlinear relationships in reservoir data, resulting in inefficient well choices. This study addresses these limitations by evaluating and comparing two ML models, ANN and ANFIS, to determine their effectiveness in

predicting post-HF production rates and supporting well candidate selection. The objective is to identify a model that not only delivers predictive accuracy but also provides interpretability and robustness in data-scarce environments, such as those found in the Bach Ho oilfield. By leveraging FL, ANFIS is particularly suited to manage uncertainty and capture complex input–output relationships, making it a practical and reliable tool for optimizing HF operations.

2. Geological Background

The study utilizes data from 41 wells in the Upper and Lower Oligocene formations of the Bach Ho oilfield, located 120 km southeast of Vung Tau City, as shown in Figure 1. The Bach Ho oilfield lies within the Cuu Long basin, a pre-Tertiary rift basin part of the center uplift belt. This belt divides the basin into Eastern and Western Blocks. The Cuu Long basin trends northeast–southwest with dimensions of 110×360 km. It borders the mainland to the northwest, the Con Son Uplift (separating it from the Nam Con Son basin) to the southeast, the Khorat-Natuna Uplift to the southwest, and the Tuy Hoa Fault Zone (separating it from the Phu Khanh basin) to the northeast. The basin fill consists of pre-Tertiary continental sediments, reaching a maximum thickness of 8–9 km in the central basin area.

Oil in the Bach Ho field is found within the Miocene Lower, Upper Oligocene, Lower Oligocene, and granite basement rocks. The basement reservoir hosts high-yield block-type oil bodies concentrated in cavernous-fractured zones. In contrast, continental sediments contain structurally complex oil bodies with varying yields, categorized based on their dome shapes and associated lithological, stratigraphic, and tectonic barriers. The oil–water boundary is clearly defined in the horizontal sedimentary oil bodies but is undetermined within the basement reservoir. Figure 2 illustrates the stratigraphic details of the Bach Ho field.

Upper Oligocene sediments, identified from core samples from 16 wells, consist of interbedded cemented siltstones, silts, and sands. The reservoir rock is a grey/brown, fine- to medium-coarse grained, cemented sandstone with high oil-saturated porosity. Some fractures are present in undersaturated zones. The average porosity of the Upper Oligocene is 12.10%, with intergranular porosity dominating over fracture porosity. Lower Oligocene sediments include Productive Zones II and III. Productive Zone II, identified from core samples in wells BH-19 and 20, exhibits entirely intergranular porosity with an average of 16.16%. Productive Zone III, identified from well BH-20, has a lower average porosity of 12.85%, also dominated by intergranular porosity over fracture porosity.

3. Methodology

3.1. Dataset Description. The dataset comprises 41 wells from the Upper Oligocene and Lower Oligocene of the Bach Ho Field, selected based on criteria ensuring optimal potential for stimulation and economic efficiency. HF was successfully conducted on these wells between 2018 and 2023. The

key parameters considered include a positive skin factor ($S > 0$), with all wells having a skin factor ranging from 0.3 to 5.05, indicating the presence of formation damage but also the potential for well stimulation. Permeability values range from 0.21 to 5.92 mD, suggesting diverse well capacities for fluid transmission and varying reservoir zones. Most wells have water cut below 10%, with some exceptions slightly exceeding this threshold, indicating potential water influx issues. Wells were categorized based on reservoir pressures into two groups. Those with current reservoir pressure greater than 0.5 times the initial reservoir pressure were identified as having unconditional potential, while wells with current reservoir pressure less than 0.5 times the initial reservoir pressure required further detailed analysis, with reservoir pressures ranging from 99.6 to 244.5 atm. Additionally, all selected wells have recoverable oil reserves exceeding 20,000 tons. The technical condition of the wells was reviewed, including the quality of cementing and the presence or absence of equipment failures at the well bottom. The dataset reflects varied reservoir and production characteristics, including effective porosity from 7.1% to 19.4%, pay thickness from 20 to 271 m, and production rates from 44 to 124 tons/day, indicating varied volumes of hydrocarbons that can be extracted, influenced by permeability, porosity, skin factor, and reservoir conditions.

3.2. Model Implementation and Comparison. The selected dataset of 41 wells, filtered based on the aforementioned initial criteria, was used for this study. It included six input variables: permeability, effective porosity, skin factor, water cut, reservoir pressure, and reservoir thickness, derived from core data and well test interpretation. The single output variable was the daily production rate, recorded once it stabilized 1 year after HF treatment. The dataset was divided into 70% for training and 30% for testing, ensuring the same sets are used for both models for reproducibility. The data was split in a way that ensures coverage of the full range of each variable, as shown in Figure 3, to maintain a representative sample across all variable distributions for both the training and testing sets. Normalization using the z-score method was implemented to handle variations in the dataset more effectively. The z-score normalization transforms data to have a mean of 0 and a standard deviation of 1, enhancing model performance by improving convergence during training.

Two predictive models, ANN and ANFIS, were implemented and compared to optimize well candidate selection for HF. Both models were evaluated based on their ability to predict postfractured production rates, ensuring a consistent basis for comparison.

The ANN architecture consisted of an input layer with six neurons, one hidden layer containing 20 neurons, and an output layer. The rectified linear unit (ReLU) activation function was used in the hidden layer to address convergence stability and reduce the risk of vanishing gradients. All layers were fully connected. Instead of the Adam optimizer, Bayesian regularization was applied during training to prevent overfitting and improve generalization on small datasets. The learning rate was set at 0.01, and the entire

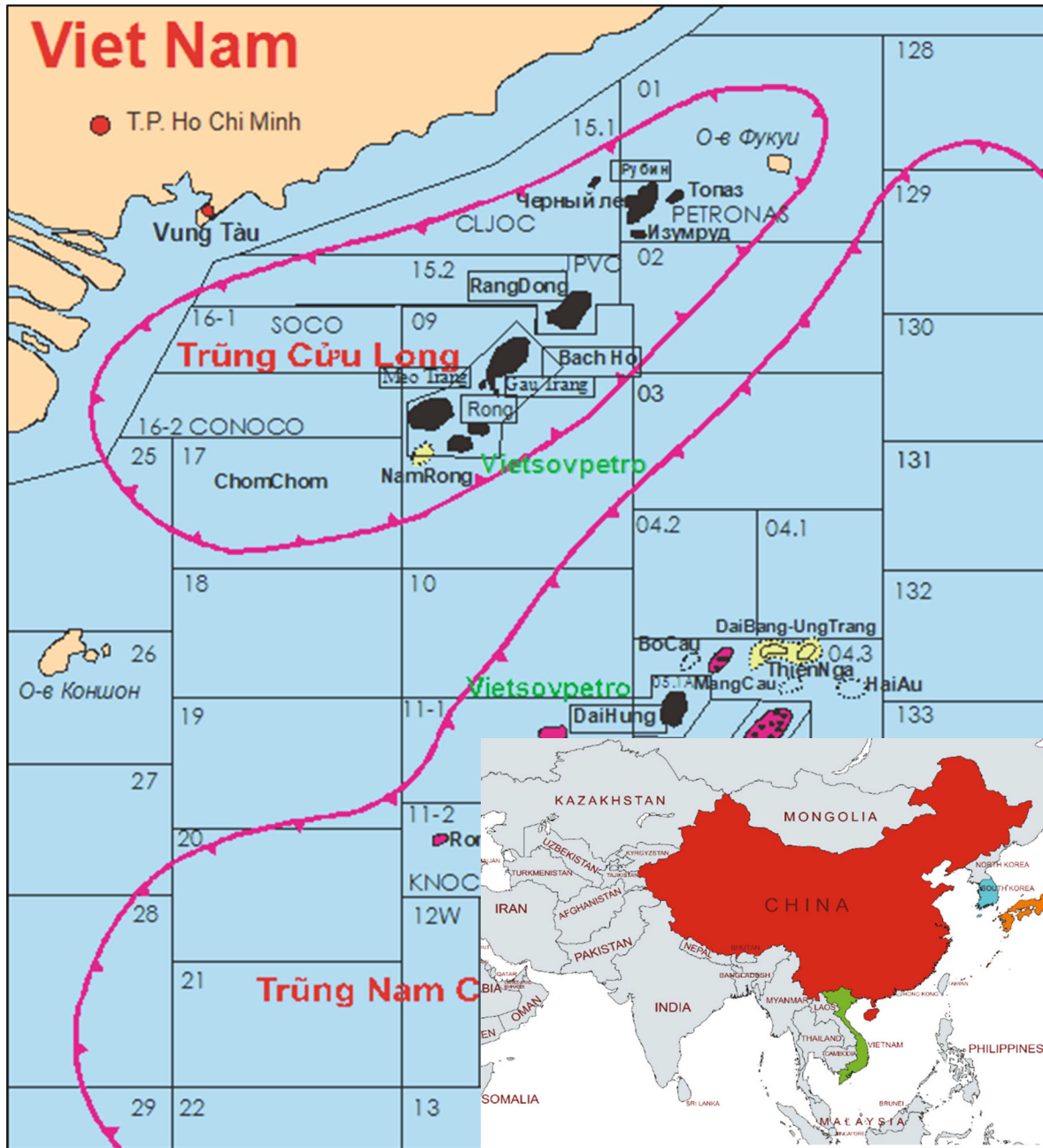


FIGURE 1: Location of Bach Ho oilfield on the southern continental shelf of Vietnam. The bottom right inset, showing Vietnam in green within a wider region, is by Ta-Thi Huong et al. and is used under CC BY 4.0 (10.3390/su132111761).

dataset was used as a batch in each epoch, which was found to be the most stable configuration given the limited number of samples.

The ANFIS model was designed to incorporate FL, which is well-suited for handling uncertainty and nonlinearity in reservoir data. Subtractive clustering was used to identify meaningful clusters in the input space, which were then translated into fuzzy inference rules. The radius of influence was set to 0.4, and the accept ratio was 0.6. Gaussian membership functions were assigned to each input parameter to ensure smooth transitions between fuzzy sets. To avoid over-

fitting and improve generalization, L2 regularization (ridge regression) was introduced during training, penalizing large membership function parameters.

ANFIS Architecture and Workflow. The ANFIS is a hybrid framework that integrates the rule-based reasoning of FL with the adaptive learning capabilities of neural networks. Its layered structure mimics a neural network but incorporates FL operations to model complex, uncertain systems.

The architecture of ANFIS comprises five interconnected layers, each performing a specific function in the fuzzy inference process. The first layer executes fuzzification,

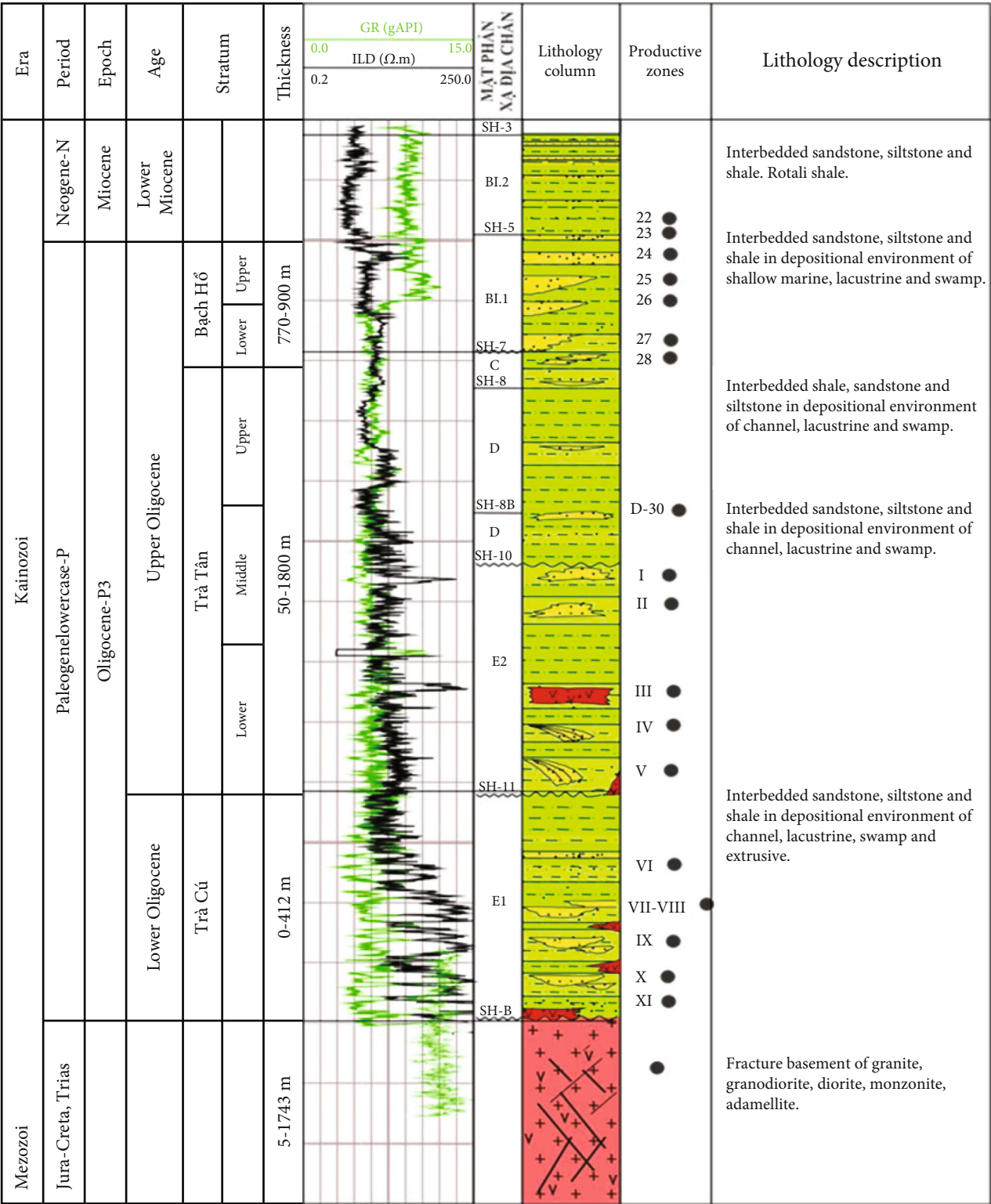


FIGURE 2: Stratigraphy of Bach Ho field.

transforming crisp input values into fuzzy membership grades. In this study, Gaussian membership functions were used for all input variables due to their smooth and continuous shape, which is ideal for capturing the gradual transitions in reservoir properties. Figure 4 illustrates the

membership functions for the six input variables: permeability, porosity, skin factor, water cut, reservoir pressure, and reservoir thickness.

In the second layer, fuzzy inference rules are generated automatically using subtractive clustering. This data-driven

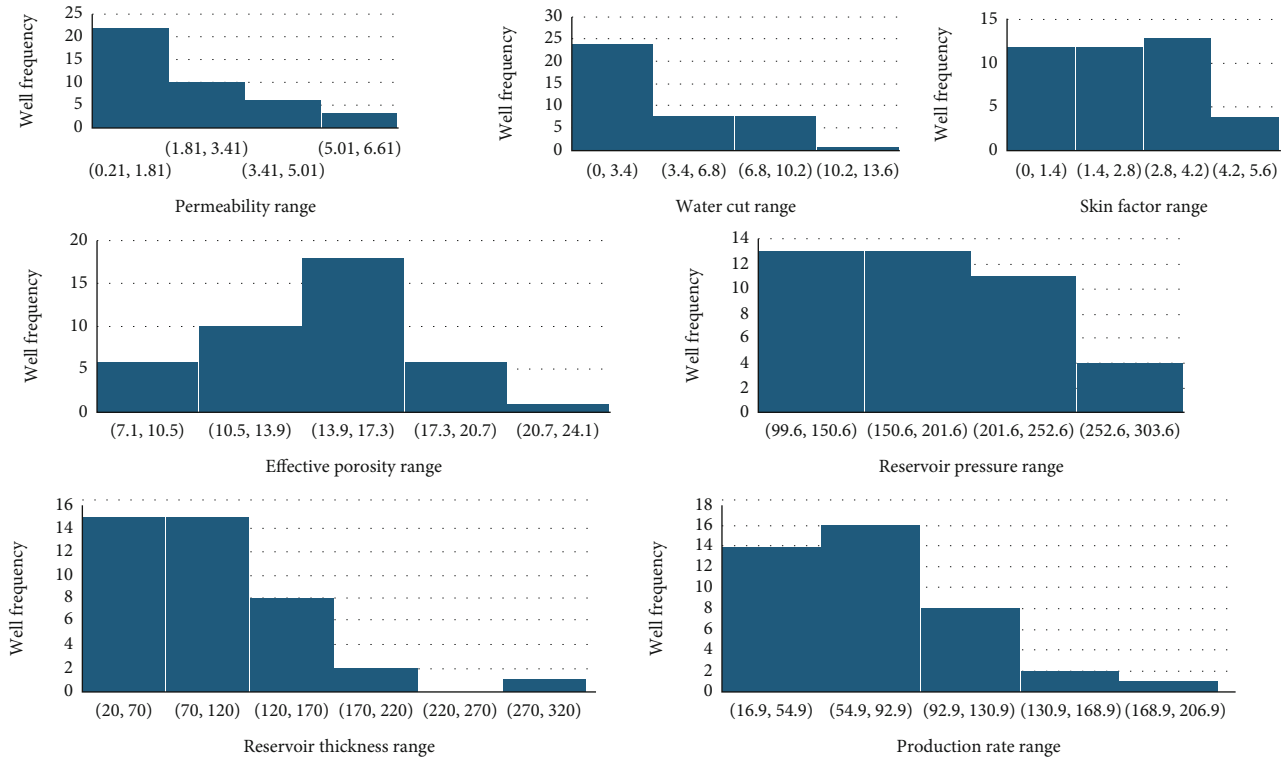


FIGURE 3: Frequency of dataset values.

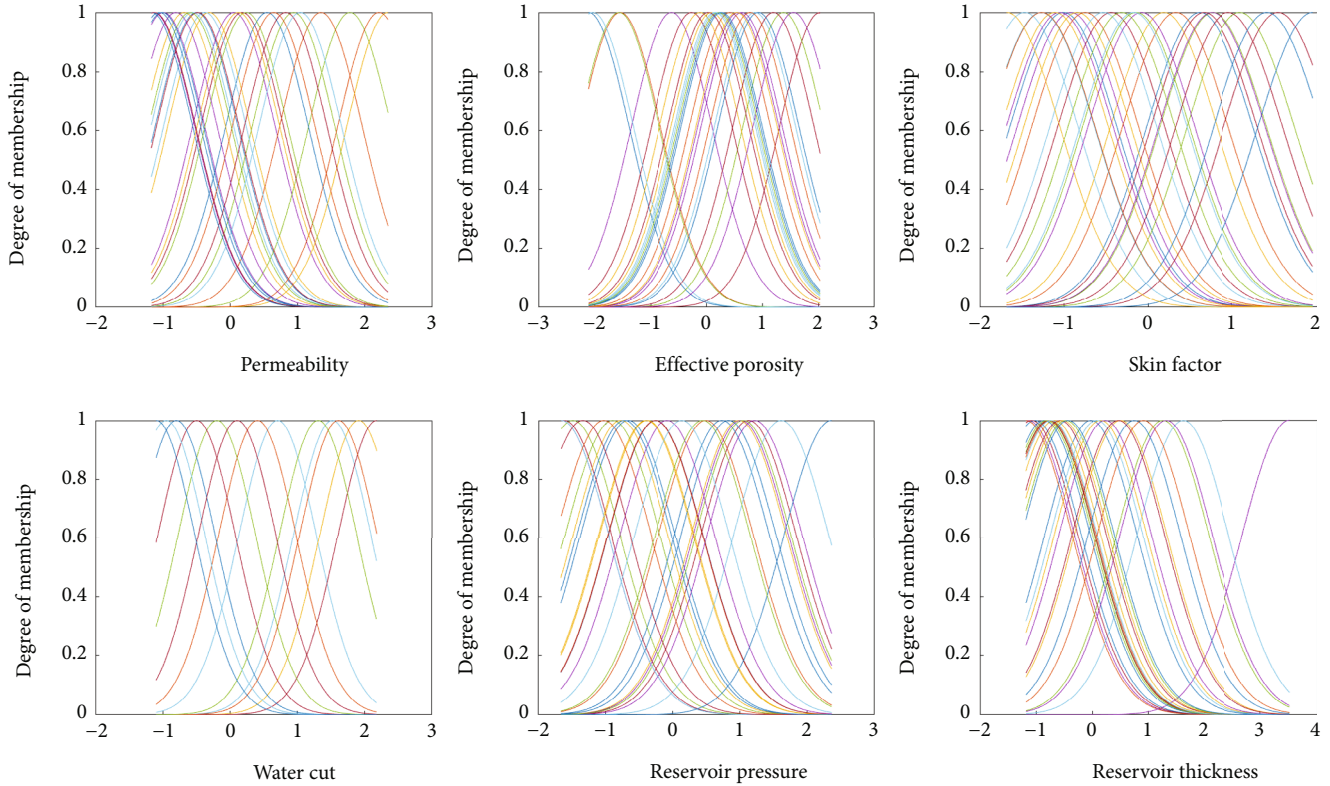


FIGURE 4: Fuzzy membership functions for the variable inputs.

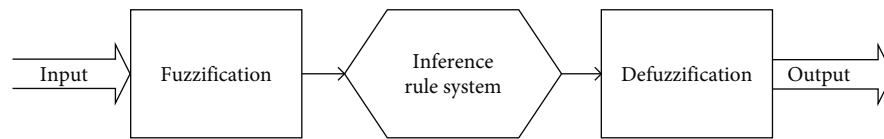


FIGURE 5: General workflow of the fuzzy inference system.

method identifies clusters in the input space, with each cluster defining the premise of a fuzzy rule. The parameters for subtractive clustering were optimized, with the radius set to 0.4 and acceptRatio set to 0.6, to ensure effective rule extraction. The third layer normalizes the firing strengths of all activated rules. The fourth layer performs defuzzification, converting the fuzzy outputs of each rule into weighted crisp values using a first-order Sugeno-type inference system. Finally, the fifth layer aggregates these values to produce a single numerical prediction.

During training, both the membership function parameters and the linear coefficients in the rule consequents are adjusted using a hybrid learning algorithm that combines least squares and backpropagation. This approach enables ANFIS to learn complex input–output mappings while preserving model interpretability.

The overall workflow of the fuzzy inference system is shown in Figure 5, highlighting the transition from input fuzzification to final output prediction. This architecture allows ANFIS to process imprecise or noisy data and generalize effectively from limited samples, common in HF scenarios where field measurements are sparse or uncertain.

Compared to ANN, ANFIS offers improved transparency and interpretability, as it represents its internal logic using fuzzy IF-THEN rules rather than opaque weights. While ANN can model highly nonlinear relationships, it often requires large datasets and lacks a mechanism for reasoning under uncertainty. ANFIS, on the other hand, provides robust performance with small datasets and allows domain experts to understand and validate the decision-making process. These features make ANFIS especially suitable for reservoir engineering applications, where expert oversight and confidence in model reasoning are essential.

The workflow of the performance evaluation between the two models for this studied dataset is illustrated in Figure 6. Both models were implemented in MATLAB and evaluated using the same training and testing datasets, with the production rate as the output variable. Performance on the training set reflects how well the model has learned from the data, demonstrating its ability to fit the training data. In contrast, performance on the testing set assesses how well the model generalizes to unseen data, indicating its predictive accuracy in real-world scenarios. By evaluating both sets, we can identify potential overfitting or underfitting issues, ensuring a thorough comparison of the models' capabilities in predicting production rates.

The evaluation metrics employed include the coefficient of determination (R^2), RMSE, mean square error (MSE), and mean absolute error (MAE). R^2 measures the proportion of variance in the output variable, such as production rate, which is explained by the independent variables, with

a range from 0 to 1 where higher values indicate better predictive accuracy. RMSE assesses the average size of the prediction error, showing how closely the predicted values match the actual values, with lower values indicating better accuracy. MSE calculates the average of the squared differences between actual and predicted values, highlighting the overall error in predictions, where lower values suggest better performance. MAE measures the average size of the absolute errors, providing a direct view of prediction accuracy without squaring the errors, with lower values indicating better performance. These metrics are chosen for their ability to provide a clear view of how well the model performs, helping to find the model that best balances accuracy and error reduction in the study.

4. Results and Discussion

The performance of the ANN and ANFIS models was evaluated using four metrics: RMSE, MSE, MAE, and R^2 , as summarized in Table 1.

The ANFIS model outperformed the ANN model in all four metrics during the training phase. It achieved an RMSE of 4.24, MSE of 73.41, and MAE of 4.24, with an R^2 of 0.93, indicating a good fit to the training data. In contrast, the ANN model had higher RMSE (27.71), MSE (767.89), and MAE (23.72), with a significantly lower R^2 of 0.63. These differences suggest that ANFIS is more effective in capturing the underlying patterns in the data and minimizing prediction errors.

In the testing phase, the ANFIS model achieved a test R^2 of 0.81, outperforming the ANN model's R^2 of 0.59. While their RMSE values were similar (40.44 for ANFIS vs. 40.43 for ANN), ANFIS showed slightly lower MAE (30.33 vs. 30.86) and MSE (1634.54 vs. 1633.33), indicating more consistent prediction accuracy. The improved R^2 demonstrates ANFIS's superior generalization capability. This performance advantage can be attributed to ANFIS's ability to handle uncertainties in the input data through FL, which is particularly beneficial in complex reservoir conditions.

The regression charts in Figures 7 and 8 illustrate the performance differences between the two models visually. The charts highlight how well the predicted values match the actual values for both models, reinforcing ANFIS's advantage in both training and testing scenarios.

Given the nature of the dataset, which includes key reservoir and production parameters such as permeability, porosity, skin factor, pay thickness, water cut, and production rate, the ANFIS model demonstrated superior capability in capturing complex relationships while managing input uncertainty through FL. This dual strength validates the

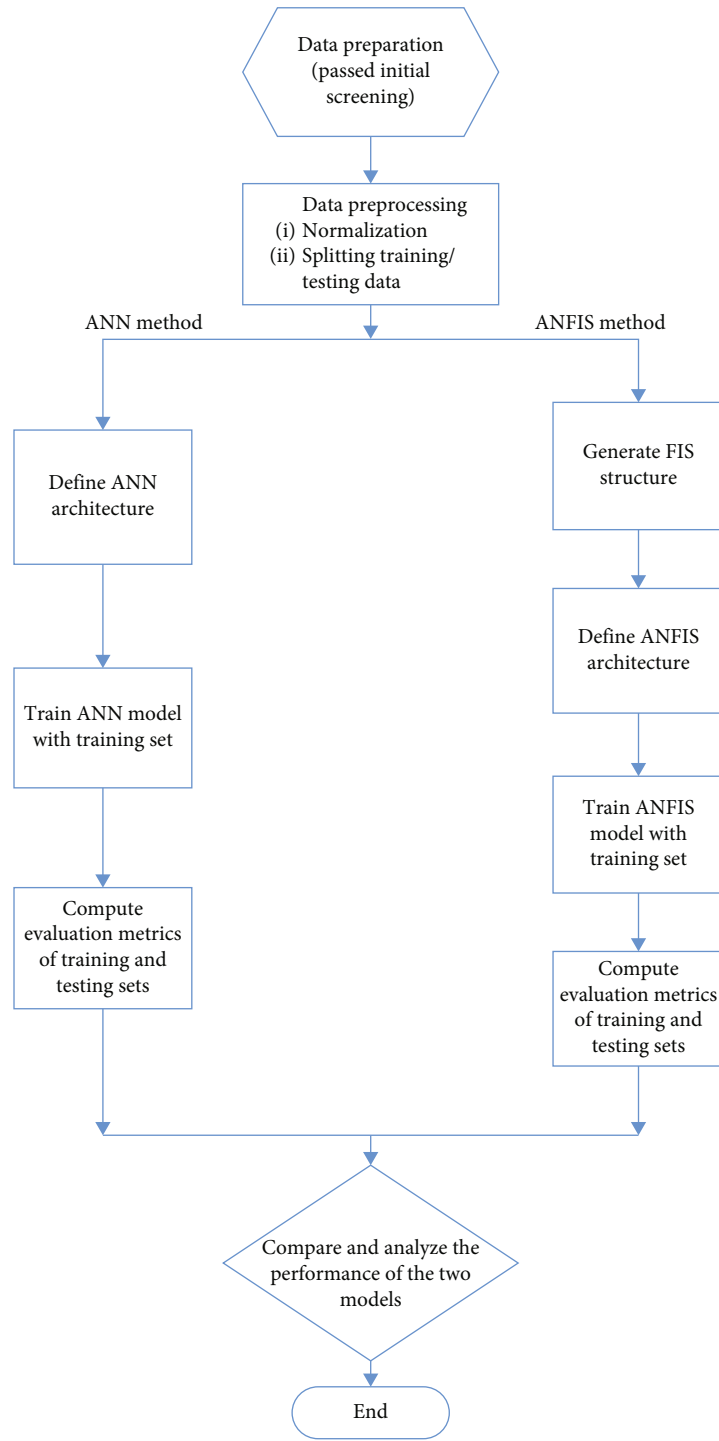


FIGURE 6: Workflow of the performance evaluation between the two models.

TABLE 1: Performance metrics for ANN and ANFIS models.

Model	Training set				Testing set			
	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2
ANN	27.71	767.89	23.72	0.63	40.43	1634.54	30.86	0.59
ANFIS	4.24	73.41	4.24	0.93	40.44	1633.33	30.33	0.81

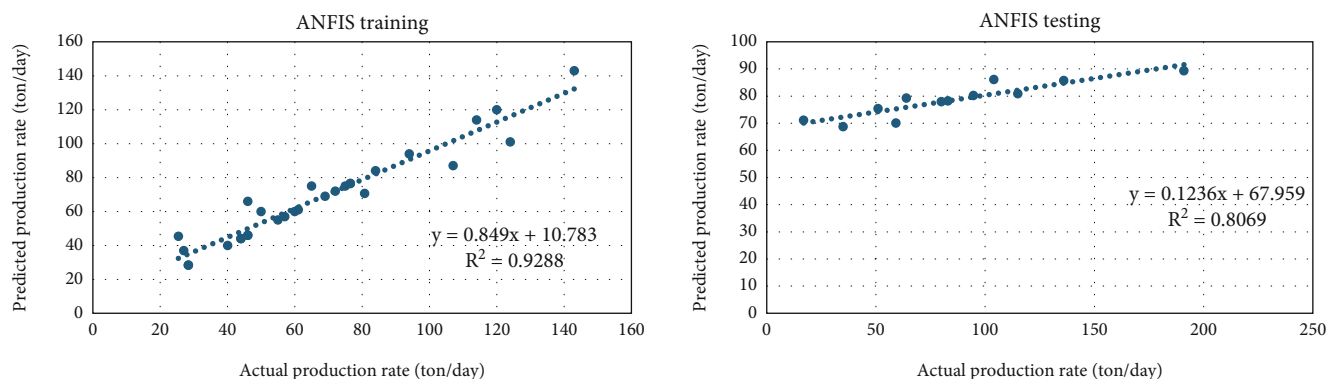


FIGURE 7: Regression plots for ANFIS model on training and testing sets.

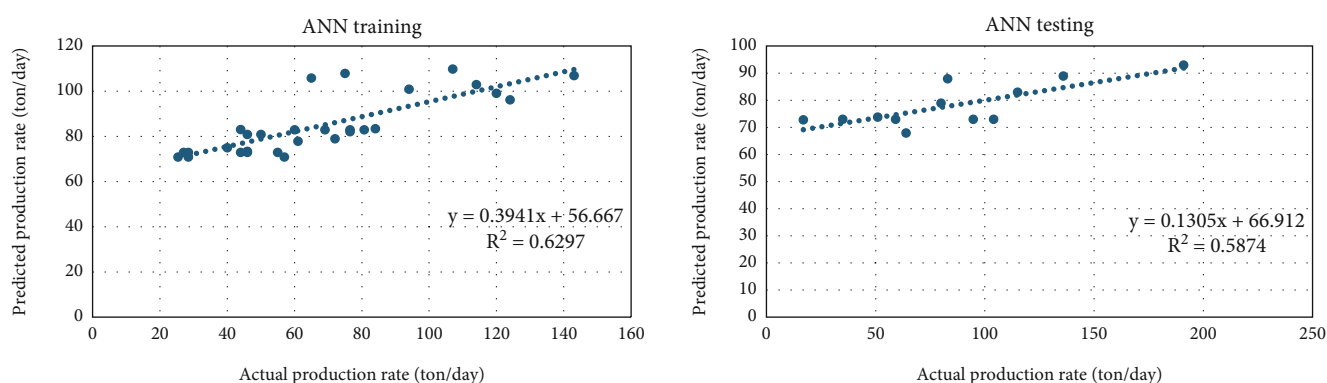


FIGURE 8: Regression plots for ANN model on training and testing sets.

integration of neuro-fuzzy systems as practical tools for HF decision-making, particularly in mature or data-scarce reservoir settings, where predictive accuracy and rule-based transparency offer critical advantages.

Although the ANFIS model performed exceptionally well during training ($R^2 = 0.93$), a drop in testing performance ($R^2 = 0.81$) indicates moderate overfitting. This discrepancy was partially mitigated through the application of ridge (L2) regularization and careful adjustment of clustering parameters during rule extraction. Nonetheless, the limited sample size (41 wells) remains a fundamental constraint, affecting the model's ability to fully generalize across unseen data.

In comparison, the ANN model showed greater difficulty in generalization, achieving a testing R^2 of only 0.59, despite the use of Bayesian regularization. This suggests that while ANN can effectively model nonlinearities under larger datasets, its performance diminishes when data is scarce or uncertain. ANFIS, by contrast, maintained better stability and interpretability under such conditions, reinforcing its suitability for real-world reservoir applications where data quality and quantity are often limited.

5. Limitations and Future Directions

Despite the overall effectiveness of the proposed models, several limitations should be acknowledged. The small sample

size used in this study restricts the models' ability to capture broader geological and operational variability, particularly in more heterogeneous or unconventional reservoirs. As a result, the findings should be interpreted in the context of this data constraint.

Additionally, the models were trained using only static reservoir and operational parameters. The exclusion of time-dependent features such as production history or pressure decline limits the ability to capture dynamic reservoir behavior and long-term production trends. The current framework also does not incorporate uncertainty quantification, which may be important for high-stakes field decisions.

Future research will aim to expand the dataset across multiple fields and geological formations to improve model generalization. The inclusion of additional input variables, such as fracturing fluid properties, proppant characteristics, and geomechanical indicators, may further enhance predictive resolution. Integrating time-series data and adopting ensemble learning or probabilistic modeling techniques may also support more robust and adaptive decision-making for HF optimization.

6. Conclusion

This study evaluated and compared the performance of ANN and ANFIS for well candidate selection in HF operations, using field data from the Bach Ho oilfield. With a

testing R^2 of 0.81 and MAE of 30.33, the ANFIS model demonstrated stronger generalization and predictive capability than ANN, which achieved a lower testing R^2 of 0.59. These results affirm the effectiveness of ANFIS in capturing complex input–output relationships under data-limited conditions.

These findings highlight the practical advantages of using interpretable ML models in HF decision-making workflows. The superior performance of ANFIS in both training and testing scenarios demonstrates its suitability for capturing complex, nonlinear relationships while handling data uncertainty, an essential capability in mature fields where data is often limited or imprecise. In this context, ANFIS serves not just as a predictive tool but also as a decision support system that can assist engineers in prioritizing candidate wells with higher confidence.

Beyond the specific case of the Bach Ho oilfield, the methodology and insights from this study have broader implications for reservoir management and production optimization in other conventional and unconventional settings. The integration of such adaptive models into field development strategies can improve resource allocation, reduce operational risk, and enhance oil recovery, particularly in data-constrained environments common to aging assets and complex geology.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The authors received no specific funding for this work.

References

- [1] F. R. Heydarabadi, J. Moghadasi, G. Safian, and R. Ashena, "Criteria for Selecting a Candidate Well for Hydraulic Fracturing," in *SPE Nigeria Annual International Conference and Exhibition* (SPE, 2010), <https://doi.org/10.2118/136988-MS>.
- [2] A. Samsuri, M. Zoveidavianpoor, and S. R. Shadizadeh, "A Review on Conventional Candidate-Well Selection for Hydraulic Fracturing in Oil and Gas Wells" 2012, <http://eprints.utm.my/30589/2/878991325419748.pdf>.
- [3] S. D. Mohaghegh, "Recent Developments in Application of Artificial Intelligence in Petroleum Engineering," *Journal of Petroleum Technology* 57, no. 4 (2005): 86–91, <https://doi.org/10.2118/89033-JPT>.
- [4] G. Li, X. Song, S. Tian, and Z. Zhu, "Intelligent Drilling and Completion: A Review," *Engineering* 18 (2022): 33–48, <https://doi.org/10.1016/j.eng.2022.07.014>.
- [5] K. Candra, M. Muhammad, and A. Azis, "Supervised Machine Learning and Multiple Regression Approach to Predict Successfulness of Matrix Acidizing in Hydraulic Fractured Sandstone Formation," *ASEAN Journal of Chemical Engineer-*
- ing* 23, no. 1 (2023): 113–113, <https://doi.org/10.22146/ajche.78255>.
- [6] D. Han and S. Kwon, "Application of Machine Learning Method of Data-Driven Deep Learning Model to Predict Well Production Rate in the Shale Gas Reservoirs," *Energies* 14, no. 12 (2021): 3629, <https://doi.org/10.3390/en14123629>.
- [7] F. S. Alakbari, M. E. Mohyaldinn, M. A. Ayoub, A. S. Muhsan, and I. A. Hussein, "A Robust Fuzzy Logic-Based Model for Predicting the Critical Total Drawdown in Sand Production in Oil and Gas Wells," *PLoS One* 16, no. 4 (2021): e0250466, <https://doi.org/10.1371/journal.pone.0250466>.
- [8] V. Van Pham, E. Fathi, and F. Belyadi, "New Hybrid Approach for Developing Automated Machine Learning Workflows: A Real Case Application in Evaluation of Marcellus Shale Gas Production," *Fuels* 17, no. 3 (2021): 286–303, <https://doi.org/10.3390/fuels2030017>.
- [9] C. Xiao, G. Wang, Y. Zhang, and Y. Deng, "Machine-Learning-Based Well Production Prediction Under Geological and Hydraulic Fracture Parameters Uncertainty for Unconventional Shale Gas Reservoirs," *Journal of Natural Gas Science and Engineering* 106 (2022): 104762, <https://doi.org/10.1016/j.jngse.2022.104762>.
- [10] H. Wang, M. Wang, S. Chen, G. Hui, and Y. Pang, "A Novel Governing Equation for Shale Gas Production Prediction via Physics-Informed Neural Networks," *Expert Systems with Applications* 248 (2024): 123387, <https://doi.org/10.1016/j.eswa.2024.123387>.
- [11] C. Sprunger, T. Muther, F. I. Syed, A. K. Dahaghi, and S. Neghabhan, "State of the Art Progress in Hydraulic Fracture Modeling Using AI/ML Techniques," *Modeling Earth Systems and Environment* 8, no. 1 (2022): 1–13, <https://doi.org/10.1007/s40808-021-01111-w>.
- [12] S. Wang and S. Chen, "Insights to Fracture Stimulation Design in Unconventional Reservoirs Based on Machine Learning Modeling," *Journal of Petroleum Science and Engineering* 174 (2019): 682–695, <https://doi.org/10.1016/j.petrol.2018.11.076>.
- [13] A. D. Morozov, D. O. Popkov, V. M. Duplyakov, et al., "Data-Driven Model for Hydraulic Fracturing Design Optimization: Focus on Building Digital Database and Production Forecast," *Journal of Petroleum Science and Engineering* 194 (2020): 107504, <https://doi.org/10.1016/j.petrol.2020.107504>.
- [14] X. Liu, T. Zhang, H. Yang, et al., "Explainable Machine Learning-Based Method for Fracturing Prediction of Horizontal Shale Oil Wells," *Pro* 11, no. 9 (2023): 2520, <https://doi.org/10.3390/pr11092520>.
- [15] W. Li, T. Zhang, X. Liu, et al., "Machine Learning-Based Fracturing Parameter Optimization for Horizontal Wells in Panke Field Shale Oil," *Scientific Reports* 14, no. 1 (2024): 6046, <https://doi.org/10.1038/s41598-024-56660-8>.
- [16] S. Mohaghegh, D. McVey, K. Aminian, and S. Ameri, "Predicting Well-Stimulation Results in a Gas-Storage Field in the Absence of Reservoir Data With Neural Networks," *SPE Reservoir Engineering* 11, no. 4 (1996): 268–272, <https://doi.org/10.2118/31159-PA>.
- [17] S. Mohaghegh, V. Platon, and S. Ameri, "Candidate Selection for Stimulation of Gas Storage Wells Using Available Data With Neural Networks and Genetic Algorithms," in *SPE Eastern Regional Meeting* (SPE, 1998), <https://doi.org/10.2118/51080-MS>.
- [18] S. Mohaghegh, S. Reeves, and D. Hill, "Development of an Intelligent Systems Approach for Restimulation Candidate

- Selection,” in *SPE/CERI Gas Technology Symposium* (SPE, 2000), <https://doi.org/10.2118/59767-MS>.
- [19] W. Yanfang and S. Salehi, “Refracture Candidate Selection Using Hybrid Simulation With Neural Network and Data Analysis Techniques,” *Journal of Petroleum Science and Engineering* 123 (2014): 138–146, <https://doi.org/10.1016/j.petrol.2014.07.036>.
- [20] A. Aryanto, S. Kasmungin, and F. Fathaddin, “Hydraulic Fracturing Candidate-Well Selection Using Artificial Intelligence Approach,” *Journal of Mechanical Engineering and Mechanics* 2, no. 2 (2018): <https://doi.org/10.33021/jmem.v2i02.322>.
- [21] H. I. Asala, J. A. Chebeir, V. Manee, I. Gupta, A. D. Taleghani, and J. A. Romagnoli, “An Integrated Machine-Learning Approach to Shale-Gas Supply-Chain Optimization and Refrac Candidate Identification,” *SPE Reservoir Evaluation & Engineering* 22, no. 4 (2019): 1201–1224, <https://doi.org/10.2118/187361-PA>.
- [22] L. A. Zadeh, “Fuzzy Sets,” *Information and Control* 8, no. 3 (1965): 338–353, [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).
- [23] Y. Erlong, “Selection of Target Wells and Layers for Fracturing With Fuzzy Mathematics Method,” in *Proceedings of the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery* (IEEE, 2009), 366–369, <https://doi.org/10.1109/FSKD.2009.808>.
- [24] X. Xiang-jun and Y. Ting, “Application of T-S Fuzzy Model in Candidate-Well Selection for Hydraulic Fracturing,” in *Fuzzy Information & Engineering and Operations Research & Management*, eds. B.-Y. Cao and H. Nasser (Springer Berlin Heidelberg, 2013), 563–570.
- [25] M. Zoveidavianpoor, A. Samsuri, and S. R. Shadizadeh, “Fuzzy Logic in Candidate-Well Selection for Hydraulic Fracturing in Oil and Gas Wells: A Critical Review,” *International Journal of Physical Sciences* 7, no. 26 (2012): 4049–4060, <https://doi.org/10.5897/IJPS12.042>.
- [26] M. Zoveidavianpoor, A. Samsuri, and S. R. Shadizadeh, “Hydraulic Fracturing Candidate-Well Selection by Interval Type-2 Fuzzy Set and System,” in *IPTC 2013: International Petroleum Technology Conference* (European Association of Geoscientists & Engineers, 2013), <https://doi.org/10.2523/IPTC-16615-MS>.
- [27] M. Zoveidavianpoor and A. Gharibi, “Applications of Type-2 Fuzzy Logic System: Handling the Uncertainty Associated With Candidate-Well Selection for Hydraulic Fracturing,” *Neural Computing and Applications* 27, no. 7 (2016): 1831–1851, <https://doi.org/10.1007/s00521-015-1977-x>.
- [28] J. Guo, L. Tao, Y. Liu, N. He, and X. Zhou, “A New Refracturing Candidate Selection Method for Multi-Fractured Horizontal Wells in Tight Oil Reservoirs,” *Journal of Geophysics and Engineering* 16, no. 4 (2019): 789–800, <https://doi.org/10.1093/jge/gxz049>.
- [29] F. Zeng, D. Wang, J. Guo, S. Zhang, P. Zhang, and B. Zhang, “Optimal Selection of Stimulation Wells Using a Fuzzy Multi-criteria Methodology,” *Mathematical Problems in Engineering* 2019, no. 1 (2019): 4084982, <https://doi.org/10.1155/2019/4084982>.
- [30] E. Artun and B. Kulga, “Selection of Candidate Wells for Re-Fracturing in Tight Gas Sand Reservoirs Using Fuzzy Inference,” *Petroleum Exploration and Development* 47, no. 2 (2020): 413–420, [https://doi.org/10.1016/S1876-3804\(20\)60058-1](https://doi.org/10.1016/S1876-3804(20)60058-1).
- [31] B. Gou, C. Wang, T. Yu, and K. Wang, “Fuzzy Logic and Grey Clustering Analysis Hybrid Intelligence Model Applied to Candidate-Well Selection for Hydraulic Fracturing in Hydrocarbon Reservoir,” *Arabian Journal of Geosciences* 13, no. 19 (2020): 975, <https://doi.org/10.1007/s12517-020-05970-y>.