ISSN: 2447-3359

REVISTA DE GEOCIÊNCIAS DO NORDESTE

Northeast Geosciences Journal

v. 11, nº 2 (2025)

https://doi.org/10.21680/2447-3359.2025v11n2ID39943

EVALUATION OF NEURAL NETWORK IN RESERVOIR CHARACTERIZATION WITH NMR DATA: PREDICTION OF PERMEABILITY, FZI, RQI, HFU PARAMETERS AND CLASSIFICATION OF POROSITY TYPES

Maurício Gabriel Lacerda Mafra[1]; Milton Morais Xavier Junior[2]

[1] UFRN, PPGG, Natal/RN, Brasil. Email: mauricio.mafra.103@ufrn.edu.br
ORCID: https://orcid.org/0000-0003-2720-4655

[2] UFRN, PPGG, Natal/RN, Brasil. Email: milton.morais@ufrn.br
ORCID: https://orcid.org/0000-0002-9072-6475

Abstract: Machine learning has advanced scientific research by enabling the analysis of large and complex datasets, including uncorrelated variables. In reservoir characterization, petrophysical parameters such as porosity and permeability are essential for calculating indicators like Flow Zone Indicator (FZI), Reservoir Quality Index (RQI), and Hydraulic Flow Units (HFU), and that can aid in pore type classification. Nuclear Magnetic Resonance (NMR) is a powerful technique in this context, as it enables direct porosity measurements and permeability estimation through models. In this study, porosity and gas permeability data from 506 carbonate samples were used to evaluate four semi-empirical models (SDR, Timur-Coates, Rios, and Han) and a deep learning model, the Multi-Layer Perceptron (MLP). The MLP outperformed the semi-empirical models, achieving an R^2 of 0.79 and $\sigma = 3.07$ for training and 0.71 and $\sigma = 3.92$ for testing. It also effectively differentiated HFUs and closely matched laboratory results. In pore type classification, the MLP model showed superior performance. These results highlight the potential of integrating NMR data with deep learning to improve HFU, FZI, and RQI predictions and support more accurate pore type characterization.

Keywords: Neural Network; Permeability; NMR

Resumo: O aprendizado de máquina tem impulsionado avanços científicos ao viabilizar a análise de grandes e complexos conjuntos de dados, mesmo entre variáveis não correlacionadas. Na caracterização de reservatórios, parâmetros petrofísicos como porosidade e permeabilidade são fundamentais para o cálculo de indicadores como o Flow Zone Indicator (FZI), Reservoir Quality Index (RQI) e Hydraulic Flow Units (HFU), e que podem auxiliar na classificação de tipos de poros. A Ressonância Magnética Nuclear (RMN) é uma técnica robusta nesse contexto, permitindo medições diretas de porosidade e estimativas de permeabilidade por meio de modelos. Neste estudo, dados de porosidade e permeabilidade a gás de 506 amostras de carbonatos foram utilizados para avaliar quatro modelos semiempíricos (SDR, Timur-Coates, Rios e Han) e um modelo de aprendizado profundo, o Perceptron Multicamadas (MLP). O MLP superou os modelos semiempíricos, com R^2 de 0,79 e $\sigma = 3{,}07$ para treinamento e 0,71 e $\sigma = 3{,}92$ para teste. O modelo também diferenciou eficientemente as HFUs e apresentou valores próximos aos obtidos em laboratório. Na classificação dos tipos de poros, o MLP apresentou desempenho superior. Assim, a integração de dados de RMN com modelos de aprendizado profundo aprimora significativamente a caracterização de HFU, FZI e RQI.

Palavras-chave: Redes Neurais; Permeabilidade; RMN

Recebido: 22/04/2025; Aceito: 23/05/2025; Publicado: 25/07/2025.

1. Introduction

High performance in reservoir exploration is essential at all stages, from the identification of zones of interest to reservoir modeling, production optimization, and planning of future operations (LUCIA et al., 2003; TIAB and DONALDSON, 2016). In reservoirs with high heterogeneity, this process can lead to several complications. Achieving such high performance requires thorough reservoir characterization, which largely depends on petrophysical modeling. The key parameters in this context are porosity (ϕ) and permeability (k) (CANNON, 2018). Using porosity and permeability data, it is possible to derive additional indicators that assist in identifying zones of interest within a reservoir, such as the Reservoir Quality Index (RQI) (1), the Normalized Porosity (ϕ_z) (2), the Flow Zone Indicators (FZI) (3), and the Hydraulic Flow Units (HFU) (AMAEFULE et al., 1993; TIAB and DONALDSON, 2016; EFTEKHARI et al., 2024). Based on these indicators, it is also possible to characterize pore types using the method proposed by Soto et al. (2010), which involves the transformation coefficient (r) (4), as well as the "polar arm" angle derived from the FZI value (θ_{Polar}) (5).

$$RQI = 0.0314 \cdot \left(\sqrt{\frac{k}{\phi}}\right) \tag{1}$$

$$\phi_z = \frac{\phi}{1-\phi} \tag{2}$$

$$FZI = \frac{RQI}{\phi_z} \tag{3}$$

$$r = \phi_z \cdot \left(\sqrt{FZI^2 + 1}\right) \tag{4}$$

$$\theta_{Polar} = \tan^{-1}(FZI) \tag{5}$$

Laboratory-based permeability measurements are not feasible across an entire reservoir, making it necessary to estimate permeability through modeling approaches. Nuclear Magnetic Resonance (NMR) techniques provide not only porosity data but also enable permeability estimation through semi-empirical models or machine learning approaches, establishing NMR as an advanced and robust petrophysical tool for acquiring essential information in reservoir exploration (KENYON et al., 1988; COATES et al., 1999; DUNN, 2002; WEI et al., 2024). Therefore, comparing the results of different permeability models based on NMR data and validating them against laboratory measurements becomes one of the key objectives in understanding and calibrating reservoir models (GAVIDIA et al., 2024). Machine learning models play a valuable role in this context, as laboratory results do not always show direct correlation with model-based estimates, which can lead to significant errors if these estimates are used in subsequent reservoir characterizations (GARIA et al., 2022).

Therefore, this study aims to obtain laboratory measurements of porosity and gas permeability from 506 samples collected from a highly heterogeneous environment, in order to compare these results with those obtained from NMR techniques. Several semi-empirical models, as well as a machine learning model, will be employed to determine which approach best approximates the laboratory data. These data will also be used to extract relevant information from the samples, enabling their characterization and the calculation of FZI and HFU. This, in turn, will allow for an evaluation of how the model predictions behave when applied to reservoir assessments.

1.1. Nuclear Resonance Magnetic

NMR provides data such as porosity and relaxation time distributions, which are related to spin concentrations and the structure of the porous medium. It is a non-destructive technique that can be applied both in the laboratory and in well environments, using core plugs, whole cores, or during drilling operations. The technique is based on the relaxation of hydrogen nuclear spins, which are present in water and hydrocarbons. Its operating principle relies on the application of an external magnetic field (B_0), which causes the nuclear spins to precess. An additional magnetic field (B_1), applied at an angle, induces a torque that alters the precession direction toward B_1, generating magnetizations (M_0 and M_1). The measurement of the transverse relaxation time (T_2) is performed using specific pulse sequences, such as the CPMG sequence (CARR and PURCELL, 1954; MEIBOOM and GILL, 1958) or the spin echo sequence (HAHN, 1950).

From this relaxation behavior, a T_2 relaxation time distribution curve can be obtained through iterative methods such as the Simultaneous Iterative Reconstruction Technique (SIRT) (CHEN et al., 2010). These relaxation time distributions are used in permeability modeling, as they are commonly interpreted in relation to pore sizes (COATES et al., 1999; HAN et al., 2018). The main models for estimating permeability using NMR are based on information extracted from the T_2 distribution curve (KENYON et al., 1988; COATES, 1999; RIOS et al., 2011; HAN et al., 2018).

1.2. Reservoir and Pore Type Characterization

With porosity and permeability data, it becomes possible to characterize the reservoir through parameters such as the RQI, FZI and HFU. The RQI reflects the reservoir's potential for exploitation, as a given location may exhibit high permeability but low porosity, or vice versa conditions that can result in low reservoir quality. However, even regions with low RQI may present hydraulic flow behavior similar to regions with high RQI, depending on their normalized porosity. For this reason, evaluating the FZI becomes essential. When different intervals present similar FZI values, they are grouped into a HFU. The clustering of HFUs can be performed arbitrarily, based on predefined thresholds, adjusted to geological information, or determined using statistical classification methods. As demonstrated by Soto et al. (2010), the r and θ_{Polar} can be used to distinguish porosity types, such as fracture/vuggy porosity and intercrystalline porosity, by applying a sigmoidal cutoff function (6).

$$S(r) = \frac{A + B}{\left\{1 + e^{\left[-\left(\frac{r-C}{D}\right)\right]}\right\}} \qquad (6)$$

$$A = -3{,}5916207, B = 5{,}06265818, C = -0{,}72243226 \; e \; D = 0{,}371324681$$

1.3. Geology

The study area is located in the Potiguar Basin (Figure 1-A), situated along the Brazilian equatorial margin. This basin was formed by rifting processes during the opening of the South Atlantic Ocean in the Jurassic–Cretaceous period (MATOS, 1992; de CASTRO, 2012), and it comprises both onshore and offshore segments. The Jandaíra Formation represents a carbonate platform deposited between the Turonian and Campanian stages (CÓRDOBA, 2001). This region was selected due to its strong analogy with pre-salt reservoirs and its surface exposure, which allows for detailed investigation of pore space properties with high accuracy and low cost. A total of 506 plug samples were analyzed, collected from various locations within the Jandaíra Formation, including both outcropping areas and core samples from wells (Figure 1-B). These locations exhibit intense karstification and complex permoporous structures with high heterogeneity (LOPES et al., 2023; ARAUJO et al., 2023), which are expected to contribute to the evaluation of porosity types as well as to the challenges involved in permeability prediction and the estimation of reservoir quality indices such as FZI, RQI, and HFU.

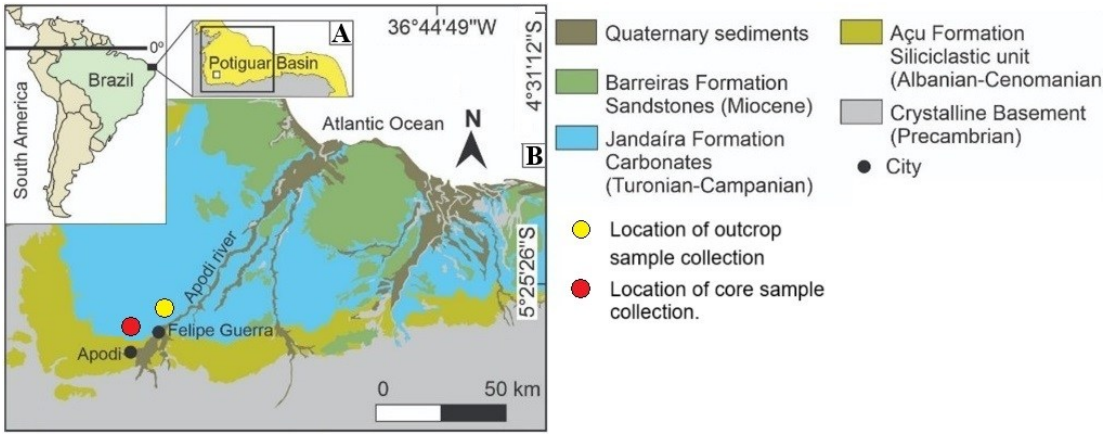

Figure 1 – A) Location of the Potiguar Basin in Northeastern Brazil; B) Regional map of the Potiguar Basin.
Source: Authors (2025)

1.4. Review of Artificial Neural Networks

Machine learning algorithms are capable of handling complex datasets and making accurate predictions, even in nonlinear contexts. In supervised learning models, the data are labelled, which guides the adjustment of model parameters to minimize prediction errors. Among the most widely used models is the Artificial Neural Network (ANN), particularly the Multilayer Perceptron (MLP) (GÉRON, 2021). The perceptron is the fundamental unit of ANNs and is a mathematical model composed of functions (w_{nm}) that modify the input data (x_i) or the outputs of other perceptrons in order to activate and propagate information toward generating the desired outcome. The combination of multiple organized perceptrons and bias terms (b) in layers allows the network to transform input data into virtually any output. In regression problems, this output corresponds to a unique estimated value (Figure 2).

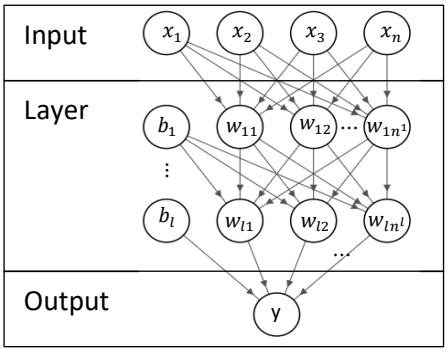

Figure 2 – The architecture of an MLP model begins with the connection of input data to the perceptrons in the first layer, which in turn connect to perceptrons in subsequent layers, propagating through the network until the input data produce an output result.
Source: Authors (2025)

The architecture of an MLP model defines its structure in terms of layers, the number of perceptrons, and activation functions, influencing how data are processed. Training adjusts weights and biases through backpropagation (HINTON et al., 2006) over multiple epochs to minimize error. The architecture involves hyperparameters set by the user, while

parameters are adjusted automatically during training. Since there is no exact rule for the optimal hyperparameters (WOLPERT, 1996), multiple configurations must be tested. GridSearchCV from Scikit-Learn (PEDREGOSA et al., 2015) automates this process by testing combinations of hyperparameters and applying cross-validation.

1.1. Semi-Empirical Models

The most common semi-empirical models used to estimate permeability from NMR data include the Schlumberger Doll Research (SDR) model (7) (KENYON et al., 1988), the Timur-Coates model (8) (COATES et al., 1999), the model proposed by Han et al. (2018) (9), and the statistical model introduced by Rios et al. (2011). These models consist of exponential functions whose coefficients are calibrated using NMR results from each sample set alongside laboratory-measured permeability, enabling reasonably accurate permeability predictions for new samples using only NMR data. The conventional models, SDR, Timur-Coates, and Han, determine specific ranges or values from the T_2 distribution curve to adjust the exponential coefficients via linear regressions. The SDR model uses only the NMR porosity (φ_{NMR}) and the logarithmic mean of the T_2 distribution ($T_{2\,\mathrm{lm}}$) (10). The Timur-Coates model (COATES et al., 1999) also uses φ_{NMR} but incorporates Bulk Volume Irreducible (BVI) and Free Fluid Index (FFI) (TIMUR, 1969) to evaluate portions of the T_2 distribution curve. Han's model (HAN et al., 2018) likewise employs φ_{NMR} but additionally uses fractions of pore sizes (S_1, S_2, S_3 and S_4), calibrated by Mercury Intrusion Capillary Pressure (MICP) measurements. The statistical model is based on identifying statistical correlations between the distribution curves of samples and their permeability, employing Partial Least Squares Regression (PLSR), which extracts components that simultaneously maximize the variance of the distribution and its covariance with permeability (GELADI and KOWALSKI, 1986; MEHMOOD et al., 2012).

$$k_{SDR} = aT_{2\,lm}{}^{b}\phi_{NMR}{}^{c} \tag{7}$$

$$k_{Timur-Coates} = \left[\left(\frac{\phi_{NMR}}{C}\right)^{a} \cdot \frac{FFI}{BVI}\right]^{b} \tag{8}$$

$$k_{Han} = a \cdot \phi_{NMR}^{b} \cdot \frac{S_3^c \cdot S_4^d}{S_1^e \cdot S_2^f} \tag{9}$$

$$T_{2\,lm} = 10^{\left[\frac{\sum LOG_{10}\left(T_2^{(i)}\right)\cdot\phi_{NMR}^{(i)}}{\sum \phi_{NMR}^{(i)}}\right]} \tag{10}$$

1.2. Workflow

The study begins with the collection of petrophysical gas data for porosity (ϕ_{Gas}) and permeability (k_{Gas}), which serve as labels to validate model predictions and their potential reservoir characterizations. Using these values of ϕ_{Gas} and k_{Gas}, five groups were identified based on their HFU. From the total of 506 samples, 70% of the samples from each group were selected for training data, and 30% were reserved for testing. After acquiring the NMR data, permeability modeling was performed using the SDR, Timur-Coates, Rios, Han models, and an MLP model specifically trained for this study. The permeability estimates from each model k_{model} along with the φ_{NMR} were then used to derive reservoir information (FZI, RQI, and HFU) for each model and compared against laboratory gas data. Finally, these data were also utilized to classify pore types according to the model proposed by Soto et al. (2010).

2. Methodology

The samples consist of cylindrical plugs up to 50 mm in length and 25 mm in diameter. Of the 506 samples collected (Figure 3 – A, B, and C), 493 are plugs extracted from the sides of well cores (Figure 1-C), and 13 are plugs obtained from outcrops (Figure 1-B), both from the Jandaíra Formation, Potiguar Basin (Figure 1-A) (ARAÚJO et al., 2023; LOPES et al., 2023). Porosity and gas permeability data were acquired simultaneously while the gas flow was still in an unstable

state, followed by Klinkenberg correction (KLINKENBERG, 1941). The MesoMr12-060H-I equipment from Niumag (Figure 3-D) was used to acquire T2 relaxation times (Figure 3-E) through the CPMG pulse sequence. The SIRT method was employed to invert the signal and obtain the T2 time distribution (Figure 3-F).

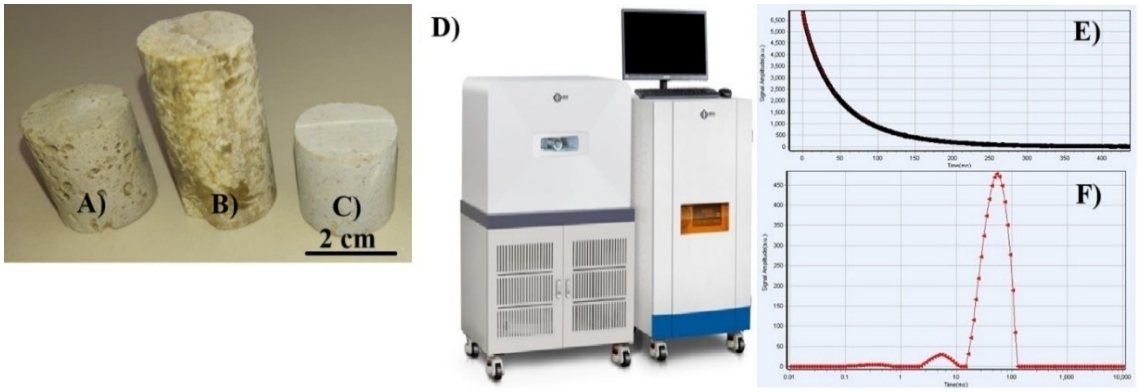

Figure 3 - Plug samples taken from: A) core sample with visible porosity, B) outcrop with stylolite, C) core sample with low porosity, D) MesoMr12-060H-I equipment from Niumag, which enables the acquisition of E) T_2 relaxation curve, and F) curve inversion using the SIRT method.
Source: Authors (2025)

To identify the HFUs, the KMeans method from the Scikit-Learn library (PEDREGOSA et al., 2015) was used, initially with 6 clusters. Since two HFUs exhibited very similar values, they were merged, resulting in a total of 5 HFUs. Cluster classification divided the data into subgroups for training and testing using the Stratified Shuffle Split method, also from Scikit-Learn. This method performs a pseudo-random split based on class labels, ensuring that 70% of the data from each HFU was allocated for training and 30% for testing, thus maintaining representation of all classes in both subsets.

Multiple linear regression, used to estimate the model coefficients, was performed through logarithmic transformation of the models. This transformation converts the exponents of each parameter into multiplicative coefficients of the logarithms, which are now summed. Given the available gas permeability values, multiple linear regression determines these multiplicative coefficients for each logarithm, thereby fitting the model to the training data. Once the adjusted coefficients are obtained from the training set, it becomes possible to predict the values in the test set. To assess the prediction errors of the permeability models, the coefficient of determination (R^2) (11) and the Root Mean Square Error in logarithmic scale (σ) (12) (Kenyon et al., 1988) were evaluated, since permeability spans several orders of magnitude.

$$R^2 = 1 - \frac{\Sigma\left(y_{real}^{(i)} - y_{predict}^{(i)}\right)^2}{\Sigma\left(y_{real}^{(i)} - \bar{y}_{real}\right)^2} \tag{11}$$

$$log_{10}(\sigma) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[log_{10}\left(k_{model}^{(i)}\right) - log_{10}\left(k_{Gas}^{(i)}\right)]^2} \tag{12}$$

To select the optimal MLP model, multiple sets of hyperparameters were defined and trained in successive stages. In each stage, possible hyperparameter combinations were provided to GridSearchCV, which performed training using 5-fold cross-validation on subsets of the training data and returned the model with the lowest R^2. The hyperparameter sets in each subsequent stage were based on the best-performing models from the previous stage. The best model from each stage was then evaluated based on the σ error on the test dataset, and the model with the lowest σ was selected as the one that best fit the dataset.

3. Results

Permeability predictions from each semi-empirical model, compared to gas-derived data, are illustrated in Figures 4 and 5 for the training and test datasets, respectively. The clusters are represented by the colors of the data points in the plots: red for cluster 0, blue for cluster 1, green for cluster 2, purple for cluster 4, and orange for cluster 5. The maximum, minimum, and average permeability values from both the gas data and the models, as well as the R^2 and σ error values for both datasets, are presented in Table 1. A total of 14 training stages were performed for the MLP model. The number of hyperparameter combinations varied between 34,560 and 11,520 per stage, totaling 283,392 trained architectures. The optimal model was configured with two hidden layers, each containing 48 perceptrons, using the hyperbolic tangent (tanh) activation function and the Adam backpropagation optimization method.

Table 1 – Results of maximum, minimum, and average values, R^2 coefficient, and σ error for each model across both datasets.

Models	*Training Data*					*Test Data*				
	Max	**Min**	**Mean**	**R^2**	***σ***	**Max**	**Min**	**Mean**	**R^2**	***σ***
Gas	786.347	0.001	4.212	1.00	1.00	517.279	0.001	4.623	1.00	1.00
SDR	4.067	0.0005	0.165	0.54	5.29	2.720	0.0001	0.176	0.52	5.86
Han	8.362	0.001	0.244	0.60	4.79	6.617	0.0002	0.260	0.58	5.24
Timur-Coates	3.400	0.001	0.177	0.58	4.92	2.273	0.0002	0.185	0.54	5.66
Rios	23278.648	0.002	66.393	0.71	3.74	15750.870	0.0025	107.808	0.56	5.49
MLP	661.253	0.002	2.904	0.79	3.07	839.999	0.0026	6.205	0.72	3.92

Source: Authors (2025)

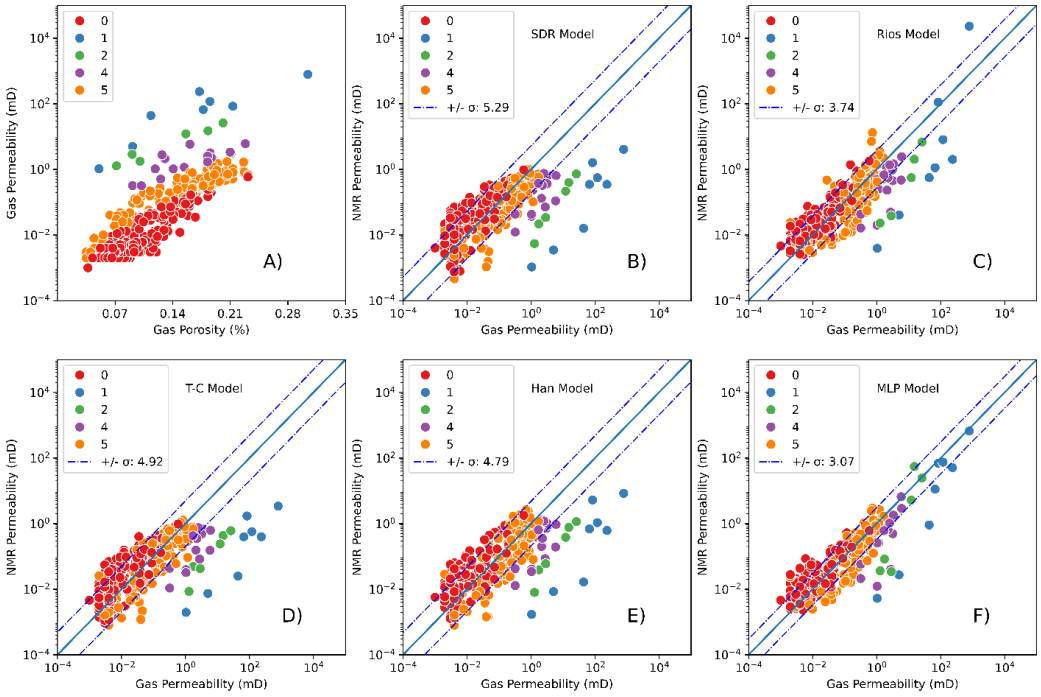

Figure 4 - Graphical representation of permeability prediction results, considering only the training dataset, for models A) SDR, B) Rios, C) Timur-Coates (T-C), D) Han, and E) MLP, in comparison with gas-derived data. The dotted lines indicate the σ error bounds. Source: Authors (2025)

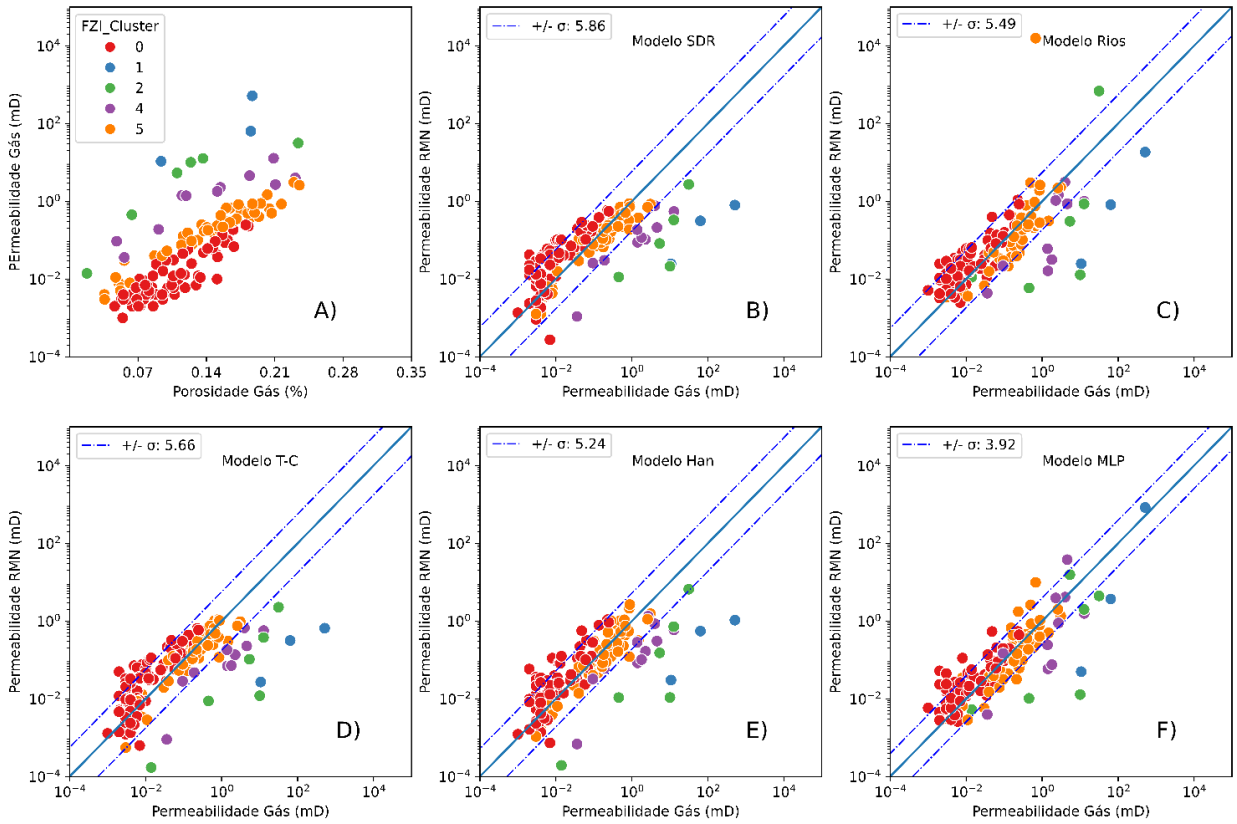

Figure 5 - Graphical representation of permeability prediction results, considering only the test dataset, for models A) SDR, B) Rios, C) Timur-Coates (T-C), D) Han, and E) MLP, in comparison with gas-derived data. The dotted lines indicate the σ error bounds. Source: Authors (2025)

The plots of the results obtained using the RQI and FZI indices, along with the HFU trend lines, are shown in Figures 6 and 7 for the training and test datasets, respectively. The models inherited the gas-based classification, and their estimates were calculated for each cluster. The HFU coefficients derived from the FZI and RQI indices, determined for each cluster based on the prior gas data classification for each model, are presented in Table 2. The separation of pore types using the sigmoidal curve (Equation 6) is illustrated in Figures 8 and 9, for the training and test datasets, respectively.

Table 2 – HFU coefficient data for each cluster in both datasets. T-C stands for Timur-Coates.

Clusters	***Training Data***						***Test Data***					
	Gas	**SDR**	**Rios**	**T-C**	**Han**	**MLP**	**Gas**	**SDR**	**Rios**	**T-C**	**Han**	**MLP**
0	0.10	0.19	0.17	0.19	0.18	0.17	0.10	0.21	0.19	0.20	0.19	0.20
1	3.46	0.23	3.05	0.25	0.32	1.79	4.33	0.25	0.63	0.24	0.30	3.14
2	1.48	0.23	0.55	0.26	0.31	1,18	1.61	0.23	1.43	0.25	0.31	1.13
4	0.58	0.20	0.27	0.21	0.24	0.36	0.65	0.21	0.29	0.22	0.24	0.68
5	0.24	0.21	0.25	0.22	0.23	0.24	0.24	0.21	0.98	0.22	0.22	0.25

Source: Authors (2025)

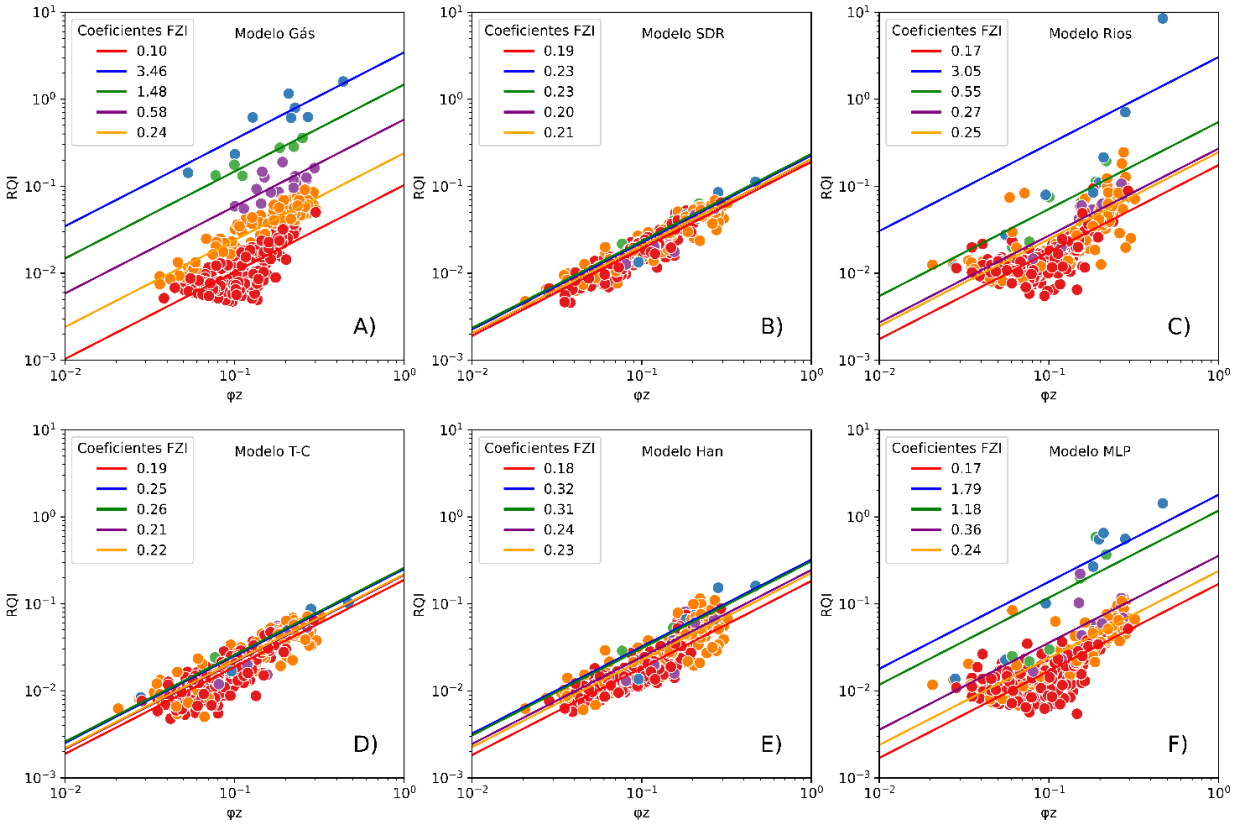

Figure 6 – Graphical representation of RQI and ϕ_z results, considering only the training dataset, for A) Gas data and models B) SDR, C) Rios, D) Timur-Coates, E) Han, and F) MLP. The lines indicate the HFU coefficients.
Source: Authors (2025)

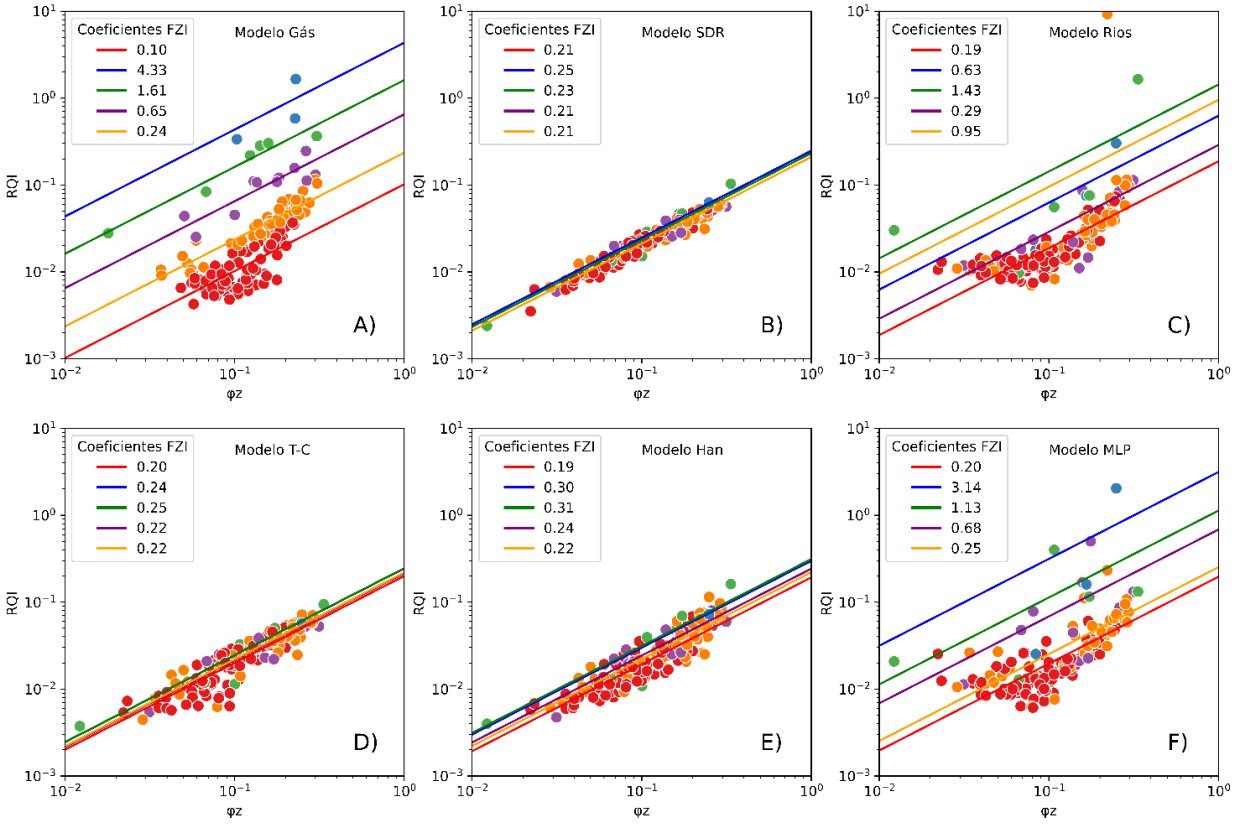

Figure 7 - Graphical representation of RQI and ϕ_z results, considering only the test dataset, for A) Gas data and models B) SDR, C) Rios, D) Timur-Coates, E) Han, and F) MLP. The lines indicate the HFU coefficients.
Source: Authors (2025)

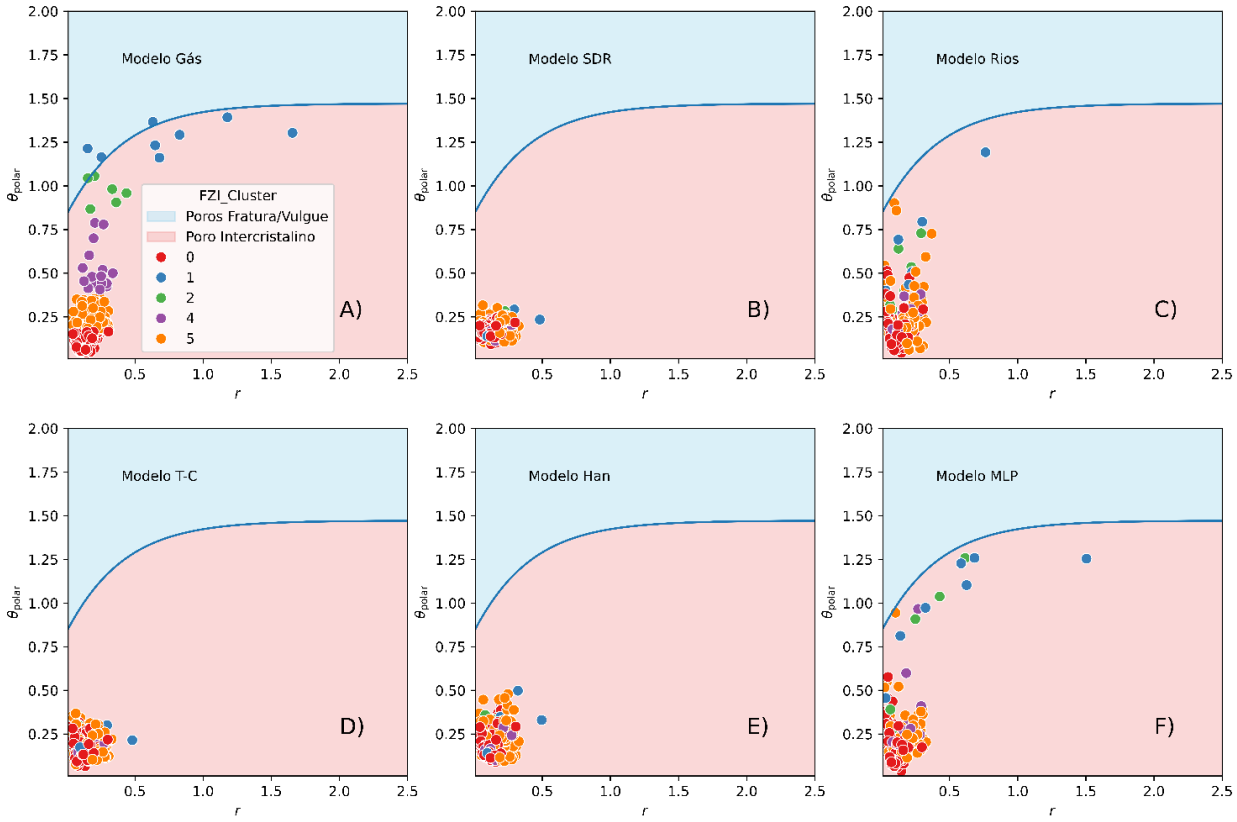

Figure 8 - Graphical representation of θ_z and r results, considering only the training dataset, for A) Gas data and models B) SDR, C) Rios, D) Timur-Coates, E) Han, and F) MLP. The line indicates the sigmoidal curve that separates pore types between fractures/vugs (above) and intercrystalline pores (below).
Source: Authors (2025)

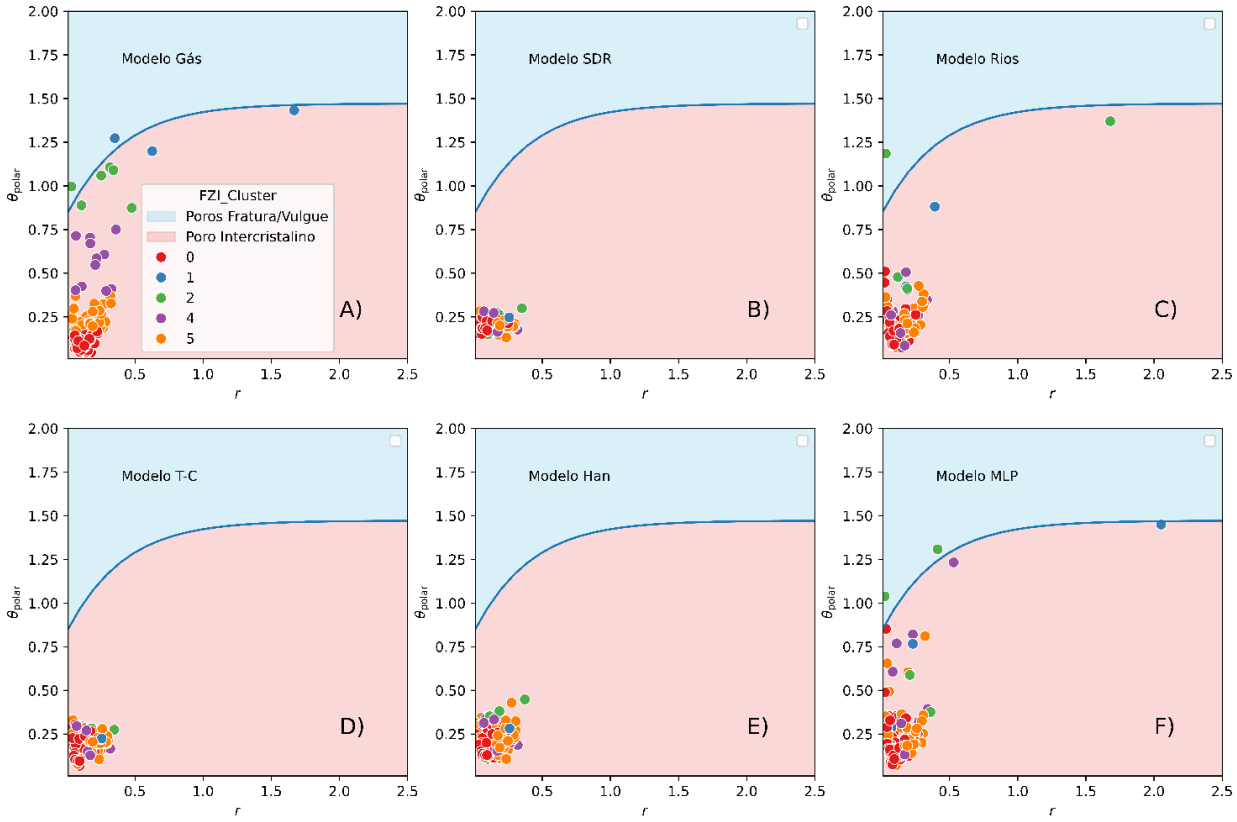

Figure 9 - Graphical representation of θ_{Polar} and r results, considering only the test dataset, for A) Gas data and models B) SDR, C) Rios, D) Timur-Coates, E) Han, and F) MLP. The line indicates the sigmoidal curve that separates pore types between fractures/vugs (above) and intercrystalline pores (below).
Source: Authors (2025)

4. Discussion

Our objective is to assess whether permeability prediction models based on NMR data can replicate the characteristics observed in laboratory measurements. The first step after obtaining the laboratory data is to classify the HFU using the K-Means method, which proved appropriate given that the study focuses on a single depositional environment with high complexity in its permoporous structures, as indicated by recent studies (ASTSAURI et al., 2024; EFTEKHARI et al., 2024). In contrast, Soto et al. (2010) used different reservoirs to separate the HFU, which allowed for a more homogeneous classification, whereas in our study (Figures 6 and 7), some HFU were less represented. The pore types, compared using the sigmoidal curve fitted by parameters r and θ_{Polar}, were also well adjusted by Soto et al. (2010) due to their choice of multiple environments; in our study, however, there was limited representation of "fracture/vug" pores based on the same sigmoidal curve. Another relevant point of this work is the computational cost of training a model, which, due to the relatively small dataset (354 samples) and limited input features (128), was not excessively high.

When evaluating permeability predictions, we observed that during training, the k_{Rios} and k_{MLP}models achieved the best accuracies, with R^2 values of 0.71 and 0.79, and σ errors of 3.74 and 3.07, respectively (Table 1). However, the k_{Rios} model tends to overestimate the maximum and average permeability values, while the k_{MLP} model provides values closer to those obtained from gas measurements. Additionally, the k_{MLP} performed well on the test set, with an R^2 of 0.72 and a σ error of 3.92, outperforming the other models, which showed an average R^2 of 0.55 and a σ error of 5.56. Although the σ error is higher compared to studies using NMR data (KENYON et al., 1988; SOUZA et al., 2013; CHEN et al., 2023),

the results are satisfactory, especially considering the high heterogeneity of the samples. These findings indicate that selecting limited parameters from the T_2 distribution curve is insufficient for accurate predictions in complex environments. Since the T_2 curve can be extensive, models such as artificial neural networks are able to capture subtle patterns and fine adjustments that conventional semi-empirical models cannot model with the same precision.

In the training dataset, only the MLP model was able to clearly separate these five zones, although it underestimated the coefficient for the blue zone (cluster 1), assigning a value of 1.79 instead of 3.46. The Rios model, which achieved the best accuracy among the semi-empirical models, separated four zones; however, clusters 4 and 5 had very similar values (0.27 and 0.25), making differentiation difficult. The other semi-empirical models showed unsatisfactory performance, with values practically invariant across zones, preventing the distinction of more than two HFUs. In the test dataset, the MLP again separated five zones with good precision, except for cluster 0, which exhibited a value approximately twice as high as expected. Although the Rios model separated five zones, it confused the cluster correspondence relative to the gas data, leading to significant errors in predictions with blind data. The remaining models also maintained unsatisfactory performance in this scenario. These results highlight the effectiveness of more sophisticated predictive models, such as the MLP, in interpreting the complexity of NMR data, significantly contributing to a more accurate reservoir characterization.

The close agreement of the MLP model with gas-derived results when evaluating the HFUs in both datasets demonstrated that the use of NMR data can significantly contribute to accurate reservoir characterization, especially when a robust model is applied for permeability prediction. This performance suggests that our dataset possesses sufficient complexity to effectively simulate challenging reservoirs, further reinforcing the potential of NMR techniques to substantially aid reservoir modeling and analysis when combined with advanced predictive approaches.

The application of the sigmoidal function for pore type separation was not very effective, as few samples fell above the curve, indicating a low correlation between permeability and cavities or fractures. Although some samples contained vugs and stylolites, factors such as poor connectivity of the vugs and recrystallization of stylolites reduce permeability (ARAÚJO et al., 2023), hindering efficient classification of these pores by the sigmoidal curve. In the gas training dataset, only four samples exhibited fracture/vug porosity, with three belonging to cluster 1 (blue) and one to cluster 2 (green). For the qualitative evaluation with the sigmoidal curve, the HFUs were well defined, with lower HFU values corresponding to lower porosity samples. However, no predictive model successfully distinguished the samples based on pore types. The MLP (Figure 8-F) and Rios (Figure 8-C) models showed partial approximation to the curve, with the MLP demonstrating better visual fitting, although both models mixed the HFUs. In the test dataset, laboratory results identified only two samples above the curve (Figure 9-A), one from each cluster. The MLP predicted two samples above the curve, both from cluster 2 (Figure 9-F), while the Rios model predicted one, also from cluster 2 (Figure 9-C). Other models continued to struggle to approximate the sigmoidal curve.

The T_2 relaxation time distribution data obtained by NMR can also be related to the size of the spin concentration and, consequently, to pore size (COATES et al., 1999; DUNN et al., 2002). Therefore, it would be expected that NMR delineates these ranges more accurately than gas-based data. However, pore type prediction is not derived directly from NMR data, but rather from permeability modeling. The conversion of T_2 time to pore size is not straightforward for all samples, meaning these variables are not necessarily correlated. As a result, semi-empirical models tend to oversimplify the NMR information, leading to characterizations that are excessively similar to each other in order to minimize errors. On the other hand, predictive models that employ more advanced statistical approaches, such as those proposed by Rios et al. (2011) and the MLP model, better capture the permeability characteristics. Thus, permeability characterization is more refined when performed with complex statistical models. In contrast, evaluating pore types for this reservoir requires a different perspective rather than relying solely on modeling or the T_2 distribution curve.

5. Conclusions

This study successfully achieved its objective by obtaining laboratory gas data and comparing them with various permeability prediction models, including three semi-empirical models, one statistical model, and one neural network-based model. The analysis was conducted in a complex geological setting, providing distinct reservoir indicators, indices, and classifications, thus allowing for a more comprehensive evaluation of the permeability predictions of these models. By addressing this challenging context, the study not only compared the accuracy of the models but also assessed their impact on reservoir characterization. Among the models analyzed, the MLP model, specifically trained for this study,

demonstrated the highest accuracy. This model outperformed the others, both quantitatively and qualitatively, and is widely used in the industry.

The results highlight a significant gap in the understanding and utilization of NMR data for permeability prediction. When these data are not correctly interpreted, they can lead to inaccurate reservoir characterizations, ultimately impacting decision-making in hydrocarbon exploration. Thus, the adoption of more sophisticated and robust models, such as the MLP, emerges as a promising alternative to improve permeability prediction, leading to more reliable geological interpretations. This not only enhances the understanding of reservoir characteristics but also assists the industry in optimizing resource exploration, increasing efficiency and reducing operational uncertainties.

It is possible that the sigmoidal curve used in this study may need to be adjusted to better represent the evaluated data, considering the specific characteristics of the samples and the relationship between porosity and permeability. Furthermore, a detailed characterization of pore types would be essential to enable more accurate quantitative assessment. An interesting approach would be to directly use the T_2 relaxation time distribution curve obtained by NMR to classify pore types, rather than relying exclusively on permeability models. This method could provide a more direct insight into pore structure and connectivity, allowing for a more detailed interpretation of rock properties.

Therefore, we offer the following suggestions for future complementary work: (1) use a larger sample set for fine adjustments; (2) evaluate pore types and adjust the sigmoidal curve to the studied data; (3) directly use NMR data to obtain HFUs and pore types.

Acknowledgments

This work was partially supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Graduate Program in Geodynamics and Geophysics (PPGG-UFRN).

The research was conducted in association with the ongoing R&D project registered as ANP 23505-1 “Porokarst Phase II – Processes and Characterization of fluid pathways in Karstified, Fractured and Silicified Reservoirs of the Presalt” (UFRN / UNB / UFPE / UFC / UFRA / IFRN / IFPB / Shell Brasil / ANP) sponsored by Shell Brasil under the ANP R&D levy program.

We also thank the High-Performance Computing Center (NPAD/UFRN) for providing relevant computational resources, and the Rock Physical Properties Laboratory (LPFR/UFRN) and the Petroleum Reservoir Engineering Laboratory (LABRES/UFRN) for collecting the petrophysical data.

Reference

AMAEFULE, J. O. *et al.* Enhanced Reservoir Description: Using Core and Log Data to Identify Hydraulic (Flow) Units and Predict Permeability in Uncored Intervals/Wells. All Days, 1993. doi: https://doi.org/10.2118/26436-ms

ARAÚJO, R. E. B. *et al.* Pore network characteristics as a function of diagenesis: Implications for epigenic karstification in shallow-water carbonates. Marine and petroleum geology, v. 149, p. 106094–106094, 2023. doi: https://doi.org/10.1016/j.marpetgeo.2022.106094

ASTSAURI, T. *et al.* Utilizing machine learning for flow zone indicators prediction and hydraulic flow unit classification. Scientific Reports, v. 14, n. 1, p. 4223, 2024. doi: https://doi.org/10.1038/s41598-024-54893-1

CANNON, S. Reservoir modelling : a practical guide. Hoboken, Nj: John Wiley & Sons, Inc, 2018.

CARR, H. Y.; PURCELL, E. M. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. Physical Review, v. 94, n. 3, p. 630–638, 1954. doi: https://doi.org/10.1103/physrev.94.630

CHEN, X. *et al.* NMR-data-driven prediction of matrix permeability in sandstone aquifers. Journal of Hydrology, v. 618, p. 129147–129147, 2023. doi: https://doi.org/10.1016/j.jhydrol.2023.129147

CHEN, S. *et al.* Study of Uncertainties in the Inversion Algorithms for Transverse Relaxation Distribution. 4th International Conference on Bioinformatics and Biomedical Engineering, 2010. doi: https://doi.org/10.1109/icbbe.2010.5514753

COATES, G. R. *et al*. NMR Logging Principles and Applications. Gulf Professional. 1999.

CÓRDOBA, V. C. A evolução da plataforma carbonática Jandaíra durante o Neocretáceo na Bacia Potiguar: Análise paleoambiental, diagenética e estratigráfica. 2001. PhD Thesis Universidade Estadual Paulista (UNESP), 2001.

DE CASTRO, D. L. *et al.* Influence of Neoproterozoic tectonic fabric on the origin of the Potiguar Basin, northeastern Brazil and its links with West Africa based on gravity and magnetic data. Journal of Geodynamics, v. 54, p. 29–42, 2012. doi: https://doi.org/10.1016/j.jog.2011.09.002

DUNN, K.-J. *et al*. Nuclear Magnetic Resonance. Elsevier, 2002.

EFTEKHARI, S. H. *et al.* Hydraulic flow unit and rock types of the Asmari Formation, an application of flow zone index and fuzzy C-means clustering methods. Scientific reports, v. 14, n. 1, 2024. doi: https://doi.org/10.1038/s41598-024-55741-y

GARIA, S. *et al.* A multivariate statistical approach in correlating the acoustic properties with petrophysics and mineralogy on sandstones. Geophysical Journal International, v. 230, n. 1, p. 160–178, 2022. doi: https://doi.org/10.1093/gji/ggac061

GAVIDIA, J. C. R. *et al.* Bridging the gap: Integrating static and dynamic data for improved permeability modeling and super k zone detection in vuggy reservoirs. Geoenergy Science and Engineering, v. 241, p. 213152–213152, 2024. doi: https://doi.org/10.1016/j.geoen.2024.213152

GELADI, P.; KOWALSKI, B. R. Partial least-squares regression: a tutorial. Analytica Chimica Acta, v. 185, p. 1–17, 1986. doi: https://doi.org/10.1016/0003-2670(86)80028-9

GÉRON, A. Mãos à Obra Aprendizado de Máquina com Scikit-Learn e TensorFlow. Alta Books, 2021.

HAHN, E. L. Spin Echoes. Physical Review, v. 80, n. 4, p. 580–594, 1950. doi: https://doi.org/10.1103/physrev.80.580

HAN, Y. *et al.* A new permeability calculation method using nuclear magnetic resonance logging based on pore sizes: A case study of bioclastic limestone reservoirs in the A oilfield of the Mid-East. Petroleum exploration and development, v. 45, n. 1, p. 183–192, 2018. doi: https://doi.org/10.1016/s1876-3804(18)30019-3

HINTON, G. E. *et al*. A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, v. 18, n. 7, p. 1527–1554, 2006. doi: https://doi.org/10.1162/neco.2006.18.7.1527

KENYON, W. E. *et al.* A Three-Part Study of NMR Longitudinal Relaxation Properties of Water-Saturated Sandstones. SPE Formation Evaluation, v. 3, n. 03, p. 622–636, 1988. doi: https://doi.org/10.2118/15643-pa

KLINKENBERG, L. The Permeability of Porous Media to Liquids and Gases. Drilling and Production Practice, American Petroleum Inst., pp. 200–213. 1941

LOPES, J. A. G. *et al.* Three-dimensional characterization of karstic dissolution zones, fracture networks, and lithostratigraphic interfaces using GPR cubes, core logs, and petrophysics: Implications for thief zones development in carbonate reservoirs. Marine and petroleum geology, v. 150, p. 106126–106126, 2023. doi: https://doi.org/10.1016/j.marpetgeo.2023.106126

LUCIA, F. J. *et al*. Carbonate Reservoir Characterization. v. 55, n. 06, p. 70–72, 2003. doi: https://doi.org/10.2118/82071-jpt

MATOS, R. M. D. de. The Northeast Brazilian Rift System. Tectonics, v. 11, n. 4, p. 766–791, 1992. doi: https://doi.org/10.1029/91tc03092

MEHMOOD, T. *et al.* A review of variable selection methods in Partial Least Squares Regression. Chemometrics and Intelligent Laboratory Systems, v. 118, p. 62–69, 2012. doi: https://doi.org/10.1016/j.chemolab.2012.07.010

MEIBOOM, S.; GILL, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. Review of Scientific Instruments, v. 29, n. 8, p. 688–691, 1958. doi: https://doi.org/10.1063/1.1716296

PEDREGOSA, F. *et al.* Scikit-learn. GetMobile: Mobile Computing and Communications, v. 19, n. 1, p. 29–33, 2015. doi: https://doi.org/10.1145/2786984.2786995

RIOS, E. H. *et al.* Modeling rock permeability from NMR relaxation data by PLS regression. Journal of Applied Geophysics, v. 75, n. 4, p. 631–637, 2011. doi: https://doi.org/10.1016/j.jappgeo.2011.09.022

SOTO, R. B. *et al.* Pore-Type Determination From Core Data Using a New Polar-Transformation Function from Hydraulic Flow Units. All Days, 2010. doi: https://doi.org/10.2118/136805-ms

SOUZA, A. *et al.* Permeability prediction improvement using 2D NMR diffusion-t2 maps. *In*: SPWLA 54TH ANNUAL LOGGING SYMPOSIUM, 2013, New Orleans, Louisiana. Society of Petrophysicists and Well Log Analysts. New Orleans, Louisiana: Society of Petrophysicists and Well Log Analysts, 2013.

TIAB, D.; DONALDSON, E. C. Petrophysics : theory and practice of measuring reservoir rock and fluid transport properties. Amsterdam; Boston: Elsevier/Gpp, Gulf Professional Publishing Is An Imprint Of Elsevier, 2016.

TIMUR, A. Pulsed Nuclear Magnetic Resonance Studies of Porosity, Movable Fluid, and Permeability of Sandstones. Journal of Petroleum Technology, v. 21, n. 06, p. 775–786, 1969. doi: https://doi.org/10.2118/2045-pa

WEI, H. *et al.* Nuclear Magnetic Resonance T_2 Distribution-Based Gas–Water Relative Permeability Prediction in Tight Sandstone Reservoirs: A Case Study on Central Sichuan Basin, China. Energy & Fuels, v. 38, n. 5, p. 3598–3608, 2024. doi: https://doi.org/10.1021/acs.energyfuels.3c03972

WOLPERT, D. H. The Lack of A Priori Distinctions Between Learning Algorithms. Neural Computation, v. 8, n. 7, p. 1341–1390, 1996. doi: https://doi.org/10.1162/neco.1996.8.7.1341