PERMEABILITY ESTIMATES OF A CARBONATE RESERVOIR IN CAMPOS BASIN, SOUTHEASTERN BRAZIL, USING WELL LOGS WITH EMPIRICAL, MULTILINEAR REGRESSION AND MACHINE LEARNING APPROACHES

Abel Carrasquilla ⓘD and Rhanderson Gomes ⓘD

Universidade Estadual do Norte Fluminense - UENF, Petroleum Engineering and Exploration Laboratory - LENEP, Imboacica, Macaé, RJ, Brazil

*Corresponding author email: abel@lenep.uenf.br

ABSTRACT. Well logging records the physical properties of geological formations and the fluids traversed by the wells. This operation is interested in parameters such as lithology, hydrocarbon presence, permeability, porosity, and fluid saturation. Generally, oil reservoirs are sandstone or carbonate rocks, and the latter characterization is a critical question in the petrophysical property distribution, mainly permeability. Estimating permeability is a complex task due to the heterogeneity of these reservoirs. Therefore, this work used conventional logs to estimate the permeability of wells A03 and A10, both belonging to the oilfield A, Campos Basin, Southeastern Brazil. Together with the logs, the permeability measured in the laboratory in rock samples was used to validate the achieved estimates. Thus, the estimates used basic logs as input and approaches such as Timur empirical equation, multilinear regression, and machine learning techniques, like fuzzy logic, neural network, and decision tree. The Pearson's coefficient of determination R was used as the comparison metric with the experimental data. The number of samples in training was 70%, with 15% in the validation and testing; the results show that the first four estimates presented bad fits (R≤0.60), while the decision tree showed good fits (R>0.60). This approach also showed that the gamma-ray and resistivity logs are the ones that have the most significant weight in the estimates.

Keywords: carbonate reservoir; permeability, well logging; multilinear regression; machine; learning approaches.

INTRODUCTION

Well logging is an operation to record the physical characteristics of the geological formations, which are used to assess the petrophysical properties of the reservoirs, such as porosity, permeability, lithology and water saturation (Luthi and Bryant, 1997). Permeability estimate is a critical question in carbonate reservoirs because they have many heterogeneities, which makes interpretation difficult (Tiab and Donaldson, 2004).

Permeability is a very complex parameter, and its value depends on many factors, including lithology, texture, cementation, etc. It can be evaluated by different methods, which have limitations: formation tests, laboratory core measurements, empirical estimates with logs, and statistical approaches. The formation tests provide average permeability values; however, they assume homogeneity in the tested zone. Direct measurements are done on cores and can be affected by factors such as the selection and representativeness of samples taken on cores, depth mismatch, and accuracy of laboratory measurements. In addition, there are differences between the measuring volumes in rock samples (centimeters) and well-logging curves (meters) (Schön, 2015). The empirical relationships of Timur and Kozeny-Carman are widely used, which provide estimated values from, in general, porosity and irreducible water saturation data. Statistical methods allow estimates using the available logs by consulting the database generated with the records of key wells and permeability data obtained from cores of these same wells. The reliability of

these statistical methods depends on the existence of a reasonable number of good-quality samples analyzed in the laboratory. All data must be perfectly correlated in depth with the logs, and the gamma-ray log recorded in the core can be used to compare it with the gamma-ray log recorded in the well. Geophysical well logs do not directly measure permeability, so it is almost always an approximate estimation rather than a parameter produced in a test (Yan, 2002).

On the other hand, machine learning evolved from studying pattern recognition and computational learning theory to artificial intelligence (Theodoridis and Koutroumbas, 2009). It can be defined as a field of study that allows computers to learn without being explicitly programmed (Cabena et al., 1998). It is about studying and constructing algorithms that can learn from their mistakes and make predictions about data. Such algorithms build a model from sample inputs to make data-driven predictions or decisions rather than simply follow inflexible and static programmed instructions (Silva et al., 2020). While in artificial intelligence, there are two types of reasoning (inductive, which extracts rules and patterns from large data sets, and deductive), machine learning is only concerned with inductive reasoning. Some parts of machine learning are closely linked (and often overlapping) with computational statistics, a discipline that focuses on making predictions using computers, with research focusing on the properties of statistical methods and their computational complexity. It has strong ties to mathematical optimization, which produces methods, theory, and application domains for this field (Mitchell, 1997).

The working script was established as follows: it started with the Geological Context to present the object of study. It continued with Materials and Methods to show how the permeability assessment in this reservoir was performed. Then, the outcomes were presented in Results and Discussions, looking for a deeper understanding of the applied concepts. Finally, the study's main results were highlighted in Conclusions and the consulted works in References.

GEOLOGICAL CONTEXT

According to Okubo et al. (2015), the origin and evolution of the Campos Basin are related to the separation of the supercontinent Gondwana (Figure 1). They are marked by the disaggregation of the plates of South America and Africa with the consequent formation of the South Atlantic Ocean. The tectonic-sedimentary evolution of this basin occurred in three phases: rift, post-rift, and

drift, which correspond, respectively, to the continental, transitional and marine sequences. The continental was deposited during the mechanical subsidence of the rift phase and includes basalts from the Cabiunas Formation and continental sediments from the Lagoa Feia Formation. The transition sequence is characterized by evaporites from the Retiro Formation deposited in shallow marine transgression pulses over continental areas and relative tectonic quiescence (Bruhn et al., 2003).

The marine sequence marks the beginning of the open marine deposition during the thermal subsidence associated with the drift phase. This stage begins with the carbonate sedimentation (Macaé Group) and grades to a mainly siliciclastic succession (Campos Group) affected by intense halokinesis. The Macaé Group sedimentation occurred during the Albian Age in an early drift context and comprised the Goitacás, Quissamã, Outeiro, Imbetiba, and Namorado formations. The Quissamã Formation is formed by grainstones and packstones, constituted by oncoids, ooids, peloids, and diverse bioclasts, associated with NE shoals deposited in high to moderate-energy environments. The Outeiro Formation is composed of fine carbonate layers, interbedded with marl and shale deposited in response to a gradual rise in sea level and the drowning of the shallow carbonate platform of the Quissamã Formation (Figure 2). These carbonate rocks have abundant pelagic microfossils, like calcispheres (pitonellids), planktonic foraminifera, and radiolarians (Guardado et al., 1989).

MATERIALS AND METHODS

Oilfield A has 27 wells with the complete set of well logs, but only two wells have the petrophysical laboratory tests (Table 1). Thus, wells A03 and A10 were used in this study to estimate the permeability of the carbonate reservoir in this basin. The complete dataset consists of the basic suite of well logs and laboratory measurements of porosity and permeability (Table 2). Figures 3 and 4, on the other hand, show, respectively, the logs for both wells. The logs are identified as gamma-ray (GR), neutron porosity (NPHI), density (RHOB), sonic (DT), and shallow (RXO) and deep (RT) resistivities, which appear, respectively, between tracks 4 and 7 of these two figures. The study involves, thus, log analysis to determine the permeability in these wells through various approaches, constantly comparing the quality of estimates with the gas permeability measured in the laboratory (Petrobras, 2012).
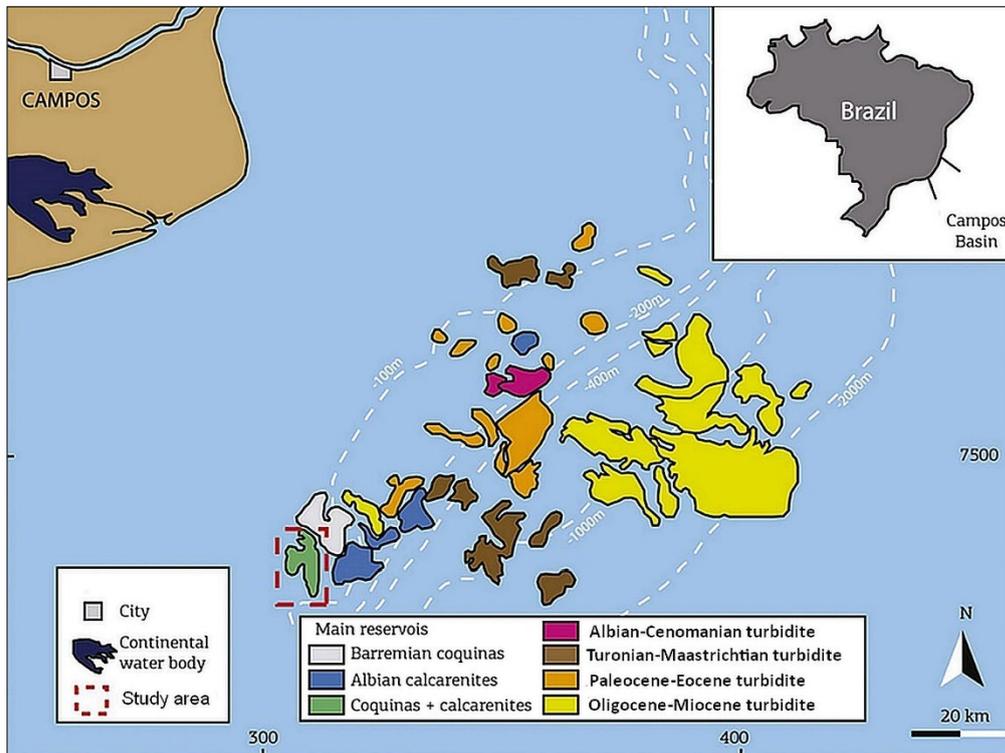
Figure 1: The location of Albian carbonate reservoirs in the Campos Basin offshore, Brazil. The Albian calcarenite reservoirs are indicated by the dashed red line (modified from Bruhn et al., 2003).
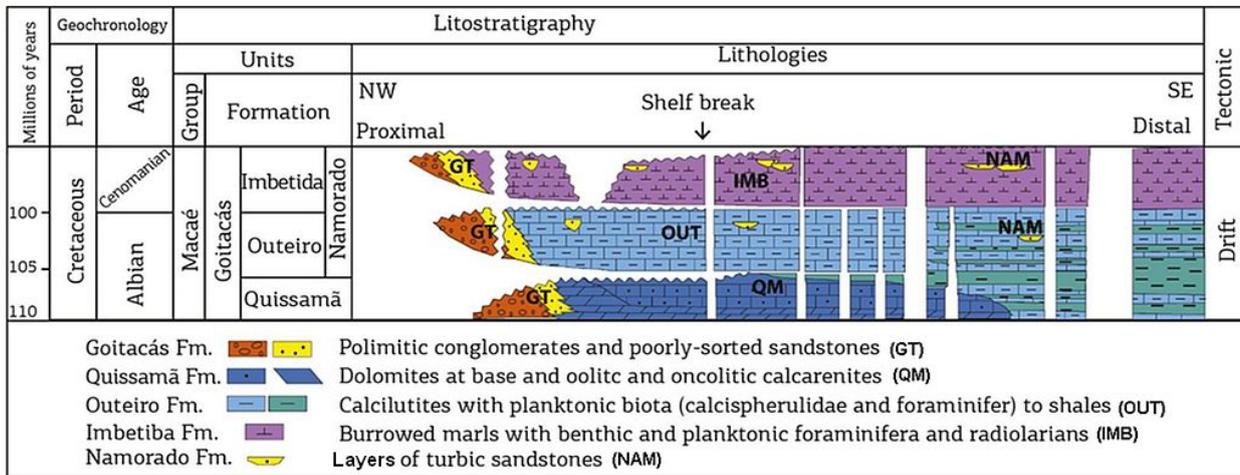
Figure 2: Stratigraphic chart of the Albian section in the Campos Basin (adapted from Winter et al., 2007).

The Interactive Petrophysics-IP (Geoactive, 2021) was used to plot and interpret, preliminary, the logs. Next, the Waikato Environment for Knowledge Analysis-Weka (Holmes et al., 1994) and Matlab (2018) were used to construct the Timur´s empirical model, the multilinear regression approach (MLR) and the machine learning simulations, as fuzzy logic (FL), neural network (NN) and the decision tree (DTree) (Russell and Norvig, 2009). Finally, Microsoft Excel was used to make the statistical analysis (Harnett and Horrell, 1998). These permeability estimates were then compared with each other, and their qualities were evaluated using the Pearson's determination coefficient-R (Cameron and Windmeijer, 1997) when matched in the adjustment with experimental laboratory data, where R is a measure of linear correlation between original and predicted values (Dell'Aversana, 2019). The workflow adopted in this work is presented in Figure 5.

Table 1: Characteristics of the carbonate reservoir ([Petrobras, 2012](#)).

Reservoir	Oilfield A
Lithology	Carbonate
Age	Albian
Area (km²)	11.6
Mean Net Pay (m)	105
Mean Porosity (%)	19
Mean Permeability (mD)	50
Mean Water Saturation (%)	22
Formation Water Salinity (ppm)	70000
Temperature (°C)	90

Table 2: Available data set of the oilfield A ([Petrobras, 2012](#)).

Logs	Curve Nemonics	Units	Wells	
			A3	**A10**
Sonic	DT	µs/ft	x	x
Density	RHOB	gr/cm³	x	x
Neutron	NPHI	%	x	x
Gamma Ray	GR	°API	x	x
Deep resistivity	RT	ohm.m	x	x
Shallow Resistivity	RXO	ohm.m	x	x
Laboratory porosity	φ	%	x	x
Laboratory permeability	k	mD	x	x

Initially, the Timur's empirical model was applied to estimate the permeability using the following equation ([Timur, 1968](#)):

$$k_{Timur} = \frac{a\phi^b}{S_{wirr}^c} \qquad (1)$$

where k_{Timur} is the permeability (mD); a, b and c are regression coefficients; φ is the porosity (%); and S_{wirr} is the irreducible water saturation (%). Generally, the used values for a, b and c are 0.136, 4.4, and 2, respectively.

Next, MLR was applied as a mathematical method that uses several explanatory variables to predict the outcome of a response variable. MLR aims to model the linear relationship between the independent variables (GR, NPHI, RHOB, DT, RXO, and RT logs) and the dependent variable (k_{MLR}). MLR is the extension of conventional least-squares regressions because it involves more than one explanatory variable. In the case of our study, permeability was estimated by entering the logs in the following equation ([Seber and Lee, 2003](#)):

$$k_{MLR} = c_1 GR + c_2 RT + c_3 RXO + c_4 NPHI + c_5 RHOB + c_6 DT + \varepsilon, \qquad (2)$$

where k_{MLR} is the permeability (mD); c_1 to c_6 are the regression coefficients; and ε is the error in the estimate.

FL is a form of many-valued logic in which the truth value of variables may be any actual number between 0 and 1. It handles partial truth, where the truth value may be entirely true or false. By contrast, in Boolean logic, the truth values of variables may only
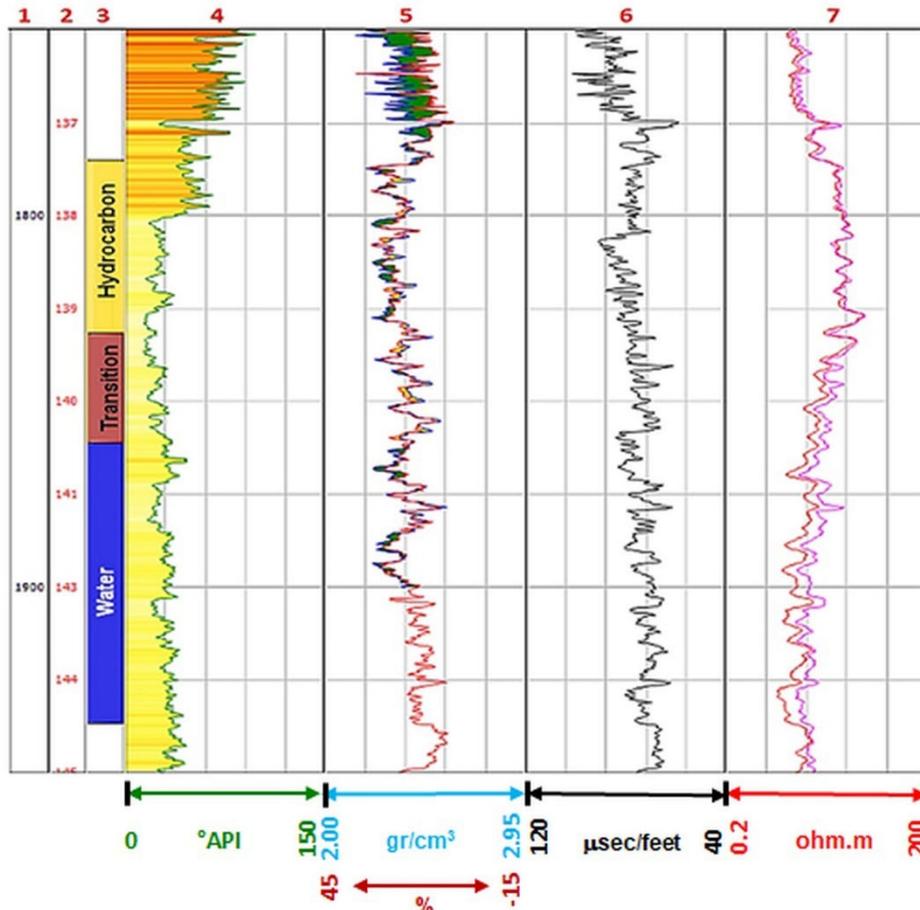
Figure 3: Well A03, tracks: 1) Depth (m); 2) Temperature (°F); 3) Fluids; 4) GR log; 5) NPHI (%, blue curve) and RHOB (gr/cm³, red curve) logs; 6) DT log (µsec/feet, black curve); 7) Shallow RXO (ohm.m, pink curve) and deep RT (ohm.m, red curve) resistivity logs, on logarithmic scale.

be the integer values 0 or 1. FL is based on the observation that people make decisions based on imprecise and non-numerical information. Fuzzy models or sets are mathematical means of representing vagueness and inaccurate information, hence the term fuzzy (Cuddy, 1997). These models can recognize, describe, manipulate, interpret, and utilize data and information that are vague and lack certainty. FL has been applied to many fields, from control theory to artificial intelligence. Fuzzy sets are often defined as triangle, trapezoid, or sigmoid-shaped curves, as each value will have a slope where the value is increasing, a peak where the value is equal to 1 (which can have a length of 0 or greater), and a slope where the value is decreasing. One typical case is the standard logistic function (S_x) defined as (Nordlund, 1996):

$$S(x) = \frac{1}{1+e^{-x}} , \qquad (3)$$

which has the following symmetry property:

$$S(x) + S(-x) = 1. \qquad (4)$$

From this, it follows that:

$$\begin{aligned}(S(x) + S(-x)).(S(y) \\ + S(-y)).(S(z) \qquad (5)\\ + S(-z)) = 1\end{aligned}$$

NN is a network or circuit of neurons, or in a modern sense, an artificial neural network composed of artificial neurons or nodes for solving artificial intelligence problems. The connections of the biological neuron are modeled in artificial neural networks as
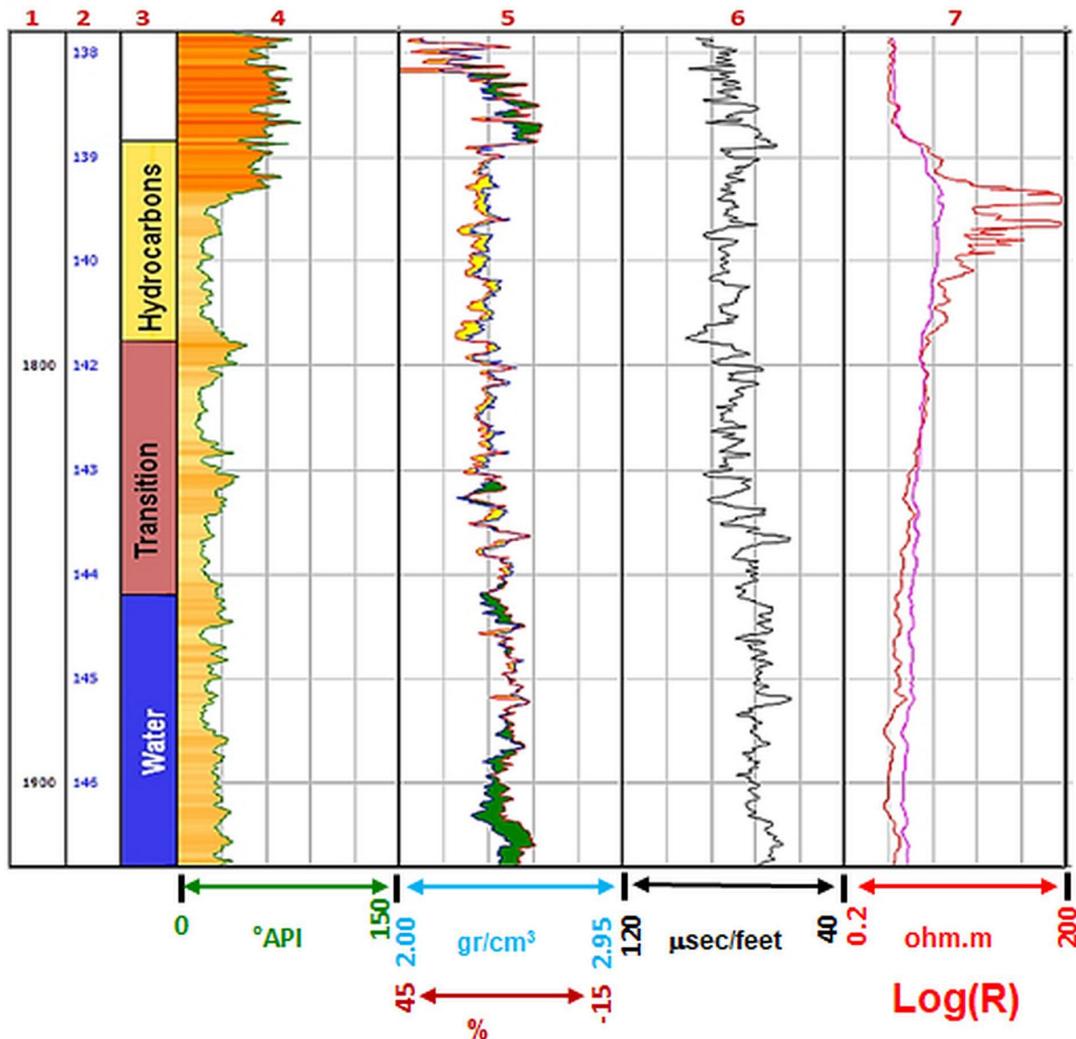
Figure 4: Well A10, tracks: 1) Depth (m); 2) Temperature (°F); 3) Fluids; 4) GR log; 5) NPHI (%, blue curve) and RHOB (gr/cm³, red curve) logs; 6) DT log (µsec/feet, black curve); 7) Shallow RXO (ohm.m, pink curve) and deep RT (ohm.m, red curve) resistivity logs, on logarithmic scale.

weights between nodes (Mohaghegh et al., 1995). A positive weight reflects an excitatory connection, while negative values mean inhibitory connections. All inputs are modified by weight and summed. This activity is referred to as a linear combination. Finally, an activation function controls the amplitude of the output. For example, an acceptable output range is usually between 0 and 1, or it could be −1 and 1 (Ahmadi et al., 2012). Artificial networks may be used for predictive modeling, adaptive control, and applications where they can be trained via a dataset. Self-learning resulting from experience can occur within networks, which can derive conclusions from a complex and seemingly unrelated set of information. As neuron activation, the NN models used, as transfer functions, a hyperbolic tangent function (equation 6) in the neurons of the hidden layer and a linear function (equation 7) in the neurons of the output layer (Silva et al., 2015):

$$\tanh(s) = \frac{2}{(1 + e^{-2s})} - 1, \tag{6}$$

where s represents the sum calculated for each neuron in each propagation. The calculation of a neuron of the hidden layer can be represented by the following function:

$$net_i = w_{0i} + \sum_{j=1}^{n} w_{ij}x_j, \tag{7}$$

where x_j are the input variables; w_{ij} are the weight factors; n is the number of hidden layer input neurons; and w_{0i} is the bias of each hidden layer neuron.
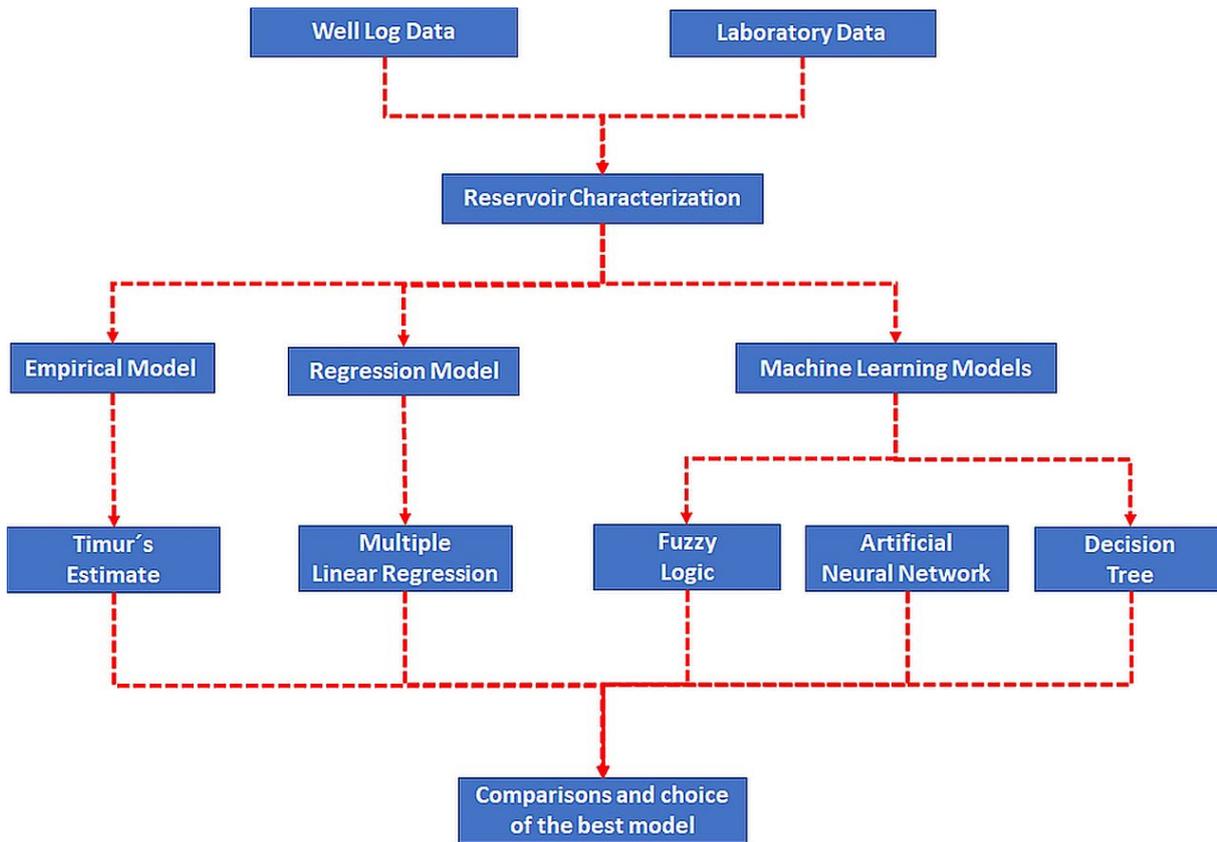
Figure 5: Workflow followed in the study.

of the output layer, the function of calculating the permeability of propagation can be represented as follows:

$$k = f\left[v_0 + \sum_{i=1}^{m} v_i f_i \left(\frac{2}{1 + e^{-2net_i}}\right)\right], \qquad (8)$$

DTree is a flowchart-like structure in which each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label. The paths from the root to the leaf represent classification rules. A DTree and the closely related influence diagram are used in decision analysis as a visual and analytical decision support tool, where the expected values of competing alternatives are calculated. A DTree consists of three types of nodes: a) decision nodes, b) chance nodes, and c) end nodes (Quinlan, 1987).

RESULTS AND DISCUSSIONS

For well A03, Figure 3 shows, from left to right, the basic set of logs: Depth (m); GR (API); NPHI (%); RHOB (gr/cm³); DT (μsec/feet); and shallow (RXO, ohm.m) and deep resistivity (RT, ohm.m). In track 4, the GR log drops at 1775 m depth and then there is a sharp drop at 1800 m depth, from around 50ºAPI to approximately 30ºAPI. This sudden change indicates a change in the type of lithology, as muddy rocks have high GR values and carbonate reservoir rock does not, suggesting a capping rock between these depths. On track 5, with the RHOB and NPHI logs, it is possible to see the intersection of the curves around 1785 m, indicating the presence of hydrocarbons (yellow shadow). The DT log in track 6 shows a drop in 1775 m in the transit time, indicating the presence of more porous rock, in this case, the reservoir rock, showing a characteristic carbonate heterogeneity. In track 7, the RT log grows in 1785 m, indicating the presence of oil, as oil has high resistivity. The RXO log is close to the RT log, showing a small invasion (low permeability) and possibly indicating that the drilling fluid is oil-based. After this, a distance between the resistivity curves suggests the presence of water and a more extensive invasion (higher permeability), the beginning of the transition zone around 1830 m. Therefore, making a cross-analysis of the values presented by each log, the hydrocarbon zone was delimited from 1785 to 1829 m, the transition zone between 1829 and 1850 m, and the water zone below 1850 m.

Figure 4 shows, from left to right, the basic set of logs for well A10. Analogously, the GR log presents an abrupt drop around 1750 m depth (track 4), indicating the beginning of the reservoir. On track 5, with the RHOB and NPHI logs, it is possible to see the intersection of the curves around 1750 m depth, indicating the presence of hydrocarbons in this region (yellow shadow). The DT log shows in track 5 a drop in 1745 m depth, indicating the presence of more porous rock, in this case, the reservoir rock. The sharp increase in RT between 1750 and 1775 m depth in track 7, in addition to the fact that the drilling fluid is water-based, also indicates the presence of a concentration of hydrocarbons. The separation between the RXO and RT curves in well A10 means more significant invasion and better permeability than in well A03. Therefore, making a cross-analysis of the values presented by each log, the hydrocarbon zone was delimited from 1746 to 1794 m depth, the transition zone from 1794 to 1855 m depth, and from 1855 m depth onwards, the aquifer.

Figures 6A and 6B present the laboratory permeability together with the permeability estimated by Timur's (1968) equation, where the red dots are the estimates (k_{Timur}), and the blue dots are the laboratory data (k_{lab}). In this, it is possible to observe a weak correlation in both wells, having R is 0.12 and -0.08 for wells A03 and A10, respectively (Table 3). This represents a solid reason to look for other types of estimates that work better.

In the case of estimating the permeability of the two wells with MLR, the DT, GR, NPHI, RT, and RHOB logs were used as input data to create the regression model. The number of samples for wells A03 and A10 is 135 and 263, respectively (Table 4). The generic equation created by the MLR has the form of Equation 2, and the value of the coefficients c_0 to c_6. The error ε represents the adjustment error, which was not calculated and replaced by R to analyze the goodness of fit. It is also observed that c_6, the coefficient for the DT log, is the most significant (9.47 and -40.58), which can mean the presence of secondary porosity for which this log is more sensitive, such as fractures or vugs. From this, the two best estimates for both wells are shown in Figures 7A and 7B, with the red being the estimates (k_{MLR}) and the blue dots being the laboratory data (k_{lab}), having R is 0.36 and 0.50 for wells A03 and A10, respectively (Table 3).

The same logs as the MLR estimate were used to estimate the permeability with FL. Figures 8A and 8B show these estimates for wells A03 and A10, respectively,

where the red dots are the estimates (k_{FL}), and the blue dots are laboratory permeability measurements (k_{lab}). Table 3 shows R=0.28 and 0.29 for wells A03 and A10, respectively, which is considered a wrong estimate. The choice of membership functions and the rules of this approach make it difficult to apply the FL technique.

The permeability estimation with NN used the Random Bayesian Regularization, utilizing the same input logs for MLR and FL estimates. For both wells, the number of iterations was 1000, with 100 neurons in the hidden layer and only one in the output layer (*k*). For well A03, Figure 9A shows the estimate (red dots, k_{NN}) and laboratory permeability measurements (blue dots, k_{lab}). The number of samples in training was 202 (70%), with 44 (15%) in the validation and testing stages, and the results show an R=0.99 (Figure 9B) in the training stage, 0.13 in the test (Figure 9C) and 0.50 for all (Figure 9D). For well A10, Figure 10A shows the estimate (red dots, k_{NN}) and laboratory permeability measurements (blue dots, k_{lab}). The number of samples in training was 333 (60%), with 71 (20%) in the validation and testing stages, and the results show an R=0.85 (Figure 10A) in the training stage, 0.46 in the test (Figure10B) and 0.76 for all (Figure10C). It is essential to point out that if the percentage of samples is decreased in the training step to increase the test and all stages, this reduces the coefficient R between the estimate and the laboratory data. Figures 9A and 10A show the bad fit after this great effort, this good result is attributed to the grain contrast and very low values in the laboratory permeability values (Table 3). The better result in well A10 is likely due to more laboratory data than in well A03.

The DTree approach results indicate R=0.71 and 0.80 for wells A03 and A10, respectively (Table 3). The percentage of samples used for this estimate at the training, test, and all steps is the same as that used with the NN estimate. Figure 11A shows the results of Dtrees estimation (red dots, k_{DTREE}) and laboratory permeability measurements (blue dots, k_{lab}) for well A03 and Figure 11B for well A10, which improves on the previous estimate. Figures 11C and 11D show that the GR log is the most critical factor in the two calculations, followed by RT and RHOB for both wells. In the case of NPHI and DT logs, DT is more important than NPHI for well A03 (Figure 11C). But in the case of well A10, the exact opposite happens, NPHI has more weight than DT (Figure11D). The difference between Figures 11C and 11D can be explained because these are different wells with similar but not identical
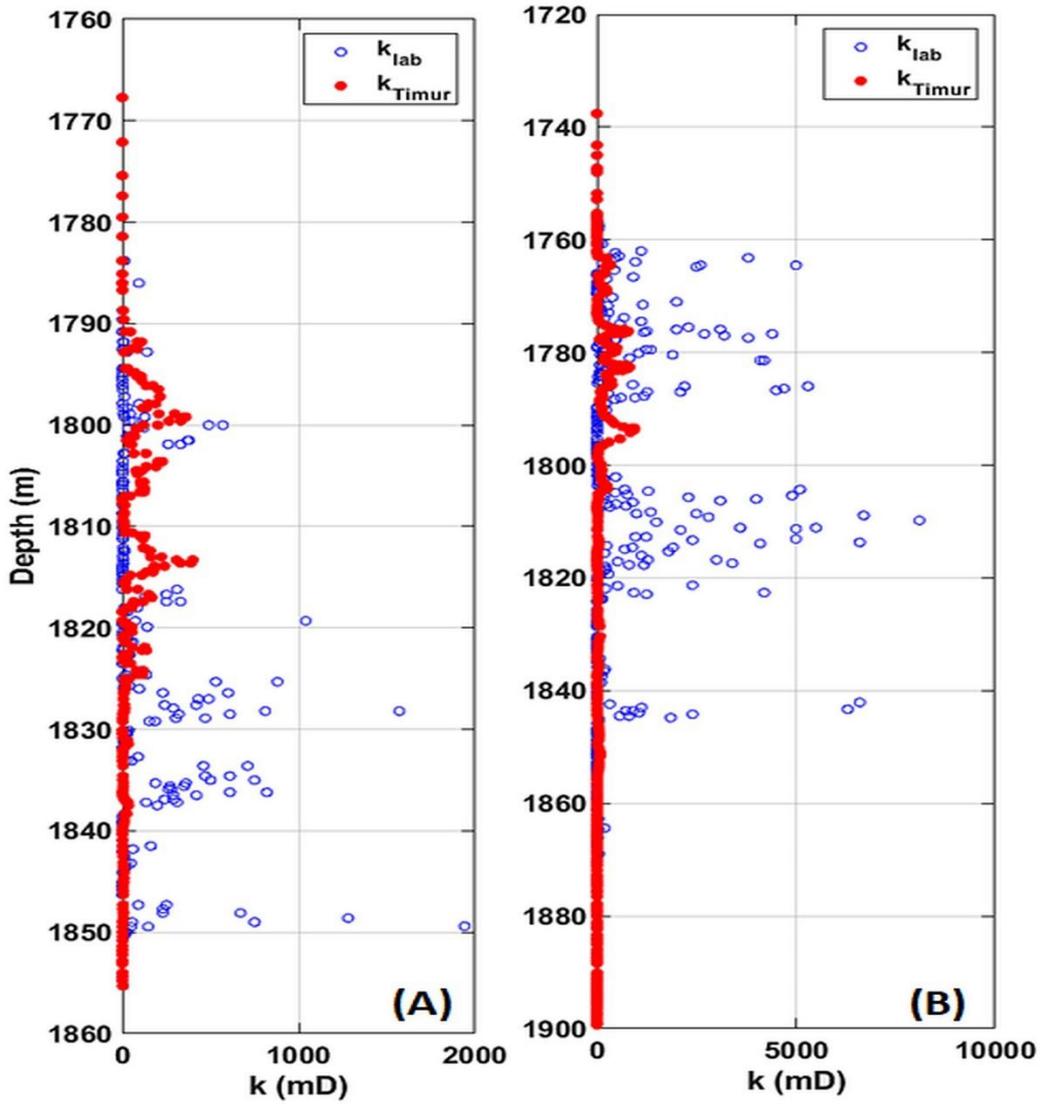
Figure 6: Timur's permeability estimates (red dots, k_{Timur}) and laboratory permeability measurements (blue dots, k_{lab}) for wells A03 (A) and A10 (B).

Table 3: Wells A03 and A10 permeability estimates for the different algorithms, showing the values for R, the Pearson's Determination Coefficient.

Wells	Estimates				
	Timur	Multilinear Regression	Fuzzy Logic	Neural Network	Decision Tree
A03	0.12	0.36	0.28	0.50	0.71
A10	-0.09	0.50	0.29	0.76	0.80

Table 4: Multilinear Regression Coefficients for the estimates of Wells A03 and A10.

Wells	No. Points	c_1	c_2	c_3	c_4	c_5	c_6
A03	135	8.32	-5.07	-2.73	0.40	2.23	9.47
A10	263	25.56	0.93	-6.30	-2.00	0.34	-40.58
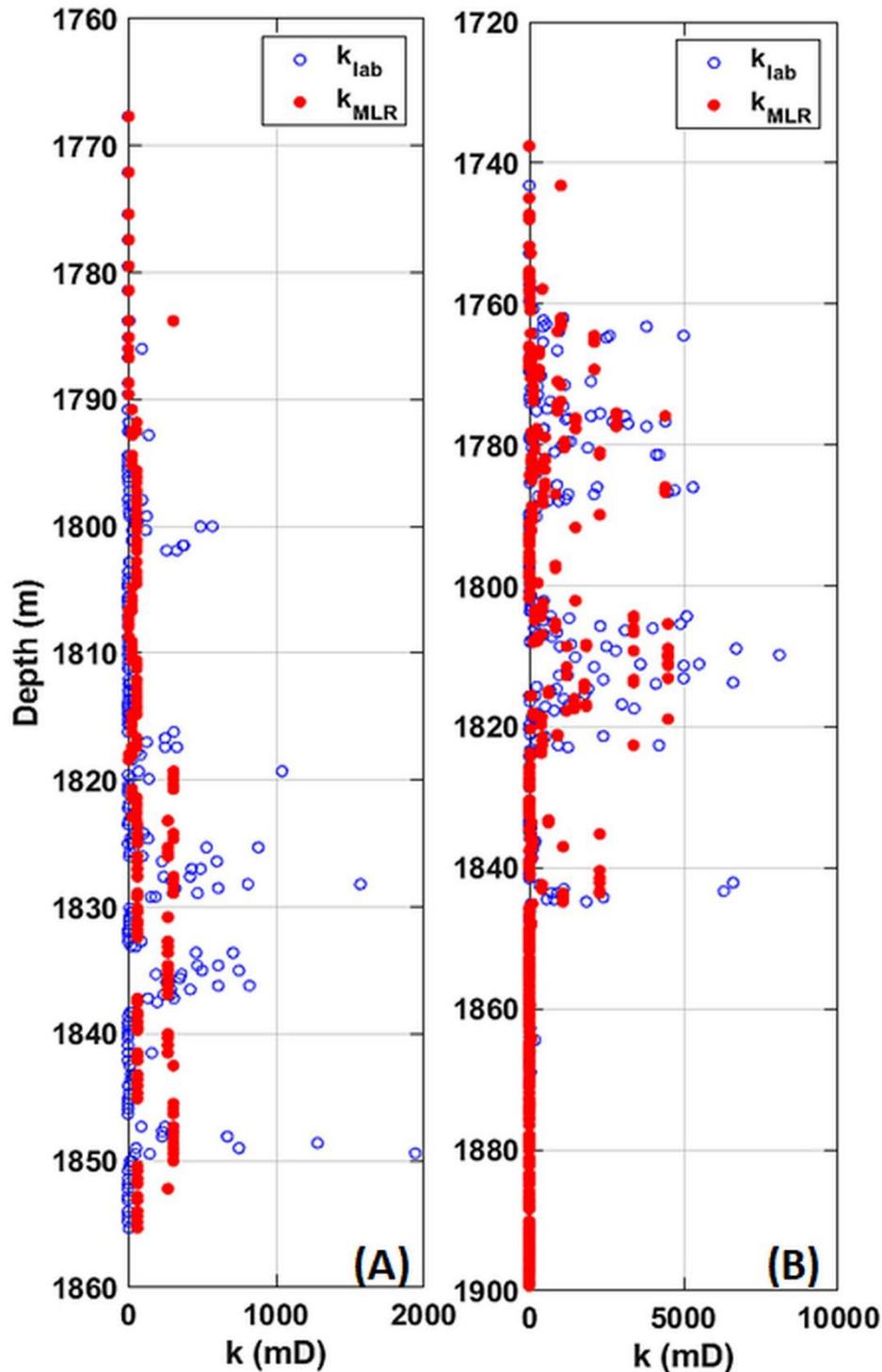
Figure 7: Multilinear regression (red dots, k_{lab}) and laboratory permeability measurements (blue dots, k_{MLR}) for wells A03 (A) and A10 (B).

lithology. The most significant influence of the GR log can be justified by the presence of carbonate mud in the pores, which hinders the flow of fluids in the porous media. The RT log also has a high correlation with permeability, as the electric current circulates along the same paths where the fluids circulate inside the

reservoir. Finally, Figure 12 shows the DTree resulting from this estimate, indicating that it is a complex process, but it is efficient in a multifaceted assessment, such as permeability in carbonate reservoirs. Only the first two final nodes or decision levels (gray rectangles) are presented among the existing twelve.
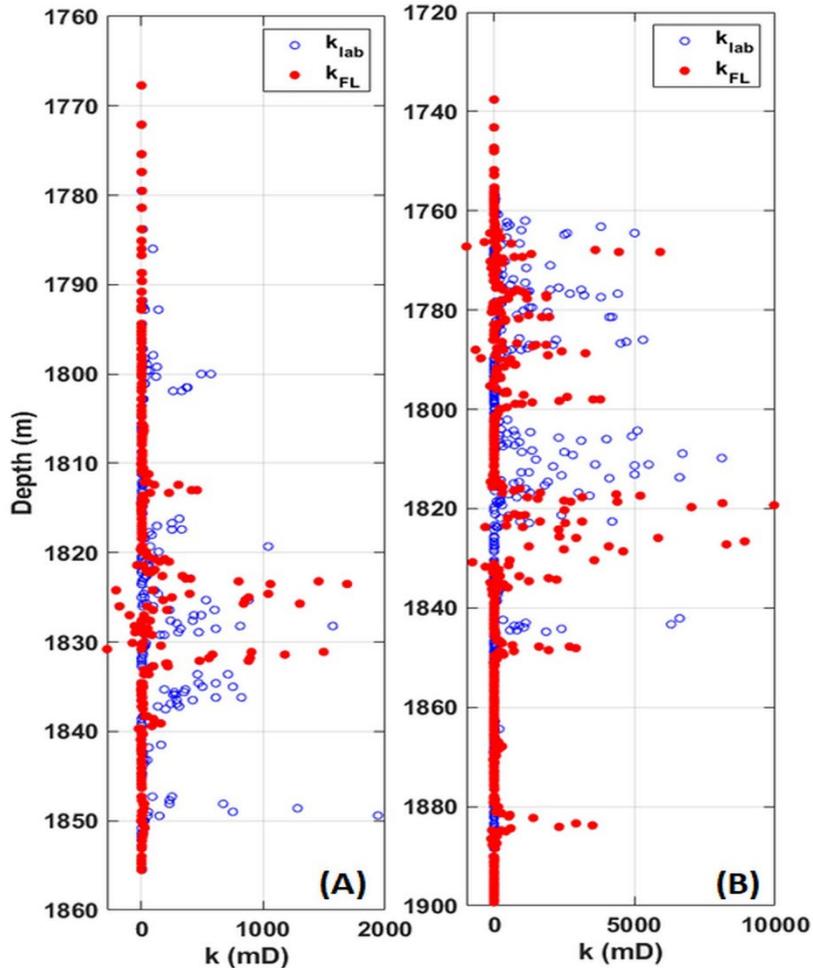
Figure 8: Fuzzy logic estimates (red dots, k_{FL}) and laboratory permeability measurements (blue dots, k_{lab}) for wells A03 (A) and A10 (B).
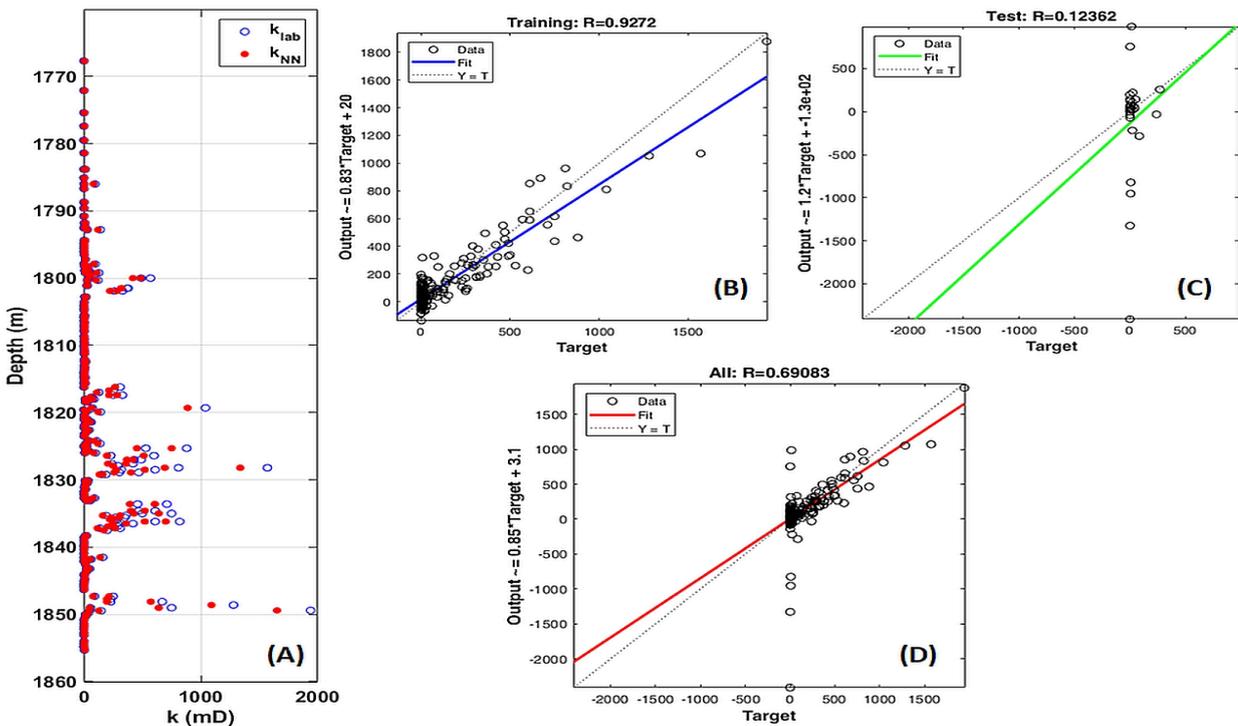
Figure 9: Artificial neural network results for well A03: (A) estimate (red dots, k_{NN}) and laboratory permeability measurements (blue dots, k_{lab}); (B) Training step; (C) test step; and (D) all steps.
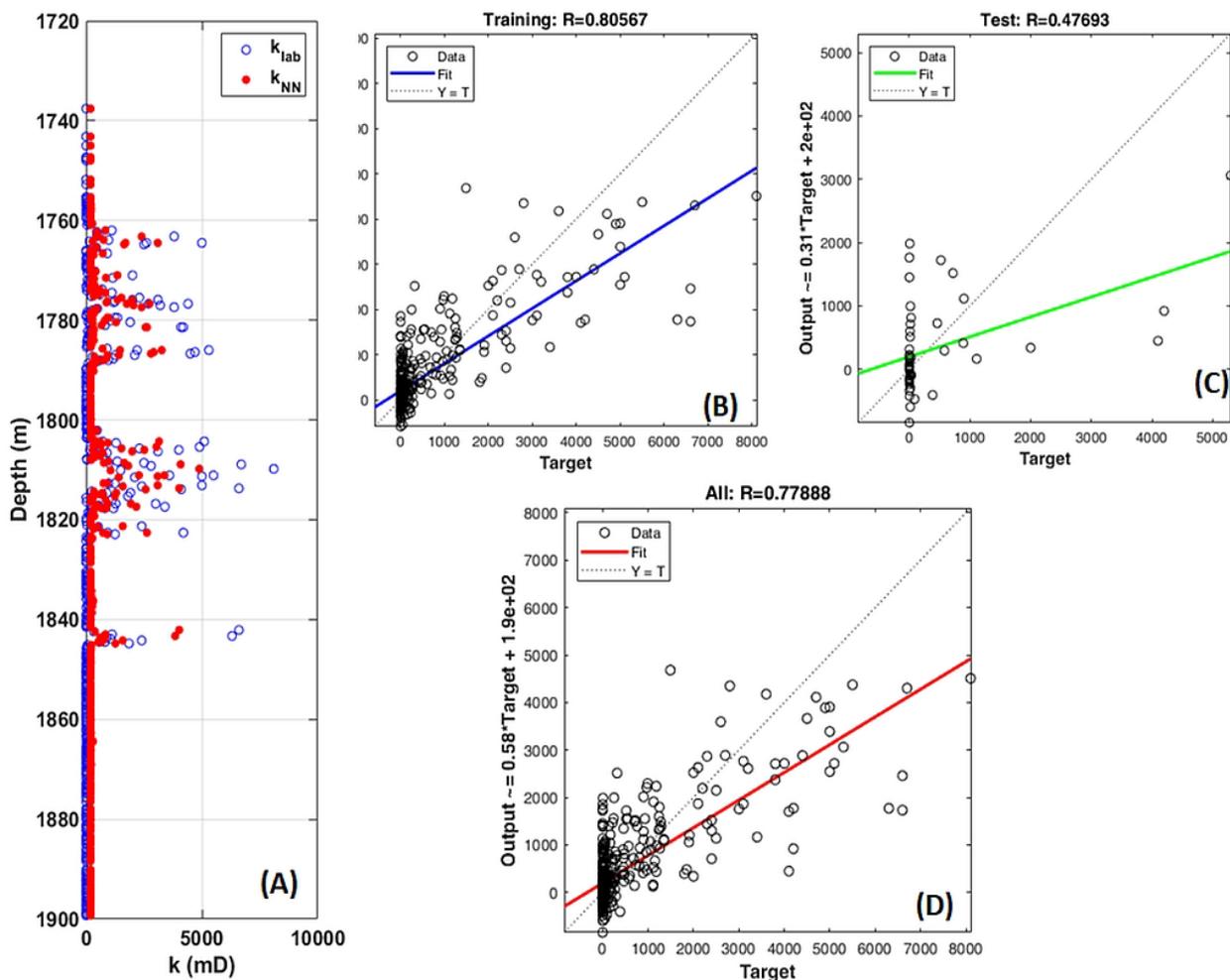
Figure 10: Artificial neural network results for well A10: (A) estimate (red dots, k_{NN}) and laboratory permeability measurements (blue dots, k_{lab}); (B) Training step; (C) test step; and (D) all steps.

CONCLUSIONS

The permeability evaluation in carbonate reservoirs requires very robust calculations due to the heterogeneity of this type of reservoir. Consequently, using well logs, the permeability estimates for wells A03 and A10 of oilfield A in Campos Basin proved difficult. In this case, Timur's empirical model was an unsatisfactory estimate, with Pearson's determination coefficients R=0.10 and 0.15 for wells A03 and A10, respectively. Multiple linear regression, a simple mathematical process, showed R=0.36 and 0.50 in the same order for the wells indicated above, which is an improvement from Timur's estimates. Also, it proved to be better than the fuzzy logic approach, which presented R=0.28 and 0.29 to those same wells. The choice of membership functions and rules makes applying the fuzzy logic approach difficult. The neural network response improved on previous estimates, with R=0.50 and 0.76 for the above wells. But it is the decision tree that performed better than all the other estimates, with R=0.71 and 0.80, in that order, for the wells mentioned above, also showing that the GR and RT logs are the ones that have the most significant weight in the estimates. Executing a decision tree approach seems to be a complicated process, as it presents many disjunctives. But, for estimating a parameter as complex as permeability, it proved to be more effective than other approaches.

ACKNOWLEDGEMENTS
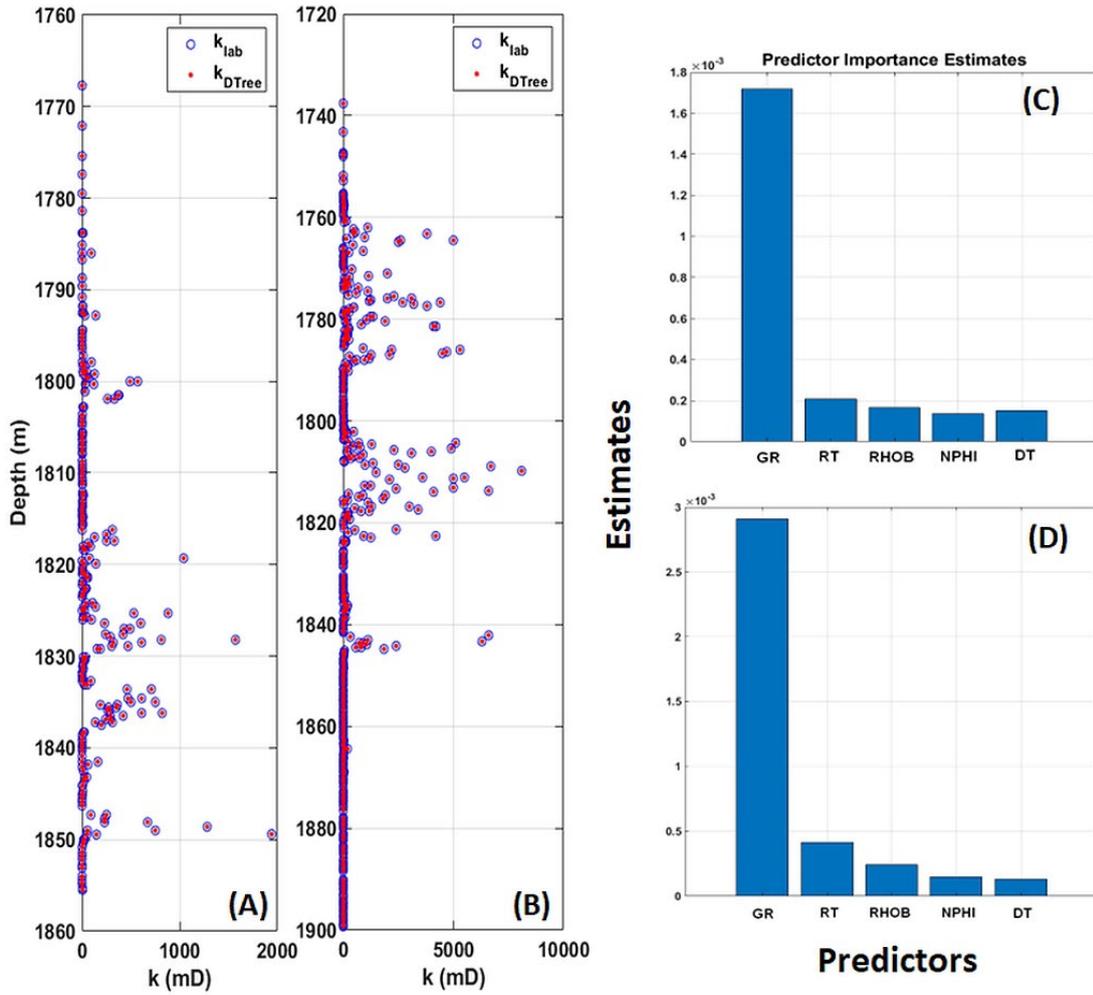
Figure 11: Decision tree estimates (red dots, k_{DTree}) and laboratory permeability measurements (blue dots, k_{lab}): (A) well A03; (B) well A10; (C) predictor importance estimates for well A03; (D) predictor importance estimates for well A10.
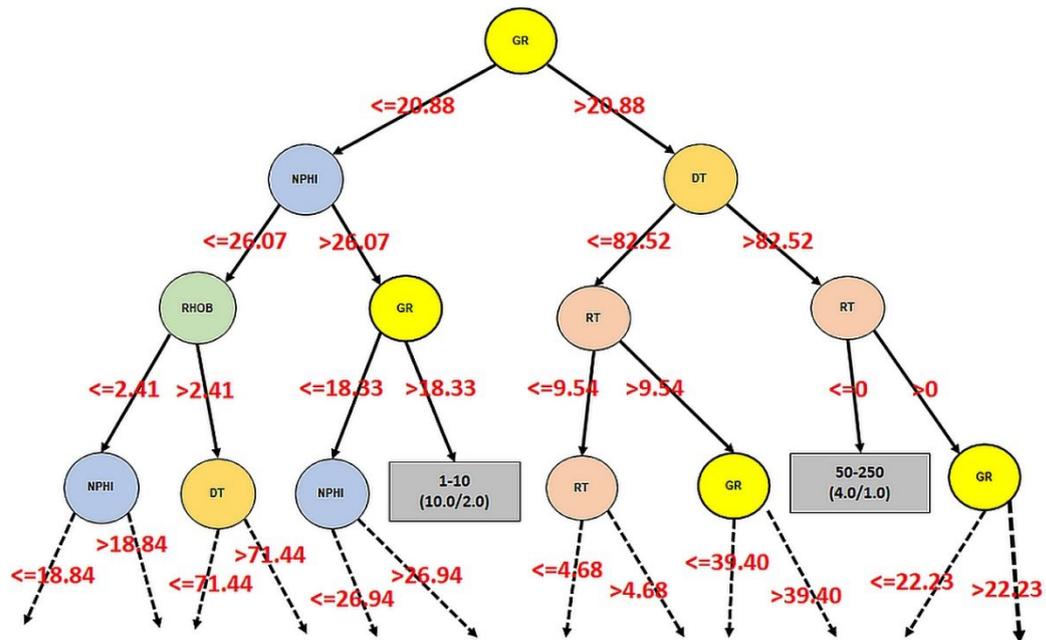
Figure 12: Decision tree scheme for estimating the permeability for wells A03 and A10.

REFERENCES

Ahmadi, M.A., S. Zendehboudi, A. Lohi, A. Elkamel, and I. Chatzis, 2012, Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization: Geophysical Prospecting, **61**, 582–598, doi: 10.1111/j.1365-2478.2012.01080.x.

Bruhn, C., J. Gomes, C. Lucchese, and P. Johann, 2003, Campos Basin: reservoir characterization and management - Historical overview and future challenges: Offshore Technology Conference, Houston, Texas, paper OTC-15220-MS, 14 p, doi: 10.4043/15220-MS.

Cabena, P., P. Hadjinian, R. Stadler, J. Verhees, and A. Zanasi, 1998, Discovering the data mining: from concept to implementation. Prentice Hall, Upper Saddle River, NJ, 195 p.

Cameron, A.C., and F. Windmeijer, 1997, An *R*-squared measure of goodness of fit for some common nonlinear regression models: Journal of Econometrics, **77**, 2, 329–342, doi: 10.1016/S0304-4076(96)01818-0.

Cuddy, S., 1997, The application of the mathematics of fuzzy logic to petrophysics: SPWLA 38th Annual Logging Symposium, Houston, TX, SPWLA-1997-S.

Dell'Aversana, P., 2019, Comparison of different Machine Learning algorithms for lithofacies classification from well logs: Bollettino di Geofisica Teorica ed Applicata, **60**, 69–80, doi: 10.4430/bgta0256.

Geoactive, 2021, Geoactive Limited, Interactive Petrophysics version 4.0, User´s Manual.

Guardado, L., L. Gamboa, and C. Lucchesi, 1989, Petroleum geology of the Campos Basin, Brazil, a model for a producing Atlantic type of basin, *in* Edwards J.D., and Santogrossi P.A. Divergent / passive margin basins: AAPG Memoir, **48**, p. 3–79, doi: 10.1306/M48508C1.

Harnett, D., and J. Horrell, 1998, Data, statistics, and decision models with Excel: John Wiley & Sons Inc., New Jersey, USA, 624 pp.

Holmes, G., A. Donkin, and I. Witten, 1994, Weka: a machine learning workbench: Proceedings of the Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, p. 357–361.

Luthi, S., and I. Bryant, 1997, Well-log correlation using a back-propagation neural network: Mathematical Geology, **29**, 413–425, doi: 10.1007/BF02769643.

Matlab, 2018, Statistics Toolbox Release: The MathWorks, Inc., Natick, Massachusetts, United States.

Mitchell, T., 1997, Machine learning: McGraw-Hill Science, Ithaca - NY, 432 p.

Mohaghegh, S., R. Arefi, S. Ameri, and D. Rose, 1995, Design and development of an artificial neural network for estimation of formation permeability: SPE Computer Applications, July 31–August 3, Dallas, SPE 28237-PA, doi: 10.2118/28237-PA.

Nordlund, U., 1996, Formalizing geological knowledge – with an example of modeling stratigraphy using fuzzy logic: Journal of Sedimentary Research, **66**, 4, 689–698, doi: 10.1306/D42683E2-2B26-11D7-8648000102C1865D.

Okubo, J., R. Lykawka, L. Warren, J. Favoreto, and D. Dias-Brito, 2015, Depositional, diagenetic, and stratigraphic aspects of Macaé Group carbonates (Albian): example from an oilfield from Campos Basin: Brazilian Journal of Geology, **45**, 2, 243–258, doi: 10.1590/23174889201500020005.

Petrobras, 2012, Oilfield A Report, Petrobras: UENF Agreement, RJ, Brazil, 150 pp. (In Portuguese).

Quinlan, J., 1987, Simplifying decision trees: International Journal of Man-Machine Studies, **27**, 3, 221–234, doi: 10.1016/S0020-7373(87)80053-6.

Russell, S., and P. Norvig, 2009, Artificial Intelligence: A Modern Approach: 3rd ed., Upper Saddle River, New Jersey, Prentice Hall, 1132 p.

Schön, J., 2015, Physical properties of rocks: Fundamentals and Principles of Petrophysics: 2nd ed., Oxford, England, Elsevier, Developments in Petroleum Science, **65**, 512 p.

Seber, G., and A. Lee, 2003, Linear regression analysis: 2nd ed., John Wiley & Sons, Inc., Hoboken, New Jersey, Wiley Series in Probability and Statistics, 583 p, doi: 10.1002/9780471722199.

Silva, A., I. Lima Neto, R. Misságia, M. Ceia, A. Carrasquilla, and N. Archilha, 2015, Artificial neural networks to support petrographic classification of carbonate-siliciclastic rocks using well logs and textural information: Journal of Applied Geophysics, **117,** 118–125, doi: 10.1016/j.jappgeo.2015.03.027.

Silva, A., M. Tavares, A. Carrasquilla, R. Misságia, and M. Ceia, 2020, Petrofacies classification using machine learning algorithms: Geophysics, **85**, 4, WA101–WA113, doi: 10.1190/geo2019-0439.1.

Theodoridis, S., and K. Koutroumbas, 2009, Pattern recognition: Academic Press, Burlington, MA, USA, 967 p.

Tiab, D., and E. Donaldson, 2004, Petrophysics: Oxford, Elsevier, 920 pp.

Timur, A., 1968, An investigation of permeability, porosity, & residual water saturation relationships for sandstone reservoirs: SPWLA, Houston, TX, USA, The Log Analyst, 9 04, Paper Number: SPWLA-1968-vIXn4a2.

Winter, W., R. Jahnert, and A. França, 2007, Campos Basin: Boletim de Geociências da Petrobras, **15**, 511–529. (In Portuguese).

Yan, J., 2002, Reservoir parameters estimation from well log and core data: a case study from the North Sea: Petroleum Geoscience, **8**, 63–69, doi: 10.1144/petgeo.8.1.63.

GOMES, R.: this work is part of this author's master's thesis; **CARRASQUILLA, A.:** advisor of the master's thesis. The authors reviewed the original work, rewriting it and adding many items, mainly figures, equations etc.