**Prediction of porosity and permeability using machine learning techniques in Tupi Oilfield wells**

Arthur R. Nunes, Gelvam A. Hartmann, Alexandre C. Sanchetta, Emilson P. Leite, Nathalia H. S. Mattos, Henrique B. Santos

Institute of Geosciences, University of Campinas (IG/UNICAMP)

Copyright 2021, SBGf - Sociedade Brasileira de Geofísica.

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 16-19 August 2021.

Contents of this paper were reviewed by the Technical Committee of the 17th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Abstract**

**Summary**

The prediction of petrophysical properties such as porosity, water saturation and lithofacies are essential for exploration and development of hydrocarbon reservoirs. Such estimates are routinely determined by constructing rock-physics models that link elastic parameters obtained from seismic data inversion with those petrophysical properties extracted from well logs or from laboratory measurements of core samples. However, seismic data are usually the only source of information available away from the wells that can be used to predict the 3D distribution of physical properties with appropriate spatial resolution. In addition, predictions in deeper and unconventional conditions are difficult to determine. The petrophysical characterization of some pre-salt carbonate rocks (such as in Brazilian pre-salt reservoirs) is revealing to be very difficult because of their heterogeneities at several scales. Here, we will predict porosity and permeability through machine learning (ML) supervised algorithms from well logs datasets of five wells from Tupi Oilfield, Santos Basin in wells “3-BRSA-369A-RJS (Petrobras)”, “9\_BRSA-716-RJS (Petrobras)”, “8-LL-9-RJS (Petrobras)”, “3-BRSA-883-RJS (Petrobras)” and “9-BRSA-908D-RJS (Petrobras)”. The predictions of porosity and permeability were performed using the Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting (GB), Decision Tree (DT), Linear Regression (LR) and K-Nearest Neighbors (KNN) methods. Input data include five different logs from each analyzed well to predict porosity and permeability. In-well analyses were performed with 30%, 50% and 70% of training data of the total available data, to predict the rest of data of each well.

**Results and Conclusions**

Input data include basic curves of GR (Gamma Ray, gAPI), RHOB (Bulk-Density, g/cm³), NPHI (Neutron Porosity), ResDeep (Deep Resistivity, Ohm.m) and PEF (Photoelectric Factor, b/e). We avoid data that could present some bias to the target petrophysical property. As the tests were carried out, the final model was defined by the best prediction performance within the five methods used. The performance of each method was analyzed using error metrics, such as the mean absolute error and Spearman's correlation coefficient. Results for permeability predictions indicated that the DT was the best method for the 30% of training data (Spearman’s correlation test of 0.99) and for the other 50% and 70% the KNN returned the best correlations (0.72 and 0.68, respectively). Porosity presented the best correlations through RF for the three training tests (Spearman’s correlation of 0.81, 0.84 and 0.84, respectively). We observed that the five ML methods are useful to predict in-well petrophysical properties for carbonate rocks of the Tupi Oilfield. The results are satisfactory considering that the input curves are not derived from the predicted one, because when inserting curves directly related to the target curve the result becomes more accurate, however, biased. Tests of inter-well in these same analyzed wells will be performed in order to evaluate the performance of these applied ML methods.