Pre-salt carbonate lithofacies classification through Random Forest

Alexandre Cruz Sanchetta, Emilson Pereira Leite, Nathalia Helena Secol Mattos, Marta Henriques Jácomo, UNICAMP

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.

Summary

Classifying lithofacies from well-log data is difficult due to the complex carbonate facies structures, strong depositional processes and particularly challenging logging conditions in pre-salt reservoirs. With the contemporary progresses in the development of powerful machine learning algorithms and the increasing of computational power, they became widely used for data regression and classification in the oil and gas industry to help in overcoming these problems.

Facies classification is one of the main applications of machine learning in the context of reservoir characterization, which is crucial in determining the net pay thickness of reservoirs that is used as an important factor in the decision making process. The automatization of the facies classification process using machine learning is a way to facilitate facies interpretation, especially when it is based on large volume data.

In this context, we present a workflow aiming at improving XRD derived facies classification in the brazilian pre-salt reservoirs based on the application of Principal Component (PC) Analysis, Independent Component (IC) Analysis and Convolution-based Smoothing (SMO); and using a balanced selection of the training dataset followed by a Random Forest (RF) classification. The workflow also allows an adequate quantification of classification errors and ensures that different scenarios are evaluated. This workflow was applied for a dataset of well 9-BRSA-716 (9-RJS-660) of the Tupi oil field, Santos Basin, Brazil. We classified four lithofacies identified from XRD derived mineralogical contents: dolomite, calcite, quartz and not-quartz.

Results and Conclusions

We used eleven curves from well 9-BRSA-716: compressional slowness (DTc), shear slowness (DTs), gamma-ray (GR), Timur-Coates permeability (KTIM), NMR effective porosity (NMReff), NMR free-fluid porosity (NMRff), NMR total porosity (NMRtt), neutron porosity (NPHI), photo-electric absorption factor (PEF), deep resistivity (RESDEEP), and bulk density (RHOB). We also created a subset with five basic curves: GR, NPHI, PEF, RESDEEP and RHOB.

We worked with the reservoir production intervals of the Barra Velha Formation (BVE100, BVE200 and BVE300) and with a non-reservoir interval of the Itapema Formation, totaling 1594 depth samples along each well log curve. To validate the RF classification, we randomly selected approximately 10% of the samples to form a training set and classified the remaining samples as a test set. The best results are observed when each training class has an approximately equal number of samples (Balanced Training Set - BTS).

The performance of the RF was monitored by its overall classification accuracy (CA) as we performed statistical analysis to understand the impact of the input curves (features) and modified the training and test sets as follows: 1) 5 features (CA = 61.1%); 2) 11 features (CA = 59.2%); 3) 5 features with BTS (CA = 80.1%); 4) 11 features with BTS (CA = 83.7%); 5) 5 PCs from 11 features with BTS (CA = 80.6%); 6) 11 PCs from 11 features with BTS (CA = 80.1%); 7) 5 ICs from 11 features with BTS (CA = 80.3%); 8) 11 ICs from 11 features (CA = 81.4%); and 9) 5 ICs from 11 features with BTS and SMO (CA = 87.6%). Therefore, we achieved the best result in case 9. Classification accuracy for each individual lithofacies is above 80%.