Petrophysical validation in the reconstruction of well logs using Machine Leanign. Case study in Colombia

Luis Alfredo Montes Vides, Andrés Mauricio Muñoz García.
Departamento de Geociencias, Universidad Nacional de Colombia.

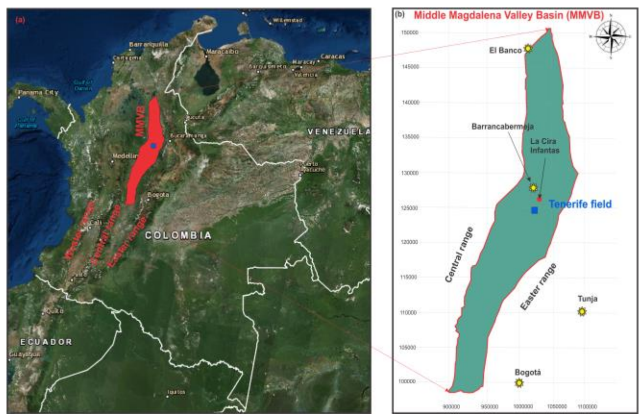
Abstract

We present the work methodology and the results of the reconstruction of the well log of the Tenerife Field in Colombia using Machine Learning. We trained Random Forest algorithms to estimate the DTSM record in the depth ranges in which these records were not taken for the Tenerife-1, Tenerife-2 wells, and the complete estimate in the Tenerife-3 well. We use for the training the P-wave (DT), spontaneous potential (SP), density (rho), deep induction resistivity (ILD), and pseudo-records porosity (phi), clay volume (Vclay), and water saturation (Sw). From the Vclay we made the classification of lithological facies, sands for Vclay ≤ 0.6 and shales for Vclay ≥ 0.6, with this classification, we determined Flood Plain deposit environments and continuous sand channels in the Tenerife field. We validated the estimated records from the petrophysical model of the field, and the results were satisfactory. This is perhaps one of the first studies in Colombia in this line of work, and is proposed as a solution to the lack of well records in new areas of oil interest, such as the case of old wells whose records were taken only in the geological formations whose resources are already exhausted.

Introduction

Exploration in Sub-Andean basins in Colombia has been traditionally challenging due to the complex geology, rough topography and inaccessible areas. The Middle Magdalena Valley Basin (MMVB) is one the most prolific petroleum basins of Colombia with a long history of hydrocarbon exploration that started with the discovery of a giant oil field called La Cira-Infantas, in 1918. The oilfields in the basin occur mainly as either structural or stratigraphic traps in Tertiary clastic reservoirs. The Tenerife Field (see Fig. 1), operated by Ecopetrol S.A., is located in the central MMVB, about 20 km south-west of the giant La Cira-Infantas Field. This field was discovered in 1971 after positive results of drilling the Tenerife-1 well. The appraisal strategy was followed by the drilling of wells Tenerife-2 and Tenerife-3. The development of the field was stopped later due to failed results in Tenerife-3. Tenerife-1 and Tenerife-2 had production of about 100 STBO/day of 22.8 API crude oil; however they are almost totally depleted today, see Ecopetrol (1983) and Sandoval (2009).

Figure 1. The MMVB. (a) Location and (b) fields La Cira-Infantas and Tenerife; the blue squere is where the Tenerife Field is located.



The continental deposits in the MMVB are composed mainly of red beds and volcanic effusive and pyroclastic deposits, although some marine facies appear locally. For the time of the Cretaceous, the basin sedimentation continued in a back-arc setting east of the Andean subduction zone (see Cooper et al., 1995).

Throughout the Aptian time, the pure extension terminated and a postrift sedimentationphasebegan, Controlledbythermal subsidence. At the same time, the following formations were deposited (see Figure 2): Tablazo, which consists of a set of limestones interbeded with black shales; Simití, which is mainly made of shales deposited in an inner to middle shelf; La Luna, which consists of interbedding of limestones and black organic-rich shales; Umir, which marks a transition into a more clastic environmentandiscomposed of gray shales and shaly sandstones with coal intervals. The major deformation phase of the CR affecting the MMVB marks a regional unconformity called the Middle Eocene unconformity (MEU). It is the most important boundary within the Cenozoic record and can be observed throughout the basin (see Figure 2). The MEU separates a premiddle Eocene deformation from a post Middle Eocene continuous deformation with thrust-controlled exhumation. For a comprehensive study in this area, see Velasquez-Espejo (2012).

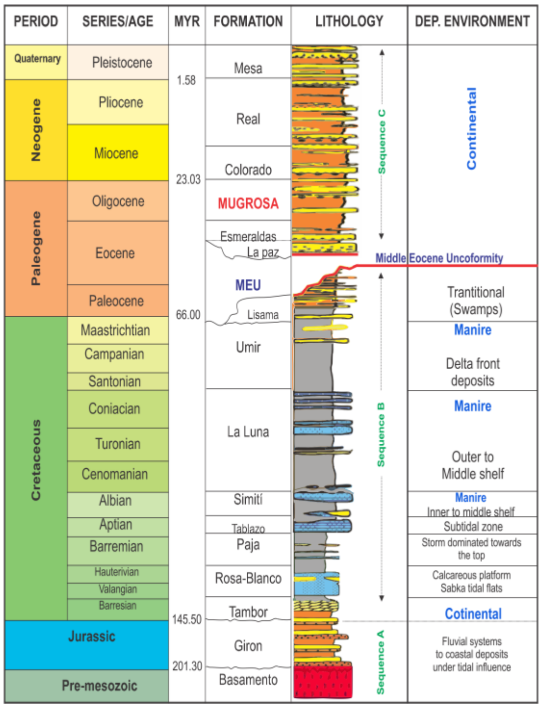


Figure 2. Stratigrafic colum of the MMVB illustrating the main petroleum system elements. The highligted red curved lines are the Cenozoic unconformities.

Method

To build a DTSM well log using Machine Learning and validate the resuls using the petrophysycal model, we implents a full procedure that is show in Figure 3 and it is descrited in the following steps.

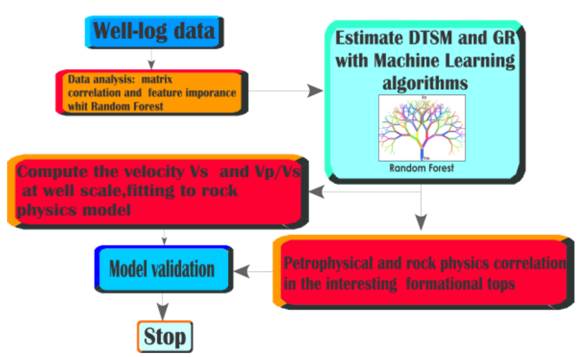
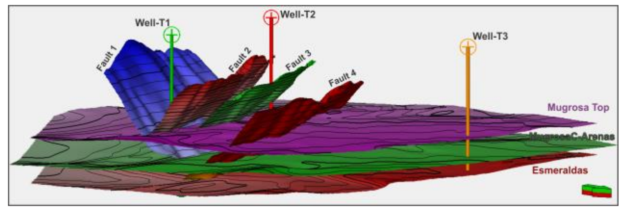


Figure 3. Procedure to validate well log estimated using Random Forest. Petrophysical and rock physics correltion in the interesting formational tops.

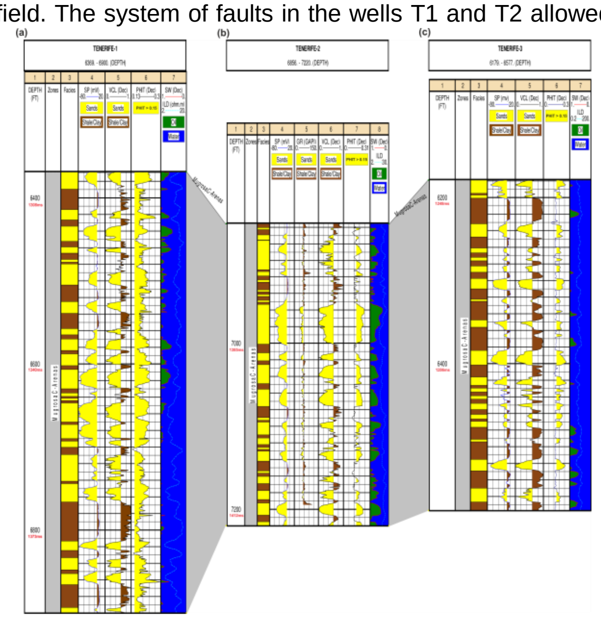
Seismic survey and well-log data

We use the available seismic data and wells Tenerife-1, Tenerife-2 and Tenerife-3. The reservoir characterization has been defined based on the seismic and geologic interpretations, petrophysical, and reservoir engineering data. These data were used to define the geologic modelo of study area. The results of the structural and stratigraphic interpretation are shown in the figure 4, and



the petrophysical model is shown in the figure 5.

Figure 3. Simplified 3D structural model of the Tenerife field. The system of faults in the wells T1 and T2 allowed



the entrapment of hydrocarbon in this area of the field, in well T3 no structural closure was observed.

Figure 4. Identification of deposit environment flood plain (Brown color) and sand channels (yellow color), interpreted as parto f a meandering fluvial system.

Analyze the data by means the correlation matrix and the importance factor using and training of random forest algorithms

The training and prediction stage are summarized in figure 5. Prior to training we selected the best combination of well records using the correlation matrix and the importance factor using Random Forest, plus we calculated the correlation factor and the mean squared error (MSE) , the results are shown in the figure 6.

We use a method derived from the algorithm known as decision trees in which different sets of them are used randomly to make decisions that lead to the best possible results (Breiman 2001). The main parameter to adjust is the number of trees or estimators that in our case was taken as one hundred. as in neural networks, the fraction of data taken for the training was 80 % and to validate the method the remaining 20 % was used.

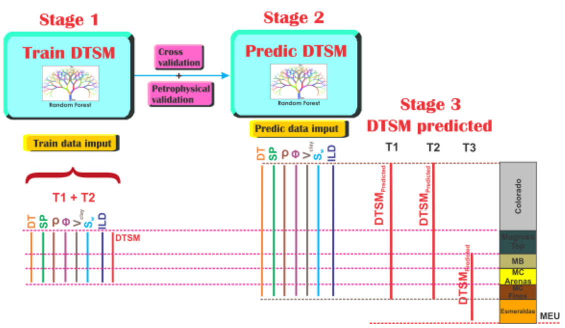


Figure 3. Stages of training and prediction of DTSM. Stage 1: The input data in the training are the records and pseudo-records of wells T1 and T2, on the other hand the output data is in this case DTSM. Stage 2: Once the Random Forest algorithms are entered, we continue with the estimation and finally in stage 3 we validate the estimated records.

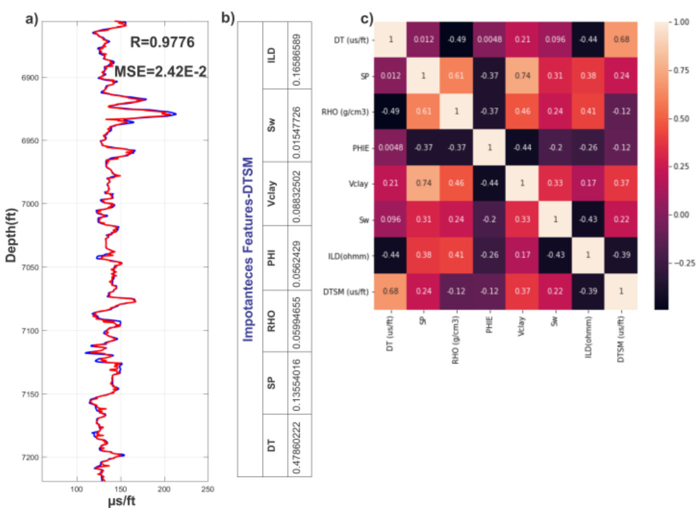


Figure 4. Estimation of DTSM using Random Forest. a) We have a correlation fator of R=0.9776 between these two paramenters and MSE=$2.42x10^{-2}$, b) Importance Features, and c) correlation matrix.

Results

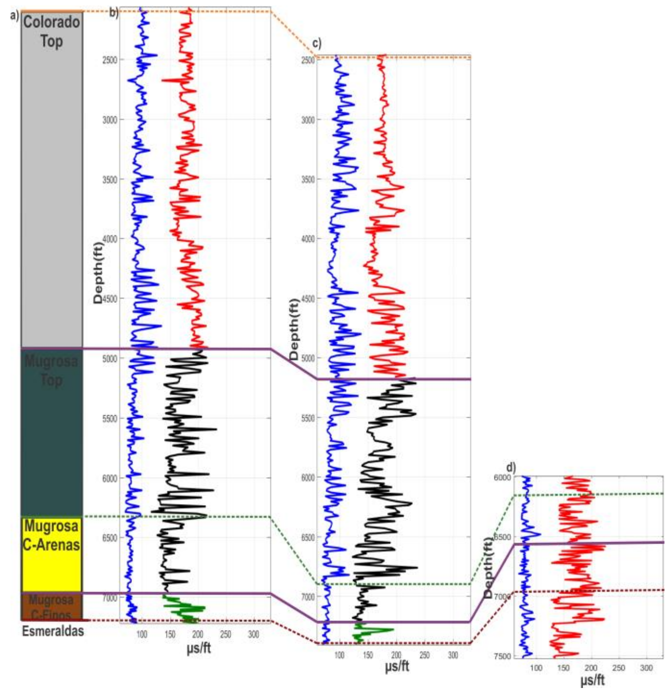


Figure 5. Estimation of DTSM using Random Forest in the wells T1 (a), T2 (b) y T3 (c).

Conclusions

The facies estimated agree with the sand bodies and its small clay intercalations identified by the well-logs. The results far away from the borehole also allow to identify continuous sand bodies with shally intercalations and in many cases it is possible to distinguish between the limits of the sand zone.

The results of the comparison between the observed and calculated DTSM well log show a reasonable fit, In addition, the results achieved by Machine Learning in the areas where it was not taken, shows a very good adjustment whit the DT well log and the generated petrophysical model.

Acknowledgments

We thank Ecopetrol S.A. within the research program 0266-2013 Colciencias Ecopetrol, for providing the well and seismic data used in this research.

References

Breiman L (2001). Random Forests. Machine Learning. 45 (1): 5–32. doi:10.1023/A:1010933404324.

Ecopetrol, S., 1983, Evaluación de yacimientos del campo Tenerife:Technical report, Ecopetrol S.A.

Sandoval, R. J., 2009, Definición de un modelo sedimento- estratigráfico, Formación Mugrosa, Campo Tenerife, Valle Medio del Magdalena, Colombia: Technical report, Ecopetrol S.A.

Velasquez-Espejo, A. J., 2012, 3D Multicomponent seismic characterization of a clastic reservoir in the Middle Magdalena Valley Basin, Colombia: PhD thesis, Colorado School of Mines, Golden, Colorado.

Cooper, M., F. Addison, R. Alvarez, M. Coral, R. Graham, A. Hayward, S. Howe, J. Martinez, J. Naar, R. Peñas, A. Pulham, and A. Taborda, 1995, Basin development and tectonic history of the Llanos basin, Eastern cordillera, and Middle Magdalena Valley, Colombia: AAPG Bulletin, 79, 1421–1443.