Log Simulations with Machine Learning and Their Applications in the Montney Evaluation

Huiju Geng, Daniel Haider
ARC Resources Ltd.

Summary

A complete suite of well logs is essential for accurate subsurface characterization. However, this ideal condition is not always met, as many wells lack the necessary logs for comprehensive petrophysical analysis. A common solution to address these missing data is to generate synthetic logs, which involves leveraging the relationships between different logs. By utilizing data from offset wells, models can be built to predict synthetic logs for wells where specific measurements are missing. Machine learning, implemented through Python libraries, has proven to be an effective tool for log simulation. This study focuses on generating synthetic logs using machine learning techniques in the Montney Formation, highlighting their applications for addressing missing data and improving subsurface characterization.

Workflow

A set of wells with good borehole conditions, high-quality logs, and complete logging suites was selected from the Attachie area to train supervised machine learning models. Published studies suggest that artificial neural networks (ANN) and random forest (RF) regression algorithms typically provide the best predictions. Log ASCII Standard (LAS) data from multiple wells were loaded and merged into a combined DataFrame. The target log for prediction was designated as the dependent variable, while other logs served as independent variables. The dataset was divided into an 80% training set and a 20% test set. The training set was used to train the model, which was subsequently tested on the test set. Model performance was evaluated by comparing predicted values against actual values in the test set using metrics such as mean absolute error, root mean square error, and correlation coefficients. Once a satisfactory model was achieved, the trained model could be applied to new wells for blind test or to generate synthetic logs where the actual log is unavailable.

Results Discussion

In our process of generating synthetic logs for several curves, random forest regression consistently provides better predictions than ANN. Figure 1 presents the test data of a sonic simulation example, where the synthetic compressional slowness (DT) was predicted from gamma ray (GR), neutron porosity, bulk density, and deep resistivity logs. The RF-predicted DT aligns more closely with the actual DT compared to the ANN-predicted DT, demonstrating smaller errors and higher correlation coefficients. The size and complexity of our data set, within the area of interest, resulted in random forest regression outperforming ANN models.

One of the most common applications of log simulation is the generation of synthetic shear sonic logs. A subset of wells with dipole sonic data was identified, and the dependent and independent variables were reconfigured to predict shear slowness (DTS) using GR, neutron porosity, bulk density, compressional Delta-T, and deep resistivity logs. To examine the impact of the training data size, predictions were conducted in two separate runs and blind-tested in the study well, *Well A*, where actual shear sonic log is available for comparison.

In the first run, data from four wells adjacent to *Well A* were used to train the model, assuming that shear slowness is influenced by local variation in depth and stress distribution. The predictions were acceptable but showed room for improvement, Figure 2. In the second run, data from four additional regional wells with shear sonic logs were included, resulting in a second version of the synthetic DTS. The second version showed better agreement with the actual DTS, with lower errors and a higher correlation coefficient, Figure 3. The two synthetic DTS curves, compared to the actual DTS, are displayed in the right two tracks of the log panel in Figure 4. This example suggests that local variations in depth and stress may not significantly impact shear slowness predictions. Instead, training data with more representative samples improves prediction accuracy. Based on these results, the shear simulation model appears applicable across the Attachie area, and can be used to generate synthetic shear logs for geophysical or geomechanical applications in the absence of actual shear logs.

Another example, illustrated in Figure 5, involves the prediction of synthetic Pef logs in *Well B*. The Pef log measures the photoelectric absorption factor, defined as $(Z/10)^{3.6}$, where Z represents the average atomic number of the formation. The Pef log is particularly valuable for determining mineralogy; however, it is highly sensitive to heavy minerals like barite in the mud cake or mud filtrate, which has an exceptionally high Pef value (267) compared to common minerals (Pef < 6). As noted in the well header, the mud density for *Well B* was recorded as 1360 kg/m^3, clearly indicating that heavy minerals were added to the mud. Consequently, the Pef log exhibited erratic readings, rendering it invalid for petrophysical analysis. To resolve this challenge, a synthetic Pef log was generated using random forest regression, with gamma ray (GR), neutron porosity, bulk density, compressional Delta-T, and deep resistivity logs as inputs. Both the synthetic Pef (black) and raw Pef (pink) curves are displayed in the right track in Figure 5.

Figure 6 compares the data distribution of the raw and synthetic Pef logs from *Well B* with those of nearby offset wells, and the synthetic Pef log falls within a similar range as the neighboring wells. The synthetic Pef log was subsequently used in a multimineral petrophysical evaluation, where the ultimate results are primarily influenced by the formation mineralogy composition. In this case, the output grain density and porosity aligned with core data, validating the effectiveness of the synthetic Pef log, Figure 7.

In conclusion, the application of machine learning provides an efficient and reliable tool for generating synthetic logs. This approach could be used to address missing data or fix poor quality logs that were affected by drilling or wellbore conditions, thereby enhancing subsurface characterization and improving the accuracy of reservoir evaluations.

References

Akinnikawe, O., Lyne, S., Roberts.J., 2018. Synthetic Well Log Generation Using Machine Learning Techniques. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, July 2018.Paper Number: URTEC-2877021-MS

Anemangely, M., Ramezanzadeh, A, Amiri, H., Hoseinpour,S., Machine learning technique for the prediction of shear wave velocity using petrophysical logs, Journal of Petroleum Science and Engineering, Volume 174,2019,Pages 306-327,ISSN 0920-4105

McDonald, A., 2023. Random Forest for Regression - Prediction of Continuous Well Logs. https://github.com/andymcdgeo/Petrophysics-Python-Series

McDonald, A., 2023. ANN for Regression - Prediction of Continuous Well Logs https://github.com/andymcdgeo/Petrophysics-Python-Series

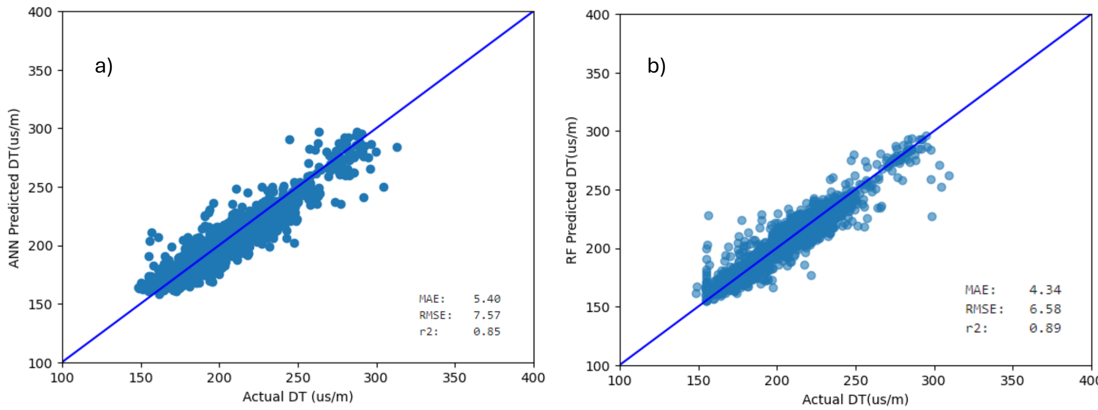

Figure 1: Performance comparison of ANN and Random Forest algorithms in simulating sonic logs: a) Comparison of ANN-predicted DT and actual DT; b) Comparison of Random Forest-predicted DT and actual DT.

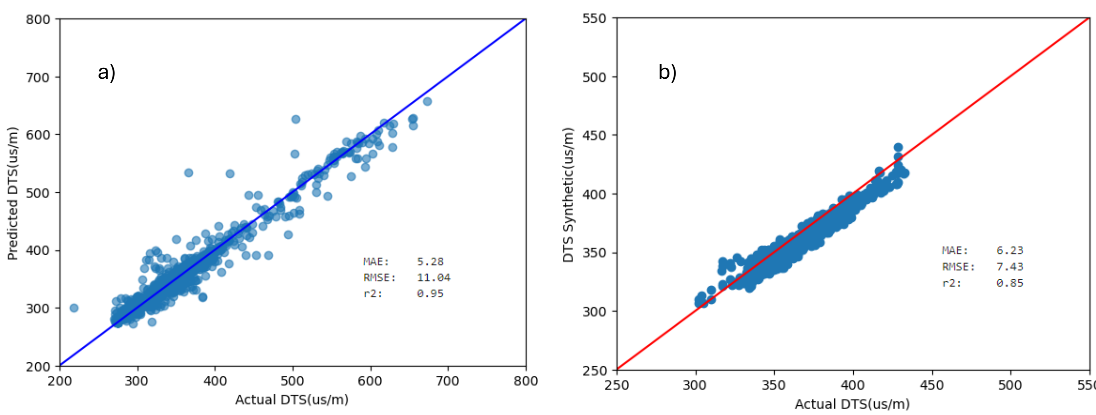

Figure 2: Results of shear sonic simulation with random forest using training data from 4 offset wells: a) Comparison of RF-predicted shear slowness and actual DTS with the test data (20% of the training data set);b) Comparison of RF-predicted shear slowness and actual DTS, blind test results in *Well A*.

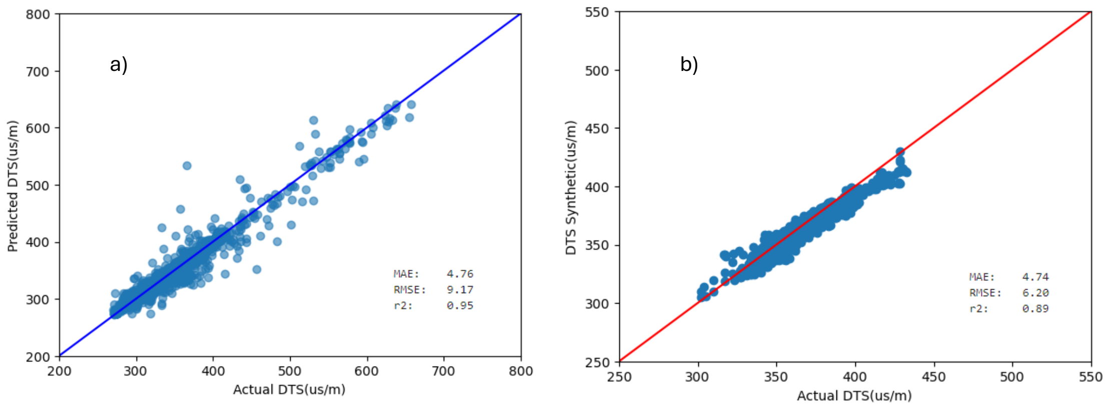

Figure 3: Results of shear sonic simulation with random forest using training data from 8 regional wells: a) Comparison of RF-predicted predicted shear slowness and actual DTS with the test data (20% of the training data set); b) Comparison of RF-predicted shear slowness and actual DTS, blind test results in *Well A*.

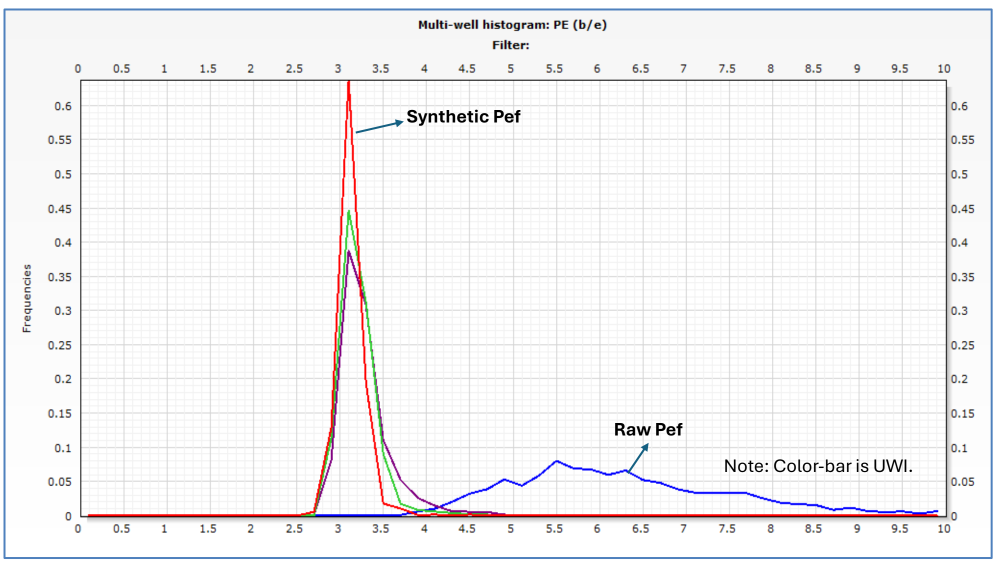

Figure 6: Comparison of the data distribution ranges for raw Pef, synthetic Pef of *Well B* and offset wells.

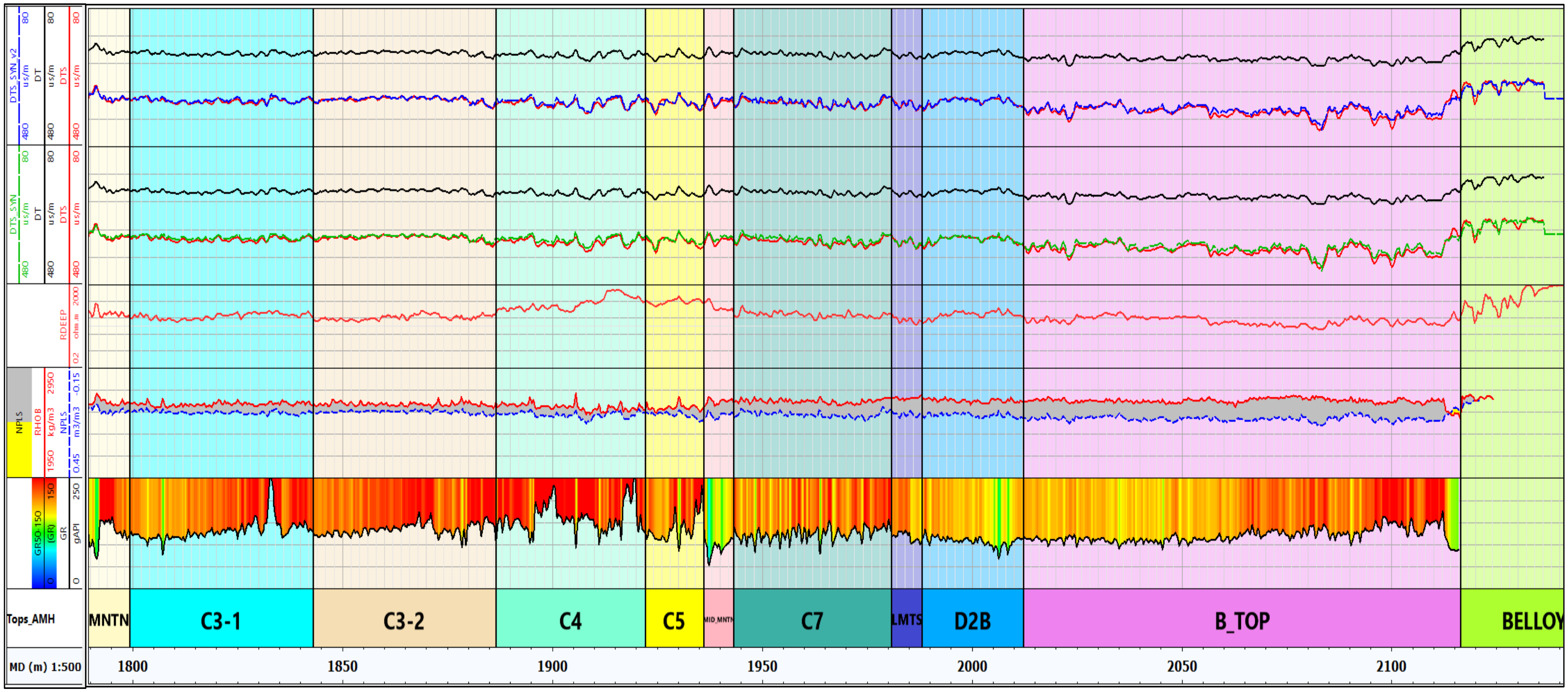

Figure 4: Log presentation of *Well A* over the Montney Formation, illustrating input logs and synthetic shear sonic logs alongside actual shear slowness.

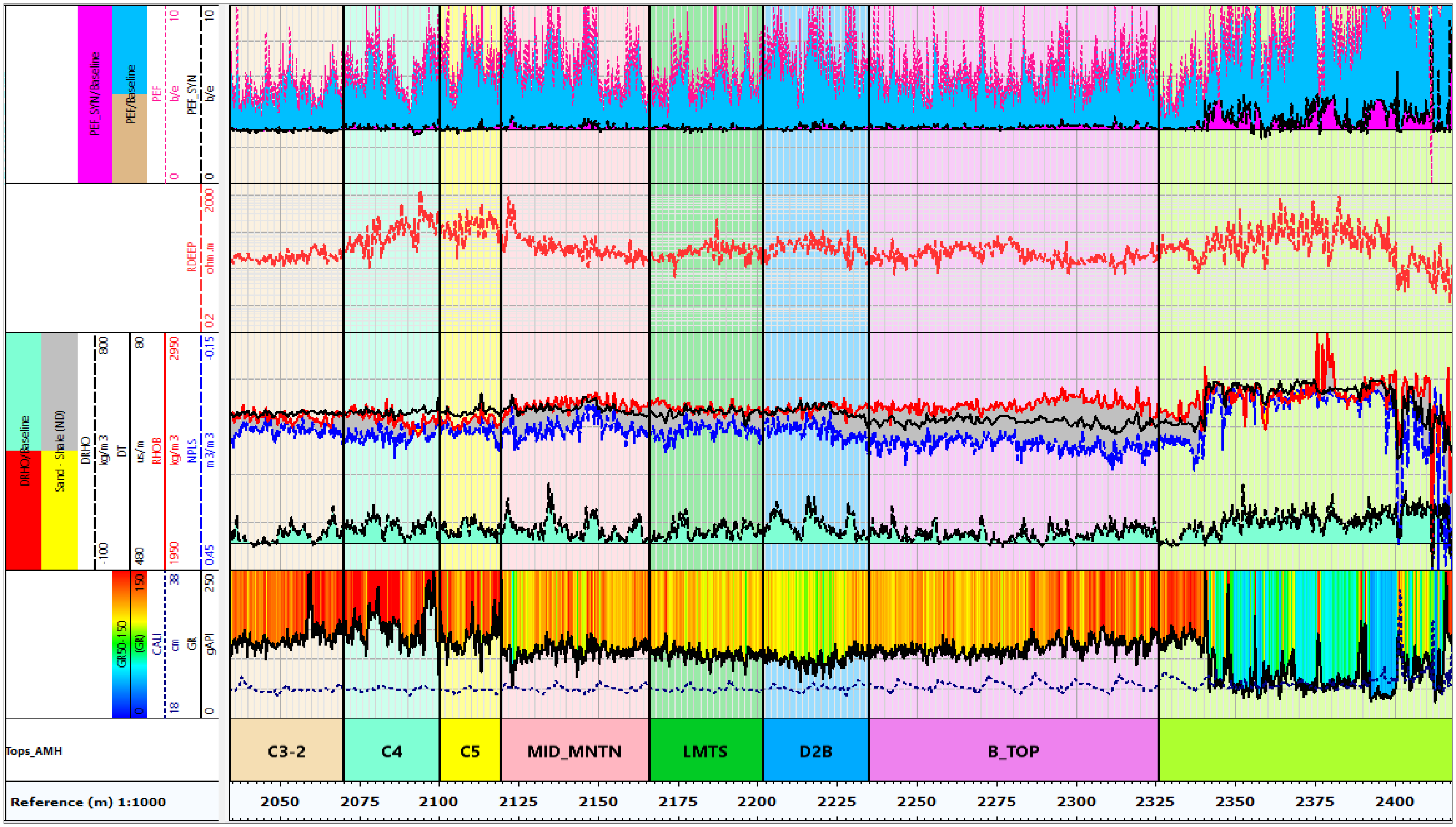

Figure 5: Example of Pef log simulation in *Well B*, showing input logs, the invalid raw Pef log, and the synthetic Pef log.

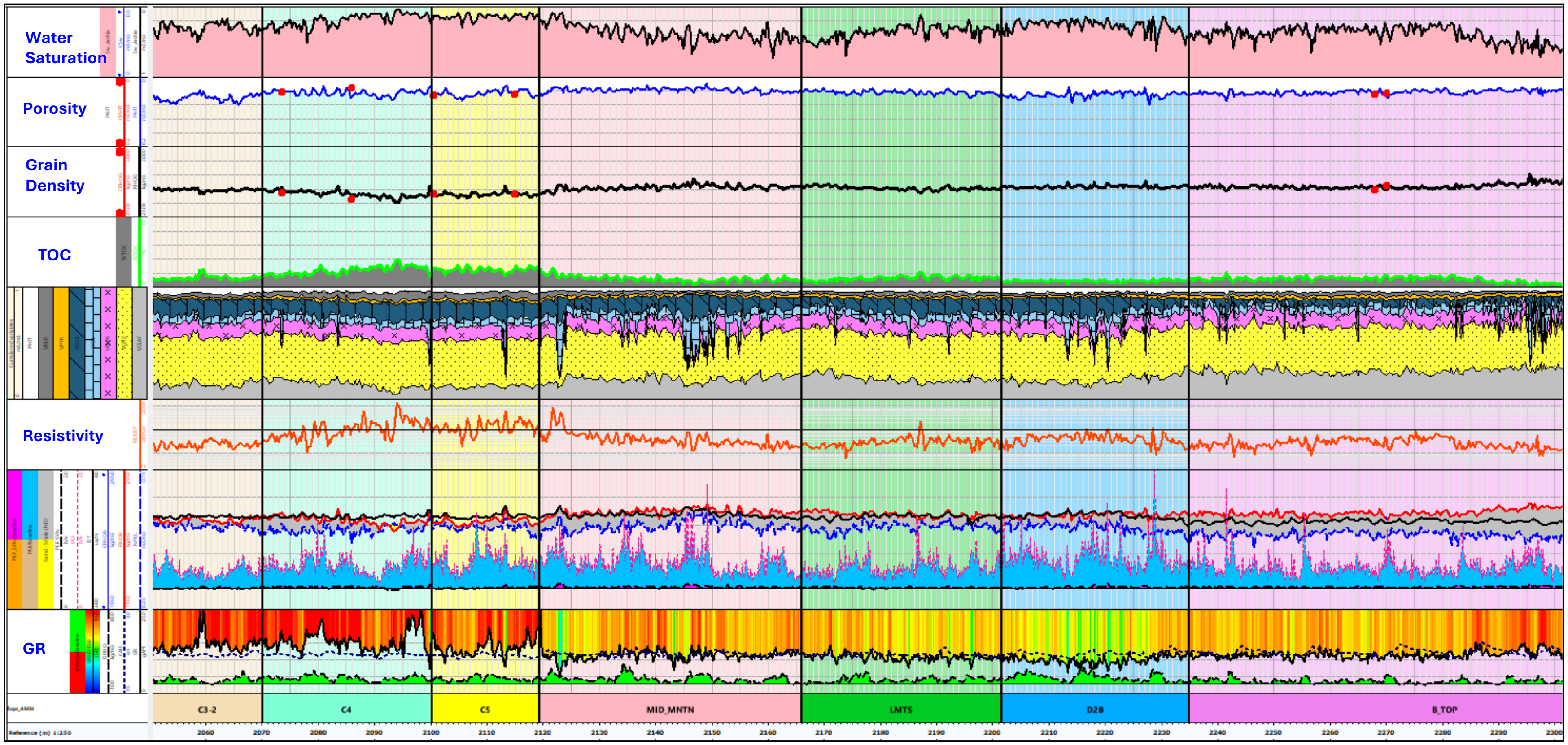

Figure 7: Comprehensive multimineral petrophysical analysis in *Well B* using synthetic Pef log as one of input logs.