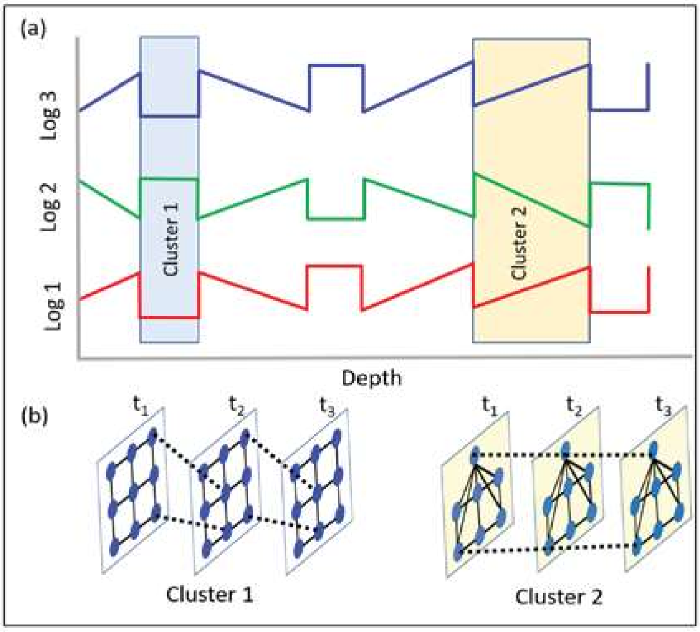
Geophysical Corner

Using Time-Series Analysis and Class-Based Machine Learning to Predict Rock Properties



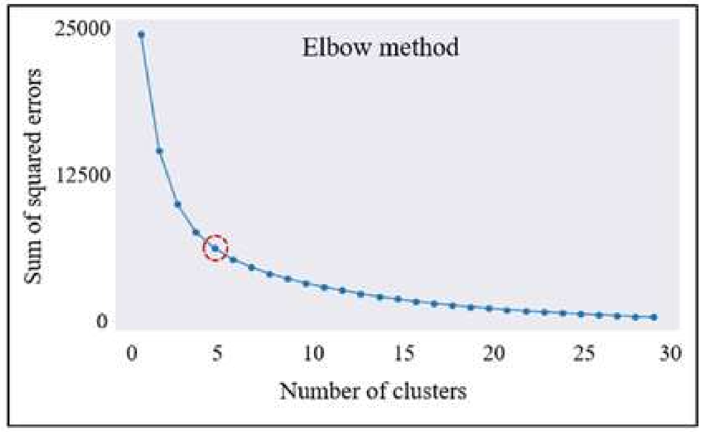
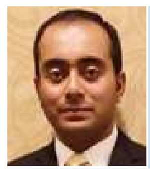


Figure 2: Optimal number of clusters (in red circle) in the current database in the Wolfcamp Formation.

Figure 1: The simplified concept of TICC in synthetic wireline logs (a) and representative Markov Random Field for two clusters (b). Wireline logs (for example, gamma-ray, porosity, and velocity) represent different motifs, such as cleaning-upward, fining- upward, blocky, across different units along the depth.



Shuvajit Bhattacharya is a researcher at the Bureau of Economic Geology at the University of Texas-Austin. He is involved in the Tight Oil Resource Assessment Consortium. His research focus includes seismic interpretation, petrophysical analysis and machine learning for geosciences. Recently, he has authored a book entitled, "A Primer on Machine Learning in Subsurface Geosciences."

With continued stress on the oil and gas industry and the expected mixed-energy economy, geoscientists are being asked to do more with fewer resources. It has become more important than ever to build powerful and insightful machine learning models to predict rock properties faster, cheaper and better, with minimal error. Currently, most published ML models are deployed on the full dataset without incorporating geologic insights into it. It is true that we often do not have prior knowledge about the rocks we drill. Still, how can we use geologic information in building predictive ML models?

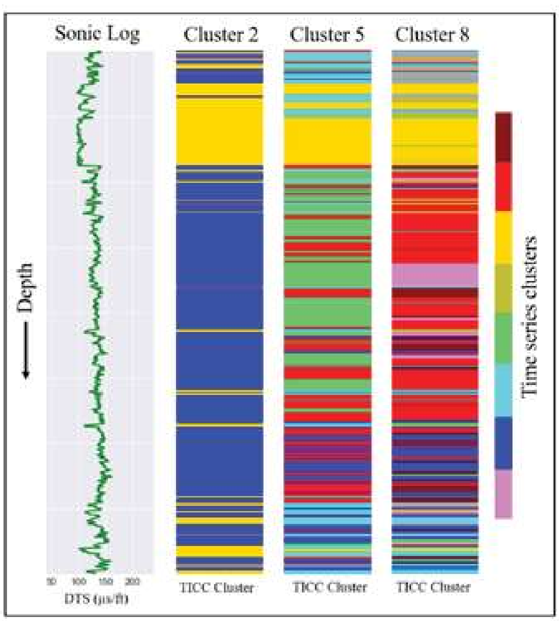
In this article, I show an example of shear sonic velocity prediction in such a situation. Prediction of shear sonic velocity (1/slowness) is of paramount importance in geophysics, petrophysics and geomechanics.

The Method

As part of the Bureau of Economic Geology's ongoing research on energy resources, new methodologies are being developed to better understand the petrophysics of the complex reservoirs and underlying controls. In this study, multivariate time-series clustering, followed by class-based ensemble ML modeling, was used to predict shear sonic slowness for several wells in the Wolfcamp formation of the Permian Basin in the United States.

Multivariate time-series analytics can significantly impact efficient geologic data interpretation, especially in wireline log-based petrophysics and sequence stratigraphy. Wireline logs are basically a form of time series (or depth series). Wireline log motifs exhibit a degree of correlation among themselves. Also, most of the samples on wireline log data are not unique and independent of the previous and subsequent samples; they are the responses of a geologic process over time. Clustering methods such as Toeplitz inverse covariance-based clustering (TICC) can be used to quantify the degree of correlation among these multivariate time-series records and cluster the data based on those graph-based patterns (figure 1). This algorithm assumes the interdependence of log attributes. Each cluster is characterized by Markov random field (MRF) structures. The method was originally developed by David Hallac, Sagar Vare, Stephen Boyd, and Jure Leskovec at Stanford University in 2017. In this case, the clusters derived from TICC contain information about facies (mineralogy and kerogen), texture, fractures and stress information.

Continued on next page ▶



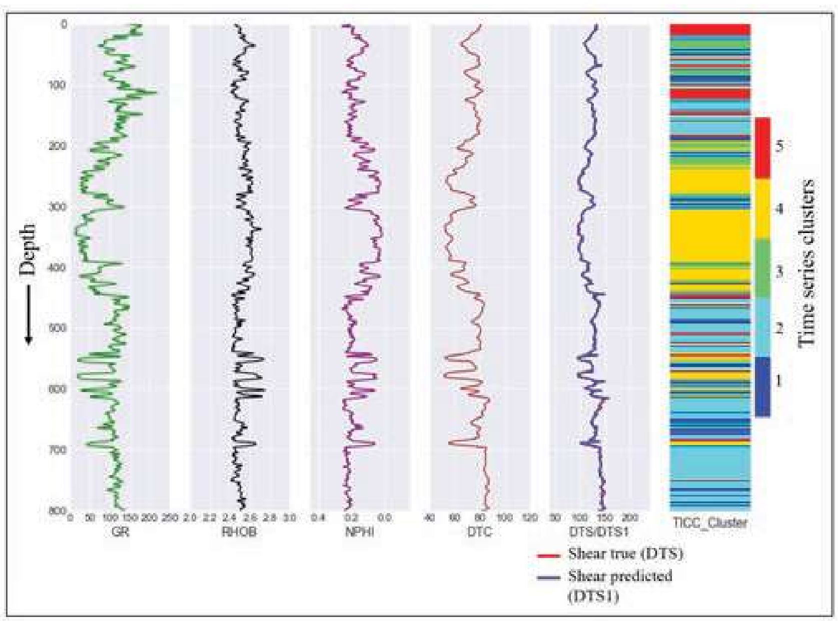


Figure 3: Clusters from various TICC models in the Wolfcamp formation in a well. Figure 4: Shear sonic slowness prediction results from a test well. Wireline logs in the first four tracks are used as an input to the ensemble ML model, and the fifth track from the left shows the shear wave slowness prediction results. Red and blue curves in the sixth track indicate original shear sonic slowness (DTS) and predicted shear sonic slowness (DTS1), respectively.

Transition

from page 15

world that strong and reliable partners like Eneva can help renewables become a reality in Brazil."

Sharing Opportunities

Petersohn and Miranda will join Rodolfo Saboia, ANP general director, Petrobras executives and 10 local and international operators in discussing Brazil's onshore and offshore potential, June 22-23, at the "Optimizing Exploration and Development Opportunities in Brazil's New Regulatory Scenario," a Virtual Technical Symposium organized by AAPG's Latin America and Caribbean Region and the Brazilian Association of Petroleum Geologists.

The symposium features technical presentations and Q&A themes organized into half-day sessions dedicated to four topics: onshore basins and mature fields; post-salt and equatorial margin basins; pre-salt basins and technological advances and applications.

To see the program and register, visit aapg.to/brazil2021. E

◄ Continued from previous page

After clustering, each cluster is trained and tested with different ML algorithms (such as random forest, convolutional neural network, and long short-term memory, among others). Then the results are aggregated to derive the final prediction results of shear wave velocity. Long-short term memory is a deep learning algorithm that uses log responses from previous depth or time to predict the response in subsequent depth. We can run both unidirectional and bidirectional LSTM.

Discussion of Results

Based on the optimal model-based clustering, at least five clusters are found in the Wolfcamp formation in the study area (figures 2 and 3). These clusters are repetitive at a local scale but distinct when the full Wolfcamp formation is considered. This indicates the combined impact of mineralogy (which is repetitive to an extent) and increasing stress gradients along the depth. The ML-based results showed an accuracy of 0.96 in predicting shear wave slowness for a test well (figure 4). Preliminary results showed consistency across several wells in the study area. This indicates that we can predict shear wave velocity with high accuracy in wells where no such log data are available, due to either vintage or cost. The combined application of time-series clustering and ensemble ML modeling makes it possible to predict rock properties successfully. It is still an area of early research. It is an understatement that we need more research on time-series clustering and ensemble class-based ML modeling approaches as it suits geology and data.

The author thanks the sponsors of the TORA consortium at BEG, UT Austin. E

(Editors Note: The Geophysical Corner is a regular column in the EXPLORER, edited by Satinder Chopra, Founder and President of SamiGeo, Calgary, Canada, and a past AAPG-SEG Joint Distinguished Lecturer.)