# Machine learning for predicting stochastic fluid and mineral volumes in complex unconventional reservoirs

*A machine learning workflow can quickly, and accurately, predict mineralogy, porosity and saturation in multiple wells to better understand productive layers in unconventional oil reservoirs.*

■ FRED JENSON, CHIRANJITH RANGANATHAN, SHI XIUPING and TED HOLDEN, CGG GeoSoftware

Determination of mineralogy is a critical step in the petrophysical analysis of many types of reservoirs. Changes in volumes of minerals indicate changes in geological deposition, diagenesis, reservoir quality and brittleness.

Particularly in shale plays, success depends on selecting leases, based on identification of hydrocarbons-in-place within potentially productive layers. Potentially productive layers in a shale play are defined as layers that are sufficiently brittle to respond to hydraulic fracture treatments. A layer's brittleness is determined predominately by its relative mineral ratios, with arenites identified as more brittle, while clay and most calcites are identified as ductile. This methodology of predicting percent brittleness utilizes the ratio of brittle minerals to total mineral volume. This method requires understanding the volume of clay, arenites, calcite, kerogen and heavy minerals, such as pyrite, that are present.

Precise determination of porosity and saturation is also needed for accurate hydrocarbons-in-place calculations. In unconventional reservoirs, the presence of less common minerals, such as kerogen and pyrite, makes it more difficult to calculate mineral and fluid volumes in these complex lithologies. Mineral volumes and porosity in these types of rocks are typically determined using core data or estimated with stochastic methods known as statistical mineral analysis, which works quite

well when using a decent set of wireline logs. Difficulty arises when older boreholes are present in the area of interest, and the log curve data required for sophisticated interpretations are not available.

## PETRO ANALYSIS

The following oil shale case study aims to demonstrate how machine learning models—aided by Python programming extensions using the machine learning ecosystem of CGG GeoSoftware's Power-Log petrophysical interpretation technology—can quickly produce high-quality well log interpretations, including mineral volumes in complex lithologies, for a large number of wells. In this context, for accurate estimates of rock mineral volumes, we propose an effective method, using a

supervised random forest regressor and PowerLog StatMin. This was applied, using tools available in the PowerLog machine learning ecosystem and data stored in a database.

The training datasets use StatMin, a stochastic analysis module in PowerLog, to determine volumes of fluids and minerals, using statistical methods for petrophysical analysis of difficult formations. Porosity, water saturation and complex mineral volumes are valuable outputs often obtained from StatMin. These are the petrophysical parameters needed to train the model.

**Methodology.** The method begins by taking a few modern wells with good curve coverage and creating a solid statistical model using statistical mineral analy-

**Fig. 1.** Representative display of the input curves used in the statistical modeling and regressive analysis.
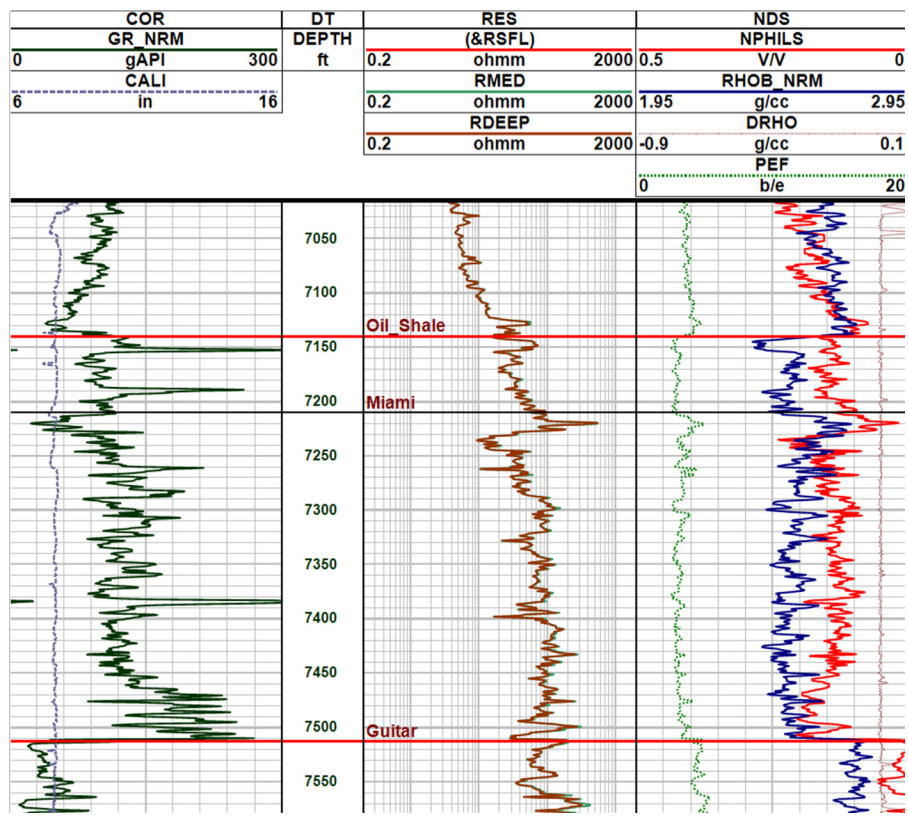
sis methods, such as StatMin in this case. These wells are then used as control wells to create the machine learning model to predict the mineral and fluid volumes for other boreholes that have a more limited set of curve data. Porosity and saturation values are encoded in the resolved fluid volumes. This workflow also could apply with core analysis to create the known mineral distributions. The method was validated using seven wells from the oil shale field:

- The supervised machine learning model was trained with only three of the available wells, while the remaining four wells were used as blind test wells for establishing the model's accuracy.
- To establish a reference, we performed stochastic modeling on all seven wells, using StatMin. Curves used as feature inputs to the regressor were gamma ray (GR), bulk density (RHOB), neutron porosity (NPHI), deep induction (ILD), and photoelectric capture cross section (PEF).

This article addresses the ability to accurately predict StatMin results, using our machine learning method, even when a photoelectric effect (PE) measurement is not available for all predicted wells. Curves used to train the machine learning regressor are GR, RHOB, NPHI, and ILD, measurements common to wells drilled after 1980. Modeled minerals are quartz (sand), calcite (lime), illite (clay), pyrite, and kerogen. Bulk volume water and bulk volume hydrocarbons are also estimated.

**Building the stochastic model.** To compare results from the machine learning model with results from the stochastic model in blind wells, the StatMin model was computed for all seven wells in the study, although only three wells were involved in creating the machine learning model. The remaining four wells not used in training comprised the blind well group. Existing log data for the seven wells are of high quality. **Figure 1** shows curve data for one of the model wells.

Results from statistical computation of mineral and fluid volumes from StatMin were used in creating the machine learning predictive models. The StatMin process uses error minimization forward modeling from known log responses to minerals and fluids to generate volumes of selected minerals and fluids. A main matrix of parameters is used in the analysis of the seven oil shale wells, both the model and the blind test boreholes. The "U" curve in the matrix is the photoelectric adsorption coefficient, calculated from the PE curve, and is a critical factor in determining mineralogy in statistical mineralogy. This PE curve is a modern curve and often not available when analyzing older boreholes. It is important to note that the machine learning models were constructed without using the PE curve. The machine learning method enables evaluation of many older wells not feasible in StatMin, due to lack of PE measurement.

**Machine learning analysis.** The next step in the process used the three selected wells to train the supervised regressor and predict the minerals and volumes on the remaining wells. We trained the model by selecting independent feature vectors including density, gamma ray, deep resistivity, and neutron curves for predicting each mineral type and fluid volume output from StatMin. Predicted volumes from each method were quite similar and very close to volumes and fluids in the control wells. A cross-plot compares modeled kerogen to statistical kerogen volume for the four blind wells, **Fig. 2.**

**Fig. 2.** Cross-plot demonstrates strong correlation of modeled kerogen to statistical kerogen for blind wells.



$$Modeled\ Kerogen = 0.000536613 + 0.941739 * Statistical\ Kerogen$$
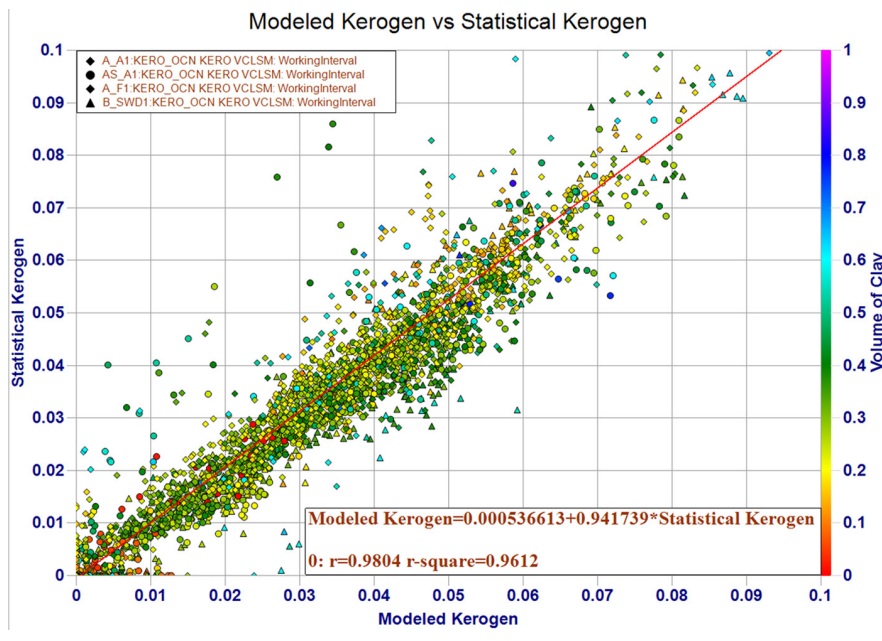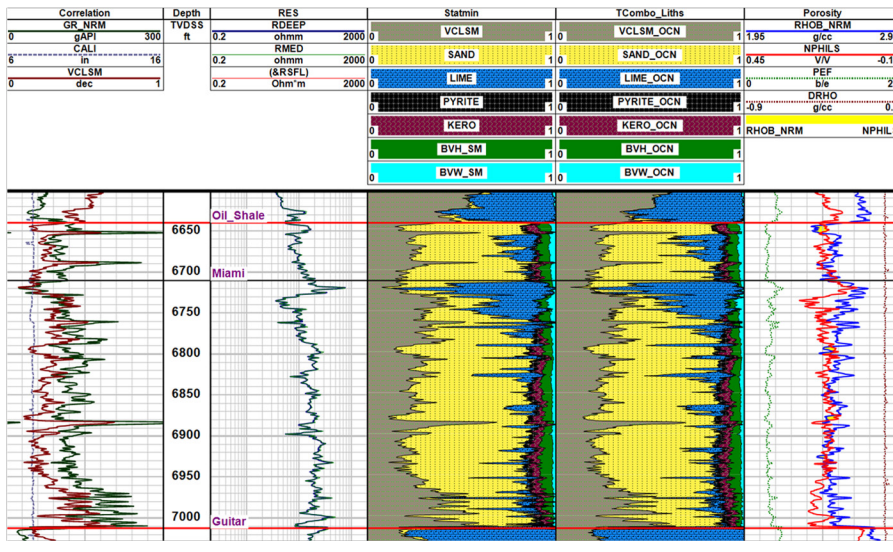
$$0: r = 0.9804\ r\text{-}square = 0.9612$$

**Fig. 3.** Stochastic modeling results compared with machine learning results on a blind test well.

Cross-correlation results between modeled mineral and fluid volumes were as good as, or better than, the match between statistical kerogen and machine learning kerogen. Kerogen was chosen as the example, as it is a critical mineral when evaluating unconventional reservoirs and is one of the lower-volume minerals, where any prediction errors would be magnified. Models for each of the minerals and fluids were generated and applied to the blind test wells. The display in **Fig. 3** compares StatMin results with the predicted results. Visual examination of the log plot demonstrates solid agreement between results from PowerLog StatMin and machine learning-generated results.

## CONCLUSION

It is a common situation when evaluating an established field to have widely disparate data types for each well. Often a few wells in the field are modern and have sophisticated interpretation, providing a high-quality volumetric analysis of fluids and minerals. Occasionally, electron capture spectroscopy tools are available that accurately measure elemental concentra-

tions and provide solid mineral volumes. Extensive core data are also valuable in analyzing complex formations to better understand lithological composition.

This case study demonstrates the value of using high-technology interpretations to build machine learning models with available high-quality log data. We have demonstrated that machine learning models can accurately predict mineralogy, porosity and saturation in unconventional reservoirs. This brings clearer understanding of productive intervals, enabling operators to maximize field-wide potential. **WO**

REFERENCES
1. Pendrel, J. and A.I. Marini, "Static models for unconventional reservoirs: A Barnett shale case study," SEG Summer Workshop Abs., San Diego, California, 2014.
2. Varga, R., R. Lotti, A. Pachos, T. Holden, I. Marini, E. Spadafora and J. Pendrel, "Seismic inversion in the Barnett shale successfully pinpoints sweet spots to optimize wellbore placement and reduce drilling risks," SEG Technical Program expanded abstracts: 1-5, 2012.

**FRED JENSON** has over 40 years of experience in the oil and gas industry. His career has included wireline field engineer, sales engineer, petrophysical consultant, and software development management. Mr. Jenson has worked on field studies in most of the major basins worldwide and has extensive experience with integrated interpretation techniques. He has spent the last ten years with Jason/CGG Geo-Software working on PowerLog development and marketing. Over the last several years, he has been working with the PowerLog Ecosystem exploring machine learning, deep learning and other Python-based workflows to enhance petrophysical interpretations.

**CHIRANJITH RANGANATHAN** is a software architect with CGG GeoSoftware responsible for design, development and enhancements to the machine learning ecosystem of PowerLog. During his 13-year tenure with CGG, he has worked on a wide variety of software problems including several machine learning workflows and applications to address petrophysical challenges. He holds a master's degree in computer science from the University of Houston Main Campus.

**SHI XIUPING** is a petrophysicist with CGG GeoSoftware, based in the China region with responsibility for consulting services as well as software sales and support. With over ten years of petrophysical and rock physics experience across more than 50 conventional and unconventional reservoir projects in the China region, she specializes in petrophysics and rock physics in complex geological conditions, special logging evaluation (image logging, dipole shear wave logging), geomechanics, and machine learning.

**TED HOLDEN** is regional technical manager for CGG GeoSoftware in North America and Latin America. Mr. Holden serves as a senior technical advisor to CGG staff and clients of CGG Geo-Software, mentoring advisors and client staff in seismic petrophysics, rock physics, seismic inversion and quantitative interpretation. During his 45-year career, he has advised client companies on numerous seismic reservoir projects around the globe focused on exploration and development of both onshore and offshore, conventional and unconventional reservoirs.