Automatic rock classification and scaling from core data to well log: Accelerating potential CO_2 storage site characterization using machine learning

Misael M. Morales, Oriyomi Raheem, Carlos Torres-Verdín, and Michael J. Pyrcz, The University of Texas at Austin; Murray Christie and Vladimir Rabinovich, S&P Global*

Summary

We propose a framework for automatic rock classification (ARC) core data to well log scale using machine learning and physics-based methods. The ARC framework is used to estimate rock classes along a well to accelerate the petrophysical interpretation of well logs using core data for rapid formation evaluation and reservoir characterization. The proposed ARC framework compares the estimated rock classes from four different unsupervised machine learning methods and three different conventional physics-based methods. The benefit of this approach is its ability to rapidly estimate rock classes at the log scale from core data without the need for manual interpretation. We validate our method with data from approximately 2400 wells in the Gulf of Mexico. The framework can be used to estimate the spatial distribution of rock classes at the basin scale to identify potential sites for CO_2 storage.

Introduction

Rock classification is a common practical approach to enhance reservoir description, evaluation, modeling, and simulation. Rock classification is also critically important to characterize potential CO_2 storage sites and identify sweet spots based on permeable units bounded by impermeable seals (Bachu et al., 2007). The definition of rock type highly depends on the objective of the characterization. Geological rock typing is based on depositional environments or lithofacies; petrophysical rock typing aims to describe units of petrographic facies or grain and pore types; reservoir and production rock typing aims to identify rock types as flow units. In a reservoir with negligible diagenesis, it is more likely for the rock types of different definitions to match each other (Ali-Nandalal and Gunter, 2003; Acosta et al., 2005). However, when dealing with highly heterogeneous reservoirs, it is common that geological facies, petrophysical rock types, and flow units significantly differ (Rushing et al., 2008; Xu and Torres-Verdin, 2013). In either case, rock classification has been proved to be a valid and effective strategy in characterizing complex subsurface environments.

Rock type can be wholly defined as a group of rock bodies that has acceptable petrophysical regression within each group and can be spatially traceable in line with the geological framework (Neo et al., 1998). Rock classification based on routine and special core analysis has been extensively studied in the literature. Pittman (1992) implements a rock classification method based on Winland's R_{35} metric of pore throat radius distribution. Amaefule et al. (1993) implement the concept of hydraulic flow units based on the Hagen-Poiseuille equation and derive a flow zone indicator to characterize heterogeneous reservoirs. For rock classification at a reservoir or basin scale, Gunter et al. (1997) implement the Stratigraphic Modified Lorenz (SML) coefficient for characterizing reservoir flow units. Other methods based on mercury injection capillary pressure (MICP) data, pore throat size distribution, and thin sections have also been studied (Al-Aruri et al., 1998; Neo et al., 1998; Clerke et al., 2008).

However, rock classification based on core measurements need to be propagated to the uncored zones along the well, to other wells in the reservoir without core data, or to a basin scale by spatially correlating the interpreted wells. For the case of uncored wells, rock classification can be based on well logs or as a spatial distribution of cored wells. Once a classification of rock types is established along a well or spatially along a reservoir or basin, petrophysicists can infer rock properties for a complete characterization of the formation (Bennis and Torres-Verdin, 2019; Raheem et al., 2023).

We propose an automatic rock classification (ARC) framework based on machine learning and physics-based methods to rapidly scale from core measurements to well log scale and detect potential CO_2 storage sites in the Gulf of Mexico. The ARC framework serves as a quick tool for rock classification and a comparison of different methods with minimal user intervention. We validate the framework on a field dataset with approximately 2400 cored wells in the Gulf of Mexico.

Method

The ARC framework is designed as a tool to aid petrophysicists and reservoir engineers in the characterization and interpretation of well, field, and basin scale modeling. Using only core measurements of porosity and absolute permeability, the ARC framework clusters the core measurements into distinct classes based on the properties of the core data and propagates the rock classes along the depth interval of the well. This allows for rapid core-to-log scaling and potential multi-well formation evaluation and field or basin-scale modeling by tracing similar zones along multiple wells.

The first step in the ARC framework is to process the core data. We have a total of 2442 wells with core measurements but select the wells with at least $N = 30$ core samples, where N is selected arbitrarily to obtain representative rock classes and correlate between wells. This results in 1379 wells with at least 30 core samples of porosity and absolute permeability. Figure 1 shows the spatial distribution of the preprocessed cored wells.

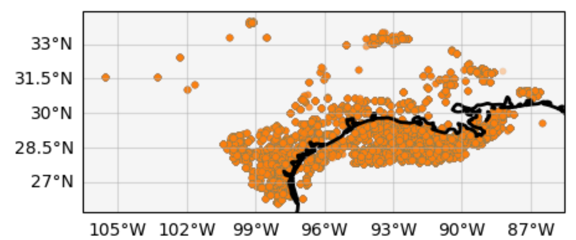

Figure 1: Spatial distribution of the 1379 cored wells with at least 30 core samples. Each orange dot is a well, and the black solid line delinates the coastline along the Gulf of Mexico.

The three physics-based rock classification methods implemented in the ARC framework are Leverett's equivalent pore throat radius, Winland's R_{35}, and the Stratigraphic Modified Lorenz Coefficient. The details of each are explained below:

Leverett's equivalent pore throat radius method separates rock classes by marking cutoffs between different sizes of pore throats, r, where $r^2 \cong k/\phi$. This assumes that each rock class possess the same pore throat radius despite variability in porosity or permeability, and the only distinction is an increase or decrease in pores. To implement Leverett's rock classification within ARC, the user must specify the cutoff values, $m_i = \left\{\sqrt{k/\phi}\right\}_{i=1}^{N_c}$, where N_c is the total number of classes.

Winland's R_{35} method separates rock classes by selecting cutoffs between different sizes of pore throats based on a pore throat size distribution function, $m_i = \{R_{35}\}_{i=1}^{N_c}$. Moving beyond Leverett's method, Winland's technique assumes that pore throats are heterogeneous in a rock and can be described with a distribution. Furthermore, it assumes that the 35th percentile of the pore throat size distribution serves as the most meaningful metric for rock typing. However, Winland's method assumes unimodal and well-behaved distributions.

The Stratigraphic Modified Lorenz (SML) coefficient quantifies the heterogeneity in a rock based on the cumulative flow ($\tilde{k}$) and cumulative storage capacity ($\tilde{\phi}$). The cutoff values are selected by comparing the change in the magnitude of the slope such that $m_i = \left\{d\tilde{k}/d\tilde{\phi}\right\}_{i=1}^{N_c}$.

When using physics-based methods for rock classification, it is possible to obtain interpretable rock properties from the inferred classes to support any petrophysical interpretation. However, physics-based methods require manual selection of the cutoff values to separate the rock classes. Figure 2 shows a crossplot of the core porosity and absolute permeability values along with the Leverett, Winland, and SML rock classes for a randomly selected well and the defined cutoff values for 3 rock classes.

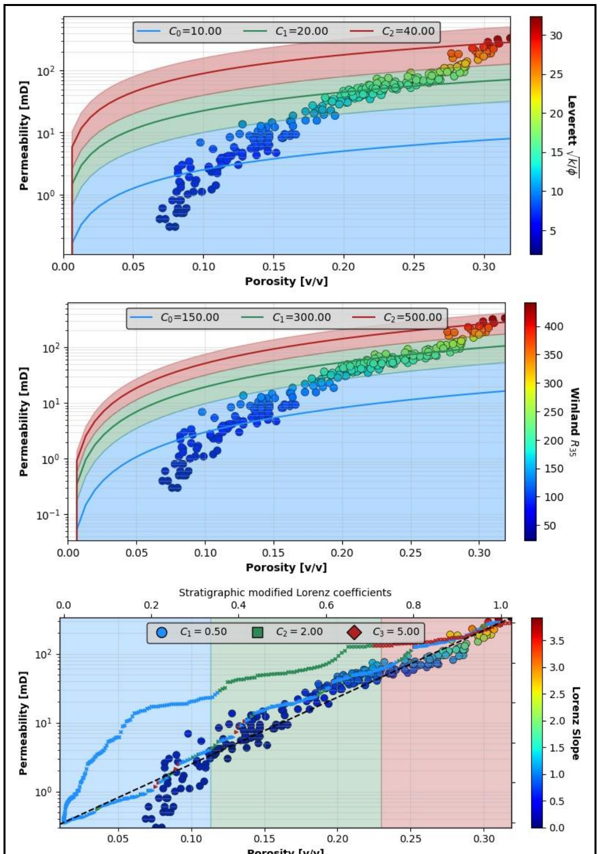

Figure 2: Crossplot of core porosity and permeability for a randomyl selected well. The physics-based rock classification methods are used to estimate 3 rock classes of low (blue), medium (green), and high (red) reservoir quality. The Leverett method (top) has cutoffs $\left\{\sqrt{k/\phi}\right\} = \{10,20,40\}$, the Winland method (middle) has cutoffs $\{R_{35}\} = \{150,300,500\}$, and the Lorenz method (bottom) has cutoffs $\left\{d\tilde{k}/d\tilde{\phi}\right\} = \{0.5,2.0,5.0\}$.

Using the estimated rock classes from core measurements, the next step is to propagate the rock classes from the core data to the log scale. Let c_i represent the class value assigned to the i^{th} data point. Let d_i represent the depth of the i^{th} data point. Let $\boldsymbol{z}$ represent the depth index along the well log. Let

$\mathcal{C}$ represent the estimated rock class value. Initialize $\boldsymbol{z} \equiv \underline{0}$. Then, for every depth along the well, z_i, we can estimate the corresponding rock class as follows:

$$c_i = \begin{cases} c_{i-1}, & if\ z_i = 0 \\ \mathcal{C}(\min|d_i - z_i|), & otherwise. \end{cases} \quad (1)$$

The well-scale rock classes, $\boldsymbol{c} = \{c_i\}_{i=1}^{N_w}$, now have a well log resolution while honoring the inferred rock classes from the core measurements. Figure 3 shows the core porosity and absolute permeability measurements along the depth of the well, along with the scaled rock classes from core to log support for the same randomly selected well using the Leverett, Winland, and Lorenz physics-based methods.

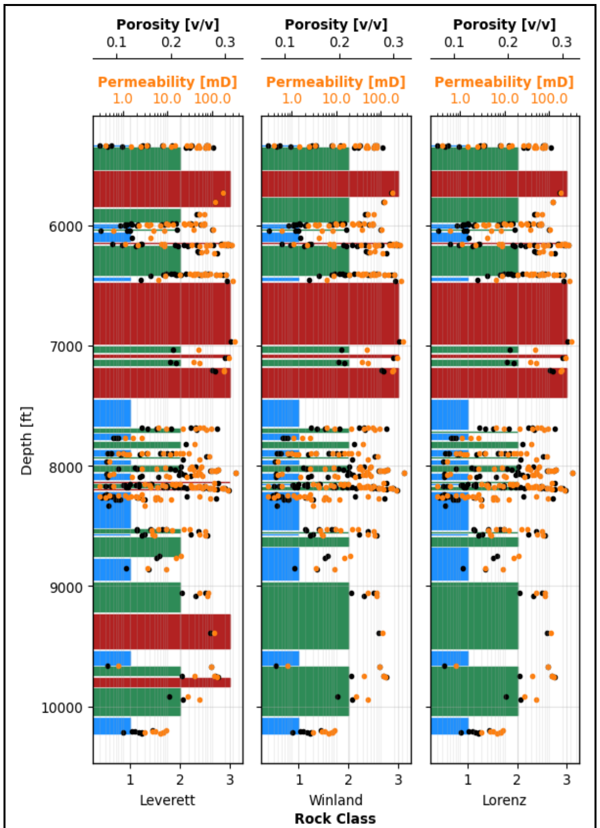

Figure 3: Estimated rock classes at the well log scale using Leverett (left), Winland (center), and Lorenz (right) physics-based methods for a randomly selected well in the Gulf of Mexico. The core measurements of porosity (black) and absolute permeability (orange) are displayed as a function of depth along the well.

However, physics-based methods require some degree of petrophysical interpretation and subjectivity in order to select the optimal cutoff points to partition the rock classes. To overcome this, the ARC framework uses unsupervised machine learning methods to estimate rock classes from core data and propagate the estimated classes to log scale.

The user can select between four different unsupervised machine learning methods to estimate the rock classes: K-Means, Bisecting K-Means, Gaussian Mixture Models (GMM), and BIRCH. K-Means clustering aims to minimize inertia, or within-cluster sum of squares by separating samples into K groups of equal variances. Bisecting K-Means is an iterative variant of K-Means that applies a divisive hierarchical clustering strategy for batches of data at a time. GMM implements the expectation-maximization algorithm to fit mixtures of Gaussian kernels to the data and estimate the probability of belonging to a class. BIRCH, which stands for balanced iterative reducing and clustering using hierarchies, lossy compresses the data into a hierarchical tree that divides into subclusters. The subclusters are then iteratively collected into K clusters based on their within-cluster similarity and between-cluster dissimilarity. Figure 4 shows the crossplot of core porosity and permeability and the estimated rock classes using the four machine learning methods, and Figure 5 shows corresponding log scale estimated rock classes.

A single machine learning method can be used to estimate rock classes or an average class from the four methods can be calculated. The ARC framework is implemented on all 1379 wells available with core data in the Gulf of Mexico to obtain a basin-scale estimation of rock classes for sweet spot detection using the average estimated classes, as shown in Figure 6. The entire process of loading, processing, normalizing, training, predicting, and saving the well-scale rock classes for all wells requires only 47.2 seconds, or approximately 0.03 seconds per well.

Further petrophysical interpretation is recommended, and the ARC framework should only be used as a screening tool for rapid identification of sweet spots along the depth of a well and at a basin scale. The physics-based methods are limited by the fact that the user has to define cutoff values to separate rock classes, and these vary well-to-well. On the other hand, machine learning methods are automatic and computationally efficient but might violate assumptions of petrophysical properties and geologic continuity. Future work includes implementing a hybrid physics-informed machine learning framework to automatically estimate the cutoff values of the physics-based methods using machine learning techniques.

Conclusions

We presented a framework for automatic rock classification (ARC) based on core measurements scaled to the well log support. The framework is used to compare physics-based techniques with unsupervised machine learning methods for rock classification using core data from the Gulf of Mexico. Furthermore, the framework allows for rapid estimation of rock classes along the depth of a well, and at a basin scale by correlating well-to-well rock classes. The framework can be implemented as a tool for rapid characterization of potential CO_2 storage sites, either vertically or spatially, by detecting sweet spots of high-quality reservoir rocks in terms of storage and flow capacity.

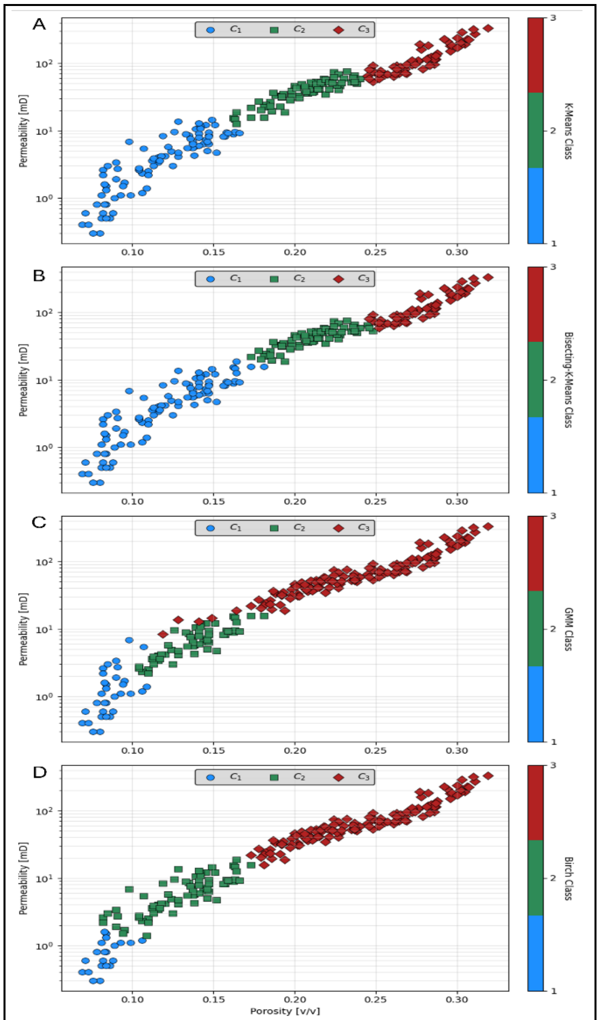

Figure 4: Crossplot of core porosity and permeability and estimated rock classes scale using (A) K-Means, (B) Bisecting K-Means, (C) GMM, and (D) BIRCH unsupervised machine learning methods for a randomly selected well in the Gulf of Mexico.

Acknowledgements

This work is supported by the Digital Reservoir Characterization (DiReCT) consortium at the University of Texas at Austin.

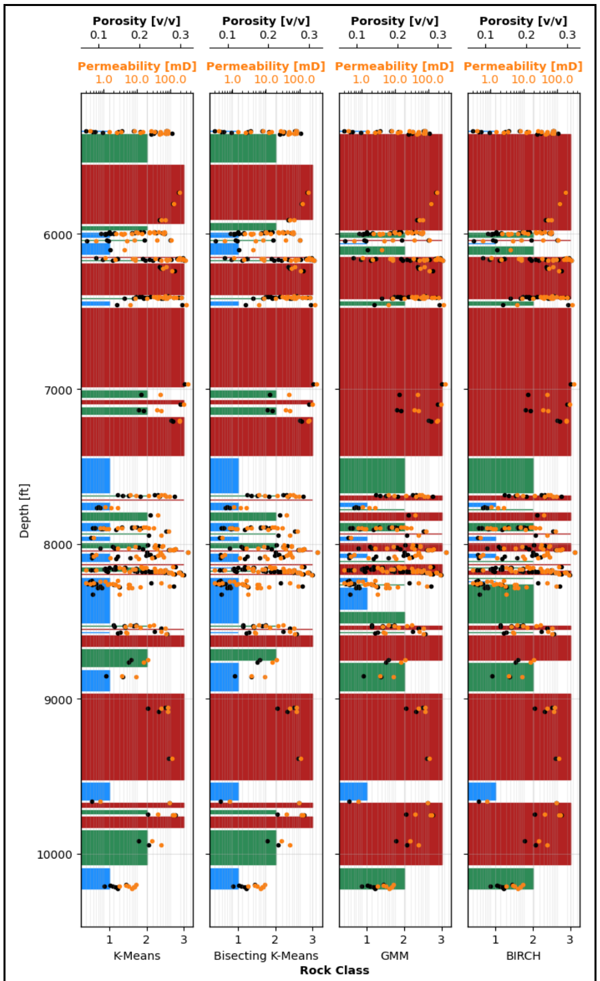

Figure 5: Estimated rock classes at the well log scale using K-Means, Bisecting K-Means, GMM, and BIRCH machine learning methods for a randomly selected well in the Gulf of Mexico. The core measurements of porosity (black) and absolute permeability (orange) are displayed as a function of depth along the well.

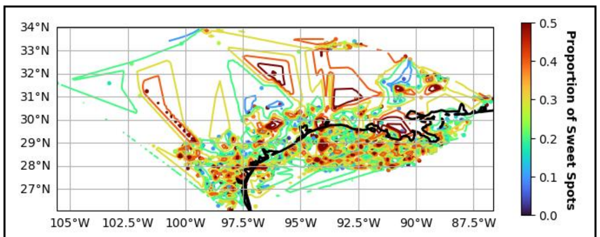

Figure 6: Spatial distribution of the proportion of sweet spots for each of the 1379 wells with at least 30 core samples using the K-Means machine learning method.

References

Acosta, L., E. Marin, E. Labastidas, J. Bello, J. Jimenez, P. Cordoba, J. C. Pascual, G. Auxiette, Y. Gou, and B. Thorsen, 2005, Reservoir Study V9 of El Furrial Field, Venezuela, *in* SPE Latin America and Caribbean Petroleum Engineering Conference: p. SPE–95047.

Al-Aruri, A., F. B. Ali, H. A. Ahmad, and S. A. Samad, 1998, Rock Type and Permeability Prediction from Mercury Injection Data: Application to a Heterogeneous Carbonate Oil Reservoir, Offshore Abu Dhabi (United Arab Emirates), *in* Abu Dhabi International Petroleum Exhibition and Conference: p. SPE–49556.

Ali-Nandalal, J., and G. Gunter, 2003, Characterising reservoir performance for the mahogany 20 gas sand based on petrophysical and rock typing methods, *in* SPE Latin American and Caribbean Petroleum Engineering Conference.

Amaefule, J. O., M. Altunbay, D. Tiab, D. G. Kersey, and D. K. Keelan, 1993, Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells, *in* SPE Annual Technical Conference and Exhibition.

Bachu, S., D. Bonijoly, J. Bradshaw, R. Burruss, S. Holloway, N. P. Christensen, and O. M. Mathiassen, 2007, CO2 storage capacity estimation: Methodology and gaps: International Journal of Greenhouse Gas Control, v. 1, no. 4, p. 430–443, doi:10.1016/S1750-5836(07)00086-2.

Bennis, M., and C. Torres-Verdin, 2019, Estimation of dynamic petrophysical properties from multiple well logs using machine learning and unsupervised rock classification, *in* SPWLA Annual Logging Symposium: p. D053S015R004.

Clerke, E. A., H. W. Mueller III, E. C. Phillips, R. Y. Eyvazzadeh, D. H. Jones, R. Ramamoorthy, and A. Srivastava, 2008, Application of Thomeer Hyperbolas to decode the pore systems, facies and reservoir properties of the Upper Jurassic Arab D Limestone, Ghawar field, Saudi Arabia: A "Rosetta Stone" approach: GeoArabia, v. 13, no. 4, p. 113–160.

Gunter, G. W., J. M. Finneran, D. J. Hartmann, and J. D. Miller, 1997, Early determination of reservoir flow units using an integrated petrophysical method, *in* SPE Annual Technical Conference and Exhibition? p. SPE–38679.

Neo, S., J. Asada, N. Fujita, S. Mohammed, and H. Arab, 1998, Geological framework modeling and rock type optimization for a giant oil field, *in* Abu Dhabi International Petroleum Exhibition and Conference.

Pittman, E. D., 1992, Relationship of porosity and permeability to various parameters derived from mercury injection-capillary pressure curves for sandstone: AAPG bulletin, v. 76, no. 2, p. 191–198.

Raheem, O., W. Pan, C. Torres-Verd\'\in, and M. M. Morales, 2023, Best Practices in Automatic Permeability Estimation: Machine-Learning Methods vs. Conventional Petrophysical Models, *in* SPWLA Annual Logging Symposium: p. D041S015R001.

Rushing, J. A., K. E. Newsham, and T. A. Blasingame, 2008, Rock typing—Keys to understanding productivity in tight gas sands, *in* SPE Unconventional Resources Conference/Gas Technology Symposium: p. SPE–114164.

Xu, C., and C. Torres-Verdin, 2013, Core-based petrophysical rock classification by quantifying pore-system orthogonality with a bimodal Gaussian density function, *in* Paper SCA2013-079 presented at International Symposium of Society of Core Analysts. Napa Valley, California, September: p. 16–19.