Application of Data Analytics and Geostatistics for Integrated Shale Facies Modeling

Shuvajit Bhattacharya¹ and Timothy R. Carr²

¹Battelle Memorial Institute ²West Virginia University

ABSTRACT

Application of data analytics and geostatistics are becoming increasingly popular to analyze subsurface data and better understand geologic history at multiple scales. The upper and lower shale members in the Bakken Formation of the Williston basin in North Dakota are selected for this study. The objective of this study is to classify, predict, and model shale lithofacies at core, well, and regional scales for better understanding of depositional history of shale lithofacies. Based on core samples and advanced geochemical spectroscopy logs, a quantitative workflow is designed to classify lithofacies, in terms of mineralogy, Total Organic Carbon, and petrophysical properties. Five different shale lithofacies are identified, such as organic mudstone, organic siliceous shale, gray siliceous shale, gray mixed shale, and gray mudstone. Support Vector Machine, a data analytics algorithm, is used to identify the pattern of different lithofacies, associated with petrophysical parameters from conventional well logs from ~500 wells. Bayesian network theory is applied to understand the causalities of petrophysical relationships to lithofacies. The classification accuracy of different lithofacies is more than 87%. After core and log-based lithofacies classification and prediction, Sequential Indicator Simulation, a geostatistical algorithm, is used to generate 3D stochastic geocellular lithofacies models spanning an area of ~13,000 sq. miles. The results show that both upper and lower Bakken shale members are vertically and laterally heterogeneous at multiple scales. Several factors, such as source of elements, paleo-redox conditions, and organic matter productivity etc. appear to have controlled the depositional pattern of different shale lithofacies.