Rapid Exploration in a Mature Area in Northwest Kansas: Improving the Definition of Key Reservoir Characteristics using Big Data

William (Bill) Full[1] and Steve James[1]

[1]*GXStat LLC*

Abstract

With the rapid changing economic environments associated with defining new viable oil and gas plays in a mature region, attention has turned to the statistical analysis of all of the available data (i.e. big data). These studies have been on basin-wide through individual play scales. For the most part, the tools used are statistical, geostatistical and multivariate in nature. Oftentimes, the user is either given a toolset found within a larger program or works with one of the many fine programs available on the market. Even with a deep understanding of the myriad of assumptions as-associated with of these approaches, it is difficult to ex-tract all but the most obvious results from the data let alone the more thorny questions on how to quantify the economic risk.

A common workflow is to gather in the best data avail-able (e.g. geologic, geophysical, geochemical, log-based), create multiple layers/surfaces of geostatistically mapped information and then perform some appropriate multivariate analysis. A great many of the assumptions associated with the common multivariate techniques are based on the necessity of the data being derived from either one or a fixed number of known populations. With big data, the verification of these assumptions are often overlooked resulting in statistically ambiguous or difficult to validate results. An extra step in the workflow needs to be added with these cases - partitioning the data in an appropriate way and analyzing each partition separately before recombining to produce a final risk map. Recognizing when this partitioning is needed requires visual, statistical, geostatistical and deterministic techniques, as described below.

This study consisted of a large data set of well-based and geophysical data (gravity and magnetic) in sever-al counties in northwest Kansas. In this area, the early Paleozoic rocks are likely dominated by basement tectonics at the time of deposition and the later Paleozoic formations appear overlie the earlier rocks including their related fault/fracture zones. After recognizing the visual hints that data partition was appropriate, computer programs designed for data partitioning (Polytopic Vector Analysis-based programs including Hyperplanar Vector Analysis, Fuzzy and Hard Clustering including Fuzzy N-Varieties) were applied. The results showed that partitioning the data produced a more refined probability of success than could be de-fined by the multivariate analysis alone. Key variables tied to reservoir quality were also defined that have been used to increase understanding of both new prospects and potentially increase reservoir production for known fields. This increased knowledge directly leads to a more confident economic risk assessment in a mature area.