Application of Principal Component Analysis in Oil-Oil Correlations of Paleozoic Oils from Williston Basin

Mark Obermajer*, Kirk G. Osadetz
Geological Survey of Canada, Calgary, AB T2L 2A7

and

Nikos Pasadakis
Technical University of Crete, Chania 73100 Greece

Several statistical models have been used in the study of well defined Paleozoic petroleum systems of the Williston Basin to investigate if Principal Component Analysis (PCA) has a potential for delineating petroleum compositional families. As applied here, this multivariate statistical analysis used variables derived from gasoline range and saturate fraction gas chromatography data. Both fractions form the bulk of oil composition and carry important information about source and post-generative alteration of petroleum.

In general, the oil families defined by unique biomarker signatures show sufficiently distinctive compositions of their light fractions to be recognized using variables derived from PCA. However, while the separation of the Ordovician-sourced family A oils is quite strong using any of the employed statistical models, multiple models, pair-wise models and additional derived variables had to be utilized to maximize distinctions among B, C and D oil families from Givetian-Tournaisian source rocks. Greater variations in light hydrocarbon distributions within these three families most likely result from their more heterogeneous sources, more complex migration patterns and compositional mixing. Furthermore, inter- and intra-familial linear and non-linear variations exhibited by PCA models are not only controlled by differences in kerogen composition and source rock depositional environment but also reflect oil mixing, especially between oil families B and C.

Despite their lack of primary diagnostic capability for familial classification, PCA of these two fractions can enhance oil-oil correlations and the interpretation of petroleum systems, especially when combined with information from more complex biological marker compounds.