Application Of Canonical Correlation Analysis In Studies Of Geochemical Compositions Of Crude Oils From Williston Basin

Nikos Pasadakis*
Mineral Resources Engineering Department Technical
University of Crete, 73100 Chania, Greece
pasadaki@mred.tuc.gr

and

Kirk G. Osadetz and Mark Obermajer
Geological Survey of Canada, Calgary, AB

ABSTRACT

The common experimental practice in geochemical characterization of crude oils comprises analyses of several oil fractions, including light hydrocarbons (C5-C8) and heavier (C12-C40) saturate and aromatic compounds. The interpretations of analytical data sets are often carried out separately what may lead to contradictory results. The possibility of inter-relationship between analytical data derived from these different compositional fractions of oil has not been studied in details.

In this work, we present the results of the Canonical Correlation Analysis (CCA) applied to compositional data of 50 family C oil samples from Mesozoic reservoirs occurring on the southwest extension of the Alberta-Saskatchewan heavy oil belt. These results are then compared with CCA data for 87 Mississippian family C oils from southeastern Saskatchewan. Biomarker and gasoline range compositional data were included in the analyses to examine if a compositional correlation exists between different oil fractions.

CCA is a multivariate technique that aims to identify the relationships among two or more sets of variables, measured on the same population of samples. This task is accomplished by deriving for each data set a weighted linear combination of maximally correlated original variables called canonical variates. The interpretation of the canonical variates is accomplished by linking together the variables in each data set, which exhibit the highest correlation to the respective canonical variates.

The CCA results indicated that within oil family C certain biomarker ratios as well as individual components are highly correlated to the compositions of the light fraction. These results can be used to refine the influence of source and post-generative processes on the composition of the light hydrocarbon fraction in family C oils.