Multivariate Analysis of Circum-Arctic Crude Oils

L. Scott Ramos and Ken E. Peters
Infometrix, Bothell, WA 98011 • USGS, Menlo Park, CA 94025

ABSTRACT

A regional study of Circum-Arctic crude oil samples was undertaken to better understand the distribution and geochemical properties of petroleum systems in this vast area. The objectives of this study were to classify crude oil samples into genetic groups using geochemical data (oil-oil correlation) and to create an automated chemometric method to classify additional oil, seep, and source-rock extract samples.

Selected source-related biomarkers and isotope ratios were used as input variables. Over 1000 oil samples from source rocks spanning the Circum-Arctic region were analyzed by GC-MS to obtain the biomarker ratios (Peters et al., 2005). Approximately 622 samples were retained for training the models, and the remainder were set aside for evaluation. Using HCA and PCA as guides, training set oil samples were separated into 31 genetically different source-rock groups. A multi-tiered decision tree, based on KNN and SIMCA models of these groups, was built to classify new samples by source-rock type.

In addition, oils from the Prudhoe Bay region on the North Slope of Alaska were studied to assess volumetric contributions from different source rocks. Contrary to conventional geochemical interpretations, results from analysis with ALS indicate that the Prudhoe Bay Field contains mixtures of oil that originated mainly from Cretaceous Hue Shale with secondary input from Triassic Shublik Formation source rocks.

OBJECTIVES

- Use source- and age-related biomarker and isotopic data to correlate crude oils into genetic groups and build predictive statistical model
- Infer source-rock age, lithology, organic matter input, and depositional environment
- Determine the extent of mixing of oil from Shublik carbonate and younger argillaceous source rocks

Figure 1 Locations of oil samples from the Circum-Arctic region above ~55° N latitude, where 31 petroleum systems were identified in this study. Yellow polygon identifies the North Slope and Beaufort-Mackenzie areas where the oil partitioning study was done.

METHODS

Oils were separated into saturated and aromatic fractions by column chromatography, using hexane and dichloromethane. Stable carbon isotope ratios were determined on a Finnigan Delta E mass spectrometer. Sterane and terpane biomarkers were derived from analysis on a HP 5890 GC (HP-2 column, 50 m x 0.2 mm, 0.11-µm film, programmed from 150 to 325°C at 2°C/min) coupled to a HP 5971 MS (selected ion mode). Chemometric analyses were performed with Pirouette® and InStep™ software (Infometrix).

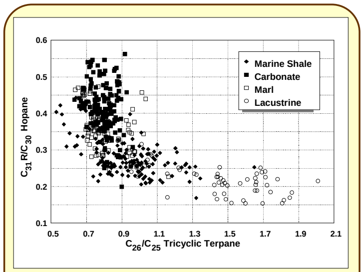

Figure 2 Biomarker ratios of oil samples from the Alaskan North Slope can be used to infer lithologies of their source rocks. For example, comparing hopane and terpane ratios distinguishes lacustrine from marine carbonate sources.

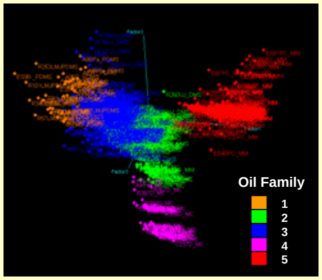

Figure 3 Principal components analysis (autoscale) was run on the training set containing data from 20 biomarkers and isotope ratios (14 terpane ratios, 4 sterane ratios and 2 isotopes). Five main families of oils were indicated by the PCA scores.

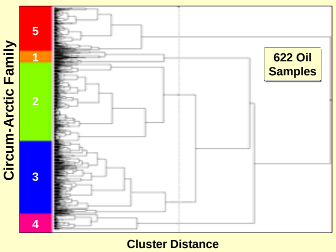

Figure 4 HCA with the same data as Figure 3 shows 5 clusters corresponding to the PCA score groups.

RESULTS

Classification

A KNN model was developed to classify an oil into one of the five major families. Each family of oils was further refined into groups and subgroups by evaluating the clusters indicated by PCA and HCA and knowledge of the geochemical parameters. Each group was analyzed by KNN, and the misclassifications helped to fine tune the group assignments and eliminate outliers. Finally, a multi-tiered hierarchical decision tree (Ramos, 1994) was developed with exclusive SIMCA models that assign a level of certainty for each group assignment. The decision tree can be used to classify newly acquired crude oil, seep, and source-rock extract samples.

Evaluation of the decision tree was performed with a separate data set composed of 290 oils extracted from the original data set in families 2 and 3. All 290 oils were correctly predicted into their families.

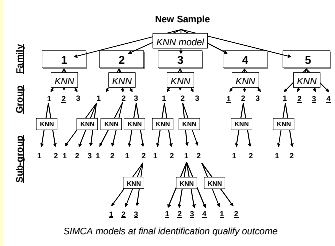

Figure 5 Multi-tiered chemometric decision tree constructed using data for 622 training set samples allowed identification of five major oil families and 31 genetically distinct oil groups within the Circum-Arctic study area.

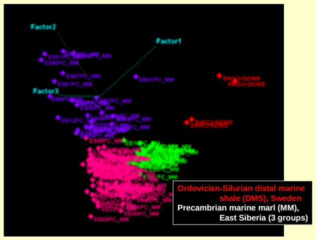

Figure 6 Oils from Family 5, when analyzed by PCA, exhibit four smaller subgroups that can be characterized by source-rock age, lithology and geographic location.

CONCLUSIONS

Selected source-related biomarker and isotope ratios can be used to organize crude oils into a predictive model. Further, this model holds true when assessing new samples, thus allowing detailed genetic classification across a vast area such as the Circum-Arctic region. The selected geochemical parameters also enable us to understand the relative proportions of input into an oil field and gives an objective means to apportion contributions from different source rocks.

Apportionment

Oil samples from the North Slope (see Figure 1) were characterized with alternating least squares (non-negativity and closure constraints). Solutions with three and four sources were evaluated using both biomarker ratios and concentrations as input.

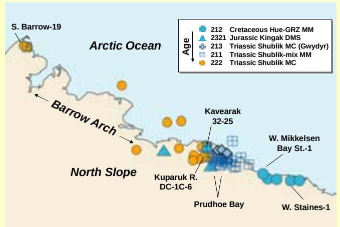

Figure 7 North Slope oil samples identified by location and genetic group (see Figure 1).

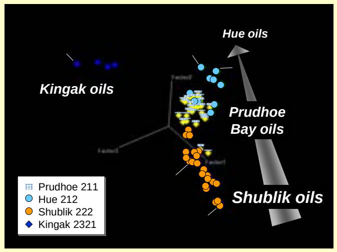

Figure 8 Principal components scores for North Slope oil samples indicate that oil in the giant Prudhoe Bay Field consists of mixtures contributed by Hue Shale and Shublik Formation source rocks with little input from the Kingak Shale.

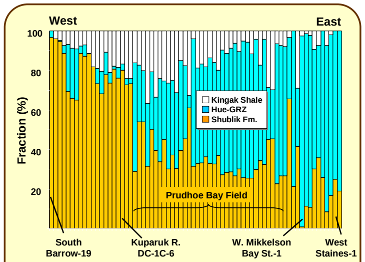

Figure 9 Amounts of oil sources inferred by alternating least squares analysis based on biomarker and isotope ratios. Samples are arranged from west to east as in Figure 7. Hue Shale contributes more input to Prudhoe Bay oils than the Shublik Formation. Similar results are obtained using concentrations rather than ratios.

REFERENCES

Peters, K.E., C.C. Walters, and J.M. Moldowan, 2005, *The Biomarker Guide*. Cambridge University Press, Cambridge, U.K., 1155 pp.
Ramos, L.S., *J. Chromatogr. Sci.*, 32:219-227 (1994).