Fuel 236 (2019) 1093–1099

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel

Full Length Article

Origin based classification of crude oils by infrared spectrometry and chemometrics

Amir Bagheri Garmarudi[a], Mohammadreza Khanmohammadi[a], Hassan Ghafoori Fard[b], Miguel de la Guardia[c,*]

[a] Chemistry Department, Faculty of Science, Imam Khomeini International University, Qazvin, Iran
[b] Electrical Engineering Department, Amir-Kabir University of Technology, Tehran, Iran
[c] Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, E-46100 Burjassot, Valencia, Spain

ARTICLE INFO

Keywords:
Crude oil
Geochemical origin
Classification
Infrared spectrometry
Chemometrics
SIMCA

ABSTRACT

Crude oil samples from different Iranian petrol resources in both, raw and mixture forms have been characterized by attenuated total reflectance mid infrared spectroscopy. Obtained spectra were classified by chemometric techniques to propose a method for geological based classification of crude oil samples. Totally 251 samples from 7 petrol fields and 3 mixtures were analyzed. Mean centering and principal component analysis (PCA) supported – leverage value based outlier detection were used as preprocessing approaches. PCA, cluster analysis and soft independent modeling of class analogy (SIMCA) were utilized to classify the spectra. Obtained results confirmed that SIMCA is a robust chemometric technique for origin classification of crude oil samples based on their IR spectra, while the mixture samples were also classified satisfactory in some cases. Root mean square error, method precision and regression coefficient for the prediction of origin of an independent validation set of 111 samples were 1.41%, 96.7% and 0.957 respectively.

1. Introduction

Crude oil (petroleum) is the main source of fossil fuels around the world. The current amount of consumed quantities of crude oil is estimated about 90 million barrels per day. Considering the important role of energy in the life stream of the society, crude oil could be mainly assumed as a valuable natural resource which is primarily consumed for production of fuels, such as diesel, gasoline, kerosene, jet fuel, etc. On the other hand, many synthetic chemicals e.g, paraffin, bitumen and also some polymers are synthesized, using crude oil as the raw material [22,27].

Crude oil from different sources is consumed in refineries directly or as a mixture of different origin oils with different ratios. Construction, development and planning of refining plants are severely related to the source of the raw material by which the refinery is fed. Different varieties of crude oils are traded all around the world, which most of them are mixtures of several sources. Thus, the input of the refineries or petrochemical plants is diverse in chemical and physical characteristics. The environmental-geological condition and local ecologic circumstances during the formation of crude oils are the variables affecting the chemical composition of crude oils. The main fraction of crude oil is hydrocarbons e.g. linear, chained, cyclo and polycylic aliphatics together with aromatic compounds. Generally the main chemicals of the crude oil are classified as paraffins, aromatics, naphtenes and asphaltenes [14]. Quality monitoring and characteristics control of crude oil is of high importance for different reasons. In order to evaluate the potential applicability of a crude oil resource, and also to explore its economical justification level, one needs to characterize the extracted samples, comparing a multivariate model of exploring cost, quality, geographic location and application. The investigation of the chemical composition and quality parameters of crude oil is also of high importance for selection of appropriate downstream processes and refining procedure. Determination of physicochemical profiles of crude oils from different origins is a common practice in oil industries. The outcome of such assay is useful for different aims, such as substituting the product of different resources or oil fields with most similarities and also for monitoring the consistency of physicochemical parameters as a measure of quality stability along time. Most of the laboratory tests for characterization of crude oil are based on standard methods developed by the American Society for Testing and Materials International (ASTM International) and Institute of Petroleum (IP). In order to investigate the similarity or differences of two or more samples from different origins, several experimental tests must be performed. However, most of standard test methods are rather slow, time consuming, environmental non-

* Corresponding author.
E-mail address: Miguel.Delaguardia@uv.es (M. de la Guardia).

https://doi.org/10.1016/j.fuel.2018.09.013
Received 11 November 2016; Received in revised form 2 September 2018; Accepted 5 September 2018

friendly, elaborate and expensive, requiring huge amounts of the crude oil sample to be analyzed. Thus it would useful to approve some supportive approaches by which rapid reliable evaluation of crude oils is possible while it is environmental friendly and it can be performed with few amounts of sample. Such approaches with be useful together with standard test methods for more detailed comparative investigations. Infrared spectroscopy (IR) has been employed already for crude oil analysis and different qualitative and quantitative methods for crude oil analysis have been developed [19,15,4]. Most of related studies deal with the quality assessment, characterization and classification of crude oil samples [13,7,16,17,24,11] and in most of the related researches, chemomectric multivariate analysis has been also utilized in order to propose reliable models for industry [12,23,10].

One of the important tasks in petroleum industry is to distinguish the characteristics of crude oils, based on their geological origin. It is well-known that the chemical characteristics of crude oil are dependent on its origin. Thus it is of high importance for the petroleum industry to employ a method by which the origin of new sample could be predicted or verified in a fast way without performing several tests. Previous efforts in this field have been started by introducing chromatographic procedures for quantitative determination of different chemical structures in crude oil by gas chromatography [8], gas chromatography with mass spectrometry detector [26] and liquid chromatography [2]. Application of infrared spectrometry for this aim has been reported for analysis of bitumen samples from 4 different origins by near infrared (NIR) and data processing by soft independent modeling of class analogy (SIMCA) and artificial neural networks (ANN) [3]. Mid infrared (MIR) spectrometry has also been utilized for qualitative and quantitative analysis of crude oil samples from Algeria, Brazil and Venezuela by chemometric data mining [1]. Recently the same strategy has been employed for classification of crude oil samples from 4 different origins of Africa and South America [9]. In this study it has been evaluated different alternative chemometric methods to identify the geochemical origin of Iranian crude oil samples based on ATR-FTIR spectrometry.

2. Experimental

2.1. Apparatus and materials

A total number of 251 crude oil samples were obtained from different Iranian petroleum fields. Samples were distributed in 10 classes (7 crude oil classes from different fields and 3 combination classes which 1 one of them was a mixture of different sources for export and 2 were mixtures used as the input feed of Iranian refineries). In this regard, classes number 1, 3, 4, 5, 6, 7 and 8 were from Ahwaz, Foruzan, Hengam, Behregan, Nowruz, Siri and Soroush petroleum resource fields of Iran respectively; class number 2 was the mixture for exportation and classes 9 and 10 were the combination mixtures, used as refinery feed. Crude oil samples were provided within a 6 months time period in order to consider the time dependent variation in chemical structure of crude oil samples obtained from a single field. Attenuated total reflectance infrared (ATR-FTIR) spectroscopy measurements were carried out at room temperature on a Tensor-27 Bruker FT-IR spectrometer aquipped with a Ge–KBr beam splitter, a DTGS detector and Beer–Norton apodization with a horizontal, fixed path ATR device (ZnSe, 45°, single reflection). The spectral resolution was 8 cm^{-1} and 64 scans were accumulated over the range from 600 to 4000 cm^{-1} for each spectrum. The spectrum of each sample was recorded 3 times and the average was used to be processed. Wavelength penetration and baseline corrections were utilized. Chemometric data processing was performed by using MATLAB Ver. 8.0.

2.2. Standard crude oil evaluation procedures

There are some standard test procedures in the petroleum industry utilized to evaluate the physico-chemical characteristics of crude oil samples. Most of the test methods are based on reliable ASTM standards. All of the samples were analyzed by a list of experimental test, summarized in Table 1. Samples obtained from all of the 7 oil fields were in consistence with the general technical data sheet in the archive of the Iranian petroleum industry, while those results of the samples from the other remaining combination classes (3 mixtures) showed some deviations.

Table 1
List of main experimental tests, performed for characterization of crude oil in the Iranian petroleum industry.

Specification	Std. Test Method
API	ASTM D-4052
Ash Content	ASTM D-482
Asphaltene Content	IP-143
Carbon Residue	ASTM D-189
H_2S Content	ASTM D7621
Kinematic Viscosity	ASTM D-445, D-341
Mercaptan	UOP-163
Pour Point	ASTM D-97
Resin Content	SARA
Sulfur Content	ASTM D-2622
Specific Gravity	ASTM D-4052
Salt Content	ASTM D-664
Total Acidity	UOP – 163
Vapor Pressure	ASTM D-323
Wax Content	BP-237
Water Content	ASTM D-95

2.3. Chemometric procedures

In the first step, the data set of 251 spectra was mean centered, detecting the outlier objects based on leverage values. The outcome after outlier detection was a set of 211 samples whole. Principal component analysis (PCA) was performed to determine the number of PCs in the set, estimating the overall distribution of samples within its plot. The applied classification approaches were cluster analysis (CA) and soft independent modeling of class analogy (SIMCA). CA was performed on the whole data set with in 10 classes. In the case of SIMCA, a calibration model was built based on 100 samples, and the remaining 111 samples were used a validation set being predicted their origin by using the calibration model.

3. Results and discussions

3.1. Mid-IR spectral characteristics of crude oil

Crude oil is a mixture of different chemical structures while the main components are hydrocarbons e.g. linear, chained, cyclo and polycylic aliphatics together with aromatic compounds. Generally the main chemicals of the crude oil are classified as paraffins, aromatics, naphtenes and asphaltenes and hence they may influence directly the characteristics of a crude oil sample spectrum. Polyaromatic molecules with aliphatic or naphthenic side chains are some of the polar components of crude oil samples, which clearly influence their IR spectra. Differences in chemical composition of crude oils affect their refining procedure, quality parameters and finally marketing. So, crude oil characterization and crude assay are important parts of petroleum industry. On the other hand, provision of knowledge about similarities and difference of crude oil samples from different origins would also enable the petroleum industry for identification of alternatives for a product. Infrared spectroscopy could be proven to be capable of this aim as a rapid, economic, and non-destructive method for identifying crude oil samples. It is known as a simple and practical approach by which the analysis of crude oil samples was investigated, based on their informative spectral features in the mid-IR region Table 2 shows the assignments of some of the typical bands which can be identified in