Chemometric differentiation of natural gas types in the northwestern Junggar Basin, NW China

Yao-Ping Wang¹,²,³, Xin Zhan⁴,⁵, Yuan Gao¹,²,³, Sibo Wang¹,³, Jia Xia¹,³ and Yan-Rong Zou²

Abstract
In recent years, the natural gas has displayed a growing significance in oil and gas exploration in the northwestern Junggar Basin (NWJB), although oil has been the main focus of exploration in the basin. Here, we systematically discuss the classification and origin of the natural gases from the NWJB based on the natural gas geochemistry and chemometric methods. The natural gases collected from the NWJB were chemometrically classified into three groups. Group A gases, defined as coal-derived gases, were likely generated from the mixing of the Jiamuhe Formation and Carboniferous strata. Group B gases, defined as the mixing of coal-derived and oil-associated gases, were restricted to the source rocks of group A and C gases. Group C gases, defined as oil-associated gases, were likely derived from both the Fengcheng and Wuerhe Formations, with a higher contribution from the latter strata. The result of this study suggests that the potential of oil generation in the Wuerhe Formation has been underestimated in the past. This is in accordance with geochemical and geological evidence. This study provides an effective chemometric method of natural gas classification and evaluation of hydrocarbon generation potential. This contributes to a better understanding of the origin of gases and distribution of oil and gas, assisting in exploration deployment in the basin.

¹College of Chemistry and Environmental Science, Guangdong Ocean University, Zhanjiang, P.R. China
²State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, P.R. China
³Shenzhen Research Institute, Guangdong Ocean University, Shenzhen, P.R. China
⁴Missouri University of Science and Technology, Rolla, USA
⁵CNOOC International Limited, Beijing, P.R. China

Corresponding author:
Yao-Ping Wang, Guangdong Ocean University, No. 1 Haida Road, Mazhang District, Zhanjiang 524088, China.
Email: wangyp@gdou.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Keywords
Junggar Basin, natural gas classification, principal component analysis, multidimensional scaling, alternating least squares

Introduction
Chemometrics is a useful tool for the analysis of multivariate data. With noise being identified and removed from the data, useful information can be extracted using chemometrics to show affinities and variations of samples (Kramer, 1988; Peters et al., 2005). Therefore, chemometric methods have been widely used in many disciplines (Bevilacqua et al., 2017; Chabukdhara and Nema, 2012; Madsen et al., 2010), especially in petroleum geochemistry, they have been used for a long time (Kvalheim et al., 1985; Øygard et al., 1984; Peters et al., 1986; Zumberge, 1987). In addition, principal component analysis (PCA) has been frequently applied to reveal the genetic families of crude oils (He et al., 2012; Peters et al., 2013, 2016) in order to make a detailed oil-source rock correlation (Mashhadi and Rabbani, 2015; Wang et al., 2014, 2018). Additionally, multidimensional scaling (MDS) and alternating least squares (ALS) are reliable tools for oil-oil correlation (Wang et al., 2016, 2018) and the quantitative evaluation of the relative contribution from each source to the mixed oil in basins with multiple source rocks (Peters et al., 2008; Zhan et al., 2016a, 2016b). Moreover, PCA has been confirmed to be an effective method for natural gas classification on the basis of chemical composition and stable carbon isotope of gases in previous studies (Wang et al., 2019).

Recently, the abundant oil and gas reserves in the northwestern Junggar Basin (NWJB) have received increasing attention of scholars (Chen et al., 2014; Tao et al., 2016), but information about the genetic types and origins of natural gases in this area remains controversial. Chen et al. (2014) recognized four different types of gases in the Zhongguai Uplift of the NWJB. These types of gases include sapropel-type gas sourced from the Fengcheng Formation (P1f) mudstone, humic-type gas derived from the Wuerhe Formation (P2w) mudstone, mixed sources of sapropel- and humic-type gases, and mixed sources of sapropel-type gas and deep inorganic gas. Gao et al. (2016) performed an analysis on the genetic types of natural gases from the Wuxia Fault Zone of the NWJB. The result suggests the existence of one gas type that is derived from the Fengcheng Formation (P1f). Chen et al. (2016b) proposed that the NWJB gases can be genetically classified into oil-associated gas, coal-derived gas, and mixed gas. The oil-associated gas may have been originated from the Fengcheng Formation (P1f), and the coal-derived gas may have been mainly sourced from the Jiamuhe Formation (P1j) while part of the contributions may have been derived from the high-maturity source rock of the Wuerhe Formation (P2w). The mixed gas may have been mainly derived from the Wuerhe Formation (P2w), mixed with the contributions of the highly mature source rock from the Fengcheng Formation (P1f). Tao et al. (2016) conducted a detailed research on the geochemistry and origin of natural gases in the Mahu Sag of the NWJB and concluded that three genetic types of natural gases exist in the study area: an oil-associated gas sourced from the Fengcheng Formation (P1f), a coal-derived gas sourced from Carboniferous strata and the Jiamuhe Formation (P1j), and a coal-derived gas sourced from the Wuerhe Formation (P2w). However, systematic studies on the natural gas classification and origin of the NWJB are available only to a limited extent (Chen et al., 2016b).
Moreover, only few studies have reported on natural gas classification using chemometric methods (Wang et al., 2019), and the applications of MDS and ALS in natural gas exploration have not been reported. Here, taking the NWJB as an example, we further study on the natural gas classification and origin of the NWJB using chemometric methods.

**Geological setting and geochemical characteristics of potential source rocks**

The Junggar Basin is in the northern part of Xinjiang Uygur autonomous region, northwestern China (Figure 1(a)). It is divided into four oil–gas zones, namely northwestern, central depression, eastern, and southern parts (Figure 1(b)), based on structural units and the relative concentration of oil and gas. At present, the major oil reservoirs in this basin, such as the Karamay Oilfield in the Mahu Sag, are in the NWJB (Tao et al., 2016). Exploration of natural gas shows promising potential in the NWJB with a total reserve of \( \sim 2.1 \times 10^{12} \) m\(^3\) estimated by the latest third-round assessment of petroleum resources, although the current focus of exploration is primarily on oil (Chen et al., 2014).

The NWJB mainly contains 13 primary sedimentary sequences (Figure 2). They are listed in ascending order as follows: Carboniferous (C); Permian Jiamuhe Formation (P\(_1j\)), Fengcheng Formation (P\(_1f\)), Xiazijie Formation (P\(_2x\)), lower Wuerhe Formation (P\(_2w\)); Triassic Baikouquan Formation (T\(_1b\)), Kelamayi Formation (T\(_2k\)), Baijiantan Formation (T\(_3b\)); Jurassic Badaowan Formation (J\(_1b\)), Sangonghe Formation (J\(_1s\)), Xishanyao Formation (J\(_2x\)), Toutunhe Formation (J\(_2t\)), and Cretaceous Tugulu group (K\(_1tg\)). The Carboniferous strata are characterized mainly by basalt and tuff. The Permian strata are characterized mainly by mudstone, sandstone, argillaceous siltstone, and tuff in the P\(_1j\); mudstone, sandstone, dolomite, and argillaceous dolomite in the P\(_1f\); mudstone, argillaceous siltstone, and sandy conglomerate in the P\(_2x\); and mudstone and sandy conglomerate in the P\(_2w\). The Triassic and Jurassic strata consist mainly of mudstone, argillaceous siltstone, siltstone, sandstone, sandy conglomerate, conglomerate, carbonaceous mudstone, and coal whereas the Cretaceous strata contain mainly mudstone, siltstone, and sandstone.

The Permian and Carboniferous strata are essential source rocks for crude oils and natural gases in the NWJB (Cao et al., 2005; Graham et al., 1990). The Permian source rock is mainly composed of a set of widespread mudstone and siltstone in various depositional sags of the basin (Wang et al., 2013). The Permian source rocks of the NWJB exist in the lower Wuerhe (P\(_2w\)), the Fengcheng (P\(_1f\)), and the Jiamuhe (P\(_1j\)) Formations (Cao et al., 2005; Pan et al., 2003). The Carboniferous source rock in the NWJB is mainly dark volcanic mudstones in marine, swamp, and lagoon environments. All these sets of source rocks have the potential for hydrocarbon generation, but the kerogen type and maturity level are slightly different. The P\(_2w\) and P\(_1f\) have great potential to generate oil with a relatively low maturity, while the P\(_1j\) and Carboniferous strata are source rocks of gas with a relatively high maturity (Chen et al., 2016a; Gao et al., 2016; Meng and Parhati, 1999; Tao et al., 2016).

**Samples and methods**

In this study, we used 71 gas samples published in previous studies (Chen et al., 2014; Tao et al., 2016; Yang et al., 2008) to classify types of natural gas in the NWJB (supplementary Appendix A; Figure 1(c)). The gas samples from previous publications were screened as
Figure 1. Geological maps of the Junggar Basin, showing (a) the locations of the basin, (b) the study area, and (c) structural units and oil and gas accumulations in the NWJB. NWJB: northwestern Junggar Basin. Source: Modified after Chen et al. (2014) and Tao et al. (2016).
Figure 2. Generalized stratigraphy of upper Paleozoic and Mesozoic strata in northwestern Juggar Basin, emphasizing prospective petroleum source rock with geochemical characteristics.
Source: Modified after Chen et al. (2016b), Gao et al. (2016), Cao et al. (2005), Meng and Parhati (1999), and Wang et al. (2013).
follows before the data set was determined: (1) gas samples with abnormal values of their geochemical composition were excluded and (2) the mean values of proxies obtained from the same well and depth were used for chemometric analysis. After data collection, PCA and MDS were applied to identify the types of natural gas in the NWJB, and ALS was introduced to de-convolute the mixed oil-associated gases. The chemometric analysis includes six geochemical parameters: C1(%), C2(%), C3(%), δ13C1, δ13C2, and δ13C3, which are similar to the parameters used in a previous study (Wang et al., 2019). The PCA was performed using commercial software (Pirouette 4.5, Infometrix Inc.), and the MDS was performed with in-house software (Wang et al., 2016). The PCA settings were programmed as follows: preprocessing = range scale, maximum factor = 3, validation method = none, and row = none. The MDS settings were the same as that for the PCA, but the Bray–Curtis distance was used for the measurement of similarity (Wang et al., 2016). The ALS settings were programmed as follows: preprocessing = range scale, maximum sources = 2, and initial estimates from = rows.

Results and discussion

Natural gas classification using PCA and MDS

The results of the PCA of the six selected proxies of chemical and carbon isotopic compositions showed three genetically distinct natural gas groups: A, B, and C (Figure 3(a)). These inferences are similar to the results of a previous study that a δ13C2 value is used to determine the genetic types of natural gases (Dai et al., 2005). Similar results are obtained from the MDS plot, indicating that this method is also effective for natural gas classification (Figure 3(b)). Gases of groups A and B occurred mainly in the Zhongguai Uplift, whereas those of group C occurred mainly in the Mahu Sag and adjacent uplifts (e.g. Ke-Bai and Wu-Xia Fault Zones). Distinct gas source rocks of each group can be inferred from different ranges of chemical and carbon isotopic compositions (Table 1).

Group A gases were composed of 12 samples, collected from the Permian reservoir in the P2w and P1j formations in the Zhongguai Uplift. The gases in this group were characterized by a relatively high content of methane, ranging from 90.51% to 95.63%, with a mean value of 93.61%. The δ13C2 and δ13C3 values of group A gases were the highest among the three groups, ranging from –27.3‰ to –23‰ and –26.6‰ to –20.1‰, respectively. Thus, group A gases were defined as coal-derived gases because the δ13C2 values of such gases are commonly higher than /C0 27.5‰ (Dai et al., 2005).

Group B gases consisted of 11 samples collected from the Triassic and Permian reservoirs in the T1b, P1t, P2w, and P1j formations in the Zhongguai Uplift, Wu-Xia Fault Zone, and Mahu Sag. In terms of the mean values of chemical and carbon isotopic compositions, group B gases had a moderate content of methane and δ13C2 and δ13C3 values similar to those of the gases in group A and C. This probably indicates a mixed source of the gases in group B (Dai et al., 2005), which is also supported by results of the PCA and MDS plots, indicating that the gas samples of this group take an intermediate position between the gases of group A and group C.

Group C gases, the most common types of gases in the NWJB, contained 48 samples. The gases of this group were distributed in the Triassic, Permian, and Carboniferous strata of the entire study area. The gas was characterized by a low content of methane from 64.68% to 95.28%, with a mean value of 82.03%. In addition, the δ13C2 and δ13C3 values of group C
gases were relatively lower in comparison with those of the gases in groups A and B, and therefore, these gases are oil-associated (Dai et al., 2005).

### Possible gas-source correlations

Overall, three groups of natural gases showed a wide range of chemical and carbon isotopic compositions, suggesting that they were likely derived from multiple sources. Notably, another significant characteristic of natural gases in the NWJB is the partial reversal in stable carbon isotopes between propane and butane (Figure 4). This can be well explained by mixed sources for the three group gases. Dai et al. (2004) proposed that partially reversed
δ^{13}C orders for natural gas alkanes may have been caused by (1) bacterial oxidation, (2) mixing of biogenic and thermogenic gases, (3) mixing of gases from two source rock intervals or one source rock unit with different maturity, and (4) mixing of sapropelic and humic gases. The δ^{13}C_{1} values for all of the gases in the NWJB were in the range of −54.4‰ to −26.6‰, indicating that the natural gases were unlikely that they originated from a significant mixing of bacterial oxidation or biogenic and thermogenic gases. Therefore, the most plausible cause for the partial isotope reversal in the NWJB gases is the mixing of gases from two source rock intervals or one source rock unit with different maturity or mixing of sapropelic and humic gases.

In accordance with the results of Tao et al. (2016), geochemical characteristics of group A gases were similar to those of the source rocks in the Jiamuhe Formation and Carboniferous strata based on biomarker proxies of the condensates from wells (e.g. the Ke75 and Ke77). This is also supported by the kerogen types of source rocks in the Jiamuhe Formation and Carboniferous strata (Figure 2). Moreover, the carbon isotopic composition of kerogens in different source rock intervals indicates a predominance of Type-II kerogen in source rocks of the Fengcheng Formation and Type-III kerogens in source rocks of the Jiamuhe Formation and Carboniferous strata (Figure 5). Therefore, group A gases were likely derived from the Jiamuhe Formation and Carboniferous strata. However, the contribution of source rocks in the Wuerhe Formation to group A gases cannot be inferred and needs further research.

Group B gases were most likely derived from the mixing of sapropelic and humic gases because of the moderate δ^{13}C_{2} values, ranging from −29.1‰ to −26.8‰ (Dai et al., 2005). In addition, group B gases were mainly distributed in the middle of group A and C gases as shown in Figures 3 and 4, suggesting that group B gases were probably derived from a mixture of source rocks from group A and C gases.

Table 1. Statistics of chemical and carbon isotopic compositions in group A, B, and C gases in the NWJB.

| Chemical compositions | CH_{4} | C_{2}H_{6} | C_{3}H_{8} |
|-----------------------|--------|-----------|-----------|
| Group A               | 90.51 to 95.63 | 1.77 to 3.74 | 0.39 to 1.09 |
|                       | 93.61 (12)    | 2.61 (12)   | 0.74 (12)  |
| Group B               | 86.64 to 94.08 | 2.10 to 5.62 | 0.68 to 3.34 |
|                       | 91.72 (11)    | 3.24 (11)   | 1.20 (11)  |
| Group C               | 64.68 to 95.28 | 1.54 to 12.42 | 0.52 to 9.12 |
|                       | 82.03 (48)    | 6.35 (48)   | 3.57 (48)  |
| Carbon isotopes       | δ^{13}C_{1} | δ^{13}C_{2} | δ^{13}C_{3} |
| Group A               | −35.1 to −26.6 | −27.3 to −23 | −26.6 to −20.1 |
|                       | −31.5 (12)    | −25.8 (12)  | −23.7 (11) |
| Group B               | −43.3 to −33.0 | −29.1 to −26.8 | −29.4 to −24.4 |
|                       | −38.6 (11)    | −28.1 (11)  | −27.4 (11) |
| Group C               | −54.4 to −36.8 | −40.9 to −29.2 | −37.7 to −27.4 |
|                       | −45.4 (48)    | −32.6 (48)  | −30.5 (48) |

Note: minimum/maximum average sample number.
NWJB: northwestern Junggar Basin.
Source: Data from Yang et al. (2008), Tao et al. (2016), and Chen et al. (2014).
As discussed above, group C gases are oil-associated gases. Group C gases have a wide range of the chemical and carbon isotopic composition as well as partially reversed $\delta^{13}C$ order between propane and butane, as a result of two periods of accumulations in the Fengcheng Formation (Tao et al., 2016). However, previous studies have confirmed that the two sets of source rocks in the Fengcheng and Wuerhe Formations have great potential to generate oil in the NWJB, and currently discovered oils are mainly derived from the

![Figure 4. Carbon isotope-type curves of natural gases in the NWJB.](source)

NWJB: northwestern Junggar Basin.
Source: Data from Yang et al. (2008), Tao et al. (2016), and Chen et al. (2014).
Nevertheless, some other studies proposed that the contribution of source rocks in the Wuerhe Formation to the oil generation in the NWJB was underestimated in the past (Chen et al., 2016c, 2016d). This can be attributed to a limited understanding of source rock distribution as the source rocks of the Wuerhe Formation revealed by drilling nowadays are mainly located at the margin of the basin with little in the center of the basin (Chen et al., 2016b). Briefly, group C gases likely consisted of mixed oil-associated gases from either the Fengcheng and Wuerhe Formations (P2w) or the two periods of accumulation in the Fengcheng Formation (P1f). To answer these inferences, we introduced ALS to de-convolute group C gases. Since the sample number for chemometrics should usually be more than 30 (Wang et al., 2016), only group C gases were discussed in this paper. In the light of the ALS results (Table 2), group C gases had two possible sources (source 1 and source 2), and the contribution of source 2 to the gases in group C was higher than that of source 1 (Figure 6). Table 3 shows the stable carbon isotopic composition of the gaseous hydrocarbons (C1–C3) from the Fengcheng Formation (P1f) under thermal simulation at different heating temperatures. The carbon isotopic composition of source 1 was similar to that of the Fengcheng Formation (P1f) at 350°C, while that of the source 2 seemed to be independent of that from the Fengcheng Formations (Chen et al., 2016b, 2016e; Yu et al., 2017). Nevertheless, some other studies proposed that the contribution of source rocks in the Wuerhe Formation to the oil generation in the NWJB was underestimated in the past (Chen et al., 2016c, 2016d). This can be attributed to a limited understanding of source rock distribution as the source rocks of the Wuerhe Formation revealed by drilling nowadays are mainly located at the margin of the basin with little in the center of the basin (Chen et al., 2016b).

Figure 5. Carbon isotopic composition of kerogens in source rocks of the P2w, P1f, and P1j Formations and Carboniferous strata in the NWJB.
NWJB: northwestern Junggar Basin.
Source: Modified after Chen et al. (2016b).
Formation (P1f). Thus, source 1 was probably derived from mature source rocks of the Fengcheng Formation (P1f), while source 2 was most likely derived from the Wuerhe Formation (P2w). This conclusion is also supported by similar ranges of the carbon isotopic composition between the P2w and P1f extracts and crude oils in the NWJB (Figure 7). Also, these results are in a good agreement with the three events of petroleum generation in the NWJB during the Middle-Late Permian, Late Triassic, and Early Cretaceous, corresponding to the source rocks in the Jiamuhe, Fengcheng, and Wuerhe Formations, respectively (Cao et al., 2005). In summary, group C gases were likely derived from the Fengcheng and Wuerhe Formations, with a higher contribution from the latter strata. Therefore, the oil generation potential of the Wuerhe Formation in the NWJB was probably underestimated in the past, and more attention should be paid in future exploration.

### Table 2. ALS-calculated two source values of group C gases.

| Sources  | CH₄   | C₂H₆ | C₃H₈ | δ¹³C₁  | δ¹³C₂  | δ¹³C₃  |
|----------|-------|------|------|-------|-------|-------|
| Source 1 | 68.21 | 11.50| 7.20 | -49.7 | -35.7 | -32.9 |
| Source 2 | 91.69 | 2.80 | 1.00 | -42.5 | -30.4 | -28.9 |

ALS: alternating least squares.

### Figure 6. ALS-calculated relative contribution of source 1 and source 2. ALS: alternating least squares.

### Table 3. Stable carbon isotopic composition of gaseous hydrocarbons (C₁–C₃) of the Fengcheng Formation (P₁f) under thermal simulation at different heating temperatures (Wang et al., 2013).

| Formation | T (°C) | δ¹³C₁ (‰) | δ¹³C₂ (‰) | δ¹³C₃ (‰) |
|-----------|--------|------------|------------|------------|
| P₁f       | 300    | -44.3      | /          | /          |
|           | 350    | -43.3      | -34.5      | -32.8      |
|           | 400    | -42.9      | -33.4      | -31.3      |
|           | 450    | -36.9      | -36.5      | -25.6      |
|           | 500    | -36.2      | -29.3      | -24.2      |
|           | 550    | -34.2      | -28.9      | /          |

"/": Not detected.
Conclusions

The results of two chemometric methods based on the PCA and MDS of six proxies of chemical compositions and isotopes signatures in natural gases in the NWJB show that the natural gases can be divided into three genetic groups: A, B, and C. Group A, B, C gases were determined as coal-derived gas, mixing of coal-derived and oil-associated gas, and oil-associated gas, respectively. All three groups of natural gases showed a wide range of their chemical and carbon isotopic composition, and there was a partial reversal of $\delta^{13}C$ values between propane and butane, indicating that these gases were derived from diverse sources. Group A gases may have been mainly derived from the mixing of sources in the Jiamuhe Formation and Carboniferous strata, while group C gases were probably derived from the Fengcheng and Wuerhe Formations. Noticeably, in accordance with the ALS results, the contribution of source rocks in the Wuerhe Formation to group C gases was higher than that of the Fengcheng Formation. This study provides an effective chemometric method of natural gas classification and evaluation of hydrocarbon generation potential. This contributes to a better understanding of origin of gases and distribution of oil and gas, assisting in exploration deployment in the basin.

Acknowledgements

We are deeply indebted to two anonymous reviewers who provided constructive reviews and comments in our original manuscript. The authors also wish to acknowledge editor-in-chief Professor Yuzhuang Sun for the timely handling of the manuscript.
Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Foundation of China (Grant No. 41602146), State Key Laboratory of Organic Geochemistry, GIGCAS (Grant No. SKLOG-201703), Doctoral Research Initiation Project of Guangdong Ocean University (Grant No. R17001), and the Special Financial Aid for Talents of Guangdong Ocean University (Grant No. 002026002004).

ORCID iD
Yao-Ping Wang  https://orcid.org/0000-0003-0994-6030

Supplemental material
Supplemental material for this article is available online.

References
Bevilacqua M, Bro R, Marini F, et al. (2017) Recent chemometrics advances for foodomics. TrAC Trends in Analytical Chemistry 96: 42–51.
Cao J, Zhang Y, Hu W, et al. (2005) The Permian hybrid petroleum system in the northwest margin of the Junggar Basin, northwest China. Marine and Petroleum Geology 22(3): 331–349.
Chabukdhara M and Nema AK (2012) Assessment of heavy metal contamination in Hindon River sediments: A chemometric and geochemical approach. Chemosphere 87(8): 945–953.
Chen J, Wang X, Deng C, et al. (2016a) Geochemical features of source rocks and crude oils in the Junggar Basin, northwest China. Acta Geologica Sinica 90(1): 37–67 (in Chinese with English abstract).
Chen J, Wang X, Deng C, et al. (2016b) Oil and gas source, occurrence and petroleum system in the Junggar Basin, northwest China. Acta Geologica Sinica 90(3): 421–450 (in Chinese with English abstract).
Chen Z, Cao Y, Ma Z, et al. (2014) Geochemistry and origins of natural gases in the Zhongguai area of Junggar Basin, China. Journal of Petroleum Science and Engineering 119: 17–27.
Chen Z, Cao Y, Wang X, et al. (2016c) Oil origin and accumulation in the Paleozoic Chepaizi–Xinguang field, Junggar Basin, China. Journal of Asian Earth Sciences 115: 1–15.
Chen Z, Liu G, Wang X, et al. (2016d) Origin and mixing of crude oils in Triassic reservoirs of Mahu slope area in Junggar Basin, NW China: Implication for control on oil distribution in basin having multiple source rocks. Marine and Petroleum Geology 78: 373–389.
Chen Z, Zha M, Liu K, et al. (2016e) Origin and accumulation mechanisms of petroleum in the Carboniferous volcanic rocks of the Kebai Fault zone, Western Junggar Basin, China. Journal of Asian Earth Sciences 127: 170–196.
Dai J, Qin S, Tao S, et al. (2005) Developing trends of natural gas industry and the significant progress on natural gas geological theories in China. Natural Gas Geoscience 16(02): 127–142 (in Chinese with English abstract).
Dai J, Xia X, Qin S, et al. (2004) Origins of partially reversed alkane δ13C values for biogenic gases in China. Organic Geochemistry 35(4): 405–411.
Gao G, Xiang B, Ren J, et al. (2016) Origin and source of natural gas from Wuxia Fault Belt in the northern Mahu Sag. *Junggar Basin. Natural Gas Geoscience* 27(4): 672–680 (in Chinese with English abstract).

Graham SA, Brassell S, Carroll AR, et al. (1990) Characteristics of selected petroleum source rocks, Xinjiang Uygur Autonomous Region, NW China. *American Association of Petroleum Geologists Bulletin* 74(4): 493–512.

He M, Moldowan JM, Nemchenko-Rovenskaya A, et al. (2012) Oil families and their inferred source rocks in the Barents Sea and northern Timan-Pechora Basin, Russia. *American Association of Petroleum Geologists Bulletin* 96(6): 1121–1146.

Kramer R (1988) *Chemometric Techniques for Quantitative Analysis*. New York: Marcel Dekker.

Kvalheim OM, Aksnes DW, Brekke T, et al. (1985) Crude oil characterization and correlation by principal component analysis of $^{13}$C nuclear magnetic resonance spectra. *Analytical Chemistry* 57(14): 2858–2864.

Madsen R, Lundstedt T and Trygg J (2010) Chemometrics in metabolomics-A review in human disease diagnosis. *Analytica Chimica Acta* 659(1–2): 23–33.

Mashhadi ZS and Rabbani AR (2015) Organic geochemistry of crude oils and Cretaceous source rocks in the Iranian sector of the Persian Gulf: An oil–oil and oil–source rock correlation study. *International Journal of Coal Geology* 146: 118–144.

Meng F and Parhati (1999) Evaluation of gas pool of Carboniferous source rock in Junggar Basin. *Journal of Xinjiang Petroleum Institute* (2): 1–60 (in Chinese with English abstract).

Øygard K, Grahl-Nielsen O and Ulvøen S (1984) Oil/oil correlation by aid of chemometrics. *Organic Geochemistry* 6(84): 561–567.

Pan C, Yang J, Fu J, et al. (2003) Molecular correlation of free oil and inclusion oil of reservoir rocks in the Junggar Basin, China. *Organic Geochemistry* 34(3): 357–374.

Peters KE, Coutrot D, Nouvelle X, et al. (2013) Chemometric differentiation of crude oil families in the San Joaquin Basin, California. *American Association of Petroleum Geologists Bulletin* 97(1): 103–143.

Peters KE, Moldowan JM, Schoell M, et al. (1986) Petroleum isotopic and biomarker composition related to source rock organic matter and depositional environment. *Organic Geochemistry* 10(1–3): 17–27.

Peters KE, Ramos LS, Zumberge JE, et al. (2008) De-convoluting mixed crude oil in Prudhoe Bay Field, North Slope, Alaska. *Organic Geochemistry* 39(6): 623–645.

Peters KE, Walters CC and Moldowan JM (2005) *The Biomarker Guide: Biomarkers and Isotopes in Petroleum and Earth History*. Cambridge: Cambridge University Press.

Peters KE, Wright TL, Ramos LS, et al. (2016) Chemometric recognition of genetically distinct oil families in the Los Angeles basin, California. *American Association of Petroleum Geologists Bulletin* 100(01): 115–135.

Tao K, Cao J, Wang Y, et al. (2016) Geochemistry and origin of natural gas in the petroliferous Mahu sag, northwestern Junggar Basin, NW China: Carboniferous marine and Permian lacustrine gas systems. *Organic Geochemistry* 100: 62–79.

Wang X, Zhi D, Wang Y, et al. (2013) *Source Rocks and Oil-Gas Geochemistry in Junggar Basin*. Beijing: Petroleum Industry Press (in Chinese).

Wang Y, Peters KE, Moldowan JM, et al. (2014) Cracking, mixing, and geochemical correlation of crude oils, North Slope, Alaska. *American Association of Petroleum Geologists Bulletin* 98(6): 1235–1267.

Wang Y-P, Zhan X, Zou Y-R, et al. (2019) Chemometric methods as a tool to reveal genetic types of natural gases – A case study from the Turpan-Hami Basin, northwestern China. *Petroleum Science and Technology* 37(3): 310–316.

Wang Y-P, Zhang F, Zou Y-R, et al. (2018) Oil source and charge in the Wuerxun Depression, Hailar Basin, northeast China: A chemometric study. *Marine and Petroleum Geology* 89(3): 665–686.
Wang Y-P, Zhang F, Zou Y-R, et al. (2016) Chemometrics reveals oil sources in the Fangzheng Fault Depression, NE China. *Organic Geochemistry* 102: 1–13.

Yang H, Wei H, Jiang X, et al. (2008) Identification of natural gas types and analysis on its distribution rules in Permian in block No. 5-8 of northwest margin of Junggar Basin. *Petroleum Geology & Oilfield Development in Daqing* 27(01): 46–50 (in Chinese with English abstract).

Yu S, Wang X, Xiang B, et al. (2017) Molecular and carbon isotopic geochemistry of crude oils and extracts from Permian source rocks in the northwestern and central Junggar Basin, China. *Organic Geochemistry* 113: 27–42.

Zhan Z-W, Tian Y, Zou Y-R, et al. (2016a) De-convoluting crude oil mixtures from Palaeozoic reservoirs in the Tabei Uplift, Tarim Basin, China. *Organic Geochemistry* 97: 78–94.

Zhan Z-W, Zou Y-R, Shi J-T, et al. (2016b) Unmixing of mixed oil using chemometrics. *Organic Geochemistry* 92: 1–15.

Zumberge JE (1987) Prediction of source rock characteristics based on terpane biomarkers in crude oils: A multivariate statistical approach. *Geochimica et Cosmochimica Acta* 51(6): 1625–1637.