

Хемометрический метод интерпретации результатов хромато-масс-спектрометрического анализа минеральных и синтетических масел для таможенных целей

***Д.В. Лякаев¹, М.О. Пастухов²**

¹Экспертно-криминалистическая служба регионального филиала Центрального экспертно-криминалистического таможенного управления г. Нижний Новгород, Российская Федерация, 603104, г. Нижний Новгород, ул. Медицинская, 24

²Центральное экспертно-криминалистическое таможенное управление Российской Федерация, 125130, г. Москва, ул. Клары Цеткин, д. 29А.

**Адрес для переписки Лякаев Денис Владимирович, E-mail: lyakaev94@mail.ru

Поступила в редакцию 15 сентября 2023 г., после доработки 07 декабря 2023 г.

В данной работе рассмотрено применение метода главных компонент (РСА) на примере различения минеральных, синтетических и полусинтетических моторных масел по результатам хромато-масс-спектрометрического анализа. Объектами исследования являлись образцы товарных масел российского и зарубежного производства. Предложен обобщенный алгоритм построения классификационной модели, в том числе описаны этапы предварительной обработки спектральной информации, а также условия постановки эксперимента. Для оценки полученных результатов был применен метод главных компонент (МГК, Principal Component Analysis, PCA) и метод регрессии на главные компоненты (МПК, Principal Component Regression, PCR). Первичными данным для хемометрической обработки служили хроматограммы исследованных образцов.

Анализ графиков счетов в модели, построенной МГК, для первой и второй главных компонент выявил, что точки, соответствующие разным группам объектов, не перекрываются и находятся в разных областях относительно первой и второй главных компонент. График нагрузок отражает характерные времена выхода, которые отвечают за разделение исследованных образцов на минеральные и синтетические масла.

Для моделирования хромато-масс-спектрометрических данных с учётом задачи количественной идентификации смесевых композиций минеральных и синтетических масел был выбран метод МПК. На основании полученных результатов установлено, что хромато-масс-спектрометрия в сочетании с хемометрическими методами анализа многофакторных зависимостей может найти применение для решения задач, связанных с установлением качественного и количественного состава моторных масел. Полученные модели демонстрирует достаточную точность для качественного и количественного определения состава смешанных моторных масел на основе минеральных и синтетических базовых компонентов.

Ключевые слова: минеральные масла, синтетические масла, хромато-масс-спектрометрия, метод главных компонент, метод регрессии на главные компоненты

For citation: *Analitika i kontrol'* [Analytics and Control], 2023, vol. 27, no. 4, pp. 299-305

DOI: 10.15826/analitika.2023.27.4.012

Chemometric interpretation of the results of chromatomass spectrometry analysis of mineral and synthetic oils for customs purposes

***D.V. Lyakaev¹, M.O. Pastukhov²**

¹Expert-Criminalistics Service – Regional Branch of the Central Expert-Criminalistics Customs Administration, Nizhny Novgorod
24 Meditsnaskay Str., Nizhny Novgorod, 603104, Russian Federation

²Central Expert-Criminalistics Customs Administration
29A Klara Tsetkin Str., Moscow, 125130, Russian Federation

*Corresponding author: Denis V. Lyakaev, E-mail: lyakaev94@mail.ru

Submitted 15 September 2023, received in revised form 07 December 2023

Application of the principal component method (PCA) using the example of distinguishing mineral, synthetic and semi-synthetic motor oils from the results of chromatography-mass spectrometric analysis is considered in the present work. Samples of commercial oils of domestic and foreign production constituted the objects investigated. A generalized algorithm for constructing a classification model, including the stages of preliminary processing of spectral information, as well as conditions for setting up the experiment, is proposed. The principal component analysis (PCA) and the principal component regression (PCR) methods were applied to evaluate the results obtained. Chromatograms of the studied samples acted as the primary data for chemometric processing.

Analysis of the graphs of scores in the model constructed by PCA for the first and second principal components revealed that the points corresponding to different groups of objects did not overlap and were located in different areas in relation to the first and second principal components. The loadings plot reflects the characteristic retention times, which are responsible for separating studied samples into mineral and synthetic oils.

The PCR method was chosen for modeling chromatography-mass spectrometry data taking into account the task of quantitative identification of mixed compositions of mineral and synthetic oils. Based on the results obtained it was found that chromatography-mass spectrometry in combination with chemometric methods of analyzing multifactorial dependencies can be applied for solving problems, related to determining qualitative and quantitative composition of motor oils. The obtained models demonstrate sufficient accuracy for qualitative and quantitative determination of the composition of mixed motor oils based on mineral and synthetic base components.

Keywords: mineral oils, synthetic oils, principal component analysis, chromatography-mass spectrometry, principal component method, principal component regression method

ВВЕДЕНИЕ

Современные углеводородные масла представляют собой композиции оптимального углеводородного состава изопарафинов с определенной длиной цепи и степенью изомеризации. Они могут быть получены выделением широкой гаммы углеводородов нефти, а также в результате использования гидрогенизационных процессов переработки ряда побочных нефтепродуктов (гидроизомеризация нормальных парафинов, гидрокрекинг нафтен и ароматики), а также целевым синтезом изопарафинов из олефинов (процесс получения синтетических поли-альфа-олефиновых масел) [1].

Ввиду большого разнообразия смазочных масел становится актуальной задача установления строения и состава масел современными инструментальными методами анализа. В целях таможенного регулирования необходимо количественное определение минерального (полученного из нефти или нефтепродуктов) и синтетического компонента в составе смазочных масел. Достаточно доступным экспрессным методом решения данной задачи может служить, например, хромато-масс-спектрометрия [2].

Метод газовой хроматографии применяют для установления химического состава смазочных масел [2]. Он позволяет однозначно различать смазочные материалы, изготовленные на базе синтетических масел: поли- α -олефиновые, сложные эфиры дикарбоновых кислот, сложные эфиры неопентиловых спиртов, полиорганосилоксаны, эфиры фосфорной кислоты. Кроме того, с помощью газовой хрома-

тографии можно устанавливать некоторые типы присадок, относящиеся, например, к стерическим затрудненным фенолам, решать классификационные и идентификационные задачи.

Вместе с тем, в связи с огромным многообразием природы базовых компонентов (минеральных или синтетических) и применяемых присадок в готовых торговых образцах анализ многомерных данных по результатам хромато-масс-спектрометрического исследования является нетривиальной задачей [2]. Перспективным подходом к установлению качественного и количественного состава смазочных масел является их анализ с последующей хемометрической обработкой полученных данных.

Целью данной работы является исследование возможности качественного и количественного определения минеральных и синтетических компонентов в составе моторных масел на основе градуировки по обучающему набору.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Объектами исследования являлись образцы товарных масел российского и зарубежного производства, как готовые к применению, так и используемые в качестве базовых компонентов – 23 образца минеральных и 6 образцов синтетических масел.

Для изучения возможности количественного определения композиций были приготовлены смеси на основе уже изученных синтетических и минеральных базовых компонентов, из которых 10 были использованы в обучающем наборе, 8 – исследованы в

Таблица 1

Описание образцов, используемых в эксперименте

Table 1

Description of the samples used in the experiment

Образцы	Количество	Массовая доля минерального масла, %
Для обучения		
- минеральные масла	23	100
- синтетические масла	6	0
- полусинтетические масла (минеральное в смеси с синтетическим)	10	20 – 80
Для проверки		
- минеральные масла	9	100
- синтетические масла	2	0
- полусинтетические масла (минеральное в смеси с синтетическим)	3	20 – 60
Неизвестные образцы	8	20 – 80

качестве неизвестных, всего 18 образцов. Разделение образцов на градуировочные и проверочные (табл. 1) проводили с помощью метода, описанного в [3].

Подготовка образцов к исследованиям состояла в приготовлении растворов с одинаковой массовой концентрацией 10 мг/мл исследуемых образцов и их смесей в хлороформе (х.ч., АО «Химреактивснаб»).

Хроматографирование и регистрацию масс-спектров проводили в условиях, указанных в [2]: хромато-масс-спектрометр Agilent GC-6850/MSD-5975C («Agilent Technologies», США), колонка HP-5MS, длина колонки 30 м, внутренний диаметр 0.25 мм, толщина неподвижной фазы 0.25 мкм; начальная температура и выдержка 70 °С, 3.33 мин; скорость нагрева 15 °С/мин до 320 °С; выдержка при макс. температуре 15 мин; скорость потока гелия (марка 6.0) 40 см/с (режим Constant Flow); температура источника 150 °С (стандартная); температура инжектора и интерфейса 280 °С; режим инжектора Splitless; режим Purge flow to Split Vent 100 мл/мин через 1.5 мин; режим экономии газа Gas Saver включен, 15 мл/мин через 2.5 мин; диапазон масс 33 – 800 а.е.м., объем пробы 1 мкл; растворитель хлороформ.

Предварительная обработка первичных хромато-масс-спектрометрических данных (рис. 1) включала вычисление усредненного масс-спектра для каждой хроматограммы путем его коррекции вычитанием масс-спектра «холостого» опыта», в качестве которого использовали средний масс-спектр хроматограммы хлороформа при соответствующем времени хроматографирования.

Статистическую (хеометрическую) обработку полученных хроматограмм проводили с использованием шаблонов Excel [4, 5]. Выходные данные для алгоритмов обработки данных представляли собой зависимость интенсивности сигнала детектора от времени (число столбцов матрицы данных равно 2444). Для оценки полученных результатов был применен метод главных компонент (МГК, Principal Component Analysis, PCA) и метод регрессии на главные компоненты (МПК, Principal Component Regression, PCR). Описание градуировочных и проверочных моделей приведено в табл. 2.

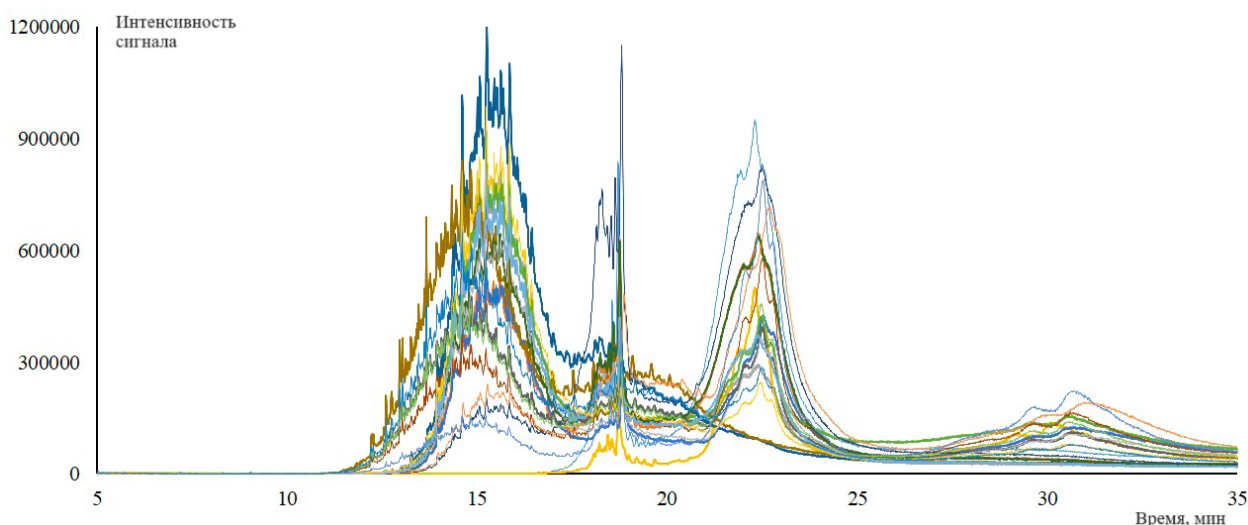


Рис. 1. Хроматограммы исследованных образцов

Fig. 1. Chromatograms of investigated samples

Таблица 2

Описание градуировочных и проверочных моделей

Table 2

Description of the calibration and test models

Описание модели	Количество образцов модели	Число ГК	СКО	Коэффициент корреляции (R ²)
Модель №1 Различение на две группы (МГК)	23 (минеральные масла) 6 (синтетические масла)	2	0.633	-
Модель №2 Количественная идентификация смесевых композиций (МРК)	градуировочный набор: 23 (минеральные масла) 6 (синтетические масла)	4	0.12	0.91
	7 (полусинтетические масла) отбор в проверочный набор: 9 (минеральные масла) 2 (синтетические масла) 3 (полусинтетические масла)		0.11	0.90
Модель №3 Количественная идентификация смесевых композиций с дополнительным набором смесевых масел (МРК)	градуировочный набор: 23 (минеральные масла) 6 (синтетические масла) 10 (полусинтетические масла)	4	0.11	0.92
Модель №4 Количественная идентификация смесевых композиций с неизвестными образцами (МРК)	градуировочный набор: 23 (минеральные масла) 6 (синтетические масла) 10 (полусинтетические масла)	4	0.13	0.87
	неизвестные образцы: 8 (смесевые масла)		0.10	0.91

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Первоначально хроматографические данные были изучены методом главных компонент (МГК, PCA) [4 – 8]. С его помощью можно увидеть, что все исследованные образцы можно разделить на две группы – синтетические и минеральные масла (Модель №1 – рис. 2).

На графике счетов исследованных образцов в координатах двух первых главных компонент ГК1 – ГК2 видно, что точки, соответствующие разным группам объектов, не перекрываются и находятся в разных областях относительно первой и второй главных компонент (principal components, далее – ГК). Это свидетельствует о достоверных различиях между группами. Расположение образцов в координатном пространстве ГК показывает, что образцы минеральных масел локализованы в зоне значений ГК2 менее 0.5, в то время как образцы синтетических масел находятся в зоне ГК2 > 0.5. Оценка полученных результатов показала, что вторая главная компонента разделяет синтетические и минеральные масла. Кроме того, при анализе графика счетов исследованных образцов можно отметить, что минеральные масла локализованы в нескольких группах, незначительно отстающих друг от друга в координатах ГК1, что, скорее всего, свидетельствует

о большом разнообразии композиционного состава базовых компонентов на минеральной основе.

График нагрузок для первой и второй главных компонент выявил характерные времена выхода, которые отвечают за разделение объектов на группы. Основные различия профилей хроматограмм минеральных и синтетических масел наблюдаются именно во временные интервалы 12.5 – 14.3 мин и

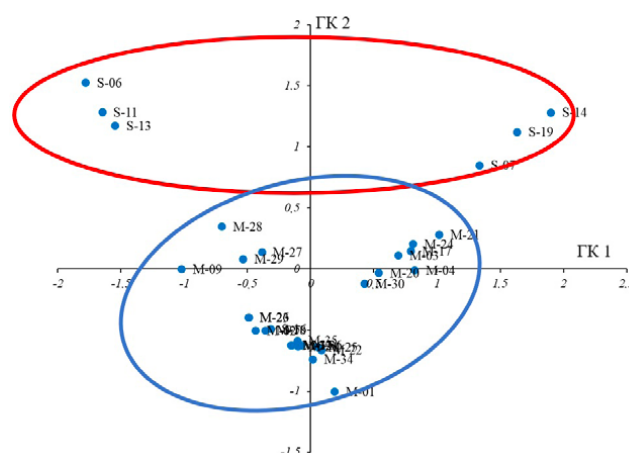


Рис. 2. График счетов исследованных образцов в координатах ГК1 - ГК2

Fig. 2. Scores plot of the studied samples in PC1 - PC2 coordinates

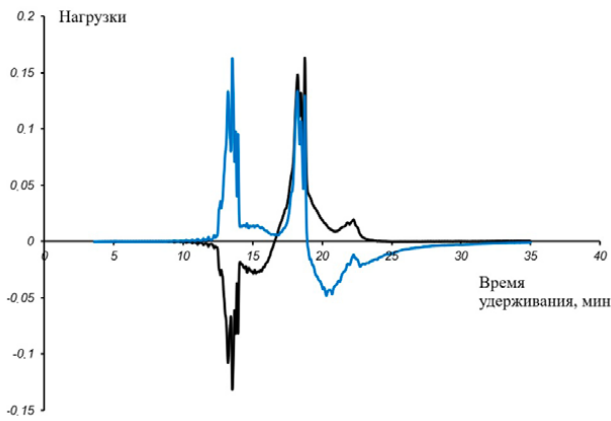


Рис. 3. График нагрузок ГК1 и ГК2 градуировочного набора образцов минеральных и синтетических масел

Fig. 3. PC1 and PC2 loading plot for the calibration set of samples of mineral and synthetic oils

16.5 – 20.9 мин в приведенных условиях хроматографирования (рис. 3).

Для моделирования хромато-масс-спектрометрических данных с учётом задачи количественной идентификации смесевых композиций минеральных и синтетических масел был выбран метод МРК [4 – 8]. В качестве определяемой величины была выбрана массовая доля минерального масла в исследуемых объектах: в случае индивидуальных минеральных масел – 100 %, синтетических – 0 %, в смесевых композициях – от 20 % до 80 %.

Для анализа данных методом МРК использовали те же данные, что и для МГК – 23 образца минеральных, 6 образцов синтетических, 7 образцов полусинтетических масел. В целях проверки был

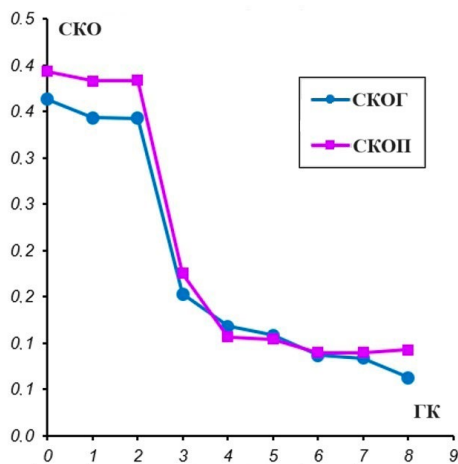


Рис. 5. Зависимость среднеквадратичного отклонения градуировки (СКОГ) и среднеквадратичного отклонения предсказания (СКОП) от количества главных компонент

Fig. 5. The dependence of the root mean square error of calibration (RMSEC) and the root mean square error of prediction (RMSEP) on the number of principal components

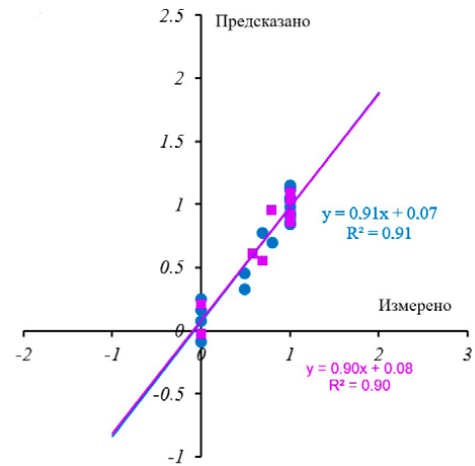


Рис. 4. Модель на основании градуировочного и проверочного наборов минеральных, синтетических и смешанных масел (ГК4)

Fig. 4. The model based on the calibration and verification sets of mineral, synthetic and mixed oils (PC4)

отобран проверочный набор, включающий в себя 9 образцов минеральных масел и 2 образца синтетических масел, а также 3 образца смешанных масел (Модель №2 – рис. 4). Значительного снижения точности предсказания определяемой величины не наблюдалось. По результатам анализа зависимости среднеквадратичного отклонения градуировки (СКОГ, *RMSEC*) и среднеквадратичного отклонения предсказания (СКОП, *RMSEP*) от количества главных компонент сделан вывод о достаточности четырех главных компонент для достижения необходимой точности и гибкости модели (рис. 5).

Следующим этапом исследовательской работы являлось дополнительное усложнение модели. В градуировочный набор были добавлены смесевые композиции на основе одного образца синтетического масла и двух образцов минеральных масел; всего в

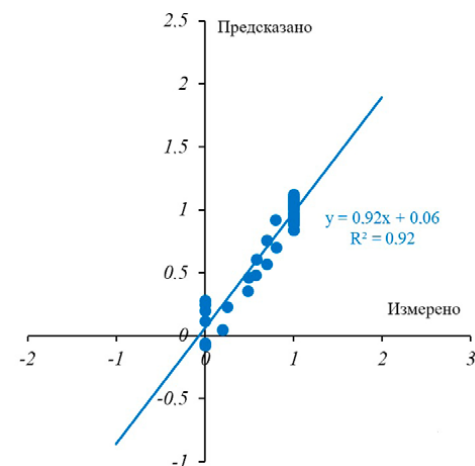


Рис. 6. Модель на основании градуировочного набора минеральных, синтетических и смешанных масел (ГК4)

Fig. 6. The model based on the calibration set of mineral, synthetic and mixed oils (PC4)

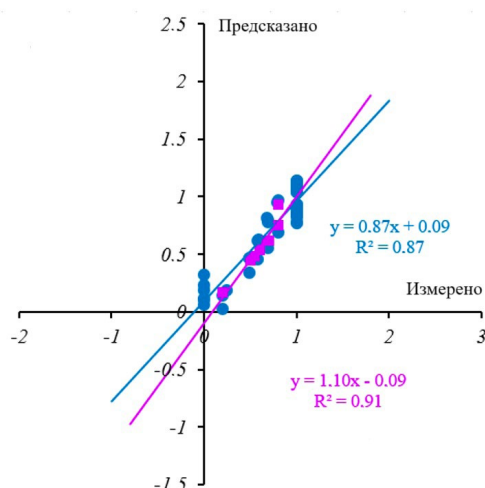


Рис. 7. Модель на основании градуировочного набора минеральных, синтетических и смешанных масел и тестового набора из смешанных масел (ГК4)

Fig. 7. The model based on a calibration set of mineral, synthetic and mixed oils and a test set of mixed oils (PC4)

градуировочном наборе использовали 10 смесевых образцов с диапазоном концентрации минерального масла от 20 % до 80 %. По результатам построения данной модели на четырех главных компонентах коэффициент корреляции составил 0.92 (Модель № 3 – рис. 6). Аналогично был составлен проверочный набор из 3 смесевых образцов с диапазоном концентраций минерального масла от 20 % до 60 %. Доказана устойчивость, а также достаточность четырех главных компонент для описания градуировочного набора хроматограмм.

Практические возможности использования созданной классификационной модели были проверены на 8 образцах смесевых композиций, составленных из двух образцов синтетического и трех образцов минерального масла, не участвовавших в построении модели и использованных в качестве неизвестных. В результате расчёта количественного состава «неизвестного» набора смешанных масел коэффициент корреляции составил более 0.85 (Модель №4 – рис. 7).

ЗАКЛЮЧЕНИЕ

Полученная модель №4 демонстрирует достаточную точность для количественного определения состава смешанных моторных масел на основе минеральных и синтетических базовых компонентов. Дополнительно можно отметить стабильность разработанного алгоритма обработки хромато-масс-спектрометрических данных, ввиду сохранения порядка показателя точности, достигнутого при анализе градуировочного набора. Таким образом, в результате проведенных исследований установлено, что хромато-масс-спектрометрия в сочетании с хемометрическими алгоритмами позволяет определить качественный, а также ко-

личественный состав смазочных моторных масел с точностью, достаточной для таможенных целей.

Для достоверного количественного анализа, увеличения показателя точности необходимым является использование дополнительного набора обучающего массива хромато-масс-спектрометрических данных исследования базовых компонентов различных производителей минеральной и синтетической природы, а также их смесевых композиций.

БЛАГОДАРНОСТИ

Авторы выражают искреннюю благодарность за помощь в подготовке настоящей работы доктору физико-математических наук, главному научному сотруднику Института химической физики им. Н.Н. Семенова Российской Академии наук Алексею Леонидовичу Померанцеву и доктору физико-математических наук, главному научному сотруднику Института химической физики им. Н.Н. Семенова Российской Академии наук Оксане Евгеньевне Родионовой.

ACKNOWLEDGEMENTS

The authors express their sincere gratitude to Doctor of Physical and Mathematical Sciences, Principal Researcher of the Institute of Chemical Physics named after N.N. Semenov of the Russian Academy of Sciences Aleksey L. Pomerantsev and Doctor of Physical and Mathematical Sciences, Principal Researcher of the Institute of Chemical Physics named after N.N. Semenov of the Russian Academy of Sciences Oksana E. Rodionova for the help in preparing this work.

ЛИТЕРАТУРА

1. Манг Т., Дрезель У. Смазки. Производство, применение, свойства. Справочник: пер. с англ. под ред. В.М. Школьникова. СПб.: ЦОП «Профессия», 2010. 944 с.
2. Тагиров Т.К., Поляков Д.Ю. Комплексное исследование смазочных материалов на основе синтетических, смешанных и нефтяных масел. М.: РФЦЭС при Минюсте России, 2008. 193 с.
3. Subset selection using Combined Analytical Signal / Pomerantsev A.L., Rodionova O.Ye. // *Microchem. J.* 2023. V. 190. Article 108654.
4. Померанцев А.Л. Хемометрика в Excel. Учебное пособие. Томск: Изд-во ТПУ, 2014. 434 с.
5. Pomerantsev A.L. *Chemometrics in Excel*. John Wiley & Sons. Hoboken NJ. 2014.
6. Principal component analysis: concept, geometrical interpretation, mathematical background, algorithms, history, practice / Esbensen H., Geladi P. // *Comprehensive Chemometrics*, Elsevier. 2009. P. 211–226.
7. Родионова О.Е. Интервальные методы обработки многоканальных экспериментов: дис. ... д-ра физ.-мат. наук: 01.04.01. М., 2008. 272 с.
8. Родионова О.Е., Померанцев А.Л. Хемометрика: достижения и перспективы / *Успехи химии*. 2006. Т. 75, № 4. С. 302-317.

REFERENCES

1. Mang T., Drezel' U. Lubricants. *Production, application, properties. Reference book* (Russ. ed.: Shkol'nikova V.M. *Smazki. Proizvodstvo, primeneniye, svoystva*. SPb.: COP «Professija», 2010. 944 p.).
2. Tagirov T.K., Poljakov D.Ju. *Kompleksnoe issledovanie smazochnykh materialov na osnove sinteticheskikh, smeshannykh i neftiannykh masel [Comprehensive study of lubricants based on synthetic, mixed and petroleum oils]*. M.: RFTSSE under the Ministry of Justice of Russia, 2008. 193 p. (in Russian).
3. Pomerantsev A.L., Rodionova O.Ye. [Subset selection using Combined Analytical Signal]. *Microchem. J.*, 2023, V. 190, article 108654.
4. Pomerantsev A.L. *Khemometrika v Excel [Chemometrics in Excel]*. Tomsk, Publishing house TPU, 2014. 435 p. (in Russian).
5. Pomerantsev A.L. *Chemometrics in Excel*. John Wiley & Sons. Hoboken NJ., 2014.
6. Esbensen H., Geladi P. Principal component analysis: concept, geometrical interpretation, mathematical background, algorithms, history, practice. *Comprehensive Chemometrics, Elsevier*, 2009, pp. 211–226. doi: 10.1016/B978-0-444-52701-1.00043-0.
7. Rodionova O.E. *Interval'nye metod obrabotki mnogokanal'nykh eksperimentov [Interval method of processing multichannel experiments. Dr. phys. and math. sci. diss.]*. Moscow, 2008. 272 p. (in Russian).
8. Rodionova O.Ye., Pomerantsev A.L. Chemometrics: achievements and prospects *Russ. Chem. Rev.*, 2006, vol. 75, no. 4, pp. 271-287. doi:10.1070/RC2006v075n04ABEH003599. (in Russian).