



УДК 541

Исходные гипотезы для распознавания многокомпонентных физико-химических систем комбинацией «высокоэффективная жидкостная хроматография – метод главных компонент»

Гаврилина В.А., Сычев С.Н.

Государственный университет учебно-научный производственный комплекс, Орел

Поступила в редакцию 6.06.2012 г.

Аннотация

Обоснованы и экспериментально подтверждены гипотезы для распознавания многокомпонентных физико-химических систем органического происхождения на основе вычисленных безразмерных признаков комбинацией «высокоэффективная жидкостная хроматография – метод главных компонент» (ВЭЖХ – МГК). Показано, что использование этих гипотез позволяет определить идентичность, подобие или не идентичность многокомпонентных систем с неизвестным или переменным составом без предварительной идентификации отдельных компонентов.

Ключевые слова: гипотеза, высокоэффективная жидкостная хроматография, метод главных компонент, распознавание на основе вычисленных признаков

Substantiated and experimentally confirmed the hypothesis for the recognition of multi-component physical and chemical systems based on organic calculated dimensionless features a combination of "high-performance liquid chromatography - principal component analysis" (HPLC - CIM). It is shown that the use of these hypotheses to determine the identity, similarity or identity of multicomponent systems with unknown or variable composition, without prior identification of individual components.

Keywords: hypothesis, high performance liquid chromatography, principal component analysis, pattern recognition, based on the calculated characteristics

Введение

Контроль идентичности и качества многокомпонентных смесей (МС) органического происхождения (вино, нефть, растительные экстракты, биожидкости) с неизвестным или переменным составом всегда бросал вызов аналитику – вызов его профессиональной подготовке, используемым методам и инструментам: чем сложнее объект, тем субъективнее контроль. В случае, например, виноградных вин попытка решения проблемы с помощью стандартных процедур ВЭЖХ не менее субъективна, чем органолептический анализ [1,2]. Кроме того, покомпонентный анализ или получение отдельных профилей многокомпонентных смесей достаточно трудоемки, очень дороги и не всегда целесообразны. Исходя из вышеизложенного,

для экспрессного, эффективного и надежного контроля МС при их поточном производстве необходима разработка методов нетипичных с точки зрения классической аналитической химии: вместо покомпонентного или профильного анализа МС желательно перейти к распознаванию МС как цельной системы.

Поиск и разработка таких методов лежит, в том числе в области комбинации ВЭЖХ (многоволновая детекция) с методом главных компонент, являющимся одним из перспективных методов распознавания многоэлементных объектов по вычисленным безразмерным величинам – факторам [3]. При использовании метода главных компонент мы предполагаем, что несколько измеряемых переменных сильно коррелируют друг с другом. Это означает, что либо они взаимно определяют друг друга, либо связь между ними обуславливается третьей величиной, которую непосредственно измерить нельзя. Модель главных компонент в большей степени связана с последним предположением и дает возможность получить числовые значения этих третьих величин в виде набора линейнонезависимых факторов F , которые описывают и воспроизводят исходную матрицу с необходимой точностью [4].

Тем более удивителен факт совершенно незначительного применения метода главных компонент в высокоэффективной жидкостной хроматографии. До сих пор известно считанное число работ по применению МГК в теории селективности жидкостной хроматографии и физической химии межмолекулярных взаимодействий [5-15]. В аналитической химии метод главных компонент нашел ограниченное применение как метод оценки гомогенности хроматографического пика при использовании многоволновой спектрофотометрической детекции [16], хотя для оценки гомогенности пика можно обойтись и без МГК [17].

Показательным, в смысле попытки использования МГК в жидкостной хроматографии, является пример приведенный в работе [3]: «моделирование» участка трехмерной хроматограммы, полученной с помощью СФД с диодной линейкой, с целью разделения плохо разделившихся пиков.

Появление СФД с диодными линейками в свое время обещало переворот в способах идентификации соединений. Считалось, что трехмерная хроматограмма дает много информации для надежной идентификации. Разберем это утверждение подробнее. Было бы замечательно, если бы для идентификации соединений можно было использовать трехмерный образ хроматографического пика и сравнивать такой образ через набор топологических характеристик с образом стандартного пика, хранящегося в памяти компьютера. Было бы еще лучше, если трехмерный образ хроматографического пика не зависел бы от хроматографической системы. Однако огибающей по оси времени такого трехмерного пика является форма двумерного пика, зависящая от состава хроматографической системы: от состава элюента и квалификации растворителей, от вида адсорбента и технологии его изготовления, от способа подготовки хроматографической колонки к работе и т.д. от параметров, изменяющихся от случая к случаю. Таким образом, использование трехмерного образа хроматографического пика не приводит к улучшению идентификации соединений [18]. Для идентификации соединений, как и в случае, например, детектора хроматографа "Милихром" [16,17], используются спектральные отношения и времена удерживания – в том числе и при использовании диодно-матричных детекторов.

Исходные гипотезы для применения метода МГК-ВЭЖХ с многоволновым детектированием

Одной из главных причин неудач в области применения МГК в хроматографии является отсутствие каких-либо гипотез по применению МГК в этой области.

В начале исследований по применению МГК в ВЭЖХ (Сычев С.Н., Гаврилина В.А.) этот недостаток был учтен; было последовательно предложено два варианта по сути одной и той же гипотезы. На первом этапе исследований (2003–2005 гг.) была выдвинута гипотеза о существовании небольшого количества линейно-независимых (базовых) электронных спектров МС с органическими компонентами сходного происхождения (вино, растительный экстракт, нефть определенного месторождения, плазма крови одного вида животных): все электронные спектры компонентов такой МС являются линейными комбинациями базовых (расчетных). Набор базовых электронных спектров многокомпонентных систем является строго индивидуальным для каждой исследуемой системы (рис. 1).

Такая гипотеза позволяла надеяться на существенную свертку информации, полученную для многокомпонентных систем методом «ВЭЖХ–МГК» [11], рассматривать многокомпонентную систему при сравнении с контрольным образцом как единое целое без идентификации отдельных компонентов (распознавание системы), а также впервые ввести количественные критерии для определения идентичности, подобия или фальсификации МС путем получения коэффициентов корреляции линейно-независимых факторов исследуемой и контрольной МС. Оценка гипотезы на право существования была проведена в работах [19-22].

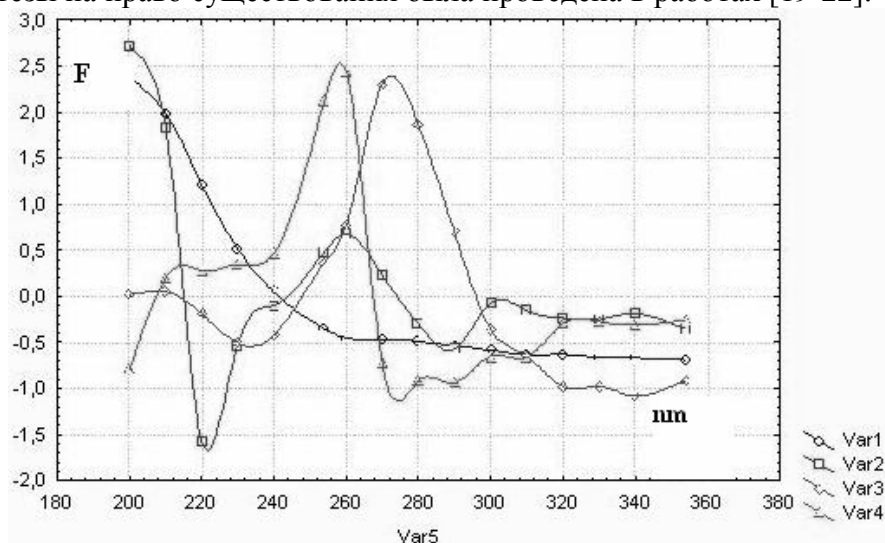


Рис. 1. Базовые линейно-независимые УФ-спектры вина «Каберне». По оси ординат – значения линейно-независимых факторов $F_1 - F_4$ (аналог оптической плотности). По оси абсцисс – длины волн, нм

Дальнейшие исследования, опубликованные в том числе в работах [19-26], позволили придать гипотезе более общий вид: существует небольшое количество линейнонезависимых (базовых, гипотетических) электронных структур многокомпонентных систем одного происхождения; электронные структуры реальных соединений, входящих в сложный объект, являются линейными комбинациями базовых.

Обобщенная гипотеза также, как и первоначальная, позволяет распознавать и сравнивать многокомпонентные системы без идентификации отдельных

компонентов системы путем сравнения наборов расчетных линейно-независимых факторов (ЛНФ) с помощью матрицы коэффициентов корреляции ЛНФ исследуемой системы и контрольного образца. Небольшое изменение гипотезы резко расширяет возможности комбинации «ВЭЖХ – МГК». Например, в теории катастроф такое расширение позволяет проследить изменение внешних условий на минорные соединения растений и выдать соответствующие рекомендации для расчета функций возмущения электронных базовых систем при решении обратной задачи уравнения Шредингера.

Отсутствие требования идентификации отдельных компонентов или профилей МС приводит к многократному ускорению, удешевлению и повышению достоверности процесса идентификации многокомпонентных систем и делает вполне приемлемым применение ВЭЖХ с многоволновым спектрофотометрическим детектированием для контроля качества многокомпонентных систем.

Выдвинутые гипотезы и процедура использования комбинации «ВЭЖХ–МГК» базируются, в том числе, на принципах аналогичных принципу линейности свободных энергий (ЛСЭ) формализованном в виде:

$$\lg (K_{\text{зам.}}/K_{\text{нез.}}) = \rho \cdot \sigma, \quad (1)$$

где $K_{\text{зам}}$ – константа реакционной способности замещенной молекулы; $K_{\text{нез.}}$ – константа реакционной способности незамещенной молекулы; ρ – коэффициент пропорциональности для серии однородных соединений; σ – вклад определенного внутримолекулярного электронного эффекта в реакционную способность.

Принцип ЛСЭ, подобно законам термодинамики, ниоткуда не следует: его справедливость успешно подтверждается экспериментально в различных областях науки и техники в течение почти 90 лет.

Применительно к выдвинутым гипотезам и процедуре применения комбинации «ВЭЖХ–МГК» принцип ЛСЭ выглядит следующим образом: хроматограмму МС, полученную на 5 длинах волн $\lambda_1 - \lambda_5$ и разбитую по времени, например, на 400 точек с интервалом Δt (рис. 5), можно представить в виде:

$$\begin{aligned} A_{i,\lambda_1}(t) &= [\lg (I_{0,\lambda_1}/ I_{\lambda_1})]_{i,\lambda_1}(t) = x_{1,\lambda_1}; x_{2,\lambda_1}; x_{3,\lambda_1}; \dots \dots \dots x_{400,\lambda_1} \\ &\dots \dots \dots \\ &\dots \dots \dots \\ &\dots \dots \dots \\ A_{i,\lambda_5}(t) &= [\lg (I_{0,\lambda_5}/ I_{\lambda_5})]_{i,\lambda_5}(t) = x_{1,\lambda_5}; x_{2,\lambda_5}; x_{3,\lambda_5}; \dots \dots \dots x_{400,\lambda_5} \end{aligned} \quad (2)$$

где I_{0,λ_i} – интенсивность света, исходящего из сравнительной кюветы при длине волны λ_i ; I_{λ_i} – интенсивность света, исходящего из рабочей кюветы при длине волны λ_i ; $A_{i,\lambda_i}(t)$ – оптическая плотность хроматограммы на временном промежутке $\Delta t \cdot i$ при длине волны λ_i ; $[\lg (I_{0,\lambda_i}/ I_{\lambda_i})]_{i,\lambda_i}(t)$ – оптическая плотность элюата как функция времени (та же хроматограмма) на временном промежутке $\Delta t \cdot i$ при длине волны λ_i ; x_{i,λ_i} – значение оптической плотности элюата при длине волны λ_i на i -том временном интервале хроматограммы. Таким образом, формируется матрица исходных оптических плотностей хроматограммы для обработки методом главных компонент.

Основная идея использования метода главных компонент при распознавании и сравнении (МС) органических соединений заключается в получении набора значимых линейно–независимых факторов из оптических плотностей многоволновых ВЭЖХ - хроматограмм вина и других МС. Оптические плотности предварительно превращены в матрицу дискретных значений оптических

плотностей элюата на разных длинах волн как временные срезы многоволновых хроматограмм (рис.2).

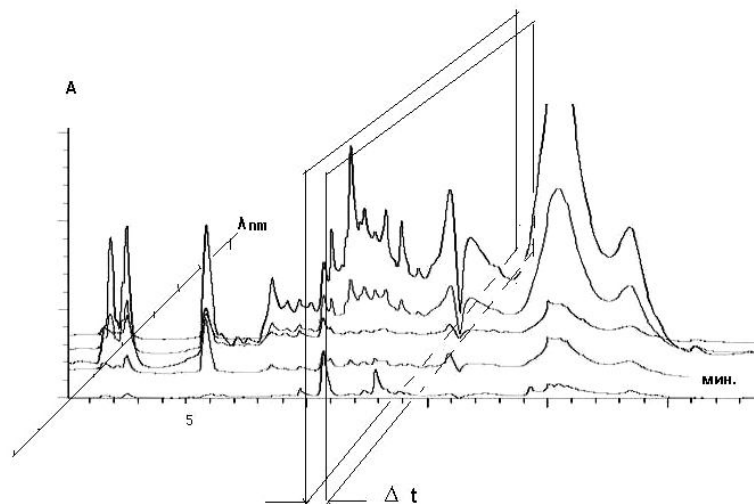


Рис. 2. Временные срезы многоволновых хроматограмм по координате длин волн: Δt – временной интервал между срезами

Подтверждение справедливости гипотезы о существовании ограниченного количества линейнонезависимых базовых спектров поглощения многокомпонентных систем

В процессе исследования вин выяснилось, что для большинства многоволновых хроматограмм вин, полученных при детектировании в УФ-диапазоне, всегда получались четыре, максимум пять, значащих линейно-независимых факторов, которые и использовались для определения идентичности, подобия и не идентичности вин. Однако не до конца разработанная математика метода не давала возможность теоретически определить: это случайность, ошибка или объективные показатели данной системы?

Прояснить этот вопрос можно следующим образом: резко увеличить число используемых длин волн внутри используемого УФ-диапазона, а, следовательно, и размерность матрицы экспериментальных данных. Если произойдет существенное увеличение числа линейно независимых факторов, то гипотеза, скорее всего, неверна или нуждается в какой-либо корректировке. Если этого не произойдет, то гипотеза имеет право на жизнь.

Для проверки исходной гипотезы была использована матрица кодов оптических плотностей вин «Каберне» (совхоз «Крымский»), красное сухое столовое вино «El Patio» и смесь вин «Каберне» и «El Patio» в соотношении 85:15 по объему, каждая размером 15x400, где 15 (количество строк) – длины волн 200, 210, 220, 240, 254, 260, 270, 280, 290, 300, 310, 320, 330, 340, 354 нм.

Так как детектор хроматографа «Милихром-5-М» может производить только пятиволновое детектирование, хроматограммы были получены при следующих условиях детектирования: 200, 210, 220, 230, 240 нм; 220, 230, 240, 254, 260 нм; 270, 280, 290, 300, 310 нм; 270, 320, 330, 340, 354 нм.

Каждая из хроматограмм была переведена в цифровой вид и нормирована к 400-м столбцам кодов оптической плотности. Таким образом, для проведения расчетов использовались матрицы кодов оптических плотностей хроматограмм вин размером от 5x400 до 15x400. Результаты обработки матрицы 15x400 приведены в табл. 1.

Таблица 1. Вклады факторов, полученных их матриц 15x400 для вин «Каберне», «El Patio» и их смеси. Вклады факторов, %

№ фактора	«Каберне»*	«El Patio»*	Смесь вин
F ₁	91.80	88.17	81.89
F ₂	4.66	6.70	11.81
F ₃	1.42	1.86	2.69
F ₄	1.14	1.13	1.57
F ₅	<u>0.42</u>	0.98	<u>0.62</u>
F ₆	0.22	0.54	0.53
F ₇	0.12	0.20	0.27
F ₈	0.10	0.13	0.23
F ₉	0.04	0.08	0.11
F ₁₀	0.03	0.07	0.10
F ₁₁	0.02	0.04	0.09
F ₁₂	0.02	0.03	0.04

Из табл. 1 следует, что в результате увеличения размерности исходных матриц с 5x400 до 15x400 четыре фактора описывают 99.02% сортового вина и 97.86 % специального (с добавками - «El Patio»). Приведенные результаты достаточно простой проверки говорят о том, что исходная (первоначальная) гипотеза верна.

Обобщенный вариант гипотезы проверялся путем испытания разнообразных по происхождению МС. Для биообъектов была выбрана достаточно спорная ситуация вокруг синдрома эндогенной интоксикации, попытки разрешения которой делятся уже около 50 лет. Применяя комбинацию «ВЭЖХ – МГК» удалось существенно продвинуться в этой области и со сдержанным оптимизмом прогнозировать определенные успехи в этой области [25, 26].

Начало работы по хемосистематике растений, являющейся подготовительной к исследованию в области теории катастроф, также пока не принесло отрицательных результатов [24].

Заключение

Дальнейшее совершенствование и применение комбинации «ВЭЖХ – МГК», базирующейся на предложенных гипотезах, существенно расширяет возможности исследователя при распознавании МС, сильно упрощая, удешевляя и ускоряя процесс контроля МС и выделения полезной информации из информационного «шума».

Авторы благодарят Варбана Пешкова, Леонида Сапрыкина и Олега Рудакова за всемерную многолетнюю помощь и поддержку выполненных по этой теме работ.

Список литературы

1. Гаврилина В.А., Сычев С.Н. Информационная поддержка контроля качества вина // Материалы V международной научно-технической конференции «Информационные технологии в науке, образовании и производстве». 2012, С.9.

2. Пешков Варбан. Отзыв на монографию Сычева С.Н. и Гаврилиной В.А. Метод контроля качества вина с помощью комбинации «ВЭЖХ-МГК» // LAMBERT Academic Publishing. 2012. 112 с.
3. Померанцев А.Л. Метод главных компонент // Российское хемометрическое общество. 2009.
4. Иберла К. Факторный анализ // М.: Статистика. 1980. 398 с.
5. Сычев С.Н. Методы совершенствования хроматографических систем и механизмы удерживания в ВЭЖХ // Орел: ОрелГТУ, 2000. 212 с.
6. Сычев С.Н., Ерилов С.П., Никитин Ю.С. Применение метода главных компонент в жидкостной хроматографии замещенных бензойных кислот // Журн. физ. химии. 1983. Т.59. № 2. С. 346-350.
7. Сычев С.Н., Ерилов С.П., Никитин Ю.С. Устойчивость линейно независимых факторов, описывающих экспериментальные данные по жидкостной хроматографии монозамещенных бензойных кислот // Журн. физ. химии. 1983. Т.59., № 3., С. 661-663.
8. Nieawdorp G.H.E. An Improved Procedure for The Description of Sabstituent Effects on Equilibrium and Rate Constants, based on Factor Analysis of the whole Body of Appropriate Data. // J.Chem.Soc. Perkin 2. 1979. N 5. P. 537 – 540.
9. Сычев С.Н., Сычев К.С., Ланин С.Н. Выявление характера межмолекулярных взаимодействий однозамещенных алифатических соединений методом главных компонент // Вестник МГУ, 2000. Сер.2. Химия. Т.41. № 1. С.28-31.
10. Lopez L.A., Rutan S.C. Comparison of methods for characterization of reversed-phase liquid chromatographic selectivity. // J. Chromatogr. A. 965. 2002. P.301-314.
11. Сычев С.Н. Свертка информации при анализе веществ и объектов // Контроль. Диагностика. 2005. № 1. С. 70 – 72.
12. Сычев С.Н. Подмастерьев К.В., Пахолкин Е.В. и др. Применение метода главных компонент для выявления межмолекулярных характеристик, входящих в явление когезии // Фундаментальные и прикладные проблемы техники и технологии.- №3, 2010, С. 99 – 105.
13. Самохин А.С., Ревельский И.А. Применение метода главных компонент для выделения «чистых» масс-спектров в газохроматографическом масс-спектральном анализе // Масспектрометрия. 2010. № 7. С. 34 – 38.
14. Davankov V.A., Sychov C.S., Ilyin M.M. et al. Hypercrosslinked polystyrene as a novel type of high-performance liquid chromatography column packing material. Mechanisms of retention // J. Chromatography A. 2003. V.987. P. 67-75.
15. Sychov C.S., Ilyin M.M., Davankov V.A. et al. Elucidation of retention mechanisms on hypercrosslinked polystyrene used as column packing material for high-performance liquid chromatography. // J. Chromatography A. 2004. V.1030. P. 17-24.
16. Kalambet Yu.A., Kozmin Yu.P., Perelroyzen M.P. Computer spectrochromatography: Principles and practice of multi-channel chromatographic data processing // Journal of Chromatography A. Volume 542. 1991. P. 247-261.
17. Барам Г.И., Болванов Ю.А., Грачев М.А. и др. Микроколоночная жидкостная хроматография с многоволновой детекцией. Микроколоночный жидкостной хроматограф «Обь-4» // Новосибирск: ИЯФ СО АН СССР. 1983. 38 с.
18. Сычев С.Н., Гаврилина В.А., Сычев К.С. Высокоэффективная жидкостная хроматография на микроколоночных жидкостных хроматографах серии «Милюхром» // Орел: ОрелГТУ, 2002. 135 с.
19. Гаврилина В.А., Мальцева О.И., Сычев С.Н. Применение метода главных компонент для идентификации и сравнения натуральных вин // Виноделие и виноградарство. 2007. № 1. С.10-12.

20. Гаврилина В.А. Мальцева О.И., Сычев С.Н. Применение метода главных компонент для идентификации и сравнения натуральных вин. Часть 2 // Виноделие и виноградарство. 2007. № 3. С.30-32.

21. Гаврилина В.А. Мальцева О.И., Сычев С.Н. Применение метода главных компонент для идентификации и сравнения натуральных вин. Часть 3 // Виноделие и виноградарство. 2007. № 4. С. 18-19.

22. Гаврилина В.А., Сычев С.Н. Комбинированная экспертная система контроля подлинности вин // Фундаментальные и прикладные проблемы техники и технологии. 2010. № 2. С.120-125.

23. Сычев С.Н., Гаврилина В.А., Полянский К.К. Получение идентификационных признаков экстракта стевии комбинацией методов многоволновой ВЭЖХ и главных компонент. // Хранение и переработка сельхозсырья. 2005. № 12. С. 36-39.

24. Гаврилина В.А., Сычев С.Н. Перспективы применения комбинированного метода «высокоэффективная жидкостная хроматография – метод главных компонент в хемосистематике растений // Фундаментальные и прикладные проблемы техники и технологии. 2011. № 6-2. С.11-15.

25. Гаврилина В.А., Сычев С.Н. Применение комбинации «высокоэффективная жидкостная хроматография – метод главных компонент» для распознавания синдрома эндогенной интоксикации. Часть 1 // Сорбционные и хроматографические процессы, 2012. Т.12. Вып. 5. С. 814-821.

26. Гаврилина В.А., Сычев С.Н. Применение комбинации «высокоэффективная жидкостная хроматография – метод главных компонент» для распознавания синдрома эндогенной интоксикации. Часть 2 // Сорбционные и хроматографические процессы, 2012. Т.12. Вып. 5 С. 822-827.

Гаврилина Вера Александровна – к.т.н., доцент кафедры «Химия», инженер лаборатории «Высокоэффективная жидкостная хроматография», Государственный университет УНПК, Орел

Сычев Сергей Николаевич – д.т.н., профессор кафедры «Химия», заведующий лабораторией «Высокоэффективная жидкостная хроматография», Государственный университет УНПК, Орел

Gavrilina Vera A. – Ph.D., Associate Professor of Department Chemistry", Engineer Lab "HPLC", UNPK State University, Orel, e-mail: chemistry@ostu.ru

Sychev Sergey N. – Doctor of Technical Sciences, Professor, Department of Chemistry, Head of Laboratory "HPLC", UNPK State University, Orel