

УДК 543.544.45:543.27.7

РАСЧЕТ ТЕМПЕРАТУР КИПЕНИЯ УГЛЕВОДОРОДОВ С ИСПОЛЬЗОВАНИЕМ ИНДЕКСОВ УДЕРЖИВАНИЯ И МОЛЕКУЛЯРНЫХ МАСС

Д.Ю. Полухин, Ю.С. Яшин, И.А. Ревельский, О.В. Напалкова, Н.А. Пасекова, П.П. Вульх

(кафедра аналитической химии)

В работе показана возможность определения температур кипения для таких классов углеводородов, как *n*-алканы, изопарафины, нафтены, олефины и арены с использованием индексов удерживания, молекулярных масс и алгоритма, основанного на принципах самоорганизации моделей на ЭВМ.

Современная капиллярная газовая хроматография (КГХ) характеризуется высокой воспроизводимостью времен удерживания. Высокая воспроизводимость этих величин позволяет использовать индексы удерживания (ИУ) для надежной идентификации веществ. Помимо этого появляется реальная возможность плодотворного изучения зависимостей между физико-химическими свойствами веществ и их хроматографическими параметрами удерживания. Известно, что порядок выхода для углеводородов (УВ) на неполярной неподвижной фазе определяется температурами кипения УВ: вещества выходят строго в порядке возрастания их температур кипения; известные отклонения лежат в пределах погрешности определения температуры кипения. Для большей части гомологических рядов считается применимой линейная зависимость

$$I_x = a \cdot T_b + b, \quad (1)$$

где коэффициенты a и b определяют по методу наименьших квадратов (МНК); I_x – индекс удерживания; T_b – температура кипения. Линейную формулу применяют только для гомологических рядов, причем в зависимость не укладываются первые несколько членов, особенно первый, и последние. Использование формулы (1) для многих соединений приводит к большой погрешности при оценке I_x и T_b , при этом авторы [1, 2, 3] особо подчеркивают принципиальную нелинейность зависимости $I_x = f(T_b)$ и недопустимость использования линейной зависимости даже для оценки ИУ. Авторы [1, 2] для оценки I_x рекомендуют использование трех- или четырехпараметрических нелинейных уравнений, например

$$\lg I_x = a \lg T_b + bA + c, \quad (2)$$

где I_x – индекс удерживания; T_b – температура ки-

пения, °С; A – любое аддитивное свойство гомологов: число атомов углерода, молекулярная масса и т.д.; a , b и c – коэффициенты, вычисляемые по МНК. В работе [4] приведены уравнения и коэффициенты для определения температур кипения в таких рядах, как парафины, изопарафины, нафтены, олефины и арены с расхождением расчетной и справочной температур кипения в среднем на 2 – 5°, однако авторами было изучено небольшое число соединений. Это объясняется тем, что исследования проводили в гомологических рядах или достаточно узких группах соединений, где количество членов невелико. В большинстве работ математическое моделирование проводили таким образом, что полученные модели или уравнения не имели оптимальной сложности (или имели, но только для данной хроматографической системы и температурной программы). Обобщение уравнений на все хроматографические системы и температурные программы может приводить к большим ошибкам. Это объясняется тем, что используемые модели не являются физическими, поэтому требуется не простой пересчет коэффициентов, а некоторый динамический подход к адаптации моделей для конкретных условий, например, изменение сложности модели или вообще пересмотр модели, следовательно, для каждой системы надо начинать весь процесс поиска моделей с нуля.

Вторым существенным недостатком известных из литературы уравнений является то, что авторы не производили оценку доверительного интервала. В большинстве работ приведены таблицы расхождения расчетных температур кипения и известных из справочника для нескольких десятков соединений, и только на основе этих данных приходилось судить об адекватности того или иного уравнения. Авторы обычно судили об эффективности модели по величине дис-

Таблица 1

Расчет температур кипения по калибровочной смеси с использованием молекулярных масс и экспериментальных индексов удерживания

Углеводород	ИУ, I	M	$T_{\text{квт}}, ^\circ\text{C}$	$T(I), ^\circ\text{C}^*$	$T - T(I)$	$T(I, M), ^\circ\text{C}^{**}$	$T - T(I, M)$
4-Метил-цис-пентен-2	573.00	84.16	58.60	56.83	1.77	58.46	0.14
4,4-Диметил-транс-пентен-2	622.90	98.19	76.73	72.87	3.86	76.00	0.73
2,2-Диметилпентан	624.85	100.25	79.19	73.48	5.71	76.92	2.27
1,1-Диметилциклогексан	672.29	98.19	87.48	88.07	0.59	88.56	1.08
Толуол	757.61	92.14	110.63	112.62	1.99	108.64	1.99
3,4-Диметилгексан	773.54	114.23	117.73	116.97	0.76	117.26	0.47
1,3-Октадиен	782.05	110.20	121.45	119.26	2.19	120.41	1.04
Октен-1	790.00	110.20	123.50	121.39	2.11	121.66	1.84
2,3,5-Триметилгексан	823.33	128.26	131.35	130.11	1.24	130.94	0.41
1цис,2-Диметилциклогексан	825.67	112.22	129.74	130.70	0.96	129.71	0.03
изо-Бутилциклопентан	892.70	126.24	147.97	147.33	0.64	147.29	0.68
5-Метилнонан	969.19	142.29	165.10	165.16	0.06	165.25	0.15
1-Метил-2-этилбензол	970.31	120.17	165.18	165.41	0.23	165.43	0.25
изо-Бутилциклогексан	986.88	140.27	171.30	169.15	2.15	169.40	1.90
1-Метил-2-н-пропилбензол	1057.50	134.22	184.97	184.85	0.12	185.92	0.95
н-Пентилбензол	1148.80	148.24	205.40	205.24	0.16	205.29	0.11

* Доверительный интервал (99%): $\pm 7.83^\circ\text{C}$.** Доверительный интервал (99%): $\pm 2.97^\circ\text{C}$.

Таблица 2

Расчет температур кипения для парафинов с использованием формальных ИУ и молекулярных масс

Соединение	ИУ, I	M	$T_{\text{кип}}$, °C	$T(I, M)$, °C*	$T - T(I, M)$
Метан	100	16.043	-161.5	-115.34	46.16
Этан	200	30.07	-88.6	-74.89	13.71
Пропан	300	44.097	-42.04	-36.55	5.49
Бутан	400	58.124	-0.5	-0.31	0.19
Пентан	500	72.151	36.06	33.22	2.84
Гексан	600	86.178	68.73	65.87	2.86
Гептан	700	100.205	98.42	95.81	2.61
Октан	800	114.232	125.68	123.64	2.04
Нонан	900	128.259	150.82	149.38	1.44
Декан	1000	142.286	174.15	173.02	1.13
Ундекан	1100	156.313	195.9	194.55	1.35
Додекан	1200	1703	216.3	213.98	2.32
Тридекан	1300	184.47	235.4	231.29	4.11
Тетрадекан	1400	198.39	250	246.54	3.46

* Доверительный интервал $\pm 2.78^\circ\text{C}$ (99%).

персии, рассчитанной на основании выборки, по которой производилось построение модели. Такой подход совершенно недопустим, так как достаточно взять модель с N коэффициентами, которая может принимать любые значения в исследуемой области, и через любые N точек можно провести кривую с нулевой дисперсией.

В настоящей работе при помощи алгоритма, основанного на принципах самоорганизации моделей на ЭВМ, изучали возможности увеличения точности определения температуры кипения различных УВ при совместном использовании данных по ИУ и молекулярным массам.

Экспериментальная часть

Работу проводили на газовом хроматографе "MEGA-2" (Fisons Ins.) с пламенно-ионизационным детектором. Для разделения УВ использовали кварцевую капиллярную колонку (100м x 0.25мм) с фазой SE-30 и толщиной пленки 0.2 мкм (Chrompack, USA). Параметры удерживания углеводородов определяли в режиме программирования температуры: изотерма 35° (в течение 13 мин), нагрев до 60° со скоростью 1.1 град/мин (изотерма при этой температуре 19 мин) и затем нагрев со скоростью 2 град/мин до 250°.

Температура детектора и испарителя составляла 250°. Скорость потока газа-носителя (гелия) через колонку составляла 1 мл/мин, воздуха – 300 мл/мин. Ввод пробы проводили в режиме деления потока 1: 200, объем пробы составлял 0.2 мкл.

Регистрацию и обработку хроматограмм осуществляли с помощью системы сбора и обработки хроматографических данных "ChromCard-1.22" (Carlo Erba, Italy), позволяющей получить данные по временам удерживания и площадям пиков компонентов смеси. Расчет индексов удерживания компонентов смеси и индивидуального состава бензина проводили с помощью программного пакета PIANO (Analytical Automation Specialist, USA), рекомендованного ASTM для определения структурно-группового и фракционного состава нефтепродуктов. Данный программный пакет содержит базу данных (ИУ, молекулярная масса, плотность, октановое число и др.) для 362 УВ, таких как парафины, изопарафины, нафтены, олефины и арены, входящих в состав бензинов. Температуры кипения известны для 241 УВ. Для расчета температур кипения нами была написана программа на BORLAND C++ , реализующая алгоритм, основанный на принципах самоорганизации моделей, модернизированный нами для данной конкретной задачи.

Обсуждение результатов

Авторами настоящей работы была написана программа, реализующая усовершенствованный алгоритм, основанный на принципах самоорганизации моделей на ЭВМ. Согласно этому алгоритму, по ряду внешних критериев, позволяющих отыскать модель оптимальной сложности, осуществляется целенаправленный перебор постепенно усложняющихся моделей. Сложность модели определяется числом отличных от нуля коэффициентов, подлежащих оценке по МНК. Например, чем больше число и степень слагаемых полиномиальной модели, тем больше сложность ее структуры. Оказывается, что при постепенном повышении сложности моделей, претендующих на выбор, внешний критерий, по которому осуществляется перебор, вначале уменьшается, проходит через минимум, а затем увеличивается [5]. Минимум критерия определяет модель оптимальной структуры, в чем и состоит принцип самоорганизации моделей. Искать оптимальную модель в области более сложных моделей уже бессмысленно.

Критерий перебора моделей называется внешним, если он основан на некоторой дополнительной информации, например на новых данных, не учитывавшихся при оценке модели. Именно внешние критерии дают

характерные минимумы. Обычно применяемый в математической статистике критерий минимума среднеквадратической ошибки, рассчитанной на всех точках выборки данных, таких минимумов не дает, поэтому для самоорганизации моделей он не пригоден. Такой критерий называют внутренним [5]. По внутренним критериям чем сложнее модель, тем точнее. Это и стало девизом имитационного моделирования. Внешний критерий можно получить, например, если часть данных использовать только для сравнения между собой моделей-претендентов.

Основным критерием при самоорганизации моделей является критерий непротиворечивости моделей. Этот критерий требует, чтобы модели, построенные по подвыборкам генеральной выборки, были не противоречивы, т.е. прогноз по одной подвыборке не должен противоречить прогнозу по другой подвыборке. Требование непротиворечивости может быть сформулировано с помощью различных математических выражений. Наиболее распространена следующая формула, называемая критерием минимума смещения:

$$n_{cm}^2 = \sum (Y_A - Y_B)^2 \rightarrow \min, \quad (3)$$

где Y_A – величина модели при оценке коэффициентов на подвыборке А; Y_B – величина модели при оценке коэффициентов на подвыборке В (А и В не пересекаются, и объединение А и В является полной выборкой с N точками). Наилучшие результаты можно получить,

если использовать не один, а несколько внешних критериев и оценку эффективности модели проводить по свертке из нескольких критериев. Свертка представляет собой сумму квадратов внешних критериев с соответствующими весами [5].

Написанная нами программа разбивала входную выборку случайным образом на две подвыборки А и В, по обеим выборкам строились модели. По модели, построенной по подвыборке А, рассчитывали сумму квадратов отклонений на подвыборке В (D_1^2); по модели построенной по подвыборке В рассчитывали сумму квадратов отклонений на подвыборке А (D_2^2); рассчитывали критерий смещения по формуле (3), системный критерий (C_r) вычисляли по формуле (4)

$$C_r = n_{cm}^2 + 0.5 (D_1^2 + D_2^2). \quad (4)$$

Модель оптимальной сложности соответствовала минимуму системного критерия, и по ней проводили необходимые расчеты. Системный критерий является сверткой внешних критериев и поэтому обладает минимумом.

Тестирование написанной нами программы проводили по базе данных коммерческого пакета PIANO. Во время тестирования проверяли правильность работы программы и принципиальную возможность определения температур кипения по имеющимся данным. При тестировании программы оказалось, что использование одного индекса в качестве исходных данных дает удовлетворительные результаты только в гомологических рядах. Поэтому в качестве второй переменной было решено использовать молекулярную массу, как наиболее доступный и достоверный параметр.

Мы не использовали разбиение УВ на классы. Это дало возможность, во-первых, использовать большее число исходных данных, что позволило увеличить точность определения температуры кипения (в качестве исходных данных мы использовали более 25 точек, что существенно больше, чем в других исследованиях), во-вторых, позволило проверить адекватность найденных моделей и оценок доверительного интервала на более чем 140 соединениях.

Программа позволила провести оценку того, насколько рассчитанные по модели температуры кипения могут отличаться от действительных в том случае, если используемая выборка представительна. В качестве оценки достоверности полученных данных использовалась оценка доверительного интервала. Ее расчет проводили следующим образом:

$$\Delta = t_n \sqrt{\frac{(n_{cm}^2 + D_1^2 + D_2^2)}{n}} \quad (5)$$

Таблица 3

Расчет температур кипения на основании времен удерживания и молекулярных масс

Соединение	$t_{уд}$, мин	M	T, °C	T($t_{уд}$,M)*	T-T($t_{уд}$,M)
2-Метилбутен-1	8.15	70.13	31.16	28.7	2.46
цис-Гексен-2	13.74	84.16	68.86	69.5	0.64
Циклогексан	17.96	84.16	80.72	82.61	1.89
3-Метилгексан	20.01	100.75	91.96	91.8	0.16
Гептен-1	21.68	98.19	93.64	95.9	2.26
2,5-Диметилгексан	27.91	114.22	109.11	109.15	0.04
цис-Октен-2	41.38	258.1	125.61	126.79	1.18
изо-Бутилциклопентан	59.95	126.23	148	147.2	0.8
n-Гексилбензол	107.73	162.26	226.11	225.87	0.24

* Доверительный интервал $\pm 3.35^\circ\text{C}$ (99%).

где t_n – коэффициент Стьюдента для n точек; в программе использовали 99% -ю вероятность. Следует отметить, что данная формула может быть использована только для оценки доверительного интервала в первом приближении.

В общем виде все модели могут быть представлены следующим образом:

$$F(I, M) = C_0 + C_1 f_1(I) + C_2 f_2(M) + C_3 f_1^2(I) + C_4 f_1(I) f_2(M) + C_5 f_2^2(M) + \dots \quad (6)$$

где C_i – коэффициенты модели, подлежащие определению по МНК, сложность модели определяется числом коэффициентов, отличных от нуля, на каждом этапе усложнения модели добавляется по одному коэффициенту; $f_1(I)$ и $f_2(M)$ – любые из функций от I и M (x , $1/x$, x^2 , $\ln(x)$, $\exp(x)$). Для вычисления необходимых функций проводили линейные преобразования координат. Дальнейшее усложнение модели происходило за счет введения произведения трех функций. Оптимальными обычно оказывались модели с 4 – 6 коэффициентами, отличными от нуля.

Примеры расчета температур кипения для УВ калибровочной смеси приведены в табл. 1. Доверительный интервал не превышает $\pm 3^\circ$ в зависимости от используемых для построения модели данных и от их количества. Нами было показано, что оптимальное количество точек составляет около 25. В табл. 1 в качестве примера приведены температуры кипения, рассчитанные только на основании ИУ. Доверительный интервал в этом случае составлял около 8° .

В табл. 2 приведены расчеты температур кипения для парафинов на основании молекулярных масс и ИУ (данные по парафинам не использовали). Из табл. 2 видно, что отклонение температур кипения от действительных велико. Данный факт можно объяснить тем,

что в связи с выбранной системой индексов, ИУ парафинов не подлежат определению (они заранее известны для любой хроматографической системы и всегда постоянны). Для парафинов определяют только времена удерживания.

Табл. 3 содержит данные по температурам кипения, рассчитанным с использованием неисправленных времен удерживания вместо ИУ. Из табл. 3 видно, что точность расчета температур кипения практически не зависит от того, используются ИУ или неисправленные времена удерживания.

Таким образом, проведенное исследование показало, что при совместном использовании данных по характеристикам удерживания и молекулярным массам и предложенной нами программы, реализующей усовершенствованный нами алгоритм, основанный на принципах самоорганизации моделей на ЭВМ, возможно определение температур кипения компонентов смеси УВ (не зависимо от класса УВ по единым моделям) с погрешностью, не превышающей $\pm 3^\circ$ (99% доверительной вероятности).

СПИСОК ЛИТЕРАТУРЫ

1. *Зенкевич И.Г., Конюхова С.В., Максимов Б.Н.* //ЖФХ. 1993. №7. С. 1474.
2. *Конюхова С.В., Зенкевич И.Г., Максимов Б.Н.* //ЖАХ. 1994. №4. С. 402.
3. *Набивач В.М., Дмитриков В.П.* // Усп. Хим. 1993. №1. С. 27.
4. *Яшин Ю.С., Напалкова О.В., Ревельский И.А. и др.* //Вестн. Моск. ун-та. Сер. 2, Химия. 1997. 38. С. 57.
5. *Ивахченко А.Г., Юрачевский Ю.П.* Моделирование сложных систем по экспериментальным данным, М., 1987.

Поступила в редакцию 19. 09. 96