

ХИМИЧЕСКАЯ ТЕРМОДИНАМИКА
И ТЕРМОХИМИЯ

УДК 544.2:541.12.036:536.4.032:547.1

КОНТРОЛЬ, КОРРЕКЦИЯ И ВОССТАНОВЛЕНИЕ ЗНАЧЕНИЙ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНЫХ ЗАВИСИМОСТЕЙ

© 2021 г. И. Г. Зенкевич^{а,*}

^аСанкт-Петербургский государственный университет, Институт химии, Санкт-Петербург, Россия

*e-mail: izenkevich@yandex.ru

Поступила в редакцию 22.05.2020 г.

После доработки 22.05.2020 г.

Принята к публикации 06.06.2020 г.

На примере нормальных температур кипения и температур плавления перфторалканкарбоновых кислот $C_nF_{2n+1}CO_2H$ рассмотрен способ контроля, коррекции и, при необходимости, восстановления значений физико-химических свойств органических соединений. Его актуальность обусловлена тем, что значения большинства свойств из различных источников информации не всегда хорошо согласуются друг с другом, причем расхождения могут быть весьма значительными. На первой стадии проверки данных алгоритм может включать отбраковку наиболее заметных аномалий за счет полиномиальной аппроксимации зависимости значений физико-химических свойств (A) от числа атомов углерода в молекуле (n). Наиболее надежные результаты обеспечивает последующая аппроксимация данных с использованием рекуррентных соотношений $A(n + \Delta n) = aA(n) + b$. Для температур кипения $\Delta n = 1$, для температур плавления $\Delta n = 2$. Такой способ означает проверку справочных данных на согласованность с данными для остальных гомологов того же ряда. Для большей части гомологов из нескольких альтернативных значений различных свойств он позволяет выбрать наиболее согласующиеся с другими данными. Реже возникает необходимость исправления или исключения данных для отдельных гомологов. При этом важно отметить, что рассматриваемый способ не вносит существенных искажений в исходные массивы физико-химических характеристик, так как необходимые поправки обычно не превышают неопределенностей многих современных методов их расчета.

Ключевые слова: физико-химические свойства, перфторалканкарбоновые кислоты, нормальные температуры кипения, температуры плавления, контроль, коррекция и восстановление справочных данных, рекуррентные соотношения

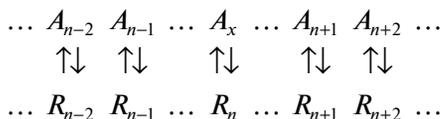
DOI: 10.31857/S0044453721040294

Определение надежных значений физико-химических свойств органических соединений во многом определило прогресс органической химии [1–5]. В число давно известных “классических” характеристик входят нормальные температуры кипения ($T_{\text{кип}}$), плавления ($T_{\text{плавл}}$), относительные плотности (d_4^{20}), показатели преломления (n_D^{20}) и другие. Две важные задачи, решаемые с их использованием – оценка чистоты органических соединений и их идентификация. Если в первом случае в настоящее время преобладают хроматографические и спектральные методы, то актуальность второй сохраняется в форме постоянного расширения и уточнения массивов справочных данных.

Такая современная база данных как Reaxys содержит информацию о приблизительно 118 млн соединений. Коллекции аналитических парамет-

ров, прежде всего масс-спектров ионизации электронами и газохроматографических индексов удерживания (например, [6]), значительно меньше. Важно заметить, что для любых характеристик создание исчерпывающе полных баз принципиально невозможно из-за многообразия объектов. В подобной ситуации допустимо использование расчетных методов (например, программного обеспечения ACD), но их точность для конкретных соединений остается невысокой (в том числе разных вариантов аддитивных схем [7–4]). Наиболее приемлемые результаты обеспечивают методы, основанные на сравнении свойств соединений разных таксономических групп (гомологических рядов) [15] между собой, или же различных гомологов в пределах одних и тех же рядов [16–18]. Если схематически изобразить алгоритмы этих типов для оценки некоторого свойства объекта A_x , то первый из них (I) основан на

взаимно-однозначном соответствии свойств группы $\{A\}$, содержащей целевой объект A_x с подробно охарактеризованной референтной группой $\{R\}$. В результате значение некоторого свойства объекта A_x может быть взаимно-однозначно пересчитано из свойств объекта R_n по единым правилам $R_i \rightleftharpoons A_i$:



Во втором подходе (II) вместо привлечения данных для групп сравнения используют закономерности вариаций различных свойств A_i в пределах тех или иных групп, в том числе гомологических рядов. Фактически, здесь для оценки свойства A_x используют данные для других членов тех же самых групп:



Однако при использовании различных сравнительных методов оценки физико-химических свойств органических соединений существует проблема, остававшаяся “недооцененной” до последнего времени [15–18]. Дело в том, что если часть данных для референтной группы $\{R\}$ в способе (I) или же для объектов той же группы $\{A_i\}$ в способе (II) ненадежна (ошибочна), то оба указанных подхода неизбежно будут приводить к неверным результатам. Таким образом, важнейшим условием применимости сравнительных методов оценки физико-химических свойств является контроль используемых массивов справочных данных с целью выявления ненадежных или ошибочных значений для их последующего исключения, коррекции или (в “идеальном” случае) – замены величинами, лучше согласующимися с данными для других соединений тех же групп.

Первое краткое дискуссионное сообщение под названием “Нуждаются ли справочные значения физико-химических свойств органических соединений в дополнительной проверке?” [19] было посвящено рассмотрению и коррекции данных по нормальным температурам кипения такого достаточно “экзотического” ряда соединений как алкилфосфины $C_nH_{2n+1}PH_2$. Проверка была основана на схеме (I) – сравнении $T_{кип}$ алкилфосфинов с соответствующими данными для гомологов подробнее охарактеризованных рядов – алкилмеркаптанов $C_nH_{2n+1}SH$ и алкиламинов $C_nH_{2n+1}NH_2$, значения $T_{кип}$ которых гораздо более надежны, с использованием трехпараметрового линейно-логарифмического уравнения [16, 17]:

$$\lg T_{кип}(\text{ряд I}) = a \lg T_{кип}(\text{ряд II}) + bY + c, \quad (1)$$

где Y – дополнительный параметр, обеспечивающий взаимно-однозначное соответствие гомоло-

гов сравниваемых рядов: молекулярная масса, число атомов углерода в молекуле, молярная рефракция и др., коэффициенты a , b , c вычисляют методом наименьших квадратов, значения $T_{кип}$ выражают в К.

В качестве заключительной стадии дополнительной проверки вычисляемых таким способом температур кипения было рекомендовано использование рекуррентной аппроксимации (2) [18, 20] с целью незначительных дополнительных вариаций значений $T_{кип}$ гомологов с числом атомов углерода в молекуле (n) и ($n + \Delta n$) по критерию максимального значения коэффициента корреляции линейной регрессии:

$$T_{кип}(n + \Delta n) = aT_{кип}(n) + b, \quad (2)$$

где $\Delta n = 1$ – разность числа атомов углерода в молекулах “соседних” гомологов, коэффициенты a , и b вычисляют методом наименьших квадратов.

Принимая во внимание важность рассматриваемой проблемы, ее обсуждение в кратком сообщении [19] нельзя признать достаточным. Прежде всего, подлежит рассмотрению схема (II) контроля массивов данных без привлечения информации для других (референтных) рядов, а сам алгоритм проверки заслуживает более детального изложения. С учетом указанных посылок настоящего сообщения посвящено проблеме контроля, коррекции и восстановления справочных значений нормальных температур кипения и (более сложная задача из-за большего межлабораторного разброса данных) – температур плавления на примере гомологов ряда перфторалканкарбоновых кислот.

ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

В качестве источников исходных значений нормальных температур кипения ($T_{кип}$) и плавления ($T_{плавл}$) перфторалканкарбоновых кислот были использованы справочные издания [21–25] и доступные интернет-источники информации. Массив найденных значений $T_{кип}$ приведен в табл. 1, а значения $T_{плавл}$ – в табл. 2.

Особенностью справочных значений многих физико-химических свойств, в том числе $T_{кип}$ и $T_{плавл}$, является различное число значащих цифр и, следовательно, их различная точность. Часть значений приведена до десятых долей градуса, особенно для летучих трифторуксусной ($T_{кип}$ 72.4°C) и пентафторпропионовой (96.4°C) кислот. Наибольшее количество данных представлено с точностью до единиц градусов, но, наряду с этим, встречаются интервалы в один (например, пентафторпропионовая кислота, 96–97°C), два (та же кислота, 95–97°C), три градуса (перфтороктановая кислота, 189–192°C) и, реже, более

Таблица 1. Исходный массив справочных значений нормальных температур кипения перфторалканкарбоновых кислот ($C_nF_{2n+1}-CO_2H$) (I) и результаты его коррекции (I) \rightarrow (IV)

n	CAS №	Исходный массив $T_{кип}$, °C (I)	Результаты коррекции (IV)	$[T_{кип}(n+1) - T_{кип}(n)]$
1	76-05-1	72, 72.4	72.4	
2	422-64-0	96–97, 96.4 , 96, 95–97, 95.5, 96–98, 97 (96.3 \pm 0.9)	96.4	24.0
3	375-22-4	119, 120 , 121.0	120	23.6
4	2706-90-3	140, 144.6 , 156, 157	140	20
5	307-24-4	157 , 190	157	17
6	375-85-9	175, 178.4	175	18
7	335-67-1	189–190, 190 , 189–192, 188, 189 (189.5 \pm 1.2)	190	15
8	375-95-1	210–218, 202.4, 218, <i>203.4</i> [203]	203	13
9	335-76-2	218, 219, 219.4	218	15
10	2058-94-8	245, 238.4	231	13
11	307-55-1	245	245	14
12	72629-94-8	–	258	13
13	376-06-7	270	270	12
14	141074-63-7	–	283	13
15	67905-19-5	294	294	11
16	57475-95-3	–	306	12
17	16517-11-6	316	316	10
18	133921-38-7	–	325*	9
19	68310-12-3	–	334*, 335**	9 (10)

Примечание. Жирным шрифтом здесь и далее выделены значения, выбранные на первой стадии проверки данных; * – значения, вычисленные рекуррентной аппроксимацией исходя из $T_{кип}$ $C_{17}F_{35}CO_2H$ (316°C); ** – значение, вычисленное по уравнению $T_{кип}(n+2) = aT_{кип}(n) + b$. Курсивом для перфторнонановой кислоты выделено расчетное значение $T_{кип}$.

широкие (перфторнонановая кислота, 210–218°C). Для некоторых высших кислот известны лишь единичные значения $T_{кип}$. При объединении литературных данных можно было бы их усреднить все значения со статистическими весами, обратно пропорциональными количеству приведенных в них десятичных значащих цифр. Однако это возможно только при выполнении условия, что все исходные данные правильные, но именно это и является предметом проверки. Следовательно, статистическая обработка всех литературных значений для каждого соединения может быть нежелательной, а задачей является выбор из них наиболее достоверных или же, при необходимости, их коррекция или замена.

Для нескольких перечисленных в табл. 1 перфторалканкарбоновых кислот достаточное количество равнозначных значений $T_{кип}$ позволяет провести их обычную статистическую обработку (при этом границы интервалов усредняли вместе с конкретными значениями). В частности, это было сделано для пентафторпропионовой (96.3 \pm 0.9°C) и перфтороктановой (189.5 \pm 1.2°C) кислот. Такая операция нужна для оценки стандарт-

ных отклонений температур кипения, которые имеют порядок около 1°C. Следовательно, как было отмечено в [20], эту оценку можно считать пределом точности результатов, так как любые расчетные методы не позволяют достичь меньшей неопределенности оценок $T_{кип}$, чем неопределенность исходных данных.

Исходный массив значений $T_{плавл}$ приведен в табл. 2. Для нескольких кислот есть возможность сравнить сведения разных источников информации, что свидетельствует об их значительно большем разбросе, чем температур кипения.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В качестве примера типичных проблем, обусловленных несопадением физико-химических свойств органических соединений из разных источников информации, можно сравнить два экспериментальных значения $T_{кип}$ 1,1,1-трихлорэтона CCl_3COCH_3 из справочника [21], основанного на базе данных Weilstein, и современного сайта ChemSpider, равные 149 и 134°C, соответственно. Выбор между ними без привлечения

Таблица 2. Исходный массив справочных значений температур плавления перфторалканкарбоновых кислот ($C_nF_{2n+1}-CO_2H$) и результаты его коррекции

<i>n</i>	Исходный массив $T_{\text{плавл}}$, °C	Результаты коррекции
1	–15.25, – 15.3 , –15.4, –15	–15.3
2	– 31.0	–31
3	– 17.5	–17.5
4	9.6*	5
5	14	19
6	33.5 , 24–30	38
7	54.9–55.6, 58–60, 55–56, 55 , 59–60, 57, 56, 55.5, 54.3	55
8	77, 71, 71–72, 65, 59.3–61.1, 69–71, 59–62, 61.5 [70]	70
9	87.4–88.2, 83.5–85.5, 76.5 [85]	85
10	97.9–100.3 [99]	99
11	112.6–114.7, 112–114 [113]	113
12	117.5–122 [120]	–
13	130.4 , 130	–

Примечание. Жирным шрифтом здесь и далее выделены значения, выбранные для первой стадии проверки данных; * – расчетная величина (при совместной обработке данных подлежит исключению в первую очередь); в квадратных скобках здесь и далее приведены отсутствующие в исходном массиве величины, близкие к средним арифметическим значениям и выбранные для дальнейшей обработки.

сведений о каких-либо структурных аналогах этого соединения невозможен. Другим примером трудно выявляемых ошибок представляется значение $T_{\text{плавл}}$ 1-октанола 16.7°C [21]. В этом случае речь идет, видимо, об опечатке (пропущен знак “минус”), так как правильное значение $T_{\text{плавл}}$ (–16.7°C) легко обнаруживается именно в результате его сравнения с данными для других 1-алканолов.

Предназначение рассматриваемого в настоящей работе алгоритма состоит в упрощении выявления подобных ошибочных или ненадежных значений с целью их коррекции или даже замены. Такая задача актуальна для многих рядов, включающих недостаточно подробно охарактеризованные гомологи, но объем журнальной публикации ограничивает выбор только одним из них. К причинам остановиться на ряде перфторалканкарбоновых кислот относятся следующие:

- перфторалканкарбоновые кислоты относятся к важным продуктам промышленного синтеза и их широко используют в качестве добавок к полимерным композициям, поверхностно-активных веществ и других целей [25–29];

- для многих фторсодержащих органических соединений типичны аномалии физико-химиче-

ских свойств и значительный разброс их значений. Это связано с высокой электроотрицательностью фтора и сложностями синтеза и очистки от примесей, что делает задачу контроля правильности их значений не только актуальной, но и достаточно наглядной;

- число охарактеризованных членов ряда не слишком велико, что упрощает рассмотрение информации; номера CAS присвоены только гомологам, содержащим от 1 (35576-88-6, “нулевой” член ряда – фтормуравьиная кислота FCO_2H – не существует при нормальных условиях [30]) до 20 атомов углерода. Экспериментальные значения $T_{\text{кип}}$ известны для гомологов, содержащих от 2 до 18 атомов углерода в молекуле (табл. 1), а $T_{\text{плавл}}$ – от 2 до 14 (табл. 2).

Коррекция температур кипения перфторалканкарбоновых кислот. Исходный массив значений $T_{\text{кип}}$ (I) для 14 гомологов рассматриваемого ряда приведен в табл. 1. Первоначальный выбор значений из исходного массива $T_{\text{кип}}$ для их последующей аппроксимации проводили по следующим правилам:

- предпочтение (по крайней мере, на первом этапе) отдавали значениям $T_{\text{кип}}$, определенным с точностью до десятых долей градуса;

- при наличии нескольких альтернативных значений выбирали значения, ближайšie к средним арифметическим величинам;

- выбираемые значения $T_{\text{кип}}$ должны быть согласованы с данными для “соседних” гомологов, т.е. больше, чем у предыдущего и меньше, чем у последующего.

Результаты нескольких стадий его последовательной коррекции целесообразно проиллюстрировать графически. На рис. 1(I) приведен график зависимости выбранных таким образом температур кипения от числа атомов углерода в молекуле, $T_{\text{кип}}(n_C)$. Такие зависимости для большинства рядов являются монотонно возрастающими нелинейными, но единых уравнений для их аппроксимации не существует. Кривая на этом рисунке соответствует простейшему полиному второй степени. Коэффициенты полинома (без стандартных отклонений), а также коэффициент корреляции (R) и значение генеральной дисперсии (S_0) приведены в подписи к рисунку. Можно заметить (особенно после представления этого рисунка в более крупном масштабе) что некоторые точки заметно отклоняются от линии регрессии и, следовательно, их целесообразно либо заменить другими значениями из исходного массива данных, либо исключить. В рассматриваемом случае выше линии регрессии расположены точки, соответствующие $T_{\text{кип}}$ перфторпентановой (144.6°C), перфторгептановой (178.4°C) и перфторундекановой кислот (238.4°C). Вполне

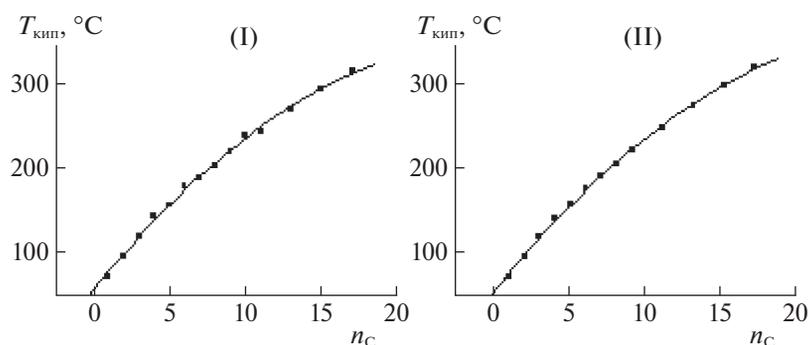


Рис. 1. Графическая иллюстрация аппроксимации исходного массива (I) значений температур кипения перфторалканкарбоновых кислот полиномом второй степени $T_{\text{кип}} = a_2 n_C^2 + a_1 n_C + a_0$ и массива после замены двух значений $T_{\text{кип}}$ альтернативными величинами (144.6 \rightarrow 140 и 178.4 \rightarrow 175) и исключения (II) одного из них (238.4). Параметры уравнений: (I) $a_2 = -0.41$, $a_1 = 22.1$, $a_0 = 55.8$, $R^2 = 0.997$, $S_0 = 3.9$; (II) $a_2 = -0.30$, $a_1 = 21.8$, $a_0 = 55.6$, $R^2 = 0.998$, $S_0 = 2.9$.

закономерно, что все три “проблемные” значения $T_{\text{кип}}$ принадлежат менее распространенным, реже используемым, и по этой причине хуже охарактеризованным кислотам с нечетным числом атомов углерода в молекуле. Первые два значения можно заменить меньшими величинами из исходного массива данных (144.6 \rightarrow 140 и 178.4 \rightarrow 175), но для последнего такой возможности нет, так что величина 238.4 на этом этапе обработки информации должна быть просто исключена. График зависимости $T_{\text{кип}}(n_C)$ после указанных преобразований приведен на рис. 1(II). Значение генеральной дисперсии S_0 , характеризующее среднюю точность аппроксимации данных в рассматриваемом диапазоне, удается незначительно снизить (с 3.9 до 2.9). Кроме того, поскольку вид зависимости $T_{\text{кип}}(n_C)$ в гомологических рядах неизвестен, дальнейшее использование полиномиальной аппроксимации не приведет к существенному улучшению результатов. Однако, как следует из ранее рассмотренного примера с алкилфосфинами [19], достижение большей точности возможно в результате применения линейных рекуррентных зависимостей вида (2).

Рекуррентные соотношения, применимые к функциям целочисленных аргументов (n), позволяют линеаризовать (с коэффициентами корреляции $R > 0.999$) нелинейные монотонные зависимости большинства физико-химических свойств органических соединений от числа атомов углерода (n_C) в молекулах гомологов, так как такой аргумент по определению не может быть нецелочисленным. Они сочетают свойства арифметических и геометрических прогрессий, а их математическими эквивалентами являются полиномы переменных степеней [16–18, 20]. Необычным свойством рекуррентных зависимостей (2) является то, что каждая точка на графике соответствует двум “соседним” значениям функции,

тогда как значения аргумента формально на таких графиках не представлены. Справедливо и обратное: вариации одного из значений функции проявляются в вариациях положений двух соседних точек на графиках рекуррентных зависимостей. Именно по этой причине они оказываются особо чувствительными к погрешностям физико-химических характеристик и могут быть эффективно использованы для контроля взаимной согласованности таких данных в пределах различных гомологических рядов. Таким образом, следующие стадии более строгого контроля массива $T_{\text{кип}}$ перфторалканкарбоновых кислот, так же как и алкилфосфинов [19], основаны на использовании именно рекуррентных соотношений.

Рекуррентная аппроксимация (2) массива данных (II) характеризуется значениями $R = 0.9993$ и $S_0 = 1.7$. Как было отмечено ранее [16–18, 20] если $R > 0.999$, то график такой зависимости визуально практически не отличим от “идеальной” прямой линии. Однако даже в таких условиях можно заметить незначительное отклонение только *одной* точки (с координатами 203–219.4) от линии регрессии, обусловленное некоторым завышением $T_{\text{кип}}$ перфтордеканкарбоновой кислоты относительно других представителей этого ряда. Этот факт предопределяет необходимость еще одной коррекции исходного массива данных: значение 219.4 следует заменить несколько меньшим, а именно другой известной для этого соединения величиной 218°C (табл. 1). Одна лишь эта замена приводит к увеличению коэффициента корреляции R до 0.9996 и уменьшению S_0 до 1.3; график полученной зависимости приведен на рис. 2(III). После этого уже не представляет особой проблемы заменить исключенную на первом этапе величину $T_{\text{кип}}$ перфторундекановой кислоты (238.4°C) значением, которое не приводит к “выбросам” соседних точек на графике (231°C). Кроме того,

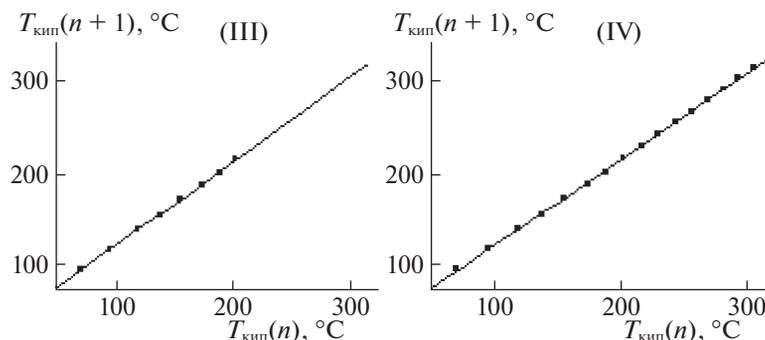


Рис. 2. Графическая иллюстрация аппроксимации температур кипения перфторалканкарбоновых кислот (на примере частично преобразованного массива данных (II)) рекуррентным уравнением $T_{\text{кип}}(n+1) = aT_{\text{кип}}(n) + b$ (III, восемь пар данных) и массива (IV) после замены одного значения $T_{\text{кип}}$ альтернативной величиной (219.4 \rightarrow 218), коррекции одного значения (231 вместо 238.4) и добавки значений $T_{\text{кип}}$ для кислот C_{13} (258), C_{15} (283) и C_{17} (306). Число пар данных увеличилось до 16. Параметры уравнений: (III) $a = 0.92 \pm 0.01$, $b = 29.7 \pm 2.1$, $R = 0.9993$, $S_0 = 1.7$; (IV) $a = 0.945 \pm 0.005$, $b = 26.5 \pm 1.2$, $R = 0.9998$, $S_0 = 1.5$.

дополнительно можно “восстановить” отсутствующие в исходном массиве значения $T_{\text{кип}}$ кислот с числом атомов углерода в молекулах 13, 15 и 17, а именно 258, 283 и 306°C, что приводит к окончательному набору (IV), график рекуррентной аппроксимации которого представлен на рис. 2(IV). Параметры уравнения линейной регрессии приведены в подписи к рисунку; имеет смысл особо отметить, что коэффициент корреляции полученных данных составляет $R = 0.9998$, а величина $S_0 = 1.5$, т. е. по сравнению с исходным массивом данных (I) уменьшилась почти в три раза.

Итоговые результаты рекуррентной аппроксимации можно использовать для вычисления других (не известных ранее) значений $T_{\text{кип}}$, например экстраполяцией для кислот C_{19} и C_{20} исходя из

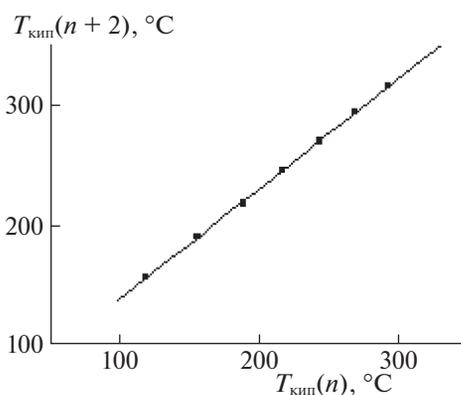


Рис. 3. Графическая иллюстрация рекуррентной аппроксимации массива (IV) температур кипения перфторалканкарбоновых кислот уравнением $T_{\text{кип}}(n+2) = aT_{\text{кип}}(n) + b$ (V, семь пар данных). Параметры уравнения: $a = 0.917 \pm 0.008$, $b = 45.8 \pm 1.7$, $R = 0.9998$, $S_0 = 1.2$.

$T_{\text{кип}}$ кислоты C_{18} (316°C), что дает для них величины 325 и 334°C, соответственно (табл. 1). Точность таких значений сопоставима с величиной S_0 , т.е. составляет около 1.5°C.

На этом этапе контроль и коррекцию значений $T_{\text{кип}}$ (I) \rightarrow (IV) можно было бы считать законченной, но целесообразно упомянуть еще два способа их дополнительной проверки. Первый из них основан на уникальных возможностях рекуррентной аппроксимации, которую можно применять для любых равноотстоящих значений аргумента, в том числе для гомологов, различающихся не на один, а на два атома углерода в молекулах [в формуле (2) $\Delta n = 2$]. В результате получаем еще один вариант линейной рекуррентной аппроксимации данных (массив V), график которого представлен на рис. 3. Из параметров уравнения регрессии (в подписи к рисунку) заслуживают упоминания высокие значение $R = 0.9998$ и точность аппроксимации $S_0 = 1.2$.

Второй способ относится к известным в математике численным методам и основан на вычислении так называемых первых конечных разностей [31], в нашем случае $\Delta_1 T_{\text{кип}}(n+1, n) = [T_{\text{кип}}(n+1) - T_{\text{кип}}(n)]$. Если зависимости $T_{\text{кип}}(n_C)$ представляют собой монотонные возрастающие выпуклые функции ($d^2 T_{\text{кип}}/dn_C^2 > 0$), то значения $\Delta_1 T_{\text{кип}}(n+1, n)$ с увеличением n_C должны монотонно убывать. Возможные отклонения от этой закономерности могут быть обусловлены погрешностями экспериментального определения температур кипения. Из приведенных в табл. 1 значений таких разностей только в четырех случаях $\Delta_1 T_{\text{кип}}(n+2, n+1)$ превышают $\Delta_1 T_{\text{кип}}(n+1, n)$ на 1°C, и лишь в одном – на два. Этот критерий дополнительно подтверждает корректность проверяемого массива данных (IV).

Резюмируя итоги проверки исходного массива температур кипения 14 перфторалканкарбоновых кислот C_2 – C_{18} можно отметить, что он включал следующие операции: выбор из нескольких альтернативных значений $T_{\text{кип}}$ – для восьми кислот, замену нескольких близких значений средними величинами – для одной, исправление – для одной и без изменений – для четырех. Исключенных данных без возможности их коррекции нет. Кроме этого, для трех кислот с отсутствующими экспериментальными данными вычислены согласованные с остальными гомологами значения $T_{\text{кип}}$. При этом особо важно подчеркнуть, что рассматриваемый алгоритм не вносит каких-либо существенных искажений в набор физико-химических характеристик (всего одно значение $T_{\text{кип}}$ для перфторундекановой кислоты исправлено на семь градусов, $238.4 \rightarrow 231^\circ\text{C}$, что значительно меньше, чем точность многих современных методов оценки температур кипения). Наиболее частым применением рассматриваемого алгоритма оказывается выбор из нескольких альтернативных значений физико-химических характеристик каждого из гомологов таких, которые лучше всего согласованы с данными для остальных членов ряда.

Коррекция температур плавления перфторалканкарбоновых кислот представляет собой гораздо более сложную задачу из-за существенно большего разброса литературных справочных значений $T_{\text{плавл}}$ (табл. 2), затрудняющего даже первоначальный выбор значений для аппроксимации. На рис. 4 представлена графическая иллюстрация зависимости температур плавления перфторалканкарбоновых кислот в зависимости от числа атомов углерода в молекуле (все доступные данные). Важно отметить, что если для $T_{\text{кип}}$ отклонения разных знаков от средних величин приблизительно равновероятны, то разброс $T_{\text{плавл}}$ в значительной степени обусловлен наличием примесей в веществах. Этот фактор объясняет преимущественные отклонения $T_{\text{плавл}}$ от значений для препаратов высокой степени чистоты исключительно в сторону их понижения [2]. Отсюда следует важное условие выбора наиболее вероятных справочных значений $T_{\text{плавл}}$: предпочтение следует отдавать наибольшему из известных величин.

Как и в случае $T_{\text{кип}}$, среди значений $T_{\text{плавл}}$ встречаются числа, приведенные с точностью до десятых долей градуса (трифторуксусная, перфтордекановая и перфтортетрадекановая кислоты), до целых градусов (большая часть данных табл. 2) и в достаточно широких интервалах (например, 24 – 30°C для $C_6F_{13}CO_2H$). Встречаются примеры значительного несоответствия данных разных источников. Так, для перфторнонановой кислоты известны значения $T_{\text{плавл}} = 59$ – 62 и 77°C , а для перфтордекановой – 87.4 – 88.2 и 76.5°C . Тем

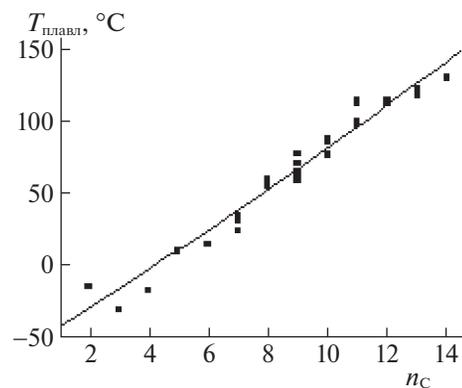


Рис. 4. Графическая иллюстрация зависимости температур плавления перфторалканкарбоновых кислот в зависимости от числа атомов углерода в молекуле (исходный набор данных).

не менее, из исходного массива данных в результате одной или нескольких попыток можно выбрать некоторый первоначальный набор значений $T_{\text{плавл}}$ для его последующей рекуррентной аппроксимации (числа, выделенные жирным шрифтом в табл. 2). Главной особенностью рекуррентной аппроксимации температур плавления является то, что уравнение (2) при $\Delta n = 1$ для этих физико-химических характеристик не выполняется; они подчиняются рекуррентному соотношению при $\Delta n = 2$. Это обусловлено необходимостью компенсации возможного альтернирования значений $T_{\text{плавл}}$ для гомологов с четным и нечетным числом атомов углерода в молекуле [32]:

$$T_{\text{плавл}}(n+2) = aT_{\text{плавл}}(n) + b. \quad (3)$$

Вторая особенность аппроксимации $T_{\text{плавл}}$ – необходимость исключения данных для простейших гомологов многих рядов, в данном случае – трифторуксусной кислоты. График зависимости (3) для исходного набора $T_{\text{плавл}}$ приведен на рис. 5а; соответствующий коэффициент корреляции $R = 0.992$ нельзя признать достаточным, что и обуславливает необходимость дополнительной коррекции некоторых данных. Такая коррекция включает замену двух экспериментальных значений большими величинами в соответствии с отмеченным выше основным свойством $T_{\text{плавл}}$ (наличие примесей приводит только к их понижению). Естественно, это условие не распространяется на расчетную величину 9.6°C для перфторпентановой кислоты, заменой которой в результате оказалось меньшее значение 5°C . Кроме того, как можно видеть из рис. 5а, две последние точки рекуррентной зависимости расположены несколько ниже линии регрессии. Это означает, что величины $T_{\text{плавл}} 120$ и 130°C для кислот C_{13} и C_{14} занижены, вероятно из-за влияния примесей,

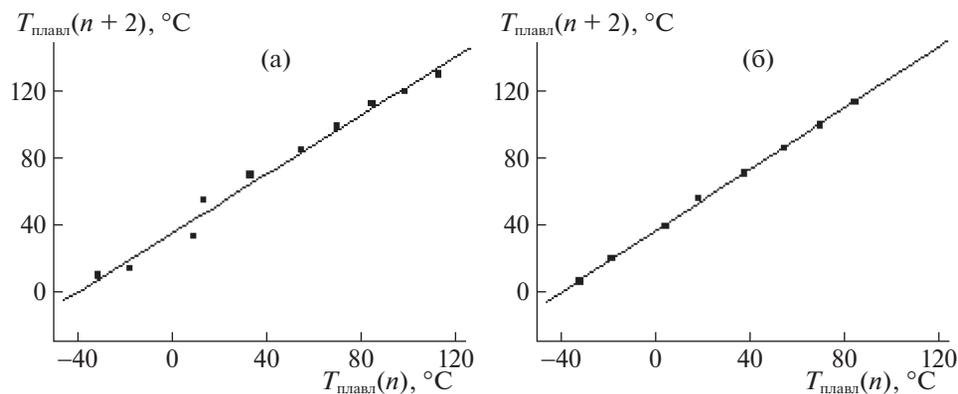


Рис. 5. Графическая иллюстрация рекуррентной аппроксимации исходного (а) массива температур плавления перфторалканкарбоновых кислот уравнением $T_{\text{плавл}}(n+2) = aT_{\text{плавл}}(n) + b$ (десять пар данных) и после его коррекции (б): два экспериментальных значения $T_{\text{плавл}}$ скорректированы в сторону увеличения, а два максимальных — исключены из рассмотрения. Параметры уравнения: (а) $a = 0.88 \pm 0.04$, $b = 34.9 \pm 2.5$, $R = 0.992$, $S_0 = 5.8$; (б) $a = 0.92 \pm 0.01$, $b = 34.7 \pm 0.6$, $R = 0.9995$, $S_0 = 1.4$.

следовательно, ненадежны и подлежат исключению. В результате получаем набор данных, рекуррентная аппроксимация которого характеризуется коэффициентом корреляции $R = 0.9995$ и средней точностью $S_0 = 1.4$; соответствующий график приведен на рис. 5б, параметры уравнения указаны в подписи к рисунку.

Если необходимо оценить $T_{\text{плавл}}$ кислот C_{13} и C_{14} , то для них можно использовать рекуррентное уравнение (3) и значения $T_{\text{плавл}}$ кислот C_{11} и C_{12} . Тогда получаем:

перфтортридекановая кислота: $0.92 \times 99 + 34.7 \approx 126^\circ\text{C}$;

перфтортетрадекановая кислота: $0.92 \times 113 + 34.7 \approx 139^\circ\text{C}$.

Средняя точность таких оценок имеет порядок величины S_0 , т.е. около 1.5°C . Таким образом, коррекция исходного массива температур плавления 13 перфторалканкарбоновых кислот C_2 – C_{14} включала в себя следующие операции: выбор из нескольких альтернативных значений $T_{\text{плавл}}$ — для двух кислот, замена нескольких близких значений средними величинами — для четырех, исправление данных — для трех (включая расчетную величину) и для трех — исключение. Два значения $T_{\text{плавл}}$ оставлены без изменений. Однако, как и при рассмотрении $T_{\text{кип}}$, важно отметить, что рассматриваемый алгоритм не вносит каких-либо существенных искажений в набор физико-химических характеристик: всего два значения $T_{\text{плавл}}$ для перфторгексановой и перфторгептановой кислот исправлены на пять градусов каждое, то есть $14 \rightarrow 19$ и $33.5 \rightarrow 38^\circ\text{C}$, что несравнимо меньше, чем точность многих известных методов оценки температур плавления [33]. Самым распространенным применением рассматриваемого

алгоритма является выбор из нескольких альтернативных значений таких физико-химических характеристик каждого из гомологов, которые лучше всего согласованы с данными для остальных членов ряда.

Рассмотренные операции контроля, коррекции и восстановления значений физико-химических характеристик органических соединений могут быть распространены на иные свойства органических соединений и на гомологи других рядов.

СПИСОК ЛИТЕРАТУРЫ

1. Fieser L.F., Fieser M. Organic Chemistry. Lexington: D.C. Heath & Co. Publ. 1950. 1125 p.
2. Органикум. Практикум по органической химии / Пер. с нем. М.: Мир, 1979. В 2 тт.
3. Jochum C., Hicks M., Sunkel J. (Eds.). Physical Property Prediction in Organic Chemistry. Berlin: Springer-Verlag, 1988. 554 p. <https://doi.org/10.1007/978-642-74140-1>
4. Reinhard M., Drefahl A. Handbook for Estimating Physicochemical Properties of Organic Compounds. Hoboken: Wiley-Intersci., 1998. 228 p.
5. Mackay D., Shin W.Y., Ma K.-C., Lee S.C. Physico-Chemical Properties and Environmental Fate for Organic Chemicals. Boca Raton: Taylor & Francis, 2006. 919 p.
6. The NIST 17 Mass Spectral Library (NIST17/2017/EPA/NIH). Software/Data Version (NIST17); NIST Standard Reference Database, Number 69, June 2017. National Institute of Standards and Technology, Gaithersburg, MD 20899: <http://webbook.nist.gov> (Accessed: June 2020).
7. Screttas C.G., Heropoulos G.A. // J. Org. Chem. 1993. V. 58. № 14. P. 3654.

8. *Stein S.E., Brown R.L.* // J. Chem. Inf. Comput. Sci. 1994. V. 34. № 3. P. 581.
9. *Katritzky A.R., Mu L., Lobanov V.S., Karelson N.* // J. Phys. Chem. 1996. V. 100. P. 10400.
10. *Rucker G., Rucker C.* // J. Chem. Inf. Comput. Sci. 1999. V. 39. P. 788.
11. *Tomovic Z., Gutman I.* // Ibid. 2001. V. 41. P. 1041.
12. *Lim T.-C.* // Acta Chim. Slov. 2004. V. 51. P. 611.
13. *Rucker C., Meringer M., Kerber A.* // J. Chem. Inf. Comput. Sci. 2004. V. 44. P. 2070.
14. *Gonzalez M.P., Toropov A.A., Duchowicz P.R., Castro E.A.* // Molecules. 2004. V. 9. P. 1019.
15. *Каранетьяни М.Х.* Методы сравнительного расчета физико-химических свойств. М.: Наука, 1965. 401 с.
16. *Зенкевич И.Г.* // Журн. физ. химии. 1996. Т. 70. № 1. С. 33.
17. *Зенкевич И.Г.* // Журн. орган. химии. 1998. Т. 34. № 10. С. 1458.
18. *Зенкевич И.Г.* // Журн. физ. химии. 2008. Т. 82. № 5. С. 695.
<https://doi.org/10.1134/S0036024408050014>
19. *Zenkevich I.G.* // Вест. СПбГУ. Сер. физ.-хим. 2016. Т. 3 (61). Вып. 4. С. 419.
<https://doi.org/10.21638/11701.spbu.2016.409>
20. *Zenkevich I.G.* // J. Chemometr. 2010. V. 24. P. 158.
<https://doi.org/10.1002/cem.1297>
21. Свойства органических соединений. Справочник. Под ред. А.А. Потехина. Л.: Химия, 1984. 519 с.
22. Handbook of Chemistry and Physics. Ed. D.R. Lide. Boca Raton, FL: Taylor & Francis, Version 2006 (CD-ROM).
23. *Dean J.A.* Lange's Handbook of Chemistry. 15th Edn. New York: McGraw-Hill, 1999. 1291 p.
24. IUPAC-NIST Solubility Database, Version 1.1. 2007. srdata.nist.gov/solubility/IUPAC/iupac.aspx (дата обращения: май 2020 г.).
<https://doi.org/10.18434/T4QC79>.
25. Long-chain (C9 – C20) perfluorocarboxylic acids, their salts, and their precursors. Draft Ecological Screening Assessment Report. Toronto: Environ. Canada, 2010. 75 p.
26. *Hudlicky M.* Organic Fluorine Chemistry. New York: Plenum Press, 1971. 197 p.
27. *Максимов Б.Н., Барабанов В.Г., Серушкин И.Л. и др.* Промышленные фторорганические продукты. Справочник. Л.: Химия, 1990. 464 с.
28. *Kissa E. (Ed.)*. Fluorinated Surfactants and Repellents. 2nd Edn. New York: Marcel Dekker Inc., 2001. 620 p.
29. *Sznajder-Katarzynska K., Surma M., Cieslik I.* // Hindawi J. Chem. 2019. Art. ID 2717528. 20 p.
<https://doi.org/10.1155/2019/2717528>
30. *Havlas Z., Kovar T., Zagradnik R.* // J. Am. Chem. Soc. 1985. V. 107. P. 7243.
31. *Демидович Б.П., Марон И.А., Шувалова Э.З.* Численные методы анализа. М.: Наука, 1967. 368 с.
32. *Зенкевич И.Г.* // Журн. общ. химии. 2006. Т. 76. Вып. 11. С. 1816.
<https://doi.org/10.1134/S1070363206110119>
33. *Столяров Е.А., Орлова Н.Г.* Расчет физико-химических свойств жидкостей. Л.: Химия, 1976. 112 с.