

НЕКОТОРЫЕ МЕТОДИЧЕСКИЕ И ВЫЧИСЛИТЕЛЬНЫЕ ПРОБЛЕМЫ КОЛИЧЕСТВЕННОГО ФИЗИКО-ХИМИЧЕСКОГО АНАЛИЗА

Ю. В. Холин¹, С. А. Мерный², Д. С. Коняев³, А. В. Пантелеймонов⁴

Количественный физико-химический анализ (КФХА) был и остается основным способом определения термодинамических характеристик реакций, протекающих в разных средах – истинных и организованных растворах, расплавах, на поверхности твердых тел. Несмотря на долгую историю развития и применения КФХА, формирование надежных стратегий его использования и наличие программно реализованных вычислительных средств, проблематика количественного физико-химического анализа не теряет актуальности. Это связано, во-первых, с распространением КФХА на новые типы объектов, значение которых постоянно растет, а области применения расширяются. Речь идет, например, об органо-минеральных гибридных материалах, объединяющих в единое целое неорганический каркас и органический модификатор, или о комплексах в организованных средах. Во-вторых, развитие методов анализа данных, включая хемометрические, позволяет существенно повысить надежность численных оценок, получаемых в результате использования КФХА, и предопределяет внедрение в практику новых расчетных процедур. В

¹ Юрий Валентинович Холин – доктор химических наук, профессор, заведующий кафедрой химического материаловедения, проректор Харьковского национального университета имени В. Н. Каразина.

² Сергей Алексеевич Мерный – кандидат химических наук, старший научный сотрудник химического факультета Харьковского национального университета имени В. Н. Каразина.

³ Дмитрий Сергеевич Коняев – кандидат химических наук, научный сотрудник химического факультета Харьковского национального университета имени В. Н. Каразина.

⁴ Антон Витальевич Пантелеймонов – аспирант химического факультета Харьковского национального университета имени В. Н. Каразина.

настоящей работе мы ограничимся обсуждением базисных аспектов КФХА и ряда вопросов, связанных с моделированием процессов на поверхности органо-минеральных гибридных материалов.

Изучая равновесные системы методами КФХА, находят число сортов, стехиометрический состав, константы устойчивости и другие физико-химические параметры химических форм. Первичными данными КФХА являются зависимости состав–свойство [1, 2]

$$A_{lk} = \varphi(\lambda_l, n_{*k}). \quad (1)$$

где A – измеряемое свойство равновесной системы, k – номер исследуемой смеси, $1 \leq k \leq N$, N – количество изученных смесей, λ – аналитическая позиция (например, длина волны поглощающего света при спектрофотометрическом исследовании растворов), l – номер аналитической позиции, $1 \leq l \leq \Lambda$, Λ – общее число позиций, n_{*k} – вектор исходных (известных по условиям смешивания) количеств вещества реагентов в k -й исследуемой смеси, φ – некоторая функция. Построение модели комплексообразования в равновесной системе состоит из следующих шагов:

1. Планирование и осуществление эксперимента – получение первичных экспериментальных данных КФХА.

2. Структурная идентификация модели – выбор такого вида функции φ в уравнении (1), связывающем начальный состав системы с измеряемыми равновесными характеристиками A , что имеющие физический смысл параметры, например, константы устойчивости комплексов, будут подгоночными параметрами модели.

3. Параметрическая идентификация модели – определение подгоночных параметров из условия «наилучшей» (в некотором смысле) аппроксимации зависимости (1) моделью.

4. Проверка статистической адекватности модели. Если модель не адекватна эксперименту, ее пересматривают, возвращаясь к шагу 3.

5. Верификация модели или дискриминация нескольких адекватных моделей. Дискриминация предусматривает выбор на основе теоретических соображений или (и) результатов дополнительных экспериментов одной модели из нескольких адекватных. При верификации наилучшими аргументами в пользу модели

признают ее способность описать данные КФХА, полученные независимым экспериментальным методом или в другой области плана эксперимента, а также независимое подтверждение состава присутствующих в системе химических форм структурно чувствительными методами исследования.

Построение модели представляет собой сложную комплексную задачу, решение которой требует на разных этапах привлекать методы термодинамики и хемометрии. В частности, необходимо обращаться к теориям планирования эксперимента, робастного оценивания, решения математически некорректных задач, использовать статистические и нестатистические методы проверки адекватности и др.

ЭТАПЫ КФХА

1. Первичные экспериментальные данные. Основные принципы планирования эксперимента для получения первичных данных КФХА хорошо известны [3, 4]. На практике, однако, далеко не всегда удается использовать оптимальный экспериментальный метод, а границы диапазона концентраций реагентов, в котором подгоночные параметры определяются с наивысшей точностью, можно точно оценить, лишь исследуя уже готовую модель. В силу этих обстоятельств мы обсудим построение модели на основе зависимостей состав–свойство, полученных по, возможно, неоптимальному плану, считая выработку рекомендаций по изменению плана эксперимента одним из результатов исследования.

2. Структурная идентификация модели. Вид функции φ определяется тремя группами уравнений [2]: 1) уравнениями связи измеряемого свойства A с равновесным составом; 2) уравнениями материального баланса; 3) уравнениями закона действия масс (ЗДМ). Основные экспериментальные методы исследования равновесий позволяют измерять линейные комбинации равновесных концентраций химических форм. Тогда

$$A_{lk} = \sum_{i=1}^S \alpha_{li} [L_i]_k, \quad (2)$$

где L_i – реагенты, S – их количество, $[L_i]_k$ – равновесная концентрация L_i в k -й смеси, α_{ji} – известный или подлежащий определению фактор интенсивности реагента L_i для аналитической позиции λ_i . При формулировке условий материального баланса удобно реакции в системе записать в канонической форме [5]:

$$\sum_{j=1}^Y v_{ij} B_j = L_i, \quad (3)$$

где v_{ij} – стехиометрические коэффициенты, B_j – подмножество реагентов, называемое независимыми компонентами. Независимые компоненты B_j друг в друга при реакциях не переходят и, следовательно, инвариантами системы являются количества вещества компонентов. Если реакции происходят в растворах, объем которых при этом не меняется, инвариантами являются и общие (аналитические) концентрации компонентов:

$$t_j = \sum_{i=1}^S v_{ij} \cdot c(L_i) = \sum_{i=1}^S v_{ij} \cdot [L_i], \quad j = 1, 2, \dots, Y, \quad (4)$$

где $c(L_i)$ – начальные (известные по условиям смешивания реагентов) концентрации L_i . Остается записать уравнения ЗДМ или, что эквивалентно, выражения для химических потенциалов реагентов. Для реакций в неполярных растворителях и водных растворах с добавками избытка фонового электролита можно принять, что коэффициенты активности химических форм при реакциях остаются неизменными, и записывать уравнения ЗДМ в виде:

$$[L_i] = \exp \left\{ \ln \beta_i + \sum_{j=1}^Y v_{ij} \ln [B_j] \right\}, \quad (5)$$

где $[B_j]$ – равновесные концентрации независимых компонентов, β_i – концентрационная константа устойчивости химической формы L_i .

Модели со структурой, заданной уравнениями (2)-(5), применяют для описания равновесий в растворах, экстракционных равновесий, равновесий в системе раствор – осадок и др.

Для ряда объектов, однако, необходима такая модификация ЗДМ, которая учитывала бы наличие эффектов, заведомо отсутствующих в случае реакций в растворах, но весьма характерных для ассоциации низкомолекулярных реагентов («лигандов») с центрами связывания (быть может, и энергетически неэквивалентными) на поверхности, взаимодействия лигандов с макромолекулами и др.

Специалисты в области адсорбции часто обсуждают эти эффекты в терминах «энергетической неоднородности поверхности» (см, например, [6-9]), преимущественно, биографической (т.е. имманентно присущей поверхности) [10]. В случае связывания лиганда М центрами \bar{Q} с образованием комплекса $M\bar{Q}$



где K – константа сродства, по экспериментальной зависимости степеней заполнения центров Q лигандами (f , $0 \leq f \leq 1$) от равновесной концентрации лиганда ($[M]$) находят плотность распределения центров по константам сродства ($p(K) \geq 0$). Для этого решают уравнение Фредгольма первого рода:

$$f([M]) = \int_0^{\infty} \theta^{\text{local}}([M], K) \cdot p(K) dK, \quad (7)$$

где ядро интегрального уравнения $\theta^{\text{local}}([M], K)$ – локальная изотерма связывания. В простейшем случае, когда электростатическими эффектами и латеральными взаимодействиями можно пренебречь, ядро $\theta^{\text{local}}([M], K)$ сводится к уравнению изотермы Ленгмюра:

$$\theta^{\text{local}}([M], K) = \frac{K \cdot [M]}{1 + K \cdot [M]}. \quad (8)$$

Кроме биографической неоднородности, на процессы связывания лигандов могут влиять эффекты эволюционной неоднородности [2], возникающие вследствие латеральных взаимодействий, электростатического отталкивания сорбата от заряженной поверхности и др. Эволюционная неоднородность меняет сродство центров связывания к лигандам с ростом степени заполнения центров. Следует

отметить, что вид изотермы связывания не позволяет заключить, какой из типов неоднородности или оба типа имеют место [11, 12].

В химии полимеров и биохимии большее распространение получили представления о «кооперативных эффектах» [13-15]. По сути, «кооперативные эффекты» – это просто иное название эволюционной неоднородности. Действительно, как указывают в [15], «кооперативные эффекты увеличивают сродство лиганда к центру связывания (положительная кооперативность) или снижают его (отрицательная кооперативность) с заполнением центров связывания». Еще в ранних работах [16-18] и фундаментальных учебниках [19, 20] отражен тот факт, что, подобно эффектам энергетической неоднородности, одинаковые кооперативные эффекты могут быть обусловлены как взаимодействиями между соседними центрами связывания, обладающими одинаковым характеристическим сродством к лигандам, так и заполнением центров, обладающих различным сродством к лигандам.

Среди подходов, позволяющих выявить и количественно описать кооперативные эффекты, особое место занимает модель полидентатного связывания [21-25]. Эта модель представляет собой частный случай общей решеточной модели [24]. В ней реакционную поверхность рассматривают как совокупность полидентатных центров \overline{Q}_Z , каждый из которых включает Z активных групп Q , а удельные концентрации центров \overline{Q}_Z равны $\frac{t_Q}{Z}$, где t_Q – общая удельная концентрация групп. Связывание лиганда с центрами \overline{Q}_Z рассматривают как ступенчатый процесс, описываемый Z константами равновесия $\beta_i^{(Z)}$, где i – номер ступени.

Модель позволяет описывать экспериментальные изотермы связывания без излишних усложнений. Построение модели начинается с испытания малого значения Z . Соответствующие константы равновесия $\beta_i^{(Z)}$ рассчитывают, минимизируя критериальную функцию, характеризующую качество аппроксимации эксперимента моделью. Если модель с малым числом подгоночных параметров неадекватна эксперименту, увеличивают Z и выполняют новые расчеты

$\beta_i^{(Z)}$. Процедуру повторяют до нахождения адекватной модели. Такая стратегия предотвращает построение избыточных моделей, не столько извлекающих из результатов измерений содержательную информацию, сколько описывающих экспериментальный шум. Кроме того, результаты моделирования позволяют легко судить о наличии и характере кооперативных эффектов (эволюционной неоднородности). В отсутствие кооперативных эффектов (идеальное связывание) присоединение лигандов к каждому активному центру Q описывает единственная характеристическая константа равновесия. В этом случае отношения ступенчатых констант¹ (параметров модели полидентатного связывания) $K_{i+1}^{(Z)} / K_i^{(Z)}$ равны статистическим факторам. При положительной кооперативности эти отношения превосходят статистические факторы, при отрицательной – меньше их. При отрицательной кооперативности или в отсутствие кооперативных эффектов справедлива система неравенств

$$K_i^{(Z)} < K_{i-1}^{(Z)} < \dots < K_1^{(Z)} . \quad (9)$$

Нарушение хотя бы одного из неравенств четко указывает на положительные кооперативные эффекты. Еще одно достоинство модели связано с ее способностью методологически объединить трактовку равновесий в терминах кооперативных эффектов и биографической энергетической неоднородности. Известно [26-28], что у многокислотного основания B , способного присоединять y ионов H^+ , кривая рН-метрического титрования такая же, как у эквимольной смеси y гипотетических однокислотных оснований. Получены формулы, связывающие константы протонизации однокислотных оснований (K_i) с общими константами протонизации (λ_n) многокислотного основания [26-28]. Например, при $y = 3$

$$\lambda_1 = K_1 + K_2 + K_3, \lambda_2 = K_1 \cdot K_2 + K_1 \cdot K_3 + K_2 \cdot K_3, \lambda_3 = K_1 \cdot K_2 \cdot K_3. \quad (10)$$

¹ *Ступенчатые константы равновесия связаны с общими соотношением*

$$\beta_i^{(Z)} = \prod_{i=1}^Z K_i^{(Z)} .$$

Таким образом, если удастся описать изотерму связывания, предполагая наличие у независимых центров Q , присутствующих в равных концентрациях (этот частный случай модели биографической неоднородности особенно часто используют в биохимии и химии гумусовых кислот [29-31]), эту изотерму столь же адекватно воспроизведет модель полидентатного связывания, рассматривающая у-ступенчатое присоединение лигандов к у-дентатному центру. Впрочем, при необходимости получить развернутую количественную информацию о биографической энергетической неоднородности, модель полидентатного связывания оказывается недостаточной.

Нелишне также отметить то непринципиальное, но серьезно облегчающее работу обстоятельство, что для параметрической идентификации модели полидентатного связывания пригодны без какой-либо адаптации практически все компьютерные программы, разработанные для моделирования равновесий в растворах.

3. Параметрическая идентификация и анализ адекватности модели. При параметрической идентификации определяют дискретные переменные – количество сортов химических форм, их стехиометрический состав L_i , значение параметра Z в модели полидентатного связывания и т.п., а также непрерывные переменные – неизвестные константы равновесия β_i , факторы интенсивности α_{il} , плотность распределения активных центров по константам связывания $p(K)$ и др.

Задача параметрической идентификации относится к классу некорректно поставленных. В таких задачах небольшие (на уровне экспериментального шума) возмущения входных данных могут вызывать сколь угодно большие возмущения решений, и возможно множество моделей, воспроизводящих экспериментальные данные в пределах их погрешностей.

Методология и численные методы параметрической идентификации моделей разного типа сильно отличаются. История развития методов и ряд современных достижений в параметрической идентификации модели биографической неоднородности представлены в работах [6, 8, 11, 32-38]. В настоящей работе обсуждается параметрическая идентификация моделей, описывающих равновесия в

растворе, а также проблемы, связанные с определением сорбционной (ионообменной) емкости материалов.

Вследствие математической некорректности задачи параметрической идентификации и необходимости определять как дискретные, так и непрерывные переменные, невозможно создать некую универсальную стратегию построения моделей. Наиболее успешными оказываются стратегии моделирования «снизу-вверх» – от неадекватной исходной модели, через более адекватные, к одной или нескольким приемлемым. По сути, такой подход похож на пошаговую регрессию, с тем только отличием, что для назначения переменной, включаемой в модель (стехиометрического состава комплекса, учет которого должен улучшить описание эксперимента), нужны особые правила. Априорное формирование до начала обработки данных эксперимента закрытого перечня возможных химических форм [39-43] хотя и помогает в ряде случаев автоматизировать построение моделей, не гарантируют, конечно, успеха во всех случаях. Более прогрессивным представляется подход, в котором состав вводимых в модель химических форм определяют по итогам исследования пробной модели. Такой подход, основанный на теории обобщенных буферных свойств, развит А. А. Бугаевским [44]. Применяя метод Бугаевского (как, впрочем, и другие способы поиска состава химических форм), каждый раз приходится делать трудный выбор, отвечая на вопрос: указывают ли большие невязки между расчетом и экспериментом на неполноту модели либо на наличие случайных промахов (выбросов) в данных?

Когда заданы дискретные переменные – число химических форм S и их стехиометрический состав (коэффициенты ν_{ij} известны), расчету подлежат две группы параметров: а) неизвестные константы устойчивости β_i , $i = 1, 2, \dots, p$; $p \leq S$, и факторы интенсивности α_{il} для Ω реагентов L_i . Таким образом, общее число искомых параметров (размерность вектора θ , содержащего все неизвестные параметры β_i и α_{il})

$$z = p + \Lambda \times \Omega < N \times \Lambda. \quad (11)$$

Расчет констант устойчивости и факторов интенсивности комплексов осуществляют численно. В качестве оценки истинных значений искомым параметров принимают такой вектор θ^* (набор подлежащих определению β_i, α_{il}), который обращает в минимум выбранную критериальную функцию U :

$$\theta^* = \arg \min U(\theta). \quad (12)$$

Со времени отхода в середине 60-х годов XX века от графических методов параметрической идентификации и вплоть до конца 90-х годов у химиков не возникало вопроса о виде критерия U : во всех компьютерных программах его задавали как взвешенную сумму квадратов невязок [3, 4, 39, 41-43, 45]:

$$U(\theta) = \sum_{l=1}^{\Lambda} \sum_{k=1}^N \rho(\xi_{lk}), \quad (13)$$

где квадратичная функция потерь $\rho(\xi_{lk}) = \xi_{lk}^2$, взвешенная невязка $\xi_{lk} = \sqrt{w_{lk}} \cdot \Delta_{lk}$, w_{lk} – статистический вес измерения A_{lk} , $w_{lk} = 1/\sigma^2(A_{lk})$, $\sigma^2(A_{lk})$ – дисперсия A_{lk} , невязка $\Delta_{lk} = A^{\text{расчет}} - A^{\text{эксп}}$.

Основной для такого выбора служила не только простота расчетной схемы метода наименьших квадратов (МНК), но и соображения, основанные на методе максимума правдоподобия. Если 1) измерения независимы, 2) все переменные, контролирующие условия эксперимента, заданы точно и 3) погрешности величин A_{lk} распределены по закону Гаусса, то при использовании в функционале (13) квадратичной функции потерь оценки параметров θ являются асимптотически эффективными, несмещенными и состоятельными [46].

Отсутствие корреляции (в частности, автокорреляции) между наблюдениями обеспечивается на стадии проведения эксперимента. Абсолютно точно выполнить второе требование невозможно, однако можно добиться, чтобы погрешность факторов (переменных, задающих условия эксперимента) была намного меньше погрешности откликов A_{lk} . В худшем случае можно воспользоваться конфлюэнтным анализом, который, как правило, дает оценки θ , мало отличающиеся от

оценок МНК [47]. Наиболее опасно нарушение третьего условия. Согласно центральной предельной теореме, закон Гаусса является предельным, что рассматривалось как теоретическое основание для задания квадратичной функции потерь. Вопрос о том, каково на самом деле распределение экспериментальных погрешностей, исследован не очень подробно. Все же, как можно судить по данным, приведенным в работах [48, 49], в очень многих случаях гипотеза о нормальном распределении не подтверждается, а реальные массивы данных всегда на 5-10% засорены промахами. Если распределение погрешностей A_{lk} имеет хвосты, более длинные, чем нормальное распределение, МНК-оценки становятся смещенными, неэффективными и несостоятельными [50]. О длине хвостов удобно судить по величине выборочного эксцесса

$$\gamma_2 = \tilde{\mu}_4 / (\tilde{\mu}_2)^2 - 3, \quad (14)$$

где $\tilde{\mu}_2$ и $\tilde{\mu}_4$ – второй (дисперсия) и четвертый выборочные центральные моменты распределения невязок ξ_{lk} . Для нормального распределения $\gamma_2 = 0$, для распределений с длинными хвостами $\gamma_2 > 0$ (например, для распределения Лапласа $\gamma_2 = 3$) [51]. Если распределение погрешностей A_{lk} имеет хвосты более длинные, чем нормальное распределение, вместо МНК для параметрической идентификации следует использовать робастные методы расчета параметров, в которых меняется вид функции потерь в критериальном функционале (13). Робастными являются метод наименьших модулей (L_1 -оценки, МНМ) [52, 53], различные виды М-оценок [46, 53-56] и др. На наш взгляд, робастный метод оценивания желательно выбирать, также основываясь на принципе максимума правдоподобия.

Как показано в [53], не только закон распределения Гаусса, но и закон Лапласа может выступать в качестве предельного. При согласии распределения погрешностей A_{lk} с законом Лапласа метод максимума правдоподобия для получения оценок параметров с оптимальными свойствами рекомендует минимизировать критерий

$$U(\boldsymbol{\theta}) = \sum_{l=1}^{\Lambda} \sum_{k=1}^N |\xi_{lk}|, \quad (15)$$

т.е. рассчитывать параметры $\boldsymbol{\theta}$ методом наименьших модулей.

Известно, что условия формирования обоих предельных законов распределения для экспериментатора практически неразличимы [53, 57], а МНК и МНМ являются равно обоснованными. Найти компромисс позволяют М-оценки Хьюбера [49, 54], удачно сочетающие свойства МНК и МНМ. М-оценки Хьюбера обосновываются методом максимума правдоподобия, если распределение погрешностей подчиняется «модели грубых промахов», когда свойства A_{lk} измерены с погрешностями ε , плотность распределения которых задается формулой

$$\rho(\varepsilon) = [(100 - \delta) \cdot \tau(0, \sigma_{\text{Гauss}}) + \delta \cdot h(\varepsilon)] / 100, \quad (16)$$

где $\tau(0, \sigma_{\text{Гauss}})$ – плотность нормального распределения с нулевым средним и стандартным отклонением $\sigma_{\text{Гauss}}$, $h(\varepsilon)$ – плотность вероятности грубых промахов (плотность распределения с длинными хвостами, в качестве которого может выступать и плотность распределения Лапласа), δ – интенсивность промахов, %.

Функция потерь для М-оценок Хьюбера принимает вид

$$\rho(\xi_i) = \begin{cases} 0.5 \cdot \xi_{lk}^2 & \text{при } |\xi_{lk}| \leq c_{out} \\ c_{out} \cdot |\xi_{lk}| - 0.5 \cdot c_{out}^2 & \text{при } |\xi_{lk}| > c_{out} \end{cases}, \quad (17)$$

где константа c_{out} зависит от значения δ . Следует отметить, что ни априорно, ни апостериорно интенсивность грубых промахов определить невозможно. Анализ данных с помощью М-оценок Хьюбера обладает повышенной гибкостью: по сравнению с методом наименьших квадратов появляется новая варьируемая переменная – процент промахов δ . Значения $\delta = 0$ отвечает МНК, а $\delta = 100\%$ – МНМ. Обработка данных при задании разных δ в интервале $0 \leq \delta \leq 100\%$ позволяет найти решение, малочувствительное к отступлению от гипотезы о нормальном распределении экспериментальных погрешностей. Испытание различных гипотез о статистических свойствах измерений (варьирование δ) – одно из важнейших средств

теории анализа данных. В рассматриваемой задаче оно адаптирует процедуру параметрической идентификации к особенностям обрабатываемых данных.

Точечные оценки θ^* , соответствующие минимуму критериального функционала, имеют не слишком высокую ценность. Вместо точечных оценок правильнее определять доверительную область R_θ , в которой вероятность нахождения θ_{true} составляет

$$P(\theta_{\text{true}} \in R_\theta) = \gamma, \quad (18)$$

где γ – некоторая задаваемая вероятность (например, 0.95). Вместе с оценками θ^* вычисляется их ковариационная матрица $D(\theta^*)$, размерностью $z \times z$. Тогда можно построить z -мерный эллипсоид, ограничивающий доверительную область R_θ с приближенной вероятностью 100 γ %. Его центр размещен в точке θ^* , а границы задаются условием [58]

$$(\theta^* - \theta) \times D(\theta^*) \times (\theta^* - \theta)^T = z \cdot F(1 - \gamma, z, f), \quad (19)$$

где $F(1 - \gamma, z, f)$ – γ -процентная точка распределения Фишера с f степенями свободы.

После расчета параметров θ проверяют адекватность модели. Если взвешенные невязки ξ_{lk} – независимые случайные нормально распределенные величины с нулевым средним и единичной дисперсией, то величина $\chi_{\text{exp}}^2 = s_0^2 \cdot f$, где остаточная дисперсия

$$s_0^2 = \frac{1}{f} \sum_{l=1}^{\Lambda} \sum_{k=1}^N \xi_{lk}^2, \quad f = N \cdot \Lambda - z, \quad (20)$$

распределена как χ^2 с f степенями свободы и математическим ожиданием 1 [58]. Модель признают адекватной, если для заданного уровня значимости α (как правило, 0.05) справедливо неравенство

$$\chi_{\text{exp}}^2 = s_0^2 \cdot f < \chi^2(f, \alpha). \quad (21)$$

где $\chi^2(f, \alpha)$ – 100 α -процентная точка распределения χ^2 для f степеней свободы. При использовании М-оценок критерий (21) модифицируем, корректируя число степеней свободы [50]:

$$f = (N \cdot \Lambda - z) \cdot \{1 + 0.5\gamma_2(N \cdot \Lambda - z)N \cdot \Lambda\}^{-1}. \quad (22)$$

По величинам взвешенных невязок ξ_{lk} можно вычислить выборочные оценки асимметрии $\tilde{A} = \mu_3 / (\mu_2)^{3/2}$, эксцесса γ_2 , среднего значения $\bar{\Delta}$ и среднего значения модулей остатков $\tilde{\Delta}$. При нормальном распределении ξ_{lk} с нулевым средним и единичной дисперсией математические ожидания указанных величин составляют: $M(\tilde{A}) = M(\gamma_2) = M(\bar{\Delta}) = 0$, $M(\tilde{\Delta}) = \sqrt{2/\pi} = 0.80$ [59]. Если для заданного уровня значимости α выборочные значения \tilde{A} , γ_2 , $\bar{\Delta}$ и $\tilde{\Delta}$ не превышают соответствующих 100 α -процентных точек [59], можно сделать вывод об адекватности модели.

Дополнительную информацию предоставляют локальные критерии адекватности ξ_{lk} . Исследование зависимости ξ_{lk} от рассчитанных величин измеренного свойства позволяет а) выявить резко выпадающие измерения (промахи); б) вскрыть неполноту модели и в) обнаружить ошибки в модели погрешностей экспериментальных данных и назначении весов. Для неполной модели ξ_{lk} меняются систематически с изменением начального состава системы. Выброс в данных проявляет себя, как правило, в невязке ξ_{out} , многократно превосходящей ξ_{lk} в соседних точках плана эксперимента и, зачастую, отличающейся от них знаком. К сожалению, анализ остатков – не панацея: в регрессионном анализе хорошо известны примеры данных с выбросами, обнаружить которые при анализе остатков не удастся.

Для оценки адекватности модели можно привлечь также метод перекрестной оценки достоверности (кросс-валидацию) [60-62]. В простейшем варианте кросс-оценивания множество N экспериментальных точек разбивают на два подмножества. Одно содержит $(N-1)$ точек и является основой для расчета параметров θ , другое, содержащее одну точку с номером g , используется для оценки достоверности модели. Отбрасывая из массива N экспериментальных

точек одну (с номером g), находят значения неизвестных параметров и взвешенные невязки кросс-оценивания

$$d_{lg} = w_{lg}^{1/2} \cdot (\hat{A}_{lg} - A_{lg}), \quad (23)$$

где величины \hat{A}_{lg} рассчитывают для условий эксперимента в g -й экспериментальной точке при значениях параметров θ , найденных по измерениям в остальных точках. Повторив эту процедуру для всех N возможных разбиений множества экспериментальных точек, находят остаточную дисперсию кросс-оценивания

$$\tilde{s}_0^2 = \frac{1}{f} \sum_{l=1}^{\Lambda} \sum_{g=1}^N d_{lg}^2. \quad (24)$$

Величина \tilde{s}_0^2 измеряет предсказательную способность модели, тогда как обычная остаточная дисперсия s_0^2 оценивает ее подгоночную способность. Дисперсия кросс-валидации \tilde{s}_0^2 всегда превосходит s_0^2 , но если различие между ними невелико, то модель почти наверняка не является избыточной. А при величине $\tilde{s}_0^2 \approx 1$ модель, с большой вероятностью, является правильной. Кросс-валидация помогает выявить ситуацию, в которой значение подгоночного параметра θ_i («риск-параметра») определяется по результатам измерений всего в одной или нескольких экспериментальных точках. В такой «точке разбалансировки» («leverage point») [49] невязки d_{lg} должны значительно превышать невязки в остальных точках. Отбрасывание этой точки из массива данных влечет за собой изменение параметра θ_i , намного превышающее оценку погрешности θ_i , которая была найдена при обработке полного массива данных.

Проиллюстрируем применение робастных М-оценок при обработке данных по исследованию протолитических свойств *m*-аминобензойной кислоты. *m*-Аминобензойная кислота (НQ) – пример двухосновной кислоты с близкими ступенчатыми константами диссоциации. Равновесия в водных растворах НQ изучили спектрофотометрически, измеряя светопоглощение растворов при

$1.77 \leq \text{pH} \leq 6.16$ и шести длинах волн в ультрафиолетовом диапазоне поглощаемого света. Приведенные в работе [63] данные о светопоглощении растворов HQ с концентрацией $t(\text{HQ}) = 1.312 \cdot 10^{-4}$ моль·л⁻¹ при 25 °С являются тестовыми для проверки различных приемов обработки результатов спектрофотометрического эксперимента. В зависимости от использованных приемов, получали сильно отличающиеся оценки констант равновесия реакций



$2.8 \leq \lg K_{\text{H}} \leq 3.3$; $4.45 \leq \text{p}K_{\text{a}} \leq 4.8$ [63]. Сложность обработки результатов измерений связана как с близостью $\lg K_{\text{H}}$ и $\text{p}K_{\text{a}}$, так и с отличием в 4-40 раз максимальных величин светопоглощения для изученных длин волн.

При параметрической идентификации модели в качестве искомым параметров выбрали $\lg K_{\text{H}}$, $\text{p}K_{\text{a}}$ и молярные коэффициенты поглощения форм H_2Q^+ , HQ и Q^- при всех изученных длинах волн. Применяли модель погрешностей, в которой среднее квадратическое отклонение светопоглощений $\sigma(A) = 10^{-2}$, а статистические веса всех измерений $w_{ik} = 10^4$.

На первом этапе обработки, задав процент грубых выбросов $\delta = 0$, нашли МНК-оценки параметров и установили адекватность модели эксперименту (табл.). Вместе с тем, оказалось, что выборочный эксцесс распределения взвешенных невязок $\gamma_2 \neq 0$, то есть распределение остатков имеет более длинные хвосты, чем нормальное распределение.

Варьируя величину процента грубых выбросов δ , повторили процедуру параметрической идентификации (табл., рис. 1). Легко видеть, что, начиная с $\delta = 20\%$, оценки параметров и γ_2 стабилизируются; можно принять $\lg K_{\text{H}} = 2.93 \pm 0.01$; $\text{p}K_{\text{a}} = 4.425 \pm 0.005$. Важно отметить, что в указанные интервалы МНК-оценки не попадают. Локальные критерии адекватности (рис. 2) показывают наличие в плане эксперимента нескольких областей со взвешенными

невязками $|\xi_{kl}| = |w_{kl}^{1/2} \Delta_{kl}|$, превосходящими уровень погрешностей экспериментальных данных. Таким образом, массив обработанных данных действительно содержит промахи; применение робастных оценщиков в данном случае обязательно.

4. Определение сорбционной (ионообменной) емкости как отдельный этап параметрической идентификации.

Как указывалось выше, при моделировании равновесий в растворах начальные количества вещества или концентрации реагентов считают известными точно. Лишь в отдельных работах [64, 65] рассматривалась задача уточнения начальных концентраций реагентов при одновременном расчете констант равновесия. Такой подход, значительно усложняющий вычислительные процедуры и увеличивающий коррелированность параметров, не нашел широкого применения на практике.

Однако при моделировании равновесий сорбции и ионного обмена часто не удается избежать необходимости определять сорбционную (ионообменную) емкость по той же изотерме сорбции (ионного обмена), которая является основой для расчета констант равновесия. Это связано с тем, что удельная концентрация активных групп сорбента или ионообменника (t_0 , «эффективная емкость» [66, 67]) может значительно отличаться от теоретической емкости, найденной по данным элементного анализа. Эффективная емкость зависит от природы сорбата и растворителя, pH и ионной силы раствора и других факторов [66-69].

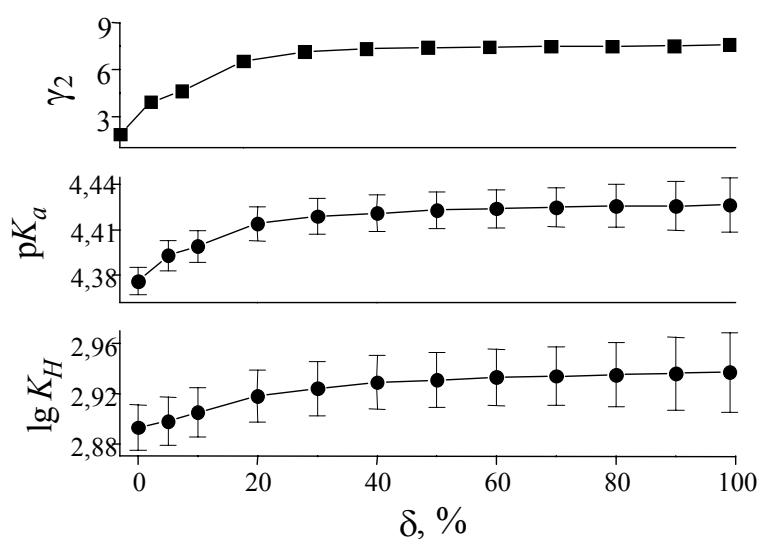


Рис. 1. Зависимость $\lg K_H$, pK_a и выборочного эксцесса γ_2 от значения параметра δ . Здесь и далее вертикальными отрезками отмечены 90%-ные доверительные интервалы Бонферони для рассчитанных параметров.

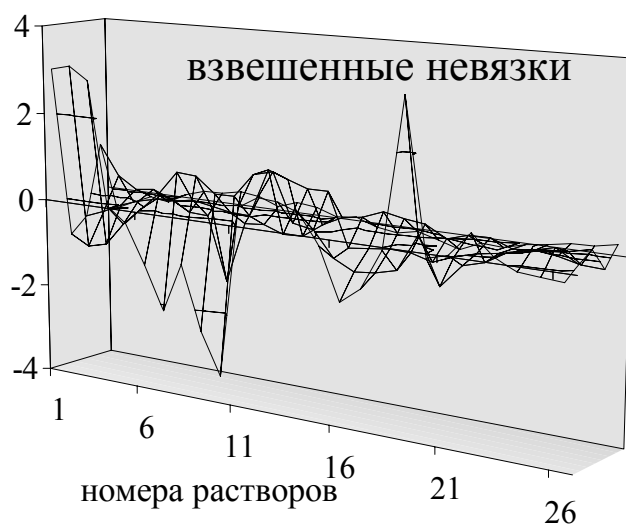


Рис. 2. Локальные критерии адекватности для модели равновесий в растворах м-аминобензойной кислоты. Расчет для $\delta = 20\%$.

Таблица. Результаты параметрической идентификации модели равновесий в растворе м-аминобензойной кислоты

δ , %	Параметр								
	$\lg K_H$	pK_a	$\bar{\Delta}$	$\tilde{\Delta}$	\tilde{A}	γ_2	f^*	χ^2_{exp}	$\chi^2(f; 0.05)$
0	2.893	4.376	$-1 \cdot 10^{-14}$	0.51	0.26	1.88	113	67.5	138.8
5	2.898	4.393	-0.03	0.49	0.51	3.94	43	27.9	59.3
10	2.905	4.399	-0.03	0.48	0.63	4.63	39	26.5	54.6

20	2.919	4.417	-0.03	0.48	0.39	6.54	30	23.3	43.8
30	2.924	4.419	-0.03	0.48	0.34	7.13	29	23.8	42.6
40	2.929	4.421	-0.02	0.47	0.30	7.33	28	23.6	41.3
50	2.931	4.423	-0.02	0.47	0.25	7.40	28	23.9	41.3
60	2.933	4.424	-0.02	0.47	0.21	7.45	28	24.1	41.3
70	2.934	4.425	-0.02	0.47	0.17	7.48	28	24.3	41.3
80	2.935	4.426	-0.01	0.47	0.14	7.49	27	23.5	40.1
90	2.936	4.426	-0.01	0.47	0.12	7.51	27	23.6	40.1
99	2.937	4.426	0.002	0.47	0.10	7.60	27	23.65	40.1

* f – здесь при $\delta > 0$ число степеней свободы исправлено по формуле (22).

Стремясь уйти от одновременного расчета t_Q и β_i , мы предлагаем подход, использующий формализм теории нечетких множеств [70]. На первом этапе находим интервал возможного варьирования t_Q (все t_Q из этого интервала допускают построение моделей, адекватных эксперименту). На втором этапе массив адекватных моделей сводится к одной окончательной модели с единственным набором параметров $|\lg \beta_i^{\text{final}}; t_Q^{\text{final}}\rangle$.

Перед началом расчетов любые значения t_Q , не превышающие теоретическую емкость T , рассматриваются как допустимые. Моделирование начинается с $t_Q = T$ и выполняется для меньших значений, равномерно распределенных в интервале $0 < t_Q < T$. Для каждого из испытываемых t_Q при варьируемой интенсивности грубых промахов δ вычисляются М-оценки p неизвестных $\lg \beta_i$ ($|\lg \beta_i\rangle$). На основании этих расчетов определяется интервал возможного варьирования t_Q (пример приведен на рис. 3), а всем испытанным t_Q в соответствие ставятся доверительные эллипсоиды $R_{|\lg \beta_i\rangle}$ и значения статистик χ_{exp}^2 . Переход от массива наборов параметров $\theta = |\lg \beta_i; t_Q\rangle$ к единственному набору $|\lg \beta_i^{\text{final}}; t_Q^{\text{final}}\rangle$ простым усреднением полученных оценок представляется крайне нежелательным, поскольку

эта процедура в рамках метода максимума правдоподобия базируется на гипотезе о нормальном распределении независимых случайных величин. Очевидно, что реальная ситуация далека от указанной гипотетической. Более того, поскольку наборы параметров $\Theta = \left\langle \lg \beta_i; t_Q \right\rangle$ получены обработкой одних и тех же экспериментальных данных, методы, основанные на концепции статистической вероятности, могут привлекаться лишь как эвристические средства.

Мы предлагаем обращаться с наборами параметров Θ как с многомерными нечеткими числами. В теории нечетких множеств вводятся функции принадлежности μ , описывающие субъективное сходство объектов. Функция принадлежности ($0 \leq \mu \leq 1$) задает степень принадлежности объекта данному множеству [70]. Единого способа задания функции принадлежности на наборе четких чисел (фаззификация), как и единого алгоритма выполнения обратной операции – перехода от нечеткого числа к четкому (дефаззификация) – не существует. Конструируя свой алгоритм, мы руководствовались следующими соображениями. Во-первых, для каждой модели со значениями статистики χ_{exp}^2 должны быть ассоциированы доверительные области R_Θ (не точечные оценки Θ^*). Во-вторых, приемлемыми следует считать все модели, адекватные эксперименту (и соответствующие этим моделям наборы параметров), но модели, обеспечивающие меньшие значения χ_{exp}^2 выглядят более привлекательно, и им следует приписать большие значение функций принадлежности.

Наконец, представляется естественным требовать, чтобы окончательные оценки параметров Θ^{final} мало зависели от способа задания функции принадлежности.

Для каждой адекватной модели функцию принадлежности задавали двумя способами:

$$\mu_v^I = G \cdot \exp(-\chi_{\text{exp},v}^2), \quad v = 1, 2, \dots, V, \quad (27)$$

где v – номер модели (набора искомых параметров), V – общее количество рассмотренных моделей, G – нормировочный множитель,

обеспечивающий $\mu = 1$ для модели с наименьшим значением χ_{exp}^2 ($\chi_{\text{exp min}}^2$), и

$$\mu_v^{\text{II}} = \exp\left(-0.5\left(\chi_v^2 - \chi_{\text{exp min}}^2\right)^2\right). \quad (28)$$

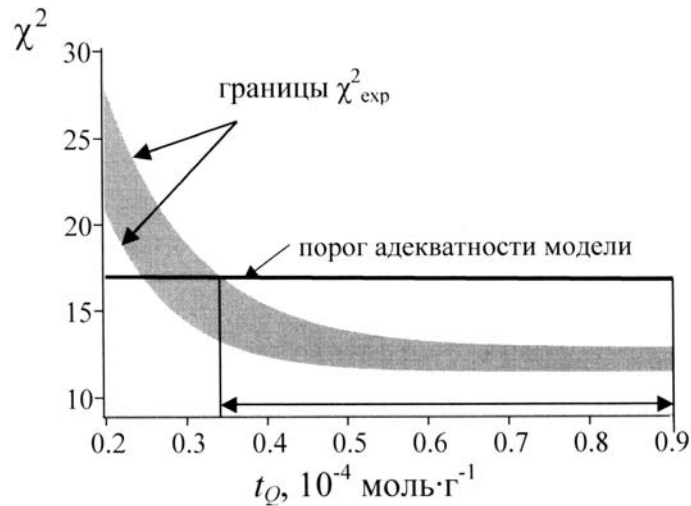


Рис. 3. Пример определения интервала возможного варьирования t_Q (обозначен стрелкой). Теоретическая емкость $T = 0.9 \cdot 10^{-4}$ моль·г $^{-1}$. Серая полоса здесь и далее отвечает значениям χ_{exp}^2 , найденным при различных гипотезах об интенсивности грубых промахов δ .

Доверительные области $R_{\theta, \nu}$ для всех моделей представляли наборами параметров $\theta_{\nu n}$, $n = 1, 2, \dots, N$, где N – число наборов параметров. Каждый набор $\theta_{\nu n}$ включал значение t_Q , отвечающее ν -й модели, и p значений $\lg \beta_i$. Последние были равномерно распределены внутри эллипсоида, границы которого устанавливает соотношение (19). Для компонентов множества $\theta_{\nu n}$ функции принадлежности задавали как

$$\mu_{vn, \lg \beta_i} = \begin{cases} \frac{1}{N} \mu_v, & \text{если } \beta_i \text{ определен,} \\ 0, & \text{если } \beta_i \text{ избыточный,} \end{cases} \quad (29)$$

$$\mu_{vn, t_Q} = \frac{1}{N} \mu_v.$$

Конечные оценки θ^{final} находили, используя метод центроидной дефаззификации [71]:

$$\theta_y^{\text{final}} = \frac{\sum_{v,n} \theta_{vn,y} \cdot \mu_{vn,y}}{\sum_{v,n} \mu_{vn,y}}, \quad y = 1, 2, \dots, Y, \quad Y = p + 1. \quad (30)$$

Предлагаемая процедура отдельного расчета $\lg \beta_i$ и t_Q не меняет, безусловно, возможную высокую коррелированность параметров, но обеспечивает численную устойчивость вычислений, а объединение в θ^{final} характеристик всех найденных адекватных моделей придает большую убедительность окончательным результатам параметрической идентификации.

Свойства предложенного алгоритма исследовали с помощью имитационного моделирования. Выполнили эксперимент по рН-метрическому титрованию при (20 ± 2) °С 20 мл раствора с начальными концентрациями глицина (НЛ) $4.455 \cdot 10^{-2}$ моль·л⁻¹ и HCl $5.60 \cdot 10^{-3}$ моль·л⁻¹ раствором щелочи с концентрацией 0.0900 моль·л⁻¹ (ионная сила раствора $5.60 \cdot 10^{-3}$ моль·л⁻¹). Погрешность измерения рН составляла 0.01. По результатам титрования искали оценки параметров $\lg \beta_1(\text{НЛ} + \text{H}^+ = \text{H}_2\text{L}^+)$ и $\lg \beta_2(\text{НЛ} = \text{H}^+ + \text{L}^-)$. Была построена адекватная модель ($\chi_{\text{exp}}^2 = 24.1 < \chi_{23, 0.05}^2 = 35.2$) с оценками параметров $\lg \beta_1 = 9.68$ (0.01), $\lg \beta_2 = 11.92$ (0.01) (в скобках указаны средние квадратические отклонения).

При имитационном моделировании считали, что начальная концентрация HCl ($c(\text{HCl})$) известна неточно и подлежит уточнению наряду с $\lg \beta_1$ и $\lg \beta_2$. Согласно описанному выше подходу искали интервал $c(\text{HCl})$, в котором удастся построить адекватные модели (рис. 4). Для каждого испытанного значения $c(\text{HCl})$ находили, варьируя

гипотезу об интенсивности грубых промахов δ , М-оценки $\lg\beta_1$ и $\lg\beta_2$ и их ковариационную матрицу. По результатам титрования находили М-оценки параметров $\lg\beta_1(\text{HL} + \text{H}^+ = \text{H}_2\text{L}^+)$ и $\lg\beta_2(\text{HL} = \text{H}^+ + \text{L}^-)$. От набора оценок параметров переходили к единственной финальной оценке $\lg\beta_1$ и $\lg\beta_2$ по методике, описанной выше.

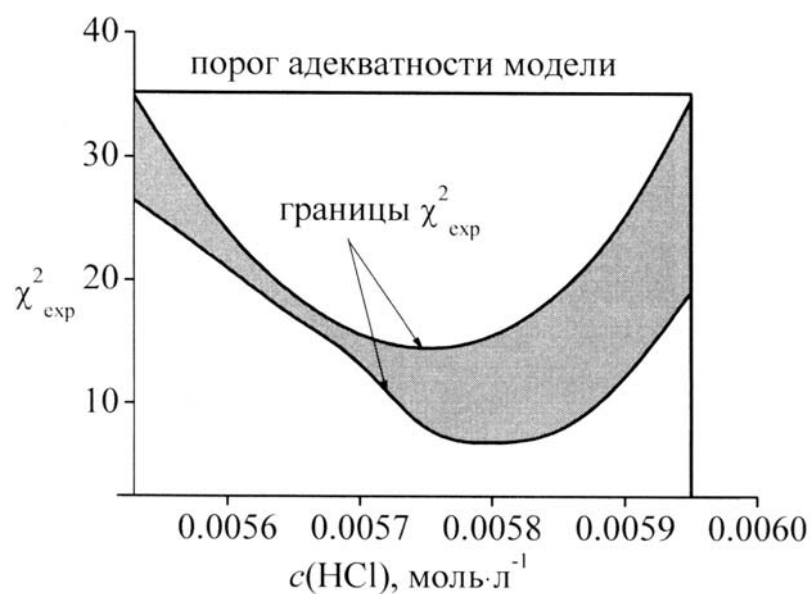


Рис. 4. Зависимость $\chi^2_{\text{эксп}}$ от $c(\text{HCl})$.

В результате были получены следующие оценки: $\lg\beta_1 = 9.71$ (0.01), $\lg\beta_2 = 11.99$ (0.01), $c(\text{HCl}) = 5.80 \cdot 10^{-3}$ моль·л⁻¹. Полученные оценки практически совпадают для обоих видов задаваемых функций принадлежности. На рис. 5 представлена зависимость $\lg\beta_1$ и $\lg\beta_2$ от $c(\text{HCl})$ и обозначены конечные оценки искомых параметров. Важно, что полученные оценки как к МНК-оценкам, так и к значениям, рекомендуемым в [72] для 20 °С и ионной силы раствора 0.001 моль·л⁻¹ ($\lg\beta_1 = 9.61$, $\lg\beta_2 = 11.96$).

Предложенный способ оценивания эффективной емкости сорбентов был недавно успешно применен при исследовании нового гибридного органо-минерального материала [25].

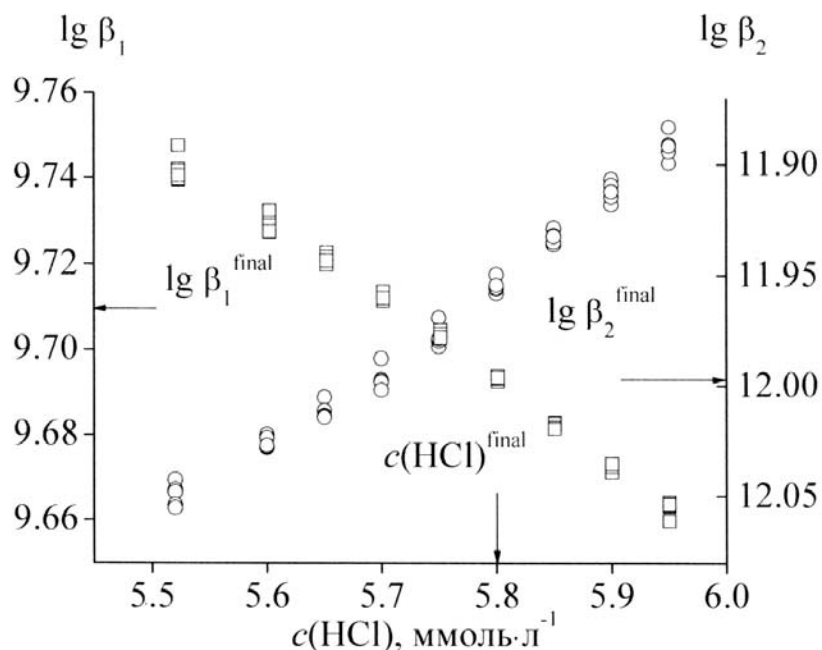


Рис. 5. Зависимость оценок $\lg \beta_1$ (\circ) и $\lg \beta_2$ (\square) от оценки $c(\text{HCl})$.

ЛИТЕРАТУРА

1. Аносов В.Я., Озерова Н.И., Фиалков Ю.Я. Основы физико-химического анализа. – М.: Наука, 1976. – 504 с.
2. Холин Ю.В. Количественный физико-химический анализ комплексообразования в растворах и на поверхности химически модифицированных кремнезёмов: содержательные модели, математические методы и их приложения. – Харьков: Фолио, 2000. – 288 с.
3. Бек М., Надьпал И. Исследование комплексообразования новейшими методами. – М.: Мир, 1989. – 413 с.
4. Meloun M., Havel J., Hogfeldt E. Computation of solution equilibria: a guide to methods in potentiometry, extraction and spectrophotometry. – Chichester: Ellis Horwood, 1988. – 298 p.
5. Бугаевский А.А., Мухина Т.П. Математика в химической термодинамике. – Новосибирск: Наука, 1980. – С. 20-35.
6. Рогинский С.З. Адсорбция и катализ на неоднородных поверхностях. – М.-Л.: Изд-во АН СССР, 1948. – 643 с.

7. Джейкок М., Парфит Д.А. Химия поверхности раздела фаз: Пер. с англ. – М.: Мир, 1984. – 269 с.
8. Jaroniec M., Madey R. Physical adsorption on heterogeneous solids. – Amsterdam, e.a.: Elsevier, 1988. – 354 p.
9. Koopal L.K. Ion adsorption on mineral oxide surface // Adsorption on new and modified inorganic sorbents. / Ed. by A.Dabrowsky, V.A.Tertykh. – Amsterdam e.a.: Elsevier, 1996. – P. 757-796.
10. Taylor H.S. Fourth report of the committee on contact catalysis // J. Phys. Chem. – 1926. – V. 30, No 2. – P. 145-171.
11. Мамлеев В.Ш., Золотарев П.П., Гладышев П.П. Неоднородность сорбентов: (феноменологические модели). – Алма-Ата: Наука, 1989. – 288 с.
12. Borcovec M., Jönsson B., Koper G. J. M. Ionization processes and proton binding in polyprotic systems: small molecules, proteins, interfaces and polyelectrolytes. In: E. Matijevic (Ed.). Colloid and Surface Science, V. 16. – New York: Plenum Press, 1999. – 287 p.
13. Sackett D.L., Saroff H.A. The multiple origins of cooperativity in binding to multi-site lattices // FEBS Letters. – 1996. – V. 397, No 1. – P. 1-6.
14. Perlmutter-Hayman B. Cooperative binding to macromolecules. A formal approach // Acc. Chem. Res. – 1986. – V. 19, No 3. – P. 90-96.
15. Braibanti A., Fiscicaro E., Compari C., Ghiozzi A., Rao R.S., Rao G.N. Calculation of site affinity, cooperativity between sites and self-association in polymer-ligand-proton complexes // Reactive&Functional Polymers. – 1998. – V. 36. – P. 245-249.
16. Katchalsky A., Spitnik P. Potentiometric titrations of poly-methacrylic acid // J. Polym. Sci. – 1947. – V. 2, No 4. – P. 432-446.
17. Marcus R.A. Titration of polyelectrolytes at higher ionic strengths // J. Phys. Chem. – 1954. – V. 58, No 8. – P. 621-623.
18. Katchalsky A., Mazur J., Spitnik P. Polybase properties of polyvinylamine // J. Polym. Sci. – 1957. – V. 23, No 104. – P. 513-532.
19. Моравец Г. Макромолекулы в растворе М., Мир, 1967. – 400 с.
20. Белинская Ф.А. Ионообменники и ионный обмен. В кн.: Физическая химия. Теоретическое и практическое руководство / Под

ред. акад. Б.П. Никольского, 2 изд. перераб. и доп. – Л.: Химия, 1987. – С. 666-705.

21. Strauss U.P., Barbieri B.W., Wong G. Analysis of ionization equilibriums of polyacids in terms of species population distributions. Examination of a "two-state" conformational transition // *J. Phys. Chem.* – 1979. – V. 83, No 22. – P. 2840-2843.

22. Strauss U.P. Representation of polyacids by subunits in the analysis of ionization equilibriums. Significance and thermodynamic applications // *Macromolecules.* – 1982. – V. 15, No 6. – P. 1567-1570.

23. Филиппов А.П. Модель реакций комплексообразования солей металлов с электронейтральными лигандами, привитыми к поверхности // *Теор. и эксп. химия.* – 1983. – Т. 19, No 4. – С. 463-470.

24. Kudryavtsev G.V., Milchenko D.V., Yagov V.V., Lopatkin A.A. Ion sorption on modified silica surface // *J. Colloid and Interf. Sci.* – 1990. – V. 140, No 1. – P. 114-122.

25. Lucho A. M. S., Panteleimonov A. V., Kholin Y. V., Gushikem Y. Simulation of adsorption equilibria on hybrid materials: binding of metal chlorides with 3-n-propylpyridiniumsilsesquioxane chloride ion exchanger // *J. Colloid and Interf. Science.* – 2007. – V. 310. – P. 47-56.

26. Simms H.S. Dissociation of polyvalent substances // *J. Am. Chem. Soc.* – 1926. – V. 48, No 5. – P. 1239-1250.

27. Бугаевский А.А. Некоторые свойства функции образования и родственных ей величин // *Докл. АН СССР.* – 1965. – Т. 161, No 1. – С. 140-142.

28. Dixon H.B.F., Clarke S.D., Smith G.A., Carne T.K. The origin of multiply sigmoid curves of pH-dependence. The partitioning of groups among titration pK values // *Biochem. J.* – 1991. – V. 278. – P. 279-284.

29. Perdue E. M., Reuter J. H., Parrish R. S. A statistical model of proton binding by humus // *Geochim. Cosmochim. Acta.* – 1984. – V. 48, No 6. – P. 1257-1263.

30. Thakur A.M., Munson P.J., Hunston D.L., Rodbard D. Characterization of ligand-binding systems by continuous affinity distribution of arbitrary shape // *Anal. Biochem.* – 1980. – V. 103, No 2. – P. 240-254.

31. Contescu C., Jaginello J., Schwarz J. A. Heterogeneity of proton binding sites at the oxide/solution interface // *Langmuir*. – 1993. – V. 9, No 7. – P. 1754-1765.

32. Sips R. On the structure of a catalyst surface // *J. Chem. Phys.* – 1948. – V. 16, No 5. – P. 490-495.

33. Годес О.М., Бондарева А.К., Хачкурузов Г.А. Теория адсорбции на неоднородных поверхностях. Определение закона распределения участков поверхности сорбента по теплотам сорбции из экспериментально наблюдаемых изотерм сорбции // *Проблемы кинетики и катализа. VII. Статистические явления в гетерогенных системах.* – М.-Л.: Изд-во АН СССР, 1949. – С. 263-296.

34. Джейкок М., Парфит Д.А. Химия поверхности раздела фаз: Пер. с англ. – М.: Мир, 1984. – 269 с.

35. Nederlof M. Analysis of binding heterogeneity. Ph. D. Thesis. – Wageningen Agricultural University, Wageningen, 1992.

36. Bushenkov V.A., Ramalho J.P.P., Smirnov G.V. Adsorption integral equation via complex approximation with constraints: the Langmuir kernel // *J. Comp. Chem.* – 2000. – V. 21, No 3. – P. 191-200.

37. Kholin Yu., Myerniy S. Energetic heterogeneity of sorbents: numerical calculation of affinity distributions // *Вісник Харківського національного університету імені В.Н. Каразіна.* – 2004. – №626. – Хімія. Вип. 11 (34). – С. 351-366.

38. Khoroshevskiy Y., Korneev S., Myerniy S., Kholin Y.V., Pavan F.A., Schifino J., Costa T.M.H., Benvenuto E.V. A mathematical simulation of H⁺ ion chemisorption by anilinepropylsilica xerogels // *J. Colloid and Interf. Sci.* – 2005. – V. 284, No 2. – P. 424-431.

39. Computational methods for the determination of formation constants. / Ed. by D.J. Legget. – New York: Plenum Press, 1985. – 478 p.

40. Sillen L.G., Warnqvist B. High-speed computers as a supplement to graphical methods. VII. Model selection and rejection with LETAGROP. Elimination of species with negative or «insignificant» equilibrium constants // *Arkiv Kemi.* – 1968. – V. 31, No 4. – P. 341-351.

41. Zekany L., Nagypal I. PSEQUAD. A comprehensive program for the evaluation of potentiometric and/or spectrophotometric equilibrium data using analytical derivatives // *Computational methods for the determination*

of formation constants. / Ed. by D.J. Legget. – New York: Plenum Press, 1985. – P. 291-353.

42. Gans P., Sabatini A., Vacca A. SUPERQUAD: An improved general program for computation of formation constants from potentiometric data // J. Chem. Soc. (Dalton). – 1985 – V. 20, No 2. – P. 141-162.

43. Gans P., Sabatini A., Vacca A. Investigation of equilibria in solution. Determination of equilibrium constants with the HYPERQUAD suite of programs // Talanta. – 1996. – V. 43. – P. 1739-1753.

44. Bugaevsky A.A., Kholin Yu.V. Computer-aided determination of the composition and stability of complex compounds in solutions with complicated equilibria // Anal. Chim. Acta. – 1991. – V. 249. – P. 353-365.

45. Meloun M., Havel J. Computation of solution equilibria. 1. Spectrophotometry. 1984. 184 p. 2. Potentiometry. 1985. 145 p. – Brno: Univerzita J.E. Purkyně-Prirodovědecká Faculta.

46. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей. Под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1985. – 487 с.

47. Бугаевский А.А., Рудная Л.Е. Конфлюэнтный анализ при расчете констант равновесия по измеренной концентрации одной из частиц // Журн. неорг. химии. – 1976. – Т. 21, No 10. – С. 2827-2831.

48. Орлов А.И. Часто ли распределение результатов наблюдений является нормальным? // Зав. лаб. – 1991. – Т. 57, No 7. – С. 64-66.

49. Хьюбер П. Робастность в статистике. – М.: Мир, 1984. – 303 с.

50. Вучков И., Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ. – М.: Финансы и статистика, 1987. – 239 с.

51. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям. – М.: Статистика, 1980. – 95 с.

52. Мудров В.И., Кушко В.Л. Метод наименьших модулей. – М.: Знание, 1971. – 64 с.

53. Мудров В.И., Кушко В.Л. Методы обработки измерений. Квазиправдоподобные оценки. – М.: Радио и связь, 1983. – 304 с.

54. Huber P. Robust statistical procedures. CBMS-NSF regional conference series in applied mathematics. – Philadelphia: Soc. for Industrial and Appl. Mathematics, 1996. – 68 p.

55. Edlund O., Ekblom H., Madsen K. Algorithms for non-linear M-estimation // Computational Statistics. – 1997. – V. 12. – P. 373-383.
56. Ekblom H., Madsen K. Algorithms for nonlinear Huber estimation // BIT. – 1989. – V. 29, No 1. – P. 60-76.
57. Корнилов А.Н. Метод наименьших квадратов и метод наименьших модулей // Математические методы в химической термодинамике. – Новосибирск: Наука, 1982. – С. 164-167.
58. Уилкс С. Математическая статистика. – М.: Наука, 1967. – 632 с.
59. Большев Х.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. – 413с.
60. Stone M. Cross-validators choice and assessment of statistical predictions // J. of the Royal Stat. Soc. – 1974. – V. 36, No 2. – P. 111-147.
61. Osten D.W. Selection of optimal regression models via cross-validation // J. of Chemometrics. – 1988. – V. 2, No 1. – P. 39-48.
62. Brünger A.T. Free R-value: cross-validation in crystallography // Methods in Enzymology. V. 277. – N.Y.: Academic Press, 1997. – P. 366-396.
63. Garcia B., Domingo P., Leal J.M. Overlapping equilibria: applications to *m*-aminobenzoic acid // Collect. Czechoslovak Chem. Commun. – 1987. – V. 52, No 5. – P. 1087-1096.
64. Havel J., Meloun M. Multiparametric curve fitting. IX // Talanta. – 1986. – V. 33, No 5. – P. 435-441.
65. Brauner P., Sillén L.G., Whiteker R. High-speed computers as a supplement to graphical methods. 9. Adjustment for systematic experimental errors and other "group parameters" in LETAGROP: Applications to potentiometric titrations // Arkiv Kemi. – 1969. – V. 31. – P. 365-376.
66. Helfferich F. Ion Exchange. – N.Y.: Dover, 1995. – 624 p.
67. Inglezakis V.J. The concept of "capacity" in zeolite ion-exchange systems // J. Colloid Interf. Sci. – 2005. – V. 281, No 1. – P. 68-79.
68. Bilba D., Bejan D., Tofan L. Chelating sorbents in inorganic chemical analysis // Croat. Chem. Acta. – 1998. – V. 71, No 1. – P. 155-178.
69. Vagheti J.C.P., Brasil J.L., Costa T.M., Benvenuti E.V. Influence of the synthesis conditions on the characteristics and metal adsorption

properties of the 3-(1,4-phenylenediamine)propylsilica xerogel // *Eclet. Quim.* – 2005. – V. 30, No 4. – P. 43-49.

70. Otto M. *Fuzzy theory explained* // *Chemometrics and Intelligent Lab. Systems.* – 1988. – V. 4. – P. 101-120.

71. Ross T.J. *Fuzzy Logic with Engineering Applications.* – N.Y.: McGraw-Hill, 1995. – 268 p.

72. Martell A.E, Smith R.M. *Critical Stability Constants*, V. 1-6. – N.Y: Plenum Press, 1974-1993.