

ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ В.Н. КАРАЗИНА

ПАНТЕЛЕЙМОНОВ АНТОН ВИТАЛИЕВИЧ

УДК 543.08+543.061+543.054

**НОВЫЕ ХЕМОМЕТРИЧЕСКИЕ СПОСОБЫ
ОБРАБОТКИ ДАННЫХ АНАЛИТИЧЕСКОГО ЭКСПЕРИМЕНТА**

02.00.02 – аналитическая химия

Автореферат
диссертации на соискание ученой степени
кандидата химических наук

Харьков – 2009

Диссертация является рукописью.

Работа выполнена в Харьковском национальном университете имени В.Н. Каразина Министерства образования и науки Украины

Научный руководитель:

доктор химических наук, профессор
ХОЛИН ЮРИЙ ВАЛЕНТИНОВИЧ,
Харьковский национальный университет
имени В.Н. Каразина,
проректор по учебно-методической работе

Официальные оппоненты:

член-корреспондент НАН Украины,
доктор химических наук, профессор
ЗАЙЦЕВ ВЛАДИМИР НИКОЛАЕВИЧ,
Киевский национальный университет имени Тараса Шевченко,
заведующий кафедрой аналитической химии

доктор химических наук, профессор
БОЛОТОВ ВАЛЕРИЙ ВАСИЛЬЕВИЧ,
Национальный фармацевтический университет
Министерства охраны здоровья Украины, г. Харьков,
заведующий кафедрой аналитической химии

Защита состоится «10» апреля 2009 г. в 16⁰⁰ час. на заседании специализированного ученого совета Д 64.051.14 Харьковского национального университета имени В.Н. Каразина (Украина, 61077, г. Харьков, пл. Свободы, 4, ауд. 7-80).

С диссертацией можно ознакомиться в Центральной научной библиотеке Харьковского национального университета имени В.Н. Каразина (Украина, 61077, г. Харьков, пл. Свободы, 4).

Автореферат разослан «5» марта 2009 г.

Ученый секретарь
специализированного ученого совета,
кандидат химических наук, доцент

В.Г. Панченко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Возрастающие требования к качеству анализа выдвигают новые требования к объему, точности и достоверности информации, извлекаемой из результатов измерений и применяемой для принятия управленческих решений. Лейтмотивом применения разнообразных аналитических методов становится перевод массивов результатов измерений в аналитические выводы с применением теории анализа данных и информационных технологий. Этим определяется возрастающее значение хемометрии как важного компонента теоретических основ современной аналитической химии и обуславливается необходимость расширения арсенала хемометрических способов обработки данных. Применению теории нечетких множеств обязаны быстрым прогрессом многие научные направления, в то время как для решения задач аналитической химии она до этого времени привлекалась лишь эпизодично. Между тем, представления о вероятности событий, лежащие в основе понятийного аппарата теории, способны существенно облегчить обработку массивов экспериментальных данных в условиях отсутствия информации о статистических свойствах результатов измерений.

Разработанные в работе хемометрические процедуры предназначены для решения задач качественного анализа (обнаружения и идентификации аналитов), а также для определения химико-аналитических характеристик твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов по данным количественного физико-химического анализа (КФХА). Область качественного анализа была выбрана в связи с ее недостаточным метрологическим обеспечением и потребностью в надежных расчетных способах для решения задач обнаружения и идентификации. Целесообразность использования хемометрических способов на основе теории нечетких множеств при определении химико-аналитических характеристик сорбентов обусловлена тем, что известные методы зачастую не обеспечивали извлечения достоверной информации из экспериментальных данных, если одновременно определению подлежат сорбционная емкость и константы сорбционных равновесий. Именно в таких условиях приходится оценивать химико-аналитические характеристики твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов, перспективных для сорбционного концентрирования компонентов растворов.

Таким образом, актуальность работы обусловлена потребностью в разработке новых хемометрических процедур анализа химико-аналитических данных, в частности для решения задач обнаружения и идентификации в качественном анализе и определения характеристик твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

Связь работы с научными программами, планами, темами. Работа выполнена в соответствии с научно-исследовательскими работами кафедры химического материаловедения ХНУ имени В.Н. Каразина «Керування

хімічними рівновагами в гетерогенних та мікрогетерогенних середовищах, перспективних для тестових та гібридних методів аналізу» (№ государственной регистрации 0103U004212); «Інформаційні технології та комп'ютерні засоби прогнозування властивостей хімічних речовин» (№ государственной регистрации 0105U002844), согласно координационным планам Научного Совета НАН Украины по проблеме «Неорганическая химия» и Дополнительному соглашению к Договору о сотрудничестве между Харьковским национальным университетом имени В.Н. Каразина и Университетом штата Сан-Пауло в Кампинасе (Бразилия).

Цель работы: разработка хемометрических процедур на основе использования теории нечетких множеств для решения задач обнаружения и идентификации в качественном анализе и определения химико-аналитических характеристик твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

Для достижения поставленных целей необходимо было решить следующие задачи: обосновать набор метрологических характеристик методик обнаружения с бинарным откликом; сформировать ограниченный набор функций для представления кривых эффективности; обосновать расчетные процедуры определения параметров кривых эффективности и нахождения на этой основе метрологических характеристик методик обнаружения; разработать алгоритм идентификации аналитов по многомерным массивам их физико-химических признаков, измеренных в условиях, отличных от условий измерения свойств эталонов; разработать метод одновременного определения химико-аналитических характеристик твердофазных аналитических реагентов (эффективной сорбционной емкости, констант сорбционных равновесий) на основе органо-кремнеземных гибридных материалов.

Объект исследования: обнаружение и идентификация аналитов по данным одно- и многооткликковых экспериментов; химико-аналитические характеристики твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

Предмет исследования: метрологические характеристики методик обнаружения с бинарным откликом; вычислительные процедуры расчета параметров кривых эффективности; критерии идентификации аналитов по данным многооткликкового эксперимента при вариабельности условий измерения; методы одновременного определения нескольких химико-аналитических характеристик твердофазных аналитических реагентов по данным количественного физико-химического анализа.

Методы исследования: теория нечетких множеств, робастное оценивание, регрессионный анализ, физико-химический анализ.

Научная новизна полученных результатов:

1. Обосновано, что в качестве метрологических характеристик методик обнаружения с бинарным откликом следует использовать:

- интервал ненадежности;

- предел обнаружения – концентрацию аналита, при превышении которой вероятность ошибки I рода (ложного заключения об отсутствии аналита) меньше 1%;
 - недостоверность – вероятность ошибки I рода.
2. На основе анализа данных для 27 методик обнаружения показано, что в качестве функций, описывающих кривые эффективности, целесообразно использовать функции логистического и экспоненциального распределений.
 3. Показано, что оценки параметров кривых эффективности, найденные нелинейным МНК, мало зависят от способа назначения статистических весов.
 4. Установлено, что при расчетах допустимо использовать приближение, согласно которому частоты обнаружения аналита в интервале ненадежности являются равнозначными случайными величинами, т.е. использовать невзвешенный МНК.
 5. На основе применения теории нечетких множеств разработан робастный алгоритм оценки параметров кривых эффективности. Показано, что робастные оценки практически совпадают с эвристическими МНК-оценками. Тем самым обосновано использование МНК в качестве вычислительного средства для решения задач обнаружения аналитов.
 6. На основе теории нечетких множеств предложен критерий идентификации аналитов по многооткликовым данным, малочувствительный к варьированию гипотез о статистических свойствах результатов измерений.
 7. Разработан алгоритм, предоставляющий объективные оценки химико-аналитических характеристик сорбентов на основе органо-кремнеземных гибридных материалов в условиях нарушения предпосылок МНК и отсутствия информации о распределении экспериментальных погрешностей.
 8. Определены химико-аналитические характеристики двух новых твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

Практическое значение полученных результатов. Формирование перечня обоснованных метрологических характеристик методик обнаружения с бинарным откликом развивает метрологию скрининга, позволяя на единой основе разрабатывать и аттестовать новые методики анализа. Выработанные рекомендации по обработке зависимостей частот обнаружения аналитов от их концентрации для нахождения параметров кривых эффективности существенно упрощают нахождение метрологических характеристик методик обнаружения. Уточнение метрологических характеристик 27 методик обнаружения обеспечивает повышение качества анализа. Разработанный в работе алгоритм позволяет идентифицировать аналиты по данным многооткликового эксперимента при различии условий измерения их свойств и свойств эталонов. Предложенный метод оценивания эффективной сорбционной емкости и констант равновесий процессов сорбции отличается большей объективностью предоставляемой информации, что обеспечивает надежное прогнозирование условий концентрирования и разделения аналитов. Определение химико-

аналитических характеристик новых твердофазных аналитических реагентов с закрепленными группами 3-*n*-пропил(4-метилпиридина) и 3-*n*-пропилпиридина позволяет расчетным путем находить условия концентрирования и разделения солей металлов.

Личный вклад автора. Анализ литературных данных, реализация, испытание и применение вычислительных алгоритмов и количественный физико-химический анализ протолитических равновесий глицина выполнены автором самостоятельно. Постановка цели и задач исследования, формулировка выводов, обобщение результатов работы выполнялись совместно с научным руководителем проф. Ю.В. Холиным. Рекомендации по метрологии бинарного тестирования обсуждались с научным руководителем, доц. Е.А. Решетняк и доц. Н.А. Никитиной (Харьковский национальный университет имени В.Н. Каразина). Экспериментальные данные о зависимости частот обнаружения аналитов от их концентрации (12 систем) предоставлены доц. Е.А. Решетняк и доц. Н.А. Никитиной, изотермы сорбции солей металлов гибридными материалами с закрепленными группами 3-*n*-пропил(4-метилпиридина) и 3-*n*-пропилпиридина – проф. Й. Гушикемом и д-ром Э. Магоссо (Институт химии Университета штата Сан-Пауло в Кампинасе).

Апробация результатов диссертации. Основные результаты работы были доложены на международных, украинских и российских научных конференциях: IV и V Всеукраинских конференциях студентов и аспирантов «Сучасні проблеми хімії» (Киев, 2003, 2004); Сессии Научного Совета НАН Украины по проблеме «Аналитическая химия» (Харьков, 2007); Международной конференции «Analytical Chemistry and Chemical Analysis» (Киев, 2005); Международной конференции «Modern Physical Chemistry for Advanced Materials» (Харьков, 2007); Международной конференции «4th Black Sea Basin Conference on Analytical Chemistry» (Солнчев Бряг, Болгария, 2007); XV Международной конференции студентов, аспирантов и молодых ученых по фундаментальным наукам «Ломоносов – 2008» (Москва, 2008).

Публикации. По материалам диссертации опубликованы: 1 монография, 5 статей и 7 тезисов докладов на научных конференциях.

Структура и объем диссертации. Диссертация состоит из введения, трех разделов, выводов и списка использованных литературных источников. Работа изложена на 183 страницах машинописного текста, содержит 58 таблиц, 55 рисунков и библиографию из 228 ссылок.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении охарактеризованы основные задачи и методы хемометрии, ее место в теоретическом фундаменте аналитической химии, обоснованы актуальность темы работы, сформулированы задачи исследования, указаны научная новизна и практическая ценность полученных результатов.

В разделе 1 «Метрологические характеристики методик качественного анализа с бинарным откликом» предложено сократить

набор метрологических характеристик методик обнаружения с бинарным откликом (в литературе описаны более 20 характеристик), ограничен набор функций, используемых для построения кривых эффективности и обоснованы вычислительные процедуры построения кривых эффективности.

Особенное внимание уделено оцениванию метрологических характеристик методик качественного анализа с бинарным откликом на основе исследования кривых эффективности, представляющих зависимость вероятности обнаружения аналита (P) от его концентрации (c). В таких методиках делают вывод о присутствии либо отсутствии аналита, сравнивая аналитический сигнал холостой пробы и анализируемой пробы. Ошибочный отрицательный результат, как правило, намного важнее ошибочного положительного, т.к. после вывода об открытии аналита проводят его количественное определение, а пробы, не показавшие присутствия аналита, количественному анализу не подвергают.

Ключевой метрологической характеристикой методики с бинарным откликом предлагается считать интервал ненадежности Δc – интервал концентраций аналита, в котором возможен как положительный так и отрицательный ответ о его присутствии в анализируемом образце.

Поскольку при скрининге массовых проб наиболее важно избежать ошибочного вывода об отсутствии аналита, границами интервала ненадежности предлагается считать такие концентрации $c_{5\%}$ и $c_{99\%}$, при превышении которых вероятности обнаружения аналита превышают 5% и 99%, соответственно. В качестве предела обнаружения целесообразно выбрать $c_{99\%}$. Вместо такой характеристики результатов качественного анализа, как неопределенность, используется достоверность обнаружения (R). Ее рекомендовано связывать с вероятностью одной, наиболее значимой ошибки (пропуск аналита) и определять как $R_{\text{обнаружения}} = 100 \cdot P(c)$, %.

На практике кривую эффективности подбирают как функцию, аппроксимирующую эмпирическую зависимость частот обнаружения $P^{\text{эмп}}$ от c . Для получения такой зависимости выбирают несколько (M) концентраций c_i внутри интервала ненадежности, для каждой из них проводят N_i испытаний, фиксируют число положительных результатов (n_i) и рассчитывают частоты обнаружения аналита $P_i^{\text{эмп}} = \frac{n_i}{N_i}$, $i = 1, 2, \dots, M$.

Для аппроксимации зависимости $P^{\text{эмп}}$ от c можно, в принципе, использовать любую неубывающую функцию, ограниченную 0 и 1. Такими свойствами обладают функции распределения случайных величин. В работе из ряда функций распределений (экстремального значения первого рода; экспоненциального; Лапласа; логистического; логнормального; нормального; Пуассона) на основе обработки зависимостей $P^{\text{эмп}}$ от c для 27 методик обнаружения для представления кривых эффективности рекомендовано использовать лишь две – логистического и экспоненциального распределения.

Проведен анализ вычислительных методов нахождения параметров кривых эффективности. Нелинейным МНК вектор подгоночных параметров $\{\theta\}$ находят минимизацией функции $\chi_{\text{эксп}}^2 = \sum_{i=1}^M \xi_i^2$, где взвешенные невязки

$\xi_i = w_i^{1/2} \cdot (P_i^{\text{эмп}} - \hat{P}_i)$, статистические веса $w_i = \frac{1}{s^2(P_i^{\text{эмп}})}$. Аппроксимацию

зависимости $P^{\text{эмп}}$ от c считали адекватной, если выполнялось неравенство $\chi_{\text{эксп}}^2 < \chi_{f, \alpha}^2$, где $\chi_{f, \alpha}^2$ – 100 α -процентная точка распределения χ^2 для $f = M - z$ степеней свободы, z – число подгоночных параметров. Анализировали также локальные критерии адекватности (взвешенные невязки) ξ_i . Погрешности определения параметров оценивали на основе расчета их ковариационных матриц.

Сравнивали различные модели погрешностей P_i для назначения статистических весов. Наиболее теоретически обоснованной является модель, учитывающая, что частоты P_i , как и количества «успехов» n_i в N_i испытаниях – случайные биномиально распределенные величины. Тогда

дисперсии частот обнаружения $s^2(P_i) = \frac{P_i(1-P_i)}{N_i}$, что приводит к формуле

для весов

$$w_i = \frac{N_i}{P_i^{\text{эмп}}(1-P_i^{\text{эмп}})}. \quad (1)$$

В другой модели считали, что все $P_i^{\text{эмп}}$ имеют одинаковые стандартные отклонения $s(P_i^{\text{эмп}}) = 0.02$. Тогда статистические веса, назначаемые как

$$w_i = \frac{1}{s(P_i^{\text{эмп}})^2}, \quad (2)$$

также одинаковы. Наконец, в третьей модели допускали, что одинаковы относительные стандартные отклонения $s_r(P_i^{\text{эмп}}) = 0.05$, а веса назначали как

$$w_i = \frac{1}{[P_i^{\text{эмп}} \cdot s_r(P_i^{\text{эмп}})]^2}. \quad (3)$$

При назначении весов по формулам (1) и (2) параметры аппроксимирующих функций и границы интервалов ненадежности (табл. 1) близки, а значения статистик $\chi_{\text{эксп}}^2$ – величины одного порядка. Это означает, что простейшая возможная модель, признающая $P_i^{\text{эмп}}$ равноточными, может использоваться наряду с моделью, учитывающей биномиальное распределение частот обнаружения, при расчете параметров кривых эффективности. Поскольку в первом случае веса одинаковы, расчеты существенно упрощаются, а список программных продуктов, которые можно использовать для вычислений, расширяется. При назначении весов по формуле (3) значения подгоночных параметров для ряда тест-систем изменяются на 30-200% (особенно часто в случае функции экспоненциального распределения), значительна и разница между оценками

пределов обнаружения. Очевидно, модель, допускающая одинаковые относительные погрешности $P_i^{\text{эмп}}$, далека от реальности.

Таблица 1. Сопоставление интервалов ненадежности при разных способах назначения статистических весов

No методики обнаружения	$\Delta c, \text{мг/л}$		
	Формула для w_i		
	(1)	(2)	(3)
1	23.9 – 213	23.9 – 220	23.7 – 257
2	$3.5 \cdot 10^{-3} - 4.6 \cdot 10^{-3}$	$3.6 \cdot 10^{-3} - 4.6 \cdot 10^{-3}$	$3.5 \cdot 10^{-3} - 4.6 \cdot 10^{-3}$
3	1.75 – 3.77	1.70 – 3.75	1.83 – 3.57
4	$2.2 \cdot 10^{-3} - 7.1 \cdot 10^{-2}$	$2.1 \cdot 10^{-3} - 7.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-3} - 7.9 \cdot 10^{-2}$
5	$8.6 \cdot 10^{-2} - 0.88$	$5.1 \cdot 10^{-2} - 1.10$	$2.6 \cdot 10^{-2} - 1.48$
6	3.21 – 8.78	2.94 – 8.99	3.38 – 7.62
7	$2.1 \cdot 10^{-2} - 9.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-2} - 0.10$	$2.2 \cdot 10^{-2} - 9.7 \cdot 10^{-2}$
8	$6.1 \cdot 10^{-2} - 0.26$	$5.8 \cdot 10^{-2} - 0.27$	$6.2 \cdot 10^{-2} - 0.26$
9	0.10 – 0.53	0.11 – 0.53	$8.1 \cdot 10^{-2} - 0.62$
10	0.12 – 0.27	0.12 – 0.27	0.12 – 0.27
11	$7.0 \cdot 10^{-2} - 0.27$	$6.2 \cdot 10^{-2} - 0.29$	$6.5 \cdot 10^{-2} - 0.28$
12	0.22 – 0.50	0.22 – 0.51	0.24 – 0.46
13	0.47 – 0.65	0.47 – 0.65	0.47 – 0.66
14	$1.7 \cdot 10^{-2} - 3.4 \cdot 10^{-2}$	$1.5 \cdot 10^{-2} - 3.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-2} - 3.6 \cdot 10^{-2}$
15	$1.1 \cdot 10^{-2} - 8.6 \cdot 10^{-2}$	$8.6 \cdot 10^{-3} - 9.2 \cdot 10^{-2}$	$1.1 \cdot 10^{-2} - 8.3 \cdot 10^{-2}$
16	$3.4 \cdot 10^{-3} - 4.3 \cdot 10^{-2}$	$3.1 \cdot 10^{-3} - 4.4 \cdot 10^{-2}$	$3.4 \cdot 10^{-3} - 4.3 \cdot 10^{-2}$
17	$2.2 \cdot 10^{-3} - 2.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3} - 3.1 \cdot 10^{-2}$	$1.6 \cdot 10^{-3} - 3.7 \cdot 10^{-2}$
18	$4.2 \cdot 10^{-3} - 2.7 \cdot 10^{-2}$	$3.0 \cdot 10^{-3} - 2.9 \cdot 10^{-2}$	$4.3 \cdot 10^{-3} - 2.4 \cdot 10^{-2}$
19	$6.0 \cdot 10^{-2} - 0.32$	$4.7 \cdot 10^{-2} - 0.31$	$7.3 \cdot 10^{-2} - 0.27$
20	$2.1 \cdot 10^{-2} - 6.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-2} - 7.0 \cdot 10^{-2}$	$2.0 \cdot 10^{-2} - 7.2 \cdot 10^{-2}$
21	$2.4 \cdot 10^{-2} - 0.48$	$2.4 \cdot 10^{-2} - 0.48$	$2.0 \cdot 10^{-2} - 0.84$
22	4.31 – 5.14	4.32 – 5.15	4.27 – 5.25
23	11.5 – 31.1	11.8 – 30.7	12.2 – 30.3
24	13.7 – 25.9	13.6 – 26.2	13.6 – 33.1
25	51.0 – 75.0	51.0 – 75.0	51.5 – 74.1
26	0.44 – 1.19	0.45 – 1.17	0.45 – 1.17
27	$3.3 \cdot 10^{-2} - 1.44$	$3.3 \cdot 10^{-2} - 1.44$	$2.6 \cdot 10^{-2} - 1.62$

Была проверена устойчивость МНК-оценок подгоночных параметров. В качестве робастного метода оценивания использовался метод, базирующийся на применении теории нечетких множеств. Проводили процедуру фаззификации (придания нечеткости) экспериментальных данных на основании таких положений: а) нечеткость определения концентрации аналита $\Delta c_i = \delta \cdot c_i$, где δ – относительная погрешность c_i ; б) нечеткость

частоты обнаружения $\Delta P_i = s(P_i)$. Каждую точку кривой эффективности интерпретировали как нечеткое измерение и рассчитывали суммарную функцию принадлежности (4), что является альтернативой обычным критериям адекватности аппроксимации:

$$\mu_{\text{sum}} = \text{card}(F) = \frac{1}{M} \sum_{i=1}^M \mu_i, \quad (4)$$

где $\text{card}(F)$ – мощность множества F функций принадлежности μ_i . Подгоночные параметры аппроксимирующей функции находили как оценки, отвечающие максимуму суммарной функции принадлежности $|\theta\rangle = \arg \max \left(\frac{1}{M} \sum_{i=1}^M \mu_i \right) = \arg \max(\text{card}(F))$. Робастные оценки и оценки МНК (с

назначением весов по формулам (1) и (2)) близки, а максимальные отличия в значениях границ интервалов ненадежности не превышают 10%.

Этот факт указывает на устойчивость МНК-оценок и возможность в практике обработки зависимостей $P^{\text{эмп}}$ от s избежать применения трудоемких робастных процедур.

С применением описанных подходов впервые определены либо уточнены метрологические характеристики 27 методик обнаружения с бинарным откликом.

Раздел 2 «Идентификация соединений – применение хемометрических подходов» посвящен разработке нестатистического метода решения задачи идентификации аналитов на основе сравнения данных о спектральных (длины волн, волновые числа и т.д.) или хроматографических (время удерживания) свойствах эталона и аналита. Известные методы оценки достоверности идентификации основаны на модельных допущениях о форме спектральных либо хроматографических пиков, законе распределения погрешностей измерений, «ширине окна» (величины, характеризующей допустимое расхождение положений пиков эталона и аналита). В данной работе с положениями пиков обращались как с нечеткими числами. Пусть наблюдалось M пиков. Вводили функции принадлежности пиков μ_i , $i = 1, 2, \dots, M$. Испробовали два типа функций μ_i : а) с линейными ветвями

$$\mu_i^L(x) = \frac{x - p_i + b_i/2}{b_i/2} \text{ при } x \leq p_i; \frac{x - p_i - b_i/2}{-b_i/2} \text{ при } x > p_i \quad (5)$$

и б) колоколообразную

$$\mu_i^G(x) = \exp \left(-0.5 \left[\frac{x - p_i}{b_i/4.7} \right]^2 \right), \quad (6)$$

где x – любая точка в интервале $[p_i - b_i/2; p_i + b_i/2]$, p_i – положение i -го пика, b_i – допустимый размах i -го пика (аналог ширины окна). Находили значения пересечений ($\mu_{a,e}$) функции принадлежности аналита (μ_a) и эталона (μ_e) для всех M пар пиков и рассчитывали суммарную функцию принадлежности

$\mu_{\text{sum}} = \frac{1}{M} \sum_{i=1}^M (\mu_{a,e})_i$. Значение μ_{sum} , при превышении которого вывод об

идентичности соединений становится приемлемым, следует устанавливать на основе обобщения результатов обработки массивов модельных и реальных данных. В этой работе принимали критическое значение $\mu_{\text{sum}} > 0.6$. Ставя в соответствие наш подход ранее разработанным методам, можно показать, что при форме пиков, которые описываются функциями плотности распределений Гаусса или Симпсона, такое значение $\mu_{\text{sum}} > 0.6$ отвечает вероятности ошибки I рода (ошибочный вывод об отсутствии тождественности аналита и эталона) $\llsim 5\%$.

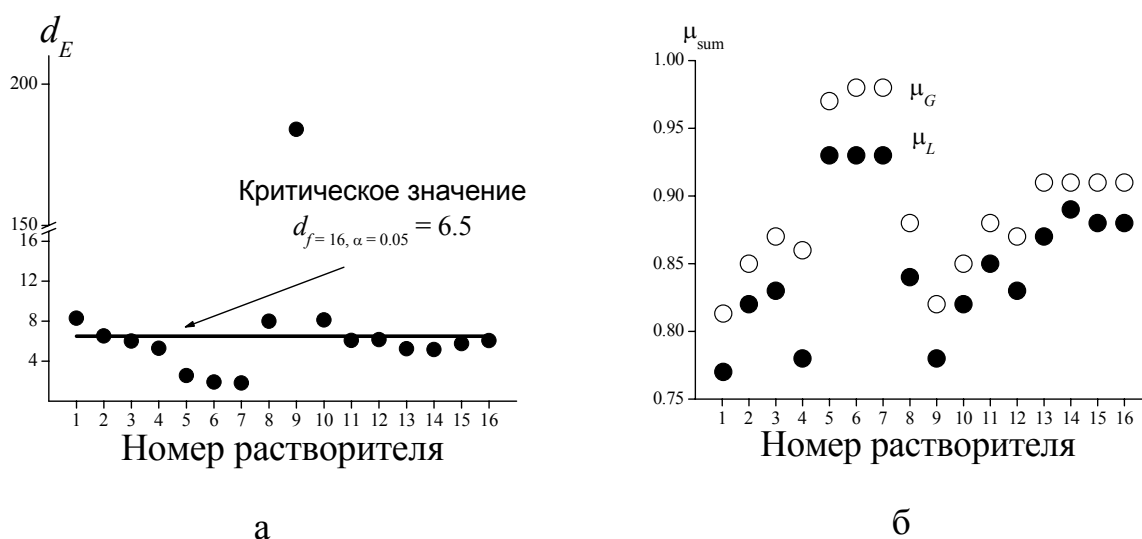


Рис. 1. Идентификация трет-бутил-2-[[2,2,2-трихлорэтанимидоил)окси]метил]-акрилата по данным ИК-спектроскопии (а): на основе определения Евклидовых расстояний между аналитами и эталоном; (б): предложенным алгоритмом.

В нашем алгоритме, в отличие от предложенных ранее, для идентификации необходимо лишь задать тип функции принадлежности и оценки допустимых размахов b_i ; информация о распределении экспериментальных погрешностей не используется. При отличии условий измерения свойств аналита и эталона особенное значение приобретает устойчивость идентификации. Установлено, что варьирование типа функций принадлежности и изменение оценок b_i в несколько раз на результатах отражаются слабо. Разработанный алгоритм идентификации апробирован на массивах хроматографических и спектральных (ИК-, УФ-спектроскопия) данных. Например, при идентификации трет-бутил-2-[[2,2,2-трихлорэтанимидоил)окси]метил]-акрилата по данным ИК-спектров, измеренных в индивидуальном состоянии (эталон) и в 16 растворителях (аналиты), найдено, что если степень совпадения спектров эталона и аналитов оценивать по Евклидовым расстояниям (d_E) между положениями полос поглощения, то вывод о тождественности аналита с эталоном справедлив в 11 случаях из 16 при вероятности ошибки I рода 5% (ширина

всех окон 6 см^{-1} , рис. 1а). При использовании разработанного алгоритма вывод о тождественности спектров сделан во всех 16 случаях (для всех полос поглощения использовали единое значение $b_i = 6 \text{ см}^{-1}$, рис. 1б).

В разделе 3 «Химико-аналитические характеристики твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов» обсуждены принципы и расчетные подходы КФХА, приведен алгоритм одновременной оценки эффективной сорбционной емкости и констант сорбционных равновесий при описании свойств твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

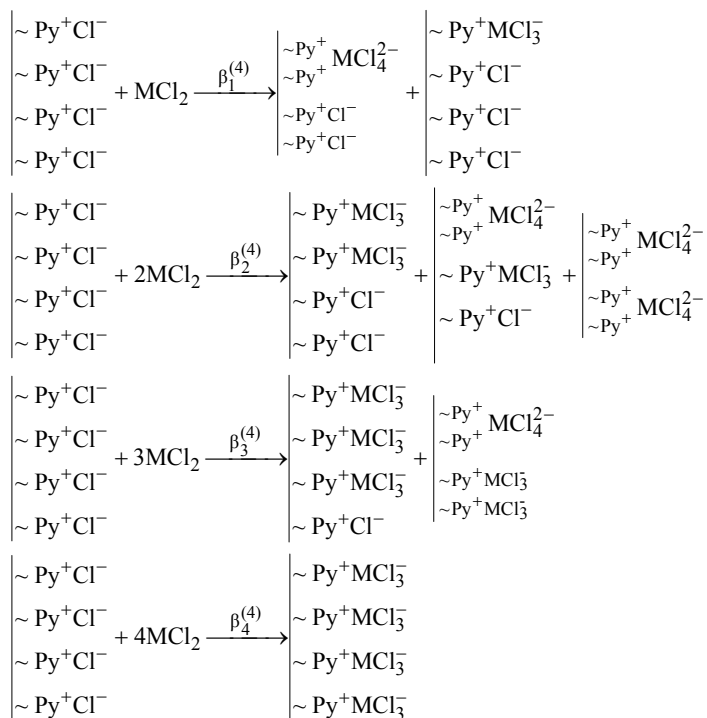


Рис. 2. Модель тетрадентатного связывания для описания адсорбции на поверхности материала SiPy^+Cl^- .

Одна из проблем моделирования сорбционных равновесий заключается в том, что эффективная емкость сорбента (t_Q , моль·г⁻¹) может значительно отличаться от теоретической (T). Если устойчивость сорбционных комплексов незначительна (константа сорбционного равновесия $\beta < 10^3 - 10^4 \text{ л} \cdot \text{моль}^{-1}$), даже в случае идеальной сорбции одновременное определение β и t_Q по изотерме адсорбции представляет собой сложную задачу. А в случае неидеальной адсорбции простые приемы определения эффективной емкости (например, по участкам насыщения на изотермах

адсорбции) вообще утрачивают применимость. Кроме того, при расчетах следует учитывать, что МНК-оценки параметров моделей могут не иметь оптимальных статистических свойств, если нарушается гипотеза о нормальном распределении погрешностей результатов измерений. В работе предложен подход, позволяющий на основе объединения принципов робастного оценивания и методов теории нечетких множеств получать более объективную информацию о химико-аналитических характеристиках сорбентов, нежели та, которую обеспечивают традиционные методы. Методической основой подхода является модель полидентатного связывания, описывающая неидеальную сорбцию как проявление эффектов кооперативности. В модели реакцию поверхность материала представляют как совокупность центров $\overline{Q_Z}$, каждый из которых содержит Z

активных групп Q , удельная концентрация центров $\overline{Q_Z}$ равна $\frac{t_Q}{Z}$. Связывание частиц сорбата центрами $\overline{Q_Z}$ рассматривают как ступенчатый процесс, а сорбционные равновесия описывают Z константами равновесия $\beta_i^{(Z)}$, где i – номер ступени (см. рис. 2).

В КФХА модели равновесных систем находят, анализируя зависимости равновесных свойств исследуемой системы A (например, адсорбции сорбата) от начального состава системы. Построение модели полидентатного связывания начинают с малого значения Z (например, $Z = 2$). Соответствующие оценки $\beta_i^{(Z)}$ рассчитывают оптимизацией подходящего функционала и оценивают адекватность модели. Если модель неадекватна, значение Z увеличивают и снова вычисляют $\beta_i^{(Z)}$. Процесс продолжают до построения адекватной модели. В предложенном алгоритме для фиксированных значений t_Q и Z параметры $\beta_i^{(Z)}$ находили методом M -оценивания по П. Хьюберу. M -оценки обосновываются методом максимума правдоподобия, если плотность распределения экспериментальных погрешностей $A(\varepsilon)$ $\rho(\varepsilon) = \frac{[(100 - \delta) \cdot \tau(0, \sigma_{\text{Gauss}}) + \delta \cdot h(\varepsilon)]}{100}$, где $\tau(0, \sigma_{\text{Gauss}})$ – плотность нормального распределения с нулевым средним и стандартным отклонением σ_{Gauss} , $h(\varepsilon)$ – плотность вероятности «грубых промахов», имеющих более длинные хвосты, чем распределение Гаусса, δ – интенсивность грубых промахов, %. Минимизируют функционал

$$U(|\beta \rangle) = \sum_{k=1}^N \rho(\xi_k), \quad (7)$$

где k – номер точки изотермы сорбции, N – их число, $|\beta \rangle$ – вектор подгоночных параметров модели, функция потерь

$$\rho(\xi_k) = \begin{cases} 0.5 \cdot \xi_k^2 & \text{при } |\xi_k| \leq c_{\text{out}} \\ c_{\text{out}} \cdot |\xi_k| - 0.5 \cdot c_{\text{out}}^2 & \text{при } |\xi_k| > c_{\text{out}} \end{cases}, \text{ взвешенная невязка } \xi_k = \sqrt{w_k} \cdot \Delta_{lk}$$

c_{out} – параметр, значение которого рассчитывают исходя из заданной интенсивности промахов, w_k – статистический вес k -го измерения, невязка $\Delta = A^{\text{модель}} - A^{\text{эксп}}$, $w_k = 1 / \sigma^2(A_k)$, $\sigma^2(A_k)$ – дисперсия A_k . Выполняют серию расчетов, варьируя значение δ от 0 (отвечает МНК) до 100% (отвечает оценкам метода наименьших модулей). При $\delta > 0$ оценки робастны.

Целесообразно вместе с точечными оценками $|\beta \rangle^*$, отвечающими минимуму функционала (7), определять доверительную область R_β , в которой вероятность нахождения $|\beta \rangle_{\text{true}}$ равна $P(|\beta \rangle_{\text{true}} \in R_\beta) = \gamma$, где γ – вероятность (например, 0.95). Z -мерный эллипсоид, ограничивающий доверительную область R_β с приближенной вероятностью γ , строили с использованием ковариационной матрицы параметров $D(|\beta \rangle^*)$.

Каждому вектору $|\beta\rangle$ сопоставляли меру адекватности аппроксимации экспериментальных данных – значение статистики $\chi_{\text{эксп}}^2 = \sum_{k=1}^N \xi_k^2$. Модель признавали адекватной, если для заданного уровня значимости α (как правило, 0.05) значение статистики $\chi_{\text{эксп}}^2$ были меньше соответствующего критического значения χ^2 для числа степеней свободы $f = N - z$. При M -оценивании f корректировали:

$$f = (N - z) \cdot \{1 + 0.5 \cdot \gamma_2 \cdot (N - z)/N\}^{-1}, \quad (8)$$

где γ_2 – выборочный коэффициент эксцесса распределения невязок ξ_k .

Оценивая по изотермам сорбции химико-аналитические характеристики сорбентов, сначала находили интервал возможного варьирования t_Q (все t_Q из этого интервала допускают построение адекватных моделей). На втором этапе набор адекватных моделей с параметрами $|\theta\rangle = |\lg \beta_i; t_Q\rangle$ сводили к одной конечной модели с набором параметров $|\theta^{\text{final}}\rangle = |\lg \beta_i^{\text{final}}; t_Q^{\text{final}}\rangle$.

Моделирование выполняли для значений t_Q , равномерно распределенных в интервале $0 < t_Q \leq T$ (T – теоретическая емкость). Для каждого из этих t_Q при разных значениях интенсивности грубых промахов δ находили M -оценки Z неизвестных $\lg \beta_i$. На основе этих расчетов определяли интервал возможного варьирования t_Q , а всем испробованным t_Q в соответствие ставили доверительные эллипсоиды $R_{|\lg \beta_i\rangle}$ и значения

статистик $\chi_{\text{эксп}}^2$. Принимали как допустимые все модели, адекватные эксперименту, а при формировании конечного набора параметров обеспечивали больший вклад моделей с меньшими значениями $\chi_{\text{эксп}}^2$. Выполнение этого требования обеспечивали, действуя с наборами параметров $|\theta\rangle$ как с многомерными нечеткими числами. Для каждой адекватной модели функции принадлежности задавали двумя способами:

$$\mu_v^I = g \cdot \exp(-\chi_{\text{эксп},v}^2), \quad v = 1, 2, \dots, V, \quad (9)$$

где v – номер модели (набора искоемых параметров), V – общее число рассмотренных моделей, g – нормировочный множитель, обеспечивающий $\mu = 1$ для модели с наименьшим значением $\chi_{\text{эксп}}^2$ ($\chi_{\text{эксп min}}^2$), и

$$\mu_v^{II} = \exp\left(-0.5(\chi_v^2 - \chi_{\text{эксп min}}^2)^2\right). \quad (10)$$

Доверительные области $R_{\theta,v}$ для всех моделей представляли наборами параметров θ_{vm} , $m = 1, 2, \dots, M$, где M – число наборов параметров. Каждый набор θ_{vm} включал значение t_Q , отвечающее v -й модели, и Z значений $\lg \beta_i$. Последние были равномерны распределены в эллипсоиде R_β . Для компонентов множества θ_{vm} функции принадлежности задавали как

$$\mu_{vm, \lg \beta_i} = \begin{cases} \frac{1}{M} \mu_v, & \text{если } \beta_i \text{ определен,} \\ 0, & \text{если } \beta_i \text{ избыточный,} \end{cases} \text{ Конечные оценки } |\theta^{\text{final}}\rangle \text{ находили, используя}$$

$$\mu_{vm, t_Q} = \frac{1}{M} \mu_v.$$

метод центроидной дефазификации:

$$\theta_y^{\text{final}} = \frac{\sum_{v,m} \theta_{vm,y} \cdot \mu_{vm,y}}{\sum_{v,m} \mu_{vm,y}}, y = 1, 2, \dots, Y, Y = Z + 1. \quad (11)$$

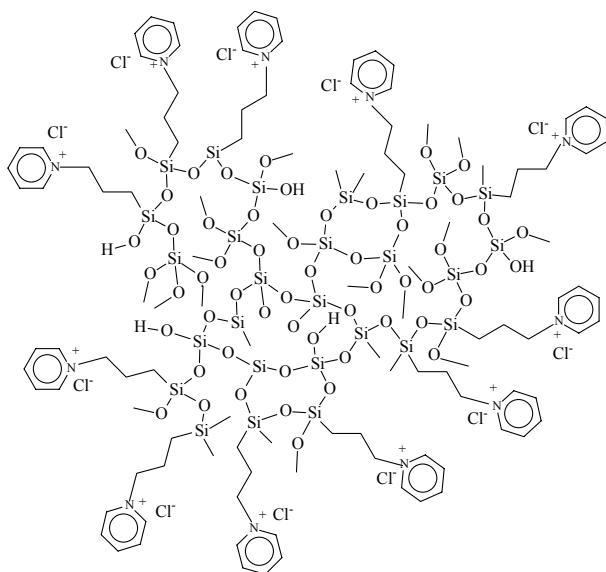


Рис. 3. Схематическое строение 3-н-пропилпиридиний силсеквиоксан хлоридного полимера.

Эффективность предложенного алгоритма проверяли при моделировании протолитических равновесий в растворе глицина по данным рН-метрического титрования. Разработанный подход использовали для оценивания химико-аналитических характеристик новых органико-кремнеземных материалов – силикагелей, покрытых 3-н-пропил(4-метилпиридиний) силсеквиоксан хлоридным полимером (SiPic⁺Cl⁻) и 3-н-пропилпиридиний силсеквиоксан хлоридным полимером (SiPy⁺Cl⁻, рис. 3).

Равновесия адсорбции хлоридов металлов MCl₂ (M = Zn, Cd, Hg) из водных растворов на поверхности SiPic⁺Cl⁻ и адсорбции FeCl₃ и MCl₂ (M = Cu, Zn, Cd, Hg) из этанольных растворов на поверхности SiPy⁺Cl⁻ описывали с помощью модели полидентатного связывания. Для описания равновесий на поверхности материала SiPic⁺Cl⁻ достаточной оказалась модель бидентатного связывания (табл. 2). В случае SiPy⁺Cl⁻ эта модель адекватно воспроизводила лишь адсорбцию CuCl₂. Адсорбция остальных хлоридов охарактеризована с помощью модели тетрадентатного связывания (рис. 2, табл. 3). Найденные модели адекватно воспроизводят экспериментальные изотермы адсорбции (примеры сравнения экспериментальных и рассчитанных изотерм адсорбции приведены на рис. 4). На адсорбцию ZnCl₂ и CdCl₂ на SiPic⁺Cl⁻ влияют эффекты отрицательной кооперативности (ступенчатые константы сорбционных равновесий $K_2^{(2)} < K_1^{(2)}$). Эти же эффекты имеют место при адсорбции CuCl₂ на SiPy⁺Cl⁻; при адсорбции остальных хлоридов металлов наблюдаются эффекты положительной кооперативности (присоединение одной или двух частиц хлоридов металлов к адсорбционному центру Q_4 облегчает дальнейшее присоединение частиц сорбата). Исследованные твердофазные реагенты перспективны для извлечения из растворов хлоридов металлов. С использованием построенных моделей рассчитывали степени извлечения

$$r = \frac{t(MCl_x) - [MCl_x]}{t(MCl_x)}, \quad (12)$$

где $t(MCl_x)$ и $[MCl_x]$ – начальная и равновесная концентрации MCl_x в растворе, соответственно. Расчеты выполнили для начальных концентраций хлоридов

10^{-5} – 10^{-3} моль/л в растворах объемом 50 мл и переменных массах навесок материалов (m , г) и выяснили, что можно достичь степеней извлечения

хлоридов металлов более 99% (пример результатов расчетов приведен на рис. 5).

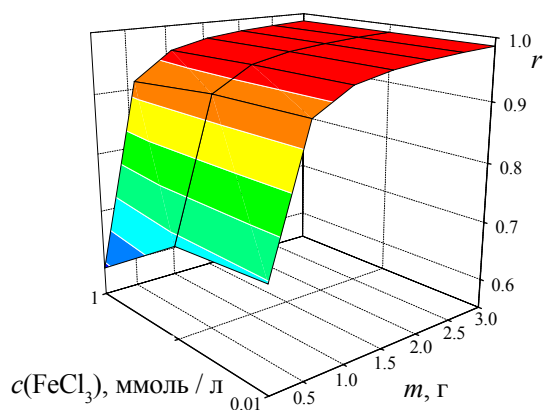
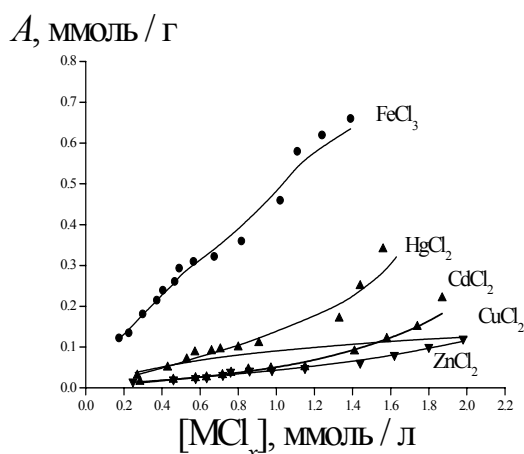


Рис. 4. Экспериментальные (точки) и рассчитанные (линии) изотермы FeCl_3 материалом SiPy^+Cl^- из адсорбции хлоридов металлов на этанольного раствора. поверхности SiPy^+Cl^- .

Таблица 2. Характеристики материала $\text{SiPic}^+\text{Cl}^-$

		Формула для функции принадлежности					
		(9)	(10)	(9)	(10)	(9)	(10)
Параметры модели	Соль	ZnCl_2		CdCl_2		HgCl_2	
	$\lg \beta_1^{(2)}$	11.8 (0.2)	11.8 (0.2)	11.2 (0.1)	11.2 (0.1)	—	—
	$\lg \beta_2^{(2)}$	18.8 (0.4)	18.8 (0.4)	19.6 (0.2)	19.6 (0.2)	6.9 (0.3)	6.9 (0.3)
	t_Q , ммоль / г	0.48	0.48	0.24	0.24	0.96	0.96

Таблица 3. Характеристики материала SiPy^+Cl^-

		Формула для функции принадлежности									
		(9)	(10)	(9)	(10)	(9)	(10)	(9)	(10)	(9)	(10)
Параметры модели	Соль	ZnCl_2		CdCl_2		HgCl_2		CuCl_2		FeCl_3	
	$\lg \beta_1^{(4)}$	2.65 (0.03)	2.63 (0.02)	2.42 (0.04)	2.41 (0.04)	2.8 (0.1)	2.8 (0.1)	3.14** (0.03)	3.14** (0.03)	3.7 (0.1)	3.7 (0.1)
	$\lg \beta_2^{(4)}$	*	*	*	*	5.4 (0.3)	5.4 (0.3)	5.20** (0.05)	5.21** (0.07)	6.7 (0.1)	6.7 (0.1)
	$\lg \beta_4^{(4)}$	10.23 (0.05)	10.25 (0.05)	10.40 (0.06)	10.41 (0.06)	11.0 (0.2)	11.0 (0.2)	—	—	12.8 (0.1)	12.8 (0.1)
t_Q , ммоль / г	0.53	0.53	0.74	0.74	0.77	0.77	0.27	0.27	0.83	0.83	

Избыточными параметрами для всех моделей тетрадентатного связывания являются $\lg \beta_3^{(4)}$, для моделей, обозначенных *, избыточными параметрами являются, также, $\lg \beta_2^{(4)}$.

** Параметры модели бидентатного связывания.

ВЫВОДЫ

В работе решена актуальная задача развития хемометрических методов обработки данных, что позволяет повысить достоверность результатов качественного химического анализа, и обеспечивает надежное оценивание химико-аналитических характеристик твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов.

1. Показано, что важнейшими метрологическими характеристиками методик обнаружения с бинарным откликом следует считать находимые по кривым эффективности предел обнаружения аналита и интервал ненадежности; для представления кривых эффективности достаточно ограничиться лишь двумя функциями логистического и экспоненциального распределений.
2. На основе теории нечетких множеств создан робастный алгоритм оценивания параметров кривых эффективности по эмпирическим зависимостям частот обнаружения от концентрации аналитов.
3. С использованием упомянутого алгоритма для обработки 27 массивов данных о зависимостях частот обнаружения аналитов от их концентраций установлено, что робастные оценки параметров кривых эффективности и метрологические характеристики методик обнаружения практически не отличаются от оценок, полученных на основе применения нелинейного взвешенного МНК.
4. Способы назначения статистических весов по моделям погрешностей, предполагающим равноточность частот обнаружения или же их биномиальное распределение, приводят к близким оценкам подгоночных параметров. Таким образом, обоснована возможность применения невзвешенного нелинейного МНК для оценивания метрологических характеристик методик обнаружения с бинарным откликом.
5. Разработан алгоритм идентификации соединений по массивам их спектральных или хроматографических характеристик, основанный на использовании аппарата теории нечетких множеств. Алгоритм минимально использует априорные допущения о свойствах экспериментальных данных и обеспечивает достаточную устойчивость выводов о тождественности аналита и эталона к варьированию условий эксперимента.
6. Использование новой расчетной процедуры при решении задачи обработки данных качественного физико-химического анализа позволяет преодолеть трудности, связанные с одновременным определением химико-аналитических характеристик органо-кремнеземных гибридных материалов (эффективной сорбционной емкости и констант сорбционных равновесий).
7. Впервые определены химико-аналитические характеристики двух новых твердофазных аналитических реагентов на основе органо-кремнеземных гибридных материалов и определены условия их использования для извлечения ряда солей металлов из водных и этанольных растворов.

СПИСОК ПУБЛИКАЦИЙ АВТОРА ПО ТЕМЕ РАБОТЫ

1. Метрологические характеристики методик обнаружения с бинарным откликом : монография / [Ю.В. Холин, Н.А. Никитина, **А.В. Пантелеймонов** и др.] – Харьков: Тимченко, 2008. – 128 с. Рекомендована к печати Ученым Советом Харьковского национального университета имени В.Н. Каразина, протокол № 7 от 27.06.2008 р.
Автором обоснованы сокращенный список метрологических характеристик методик обнаружения с бинарным откликом и выбор функций логистического и экспоненциального распределений для представления кривых эффективности, сопоставлены способы назначения статистических весов при вычислении параметров кривых эффективности, испытан метод кривых Пирсона, разработан робастный алгоритм оценивания параметров на основе теории нечетких множеств, оценены метрологические характеристики 27 методик обнаружения аналитов; автор участвовал в обсуждении всех результатов и написании монографии.
2. Островская В. М. Тест-метод определения суммы металлов реактивной индикаторной бумагой и его метрологические характеристики / В. М. Островская, Е. А. Решетняк, Н. А. Никитина, **А. В. Пантелеймонов**, Ю. В. Холин // Журн. аналит. химии. – 2004. – Т. 59, № 10. – С. 995-1002.
Автор реализовал алгоритм оценки метрологических характеристик систем обнаружения с бинарным откликом, участвовал в обсуждении результатов и подготовке публикации.
3. Пантелеймонов А. В. К разработке нечетких критериев идентификации в качественном химическом анализе / **А. В. Пантелеймонов**, Ю. В. Холин // Вісник Харківського національного університету. – 2005. – № 669. Хімія. Вип. 13(36). – С. 128-133.
4. Magosso H. A. Synthesis, characterization and metal adsorption properties of the new ion exchanger polymer 3-n-propyl(4-methylpyridinium) silsesquioxane chloride / H. A. Magosso, **A. V. Panteleimonov**, Y. V. Kholin, Y. Gushikem // J. Colloid and Interf. Sci. – 2006. – V. 303. – P. 18-24.
Автор провел моделирование сорбционных свойств материала, определил его химико-аналитические характеристики, принял участие в обсуждении результатов и подготовке публикации.
5. Lucho A. M. S. Simulation of adsorption equilibria on hybrid materials: Binding of metal chlorides with 3-n-propylpyridiniumsilsesquioxane Chloride Ion Exchanger / A. M. S. Lucho, **A. V. Panteleimonov**, Y. V. Kholin, Y. Gushikem // J. Colloid and Interf. Sci. – 2007. – V. 310. – P. 47-56.
Автор реализовал алгоритм одновременной оценки химико-аналитических характеристик материала, провел моделирование сорбционных равновесий, принял участие в обсуждении результатов и подготовке публикации.
6. Холин Ю. В. Некоторые методические и вычислительные проблемы количественного физико-химического анализа / Ю. В. Холин,

С. А. Мерный, Д. С. Коняев, **А. В. Пантелеймонов** // В кн.: Научное наследие Н. А. Измайлова и актуальные проблемы физической химии. – Харків: ХНУ імені В. Н. Каразіна, 2007. – С. 525-553.

Автор испытал разработанный алгоритм одновременного оценивания констант равновесия и начальных (общих) концентраций реагентов в задачах физико-химического анализа для исследования протолитических равновесий в водном растворе глицина, принял участие в обсуждении результатов и подготовке публикации.

7. Пантелеймонов А. В. Математические методы проверки гипотез в тестовом анализе / **А. В. Пантелеймонов**, Е. А. Решетняк, Ю. В. Холин // Тези доп. Четвертої Всеукр. конф. студентів та аспірантів «Сучасні проблеми хімії». – Київ. – 2003. – С. 13-14.

Автор реализовал алгоритм оценки параметров кривых эффективности нелинейным МНК, выступил с устным докладом.

8. Пантелеймонов А. В. Усовершенствованный метод исследования метрологических характеристик методик визуального тестового анализа / **А. В. Пантелеймонов**, Ю. В. Холин // Тези доп. П'ятої Всеукр. конф. студентів та аспірантів «Сучасні проблеми хімії». – Київ. – 2004. – С. 151.

9. Panteleimonov A. V. Application of fuzzy sets theory to solving tasks of multicomponent qualitative analysis / **A. V. Panteleimonov**, Yu. V. Kholin. // Book of abstracts International conference «AC&CA». – Kyiv. – 2005. – P. 48.

10. Пантелеймонов А. В. Вычислительные аспекты определения метрологических характеристик методик визуального тестового анализа / **А. В. Пантелеймонов**, Ю. В. Холин // Матеріали Сесії Наукової Ради НАНУ з проблеми «Аналітична хімія». – Харків. – 2007. – С. 32.

11. Kholin Yu. Chemometric tools at building models of equilibria on surfaces of hybrid silica-organic materials / Yu. Kholin, **A. Panteleimonov** // Book of abstracts International conference «Modern Physical Chemistry for Advanced Materials». – Kharkiv. – 2007. – P. 304-306.

12. Kholin Yu. Metrological characteristics of visual test analysis: chemometric consideration and new computational tools / Yu. Kholin, **A. Panteleimonov** // Book of abstracts 4th Black Sea Basin Conference on Analytical Chemistry. – Sunny Beach, Bulgaria. – 2007. – P141.

13. Пантелеймонов А. В. Скрининг с помощью тест-систем: метрологические характеристики и вычислительные аспекты их определения / А. В. Пантелеймонов // Материалы XV Международной конференции студентов, аспирантов и молодых ученых по фундаментальным наукам «Ломоносов – 2008». Химия. М.: МГУ имени М. В. Ломоносова. – 2008. – С. 60.

АНОТАЦІЯ

Пантелеймонов А.В. Нові хемометричні засоби обробки даних аналітичного експерименту. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата хімічних наук за спеціальністю 02.00.02 – аналітична хімія. – Харківський національний університет імені В.Н. Каразіна, Харків, 2008.

В роботі надано рекомендації з метрології бінарного тестування, обґрунтовано набір метрологічних характеристик методик виявлення з бінарним відгуком. В якості функцій, що подають криві ефективності, рекомендовано застосовувати функції логістичного та експоненційного розподілу. Нелінійний незважений МНК рекомендовано як розрахунковий засіб для визначення підгінних параметрів апроксимуючих функцій. Показано, що він дає оцінки, що збігаються з робастними, одержаними на основі використання теорії нечітких множин.

Розроблено алгоритм ідентифікації сполук, який базується на розрахунку ступеню приналежності аналіту та еталону одній сполуці за допомогою теорії нечітких множин.

Вперше визначено хіміко-аналітичні характеристики двох нових органо-кремнеземних матеріалів за допомогою нового методу, що описує неідеальну адсорбцію в рамках моделі полідентатного зв'язування та як розрахункові засоби використовує методи робастного *M*-оцінювання за Хьюбером та засоби теорії нечітких множин.

Ключові слова: хемометрія, методики виявлення з бінарним відгуком, ідентифікація, органо-кремнеземні матеріали, теорія нечітких множин.

АННОТАЦИЯ

Пантелеймонов А.В. Новые хемометрические способы обработки данных аналитического эксперимента. – Рукопись.

Диссертация на соискание ученой степени кандидата химических наук по специальности 02.00.02 – аналитическая химия. – Харьковский национальный университет имени В.Н. Каразина, Харьков, 2008.

В работе даны рекомендации по метрологии бинарного тестирования, обоснован набор метрологических характеристик методик обнаружения с бинарным откликом, приведены соответствующие расчетные алгоритмы и процедуры. Показано, что определение метрологических характеристик методик с бинарным откликом целесообразно проводить путем исследования кривых эффективности. В качестве функций, аппроксимирующих кривые эффективности, рекомендовано применять функции логистического и экспоненциального распределений. Изучено влияние способа назначения статистических весов на искомые характеристики и доказана возможность проведения расчетов в рамках модели, предполагающей равнозначность эмпирических частот обнаружения аналита. В качестве расчетной процедуры для определения подгонных параметров аппроксимирующих функций рекомендован нелинейный невзвешенный МНК.

Разработан алгоритм идентификации соединений, основанный на вычислении степени принадлежности аналита и эталона одному соединению

с помощью расчетной процедуры, использующей формализм теории нечетких множеств. Исследованы свойства предлагаемого критерия и показана его малая чувствительность к возможному варьированию характеристик аналита и эталона в зависимости от экспериментальных условий.

Впервые определены химико-аналитические характеристики (константы адсорбционных равновесий и эффективные емкости) двух новых органо-кремнеземных материалов – силикагелей, покрытых 3-*n*-пропил(4-метилпиридиний) силсесквиоксан хлоридным полимером и 3-*n*-пропилпиридиний силсесквиоксан хлоридным полимером. Неидеальная адсорбция описана в рамках модели полидентатного связывания. В качестве расчетных средств использованы методы робастного *M*-оценивания по Хьюберу и алгоритмы теории нечетких множеств. Определены условия извлечения хлоридов металлов из водных и этанольных растворов с использованием органо-кремнеземных материалов в качестве сорбентов.

Ключевые слова: хемометрия, методики обнаружения с бинарным откликом, идентификация, органо-кремнеземные материалы, теория нечетких множеств.

SUMMARY

Panteleimonov A.V. New chemometrical tools for handling data of analytical experiment. – Manuscript.

The thesis for a candidate's degree in chemical science by speciality 02.00.02 – analytical chemistry. – V.N. Karazin Kharkiv National University, Kharkiv, 2008.

The thesis deals with metrology of binary testing. A certain set of metrological characteristics is proven to be useful in this context. The logistic and exponential distribution functions are recommended for approximation of performance characteristic curves. The non-linear LSM is recommended as a method for determining fitting parameters of the approximation functions. It is shown that LSM-estimations are closely to robust estimations obtained with the use of the fuzzy sets theory.

An algorithm for identification of compounds has been developed. The algorithm is based on the fuzzy sets theory. It allows labeling the analyte and the standard as the same compound.

The analytical characteristics of two new silica hybrid materials have been determined for the first time. This is done with the aid of a new method which describes a non-ideal adsorption within the polydentate binding model. Methods of Huber

M-estimations and the fuzzy sets theory provided the background of the algorithm.

Keywords: chemometrics, binary testing, identification, silica-organic materials, fuzzy sets theory.