

И.В. Суровцев, А.Э. Татаринов, С.К. Галимов

Моделирование дифференциальных хронопотенциограмм суммой нормальных распределений

Предложены алгоритмы моделирования сигнала дифференциальной хронопотенциограммы суммой нормальных плотностей распределений для увеличения точности измерения концентрации токсичных химических элементов.

The algorithms of the modeling of a differential chronopotentiogram signal by the sum of normal densities of distributions for increasing the concentration measurement accuracy of different chemical elements are suggested.

Запропоновано алгоритми моделювання сигналу диференційної хронопотенціограми сумою нормальних щільностей розподілів для збільшення точності вимірювання токсичних хімічних елементів.

Введение. Вопросам экологического мониторинга окружающей среды в последнее время уделяется повышенное внимание. Один из факторов безопасности – определение концентрации токсических химических элементов (ртуть, мышьяк, свинец, кадмий и другие металлы) в разных объектах окружающей среды.

Для определения массовой концентрации токсических элементов используются различные методы анализа, в том числе инверсионная хронопотенциометрия (ИХП) как разновидность полярографического электрохимического метода.

Метод анализа ИХП использован в анализаторе М-ХА1000-5, разработанном Международным научно-учебным центром информационных технологий и систем НАН и МОН Украины совместно с Национальным университетом биоресурсов и природопользования и ООО «Мириада» (г. Киев), который их выпускает. Анализаторы используются для измерения концентрации токсических элементов при анализе продуктов питания, воды, воздуха, почвы и прочего в лабораториях санитарно-эпидемиологической службы и в других экологических лабораториях. Анализатор внесен в Государственный реестр средств измерительной техники Украины и выпускается серийно [1, 2].

Прежде чем перейти к конкретному изложению постановки задачи, необходимо остановиться на физико-химических аспектах применения метода ИХП для измерения концентрации химических элементов.

Суть метода ИХП состоит в электрохимическом накоплении на измерительном электроде

ионов элементов, содержащихся в жидкой пробе объекта, и последующем измерении времени их электролитического растворения. При постоянных условиях измерения время инверсии элемента (время растворения) прямо пропорционально его концентрации в анализируемой пробе [3, 4].

В процессе работы анализатора автоматически строится кривая растворения или хронопотенциограмма инверсии, которая является функцией потенциала инверсии во времени $F(t)$. По хронопотенциограмме проводится идентификация химических элементов в соответствии с их диапазоном потенциалов растворения. Для определения концентрации химических элементов используется метод «добавки» известной массы ионов анализируемых элементов. Вначале измеряется хронопотенциограмма жидкой пробы объекта, а затем в пробу вносится незначительный объем стандартного раствора с известной концентрацией определяемого элемента, и снова измеряется хронопотенциограмма пробы с добавкой. По разнице времени инверсии элемента определяется его концентрация.

Математически сложно анализировать интегральную кривую, которой является хронопотенциограмма. Особенно это проявляется при анализе концентраций на уровне нижней границы диапазона измерений, когда точки перегиба функции $F(t)$ на хронопотенциограмме найти практически невозможно. Поэтому для определения значений времени инверсии используют не хронопотенциограмму, а дифференциальный график хронопотенциограммы, т.е. функцию $\Delta T(f)$:

$$\Delta T(f) = \frac{dt}{df} = \frac{dF(t)}{df}. \quad (1)$$

Фактически, это интенсивность инверсии во времени (в мс) каждого интервала потенциала (0,25 мВ). Общий вид дифференциальной кривой или графика интенсивности, построенный по хронопотенциограмме инверсии четырех химических элементов, представлен на рис. 1.

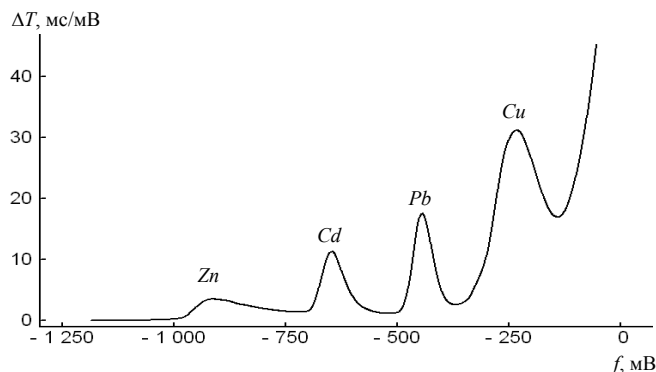


Рис. 1. Зависимость интенсивности ΔT от потенциала f инверсии элементов: Zn – цинк, Cd – кадмий, Pb – свинец, Cu – медь

Каждая точка максимума графика соответствует одному химическому элементу. Точки минимума указывают на то, что закончилась инверсия одного элемента и началась инверсия следующего элемента, а диапазон этих значений называют диапазоном измерения. Значение точки максимума и диапазона измерения – качественная характеристика анализируемого элемента, а количественная характеристика – время инверсии отдельного элемента.

Время инверсии определяется как площадь кривой интенсивности в диапазоне его измерения

$$T_j = \int_{f_{1j}}^{f_{2j}} \Delta T(f) df = \sum_{i=f_{1j}}^{f_{2j}} \Delta T(f_i), \quad (2)$$

где T_j – время инверсии; f_{1j}, f_{2j} – значения потенциалов диапазона измерения j -го элемента.

Первоначально во всех приборах, использующих метод ИХП, время инверсии определялось из графика интенсивности по формуле (2). Ошибка учета площади кривой интенсивности, расположенной под графиком в пределах диапазона измерения отдельного элемента, при больших значениях концентрации (более 0,01 мкг/мл)

была незначительной. Но в новых модификациях анализатора чувствительность прибора достигла значения 0,0005 мкг/мл, при этом ошибка стала сопоставимой с измеряемым значением времени инверсии, что вынуждает искать иные методы анализа кривой. Кроме того, диапазоны измерения «соседних» элементов частично перекрываются, что также вносит ошибку в значение определяемого времени инверсии.

Как ясно из рис. 1, вид кривой для одного элемента очень похож на нормальную плотность распределения случайной величины, а график хронопотенциограммы – на сумму выборочных распределений.

Для повышения точности определения времени инверсии химических элементов при измерении их концентрации необходимо представить дифференциальную хронопотенциограмму в виде математической модели, состоящей из нижней огибающей кривой интенсивности и суммы плотностей нормальных распределений.

Анализ исследований и публикаций по использованию метода ИХП показал, что ранее не ставилась задача моделирования и аппроксимации дифференциальной хронопотенциограммы.

Прежде чем решить общую задачу моделирования, необходимо научиться определять параметры отдельной плотности нормального распределения и суммы плотностей нормальных распределений по исходным данным, используя общедоступные методы аппроксимации. Учитывая, что нижняя огибающая кривой интенсивности строится методом кусочно-линейной аппроксимации, основания плотностей нормальных распределений могут быть расположены с некоторым сдвигом по отношению к оси ординат.

Постановка задачи

Найти простые алгоритмы определения параметров отдельной плотности нормального распределения по исходным данным, учитывая сдвиг основания плотности по оси ординат, применяя метод наименьших квадратов (МНК), использовать их для определения суммы плотностей нормальных распределений.

Решение задачи

Если провести нижнюю огибающую для кривой интенсивности, а затем рассмотреть разность, то полученные значения очень похожи на обычную спектральную кривую спектральных методов анализа. График разности, построенный для указанной кривой интенсивности, представлен на рис. 2.

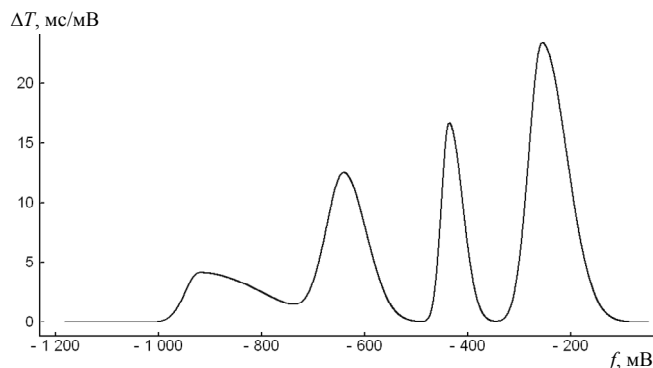


Рис. 2. Зависимость разности между кривой интенсивности и ее нижней огибающей

Из рис. 2 очевидно, что вид кривой для одного элемента очень похож на плотность распределения случайной величины, а график хронопотенциограммы — на сумму выборочных распределений с неизвестными плотностями. График разности между кривой интенсивности и нижней огибающей назовем графиком плотностей распределений, а кривую для одного элемента — выборочной плотностью распределения.

Если представить кривую разности в виде суммы выборочных плотностей распределения, то, согласно центральной предельной теореме [5], отдельное выборочное распределение с неизвестной плотностью $p(f)$ при $N \rightarrow \infty$, где N — равно числу точек, стремится к нормальному распределению со средним f_0 , дисперсией S_f^2 и нормальной плотностью распределения $p(f)$

$$f_0 = \frac{1}{N} \sum_{i=1}^N f_i, \quad S_f^2 = \frac{1}{N-1} \sum_{i=1}^N (f_i - f_0)^2,$$

$$p(f) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{(f - f_0)^2}{2\sigma^2}\right], \quad (3)$$

где f_0 — абсцисса точки максимума, а σ — произвольная положительная постоянная, которую назовем дисперсией выборочной плотности распределения.

Анализ реальных хронопотенциограмм показал, что плотности распределения отдельных химических элементов практически всегда несимметричны и имеют сдвиг основания плотности по оси ординат.

С учетом изложенного, математически задача моделирования ставится так. Необходимо аппроксимировать кривую суммой несимметричных плотностей в виде функций плотности распределения

$$p(f) = \Delta t + H \cdot \exp\left[-\frac{(f - f_0)^2}{2\sigma^2}\right], \quad (4)$$

где Δt — смещение плотности распределения по ординате, H — высота плотности распределения. При $\Delta t = 0$ и равенстве левой и правой дисперсии выборочная плотность совпадает с нормальной плотностью распределения. Критерием точности аппроксимации служит соответствие исходной и расчетной общей площади кривой распределения, с учетом возможной корректировки нижней огибающей кривой интенсивности в процессе моделирования.

Рассмотрим вначале моделирование отдельной плотности распределения без смещения по оси ординат, когда $\Delta t \equiv 0$, в виде нормальной функции плотности распределения

$$p(f) = H \cdot \exp\left[-\frac{(f - f_0)^2}{2\sigma^2}\right], \quad (5)$$

поскольку алгоритмы и точность аппроксимации выборочной плотности для функций (4) и (5) отличаются.

Моделирование отдельной плотности распределения без смещения

Создадим модель несимметричной выборочной плотности с двумя заданными левой и правой дисперсиями σ_1, σ_2 , высотой H и абсциссой точки максимума f_0 по уравнению (5).

Задача аппроксимации ставится следующим образом: по значениям точек модельной выборочной плотности $(f_1, t_1), (f_2, t_2), \dots, (f_i, t_i)$, где

$(f_i, t_i) \in p(f)$ и $(t_i > 0)$ необходимо найти параметры σ_1, σ_2, H и f_0 несимметричной плотности, используя метод МНК.

Ищем решение.

Алгоритм 1. Прологарифмируем функцию плотности $p(f)$

$$\ln(p(f)) = \left[\ln(H) - \frac{f_0^2}{2\sigma^2} \right] + \left[\frac{f_0}{\sigma^2} \right] \cdot f + \left[-\frac{1}{2\sigma^2} \right] \cdot f^2. \quad (6)$$

Аппроксимируем функцию (6) полиномом 2-й степени методом МНК отдельно для левой и правой половины выборочной плотности

$$\ln(t_i) = A + B \cdot f_i + C \cdot f_i^2; \quad i = 1, 2, \dots, N, \quad (7)$$

тогда $A = \ln(H) - \frac{f_0^2}{2\sigma^2}; \quad B = \frac{f_0}{\sigma^2}; \quad C = -\frac{1}{2\sigma^2},$

откуда $\sigma_j^2 = \frac{1}{2 \cdot |C|}; \quad \sigma_j = \sqrt{\sigma_j^2}; \quad f_{0j} = B \cdot \sigma_j^2;$

$$H_j = \exp \left[A + \frac{f_{0j}^2}{2 \cdot \sigma_j^2} \right]; \quad j = 1, 2;$$

$$H = (H_1 + H_2) / 2; \quad f_0 = (f_{01} + f_{02}) / 2. \quad (8)$$

Особенность данного алгоритма – возможность определения параметров плотности распределения не по всем исходным точкам выборочной плотности, а только по точкам близлежащим к вершине исходной кривой, что позволяет использовать этот алгоритм при определении суммы плотностей или для анализа других спектральных кривых.

Достаточное количество точек N , необходимых для аппроксимации, определяется выражением

$$N = \text{целое значение} (20 \div 30 / \Delta f), \quad (9)$$

где Δf – шаг исходных данных по оси ординат.

Моделирование отдельной плотности распределения со смещением по оси ординат

Создадим модель несимметричной выборочной плотности с двумя заданными левой и правой дисперсиями σ_1, σ_2 , высотой H , абсциссой

точки максимума f_0 и смещением Δt по уравнению (4).

Задача аппроксимации ставится следующим образом: для значений точек модельной выборочной плотности $(f_1, t_1), (f_2, t_2), \dots, (f_i, t_i)$, где $(f_i, t_i) \in p(f)$ и $(t_i > 0)$, необходимо найти параметры $\sigma_1, \sigma_2, H, f_0$ и Δt несимметричной плотности, используя метод МНК.

Ищем решение.

Алгоритм 2. По значениям модели смещенной несимметричной выборочной плотности найдем координаты точки максимума (t_0, f_0) , а также производную от функции $p(f)$ по переменной f

$$p'(f) = \frac{H \cdot (f - f_0)}{\sigma^2} \cdot \exp \left[-\frac{(f - f_0)^2}{2\sigma^2} \right]. \quad (10)$$

Введем новую функцию $\lambda(f)$, поделив левую и правую часть (10) на выражение $(f - f_0)$:

$$\lambda(f) = \frac{p'(f)}{f - f_0} = \frac{H}{\sigma^2} \cdot \exp \left[-\frac{(f - f_0)^2}{2\sigma^2} \right]. \quad (11)$$

Прологарифмируем зависимость (11):

$$\ln(\lambda(f)) = \left[\ln \left(\frac{H}{\sigma^2} \right) - \frac{f_0^2}{2\sigma^2} \right] + \left[\frac{f_0}{\sigma^2} \right] \cdot f + \left[-\frac{1}{2\sigma^2} \right] \cdot f^2. \quad (12)$$

Аппроксимируем функцию (12) полиномом 2-й степени методом МНК отдельно для левой и правой половины выборочной плотности. Учитывая, что для дискретных значений выборочной плотности $p'(f_i) = t_i - t_{i-1}$, получим

$$\ln \left(\frac{t_i - t_{i-1}}{f_i - f_0} \right) = A + B \cdot f_i + C \cdot f_i^2, \quad (13)$$

$$i = 1, 2, \dots, N,$$

тогда $A = \ln \left(\frac{H}{\sigma^2} \right) - \frac{f_0^2}{2\sigma^2}; \quad B = \frac{f_0}{\sigma^2}; \quad C = -\frac{1}{2\sigma^2},$

откуда $\sigma_j^2 = \frac{1}{2 \cdot |C|}; \quad \sigma_j = \sqrt{\sigma_j^2}; \quad f_{0j} = B \cdot \sigma_j^2;$

$$H_j = \sigma_j^2 \cdot \exp \left[A + \frac{f_{0j}^2}{2 \cdot \sigma_j^2} \right]; j = 1, 2;$$

$$H = (H_1 + H_2) / 2; f_0 = (f_{01} + f_{02}) / 2;$$

$$\Delta t = t_0 - H. \quad (14)$$

Примеры моделирования плотности распределения

Пример 1. Создадим модельную несимметричную выборочную плотность для следующих значений параметров: $\sigma_1 = 40,567$, $\sigma_2 = 60,234$, $H = 5,051$, $f_0 = -470,4$, $f_i \in [-1500, 0]$, $\Delta f = 0,25$, $N = 120$. Аппроксимируем модельную выборочную плотность без смещения по оси ординат при помощи алгоритма 1.

На рис. 3 приведены графические результаты моделирования выборочной плотности (сплошная линия – модельная выборочная плотность, точки – расчетная нормальная плотность распределения).

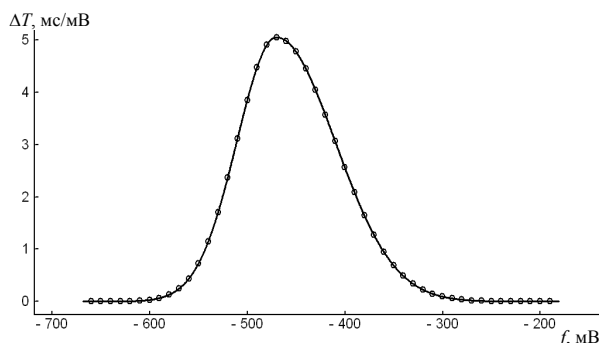


Рис. 3. Моделирование плотности распределения без смещения по ординате

Пример 2. Создадим модельные несимметричные выборочные плотности со значениями параметров σ_1 , σ_2 , H , f_0 , приведенными в примере 1, но для разных значений смещения по оси ординат: $\Delta t_1 = +0,814$, $\Delta t_2 = -0,814$, $\Delta t_3 = 0$. Аппроксимируем модельные выборочные плотности при помощи алгоритма 2.

На рис. 4 представлены графические результаты моделирования выборочной плотности для положительного значения смещения по оси ординат, а на рис. 5 – для отрицательного значения смещения.

В таблице приведены значения исходных и расчетных данных всех моделей: σ_1 – левая дисперсия; σ_2 – правая дисперсия; H – высота

плотности; f_0 – абсцисса точки максимума; Δt_m , Δt – смещение плотности по оси ординат модельное и расчетное; CKO_1 , CKO_2 – среднеквадратические ошибки аппроксимации по методу МНК соответственно левой и правой половины плотности распределения для N точек (9); CKO – среднеквадратическая ошибка аппроксимации всех точек распределения; T_m , T – площадь модельной и расчетной плотности распределения, определяемых по (2).

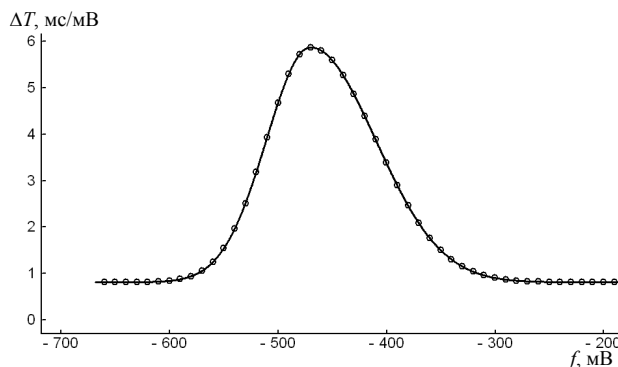


Рис. 4. Моделирование плотности распределения с положительным смещением по оси ординат

Т а б л и ц а. Результаты моделирования плотностей распределения

Наименование параметра	Значения параметров				
	Исходные параметры	Модель без смещения	Модель со смещением по оси ординат		
			$\Delta t_m \Delta t \equiv 0$	$\Delta t_{m1} = +0,814$	$\Delta t_{m2} = -0,814$
σ_1	40,567	40,567	40,579	40,649	40,579
CKO_1	–	$5,19 \cdot 10^{-17}$	$4,37 \cdot 10^{-5}$	$2,46 \cdot 10^{-6}$	$4,37 \cdot 10^{-5}$
σ_2	60,234	60,234	60,217	60,108	60,217
CKO_2	–	$1,45 \cdot 10^{-17}$	$7,08 \cdot 10^{-5}$	$1,40 \cdot 10^{-5}$	$7,08 \cdot 10^{-5}$
H	5,051	5,051	5,051	5,051	5,051
f_0	-470,400	-470,400	-470,346	-470,245	-470,346
Δt	–	–	+0,814	-0,814	0,000
T_m	–	638,12	1034,74	445,58	638,12
T	–	638,12	1034,70	445,40	638,08
CKO	–	$2,00 \cdot 10^{-6}$	$1,20 \cdot 10^{-3}$	$3,20 \cdot 10^{-3}$	$1,20 \cdot 10^{-3}$

Среднеквадратическая ошибка аппроксимации рассчитывается по формуле

$$CKO = \sqrt{\frac{\sum_{i=1}^N (t_{m_i} - t_i)^2}{N-1}}, i = 1, \dots, N, \quad (15)$$

где t_{m_i} , t_i – модельные и расчетные значения ординат плотности распределения, N – число точек.

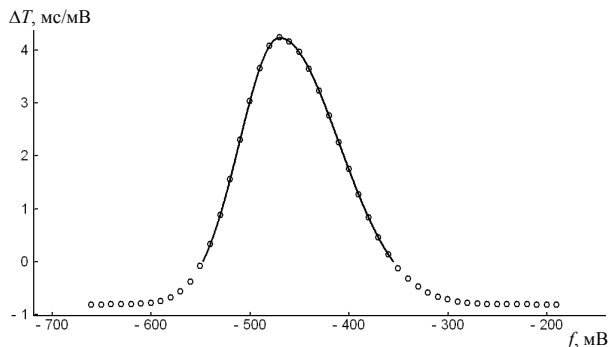


Рис. 5. Моделирование плотности распределения с отрицательным смещением по оси ординат

Моделирование суммы плотностей распределения без смещения

Аппроксимируем каждую выборочную плотность суммы плотностей распределения при помощи алгоритма 2 с последовательным вычитанием крайней левой расчетной плотности при движении «вперед» и с последовательным вычитанием крайней правой плотности при движении «назад». Число циклов аппроксимации суммы обычно не более пяти. Критерием остановки цикла аппроксимации служит достижение заданного значения относительной ошибки определения площади расчетной суммы к исходной. Как правило, для инженерных расчетов это значение порядка одного процента.

Пример 3. Создадим сумму из четырех модельных несимметричных выборочных плотностей, имеющих распределение похожее на реальный график плотностей распределений, представленный на рис. 2, с нулевым смещением по оси ординат.

На рис. 6 приведены графические результаты моделирования суммы несимметричных выборочных плотностей. Численные значения параметров модели и их расчетные значения практически совпадают и в статье не приводятся, отметим только, что значение общей площади модели равно 4944,9, а расчетное значение – 4939,9 при относительной ошибке – 0,101% и СКО моделирования всех точек суммы распределений – 0,0176.

Заключение. Результаты моделирования отдельных выборочных плотностей и суммы распределений, приведенные в примерах 1–3, показали, что алгоритмы 1–2 позволяют очень качественно аппроксимировать выборочные плотности распределений и могут быть успешно использованы как для определения времени инверсии по кривой распределения дифференциальной хронопотенциограммы, так и для анализа других спектральных кривых. Причем по точности аппроксимации алгоритм 2 хуже алгоритма 1, поскольку использует данные численного дифференцирования ординат исходных точек.

Однако применение теоретических результатов при анализе реальных кривых распределения дифференциальной хронопотенциограммы показало, что имеются значительные сложности при определении последовательного смещения по оси ординат выборочных плотно-

стей суммы распределений, поскольку нижняя огибающая кривой интенсивности представляет собой нелинейную экспоненциальную кривую, поэтому ее реальное расположение на плоскости может быть определено только итерационным методом.

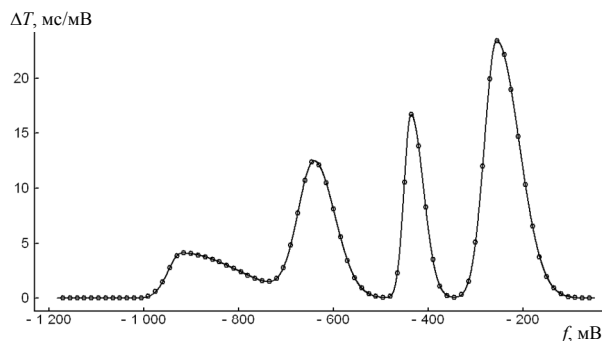


Рис. 6. Моделирование суммы несимметричных выборочных плотностей

Кроме того, порядок кривизны реальных выборочных плотностей иногда выше (степень 3 или 4) порядка кривизны нормального распределения по формулам (4–5), равного степени 2. А это значит, что добавляются новые алгоритмы аппроксимации и усложняется процесс анализа суммы выборочных распределений.

Дальнейшие исследования направлены на изучение данной проблемы и на математическое моделирование всего процесса определения концентрации химических элементов по методу ИХП.

1. *Інверсійно-хронопотенціометричне визначення важких металів в об'єктах навколишнього середовища / О.І. Карнаухов, О.М. Полумбрик, А.Т. Безніс та ін. // Наук.-методич. розробка. – К.: УГУПТ, 1997. – 88 с.*
2. *Визначення вмісту рухомих форм важких металів ртуті та миш'яку у воді, ґрунтах, донних відкладеннях, опадах стічних вод методом інверсійної хронопотенціометрії / О.І. Карнаухов, В.М. Галімова, І.В. Суровцев та ін. // Наук.-методич. розробка. – К.: НУБіПУ, 2008. – 26 с.*
3. *Карнаухов О.І., Галімова В.М., Галімов К.Р. Теорія інверсійної хронопотенціометрії із заданим опором ланцюга // Наук. вісн. НАУ. – 2000. – № 32. – С. 204–209.*
4. *Татарінов А.Э., Суровцев И.В. Использование методов импульсной хронопотенциометрии при измерении концентрации тяжелых металлов // Вісн. Вінницького політехн. ін-ту. – 2006. – 6(69). – С. 101–105.*
5. *Бендат Дж., Пирсол А. Прикладной анализ случайных данных. – М.: Мир. – 1989. – 540с.*

Поступила 18.09.2009
Тел. для справок: (044) 502-6337, (Київ)
E-mail: dep175@irts.org.ua

© И.В. Суровцев, А.Э. Татаринев, С.К. Галимов, 2009