

ЧТО НЕ ТАК С ФАРМАКОПЕЙНЫМ ОПРЕДЕЛЕНИЕМ ОТНОШЕНИЯ СИГНАЛ/ШУМ?

Каламбет Ю. А.

ООО Амперсанд, 117437, город Москва, улица Островитянова, дом 25, корпус 1,
квартира 121
support@ampersand.ru

Фармакопеи определяют шум как показано на Рисунке 1 [1].

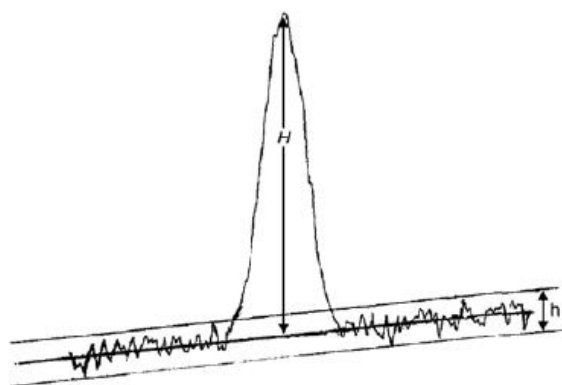


Рисунок 1. Иллюстрация определения сигнал/шум из Российской фармакопеи 13.
 $S/N = 2H/h = H/(h/2)$

Как видно из рисунка, мерой шума выступает его размах пик-к-пику. Альтернативой, предлагаемой во всех программах обработки данных, является пяти- или шестикратное стандартное отклонение распределения шума (СКО). Далее мы будем полагать этот коэффициент равным шести, как принято в программе МультиХром. Согласно определению ASTM [2] хроматограмма базовой линии разделяется на короткие части и вычисляется среднее значение шума пик-к-пику для каждой части. Следует отметить, что размах шума пик-к-пику – плохая статистическая мера шума. Рисунок 2 показывает зависимость шума пик-к-пику от числа точек данных во фрагменте [3].

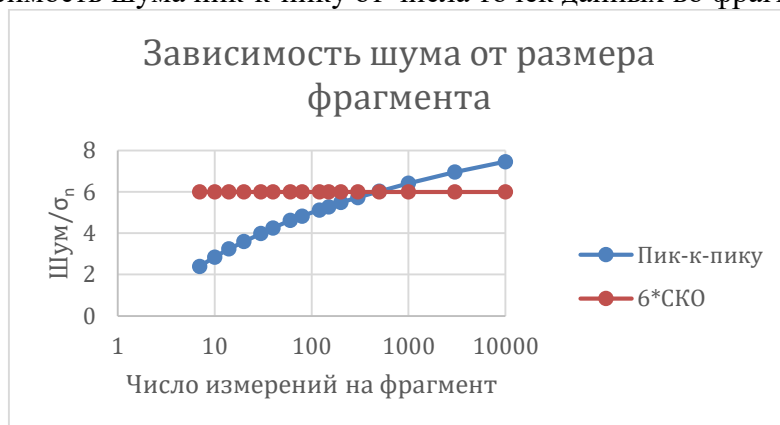


Рисунок 2. Базовая линия моделирована 120 тысячами измерений, распределенных как нормальный шум со стандартным отклонением σ_n . Массив данных был разделен на фрагменты с равным числом измерений, размер фрагмента на оси абсцисс; по оси ординат отложен средний шум, деленный на σ_n для соответствующих фрагментов.

Глядя на график рис.2 нетрудно заметить, что определение шума детектора через размах пик-к-пику дает смещенную и несостоятельную оценку уровня шума: состоятельная оценка должна стремиться к оцениваемому параметру при увеличении числа измерений в

выборке (фрагменте), тогда как размах пик-к-пику при увеличении числа измерений неограниченно растет. Несмещенная оценка не должна зависеть от размера выборки (фрагмента), как это происходит при оценке шума по СКО, а математическое ожидание размаха зависит от размера выборки (количества измерений во фрагменте). Метод $6 \cdot \text{СКО}$ обеспечивает несмещенную и состоятельную оценку: уровень шума не зависит от размера фрагмента и при большом числе измерений выборки приближается к истинному значению оцениваемого параметра (шума). Все расчеты производились с использованием программного обеспечения Microsoft Excel и/или программы «МультиХром» © ООО «Амперсенд».

Эта проблема может быть весьма актуальной: минимальное число измерений, требуемое для оценки параметров пика путем аппроксимации, составляет около 15 на FWHM, размер области базовой линии, который, согласно фармакопеям, должен использоваться для оценки шума, равен $5 \cdot \text{FWHM} \approx 75$ точек. Оценка шума в этом случае составляет около $4,7\sigma$ и она, по крайней мере, на 20% меньше, чем несмещенная оценка по формуле $6 \cdot \text{SD}$. Ситуация с очень узкими пиками еще хуже, так как формальная область, по которой оценивается шум базовой линии может уменьшиться до 6...15 измерений, и среднее значение погрешности пик-к-пику окажется приблизительно в два раза ниже несмещенной оценки погрешности по формуле $6 \cdot \text{SD}$.

Что делать?

Оценивать значение аналитического сигнала и границы доверительных областей, фильтруя шум алгоритмом ЛОПФ [4,5]. По результатам фильтрации можно оценить хроматограмму как аналитический сигнал, вычислить погрешность для измерения с любым порядковым номером, и по этим данным оценить погрешности как высоты, так и площади пика. Кроме того:

Строить градуировочный график с оценкой границ доверительных областей и оценивать пределы определения и обнаружения по градуировочному графику. Доверительные области градуировочного графика позволяют применять любой алгоритм фильтрации шумов, но при применении ЛОПФ результат имеет меньшую погрешность.

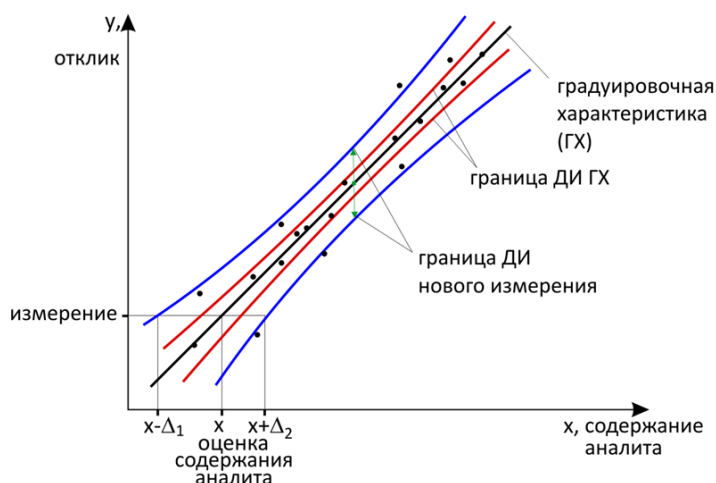


Рисунок 3. Градуировочный график с границами доверительных областей

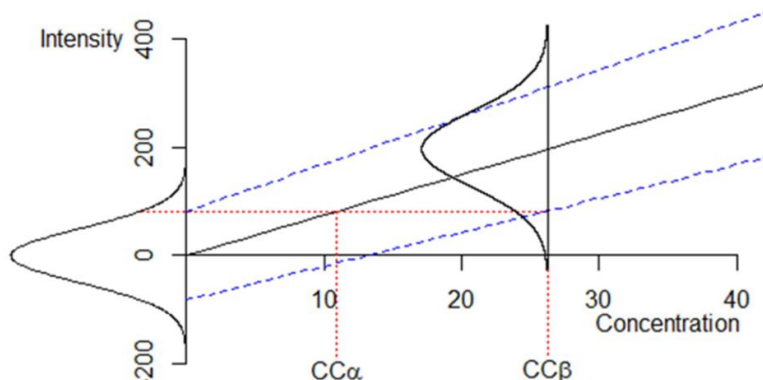


Рисунок 4. Пределы обнаружения и определения по градуировочному графику

Итак, что не так с фармакопейным определением отношения сигнал/шум:

- не строится кривая, описывающая аналитический сигнал или эта кривая считается проведенной по измерениям «точка к точке». Оценка высоты пика производится без аппроксимации его вершины;

- не строится доверительный интервал аналитического сигнала, в том числе вершины пика и точек начала/конца базовой линии;

- мера погрешности, предлагаемая к использованию (размах шума), даёт смещённую и несостоятельную оценку уровня шума.

Единственное преимущество шума пик-к-пику заключается в том, что его можно измерить линейкой, используя хроматограмму, записанную на ленте самописца.

Литература

1. Государственная фармакопея Российской Федерации. XIII издание. Федеральная электронная медицинская библиотека, 2015.
2. KITAJIMA A. et al. Baseline Noise and Measurement Uncertainty in Liquid Chromatography // Anal. Sci. 2007. Vol. 23, № 9. P. 1077–1080.
3. Kalambet Y.A. Data acquisition and integration // Gas Chromatography. Second Edi / ed. Poole K.F. Amsterdam, Netherlands: Elsevier, 2021. P. 505–524.
4. Каламбет Ю.А., Мальцев С.А., Козьмин Ю.П. Фильтрация шумов: окончательное решение проблемы // Аналитика. 2011. Vol. 1, № 1. P. 50–55.
5. Каламбет Ю.А. Оптимизация параметров линейного сглаживания хроматографических пиков // Научное приборостроение. 2019. Vol. 29, № 3. P. 51–62.

УДК 543.544

ХРОМАТОГРАФИЯ В НОВОКУЙБЫШЕВСКЕ

Занозин И.Ю.¹, Занозина И.И.^{1,2}

¹АО «Средневолжский научно-исследовательский институт по нефтепереработке»,
446200, г. Новокуйбышевск, ул. Научная, 1

²ФГБОУ ВО «Самарский Государственный Технический Университет»,
443100, Самарская область, г. Самара, ул. Молодогвардейская, 244
zanozinaii@sni.rosneft.ru

Новокуйбышевск – известный в России и за ее рубежами город нефтепереработки и нефтехимии, получивший свое название как город-спутник Куйбышева в феврале 1952 года. С этого момента можно и начать отсчет становления и развития хроматографии в Новокуйбышевске изначально на промышленных предприятиях, когда при строительстве и пуске новых установок вводили в действие лаборатории для оперативного аналитического контроля технологических процессов, а затем в крупных отраслевых научно-исследовательских институтах и филиалах всесоюзного значения, получивших «прописку» в городе Новокуйбышевск.