

Робастный метод определения количественного состава многокомпонентных стационарных газовых смесей

12, декабрь 2014

Белов М. Л., Антонюк Т. К., Городничев В. А.

УДК: 621.378:551.508

Россия, МГТУ им. Н.Э. Баумана

belov@bmstu.ru

Введение

Одной из важнейших в настоящее время является проблема охраны окружающей среды и, в первую очередь, атмосферы. Постоянно возрастает количество выбрасываемых в атмосферу газообразных загрязняющих веществ, многие из которых сильно влияют на естественные химические и физические процессы в природе и представляют опасность для жизнедеятельности человека.

Для экологического мониторинга атмосферы и оперативного контроля выбросов различных источников загрязнений требуется создание высокочувствительных быстродействующих газоанализаторов, способных контролировать содержание в атмосферном воздухе газовых загрязнителей. Лазерные методы являются наиболее перспективными для оперативного дистанционного и локального газоанализа многокомпонентных загрязнений атмосферы [1-3].

В настоящее время одной из проблем, возникающих при использовании лазерных методов газоанализа, является неустойчивость результатов определения количественного состава газов при контроле многокомпонентных смесей. Это вызывает необходимость применения специальных алгоритмов обработки результатов лазерных измерений.

Для многокомпонентных газовых смесей при решении задачи определения концентраций газов из результатов многоспектральных лазерных измерений могут быть эффективно использованы методы решения некорректных математических задач (см., например, [2,3]).

Эти методы позволяют по данным одиночных (на каждой длине волны) измерений восстанавливать концентрации газов в многокомпонентных газовых смесях, даже если линии поглощения газов перекрываются.

Ниже рассматривается задача лазерного газанализа многокомпонентных стационарных газовых смесей, для которых существует возможность проведения серии измерений на каждой длине волны.

1. Постановка задачи

Если качественный состав газовой смеси известен, то измерение концентраций компонент смеси может быть проведено с использованием перестраиваемого по длине волны источника лазерного излучения, путем регистрации поглощения излучения анализируемой газовой смесью для определенного набора спектральных каналов измерения.

Для определенности будем считать, что анализируемая газовая смесь содержит N газовых компонент и измерения проводятся на M длинах волн, а ширина линий генерации лазера много меньше ширины линий поглощения анализируемых газов. Тогда задача нахождения концентраций газов из результатов многоспектральных измерений сводится к решению системы линейных алгебраических уравнений вида (см., например, [2,3])

$$\left\{ \begin{array}{l} k_a(\lambda_1) + \sum_j^N c_j k_j(\lambda_1) = y(\lambda_1) \\ \\ k_a(\lambda_M) + \sum_j^N c_j k_j(\lambda_M) = y(\lambda_M) \end{array} \right., \quad (1)$$

где

$y(\lambda_i)$ – приведенный измеряемый сигнал на длине волны λ_i ;

$k_a(\lambda_i)$ - коэффициент неселективного ослабления на длине волны λ_i ;

$k_j(\lambda_i)$ - коэффициент поглощения j-ой газовой компоненты смеси на длине волны λ_i ;

C_j - концентрация j-ой газовой компоненты смеси;

N - полное число газовых компонент в анализируемой смеси.

В матричной форме система уравнений (1) лазерного газоанализа имеет вид

$$W\vec{x} = \vec{k}_q + K\vec{c} = \vec{y}, \quad (2)$$

где

W - матрица системы (1) размерностью $M \times M$;

$\vec{X}$ - M -мерный искомый вектор (N компонент этого вектора соответствуют концентрациям газов);

 $\vec{k}_a$ - M-мерный вектор коэффициентов неселективного поглощения;

K - матрица коэффициентов поглощения компонент газовой смеси размерностью $M \times N$;

$\vec{c}$ - N -мерный вектор концентраций газов;

$\vec{y}$ - M -мерный вектор приведенных измеряемых сигналов (показателей поглощения исследуемой смеси).

Неизвестными величинами в системе уравнений (1) являются концентрации c_j и коэффициенты неселективного поглощения $k_a(\lambda_i)$.

Величины $k_a(\lambda_i)$ слабо зависят от длины волны. Для устранения их влияния при определении концентрации каждого газа измерения выполняются на паре лазерных длин волн: $\lambda_{\text{он}}$ - длине волны, на которой поглощение исследуемым газом максимально, λ_{off} - длине волны, на которой поглощение исследуемым газом практически отсутствует. Длины волн $\lambda_{\text{он}}$ и λ_{off} выбирают расположенными попарно (для каждого газа) достаточно близко. Поэтому обычно считают, что для каждой пары длин волн $\lambda_{\text{он}}$ и λ_{off} коэффициенты k_a можно положить равными константе. В этом случае из M спектральных каналов, необходимых для зондирования газовой смеси, информация $M/2$ каналов требуется для определения коэффициентов k_a .

Трудность решения системы уравнений (1) заключается в том, что правая часть уравнения всегда известна со случайной ошибкой, обусловленной погрешностями измерения, шумами аппаратуры и т.п. В этих условиях попытки непосредственно обратить систему уравнений (1) приводят к тому, что найденный обратный оператор не обладает свойством устойчивости и малые вариации данных измерений приводят к большим вариациям искомых величин. Выход из этой трудности заключается в использовании алгоритмов обработки, основанных на методах решения некорректных математических задач (см., например, [2-6]).

В настоящее время для определения концентраций газов в многокомпонентных смесях по данным многоспектрального измерений (при одном измерении на каждой длине волны) эффективно используются метод поиска квазирешений, байесовский метод или метод регуляризации Тихонова с применением различных способов (как детерминистических, так и статистических) выбора параметра регуляризации (см., например, [2,3,7-9]).

Однако, если смесь стационарная и существует возможность проведения серии измерений на каждой длине волны можно использовать гораздо более простой метод определения концентраций газов - метод наименьших квадратов (МНК), основанный на минимизации суммы квадратов некоторых функций от искомых переменных.

Пусть c_j^l - концентрация j -ой газовой компоненты смеси определенная (стандартным методом решения системы (1) линейных алгебраических уравнений) по результатам l -ого измерения, n - количество измерений в серии, $\hat{c}_j$ - оценка концентрации j -ой газовой компоненты смеси определенная методом наименьших квадратов.

Введем функцию невязки $E(d)$

$$E(d) = \sum_{l=1}^n [c_j^l - \hat{c}_j]^2. \quad (3)$$

Значение $\hat{c}_j$, минимизирующее функцию $E(d)$, будет искомой оценкой концентрации j -ой газовой компоненты смеси, определенной методом наименьших квадратов.

Оценки, полученные методом наименьших квадратов, являются эффективными, если погрешности измерений распределены по нормальному закону. Однако на практике часто нормальность закона распределения погрешностей нарушается. Нарушения нормальности закона распределения могут приводить к значительной потере эффективности оценок, полученных методом наименьших квадратов [10,11].

Особенно большая потеря эффективности оценок, полученных методом наименьших квадратов, происходит при наличии даже небольшой доли выбросов. На практике большим выбросам соответствуют измерения, реальная погрешность которых значительно превосходит приписываемое им среднеквадратическое отклонение, причем априорно ни реальная величина таких погрешностей, ни моменты появления выбросов не известны [10,11].

Поэтому актуальной является разработка метода определения количественного состава многокомпонентных стационарных газовых смесей, устойчивого к выбросам измеряемого сигнала.

2. Робастный метод определения количественного состава многокомпонентных стационарных газовых смесей

При наличии больших выбросов необходимо применять робастные (устойчивые) методы оценивания [10,11], позволяющие значительно снизить влияние на оценку больших выбросов и получить приемлемую итоговую оценку искомых параметров.

Существуют разные робастные методы оценивания [10,11]. Запишем функцию невязки $E(d)$ в более общем (чем формула (3)) виде

$$E(d) = \sum_{l=1}^n \rho(c_j^l - \hat{c}_j), \quad (4)$$

Здесь $\rho(x)$ - выпуклая, симметричная, положительная функция.

Общий подход к выбору функции $\rho(x)$ основан на том, что оценка доставляющая минимум функции невязки (4), должна обладать робастностью по отношению к большим выбросам и «не портить» измерения, не являющиеся большими выбросами.

Например, функция $\rho(x)$ должна быть (желательно) квадратичной при малых значениях $|x|$ и менее возрастающей по сравнению с квадратичной функцией при больших значениях $|x|$. В качестве конкретных функций, порождающих конкретные робастные оценки, могут быть взяты различные функции [10-12]:

- для робастной оценки Хьюберта (Huber)

$$\rho(x) = \begin{cases} x^2 / 2, & |x| < K \\ -K^2 / 2 + K|x|, & |x| > K \end{cases},$$

где K - параметр Хьюберта.

- для робастной оценки Германа-Маклара (German-Maclure)

$$\rho(x) = \frac{x^2 / 2}{1 + x^2}.$$

- для робастной оценки Уэлша (Welsch)

$$\rho(x) = \frac{c^2}{2} [1 - \exp(-(\frac{x}{c})^2)].$$

Другой подход к формированию робастной оценки, основан на замене квадратичной формы в функции (3) невязки $E(d)$ на L_p -форму:

$$\rho(x) = |x|^p.$$

В рамках последнего метода вместо задачи минимизации невязки $E(d)$ квадратичной формы (3) рассматривается задача минимизации невязки $\hat{E}(d)$:

$$\hat{E}(d) = \sum_{l=1}^n |c_j^l - \hat{c}_j|^p. \quad (5)$$

Здесь $1 \leq p \leq 2$.

При уменьшении p вплоть до единицы «степень робастности» оценки (полученной при решении задачи минимизации $\hat{E}(d)$) увеличивается. Причем, поскольку при $p \geq 1$ функционал правой части (6) выпуклый, то существует единственная робастная оценка в задаче минимизации $\hat{E}(d)$.

Численное нахождение самой робастной оценки (при решении задачи минимизации $\hat{E}(d)$) можно осуществить любыми стандартными методами.

4. Математическое моделирование

Для оценки эффективности робастного метода определения количественного состава многокомпонентных стационарных газовых смесей из многоспектральных лазерных измерений было проведено математическое моделирование. Моделирование проводилось для газовых смесей с числом компонент от 4 до 6.

Считалась, что газовая смесь стационарная и для определения количественного состава газов в смеси проводится n измерений (n задавалось от 2 до 6). Концентрации газов при единичных измерениях c_j^l находились стандартным методом решения матричного уравнения (2). Для нахождения концентраций газов по n измерениям использовалась функция невязки (5).

Найденные значения концентраций газов $\tilde{c}_j$ сравнивались с исходными (заданными) значениям концентраций c_j и определялась погрешность определения концентраций газов:

$$\delta_j = \frac{|c_j - \tilde{c}_j|}{c_j}.$$

Шум измерения задавался с пуассоновским распределением, моделирующим редкие выбросы, с параметром распределения $\lambda = 0,1$. Измеряемый сигнал $\tilde{y}_k$ в k -ом спектральном канале моделировался как:

$$\tilde{y}_k = y_k(1 + a\xi_i),$$

где

y_k - сигнала в случае отсутствия шума измерения;

a - вес (в величине сигнала) пуассоновского шума ξ_i .

На рисунках 1-7 приведены примеры результатов математического моделирования для шестикомпонентной смеси этилен-фреон-12-гидразин-аммиак-метанол-этилакрилат. Для контроля количественного состава смеси были использованы следующие спектральные каналы измерения: 10P14, 10P10, 10R18, 10R22, 10P22, 10P28, 10R6, 10R12, 9P14, 9P20, 9R14, 9R22 (они соответствуют дискретно перестраиваемому по длине волны лазеру на углекислом газе [3]). Концентрации газов задавались равными: 10^{-4} ; $2 - 4,6 \cdot 10^{-3}$; $3 - 8 \cdot 10^{-3}$; $4 - 8,5 \cdot 10^{-3}$; $5 - 6,6 \cdot 10^{-3}$; $6 - 9,2 \cdot 10^{-3}$. Число измерений n полагалось равным 6.

На рисунке 1 показана реализация моделируемого «измеренного» сигнала в первом спектральном канале. На рисунке i - номер шумовой реализации.

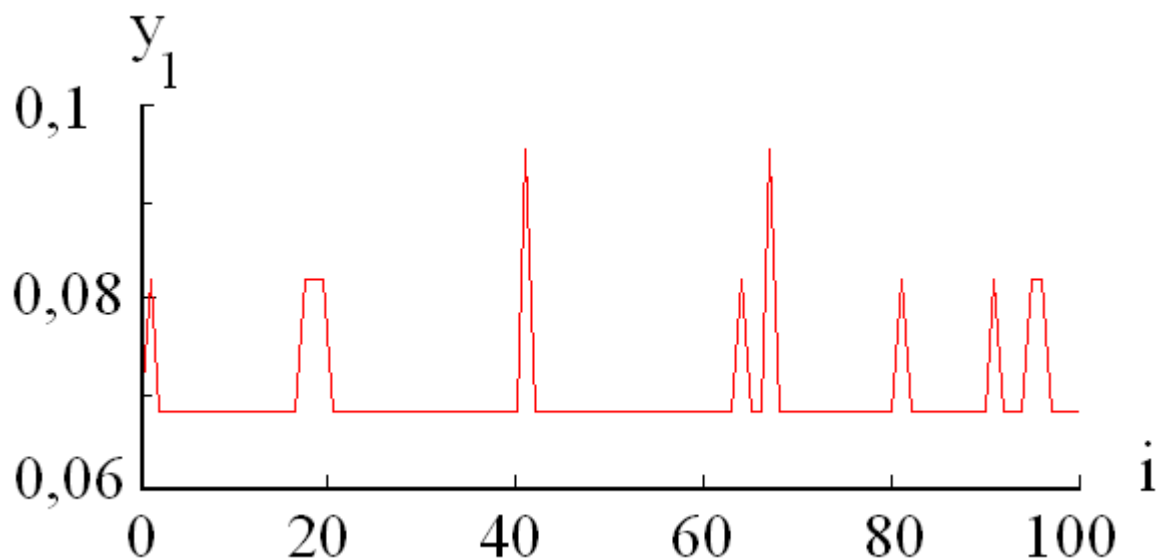


Рис.1. Реализация моделируемого сигнала в первом спектральном канале

На рисунках 2-7 приведены средние (по серии из 1000 шумовых реализаций) погрешности δ (в процентах) определения концентрации компонент газовой смеси - этилена (рисунки 2-4 для разного веса α пуассоновского шума) и метанола (рисунки 5-7 для разного веса α пуассоновского шума).

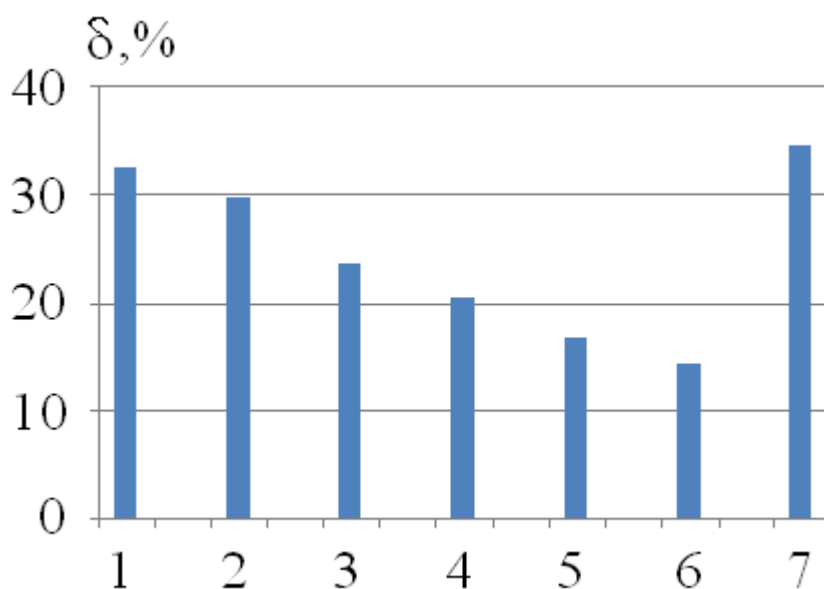


Рис.2 Погрешности определения концентрации этилена. $\alpha = 0,1$.

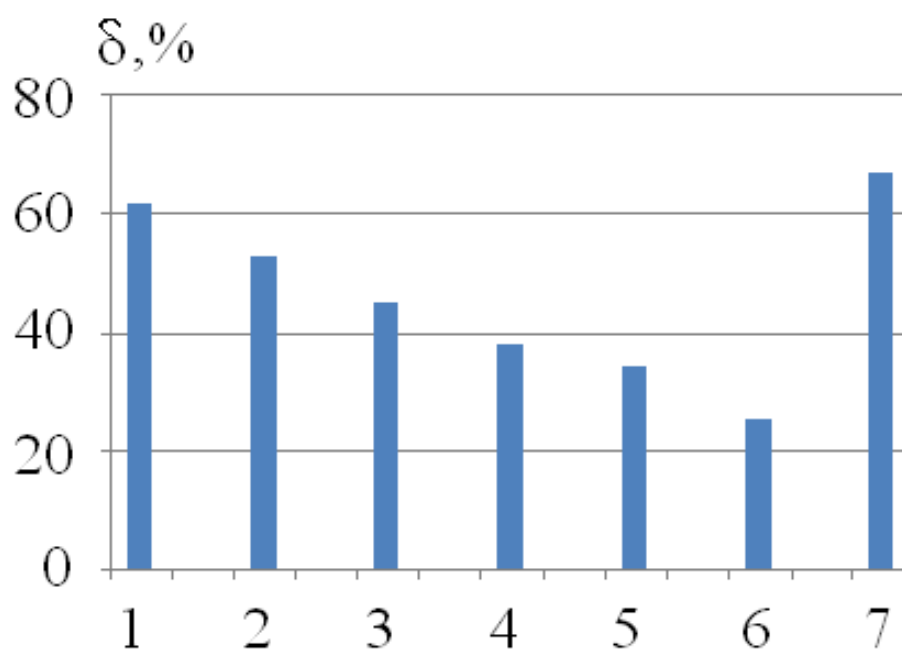


Рис.3 Погрешности определения концентрации этилена. $a = 0,2$.

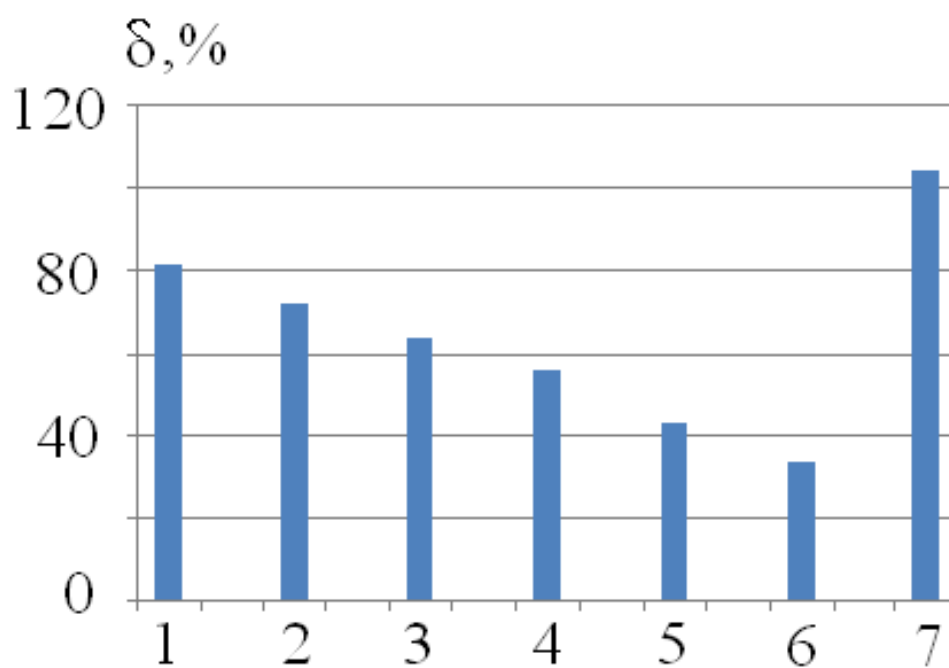


Рис.4 Погрешности определения концентрации этилена. $a = 0,3$.

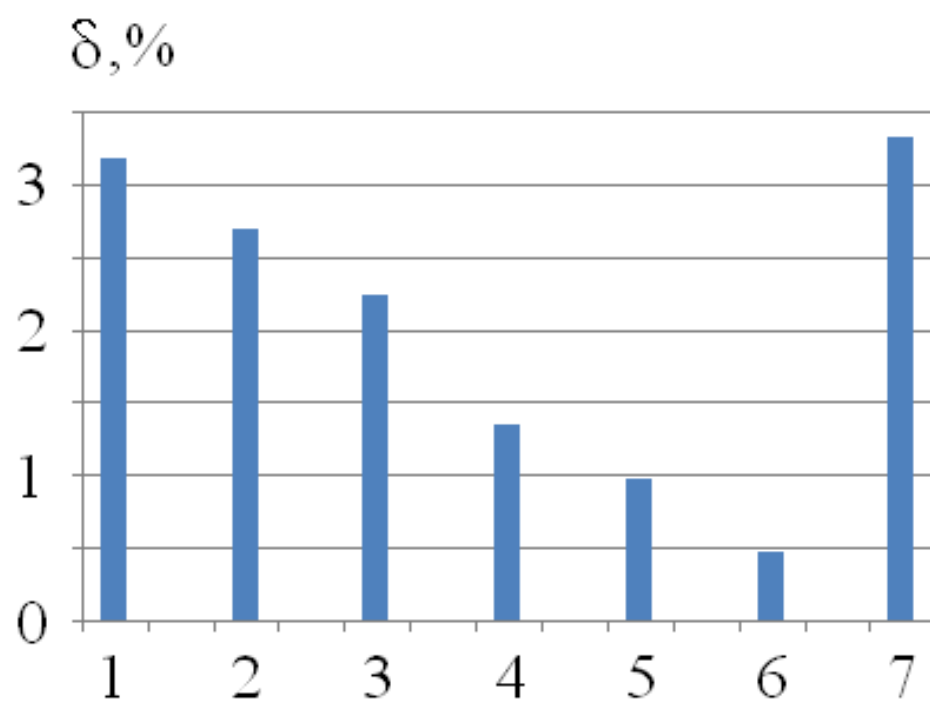


Рис.5 Погрешности определения концентрации метанола. $a = 0,1$.

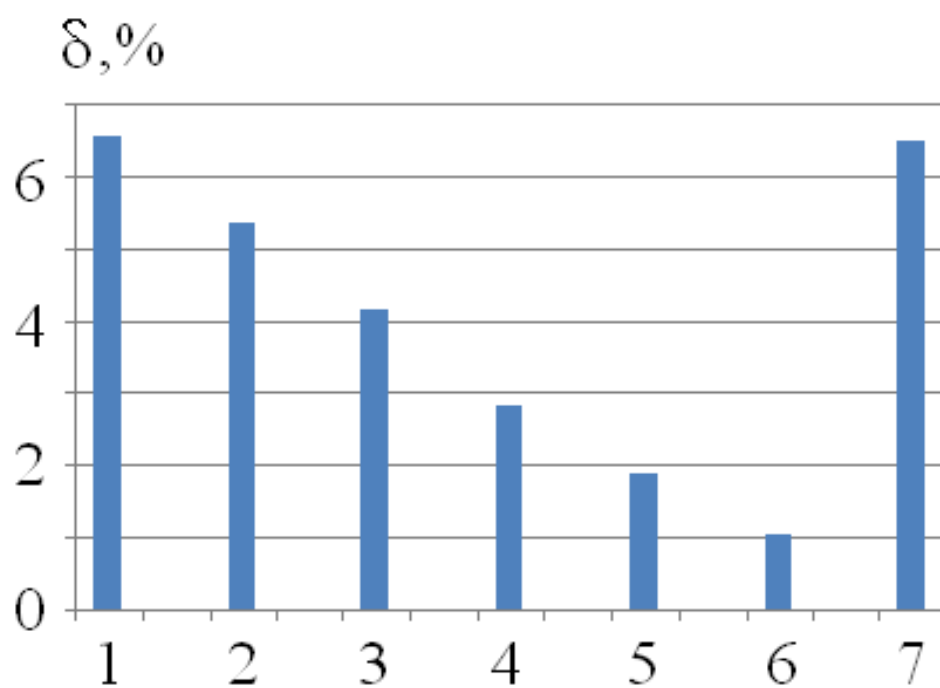


Рис.6 Погрешности определения концентрации метанола. $a = 0,2$.

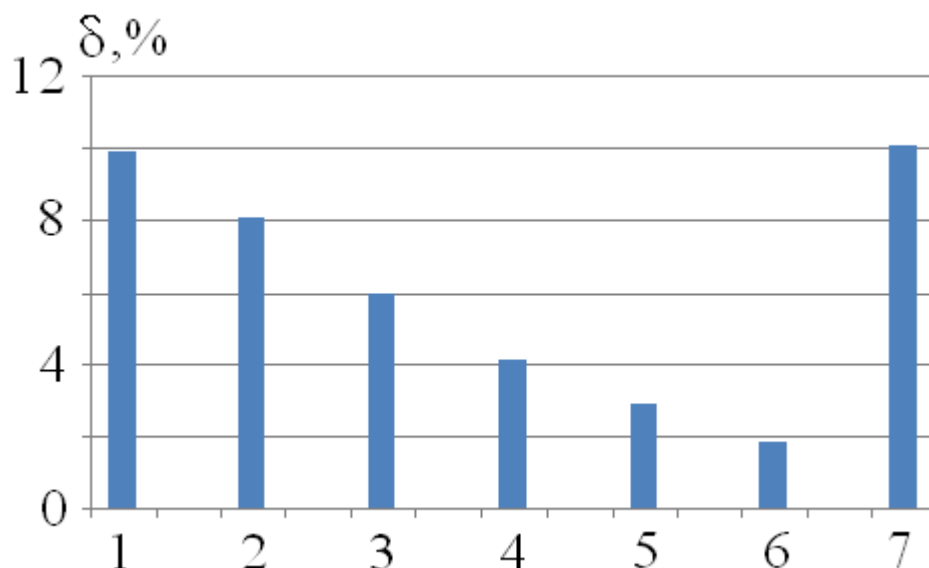


Рис.7 Погрешности определения концентрации метанола. $\alpha = 0,3$.

На рисунках 2-7: столбец 1 соответствует восстановленной концентрации для параметра p функции невязки (5) равного 2; столбец 2- 1,8; столбец 3 - 1,6; столбец 4 –1,4; столбец 5 – 1,2; столбец 6 – 1; столбец 7 - соответствует концентрации определенной как среднее арифметическое из результатов 6 измерений.

Из рисунков видно, что, несмотря на сильное отличие величины погрешностей этилена и метанола (погрешности восстановления концентрации этилена гораздо больше – это связано с очень малым содержанием его в смеси), характер зависимости от величины параметра p функции невязки одинаков – у уменьшением параметра p от 2 до 1 погрешности определения концентраций газов уменьшаются (в 2, 4, а то и 6 раз). Видно также, что определение концентраций газов как среднее арифметическое из результатов одиночных измерений дает такую же погрешности как метод наименьших квадратов (3) (т.е. использование функции невязки (5) для $p=2$).

Заключение

Таким образом, в статье рассмотрена задача количественного анализа состава многокомпонентных стационарных газовых смесей, для которых существует возможность проведения серии измерений на каждой длине волны. Показано, что использование робастной оценки, основанной на замене квадратичной формы в функции невязки на L_p -форму, позволяет в условиях выбросов изменяемого сигнала уменьшить в несколько раз погрешность определения концентраций компонент газовых смесей.

Список литературы

1. Лазерная оптико-акустическая спектроскопия межмолекулярных взаимодействий в газах / Ю.Н. Пономарев, Б.Г. Агеев, М.В. Зигрист, В.А. Капитанов, Д. Куртуа, О.Ю. Никифорова. Томск: МГП «РАСКО», 2000. 200 с.
2. Основы количественного лазерного анализа / В.И. Козинцев, М.Л. Белов, В.А. Городничев, Ю.В. Федотов. М.: Из-во МГТУ, 2006. 464 с.
3. Козинцев В.И., Белов М.Л., Городничев В.А., Федотов Ю.В. Лазерный оптико-акустический анализ многокомпонентных газовых смесей. М.: Изд-во МГТУ им. Н. Э. Баумана, 2003. 352 с.
4. Существующие подходы к решению некорректных задач. Режим доступа: <http://pelbook.narod.ru/2004/02.htm> (дата обращения 20.11.2014).
5. Engl H.W., Hanke M., Neubauer A. Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.
6. Neubauer A. Inverse and Ill-Posed Problems. 2008. Режим доступа: <http://www.impan.pl/BC/Arch/2008/08NeubauerCourse.pdf> (дата обращения 20.11.2014).
7. Еременко Л.Н., Козинцев В.И., Городничев В.А. Метод байесовских оценок в задаче лазерного газоанализа // Известия вузов. Физика. 2008. №9. С. 29-35.
8. Городничев В.А. Методы восстановления концентраций газов для задачи лидарного контроля газовых загрязнителей атмосферы // Биомедицинская радиоэлектроника. 2003. №9. С.52-56.
9. Катаев М.Ю., Бойченко И.В. Программное и методическое обеспечение задач лидарного зондирования атмосферы. Томск:СТТ, 2007. 246 с.
10. Крянев А.В., Лукин Г.В.. Математические методы обработки неопределенных данных. М.: Физматлит, 2003. 212 с.
11. Крянев А.В. Лукин Г.В. Метрический анализ и обработка данных. М.:Физматлит, 2010. 280 с.
12. Грауэр Л.В., Архипова О.А. Робастные регрессионные модели. Режим доступа: http://compscicenter.ru/media/slides/math_stat_2014_spring/2014_04_18_math_stat_2014_spring_1.pdf. (дата обращения 20.11.2014).