

При использовании детерминированных методов регуляризации для системы уравнений (1) регуляризованное решение определяется выражением [1, 11, 12]:

$$\mathbf{x}_\alpha = (W^T W + \alpha E)^{-1} W^T \mathbf{y}, \quad (3)$$

где α – параметр регуляризации; \mathbf{x}_α – регуляризованное решение уравнения (1); E – единичная матрица; $(W^T W + \alpha E)^{-1}$ – матрица, обратная матрице $W^T W + \alpha E$. Значок «т» означает транспонирование матрицы.

Основная сложность при практическом построении регуляризованных решений заключается в выборе параметра регуляризации α . Для получения решения использовалось несколько способов выбора параметра регуляризации.

Пусть вместо точной правой части уравнения (1) у задано ее значение $\tilde{\mathbf{y}}$, удовлетворяющее условию

$$\|\mathbf{y} - \tilde{\mathbf{y}}\| \leq \delta,$$

где $\|\mathbf{y}\|$ – норма вектора.

Введем функцию [11, 14]:

$$r(\alpha) = \rho(W\mathbf{x}_\alpha, \tilde{\mathbf{y}}). \quad (4)$$

Численное решение уравнения

$$r(\alpha) = \delta^2 \quad (5)$$

дает значение параметра регуляризации, найденное на основе принципа невязки.

По второму способу квазиоптимальный параметр регуляризации находился из условия [11]:

$$\inf_\alpha \|\eta_\alpha\|^2, \quad (6)$$

где $\eta_\alpha = \alpha \frac{dx_\alpha}{d\alpha}$.

В качестве третьего способа выбора параметра использовалась модификация метода невязки. Шумы измерения в различных спектральных каналах могут сильно отличаться между собой. Поэтому при восстановлении концентрации каждого газа использовался свой параметр регуляризации, который находился в два этапа. На первом этапе для каждого газа параметр регуляризации α_i получили из уравнения (5), записанного для спектрального канала, в котором данный газовый компонент имеет пик поглощения. На втором этапе в окрестности найденного значения α_i (в интервале между α_i и α , найденным из (5) для всех спектральных каналов) из условия (6) получалось окончательное значение параметра регуляризации.

В четвертом способе использовалось более строгое выражение для квазиоптимального параметра регуляризации, когда α получается из условия [11]:

$$\inf_\alpha \sup_y \|\eta_\alpha\|^2, \quad (7)$$

где \sup_y находится по набору реализаций (вычисления проводились по 10 реализациям) правой части (1).

Для выбора параметра α по пятому способу использовался метод отношений [11], для которого в качестве параметра регуляризации принималось наименьшее значение из множества $\{\alpha_i\}$ значений α , доставляющих минимум функционалу:

$$r_{rel}(\alpha) = \frac{r_1(\alpha)}{r(\alpha)}, \quad (8)$$

где $r_1(\alpha) = \|W\eta_\alpha - (W\mathbf{x}_\alpha - \tilde{\mathbf{y}})\|^2$.

В шестом способе выбор α проводился по минимуму невязки при использовании независимой информации. Параметр α выбирается из условия [13]:

$$\min_\alpha \|\tilde{\mathbf{y}}_\alpha - \tilde{\mathbf{y}}_\delta\|, \quad (9)$$

где $\tilde{\mathbf{y}}_\alpha = W\mathbf{x}_\alpha$; $\tilde{\mathbf{y}}_\delta$ – вектор правых частей уравнения (1), полученный при дополнительных независимых измерениях.

При наличии серии измерений (набора реализаций) из одной реализации находится \mathbf{x}_α и $\tilde{\mathbf{y}}_\alpha$, а другие могут быть использованы в качестве дополнительной независимой информации [13]. Для серии из 10 опытов мы использовали результаты одного опыта для нахождения \mathbf{x}_α и $\tilde{\mathbf{y}}_\alpha$, а результаты оставшихся опытов – для нахождения среднего (по этим опытам) значения $\tilde{\mathbf{y}}_\delta$.

В седьмом и восьмом способах использовался статистический метод выбора параметра регуляризации из критерия оптимальности регуляризованного решения. Статистически регуляризованное решение \mathbf{x}_α имеет вид [11, 14]:

$$\mathbf{x}_\alpha = (\alpha G + W^T V_\xi^{-1} W)^{-1} W^T V_\xi^{-1} \tilde{\mathbf{y}}, \quad (10)$$

где G – неотрицательно определенная симметричная матрица.

В качестве параметра α выбирается такое его значение α_W , которое статистически не противоречит оптимальности регуляризованного решения [11]:

$$V_e(\alpha_W) = V_\xi P^T(\alpha_W), \quad (11)$$

где V_e – матрица вторых моментов вектора невязки; $V_e = M[\mathbf{e} \times \mathbf{e}^T]$; \mathbf{e} – вектор невязки, $\mathbf{e} = \tilde{\mathbf{y}} -$

$W\mathbf{x}_\alpha$; $\mathbf{x}_\alpha = T \tilde{\mathbf{y}}$; T – регуляризирующий оператор; M означает математическое ожидание; P – оператор невязки, $P = E - WT$; E – единичная матрица.

Гипотеза об оптимальности решения будет отвергаться, если невыполнение тождества (11) обусловлено не случайными ошибками, возникающими из-за конечного объема выборки, а систематическими ошибками, связанными с неоптимальностью принятого значения параметра регуляризации. Следовательно, если при $\alpha = \alpha_W$ принимается гипотеза (11), то значение α_W может рассматриваться как оценка оптимального значения параметра регуляризации [11].

Для проверки гипотезы (11) вводится статистика [11]:

$$r_W(\alpha) = \mathbf{e}^T(\alpha) K^{-1}(\alpha) \mathbf{e}(\alpha), \quad (12)$$

где

$$K(\alpha) = V_\xi P^T(\alpha).$$

Для значений α , которые не противоречат (12), свойства статистики $r_W(\alpha)$ позволяют аппроксимировать распределение $r_W(\alpha)$ χ^2 -распределением с M степенями свободы [11]. Тогда проверка гипотезы (11) сводится к выяснению, подчиняется ли $r_W(\alpha)$ χ^2 -распределению с M степенями свободы. Проверка этой гипотезы осуществляется следующим способом. Построим интервал

$$\theta_M(\beta) = \left[\vartheta_M\left(\frac{\beta}{2}\right), \vartheta_M\left(1 - \frac{\beta}{2}\right) \right], \quad (13)$$

где $\vartheta_M(\beta/2)$ – квантиль χ^2 -распределения уровня $\beta/2$; β – вероятность ошибки первого рода (вероятность отвергнуть гипотезу согласно критерию, когда гипотеза верна).

Если величина $r_W(\alpha)$ попадает в интервал $\theta_M(\beta)$, то гипотеза (11) может быть принята с вероятностью ошибки первого рода, равной β , и в качестве значения параметра регуляризации принимают такое значение α , при котором $r_W(\alpha) \in \theta_M(\beta)$.

Если $M \leq 30$, то значения квантилей приводят в таблице вероятностей χ^2 -распределения [15]. При $M > 30$ χ^2 -распределение хорошо аппроксимируется нормальным распределением с математическим ожиданием M и дисперсией $2M$.

Таким образом, решение (10) и использование процедуры нахождения параметра регуляризации дают искомое значение вектора \mathbf{x}_α . Разница между вектором \mathbf{x}_α и истинным решением \mathbf{x} может быть представлена в виде [11, 14]:

$$\mathbf{x}_\alpha - \mathbf{x} = \mathbf{b}_\alpha + \xi_\alpha.$$

Здесь случайный вектор ξ_α имеет нулевое среднее значение и определяет влияние шума измерения на случайную составляющую ошибки регуляризованного решения. Вектор \mathbf{b}_α характеризует систематическую ошибку (смещение) регуляризованного решения и находится из следующего уравнения [11, 14]:

$$W^T V_\xi^{-1} W + \alpha G) \mathbf{b}_\alpha = -\alpha G \mathbf{x}. \quad (14)$$

Из-за незнания истинного решения \mathbf{x} , входящего в правую часть уравнения (14), возможно только определение оценки $\hat{\mathbf{b}}_\alpha$ вектора \mathbf{b}_α при подстановке в (14) вместо \mathbf{x} регуляризованного решения \mathbf{x}_α . Нахождение вектора $\hat{\mathbf{b}}_\alpha$ позволяет ввести новое решение $\tilde{\mathbf{x}}_\alpha$, которое может частично компенсировать смещение решения \mathbf{x}_α [11, 14]:

$$\tilde{\mathbf{x}}_\alpha = \mathbf{x}_\alpha - \hat{\mathbf{b}}_\alpha. \quad (15)$$

При расчетах полагалось, что шум измерения некоррелирован и корреляционная матрица шума имеет вид $V_\xi = \sigma_\xi^2 E$ (где σ_ξ^2 – дисперсия шума,

E – единичная матрица). В качестве матрицы G использовалась матрица V_ξ^{-1} .

Для проверки работоспособности алгоритмов обработки и оценки точности восстановления концентраций газов в многокомпонентных смесях проводились математическое моделирование и обработка данных измерений ЛОАГ. Математическое моделирование работы алгоритмов восстановления концентраций газов проводилось для многокомпонентных смесей с числом компонентов 3–10. Результаты математического моделирования показывают, что в отсутствие шума измерения концентрации газовых компонентов точно восстанавливаются как стандартными методами решения системы линейных алгебраических уравнений, так и с использованием процедур построения регуляризованного решения. С увеличением шума измерения ошибки восстановления возрастают.

Для многокомпонентных смесей (с числом компонентов больше 5) процедуры обработки, основанные на построении регуляризованного решения, обеспечивают гораздо более низкий уровень ошибок восстановления концентраций газов. Причем значительное преимущество этих методов проявляется как при одиночных измерениях, так и при использовании для восстановления концентраций газов предварительно усредненных (по 10 опытам) значений сигналов.

Обработка данных измерений ЛОАГ проводилась для газовых смесей с числом компонентов от 3 до 6. Результаты обработки показывают, что для трех- и четырехкомпонентных смесей ошибки восстановления концентраций газов небольшие и использование регуляризирующих процедур обработки не дает повышения точности. Для пятикомпонентной смеси наблюдается небольшое повышение точности, а для шестикомпонентной смеси точность восстановления значительно увеличивается при использовании регуляризирующих процедур обработки.

На рис. 1, 2 показаны результаты восстановления концентраций C газов (этилена – рис. 1, изопропанола – рис. 2) в шестикомпонентной газовой

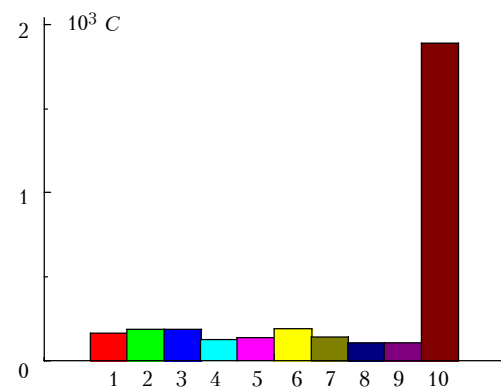


Рис. 1. Результаты восстановления концентрации этилена из данных измерений

смеси этилен – углекислый газ – аммиак – метанол – этанол – изопропанол в случае, когда при обра-

ботке использовались средние (по 10 опытам) значения сигналов газоанализатора. Измерения проводились для спектральных каналов 10,532; 10,591; 9,329; 9,282; 9,317; 9,292; 9,658; 9,676; 9,753; 9,733; 10,303; 10,182 мкм.

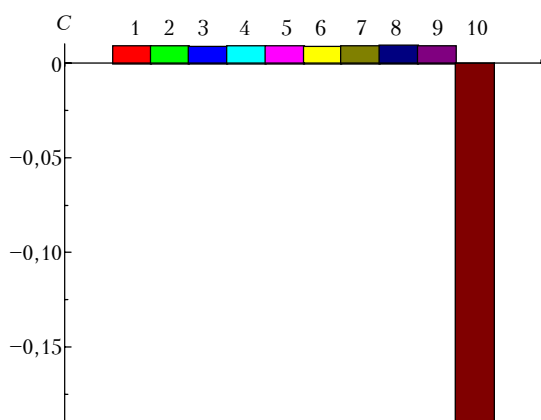


Рис. 2. Результаты восстановления концентрации изопропанола из данных измерений

Методика измерений и лазерный оптико-акустический газоанализатор описаны в [8]. Дисперсии шума измерения в спектральных каналах оценивались по результатам серии тестовых измерений. При обработке задавались следующие относительные среднеквадратические значения шума измерения: 0,01734 (10,532 мкм), 0,03296 (9,292 мкм), 0,02229 (9,676 мкм), 0,00905 (9,733 мкм), 0,03537 (10,303 мкм).

Столбец 1 дает действительное значение концентрации газового компонента, определенное по парциальному давлению при заполнении кюветы. Выбор параметра регуляризации осуществлялся на основе принципа невязки (столбец 3), модифицированного принципа невязки (столбец 2), выбора квазиоптимального параметра регуляризации по формулам (6) и (7) (столбцы 4 и 5), метода отношений (столбец 6), принципа невязки при использовании дополнительной независимой информации (столбец 7), метода статистической регуляризации (столбцы 8, 9; регуляризованное решение для столбца 8 находилось из выражения (10), для столбца 9 – из (15)). Столбец 10 соответствует результатам восстановления концентраций газовых компонентов стандартным методом решения системы алгебраических линейных уравнений (для решения был использован метод Гауссова исключения).

Из рис. 1, 2 видно, что даже при использовании предварительного усреднения по 10 опытам ошибки восстановления концентраций газов значительно уменьшаются при использовании процедур обработки, основанных на методах регуляризации. Концентрации газов, восстановленные стандартным методом решения системы линейных алгебраических уравнений, могут не только сильно отличаться от действительных концентраций, но и принимать отрицательные значения.

Для сравнения различных алгоритмов восстановления концентраций газов, основанных на де-

терминированных и статистических методах регуляризации, на рис. 3 показаны относительные ошибки восстановления концентраций δ (модуль разности действительного и восстановленного значения концентрации, деленный на действительное значение концентрации) этилена (I), аммиака (II),

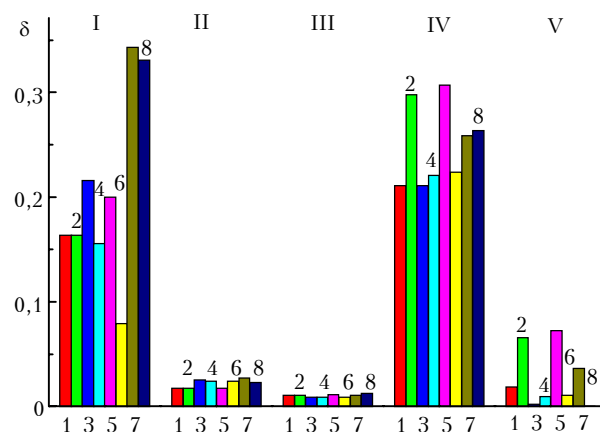


Рис. 3. Относительные ошибки восстановления концентраций компонентов шестикомпонентной газовой смеси

метанола (III), этанола (IV) и изопропанола (V) (для той же газовой смеси, что и данные рис. 1, 2). Выбор параметра регуляризации осуществлялся на основе принципа невязки (столбцы 2), модифицированного принципа невязки (столбцы 1), выбора квазиоптимального параметра регуляризации по формулам (6) и (7) (столбцы 3 и 4), метода отношений (столбцы 5), принципа невязки при использовании дополнительной независимой информации (столбцы 6), метода статистической регуляризации (столбцы 7, 8; регуляризованное решение для столбца 7 находилось из выражения (10), для столбца 8 – из (15)). Ошибки восстановления концентраций стандартным методом решения системы алгебраических линейных уравнений на рис. 3 не показаны, так как они очень велики.

Результаты математического моделирования и обработки данных измерений показывают, что использование регуляризирующих процедур обработки для многокомпонентной смеси обеспечивает в большинстве случаев достаточно низкий уровень ошибок восстановления концентраций газов при использовании как статистических, так и детерминированных методов регуляризации. Значительное преимущество (перед стандартными методами) методов, основанных на построении регуляризованных решений, проявляется как при одиночных измерениях, так и при использовании предварительно усредненных значений сигналов. Различные способы выбора параметра регуляризации для разных газов могут давать как близкие значения восстановленных концентраций, так и существенно различающиеся между собой. Точность восстановления концентраций также существенно различается для разных газовых компонентов смеси. В большинстве случаев наименьшие ошибки восстановления обеспечивают выбор квази-

оптимального параметра регуляризации, модифицированный метод невязки и метод невязки при использовании дополнительной независимой информации.

1. Макушкин Ю.С., Мицель А.А., Хмельницкий Г.С. // Журнал прикладной спектроскопии. 1981. Т. 35. Вып. 5. С. 785–790.
2. Межерис Р. Лазерное дистанционное зондирование. М.: Мир, 1987. 550 с.
3. Иванов С.В., Панченко В.Я., Разумихина Т.Б. // Оптика атмосферы и океана. 1993. Т. 6. № 8. С. 1023–1029.
4. Пономарев Ю.Н. // Оптика атмосферы и океана. 1995. Т. 8. № 1–2. С. 224–241.
5. Зигрист М., Катаев М.Ю., Мицель А.А. и др. // Оптика атмосферы и океана. 1994. Т. 7. № 11–12. С. 1471–1477.
6. Катаев М.Ю., Мицель А.А. // Автометрия. 1985. № 4. С. 15–20.
7. Катаев М.Ю., Мицель А.А., Тарасова С.Р. // Оптика атмосферы. 1990. Т. 3. № 8. С. 832–841.
8. Козинцев В.И. // Оптика атмосферы и океана. 1996. Т. 9. № 8. С. 1087–1091.
9. Козинцев В.И. // Оптика атмосферы и океана. 1996. Т. 9. № 10. С. 1373–1378.
10. Авдиенко В.В., Белов М.Л., Городничев В.А., Козинцев В.И., Добрица Д.Б. // Журнал прикладной спектроскопии. 1996. Т. 63, № 5. С. 755–759.
11. Воскобойников Ю.Э., Преображенский Н.Г., Седельников А.Н. Математическая обработка эксперимента в молекулярной газодинамике. Новосибирск: Наука, 1984. 238 с.
12. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979. 288 с.
13. Кочетов И.И. // Журнал вычислительной математики и математической физики. 1976. Т. 16. № 2. С. 499–503.
14. Воскобойников Ю.Э., Мицель А.А. // Автометрия. 1982. № 2. С. 67–72.
15. Корн Г., Корн Т. Справочник по математике. М.: Наука, 1974. 831 с.

M.L. Belov, V.A. Gorodnichenov, D.B. Dobritsa, V.I. Kozintsev. Comparative analysis of methods for reconstructing gas densities in multicomponent gaseous mixtures from data of laser photoacoustic gas-analyser.

The comparative analysis of algorithms for reconstructing gas densities is considered for the problem of gas-analysis of multicomponent gas mixtures. Processing of signals, based on methods of regularization of solving incorrect mathematical problems, is described for laser method of the differential absorption. The results of experimental investigations are presented. It is shown that the use of regularized solutions gives rather low level of reconstructing errors of gas densities for different methods of choice of the regularization constant. In most cases the least errors of the restoration is provided by a choice of quasioptimum regularization constant, modified method of discrepancy, and method of discrepancy at use of additional independent information.