

В.Ф. Рапута, К.П. Куценогий, Г.Т. Яковенко

ОПТИМИЗАЦИЯ ИЗМЕРЕНИЙ СПЕКТРА СУБМИКРОННЫХ
АТМОСФЕРНЫХ АЭРОЗОЛЕЙ

ИВМ и МГ СО РАН, ИХКГ СО РАН, г. Новосибирск

Поступила в редакцию 3.03.99 г.

Принята к печати 15.03.99 г.

Для многомодальных функций распределений спектра субмикронного аэрозоля предложен метод построения дискретных оптимальных схем измерений. На его основе проведено численное моделирование насыщенных планов наблюдения в случае двухмодального распределения. Определены некоторые закономерности формирования оптимальных сеточных наборов. Обсуждаются причины возможной неустойчивости решения обратной задачи оценивания параметров функции распределения спектра размеров аэрозоля.

Введение

Точность восстановления параметров функций распределения спектра атмосферных аэрозолей существенно зависит от выбора той или иной схемы проведения измерений. Вопросам оптимального планирования эксперимента применительно к сетчатой диффузионной батарее посвящены работы [1, 2].

На основе предложенной модели оптимизации измерений в случае одномодальной функции распределения логнормального вида получены непрерывные оптимальные планы наблюдений и их аналитические аппроксимации. Количество сеток в полученных планах в основном зависит от среднегеометрического радиуса частиц и представляет собой монотонно возрастающие функции от этого параметра.

В статье обсуждаются возможности предложенного выше алгоритма построения оптимальных планов для случая многомодальных функций распределения спектра размеров атмосферных аэрозолей.

Существенно возрастающая размерность оптимизационной задачи планирования наблюдений вызывает затруднения при непосредственном округлении промоделированных непрерывных планов. В связи с этим процедура планирования разбивается на два этапа: строится непрерывный оптимальный план [3, 4], а затем проводится его округление до дискретного путем перебора по вычисленному спектру уже полученного непрерывного плана.

Анализ результатов численного моделирования для бимодальных функций распределения позволил выявить дополнительные качественные и количественные закономерности в распределении сеточных наборов дискретных пятиточечных планов.

Планирование измерений

Проскок аэрозоля $p(y)$ через y сеток, измеряемый с помощью сетчатой диффузионной батареи, можно описать с помощью следующего выражения [5]:

$$p(y) = \int_{-\infty}^{\infty} K(y, x) f(x) dx, \quad c \leq y \leq d, \quad (1)$$

где $f(x)$ – функция распределения размеров частиц атмосферного аэрозоля; $x = \log r$, r – радиус частиц; $K(y, x)$ – аппаратная функция, имеющая вид

$$K(y, x) = \exp(-y t(x)), \quad t(x) = c (B_1 \cdot 10^{-x} + B_2 \cdot 10^{-2x})^{2/3}, \quad (2)$$

где $B_1 = 8,53 \cdot 10^{-9}$, $B_2 = 1,27 \cdot 10^{-6}$; c – константа, зависящая от скорости потока и параметров сеток.

Соотношение (1) является линейным интегральным уравнением первого рода относительно функции $f(x)$, и нахождение его решения по измерениям функции $p(y)$ относится к классу некорректных обратных задач. Применение математических методов планирования эксперимента при самых общих предположениях относительно вида функции распределения $f(x)$ весьма проблематично [3]. Поэтому необходима дополнительная априорная информация о свойствах функции $f(x)$.

Эффективность планирования наблюдений существенно возрастает в случае задания $f(x)$ с точностью до набора параметров. В связи с этим в дальнейшем ограничимся представлением $f(x)$ в виде суммы логарифмически нормальных распределений

$$f(x, \theta) = \sum_{i=1}^m q_i f_i(x, r_{50}^i, \sigma_g^i) \quad (3)$$

с условием нормировки

$$\sum_{i=1}^m q_i = 1. \quad (4)$$

Здесь

$$f_i(x) = (2\pi \log^2 \sigma_g^i)^{-1/2} \exp\left\{-\frac{(x - \log r_{50}^i)^2}{2 \log^2 \sigma_g^i}\right\}, \quad i = \overline{1, m}, \quad (5)$$

где r_{50}^i и σ_g^i – среднегеометрический радиус частиц и стандартное геометрическое отклонение для i -й моды атмосферного аэрозоля; $\theta = (r_{50}^1, \sigma_g^1, \dots, r_{50}^m, \sigma_g^m)$ – вектор неизвестных параметров.

Оценки вектора θ могут быть получены на основе измерений функции $p(y)$ в конечном наборе точек y_k , $k = \overline{1, n}$, $n \geq 3m - 1$, с помощью, например, метода наименьших квадратов. Тогда для численного нахождения искомого вектора θ можно применить процедуру градиентного спуска [2, 3]:

$$\theta_{j+1} = \theta_j + M^{-1}(\theta_j) \cdot Y(\theta_j), \quad (6)$$

где

$$M(\theta) = \sum_{k=1}^n \sigma_k^{-2} \varphi(y_k, \theta) \cdot \varphi^T(y_k, \theta); \quad (7)$$

$$Y(\theta) = \sum_{k=1}^n \sigma_k^{-2} \varphi(y_k, \theta) \lambda_k;$$

$$\varphi(y_k, \theta) = \nabla_{\theta} p; \quad (8)$$

λ_k – измеренные значения проскока; σ_k^{-2} – дисперсия измерений; T – транспонирование.

Из (6) вытекает, что точность оценивания зависит от характеристик информационной матрицы (7), которые, в свою очередь, зависят от выбора плана измерений точек y_k и их весов s_k , $k = \overline{1, n}$. Если ограничиться рассмотрением планов измерений, максимизирующих определитель информационной матрицы M , то численное построение D -оптимального плана можно провести с помощью поэтапной процедуры последовательного анализа и планирования измерений, описанной в [2, 3].

В случае одномодальной функции распределения эта процедура позволила численно промоделировать локально-оптимальные планы для допустимых значений искомых параметров. Использование процедуры планирования при многомодальной функции распределения наталкивается на ряд трудностей. Одна из них связана со сложностью округления полученных непрерывных планов до насыщенных дискретных, число точек в которых равно количеству неизвестных параметров. В связи с этим построение насыщенного дискретного плана рекомендуется проводить на спектре непрерывного плана следующим образом [4]:

1. Задается произвольная выборка объема $N = k$ из спектра непрерывного D -оптимального плана.

2. Каждая точка этой выборки последовательно заменяется на одну из оставшихся $n - k$ точек спектра. После каждой замены вычисляется определитель информационной матрицы полученного плана и окончательно в план включается точка, дающая максимальное значение определителя.

3. Действия п. 2 повторяются с выборкой, полученной на предыдущем этапе, до тех пор, пока происходит увеличение определителя информационной матрицы.

4. Данные операции выполняются с различными начальными выборками из спектра непрерывного D -оптимального плана, и выбирается план, в наибольшей степени удовлетворяющий критерию D -оптимальности.

Численные эксперименты

В силу большой трудоемкости вычислений ограничимся рассмотрением двухмодальных распределений. С учетом условия нормировки (4) число неизвестных параметров равно пяти. В этом случае, так же как для одномодального распределения, удобнее перейти в (3) к другому набору неизвестных параметров:

$$\theta_1 = q_1/\sqrt{2p}, \quad \theta_2 = 1/\log \sigma_g^{(1)}, \quad \theta_3 = \log r_{50}^{(1)}/\log \sigma_g^{(1)},$$

$$\theta_4 = 1/\log \sigma_g^{(2)}, \quad \theta_5 = \log r_{50}^{(2)}/\log \sigma_g^{(2)}. \quad (9)$$

Тогда с учетом обозначений (8), (9) вектор базисных функций $\varphi(y, \theta) = \nabla_{\theta} p$ приобретает менее громоздкий вид:

$$\varphi_1 = \int_{-\infty}^{\infty} (\theta_2 e^A - \theta_4 e^B) dx,$$

$$\varphi_2 = \int_{-\infty}^{\infty} \theta_1 (1 - \theta_2^2 x^2 + \theta_2 \theta_3 x) e^A dx,$$

$$\varphi_3 = \int_{-\infty}^{\infty} \theta_1 (\theta_2^2 x - \theta_2 \theta_3) e^A dx, \quad (10)$$

$$\varphi_4 = \int_{-\infty}^{\infty} (1 - \theta_1) (1 + \theta_4 \theta_5 x - \theta_4^2 x^2) e^B dx,$$

$$\varphi_5 = \int_{-\infty}^{\infty} (1 - \theta_1) (\theta_4^2 x - \theta_4 \theta_5) e^B dx,$$

где

$$A(x, y, \theta) = -y f(x) - \frac{1}{2} (x\theta_2 - \theta_3)^2;$$

$$B(x, y, \theta) = -y f(x) - \frac{1}{2} (x\theta_4 - \theta_5)^2.$$

Свойства функции $\varphi_1, \dots, \varphi_5$ могут быть исследованы с помощью численного моделирования. На рис. 1 приведены графики базисных функций (10) для набора возможных значений параметров логнормальных распределений. Анализ рис. 1 позволяет сделать определенные выводы о природе неустойчивости оценивания параметров функции распределения. Причина возможной неустойчивости решения обратной задачи состоит в подобном характере поведения кривых $\varphi_1, \dots, \varphi_5$, которые порождают плохую обусловленность информационной матрицы (7). Одним из способов уменьшения влияния вычислительной неустойчивости является использование при проведении измерений оптимальных сеточных наборов.

На рис. 2 представлены примеры расчета оптимальных схем измерений сначала с помощью алгоритма непрерывного планирования эксперимента, а потом наилучшее распределение числа сеток производилось с использованием процедуры (п. 1–4) дискретного перебора сеточных наборов по спектру полученного непрерывного плана.

Численное моделирование оптимальных схем наблюдений для бимодальных функций распределения позволяет выявить дополнительные качественные и количественные закономерности в положениях точек планов. На рис. 2, б происходит скачкообразный переход точек локально-оптимального плана из одного состояния в другое, что связано со взаимным расположением среднегеометрических радиусов частиц $r_{50}^{(1)}$ и $r_{50}^{(2)}$. В распределении точек найденных планов наблюдается степенной рост числа сеток в соответствующих сеточных наборах.

В связи с этим следует отметить, что численно промоделированное распределение числа сеток в сеточных наборах

согласуется с рекомендуемыми распределениями, полученными эмпирическим способом.

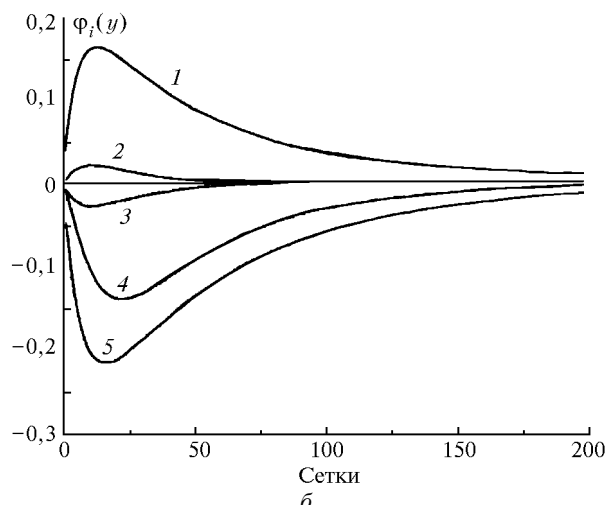
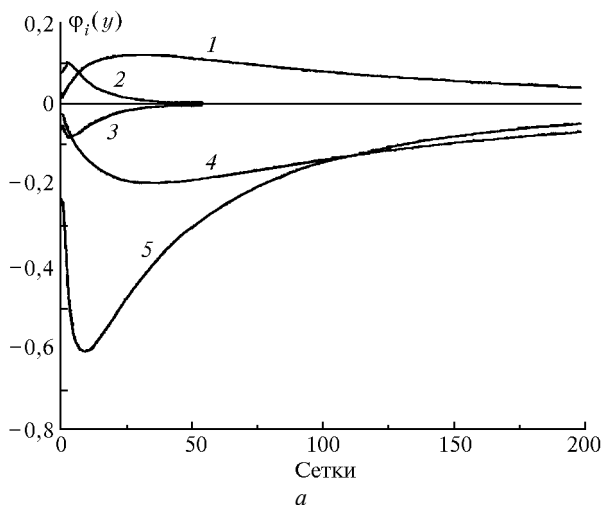


Рис. 1. Базисные функции $\Phi_i(y, \theta)$, $i = \overline{1,5}$, при $a - q_1 = 0,5$, $r_{50}^{(1)} = 6$ нм, $\sigma_g^{(1)} = 1,7$, $r_{50}^{(2)} = 40$ нм, $\sigma_g^{(2)} = 2$, 1; б) $q_1 = 0,27$, $r_{50}^{(1)} = 15$ нм, $\sigma_g^{(1)} = 1,2$, $r_{50}^{(2)} = 20$ нм, $\sigma_g^{(2)} = 1,86$

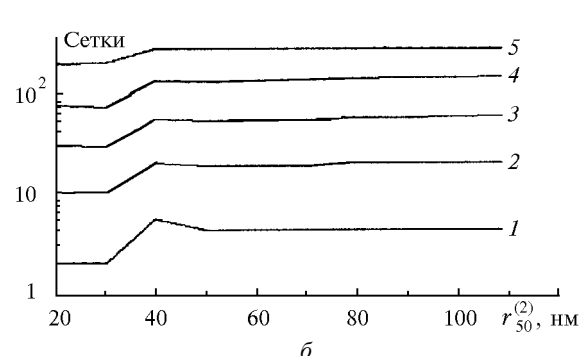
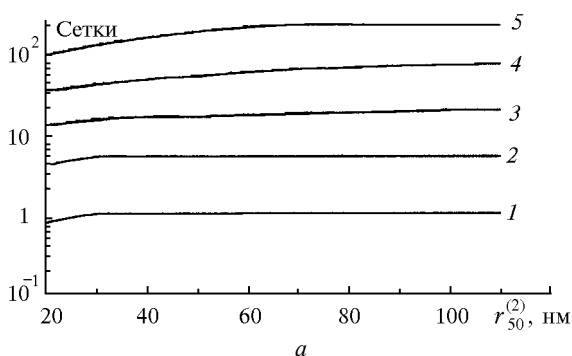


Рис. 2. Дискретные локально-оптимальные планы наблюдений для $\sigma_g^{(1)} = 1,4$, $\sigma_g^{(2)} = 1,86$: а) $r_{50}^{(1)} = 5$ нм; б) $r_{50}^{(1)} = 15$ нм

Заключение

Проведенное численное построение непрерывных и дискретных планов измерений применительно к сетчатой диффузионной батарее позволило выявить ряд дополнительных особенностей их размещения для случая двухмодальной функции распределения по сравнению с одномодальной. Численный анализ поведения базисных функций позволяет выявить возможные причины плохой обусловленности рассматриваемых обратных задач.

Существенно возросли требования к вычислительным ресурсам. С одной стороны, то или иное взаимное расположение мод в функции распределения может привести к скачкообразному изменению схемы наблюдений. С другой стороны, зависимость положения точек планов от искомых параметров имеет достаточно простой вид. В связи с этим

возникает возможность построения аналитических аппроксимаций численно промоделированных планов измерений в широком диапазоне изменения допустимых значений параметров функций распределения спектра размеров атмосферного аэрозоля.

1. Raputa V.F., Koutsenogii K.P., Jakovenko G.T. // J. Aeros. Sci. 1997.V. 28. P. 259–260.
2. Рапута В.Ф., Куценюгий К.П., Яковенко Г.Т. // Оптика атмосферы и океана. 1998. Т. 11. N 9. С. 1022–1025.
3. Fedorov V. // Theory of Optimal Experiments. New York: Academic Press, 1972. 312 с.
4. Круг Г.К., Сосулин Ю.А., Фатуев В.А. Планирование эксперимента в задачах идентификации и экстраполяции. М.: Наука, 1977. 208 с.
5. Bashurova V.S. et al. // J. Aerosol Sci. 1991. V. 22. P. 373–388.

V.F. Raputa, K.P. Koutsenogii, G.T. Jakovenko. **Optimization of Submicron Aerosol Spectrum Measurements.**

A method for construction of discrete optimal measuring schemes is proposed for multimodal distribution function of submicron aerosol spectrum. It is used to perform numerical modeling of saturated observation plans for two-modal distribution. Some regularities of the formation of optimal screen-type sets have been determined. The reasons are discussed for possible instability of the inverse problem solution of the distribution function parameters of aerosol size spectrum estimation.