

## Улучшенный генетический алгоритм многоволнового лидарного зондирования атмосферного аэрозоля

Г.М. Креков, А.Я. Суханов\*

*Институт оптики атмосферы им. В.Е. Зуева СО РАН  
634021, г. Томск, пл. Академика Зуева, 1*

Поступила в редакцию 18.05.2011 г.

Приводится алгоритм восстановления микрофизических характеристик облачности и аэрозольных образований при использовании лидаров белого света на основе лазеров с фемтосекундной длительностью импульса. Решается прямая задача переноса широкополосного излучения методом Монте-Карло, а затем обратная задача с применением модифицированного генетического алгоритма, позволяющего одновременно восстанавливать физические и оптические характеристики аэрозольных образований.

*Ключевые слова:* лидар, фемтосекундный лазер, аэрозоль, генетический алгоритм; lidar, femtosecond laser, aerosol, genetic algorithm.

### Введение

Появление широкополосных лазерных источников излучения позволяет найти новое их применение в задачах дистанционного зондирования атмосферы, в том числе в задачах определения микрофизических параметров атмосферы, включая облачность. Эффект филаментации и последующее формирование направленного излучения с узким углом расхождения пучка, имеющего широкий спектр длин волн, простирающийся в диапазоне от 0,4 до 2 мкм, охватывают значительную часть линий поглощения атмосферных газов, а также полос, попадающих в окна прозрачности атмосферы. В свою очередь это открывает возможность селекции длин волн для зондирования атмосферных аэрозолей.

В работах [1, 2] проанализированы спектры поглощения атмосферных газов и проведен численный эксперимент по лидарному зондированию слоисто-неоднородной модели облака методом Монте-Карло.

Для восстановления оптических характеристик облачности, точнее вертикальных профилей коэффициентов экстинкции и обратного рассеяния, использовался параметризованный итерационный алгоритм [3], а затем по коэффициентам обратного рассеяния с использованием генетического алгоритма решалась задача восстановления функций распределения частиц по размерам также в зависимости от глубины проникновения в облачность. Данный подход требует двух этапов работы, причем итерационный метод восстановления коэффициентов обратного рассеяния также требует знания

базовых коэффициентов обратного рассеяния в начальной точке итераций и некоторых упрощений на основе априорной информации. Кроме того, применяемый генетический алгоритм в своем исходном варианте — довольно медленная процедура, восстанавливающая данные на основе предыдущего этапа и зависящая от точности итерационного метода.

### 1. Сравнительный анализ эффективности статистических и интеллектуальных методов восстановления микроструктурных параметров

Ранее в [1–4] авторами выполнен сравнительный анализ эффективности обратного метода Монте-Карло и интеллектуальных методов восстановления микроструктурных параметров облачного аэрозоля в схеме мультиволнового лидарного зондирования. Анализ выполнен в рамках замкнутого вычислительного эксперимента. Решение прямой задачи оценки сигналов обратного рассеяния в рамках нестационарной линейной теории переноса осуществлено полуаналитическим методом Монте-Карло, позволяющим достаточно строго учесть граничные условия реального физического эксперимента. В данном случае физическим прототипом лидара служила известная франко-германская лидарная станция «Teramobile», тактико-технические характеристики которой подробно освещены в литературе [5, 6].

Результаты решения прямой задачи, т.е. временные развертки сигналов, полученные для задаваемого набора длин волн, удовлетворяющих требованиям информативности, и априори известной модели облачной атмосферы, использовались для

\* Георгий Михайлович Креков (gm@iao.ru); Александр Яковлевич Суханов (say@iao.ru).

восстановления заложенных в модель оптических и микрофизических параметров атмосферы. Расчеты выполнены для четырех длин волн зондирования: 1,285; 1,557; 1,629 и 2,130 мкм, попадающих, по нашим оценкам, в микроокна прозрачности ближнего ИК-участка спектра. В качестве оптической модели атмосферы выбрана среднециклическая модель континентального аэрозоля [7]. Облачный слой толщиной 100 м задавался на высоте 200 м, микрофизика слоисто-неоднородного облака соответствовала общепринятой тестовой модели «Cloud C1» по Deirmendjan [8].

Результаты, полученные с помощью метода Монте-Карло, показали, что данный метод при больших погрешностях в коэффициентах обратного рассеяния начинает давать сильное расхождение. Кроме того, в реальных экспериментах модальный радиус известен лишь приблизительно и, самое главное, для повышения точности метода требуется использование значительно большего числа длин волн зондирования, что в аппаратурном плане пока нереально. Далее для решения задачи был применен генетический алгоритм. Данный алгоритм позволил повысить точность восстановления функции распределения.

Однако оказывается, что при сильном искажении восстановленных профилей коэффициента обратного рассеяния достаточно сложно выбрать надлежащий критерий останова итераций генетического алгоритма, поскольку получаемые решения ограничены классом гамма-функций. Также определенным недостатком генетического алгоритма является его медленная сходимость, что затрудняет его использование в режиме реального времени. Тем не менее точность восстановления достаточно высока в связи с тем, что модельная функция зависимости коэффициента обратного рассеяния от длины волны ограничена некоторым классом функций.

Этот метод может успешно применяться и при обработке реальных данных, поскольку моно-модальные функции распределения  $N(r)$  в большинстве случаев хорошо аппроксимируются обобщенным гамма-распределением. В случае сложных композитных распределений  $N(r)$  также необходимо привлечение большего количества длин волн зондирования. Однако, как показали численные эксперименты, даже при использовании многих длин волн в ограниченном классе функций  $N(r)$  сохраняется вероятность получить несколько решений задачи.

Метод нейронных сетей оказался [4] несколько более чувствителен к погрешностям входных значений профилей коэффициента обратного рассеяния по сравнению с генетическим алгоритмом. Кроме того, возможности нейронной сети ограничены областью обучающих примеров. Если реальные или тестовые значения выйдут за пределы обучающей выборки, сеть может давать неадекватные решения или несоответствующие реальности, но во многом это преодолевается путем создания наиболее полной информативной обучающей выборки. При этом

нейронная сеть способна решать задачу восстановления в режиме реального времени, а также автоматически учитывает проблему множества решений задачи, усредняя их.

Общим недостатком интеллектуальных методов, использующих в качестве входной информации восстановленные с нарастающей ошибкой вертикальные профили коэффициентов обратного рассеяния  $\beta_\pi(\lambda_i, h)$ ,  $i = 1, \dots, 4$ , как показал проведенный в [3, 4] анализ результатов численного моделирования, остается недостаточная по глубине облака область устойчивых решений.

## 2. Модификация генетического алгоритма

На данном этапе для преодоления отмеченных недостатков представляется целесообразным объединить в едином алгоритме процедуру восстановления оптических и микрофизических характеристик облачности и аэрозольных образований из сигналов, полученных на нескольких оптимально выбранных длинах волн.

Оказывается, что уравнение функционала для решения задачи оптимизации с неизвестными параметрами можно записать в следующем виде:

$$\sum_{i=1}^M \sum_{j=2}^N \sum_{n=1}^{j-1} \left( \frac{S_j(\lambda_i)}{S_n(\lambda_i)} - \frac{\beta_{\pi_j}(\lambda_i)(h_n)^2 G(\lambda_i, h_j)}{\beta_{\pi_n}(\lambda_i)(h_j)^2 G(\lambda_i, h_n)} \right) \times \exp \left( - \sum_{k=n}^{j-1} (\beta_{extk} + \beta_{extk+1}) \Delta h \right)^2, \quad (1)$$

где

$$\beta_{\pi_j}(\lambda_i) = \int_{\eta}^{r_j} f_j(r) K_\pi(r, n(\lambda_i), \lambda_i) dr$$

— коэффициент обратного рассеяния на длине волны  $\lambda_i$  и на высоте  $h_j$  в предположении слоисто-неоднородной структуры;

$$\beta_{extj}(\lambda_i) = \int_{\eta}^{r_j} f_j(r) K_{ext}(r, n(\lambda_i), \lambda_i) dr$$

— коэффициент ослабления;  $S_j(\lambda_i)$  — сигнал на длине волны  $\lambda_i$ , полученный на высоте  $h_j$ ;  $f_j(r)$  — функция распределения частиц по размерам на высоте  $h_j$ ;  $K_\pi(r, n(\lambda_i), \lambda_i)$ ,  $K_{ext}(r, n(\lambda_i), \lambda_i)$  — коэффициенты обратного рассеяния и ослабления для сферической частицы радиусом  $r$  при коэффициенте преломления вещества частицы  $n(\lambda_i)$ ;  $G(\lambda, h)$  — геометрический фактор лидара.

Генетический алгоритм представляет собой, как отмечалось выше, эвристический алгоритм, моделирующий процесс природной эволюции для решения задач оптимизации, он позволяет найти глобальный оптимум. Общий принцип генетического алгоритма заключается в формировании множества решений

называемых популяцией. Из данной популяции выбираются два решения, над которыми проводится операция перекombинации – кроссинговера, что соответствует обычно обмену битовых последовательностей в элементах вектора решений или следующей операции:  $x_i = x_{1i}a + x_{2i}(1-a)$ , где  $a$  – случайное значение от 0 до 1;  $x_i, x_{1i}, x_{2i}$  – элементы нового вектора решения и родительских пар. После этого над новым решением проводится операция мутации, обычно являющаяся изменением битового представления или элемента вектора решения  $x_i$ . Полученное решение включается в популяцию решений.

Операция выбора родительских пар и создание новой особи могут состоять из определенного числа шагов. Через определенное число итераций из популяции удаляются несколько или одна особь, являющиеся наихудшими решениями, или с максимальной вероятностью удаляются решения, являющиеся наихудшими. Наихудшее решение определяется с помощью так называемой функции фитнеса, обычно в роли такой функции выступает оптимизируемый функционал. В каноническом виде такой алгоритм часто приводит к тому, что локальный минимум принимается за глобальный, что соответствует вырождению популяции решений, при этом сходимость достаточно медленная. Было решено модифицировать функцию фитнеса, а также сам генетический алгоритм для ускорения сходимости к искомому решению.

В качестве неизвестного решения принимается функция распределения частиц по размерам, представленная в виде суммы трех гауссовых функций, заданных на каждой высоте:

$$f_i(r) = \sum_{k=1}^3 A_{i,k} \exp\left(-\frac{(r-r_{i,k})^2}{2\sigma_{i,k}^2}\right). \quad (2)$$

В качестве коэффициента преломления среды берется коэффициент преломления воды на четырех длинах волн, на которых производится зондирование. В этом случае коэффициенты оптического взаимодействия  $\beta_{optj}(\lambda_i)$  можно задать как

$$\beta_{optj}(\lambda_i)_j(\lambda_i) = \sum_{l=1}^L \sum_{k=1}^3 A_{i,k} \exp\left(-\frac{(r-r_{i,k})^2}{2\sigma_{i,k}^2}\right) K_{opt}(r_l, \lambda_i) \Delta r, \quad (3)$$

где  $r_l$  – равномерная сетка от 0,001 до 15 мкм, с шагом  $\Delta r$  и числом элементов  $L$ . Функция фитнеса в нашем случае задана в следующем виде:

$$\text{fit}(A_{i,j}, r_k, \sigma_k) = \sum_{i=1}^M \sum_{j=2}^N \sum_{l=1}^{j-1} \left( \frac{(S_{i,j,l} - S_{a_{i,j,l}}) \cdot 2}{S_{i,j,l} + S_{a_{i,j,l}}} + 1 \right)^4, \quad (4)$$

$i \in 1 \dots M = 4; j \in 1 \dots N; k \in 1 \dots 3,$

где

$$S_{i,j,l} = \frac{S_j(\lambda_i)}{S_l(\lambda_i)}, \quad l \in 1 \dots j-1; \quad (5)$$

$$S_{a_{i,j,l}} = \frac{\beta_{\pi j}(\lambda_i)(h_l)^2}{\beta_{\pi l}(\lambda_i)(h_j)^2} \exp\left(-\sum_{p=l}^{j-1} (\beta_{extp}(\lambda_i) + \beta_{extp+1}(\lambda_i)) \Delta h\right). \quad (6)$$

Предполагается, что геометрический фактор лидара  $G(\lambda_i, h)$  незначительно меняется с высотой и, соответственно, отношение геометрических факторов в (1) близко к единице.

Для улучшения решения и нормировки коэффициентов оптического взаимодействия при априорно известных коэффициентах на какой-либо высоте можно добавить к функции фитнеса дополнительный член:

$$\sum_{i=1}^M \left( \frac{2(\beta_{opt0}(\lambda_i) - \beta_{opta}(\lambda_i))}{\beta_{opt0}(\lambda_i) + \beta_{opta}(\lambda_i)} + 1 \right)^4, \quad (7)$$

где  $\beta_{opta}$  – известное значение коэффициента оптического взаимодействия на высоте  $h_0$ .

Высокая степень обуславливает более крутой спуск к искомому решению, нежели использование низких степеней, что влечет возможное вырождение популяции в процессе итераций. В качестве искомого решения рассматриваются параметры  $A_{i,j}, r_k, \sigma_k$  функции фитнеса, суть параметров функций распределения частиц по размерам. Модификация генетического алгоритма заключается в модификации оператора кроссинговера, в данном случае используется следующая процедура. Случайным образом из множества решений выбираются три решения,  $x_1, x_2, x_3$ , среди которых решение  $x_1$  берется как наиболее оптимальное. Новое решение получается путем преобразования

$$x = x_1 - (x_3 - x_2) \text{sign}(\text{fit}(x_3) - \text{fit}(x_2)) \cdot 0,001,$$

что с большой вероятностью является шагом в направлении убывания функции.

Оператор мутации представляет собой следующую стохастическую процедуру для произвольного числа произвольных элементов вектора, сформированного путем кроссинговера, а именно:

$$\tilde{x}_i = x_i a \cdot 2.$$

Кроме того, существуют дополнительно две популяции. В первую заносятся решения, функции фитнеса для которых отличаются от всех имеющих в популяции на 1%, такая особь заменяет собой самую плохо приспособленную. В другую популяцию заносятся особь, имеющая большее расстояние между всеми особями по сравнению с наименее приспособленной особью в данной популяции и лучшую приспособленность. С малой вероятностью для операции кроссинговера выбираются также особи из данных двух популяций. Это необхо-

димо для того, чтобы популяция обладала высоким уровнем разнообразия особей и не вырождалась в локальные минимумы или худшие решения.

### 3. Численный эксперимент по зондированию оптических и микрофизических параметров облачности на основе модифицированного генетического алгоритма

Для решения прямой задачи мультиволнового лидарного зондирования облачного аэрозоля использовалась техника статистического моделирования, идентичная [3, 4]. Как и ранее, начальные и граничные условия решения интегрального уравнения переноса соответствуют известной схеме моностатического лидара белого света [5, 6]. Предполагалось, что источник излучает ультракороткий световой импульс в конусе направлений  $2\pi(1 - \cos\varphi_s)$ , где  $\varphi_s = 0,1$  мрад – полный угол расходимости источника. Эхосигнал регистрируется приемником в угловых конусах  $2\pi(1 - \cos\varphi_d^k)$ , где  $\varphi_d^i$  – набор полных углов приема,  $k = 1, 2, \dots$ . Облачный слой толщиной 0,5 км задавался на высоте 1 км.

Для решения прямой задачи был выполнен предварительный расчет по формулам Ми оптических параметров выбранной модели облака, включая коэффициенты рассеяния, ослабления, обратного рассеяния, вероятности выживания кванта и индикатрисы рассеяния, для длин волн оптического зондирования 1,285; 1,557; 1,629 и 2,130 мкм. Подобный набор длин волн обусловлен, как отмечено выше, необходимостью работы лидара в окнах прозрачности атмосферы и информативностью принятого набора волн относительно искомым параметров облачного аэрозоля. Традиционные оптические характеристики атмосферы, включая облачный слой, задавались в данном примере как кусочно-линейные функции высоты  $h$  (см. ниже рис. 2).

Рассчитанные методом Монте-Карло сигналы обратного рассеяния для выбранной схемы модельного вычислительного эксперимента показаны на рис. 1 для одной длины волны  $\lambda = 1,285$  мкм в зависимости от угла поля зрения предполагаемого детектора  $\varphi_d^i$ . Значения реальных сигналов, включающие в себя помеху многократного рассеяния, используются далее в качестве входных параметров решения обратной задачи, т.е. восстановления дискретных вертикальных профилей априори заданных оптических и микрофизических характеристик облачного аэрозоля на основе предложенного генетического алгоритма. В этом состоит принципиальное отличие нового подхода: промежуточный этап восстановления профилей  $\beta_\pi(\lambda_i, h)$ , наиболее подверженный влиянию помехи многократного рассеяния, устраняется, а входными функционалами в процедуре генетического поиска становятся, как следует из (1), непосредственно значения сигналов  $S_j(\lambda_i)$ .

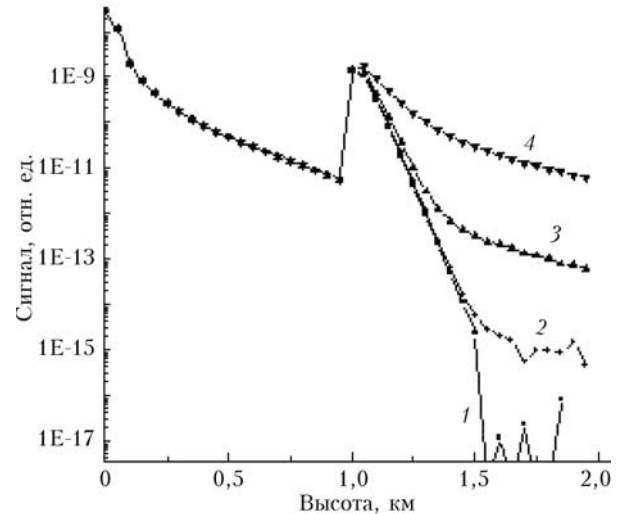


Рис. 1. Рассчитанные методом Монте-Карло лидарные сигналы на  $\lambda = 1,285$  мкм в зависимости от угла поля зрения детектора  $\varphi_d = 0,0001; 0,001; 0,01$  и  $0,1$  рад (кривые 1–4 соответственно)

Более того, использование отношения сигналов в (1) позволяет, как хорошо известно, существенно снизить нагрузку многократного рассеяния. Восстановленные значения коэффициентов ослабления и обратного рассеяния для наиболее реалистичного значения  $\varphi_d = 0,01$  рад приведены на рис. 2 в сопоставлении с модельными значениями. Интегральная оптическая толщина  $\tau$  облака значительна ( $\tau = 5-6$ ), следовательно, результаты восстановления  $\beta_{ext}(\lambda_i, h)$  и  $\beta_\pi(\lambda_i, h)$  весьма состоятельны.

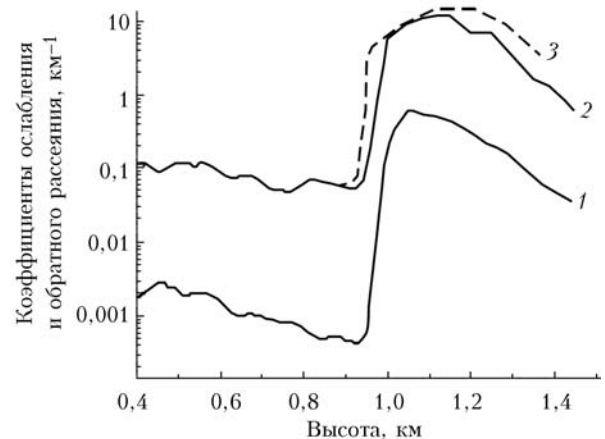


Рис. 2. Пример восстановления профилей коэффициентов обратного рассеяния  $\beta_\pi(\lambda, h)$  (кривая 1) и ослабления  $\beta_{ext}(\lambda, h)$  (2) модифицированным генетическим алгоритмом в сопоставлении с модельным профилем (3) для длины волны  $\lambda = 1,285$  мкм при угле поля зрения детектора  $\varphi_d = 0,01$  рад

Модельный высотный в пределах облака разрез функций распределения счетной концентрации облачных капель по размерам  $N(r)$  в данном численном эксперименте задавался в формате обобщенного  $\gamma$ -распределения [8] с постоянным модальным радиусом  $r_m = 4$  мкм и переменной по высоте

функцией  $N(r)$ , значения которой косвенно отражаются в модельных значениях коэффициентов ослабления  $\beta_{ext}(\lambda_i, h)$ , показанных на рис. 2. Результаты восстановления  $N(r)$  методом генетического поиска (1)–(7) приведены на рис. 3.

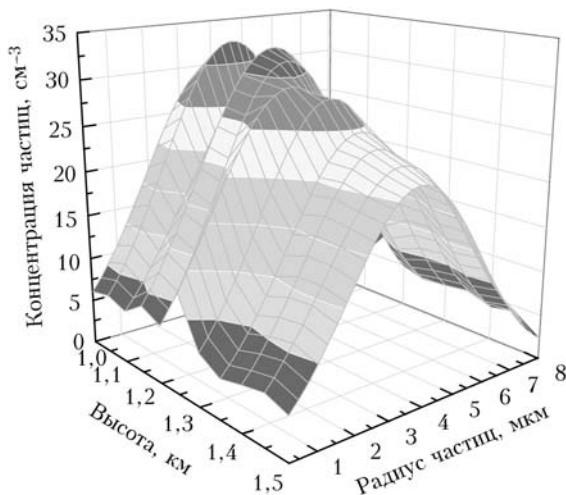


Рис. 3. Высотный профиль функции распределения счетной концентрации частиц  $N(r)$  по размерам, полученный на основе модифицированного генетического алгоритма с использованием четырех длин волн (1,285; 1,557; 1,629 и 2,130 мкм)

Можно утверждать, что предложенный алгоритм в достаточной степени воспроизводит унимодальный характер заданного профиля  $N(r)$  в классе функций Гаусса (2), в то же время по мере увеличения глубины зондирования оценочная величина модального радиуса испытывает заметные смещения, величина которых постепенно нарастает. В целом, однако, предложенная модификация генетического алгоритма показала повышенную по сравнению с известными методами устойчивость к помехе многократного рассеяния и позволила дос-

тичь рекордных ( $\tau \leq 5$ ) оптических глубин зондирования. Дальнейшее повышение эффективности алгоритма может быть связано с разумным, в рамках технических возможностей, увеличением количества длин волн зондирования.

Работа выполнена при финансовой поддержке РФФИ (гранты № 09-01-00698-а и 11-07-00257-а).

1. Krekov G.M., Sukhanov A.Ya. Application of artificial intelligence methods in remote sensing problems // XV Int. Sympos. «Atmospheric and Ocean Optics. Atmospheric Physics». June 22–29, 2008. Krasnoyarsk, 2008. С. 100. СО–05.
2. Креков Г.М., Крекова М.М., Суханов А.Я. Оценка эффективности использования перспективных лидаров белого света для зондирования микрофизических параметров слоистой облачности: 1. Аналитический обзор // Оптика атмосфер. и океана. 2009. Т. 22, № 7. С. 661–670.
3. Креков Г.М., Крекова М.М., Суханов А.Я. Оценка эффективности использования перспективных лидаров белого света для зондирования микрофизических параметров слоистой облачности: 2. Параметрическая модификация итерационного метода решения лидарного уравнения // Оптика атмосфер. и океана. 2009. Т. 22, № 8. С. 795–802.
4. Креков Г.М., Крекова М.М., Суханов А.Я. Оценка эффективности использования перспективных лидаров белого света для зондирования микрофизических параметров слоистой облачности: 3. Решение обратной задачи // Оптика атмосфер. и океана. 2009. Т. 5, № 9. С. 862–872.
5. Wolf J.-P., Bourayou R., Boutou V., Favre C., Kasparian J., Megian G., Mondelain D., Masyrowich A., Rodrigues M., Woster L., Yin Yu. Teramobile: a Nonlinear Femtosecond Terawatt Lidar // Proc. ILRC 21, Quebec City, Canada. Part 1. 2002. P. 47–50.
6. LIDAR: rang-resolved optical remote sensing of the atmosphere / Claus Weitkamp, ed. Singapore: Springer Science Media Inc., 2005. 451 p.
7. Зуев В.Е., Креков Г.М. Оптические модели атмосферы. Л.: Гидрометеиздат, 1986. 256 с.
8. Deirmendjan D. Electromagnetic Scattering of Spherical Polydispersions. N.Y.: Elsevier Publ. Co, 1969. 157 p.

*G.M. Krekov, A.Ya. Sukhanov. Improved genetic algorithm of multiwavelength lidar sounding of the atmospheric aerosol.*

Algorithm of cloud and aerosol microphysical characteristic restoration with the use of white light lidar systems based on femtosecond lasers is presented. The direct task of wideband radiation transport is solved by the Monte Carlo simulation method, and then inverse task simultaneous restoration of aerosols physical and optical characteristics are also solved using the modified genetic algorithm.