

УДК 004.021

ПРЕДСТАВЛЕНИЕ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ОБ ОПТИЧЕСКИХ ПАРАМЕТРАХ АТМОСФЕРЫ АНАЛИТИЧЕСКИМИ ЗАВИСИМОСТЯМИ

А.В. Демин, М.И. Монсева

Предлагается метод получения эмпирического выражения для расчета коэффициента пропускания атмосферы на основе экспериментальных данных для двух спектральных диапазонов. Описан алгоритм разбиения множества известных экспериментальных данных на группы, представимые в виде различных аналитических выражений. На основании этого определена эмпирическая формула вычисления коэффициента пропускания атмосферы.

Ключевые слова: коэффициент пропускания атмосферы, алгоритм, спектральный диапазон, дистанционное зондирование Земли.

Введение

Дистанционное зондирование Земли из космоса с помощью оптико-цифровых систем и комплексов сверхвысокого разрешения открыло новые возможности по повышению эффективности управления сферами человеческой жизнедеятельности и научного познания. Дистанционное зондирование Земли из космоса (ДЗЗ) в оптическом диапазоне длин волн обеспечивает получение информации как о географических и геофизических параметрах Земли, так и о техногенных процессах, происходящих на ее поверхности, что увеличивает ценность этой информации. ДЗЗ в зависимости от задач, для решения которых оно предназначено, осуществляется такими методами, как маршрутная, объектовая, стереоскопическая, статическая, динамическая, топографическая и спектрометрическая съемки. В оптическом диапазоне ДЗЗ осуществляется с помощью оптико-цифровых систем и комплексов, устанавливаемых на борт космических аппаратов.

Для повышения достоверности получаемых при ДЗЗ данных необходимо учитывать в процессе съемки состояние атмосферы, что накладывает дополнительную нагрузку на служебные системы и требует введения дополнительных систем, обеспечивающих учет параметрического состояния атмосферы в момент съемки. Имеются экспериментальные данные, однако условия их получения могут не соответствовать условиям текущей съемки. В этой связи актуальным является вопрос об аналитическом представлении состояния атмосферы в реальном масштабе времени с целью повышения эффективности и достоверности работы системы ДЗЗ.

Параметрическое состояние атмосферы в части ее пропускания определяется, помимо химического состава, климатической и географической привязкой. Искомая аналитическая зависимость $F(\tau_a)$ включает в себя кроме основной компоненты $\{F(\lambda)\}$, учитывающей спектральное пропускание атмосферы, географическую $\{F(\varphi, \xi)\}$ и климатическую $\{F(T)\}$ компоненты, уточняющие параметрическое состояние атмосферы при конкретных условиях съемки, а именно:

$$F(\tau_a) = \{F(\lambda) \circ F(\varphi, \xi) \circ F(T)\}. \tag{1}$$

Основным расчетным параметром, который необходимо учитывать при проведении ДЗЗ, является коэффициент пропускания атмосферы τ_a .

В [1, 2] имеются экспериментальные данные о значениях коэффициента пропускания атмосферы для диапазонов длин волн 3–5,2 мкм и 8–14 мкм с шагом 0,02 мкм. Экспериментальные данные получены для различных климатических широт и различных значений дальностей видимости (трасса не указана). Следует заметить, что исходные данные содержат множество локальных минимумов и максимумов, что затрудняет подбор единой эмпирической формулы для их представления в аналитической форме.

Таким образом, для учета значения τ_a при проведении ДЗЗ необходимо на основе имеющихся экспериментальных данных выявить функциональную зависимость, т.е. найти алгоритм для вычисления коэффициента пропускания атмосферы для произвольного значения длины волны λ из указанных выше спектральных диапазонов без хранения в памяти множества полученных ранее экспериментальных данных.

Целью данной работы является разработка алгоритма построения эмпирической функции $\tau_a(\lambda)$ на основе экспериментальных данных, полученных ранее.

Решение поставленной задачи

Известны лишь результаты измерений τ_a в некоторых дискретных точках рабочих диапазонов длин волн, полученных в экспериментах, условия проведения которых могут не совпадать с условиями конкретной съемки. Чтобы предсказать результат в любой произвольной точке рассматриваемого диапазона, необходимо построить график зависимости $\tau_a(\lambda)$. Иначе говоря, нужно получить эмпирическую формулу $\tau_a = F(\lambda)$.

Для построения эмпирической формулы необходимо решить две задачи [3]: выявить общий вид формулы и определить значения параметров для полученной формулы. Следует заметить, что формула, которая строится для расчета коэффициента пропускания атмосферы τ_a , является приближенной, поскольку основана на экспериментальных данных.

Для решения данной задачи выберем из [1, 2] в качестве исходных данных результаты измерений для диапазонов 3–5,2 мкм и 8–14 мкм при дальности видимости 5 км в климатической зоне тропиков. Будем предполагать, что искомая зависимость $\tau_a(\lambda)$ определена в каждой точке диапазона и может быть представлена всюду дифференцируемой аналитической функцией. Тогда эта зависимость может быть выражена целой функцией, т.е. функцией вида

$$F(\lambda) = c_0 + c_1 \cdot \lambda + \dots + c_n \cdot \lambda^n + \dots, \quad (2)$$

где c_i – константы, $i = 0, 1, 2 \dots$ [4].

Важно отметить, что искомая эмпирическая формула должна быть пригодна для проведения расчетов на борту космического аппарата, т.е. обеспечивать достаточную точность и скорость вычислений, а также быть удобной для хранения. Кроме того, как было отмечено выше, искомая функция не является монотонно возрастающей или монотонно убывающей, а содержит множество экстремумов. С учетом сказанного представим искомую функцию в виде совокупности элементарных функций. К классу элементарных функций относятся полиномы, показательные функции, логарифмические, тригонометрические и обратные тригонометрические функции, а также функции, получаемые из перечисленных выше с помощью четырех арифметических действий и суперпозиций, примененных конечное число раз [4].

Для представления искомой зависимости совокупностью элементарных функций разобьем исходный диапазон данных на несколько поддиапазонов и на каждом поддиапазоне будем искать свою эмпирическую формулу. При этом возникает проблема нахождения компромисса между выбором минимального числа интервалов разбиения и обеспечением достаточной точности и простоты получаемых для них формул.

Одним из простейших вариантов разбиения диапазона является разбиение по точкам экстремумов. При этом, поскольку в качестве исходных данных используется конечный массив дискретных данных, минимумы и максимумы искомой функции можно определить путем последовательного сравнения значений соседних элементов этого массива.

Другим вариантом разбиения может быть разбиение по точкам перегиба. В случае аналитической функции точки перегиба можно найти путем вычисления нулей второй производной функции. В данном же случае, когда известны лишь дискретные значения функции, вместо вычисления производной функции можно найти приближенные значения точек перегиба через вычисления приращений функции в ее соседних точках. Например, приближенное значение производной можно вычислить по формуле

$$f'(x_i) = \frac{1}{2} \cdot \left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} + \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \right). \quad (3)$$

Однако исходные экспериментальные данные имеют множество экстремумов и точек перегиба, а число интервалов разбиения должно быть минимально, как было сказано выше. Для решения поставленной задачи будем стремиться при разбиении выбрать один или два наиболее выдающихся экстремума. Также заметим, что каждая точка разбиения входит только в один из разделяемых ею поддиапазонов.

Данные из диапазона 8–14 мкм для поиска аналитических зависимостей разобьем на два участка (рис. 1).

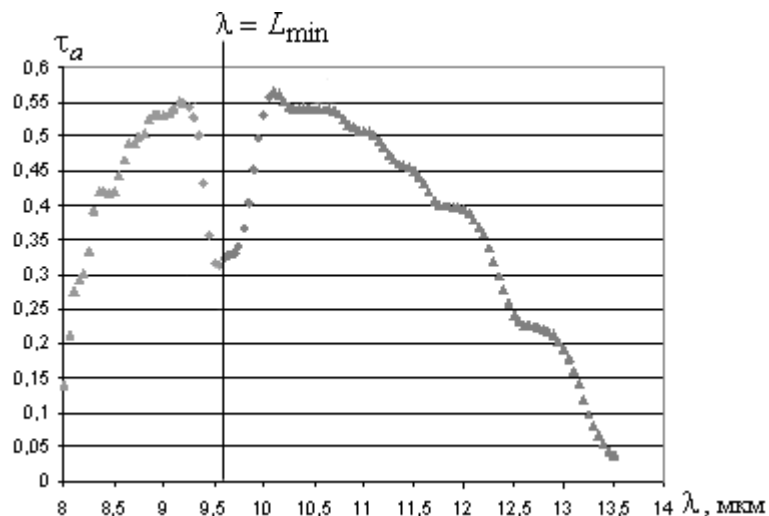


Рис. 1. Разбиение диапазона 8–14 мкм на участки для поиска зависимостей

Алгоритм разбиения – следующий:

1. Вычисляется приближенно производная $\tau'_a(\lambda)$ по формуле (3);
2. Находится абсцисса ее наименьшего минимума $\lambda_{d \min}$ (рис. 2);

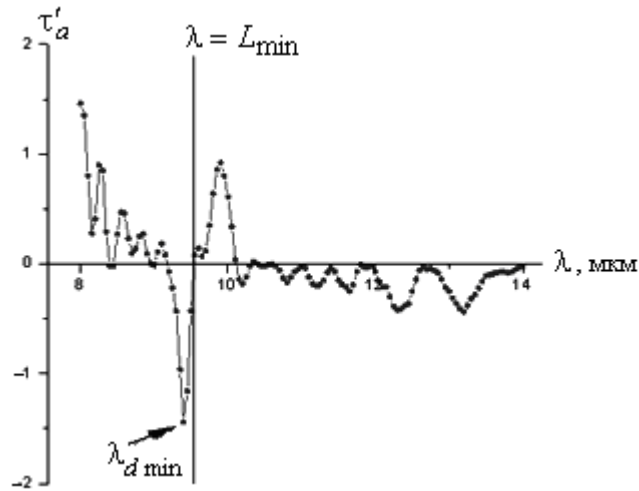


Рис. 2. График производной $\tau'_a(\lambda)$

3. Определяются экстремумы исходной функции путем поэлементного сравнения массива исходных данных;
4. Выбирается точка локального минимума исходной функции, чья абсцисса расположена наиболее близко к $\lambda_{d \min}$;
5. Абсцисса полученной точки обозначается как L_{\min} ;
6. Получаем два участка разбиения: $[\lambda_n; L_{\min}]$ и $(L_{\min}; \lambda_k]$. Здесь λ_n и λ_k – начальное и конечное значения диапазона, соответственно.

Данные из диапазона 3–5,2 мкм для поиска аналитических зависимостей разобьем на три участка (рис. 3) путем выделения наиболее выдающихся пиков. Алгоритм разбиения – следующий:

1. Находится «нулевая область» функции – наибольшая совокупность соседних абсцисс, в которых функция принимает нулевое значение. Абсцисса, ближайшая слева к нулевой области, обозначается как λ_{m2} ;
2. Находится λ_{m1} – абсцисса наименьшего минимума на интервале $[\lambda_n; \lambda_{m2}]$;
3. Приближение производится для трех участков: $[\lambda_n; \lambda_{m1})$, $[\lambda_{m1}; \lambda_{m2})$ и $(\lambda_{m2}; \lambda_k]$. Здесь λ_n и λ_k – начальное и конечное значения диапазона соответственно.

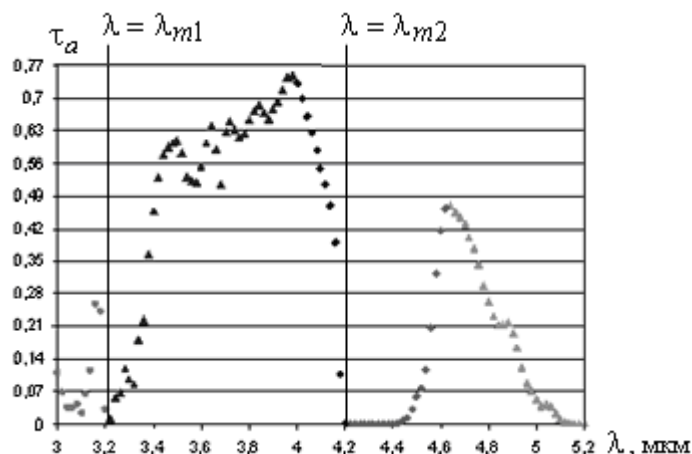


Рис. 3. Разбиение диапазона 3–5,2 мкм на участки для поиска зависимостей

Различия в разбиении диапазонов на участки обусловлены различиями в характере исходных экспериментальных данных. Также вследствие физических причин на каждом участке определяется свой вид кривой регрессии.

В результате перебора различных известных зависимостей (полиномиальной, экспоненциальной, логарифмической, кривой Гаусса и других) в качестве формулы приближения для всех участков диапазонов были выбраны полином пятой степени

$$F(\lambda) = c_0 + c_1 \cdot \lambda + c_2 \cdot \lambda^2 + c_3 \cdot \lambda^3 + c_4 \cdot \lambda^4 + c_5 \cdot \lambda^5 \quad (4)$$

с параметрами $c_0, c_1, c_2, c_3, c_4, c_5$ и функция вида

$$F(\lambda) = \tau_0 + A \cdot \frac{1}{1 + \exp\left(-\frac{\lambda - \lambda_c + \frac{1}{2} \cdot w_1}{w_2}\right)} \cdot \left(1 - \frac{1}{1 + \exp\left(-\frac{\lambda - \lambda_c - \frac{1}{2} \cdot w_1}{w_3}\right)} \right) \quad (5)$$

с параметрами $\tau_0, A, \lambda_c, w_1, w_2, w_3$.

Выбор зависимости проводился с использованием специальных математических пакетов, позволяющих провести аппроксимацию заданной функцией по методу наименьших квадратов и рассчитать коэффициент детерминации R^2 . Коэффициент детерминации вычисляется как отношение суммы квадратов регрессии к общей сумме квадратов. Таким образом, $0 < R^2 \leq 1$. $R^2 = 1$ в случае идеальной модели, когда все экспериментальные точки исходных данных лежат на полученной кривой регрессии. Для аппроксимации в итоге были выбраны функции, для которых коэффициент детерминации был наиболее близок к единице ($0,9 < R^2 \leq 1$).

После расчета значений параметров обеих формул приближения для каждого из участков выбирается та формула, для которой значение R^2 больше.

Определение наилучших параметров для полученных формул производится путем регрессионного анализа – с помощью метода наименьших квадратов. Поскольку аппроксимация производится нелинейными зависимостями, для минимизации величины отклонения воспользуемся итерационным алгоритмом. К итерационным методам относятся метод градиентного спуска, метод Ньютона, метод Ньютона–Гаусса и другие [5]. При решении поставленной задачи для нахождения параметров эмпирической формулы $\tau_a = F(\lambda)$ будем использовать алгоритм Левенберга–Маркуардта, представляющий собой «компромисс между методом линеаризации и методом наискорейшего спуска» [6].

Графики найденных зависимостей $\tau_a(\lambda)$ для дальности видимости 5 км в тропиках изображены на рис. 4. Для всех полученных кривых $R^2 > 0,94$.

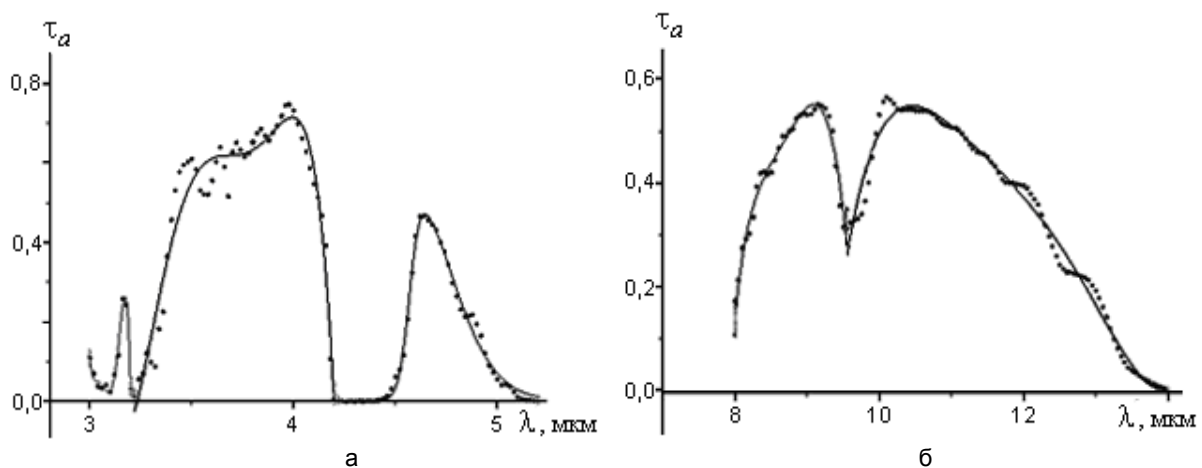


Рис. 4. Графики полученных аппроксимирующих зависимостей: диапазон 3–5,2 мкм (а); диапазон 8–14 мкм (б)

Для других исходных данных в этих же спектральных диапазонах общий вид эмпирических формул будет тот же. Различаться будут значения параметров формулы. Это обусловлено тем, что экспериментальные данные имеют географическую и климатическую привязку.

Точность приближения зависимости $\tau_a(\lambda)$, как было отмечено выше, оценивается коэффициентом детерминации R^2 . В настоящей задаче приемлемым значением R^2 будем считать значение, большее 0,9; оптимальным – большее 0,94.

Заключение

Предложен алгоритм разбиения множества известных экспериментальных данных для диапазонов 3–5,2 мкм и 8–14 мкм на группы для их представления аналитическими зависимостями. Объединение этих зависимостей позволит получить обобщенную эмпирическую формулу для расчета коэффициента пропускания атмосферы, что является предметом следующей работы.

Литература

1. Справочник по инфракрасной технике / Под ред. У. Волф, Г. Цисис. Т. 1. Физика ИК излучения: Пер. с англ. – М.: Мир, 1995. – 606 с.
2. Ллойд Дж. Системы тепловидения: Пер. с англ. – М.: Мир, 1978. – 417 с.
3. Демидович Б.П., Шувалова Э.З., Марон И.А. Численные методы анализа: Приближение функций, дифференциальные и интегральные уравнения / Под ред. Б.П. Демидовича. – 3-е изд. – М.: Наука, 1967. – С. 79–81.
4. Математическая энциклопедия в 5 томах / Под ред. И.М. Виноградова. – М.: Советская энциклопедия, 1985. – Т. 5. – С. 797–799.
5. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей: Справочное издание. – М.: Финансы и статистика, 1985. – С. 301–306.
6. Дрейпер Н., Смит Г. Прикладной регрессионный анализ: в 2-х кн. Пер. с англ. – 2-е изд. – М.: Финансы и статистика, 1986. – Т. 2. – С. 206–207.

- Демин Анатолий Владимирович* – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, dav_60@mail.ru
- Моисеева Мария Игоревна* – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, mim14@mail.ru