

**В.К. Ошлаков**

## **Интерпретация результатов фотометрирования облачных полей с применением модели составного сигнала. Учет нестационарности рядов наблюдений**

*Институт оптики атмосферы СО РАН, г. Томск*

Поступила в редакцию 7.05.2004 г.

Обсуждаются проблемы учета нестационарности в рядах наблюдений яркости солнечного излучения в условиях облачности. Рассматривается обоснованность применения модели сигналов с временным уплотнением при аппроксимации статистических характеристик флуктуаций яркости, наблюдаемых на разрезе облачного слоя.

### **Введение**

Одной из проблем, связанных с интерпретацией результатов наблюдений флуктуаций поля яркости нисходящего солнечного излучения в условиях облачности, является проблема нестационарности рядов наблюдений. Данная статья является продолжением публикации [1] и посвящена проблеме учета нестационарности в рядах наблюдений флуктуаций яркости солнечного излучения и обоснованности применения модели сигналов с временным уплотнением в задачах интерпретации результатов наблюдений в условиях облачности.

### **Учет влияния трендов**

При регистрации яркости в определенном угловом направлении и в узком спектральном интервале в рядах наблюдений существенно сказывается влияние линейных, апериодических или псевдопериодических трендов средних значений и дисперсий яркости. Основными причинами трендов яркости являются: суточный ход координат Солнца, пространственно-временная изменчивость характеристик облачного слоя, изменчивость оптической толщи и фазового состава элементов облачного слоя. Характер тренда в конкретной реализации зависит от длительности реализации и условий наблюдения [2, 3].

Традиционно задача выделения тренда решается с помощью известных процедур сглаживания и цифровой фильтрации временных рядов [2, 4]. В то же время исключить влияние, например, изменения положения Солнца можно, проводя наблюдения яркости в нескольких спектральных интервалах, изменяя тем самым мерность информационного пространства и анализируя изменчивость цветовых координат на трехкомпонентной диаграмме (ТКД). В этом случае намного упрощается процедура обработки результатов наблюдений [5–7]. Речь идет о проблеме выбора информационного признака

в задачах классификации, что является самостоятельной задачей, и эта проблема будет обсуждена в последующих публикациях.

Известно, что модальный размер капель воды, из которых состоят жидкокапельные облака, — 4–5 мкм [8]. Следовательно, если наблюдатель сравнивает соотношения амплитуд отсчетов яркости  $a(\lambda_i)$ , например зенита облачного и безоблачного неба в различных интервалах видимого диапазона спектра (на границах и в центре), то в соответствии с теорией Ми для безоблачного неба должно реализовываться соотношение  $a(\lambda_1) > a(\lambda_2) > a(\lambda_3)$  (синее небо), а для облачного —  $a(\lambda_1) < a(\lambda_2) > a(\lambda_3)$  («белый» цвет), так как такое соотношение характерно для источника излучения — Солнца. Качественно различия реализуются в смещении максимума спектрального распределения энергии из «зеленой» части спектра (для облачных частиц) в сторону более коротких длин волн (для безоблачного неба).

Изменчивость положения цветовых координат контролируемого информационного признака — спектральной яркости — на ТКД объясняется изменчивостью оптических характеристик облачности и безоблачного неба (например, оптической толщи), связанных с изменением их состава (в том числе фазового состава) и, следовательно, изменчивостью спектрального распределения рассеянного атмосферными частицами излучения от источника с постоянной спектральной характеристикой, в данном случае Солнца [6, 7]. Вывод о возможности классификации сигналов спектральной яркости по признаку «облачность» базируется на результатах анализа спектральных распределений яркости, приведенных в [3, 6].

В качестве примера исключения влияния тренда, связанного с изменением положения Солнца, можно привести результаты статистической обработки одной из реализаций. Наблюдения яркости зенита проводились на трех длинах волн:  $\lambda_1 = 0,42$ ,

$\lambda_2 = 0,53$  и  $\lambda_3 = 0,69$  мкм, в зимний сезон тип облачности  $S_c$  10 баллов. Время наблюдения — с 10 до 17 ч 30 мин. В полуденное время отмечалось выпадение осадков. Дискретность поступления данных — 36 с. Предварительный анализ производился с учетом данных Гидрометеослужбы и собственного поста наблюдений, оборудованного в соответствии с необходимыми требованиями. Пример выбирался в предположении, что при прохождении облачного слоя над пунктом наблюдения существенно изменялась лишь мощность слоя, а фазовый состав облачности оставался постоянным. В процессе обработки реализация была разбита на отрезки, в соответствии с данными расчета угловых координат Солнца для данного пункта наблюдений: до полудня, кульминация, после полудня (зенитные углы Солнца менее  $85^\circ$ ). С учетом априорной информации предполагалось, что тренды в выделенных участках реализаций близки к линейным.

Для последовательностей амплитуд сигналов в каждой длине волны  $a(\lambda_1)$ ,  $a(\lambda_2)$ ,  $a(\lambda_3)$ , суммарных сигналов  $S$  (суммы амплитуд) и цветоделенных сигналов  $m$ ,  $n$ ,  $l$  (отношение амплитуды на каждой длине волны к суммарному сигналу) были оценены, в качестве критерия к линейному тренду, коэффициенты ранговой корреляции  $\hat{\tau}$ . Фактически — это коэффициенты между порядком амплитуд сигналов в последовательности отсчетов и их порядком по величине. Затем были рассчитаны параметры линии тренда и определена величина  $\text{tg}\varphi$  — тангенс угла наклона линии тренда. В [2] указывается, что коэффициент ранговой корреляции  $\tau$  является довольно «мощным» критерием тенденции к линейному тренду: его асимптотическая относительная эффективность равна 0,98. Математическое ожидание  $\tau$  равно 0, дисперсия  $\sigma^2(\tau) = \frac{2(2n+5)}{9n(n-1)}$ , распре-

деление быстро стремится к нормальному. Для оценки значимости вычисленного значения  $|\hat{\tau}|$  достаточно сравнить его с теоретической величиной  $|3\sigma(\tau)|$  (проверка нулевой гипотезы о величине среднего значения  $\tau = 0$  против альтернативы: среднее значение  $\tau$  равно вычисленному значению  $\hat{\tau}$ ). В данном случае величина  $3\sigma(\tau)$  равна 0,11. В таблице приведены значения вычисленных коэффициентов ранговой корреляции и значения  $\text{tg}\varphi$  линий тренда для исходных последовательностей на трех длинах волн, суммарного сигнала и для цветоделенных сигналов.

Последовательность сигналов	До полудня		После полудня	
	$\hat{\tau}$	$\text{tg}\varphi$	$\hat{\tau}$	$\text{tg}\varphi$
$a(\lambda_1)$	0,431	0,172	-0,854	-0,287
$a(\lambda_2)$	0,354	0,508	-0,819	-0,948
$a(\lambda_3)$	0,33	0,301	-0,787	-0,562
$S = \sum a(\lambda_i)$	0,354	0,979	-0,353	-0,158
$m = a(\lambda_1)/S$	0,304	0,000017	0,39	0,000017
$n = a(\lambda_2)/S$	0,166	0,000018	-0,684	-0,000047
$l = a(\lambda_3)/S$	-0,171	-0,000015	0,635	0,000032

Для последовательностей цветоделенных сигналов значения тангенсов углов наклона практически равны 0 при значимых оценках тенденции к линейному тренду.

## Проверка гипотез о дисперсиях и средних наблюдаемых значениях спектральной яркости

С целью проверки обоснованности выбора модели проверялись гипотезы о различии средних значений дисперсий и средних значений наблюдаемой яркости облачного и безоблачного неба. Наблюдения проводились на трех длинах волн — 0,42; 0,53; 0,69 мкм. Одновременно контролировались метеопараметры, оценивались высота нижней границы, направление и скорость перемещения облачности и выбирались реализации с совпадающими параметрами. При проверке гипотез было проанализировано 12 реализаций длительностью от 3,5 до 7,5 ч (направление наблюдения — зенит): 7 — для разрывной облачности, 5 — для безоблачных ситуаций.

При анализе предполагалось, что яркость разрыва в облачном слое стремится к яркости безоблачного неба в том же угловом направлении и при том же положении Солнца и использовалось представление последовательности отсчетов с помощью модели [3, 9, 11]. Напомним, что в этой модели среднее значение  $\bar{B}_n$  и средний квадрат наблюдаемой яркости представлены линейными зависимостями:

$$\bar{B}_n(p) = p\bar{B}_1 + (1-p)\bar{B}_0 = p(\bar{B}_1 - \bar{B}_0) + \bar{B}_0, \quad (1)$$

$$\overline{B_n^2}(p) = p\overline{B_1^2} + (1-p)\overline{B_0^2} = p(\overline{B_1^2} - \overline{B_0^2}) + \overline{B_0^2}, \quad (2)$$

а дисперсия  $\sigma_n^2$  — квадратичной зависимостью от вероятности наличия облака на линии визирования  $p$  [6]:

$$\sigma_n^2(p) = -p^2(\bar{B}_1 - \bar{B}_0)^2 + p[(\overline{B_1^2} - \overline{B_0^2}) - 2\bar{B}_0(\bar{B}_1 - \bar{B}_0)] + \sigma_0^2. \quad (3)$$

В (1)–(3) индексы  $n, 1, 0$  относят соответствующие величины к классам «наблюдаемое значение», «облако» и «разрыв в облачном слое».

Предполагается, что анализируемые последовательности освобождены от трендов, связанных с изменением положения Солнца.

## Соотношение выборочных дисперсий

При проверке гипотез о равенстве дисперсий двух выборок — наблюдаемых в условиях облачности и в отсутствие облачности — статистикой критерия (в предположении о нормальности распределений отсчетов яркости в двух классах значений) будет отношение выборочных дисперсий [10]:

$$F^* = \sigma_n^2 / \sigma_0^2 > 1, \quad (4)$$

т.е. бóльшую из выборочных дисперсий помещают в числитель дроби.  $F^*$  сравнивают с критическим значением из  $F$ -распределения Фишера:  $F_{v_n, v_0, 1-\alpha/2}$ , где  $\alpha$  — уровень значимости;  $v_n = N_n - 1$  и  $v_0 = N_0 - 1$  — число степеней свободы, соответствующее оценкам  $\sigma_n^2$  и  $\sigma_0^2$ ,  $N_n$  и  $N_0$  — объемы выборок. Нулевая гипотеза отвергается, если отношение двух выборочных дисперсий больше критического.

В обсуждаемом представлении отношение дисперсий с учетом (3) запишется в виде

$$F^*(p) = \frac{\sigma_n^2(p)}{\sigma_0^2} = 1 + \frac{-p^2(\bar{B}_1 - \bar{B}_0)^2 + p[(\bar{B}_1^2 - \bar{B}_0^2) - 2\bar{B}_0(\bar{B}_1 - \bar{B}_0)]}{\bar{B}_0^2 - (\bar{B}_0)^2}. \quad (5)$$

Так как дисперсия яркости в условиях облачности зависит от вероятности наличия облака на линии визирования, то и  $F^*$  зависит от  $p$ . Если число степеней свободы много больше 120, то теоретические значения  $F$ -распределения Фишера близки к единице. В этом случае нулевая гипотеза с использованием критерия (5) принимается, если выполняется одно из двух условий: или в проверяемых выборках  $p = 0$  или если  $p \neq 0$ , то одновременно  $\bar{B}_1 = \bar{B}_0$  и  $\bar{B}_1^2 = \bar{B}_0^2$ . При выполнении второго условия модель (1)–(3) теряет смысл. Кроме того, из (1) следует, что если  $p \neq 0$ , то для всех  $\bar{B}_n \neq \bar{B}_0$  значения  $\bar{B}_1 \neq \bar{B}_0$  и отношение дисперсий в (5) больше единицы. Ограничения на соотношение средних квадратов не определены. Следовательно, критерий (5) должен быть чувствителен при всех значениях  $p > 0$  и при соблюдении условия неравенства средних значений яркости облака и разрыва в облачном слое.

### Соотношения выборочных средних значений

Рассмотрим процедуру сравнения средних значений двух совокупностей: контрольной — «чистое небо» — и «наблюдаемой в условиях облачности».

Для проверки средних значений двух совокупностей в случае, когда обе дисперсии неизвестны и не предполагается, что они равны, статистикой критерия является

$$T = \bar{B}_n - \bar{B}_0 / \sqrt{\frac{\sigma_n^2}{N_n} + \frac{\sigma_0^2}{N_0}}, \quad (6)$$

ее распределение близко к  $t$ -распределению Стьюдента с числом степеней свободы  $v$ , лежащим между наименьшим из  $N_n - 1$  и  $N_0 - 1$  и их суммой  $(N_0 + N_n - 2)$  [10].

Подставляя в (6) значение  $\bar{B}_n(p)$  и  $\sigma_n^2(p)$  из (1), (3), получим:

$$T(p) = p(\bar{B}_1 - \bar{B}_0) / \sqrt{\frac{\sigma_n^2(p)}{N_n} + \frac{\sigma_0^2}{N_0}}. \quad (7)$$

Так как и среднее значение, и дисперсия наблюдаемых флуктуаций яркости зависят от вероятности наличия облака на линии визирования, то и статистика критерия проверки средних значений зависит от этой величины. Из (7) следует, что проверка гипотезы о равенстве средних значений сводится к проверке условия равенства средних значений яркости облака и разрыва в облачном слое. Условие проверяется с учетом значения вероятности наличия облака на линии визирования  $p \neq 0$  и оценок дисперсий наблюдаемой яркости  $\sigma_n^2(p)$  и чистого неба  $\sigma_0^2$ , причем оценки дисперсий не могут быть одновременно равны нулю.

На практике процедура проверки гипотезы сводится к проверке условия  $T > t_{v, 1-\alpha/2}$ . Если условие выполняется, то нулевая гипотеза о равенстве средних значений отвергается, следовательно, наблюдения проводятся в условиях облачности. Для заданных значений  $\alpha$ ,  $B_0$ ,  $\sigma_0^2$  и наблюдаемых  $\bar{B}_n$ ,  $\sigma_n^2$ , после небольших преобразований с учетом (3), это условие запишется в следующем виде: если

$$N_n \left\{ \left[ \frac{(\bar{B}_n - \bar{B}_0)}{t_{v, 1-\alpha/2}} \right]^2 - \frac{\sigma_0^2}{N_0} \right\} > \sigma_n^2, \quad (8)$$

или с учетом (1) для заданных значений  $p$ , если

$$N_n \left\{ \left[ \frac{p(\bar{B}_1 - \bar{B}_0)}{t_{v, 1-\alpha/2}} \right]^2 - \frac{\sigma_0^2}{N_0} \right\} > \sigma_n^2(p), \quad (9)$$

то для (8) наблюдаемые значения яркости не равны значениям яркости безоблачного неба, а для (9) — при данном значении вероятности наличия облака на линии визирования яркость облака не равна яркости безоблачного неба. Следовательно, выборка наблюдаемых значений относится к наблюдениям в условиях облачности.

### Совместные распределения отсчетов относительно средних значений

Из вышеизложенного следует, что результаты наблюдения яркости в заданном направлении могут быть представлены смесью нескольких классов сигналов, отличающихся средними значениями и, возможно, дисперсиями наблюдаемых значений. Известно, что разделение смеси сигналов на заданное количество классов выполняется при условии неравенства средних значений в разных классах сигналов. Задача обнаружения полезных сигналов может быть решена в результате анализа эмпирических распределений сигналов в смешанной выборке. Выборка наблюдаемых значений может быть описана

бимодальным распределением вероятностей смеси сигналов от облака и разрыва в облачном слое:

$$f_p(x) = p_1 f_1(x) + p_0 f_0(x), \quad (10)$$

где  $f_1$  и  $f_0$  — законы распределения сигналов в классе «облако» и «разрыв в облачном слое»; вероятности  $p_1$  и  $p_0 = (1 - p_1)$  априори неизвестны [12].

Из анализа имеющихся эмпирических распределений можно сделать следующие выводы: при отсутствии трендов средних значений и дисперсий распределение, например, результатов обработки рядов сигналов цветности зенита в условиях разрывной облачности подчиняется распределению смеси сигналов (10) (рис. 1). Аналогичные выводы сделаны и при обработке результатов наблюдений в одной длине волны (рис. 2).

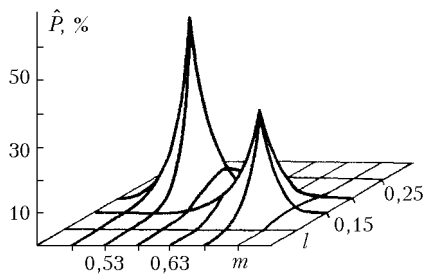


Рис. 1. Распределение вероятностей координат цветности зенита в условиях кучевой облачности [6]

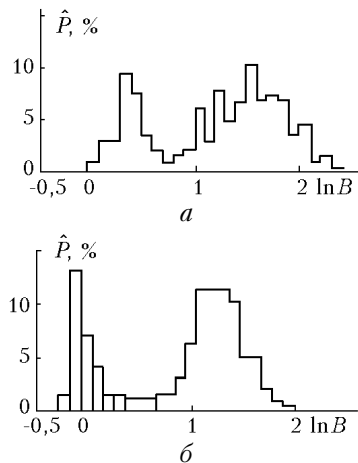


Рис. 2. Гистограммы распределения фотометрических отсчетов от облачного зенита ( $\lambda = 0,69$  мкм): а — до введения, б — после введения поправок на высоту нижней кромки облачности и спектральную чувствительность фотометра [3]

Напоминаем, что при обработке исключено влияние тренда, связанного с изменением положения Солнца и, при необходимости, были введены поправки на спектральную чувствительность фотометра и учтены изменения высоты нижней кромки облачности [3, 6]. В приведенных примерах (рис. 1 и 2) гипотезы о равенстве средних значений сигналов, а также их дисперсий отвергнуты на уровне значимости 0,01.

Изменчивость положения цветовых координат на ТКД (см. рис. 1) объясняется различием оптических характеристик облачности и безоблачного неба, связанных с их составом и, следовательно, с различием в спектральном распределении рассеянного излучения от источника с постоянной спектральной характеристикой (Солнца).

Как правило, ряды наблюдений освобождают от детерминированной составляющей (центрируют средним значением), а затем — нормируют на средноквадратическое отклонение с целью приведения наблюдений к рядам с нулевым средним и единичной дисперсией. Дальнейшие выводы о флуктуационных составляющих делают на основе анализа «остатков». В [1] приведен общий вид корреляционной функции флуктуаций яркости солнечного (оптического) излучения в условиях разрывной облачности. Из анализа аналитического представления корреляционной функции [1] следует, что если произвести центрирование исходного ряда наблюдаемой яркости средним значением (1), то в корреляционной функции и, следовательно, в спектральной плотности останутся составляющие, связанные с пространственно-оптической неоднородностью как элементов облачного поля, так и самого поля. С целью исключения влияния «внутренних» неоднородностей элементов облачного поля центрирование средним значением и нормирование на средноквадратическое отклонение следует производить в каждом классе сигналов отдельно. В этом случае описание наблюдаемых процессов может быть сведено к комбинациям их средних значений и средноквадратических отклонений, что подтверждается видом эмпирических распределений наблюдаемых флуктуаций яркости (см. рис. 1, 2).

## Выводы

По результатам обработки наблюдений была проверена обоснованность использования модели сигналов с временным уплотнением в задачах интерпретации результатов наблюдения флуктуаций яркости в условиях облачности. Сделаны выводы о значимости различий средних значений и дисперсий наблюдаемой яркости солнечного излучения, рассеянного облачным полем в выделенном угловом направлении.

Приведенные соотношения использовались в следующих случаях:

- при классификации предварительно освобожденных от тренда выборок,
- для проверки гипотез о соотношении средних значений и дисперсий для выборок, полученных в совпадающих условиях наблюдения,
- для разработки правил принятия решения о наличии облачности на линии визирования.

Рассматривались два класса сигналов в следующих комбинациях: безоблачное небо и кучевая облачность (40 реализаций длительностью 3,5 ч каждая с шагом по времени 15–20 с, с предварительной классификацией по признакам: сезон наблюдения, тип облачности, балл облачности, поло-

жение Солнца), безоблачное небо и облачность любого типа – во всех остальных случаях. Отмечена возможность классификации сигналов на подмножества: облака жидкокапельной структуры и облака кристаллической структуры [6].

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 02-05-65169-а).

1. *Ошлаков В.К.* Интерпретация результатов фотометрирования облачных полей с применением модели составного сигнала // Оптика атмосф. и океана. 2004. Т. 17. № 4. С. 345–349.
2. *Кендал М., Стьюарт А.* Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
3. *Ошлаков В.К.* Модель составного сигнала в задачах интерпретации результатов фотометрирования облачных полей // Оптика атмосф. и океана. 1990. Т. 3. № 3. С. 289–283.
4. *Введение в цифровую фильтрацию* / Под ред. Р. Богнера и А. Константидиса. М.: Мир, 1976. 216 с.
5. *Быков Р.Е., Гуревич С.Б.* Анализ и обработка цветных и объемных изображений. М.: Радио и связь, 1984. 248 с.
6. *Ошлаков В.К.* Определение наличия облачности на линии визирования по результатам фотометрирования // Оптика атмосф. и океана. 1990. Т. 3. № 4. С. 431–435.
7. *Галилейский В.П., Морозов А.М., Ошлаков В.К.* Панорамный фотометрический комплекс для контроля прямой и рассеянной солнечной радиации // Региональный мониторинг атмосферы. Ч. 2. Новые приборы и методики измерений / Под ред. М.В. Кабанова. Томск: Изд-во «Спектр» Института оптики атмосферы СО РАН, 1997. 295 с.
8. *Облака и облачная атмосфера: Справочник* / Под ред. И.П. Мазина, А.Х. Хргана Л.: Гидрометеониздат, 1989. 647 с.
9. *Френкс Л.* Теория сигналов. М.: Сов. радио, 1974. 344 с.
10. *Джонсон Н., Лион Ф.* Статистика и планирование эксперимента в технике и науке. Методы обработки данных. М.: Мир, 1980. 510 с.
11. *Шалыгин А.С., Палагин Ю.И.* Прикладные методы статистического моделирования. Л.: Машиностроение, 1986. 320 с.
12. *Миленский А.В.* Классификация сигналов в условиях неопределенности. Статистические методы самообучения в распознавании образов. М.: Сов. радио, 1975. 328 с.

***Oshlakov V.K. Interpretation of the results of photometry of cloud fields applying the model of compound signal. Accounting for nonstationarity of observation series.***

The problems of taking nonstationarity into account in the series of observations of solar radiation brightness under cloudy conditions are discussed. Application of the model of signals with time compression to approximation of statistical characteristics of brightness fluctuations observed at a cloud layer cross section is justified.