

Факторный анализ многоканальных спутниковых изображений океана для исследования эффектов поверхностного волнения

Д.М. Ермаков^{1,2}, М.Т. Смирнов¹

¹*Институт радиотехники и электроники РАН, Фрязинский филиал,
141190 Фрязино Московской области, пл. Введенского, 1
E-mail: dima@ire.rssi.ru; smirnov@ire.rssi.ru;*
²*Институт космических исследований РАН,
117997, Москва, Профсоюзная, 84/32*

Предложен метод совместной обработки дистанционных спутниковых данных, направленный на выявление и оценку влияния мелкомасштабного ветрового возмущения поверхности на регистрируемый цвет океана. Метод основан на оригинальной нелинейной модели факторного анализа. Численные эксперименты и обработка спутниковой информации показали перспективность метода для решения поставленной задачи и позволили дать первые оценки «цветовой чувствительности» океана к скорости приповерхностного ветра.

Ключевые слова: оптика океана, гравитационно-капиллярное волнение, радиометрия, совместный анализ.

Введение

В работах, посвященных уточнению алгоритма атмосферной коррекции данных спутниковых сканеров цвета океана, (Лихачева и др., 2009; Шеберстов и др., 2009) показана важность учета «мгновенного» состояния поверхности океана в связи с влиянием малоразмерных бликов на гравитационно-капиллярных волнах (ГКВ). Из анализа данных наземных экспериментов известны и другие эффекты, связывающие вариации воспринимаемого цвета океана с мелкомасштабными возмущениями его поверхности (Бородина, Салин, 2010; Ермаков et al., 2010). Моделирование этих связей позволяет, при определенных условиях наблюдения, восстанавливать спектры ГКВ по данным *in situ* (Титов и др., 2010). Однако, непосредственный перенос результатов такого моделирования на спутниковые данные невозможен из-за огромных размеров пятна разрешения спутниковых приборов. С другой стороны, остается открытым вопрос, в какой мере сохраняется искажающее воздействие ГКВ на регистрируемый цвет океана после процедуры атмосферной коррекции, призванной, в частности, компенсировать эффекты, связанные с вариациями состояния поверхности.

В настоящей работе предложен метод совместной обработки данных оптического сканера цвета с данными синхронных СВЧ радиометрических спутниковых измерений, направленный на выявление в оптических данных информации о состоянии поверхности. Метод позволяет оценить «чувствительность» восстановленной характеристики цветности океана к вариациям состояния поверхности. После накопления статистики по большому объему спутниковых измерений метод открывает возможности как для компенсации влияния состояния поверхности на регистрируемый цвет океана, так и для привлечения оптических данных к анализу состояния поверхности океана дистанционными методами.

Математическая модель метода совместной обработки

Рассмотрим спутниковые данные двух типов: некоторую характеристику цвета океана, рассчитанную по данным оптического сканера цвета (О-данные), и скалярное поле скоростей приповерхностного ветра по измерениям СВЧ-радиометрического сканера (М-данные). Предполагаем, что измерения выполнены синхронно обоими сканерами и данные наложены на общую координатную сетку, так что i -й точке измерений соответствуют величины o_i О-данных и m_i М-данных. Требуется проверить гипотезу о том, что в О-данных, помимо «истинного» цвета океана c_i содержится ненулевая добавка $h(w_i)$, зависящая от скорости ветра w_i , известного из М-данных с погрешностью e_i :

$$\begin{cases} o_i = h(w_i) + c_i \\ m_i = w_i + e_i \end{cases} \quad (1)$$

Вид функции h неизвестен. Будем считать ее нелинейной непрерывной функцией от w , а вклад $h(w_i)$ в o_i существенно меньшим, либо сравнимым с c_i . Таким образом, (1) ставит нелинейную факторную задачу с общим фактором w_i и характерными факторами c_i и e_i . Для ее полного решения необходимо сначала найти переменную факторную нагрузку, функцию h (т.е. проверить гипотезу о влиянии ветра на регистрируемый цвет), а затем оценить вклады общего и характерного факторов (т.е. восстановить «истинный» цвет c_i). Сделаем дополнительное предположение о качестве восстановления ветра по М-данным:

$$\bar{e} = \frac{1}{N} \sum_i e_i = 0, \quad \sigma_e^2 = \frac{1}{N} \sum_i e_i^2 \ll \sigma_w^2 \quad (2)$$

Здесь и далее σ обозначает выборочную вариацию, N – общее число измерений. Условие (2) означает, что оценка ветра m_i по М-данным – несмещенная, с малой относительной погрешностью.

Разобьем весь диапазон W возможных значений w_i на интервалы $W_j = [w_j^-, w_j^+)$, на каждом из которых h может быть аппроксимирована линейной зависимостью от w_i :

$$h(w_i) = A_j w_i + B_j, \quad w_i \in W_j, \quad \bigcup_j W_j = W \quad (3)$$

Чтобы оценить вклад общего фактора в o_i , необходимо вычислить A_j и B_j для каждого интервала W_j . Обозначим через I_j множество (мощности N_j) таких индексов i , для которых значения w_i принадлежат W_j . Тогда из (1) и (3) следует:

$$o_i = A_j w_i + B_j + c_i, \quad i \in I_j \quad (4)$$

Нормализуя w_i и o_i (вычитая выборочное среднее и деля на выборочную вариацию)

$$x_i = \frac{w_i - \bar{w}}{\sigma_w}, \quad y_i = \frac{o_i - \bar{o}}{\sigma_o}, \quad i \in I_j, \quad (5)$$

получим, подставив (5) в (4):

$$\sigma_o y_i = A_j \sigma_w x_i + B_j + A_j \bar{w} - \bar{o} + c_i, \quad i \in I_j \quad (6)$$

Умножая обе части (6) на x_i и суммируя по всем индексам I_j , получим, с учетом того, что x_i и y_i имеют нулевое среднее и единичную дисперсию:

$$\sigma_o \rho_j = A_j \sigma_w + \frac{1}{N_j} \sum_i x_i c_i, \quad \rho_j = \frac{1}{N_j} \sum_i x_i y_i \quad (7)$$

В соответствии с факторной моделью полагаем, что значения общего фактора w_i (истинной скорости ветра) не коррелируют со значениями характерного фактора c_i («истинного» цвета океана). Тогда, с учетом нулевого среднего x_i , имеем:

$$\sum_i x_i c_i = 0, \quad i \in I_j \quad (8)$$

Из (7) и (8) следует:

$$A_j = \frac{\sigma_o \rho_j}{\sigma_w} \quad (9)$$

Значения σ_g и ρ_j неизвестны, но, принимая во внимание (1) и (2), получим:

$$m_i = w_i, \quad \sigma_m^2 = \sigma_w^2 + \sigma_e^2, \quad \rho_j = \frac{\sigma_o}{\sigma_w} \rho'_j, \quad (10)$$

где ρ'_j – выборочная корреляция o_i и m_i , вычисленная для индексов i , составляющих I_j . Из (9) и (10) следует оценка для A_j :

$$A_j = \frac{\sigma_o \sigma_m}{\sigma_w^2} \rho'_j = \frac{\sigma_o}{\sigma_m} \left(1 + \frac{\sigma_e^2}{\sigma_w^2} \right) \rho'_j \approx \frac{\sigma_o}{\sigma_m} \rho'_j, \quad (11)$$

которая может быть вычислена по исходным данным.

Чтобы оценить B_j , рассмотрим два соседних интервала:

$$W_{j-1} = [w_{j-1}^-, w_{j-1}^+), \quad W_j = [w_j^-, w_j^+), \quad w_{j-1}^+ = w_j^- \quad (12)$$

Если (3) (в силу непрерывности h) построена как непрерывная кусочно-линейная аппроксимация, то значения B_{j-1} и B_j для интервалов (12) связаны соотношением:

$$B_j = B_{j-1} + A_{j-1} w_{j-1}^+ - A_j w_{j-1}^- \quad (13)$$

По соотношению (13) можно последовательно рассчитать значения B_j для всех интервалов W_j , если известно значение B_k для одного интервала W_k . Учитывая физический смысл введенных переменных, следует считать, что при нулевом значении общего фактора (скорости ветра), его вклад в О-данные также будет нулевым. Тогда из (3):

$$B_0 = 0 \quad (14)$$

Система (11), (13), (14) позволяет вычислить все A_j, B_j из входных данных и тем самым восстановить факторную нагрузку h в кусочно-линейном приближении. Однако, в оценку вклада $h(w_i)$ в o_i входит неизвестное значение w_i . Если w_i принадлежит интервалу W_j , то из (1) и (3) следует:

$$o_i = A_j m_j + B_j - A_j e_j + c_i \quad (15)$$

Учитывая (2), можно рассматривать член $A_j m_j + B_j$ как несмещенную оценку вклада общего фактора в О-данные с погрешностью $A_j e_j$. Тогда $o_i - (A_j m_j + B_j)$ является несмещен-

ной оценкой c_p , полученной с той же погрешностью. Остается вопрос поиска интервала W_j , в котором лежит w_i с заданным индексом i . Простейший подход состоит в том, чтобы принять в качестве W_j интервал W_k , которому принадлежит оценка m_i , т.к. $m_i \approx w_i$. Если w_i лежит вблизи границы интервала W_j то в некоторых случаях (в зависимости от величины e_i) будет ошибочно выбран соседний интервал W_{j-1} (W_{j+1}). Однако представляется, что эти случаи не вносят больших ошибок (что подтверждает численное моделирование), т.к. могут рассматриваться как линейная экстраполяция непрерывной функции h на небольшие области (с характерным размером σ_e) за пределы интервала W_j .

Алгоритм совместной обработки спутниковых данных

Входными данными алгоритма являются О-данные и М-данные (см. предыдущий раздел). Требуется разложить О-данные в сумму Н-вклада (ветровой фактор) и С-вклада («истинный» цвет) с помощью М-данных в смысле задачи (1), т.е. вычислить значения Н- и С-вкладов во всех точках по формуле:

$$h_i = A_j m_i + B_j, c_i = o_i - h_i \quad (16)$$

где A_j, B_j – решения системы (11), (13), (14), построенные для некоторого разбиения $\{W_j\}$ динамического диапазона $[0, mlim)$ М-данных. Выбор разбиения должен учитывать два требования: каждый интервал W_j должен быть достаточно широк, чтобы обеспечить существенно положительную σ_m (11), но достаточно узок, чтобы приближение (3) выполнялось с хорошей точностью. Например, разбиение на n интервалов равной длины Δ задается следующим образом:

$$W_j = [j \Delta, (j + 1) \Delta), j = 0..n-1 \quad (17)$$

При заданном разбиении алгоритм выполняет n (по числу интервалов) одинаковых шагов. В начале j -го шага формируется множество индексов точек измерений I_j по критерию

$$i \in I_j \Leftrightarrow m_i \in W_j. \quad (18)$$

Затем для данных с индексами i из I_j вычисляются значения σ_o, σ_m и ρ'_j , см. (10), и рассчитывается величина A_j по (11), а величина B_j – по (13) с использованием вычисленных на предыдущих шагах значений и условия (14).

После n описанных шагов известны все значения A_j и B_j , и алгоритм завершается вычислением Н- и С- вкладов согласно (16). Отметим, что совокупность значений A_j и B_j представляет самостоятельный интерес, т.к. описывает воздействие ветрового возмущения на регистрируемый цвет океана в виде кусочно-линейной функции от скорости ветра.

Тестирование метода обработки численным моделированием

Для анализа корректности предложенного метода совместной обработки дистанционных данных были проведены серии тестов с применением численного моделирования. Каж-

дый тест был организован следующим образом. Сначала задавали массив w_i значений общего фактора (аналитически или в виде растрового изображения) и функцию «ветрового вклада в цвет» $h(w)$, как кусочно-линейную зависимость, определенную на всем диапазоне w_i . Далее по заданным w_i и $h(w)$ рассчитывали Н-вклад (см. предыдущий раздел) как $h(w_i)$, а в качестве характеристических факторов e_i и c_i использовали реализации белого шума с разными значениями дисперсий и средних (среднее e_i всегда было равно 0). В результате, по (1) генерировали О- и М-данные, являющиеся входной информацией для алгоритма совместной обработки.

Алгоритм обработки рассчитывал последовательность значений A_j, B_j (см. выше), задающих приближение $h^*(w)$ исходной функции $h(w)$, и вычислял оценки Н* Н-вклада и С* С-вклада. При анализе качества восстановления исследовалась близость заданных и восстановленных значений; примеры приведены в разделе обсуждения результатов.

Обработка спутниковых данных

При переходе к обработке реальных дистанционных данных принципиальным вопросом является адекватность их описания моделью (1) и, прежде всего, возможность статистически достоверной проверки гипотезы о не тождественном равенстве нулю функции «ветрового вклада в цвет» $h(w)$. Анализ метода показывает, что построение кусочно-линейного приближения $h(w)$ выполняется путем оценки значений A_j (11) ее производной по w и численным интегрированием (13) с начальным условием (14). Иными словами, выражение (11) аппроксимирует «цветовую чувствительность» к ветру $s(w)$ коэффициентом ковариации О- и М-данных в узких диапазонах значений М-данных:

$$s(w) = \frac{dh}{dw} \approx \lim_{\Delta w \rightarrow 0} \text{cov}(o_i, m_i), w - \Delta w \leq m_i \leq w + \Delta w \quad (19)$$

Приближение (19) тем точнее, чем меньше Δw , т.е., чем уже интервалы разбиения (3) диапазона значений М-данных. С другой стороны, чем меньше Δw , тем меньше выборка соответствующих пар О- и М-данных, и тем менее достоверна оценка их ковариации выборочным значением (11). Таким образом, актуальна задача формирования больших однородных массивов синхронных О- и М-данных и их совместной обработки для статистически достоверной оценки $s(w)$ и $h(w)$ (в частности, сравнения с нулем) в разных диапазонах значений w .

В настоящей работе приведены некоторые результаты пробной совместной обработки информации приборов спутника TERRA: О-данных прибора MODIS (<http://oceancolor.gsfc.nasa.gov/>) и М-данных AMSR-E (http://nsidc.org/data/ae_ocean.html).

Обсуждение результатов численного моделирования

Выше была описана методика численного моделирования для анализа качества работы предложенного алгоритма совместной обработки. Рассмотрим пример, в котором распределение общего фактора задается аналитически в виде двумерной функции от x, y :

$$w_i = f(x, y) = 127.5 \left(1 + \cos \left(\frac{2\pi \sqrt{(x-300)^2 + \frac{1}{2}(y-300)^2}}{400} \right) \right), x = 0..399, y = 0..399, i = 400y + x \quad (20)$$

Функция $h(w)$ задана в виде ломаной, определенной для $w = 0..255$ (рис. 1); c_i – реализация белого шума с равномерным распределением в диапазоне $0..128$; для e_i рассмотрены два случая: тест 1) $e_i = 0$ и тест 2) e_i – белый шум с равномерным распределением в диапазоне $-16..16$. Н-вклад построен как $h_i = h(w_i)$; М-данные – как $m_i = w_i + e_i$; О-данные – как $o_i = h_i + c_i$. Данные визуализировали полутоновыми изображениями (256 градаций яркости), яркость в точке (x,y) численно равнялась визуализируемой величине с соответствующим индексом i (20). Случай теста 2) представлен на рис. 2.

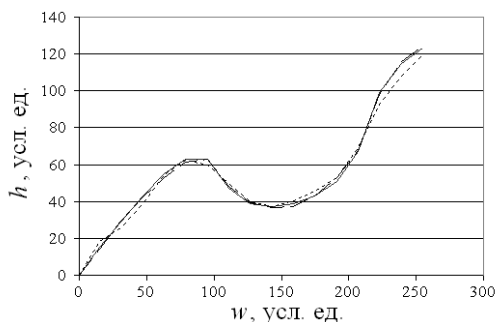


Рис. 1. Модельная функция «ветрового вклада» в цвет h (сплошная линия) и восстановленные функции h^* : длинный пунктир – тест 1); короткий пунктир – тест 2). Условные единицы

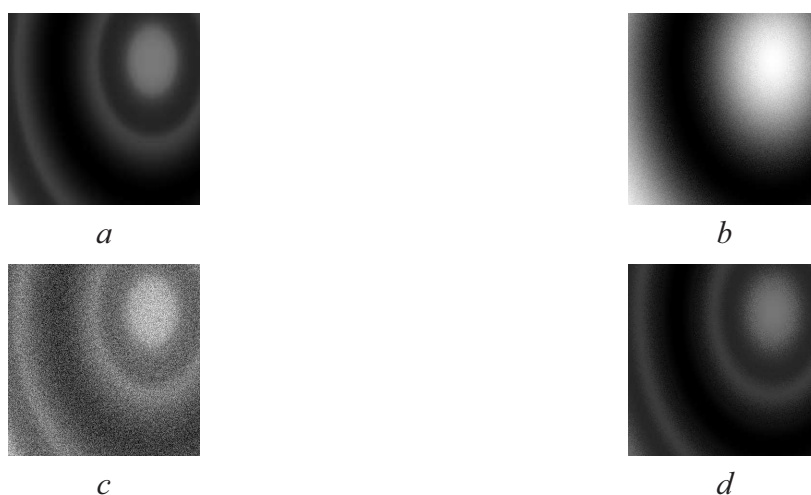


Рис. 2. Визуализация данных тестового случая 2): a) модельный Н-вклад; b) М-данные; c) О-данные; d) восстановленный Н*-вклад

Используя построенные М- и О-данные в качестве входных для алгоритма совместной обработки восстановили функцию ветрового вклада в цвет $h^*(w)$ (рис. 1) и распределение ветрового вклада H^* (рис. 2). Для оценки близости h и h^* рассчитаны вариации σ_h и средние значения m_h разностей $h(j) - h^*(j)$, $j = 0..255$. Для оценки близости Н- и Н*-вкладов рассчитаны вариации σ_H and средние m_H разностей для всех точек. Результаты оценок приведены в Таблице 1.

Таблица 1. Оценки точности восстановления h^* и H^*

Тест 1)	σ_h	m_h	σ_H	m_H	Тест 2)	σ_h	m_h	σ_H	m_H
	1.0	3.1	0.9	0.9		3.1	1.3	5.9	0.9

В целом результаты численного моделирования свидетельствуют о корректности предложенного подхода.

Результаты совместной обработки спутниковых данных

В результате пробной обработки спутниковых данных MODIS (в качестве характеристики цвета использовалось отношение нормализованных яркостей на длинах волн 443 и 555 нм) и ASMR-E (поля ветра), накопленных над акваторией Тихого океана за 11.10.2010–21.10.2010, были рассчитаны функции «цветовой чувствительности» $s(w)$ (19) и «ветрового вклада в цвет» $h(w)$ (1). Для исследования устойчивости работы алгоритма были применены четыре разных разбиения диапазона значений скорости ветра 0–17 м/с, восстановленных по данным ASMR-E, на узкие интервалы W_j (3), см. Таблицу 2.

Таблица 2. Интервалы W_j : минимальное/максимальное значение, м/с (варианты разбиения №1..4)

№	j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
		W_0	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{11}	W_{12}	W_{13}	W_{14}
1		0.0/	1.1/	2.3/	3.4/	4.6/	5.7/	6.8/	8.0/	9.1/	10.2/	11.4/	12.5/	13.7/	14.8/	15.9/
		1.1	2.3	3.4	4.6	5.7	6.8	8.0	9.1	10.2	11.4	12.5	13.7	14.8	15.9	17.0
2		0.0/	2.3/	4.6/	6.8/	9.1/	11.4/	13.7/	15.9/	-	-	-	-	-	-	-
		2.3	4.6	6.8	9.1	11.4	13.7	15.9	17.0							
3		0.0/	1.8/	2.9/	4.1/	5.2/	6.4/	7.5/	8.6/	9.8/	10.9/	12.0/	13.2/	14.3/	15.5/	-
		1.8	2.9	4.1	5.2	6.4	7.5	8.6	9.8	10.9	12.0	13.2	14.3	15.5	17.0	
4		0.0/	2.9/	5.2/	7.5/	9.8/	12.0/	14.3/	-	-	-	-	-	-	-	-
		2.9	5.2	7.5	9.8	12.0	14.3	17.0								

Результаты восстановления «цветовой чувствительности» $s(w)$ показаны на рис. 3. Вертикальными линиями обозначены доверительные интервалы оцененных значений по уровню вероятности 95%. Видно, что $s(w)$ положительна в диапазоне скоростей ветра 1.0–2.5 м/с и отрицательна при скоростях ветра 3.5–5.5 м/с, т.е. статистически достоверно доказано влияние ветрового волнения на регистрируемый цвет океана. Однако имеющийся разброс значений в пределах доверительных интервалов обуславливает неустойчивость восстановления функции ветрового вклада. Примеры восстановленных функций $h(w)$ для четырех разбиений (3) (см. Таблицу 2) приведены на рис. 4. Для улучшения оценок $h(w)$ необходимо сузить доверительные интервалы оценок $s(w)$ путем накопления и обработки большего массива спутниковых данных.

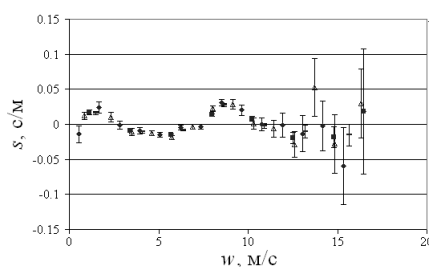


Рис. 3. Оценки «цветовой чувствительности» $s(w)$ к ветру w на разных интервалах скоростей ветра: ромбы – разбиение №1; квадраты – №2; треугольники – №3; планка – №4 (см. Таблицу 2). Вертикальные полосы – доверительные интервалы по уровню 95%

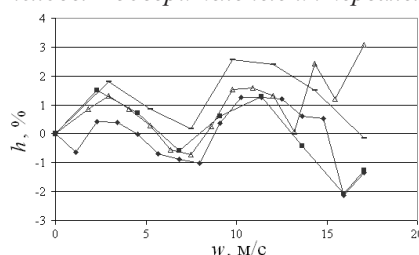


Рис. 4. Оценки «ветрового вклада» $h(w)$ в цвет океана (в % от полной величины) на разных интервалах скорости ветра w (по разбиениям №1..4, как на рис. 3)

Заключение

Численные эксперименты и обработка спутниковой информации показали перспективность применения предложенного метода совместной обработки спутниковых данных в задаче исследования влияния ветрового волнения на регистрируемый цвет океана и позволили получить первые оценки «цветовой чувствительности» океана к скорости приповерхностного ветра. Развитие данного подхода требует уточнения полученных оценок, которое принципиально достижимо путем накопления и обработки больших однородных массивов спутниковых данных.

Литература

1. *Ermakov D.M., Smirnov M.T., Polyakov V.M.* Integrating optical remote data into passive microwave remote sensing of the ocean // Proc. ESA Living Planet Symposium, Bergen, Norway (Ed. H. Lacoste-Francis), ESA SP-686 (CD ROM), ESA Communications, European Space Agency, ESTEC, Noordwijk, The Netherlands, December 2010.
2. *Бородин Е.Л., Салин М.Б.* Оценка пространственно-временных характеристик поверхностного волнения по видеоизображению // Известия РАН. Физика атмосферы и океана. 2010. Т.46. №2. С. 263-273.
3. *Лихачева М.В., Копелевич О.В., Шеберстов С.В.* Коррекция данных спутникового сканера MODIS на солнечный блик без использования дополнительных данных о скорости ветра // Современные проблемы дистанционного зондирования Земли из космоса. 2009. Вып.6. Т.1. С. 412-428.
4. *Титов В.И., Баханов В.В., Зуйкова Э.М., Лучинин А.Г., Троицкая Ю.И.* Исследование динамики двумерных спектров морского волнения // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т.7. №3. С. 197-207.
5. *Шеберстов С.В., Вазюля С.В., Копелевич О.В.* Алгоритм использования информации, получаемой при сканировании вдоль трассы полета, для атмосферной коррекции данных спутникового сканера цвета // Современные проблемы дистанционного зондирования Земли из космоса. 2009. Вып.6. Т.1. С. 234-242.

Factor analysis of multi-channel satellite images of the ocean to study the effects of surface roughness

D.M. Ermakov^{1,2}, M.T. Smirnov¹

¹*Institute of Radioengineering and Electronics of RAS, Fryazino department,
141190, Fryazino, Moscow region, 1 Vvedenskogo sq.,
E-Mail: dima@ire.rssi.ru;*

²*Space Research Institute of RAS
117997, Moscow, 84/32, Profsoyuznaya Str.*

A method of joint processing of remote satellite data is proposed, aimed at identifying and assessing the impact of small-scale surface roughness into the registered ocean color. The method is based on the original nonlinear model of factor analysis. Numerical simulations and processing of satellite data showed the promise of the proposed method for the task and allowed to give first estimates of the “ocean color sensitivity” to the surface wind speed.

Keywords: ocean optics, gravity-capillary waves, radiometry, joint analysis.