Remote Sensing of Environment 112 (2008) 4186–4203

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development

Catherine A. Brown [a,b,*], Yannick Huot [a,b], P. Jeremy Werdell [c], Bernard Gentili [a,b], Hervé Claustre [a,b]

[a] CNRS, UMR 7093, Laboratoire d'Océanographie de Villefranche, 06230 Villefranche-sur-Mer, France
[b] Université Pierre et Marie Curie-Paris 6, UMR 7093, Laboratoire d'Océanographie de Villefranche, 06230 Villefranche-sur-Mer, France
[c] NASA Goddard Space Flight Center, Code 614.2, Greenbelt, MD 20771 USA

ARTICLE INFO

Article history:
Received 23 November 2007
Received in revised form 4 June 2008
Accepted 20 June 2008

Keywords:
Remote sensing
MODIS
Optical properties
Backscattering
CDOM
Ocean color
Bio-optics

ABSTRACT

Empirical algorithms based on first order relationships between ocean color and the chlorophyll concentration ([Chl]; mg m^{-3}) are widely used, but cannot explain the statistical dispersion or "anomalies" around the mean trends. We use an empirical approach that removes the first order effects of [Chl] from satellite ocean color, thus allowing us to quantify the impact on the ocean color signal of optical anomalies that vary independently of the global mean trends with remotely sensed [Chl]. We then present statistical and modeling analyses to interpret the observed anomalies in terms of their optical sources (i.e. absorption and backscattering coefficients). We identify two main sources of second order variability for a given [Chl]: 1) the amount of non-algal absorption, especially due to colored dissolved organic matter; and 2) the amplitude of the backscattering coefficient of particles. The global distribution of the anomalies displays significant regional and seasonal trends, providing important information for characterizing the marine optical environment and for inferring biogeochemical influences. We subsequently use our empirically determined anomalies to estimate the backscattering coefficient of particles and the combined absorption coefficient for colored detrital and dissolved materials. This purely empirical approach provides an independent assessment of second order optical variability for comparison with existing methods that are generally based on semi-analytical models.

1. Introduction

Ocean color can be quantified as the normalized water-leaving radiance (nLw; in mW cm^{-2} μm^{-1} sr^{-1}) at several wavelengths in the visible spectrum (e.g. 412, 443, 488, 531 and 551 nm for the MODIS instrument on the Aqua platform). The nLw is defined as the radiance that would be measured exiting a flat ocean surface with the sun at zenith and the atmosphere absent (Gordon et al., 1988). If correctly normalized (Morel et al., 2002), nLw is theoretically independent of the observational conditions and varies only with the amount and composition of constituents that absorb and backscatter light according to a relationship of the form

$$nLw(\lambda) = f(b_b(\lambda), a(\lambda)) \quad (1)$$

where f represents a function, $b_b(\lambda)$ the backscattering coefficient (m^{-1}), $a(\lambda)$ the absorption coefficient (m^{-1}), and λ the wavelength (nm) (Morel and Prieur, 1977).

In open ocean waters far from the influence of terrestrial runoff, generally referred to as Case 1 waters, the spectral quality and intensity of light leaving the ocean depends to the first order on the concentration of phytoplankton (Morel and Prieur, 1977; Morel, 1980). Since phytoplankton absorb strongly in the blue and weakly in the green, empirical algorithms of ocean color based on blue-to-green ratios are used to estimate the chlorophyll *a* concentration ([Chl], in mg m^{-3}) within the upper layer of the water column (McClain et al., 1998; O'Reilly et al., 1998; McClain et al., 2004). These algorithms inherently account for the average effect of all in-water constituents on the measured optical signal, insofar as they co-vary predictably with the algal biomass. Optically significant substances that influence ocean color other than phytoplankton and the water medium include: 1) organic particles, 2) inorganic particles, and 3) colored dissolved organic matter (CDOM) (Sathyendranath et al., 1989; Carder et al., 1991; IOCCG, 2000; Stramski et al., 2001). Standard empirical algorithms, however, do not explain the statistical dispersion around the mean trends, which we refer to here interchangeably as "second order variability" and "anomalies" in ocean color.

Advances in the analysis of ocean color data have led to inverse methods based upon models such as Eq. (1) for the separation of light absorbing and backscattering components in seawater. Semi-

* Corresponding author. CNRS, UMR 7093, Laboratoire d'Océanographie de Villefranche, 06230 Villefranche-sur-Mer, France.
E-mail address: brown@obs-vlfr.fr (C.A. Brown).

0034-4257/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2008.06.008

analytical inverse models provide global estimates of [Chl] of comparable accuracy to the empirical band ratio algorithms, with the added benefit of simultaneously retrieving other optical constituents such as absorption coefficients (e.g. for CDOM, non-algal particulate matter and phytoplankton) and particulate backscattering coefficients (b_{bp}; m^{-1}) (e.g. Roesler and Perry, 1995; Hoge and Lyon, 1996; Lee et al., 2002; Maritorena et al., 2002). Many of these models require that spectral shapes, based on empirical datasets, be pre-assigned to the various optical components in order to subsequently solve for the constituents' magnitudes using an optimization procedure. Other approaches rely more on manipulating analytical radiative transfer equations. The application of such models to satellite imagery produced the first global analyses of second order optical variability in open ocean waters showing that non-algal absorption and b_{bp} vary to some extent independently of the [Chl] distribution (Loisel et al., 2002; Siegel et al., 2002, 2005). The term non-algal absorption encompasses the effects of both CDOM and non-algal particulate absorption, but the proportion of the former in the open ocean is often considered to be significantly greater (Siegel et al., 2002); here, we will refer to these optical constituents collectively as colored detrital and dissolved materials (CDM) and to their absorption coefficient as a_{CDM} (m^{-1}). The uncertainty in the global distributions of a_{CDM} and b_{bp} derived from semi-analytical models is largely due to: 1) the paucity of field observations for ground truth validation, and 2) the similar optical parameterizations used by many models that can limit efforts for the independent evaluation of their results (IOCCG, 2006).

Natural variability around the mean statistical relationships for Case 1 waters is a well-recognized feature (Gordon and Morel, 1983; Morel, 1988; Bricaud et al., 1995, 1998; Lee and Hu, 2006). In some cases, specific regions have been compared (Morel et al., 2007a) to examine locations where second order variability consistently differs from the global average. Large in situ datasets have also been used to systematically analyze this dispersion (Bricaud et al., 2004). Alvain et al. (2005) empirically relate second order variability in SeaWiFS spectra to four phytoplankton groups based on a dataset of coincident nLw observations and pigment measurements. Ocean color anomalies were calculated by dividing SeaWiFS nLw observations by empirically derived reference spectra from a look-up table of SeaWiFS [Chl] and nLw, thus removing the first order variability in ocean color. The method does not require any assumptions about the characteristics of the optically significant constituents. Instead, it simply associates the dominant algal group in a sample with a range of average anomaly. Furthermore, it avoids potential biases present in a particular ocean color sensor measurement or processing scheme.

For ocean color data, it can be proposed that departures from the established global mean trends with [Chl] generally arise from two main conditions.

Firstly, the optical constituents may vary completely independently of each other or the algal biomass; these are the so-called Case 2 waters where variability comes from substances other than the algal retinue (i.e. “algal retinue” refers to all material that covaries with [Chl]). Examples include the episodic upwelling of CDM rich water and aeolian deposition of desert dust, as well as the effect of terrestrial runoff in coastal waters.

Secondly, the optical constituents may co-vary with [Chl] over a given region, but the relationships could be different from the global mean trends. These conditions give rise to ocean color nuances in so-called Case 1 waters and are the focus of this paper. Two main sources of this variability can be identified, the first is variability in chlorophyll specific coefficients (i.e. absorption and backscattering coefficients normalized to [Chl]) for the *non-algal material*; that is, variability in the algal retinue that is not well represented by the established mean trends. For example, for a given [Chl] compared to the global average, the Mediterranean Sea has higher CDM absorption, whereas the subtropical Pacific Ocean has lower CDM absorption (Morel et al., 2007a). It has furthermore been hypothesized that unique biota and associated detrital material occurring in the presence of specific algal groups may lead to this type of nuance in ocean color. The second source of variability is due to the chlorophyll specific coefficients of *phytoplankton*, which may also differ from the global mean trends. Variability in algal properties occurs, for example, during coccolithophore blooms (e.g. Balch et al., 1996) and in other high [Chl] conditions associated with changes in the algal absorption spectra (e.g. changes in pigment packaging or pigment composition) (Bricaud et al., 1988; Morel, 1988; Loisel and Morel, 1998).

In this study, we use an empirical approach to examine second order variability in ocean color; our aim is to broadly analyze the sources of this variability in terms of potential optically significant substances and without introducing *a priori* assumptions about the influence of algal groups. To do so, differences between satellite measured nLw and the global mean nLw are calculated for a common ocean color index (e.g. maximum band ratio, MBR, which serves as a proxy for [Chl] in many operational remote-sensing algorithms, O'Reilly et al., 1998). By removing the information contained in the biomass field, the relative influence of other optical constituents and hence their distributions are observed. We obtain estimates of second order variability in an approach completely independent of semi-analytical models and with some significant advantages. Most notably, our method does not require various assumptions or parameterizations for characterizing the spectral shapes of the optical constituents and is not influenced by potential biases in the underlying physical model.

Our study will be presented in four steps. Firstly, a statistical analysis will be used to identify the main types of anomalies observed in satellite ocean color. Secondly, a modeling analysis is undertaken to interpret the identified anomalies in terms of optically significant constituents. Thirdly, various in situ datasets are used to evaluate our results. Finally, global algorithms for b_{bp} and a_{CDM} are developed based upon anomalies observed in satellite ocean color.

2. Methods and data

2.1. Statistical analysis

2.1.1. First order variability

In a manner analogous to Alvain et al. (2005) we use an empirical approach to remove first order variability from satellite ocean color thereby allowing us to identify second order effects.

Table 1
Description of the datasets

Dataset Name	Location	Description
nLw database	Global	540,000 MODIS Aqua nLw and anomaly spectra
Statistical database	Global	10,800 anomaly spectra extracted randomly from the nLw database
Modeling results	Applicable to Case 1 waters	Simulated anomaly spectra using an ocean color model for Case 1 waters
Match-up database	Global	334 MODIS nLw spectra and match-up [Chl]
BBOP dataset	Bermuda Atlantic Timeseries Study (near Bermuda)	Five year timeseries of surface absorption and radiometric data
BIOSOPE dataset	Southeastern Pacific	In situ b_{bp} and [Chl] data measured in surface waters and MODIS Aqua data extracted along the BIOSOPE transect
IOCCG dataset	Global, but biased to coastal waters	656 stations providing *in situ* measurements of R_{rs}, [Chl] and $a_{CDM}(412)$

Twelve global eight-day composites (MODIS-Aqua, Level-3 binned data) spanning the year 2005 (i.e. one image per month) were downloaded from NASA's ocean color website (http://oceancolor.gsfc.nasa.gov/). To obtain a representative sample of the variability present in remotely sensed nLw, 45,000 pixels were randomly selected from each of the twelve eight-day composite images, thus providing 540,000 nLw spectra (the "nLw database"; see Table 1. for a description of the different datasets used herein). No attempts were made to exclude coastal (Case 2) waters, which form only a small fraction of the total oceanic surface and thus have a limited influence on the results. For each spectrum, the maximum band ratio (MBR) was computed as the maximum of two ratios, ${}^{443}_{551}R_{\mathrm{Lw}}$ (i.e. ${}^{443}_{551}R_{\mathrm{Lw}} = nLw_{443}/nLw_{551}$) and ${}^{488}_{551}R_{\mathrm{Lw}}$, such that MBR = max $({}^{443}_{551}R_{\mathrm{Lw}}, {}^{488}_{551}R_{\mathrm{Lw}})$. We retain the original 551 nm notation for band 4, while noting that 547 nm is the correct waveband center. Eleven polynomial fits were obtained for the "nLw database" between: 1) nLw (i.e. 412, 443, 488, 531 and 551 nm) vs. MBR; and 2) nLw ratios (i.e. ${}^{412}_{443}R_{\mathrm{Lw}}$, ${}^{412}_{488}R_{\mathrm{Lw}}$, ${}^{412}_{531}R_{\mathrm{Lw}}$, ${}^{443}_{488}R_{\mathrm{Lw}}$, ${}^{443}_{531}R_{\mathrm{Lw}}$ and ${}^{488}_{531}R_{\mathrm{Lw}}$,) vs. MBR. While the fits could have been carried out using [Chl] as the independent variable, to avoid limiting the approach to a specific [Chl] algorithm or version we instead opted to use the MBR. We represent the results of these fits as $\hat{nLw}_{\lambda}$ and ${}^{i}_{j}\hat{R}_{Lw}$, where i and j correspond to the wavebands of the numerator and denominator, respectively. These fits are empirical measures of first order variability, which we will subtract from our ocean color observations in order to examine only second order effects.

2.1.2. Second order variability

Each of the 540,000 pixels in the "nLw database" is assigned a vector, or anomaly spectrum (AS), with eleven elements representing differences relative to a hypothetical "average pixel" with the same MBR (represented by $\hat{nLw}_{\lambda}$ and ${}^{i}_{j}\hat{R}_{Lw}$),

$$AS = \begin{bmatrix} nLw_{412} - \hat{nLw}_{412} \\ \vdots \\ nLw_{551} - \hat{nLw}_{551} \\ {}^{412}_{443}R_{Lw} - {}^{412}_{443}\hat{R}_{Lw} \\ \vdots \\ {}^{488}_{531}R_{Lw} - {}^{488}_{531}\hat{R}_{Lw} \end{bmatrix} = \begin{bmatrix} AS_{412} \\ \vdots \\ AS_{551} \\ {}^{412}_{443}AS \\ \vdots \\ {}^{488}_{531}AS \end{bmatrix} \quad (2)$$

The notation for the AS vector uses AS_k to represent an element corresponding to a single nLw band (k is the waveband center, e.g. AS_{412}), whereas ${}^{i}_{j}AS$ corresponds to a ratio of nLw bands (i and j are the wavebands for the numerator and denominator, respectively, e.g. ${}^{412}_{443}AS$). To normalize AS for the subsequent clustering analysis, for each element we subtract the mean ($\overline{AS}$) anomaly for all pixels for a given band or ratio of bands and divide by their standard deviation (σ_{AS}) referring to this normalized vector as AS^{N} (e.g. for AS_{551}, $AS^{N}_{551} = \left(AS_{551} - \overline{AS}_{551}\right)/\sigma_{\mathrm{AS}}(551)$).

2.1.3. Data mining

Statistical techniques for data mining are designed to extract useful information from large datasets and include methods such as

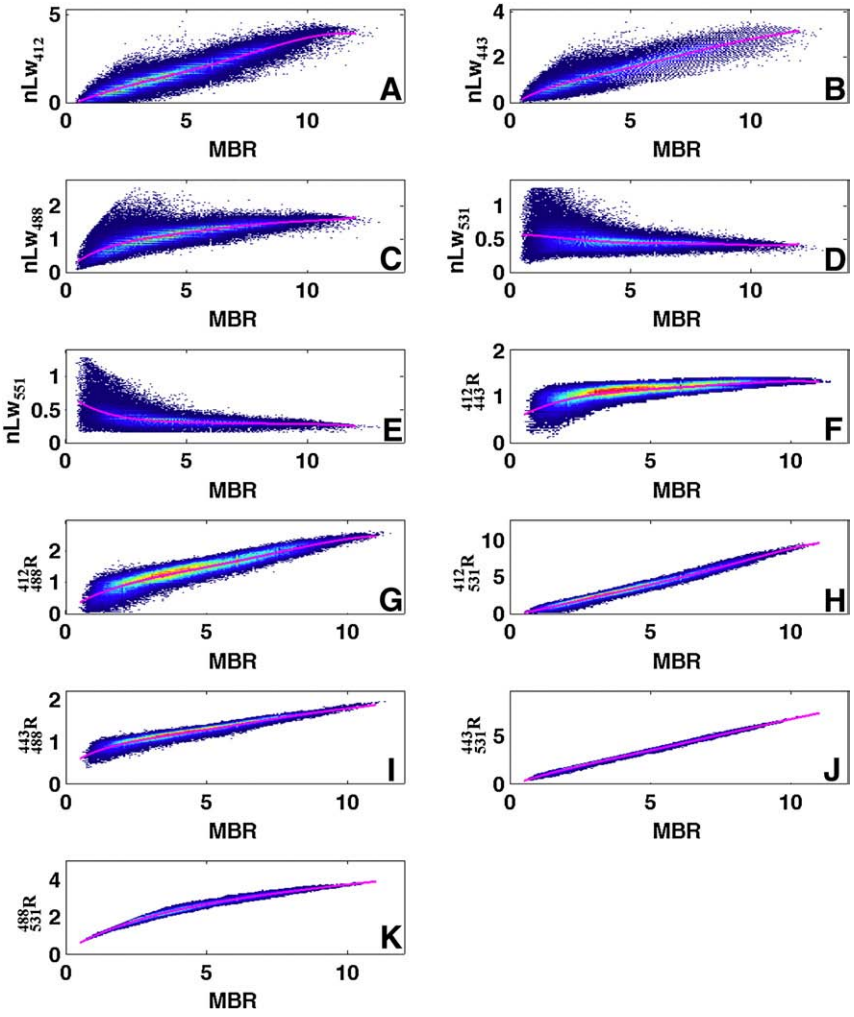

Fig. 1. Polynomial fits for different ocean color measures (see the ordinate label) and the maximum band ratio (MBR, the maximum of either $nLw(443)/nLw(551)$ or $nLw(488)/nLw(551)$).

clustering and principal component analysis (PCA). These two methods are used in conjunction to identify the main types of variability in the AS^N vectors. We randomly select 1 out of every 50 AS^N vectors from the total of 540,000 contained in the "nLw database" and are left with 10,800 AS^N vectors for the statistical analysis (the "statistical database"). For the cluster analysis, we use a standard K-means algorithm (MacQueen, 1967) that iteratively relocates the "centroid" representing the theoretically average AS^N spectrum associated with each cluster. The algorithm aims to minimize the total intra-cluster variance, in this case a squared error function

$$V = \sum_{p=1}^{q} \sum_{x_n \in S_p} |x_n - \mu_p|^2, \quad (3)$$

where there are q clusters identified here as S_p (p=1,2,…,q), x_n is a vector representing the nth data point, and μ_p is the centroid or mean point of all the points $x_n \in S_p$. This algorithm categorizes variability in large datasets by matching each data point to a centroid. This unsupervised classification method requires that the number of groups be pre-determined. This is generally done by running the classification method for different numbers of clusters and then using standard indices to evaluate the "correct" number of clusters; these indices usually compare the changes in the variance explained by the addition of a new cluster. We examined several of these indices in order to choose the final number of clusters. We found that the number of clusters varies slightly (generally ±1) amongst the indices chosen. We also verified that our results and conclusions are not affected by the number of clusters within the ranges provided by the different indices. Three different sets of anomalies were obtained for: 1) normalized water-leaving radiance only (i.e. AS_k; "nLw clusters"); 2) nLw band ratios only (i.e. ${}^{i}_{j}AS$; "band ratio clusters"); and 3) combined nLw and nLw ratios (i.e. AS_k and ${}^{i}_{j}AS$; "mixed clusters"). To avoid local minima during the minimization step, we chose the sets of clusters that explained the most variance in the dataset after reinitializing the K-means algorithm 20 times with randomly selected initial values.

The cluster analysis can identify the main types of variability in the "statistical database" but it does not provide a description of the sources of this variability in terms of optical constituents. This biogeochemical interpretation is developed using a modeling study described below.

2.2. Modeling study

The method developed in this study utilizes anomalies from the mean ocean color relationships as a function of the MBR. To develop a theoretical framework for the interpretation of these anomalies, we examine a series of radiative transfer simulations (described in detail in Appendix 1 and summarized in Table A1) with a diverse range of optical constituents that result in the same MBR, and hence [Chl] as retrieved from empirical algorithms. In this way, the modeling study allows us to simulate and compare the influence that various optical conditions with the same MBR have on the AS vectors. For four different [Chl] (see Appendix 1), we simulate the effect of both increasing and decreasing: 1) a_{CDM}; 2) b_{bp} amplitude; and 3) b_{bp} slope. The b_{bp} slope is the exponent of a power law describing the spectral dependency of the particulate backscattering coefficient. The lowest (0.03 mg m^{-3}) and highest (0.5 mg m^{-3}) [Chl] were chosen to bound ~97% of the surface [Chl] observed in the world ocean by satellite remote sensing (Antoine et al., 2005).

2.3. Match-up database

We acquired MODIS-Aqua Level-2 [Chl] validation results from the NASA Ocean Biology Processing Group. This dataset consists of 334 coincident pairs of satellite remote sensing reflectance (Rrs, the ratio of water-leaving radiance to surface irradiance; calculated using nLw/F_0, where F_0 is the mean extraterrestrial solar irradiance) and in situ [Chl], where 271 in situ values were obtained by fluorometry and 100 by HPLC (when both values are available we chose the HPLC values). Bailey and Werdell (2006) provide a complete description of the satellite-to-in situ validation process. Briefly, Level-2 files were generated using the MODIS-Aqua Reprocessing 1.1 processing configuration, which includes the Gordon and Wang (1994) atmospheric correction approach, plus corrections for near infrared water-leaving radiances, bi-directional reflectance, and spectral band-pass effects (Morel et al., 2002; Patt et al., 2003). Pixels including land, clouds, ice, stray light, sun glint, high top-of-the-atmosphere radiances, or cloud shadows were masked. Only pixels with satellite zenith and solar zenith angles less than 60° and 75°, respectively, were retained, as this geometry corresponds to the limits of reliability for the atmospheric correction process. In generating the "match-up" dataset, temporal coincidence is defined as ±3 h and satellite values are determined as the median of all valid pixels within a 5×5 pixel box centered on the in situ target after outliers are rejected. Finally, to ensure statistical significance and moderate homogeneity, satellite values were rejected when fewer than 50% of non-land pixels in the satellite box were unmasked or when the coefficient of variation of the satellite radiances within this box exceeded 0.15. Additional satellite pixel exclusion criteria and homogeneity tests are graphically described in Fig. 1 of Bailey and Werdell (2006). For this analysis, to remain as close as possible to prevailing remote sensing conditions, we decided to keep 43 nLw spectra, which have negative nLw at 412 nm; the interpretation of the results do not change if these points are excluded.

2.4. *In situ* data

We evaluate our backscattering and CDM anomalies using two regional datasets: the first is from the Bermuda Bio-Optics Project (BBOP) associated with the Bermuda Atlantic Times series Study (BATS), and the second from the Biogeochemistry and Optics South Pacific Experiment (BIOSOPE). The first consists of a five-year time series of surface (we restricted the analysis to the top 20 m of the water column) radiometric and spectrophotometric absorption data that

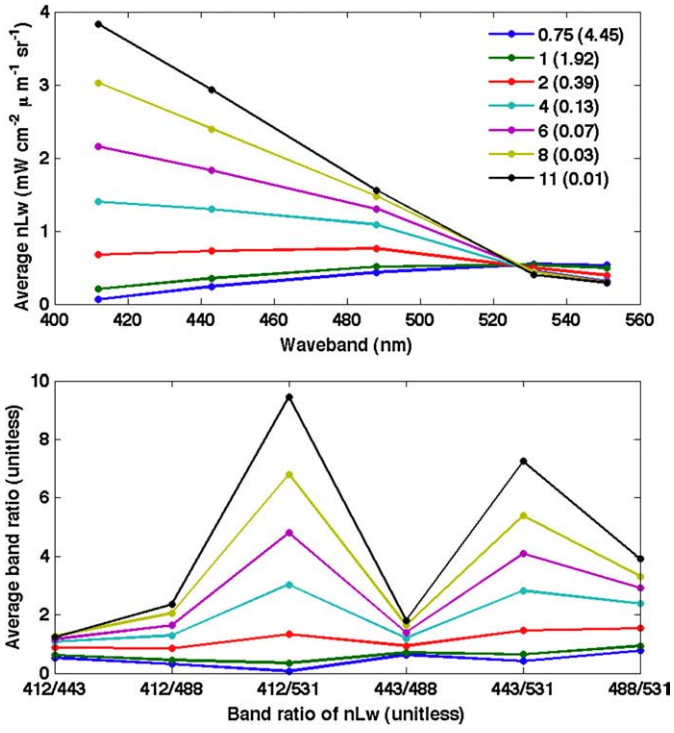

Fig. 2. Examples of average spectra of $n\hat{L}w$ (top panel) and curves for ${}^{i}_{j}\hat{R}_{Lw}$ (bottom panel) for different MBR (see legend in top panel). The numbers in parentheses in the legend represent [Chl] as computed using the OC3M algorithm.

were obtained from the BBOP website. The second is from a transect in the Southeastern Pacific; a more detailed description of this dataset is given in Huot et al. (2008) and further details on the collection and processing of the *in situ* surface backscattering data can be found in Stramski et al. (2007).

We evaluate our CDM absorption algorithm using the IOCCG in situ dataset (IOCCG, 2006). Briefly, it contains 656 stations with matched measurements of [Chl], spectral remote sensing reflectance (we used 555 nm instead of 551 nm as used by MODIS for our computations), and CDM absorption. These measurements are compiled from the work of different researchers throughout the world ocean. They are, however, mostly measured relatively near the coast. For our application, R_{rs} from the dataset was multiplied by the extraterrestrial solar spectrum used by MODIS and obtained from the NASA ocean color website.

3. Results and discussion

3.1. First order variability

The first order variability in ocean color is captured by fitting fifth or sixth order polynomials to nLw and ratios of nLw in the "nLw database" as a function of the MBR (Fig. 1). These fits are respectively represented by $n\hat{L}w_\lambda$ and ${}^{j}_{i}\hat{R}_{Lw}$. From these mean relationships, we can calculate the $n\hat{L}w_\lambda$ spectra (Fig. 2A) and ${}^{j}_{i}\hat{R}_{Lw}$ curves (Fig. 2B) for any MBR. The $n\hat{L}w_\lambda$ spectra illustrate the expected relationships with changing [Chl], and are generally consistent with the trends produced by bio-optical Case 1 models (e.g. Fig. 8 in Morel and Maritorena, 2001).

3.2. Statistical analysis

The first analysis consists of clustering the MODIS anomaly spectra from the "statistical database". Three types of cluster centroids (Fig. 3, left column) are obtained corresponding to "nLw clusters", "band ratio clusters" and "mixed clusters". For comparison, PCA (Fig. 3, middle column) are presented as an alternative examination of variability in the dataset; they are found to be consistently in agreement with the clustering results. The PCA for the "nLw clusters" (Fig. 3B) shows that most of the variability (89%) lies in a spectrally flat component. This is reflected in the clusters as a series of flat spectra of various magnitudes (Fig. 3A); an exception is the uppermost spectrum exhibiting a slope that reveals the effect of the second principal component (Fig. 3B). Fig. 3D shows that relatively flat curves are also representative of the "band ratio clusters", with an exception for ${}^{488}_{531}R_{Lw}$. The corresponding PCA (Fig. 3E) describes 67% of the variability by a flat principal component, also with an exception for ${}^{488}_{531}R_{Lw}$. The second principal component represents 21% of the variability; its effect is most apparent in two centroids in Fig. 3D (second and third curves from the top showing a corresponding inflection at ${}^{412}_{531}R_{Lw}$). The "mixed cluster" centroids (Fig. 3G) combine information from the two previous analyses (i.e. five nLw and six nLw ratios as shown along the abscissa). Except for the uppermost spectrum that consistently shows a positive slope for which the first five points (i.e. 412 to 555 nm) are similar to the top cluster in Fig. 3A, G illustrates the influence of nLw ratios on the "mixed clusters". Indeed the first three data points of each spectrum (i.e. 412, 443, 488 nm) now have a positively or negatively sloping trend compared to the same wavebands in Fig. 3A

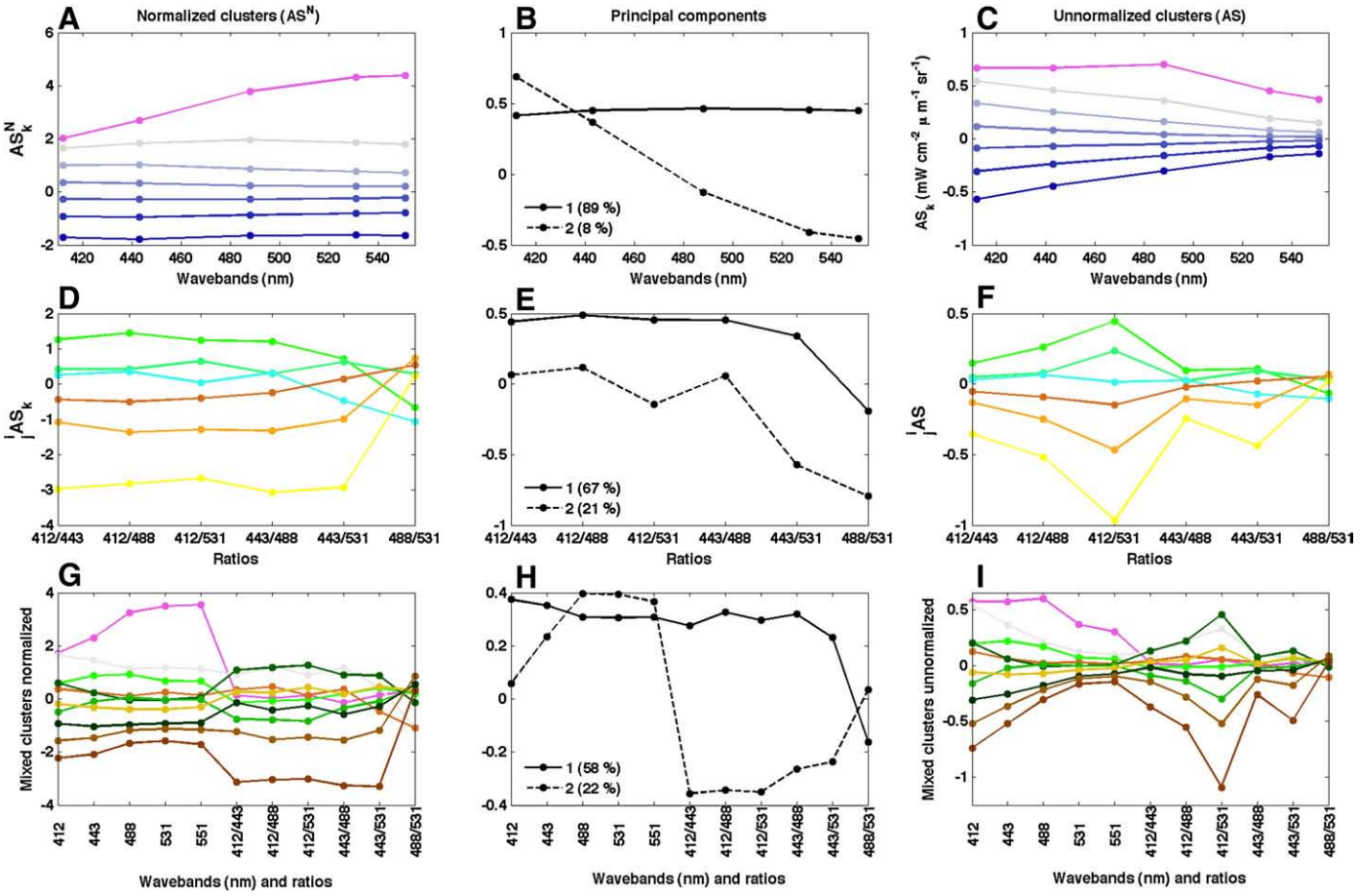

Fig. 3. Three types of clusters and associated PCA for the anomaly spectra. Left column shows the normalized cluster centroids, middle column the principal component weights, and right column the unnormalized centroids. Top row corresponding to "nLw clusters" (seven centroids), middle row to "band ratio clusters" (six centroids), and bottom row to "mixed clusters" (ten centroids). The percentage of variability explained by each principal component is provided in parentheses in the legend for each panel in the middle column.

which formed a flat line. This implies that most of the variability observed in the two types of anomalies (i.e. AS_k^N and ${}_j^iAS^N$) originates from different sources.

3.3. Interpretation of statistical analysis using the modeling results

We now turn to the modeling study to interpret the results of the statistical analysis presented in Section 3.2. Our sensitivity analysis allows us to model anomalies that are expected from variability in different optical components relative to the mean trend. If the shapes and (to a lesser extent) the absolute value of these modeled anomalies reproduce the centroids observed in Section 3.2, it is a strong indication that these anomalies originate from the optical constituent that is varied in the sensitivity analysis. We will focus specifically on the nLw and band ratio anomalies, leaving aside the "mixed clusters".

The first step in order to compare the cluster centroids to the modeling study is to remove the normalization of the clusters (i.e. the normalized centroids are multiplied by the standard deviation and the mean is added; Fig. 3, right column). Since we have already subtracted the best fit by nature of our method, the effect of adding the mean to all bands is very small. In contrast, multiplying by the standard deviation significantly modifies the spectral shapes compared to the normalized centroids (Fig. 3, compare left and right columns).

The striking resemblance between Fig. 4D and 3F leaves little doubt that CDM absorption is the most important effect on the "band

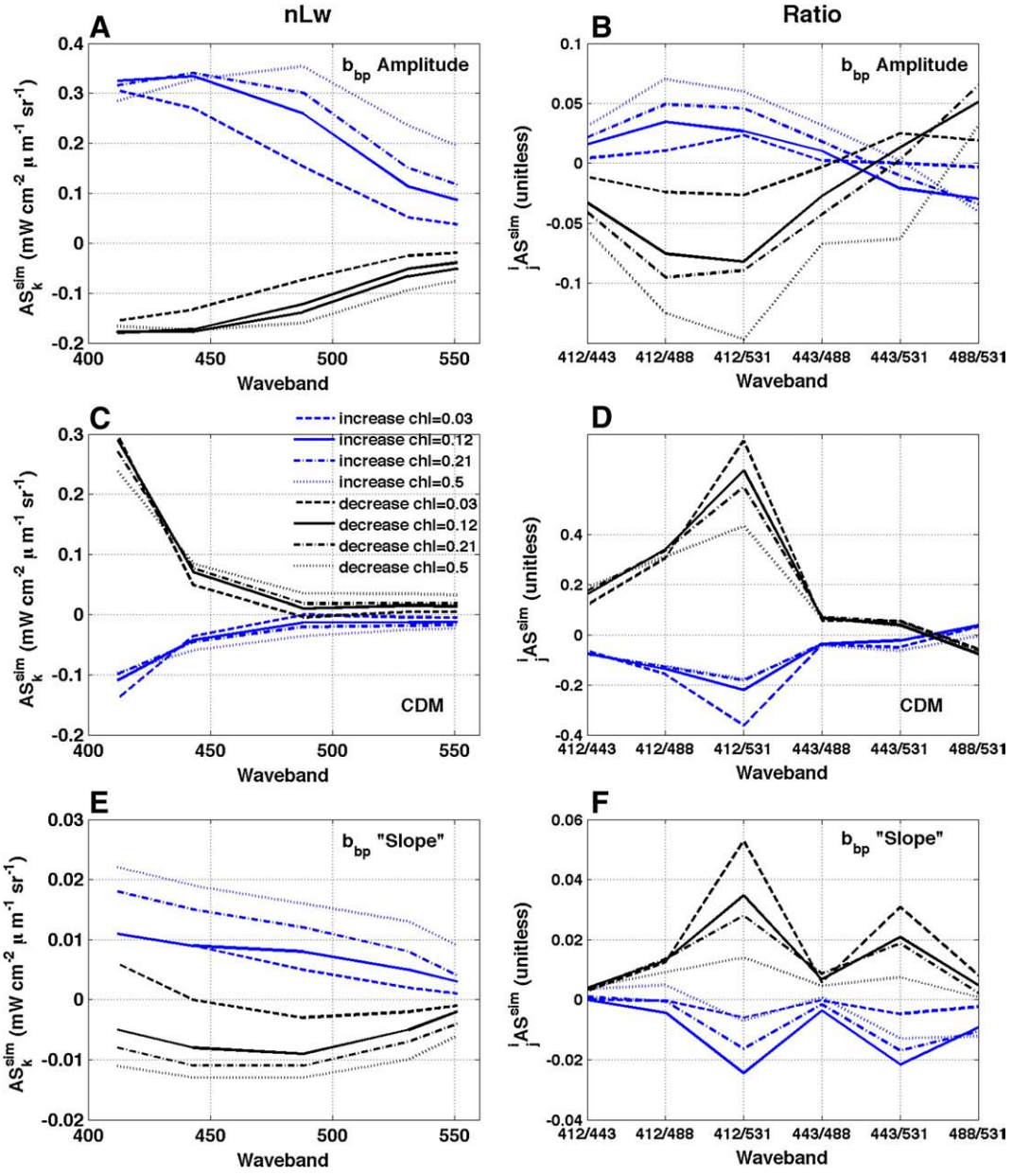

Fig. 4. The modeling analysis demonstrates the effect of different optical components on the amplitude of *nLw* (AS_k^{sim}) and spectral ratios of nLw (${}_j^iAS^{sim}$), as shown in the left and right columns, respectively. Each row corresponds to a different set of departures from the standard model. The top row simulates variations for the b_{bp} amplitude, middle row for the CDM absorption, and bottom row for the slope of b_{bp}. The line type corresponds to different [Chl] (see legend in panel C), while the blue lines indicate an increase of the parameters (whether a_{CDM}, b_{bp}, or the slope of b_{bp}), and the black lines represent a decrease in the same parameters. For each of the three [Chl] cases, we obtain two sub-cases: 1) to examine the effect of CDM absorption, we modify $a_{nw}(\lambda)$ by adding and subtracting $0.5 * a_{nw}(400)\exp(-0.017*(\lambda-400))$; 2) to examine the effect of the backscattering amplitude, we alter b_{bp} by multiplying and dividing the absolute value by 2; and 3) to examine the effect of the backscattering shape, we change the slope parameter for backscattering (v in MM01v07, see Appendix 1) by adding and subtracting 0.5 to this value.

ratio clusters" ($^{i}_{j}AS$). Not only the shapes of the curves but also the absolute values of the anomalous nLw ratios are consistent between the statistical and modeling results. Varying the CDM slope does not change this interpretation; CDM spectra with less negative slopes have more flattened anomaly spectra but retain their distinct spectral signature (Fig. 5).

The effect of the b_{bp} amplitude on the "nLw clusters" (AS_k) is strongly suggested by a comparison between Fig. 4A and 3C. The varying amplitudes of the clusters observed in Fig. 3C can be easily reproduced by the model (not shown) by changing the amount of backscattering added or subtracted from the standard model.

The fact that the shapes of the b_{bp} anomaly spectra and CDM anomaly curves (e.g. Fig. 4A and D) change with [Chl] means that the shapes of the spectra for the cluster centroids are not expected to be "independent" of first order variability. While this feature could be exploited to separate the effects of variable b_{bp} or CDM signatures at different [Chl], herein, we will attempt to avoid these effects.

In Fig. 4E and F, an interesting feature of the modeling results is that despite a relatively large (compare for example with Loisel et al., 2006; Snyder et al., 2008) variation in the slope of b_{bp} (±1) for a given MBR its influence on the spectral nLw anomalies is small. The effect of varying the slope of b_{bp} is generally ten times less than that of the other parameters tested (but becoming more important at longer wavelengths). It is thus not expected that the effect of a variable slope will be detectable from space using our approach; other methods have been proposed to accomplish this (Loisel et al., 2006).

A closer inspection of Figs. 3 and 4 suggests that both the 531 and 551 nm wavebands provide a good measure of the b_{bp} anomaly for the "nLw clusters". The variations in b_{bp} are apparent in the anomalies because their magnitudes are proportional to the addition or removal of backscattering. Therefore, to simplify our approach in the remainder of this paper we will use only the 551 nm waveband (AS_{551}) to represent the b_{bp} anomaly (this waveband, or 555 nm, was also chosen for similar purposes by other authors, e.g. Carder et al., 1999; Stramski et al., 1999; Lee and Hu, 2006; Morel and Gentili, 2008). This band tends to be the least affected by variability in absorption by CDM (e.g. Fig. 4C) or phytoplankton. For the CDM anomaly, both the ratios 412/443 and 412/488 provide consistent measures of the CDM effect for the "band ratio clusters" (Fig. 4D). Note that for all [Chl] we added the same relative amount of CDM absorption to a_{nw} (see Appendix 1), and thus expect a similar effect on the band ratios (and CDM anomaly) for all [Chl]. To simplify the utilization of the CDM anomaly we will use only one ratio; we arbitrarily chose to use the 412/488 ratio.

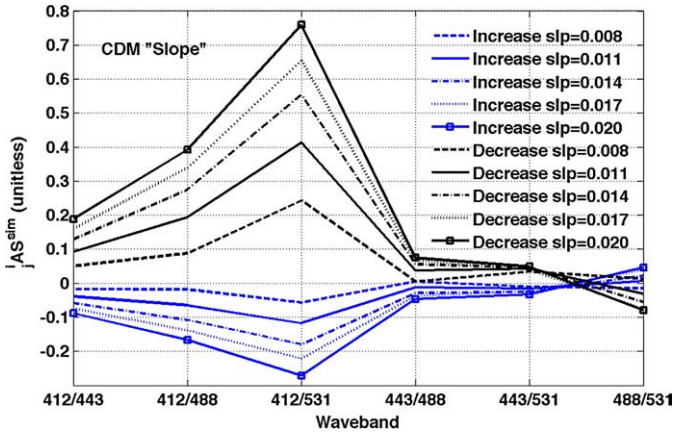

Fig. 5. Modeling results showing the effect of changing the slope of the CDM spectra on the spectral ratios of nLw ($^{i}_{j}AS^{sim}$). The simulation was carried out at 0.12 mg chl m^{-3}. "Increase" and "decrease" in the legend refer to an increase or decrease of the amount of a_{CDM} from the standard model (i.e. $\pm 0.5 * a_{nw}(400)\exp[-\text{slp} * (\lambda - 400)]$, where slp is given in the legend). These results complement the results of Fig. 4D, which were carried out at different [Chl] but for only one slope for CDM (0.017).

3.4. Evaluation

Thus far, the modeling study indicates that the origin of anomalies in global ocean color remote sensing data lies in the variability of two main optical properties for any given [Chl]: 1) the absorption coefficient of colored detrital and dissolved matter, and 2) the backscattering coefficient of particulate matter. This finding is unlikely to surprise optical oceanographers, as these two optical properties have been included in some of the earliest ocean color models in addition to [Chl] (e.g. Sathyendranath et al., 1989).

In the next few sections, we describe a variety of analyses using in situ data that support the results of our statistical and modeling studies, and hence our interpretation of ocean color anomalies in terms of a_{CDM} and b_{bp}. It is more difficult to validate interpretations of second order variability in ocean color compared to first order effects because the signal is considerably weaker and thus requires more accurate measurements.

3.4.1. Global distribution of anomalies

The b_{bp} anomaly represents variability in b_{bp} that is independent of the global mean trends with [Chl]. Indeed, because the mean trend is subtracted, the distribution of positive and negative values must differ from the MBR or [Chl] distributions such that roughly half of the oceanic areas with a given [Chl] will have positive anomalies, the other half, negative anomalies. The temporal changes in the spatial distributions shown in Fig. 6. (left column) display strong seasonal patterns in the mid and high latitudes characterized by the progressive intensification of the b_{bp} anomaly with the local onset of spring/summer and then decreasing throughout the autumn and into winter. The same seasonal progression occurs in both the northern and southern hemispheres. The high b_{bp} anomalies in the northern hemisphere, occurring from approximately June to July in the Atlantic Ocean and June to September in the Pacific Ocean (Fig. 6 shows data for only June and September), could in part be due to coccolithophore blooms. In the austral ocean during summer, a zonal band associated with the Antarctic Convergence displays high b_{bp} anomalies. By contrast, low b_{bp} anomalies are concurrently found in the Ross and Weddell Seas, and west of the Antarctic Peninsula, consistent with in situ studies (regional maps not shown, e.g. Stramski et al., 1999; Dierssen and Smith, 2000; Reynolds et al., 2001). In the Mediterranean Sea, more oligotrophic regions in the southeast tend to show higher b_{bp} anomalies compared to higher chlorophyll regions in the northwest (Claustre et al., 2002). While the nLw anomalies provide a means to view the global distribution and temporal evolution of bulk backscattering, their specific origin (i.e. silica frustules, coccoliths, sea foam, desert dust, sulfur plumes or organic detrital material) cannot be identified using this approach.

The CDM anomalies are shown in Fig. 6 (right column) and represent variability in the proportion of CDM to chlorophyll that differs from the global mean for a given [Chl]. In general, the Pacific Ocean has lower CDM anomalies in the tropics and subtropics compared to the Atlantic Ocean. A seasonal pattern in the northern hemisphere is observed when the CDM anomaly is low coincident with the spring bloom (June in right column of Fig. 6 and 7), particularly in the Atlantic Ocean. The decrease in the CDM anomaly is likely due to algal blooms occurring in waters with a relatively constant CDM background. In the very high latitude northern hemisphere, CDM anomalies appear to be almost continually high. A series of studies have reported that trends in the bio-optical properties of polar waters differ significantly from those of lower latitudes (Mitchell and Holm-hansen, 1991; Mitchell, 1992; Arrigo et al., 1998; Sathyendranath et al., 2001; Cota et al., 2003; Stramska et al., 2003). The difference has often been attributed to highly packaged pigments within large phytoplankton such as diatoms leading to underestimates of [Chl] by global algorithms (e.g. Cota

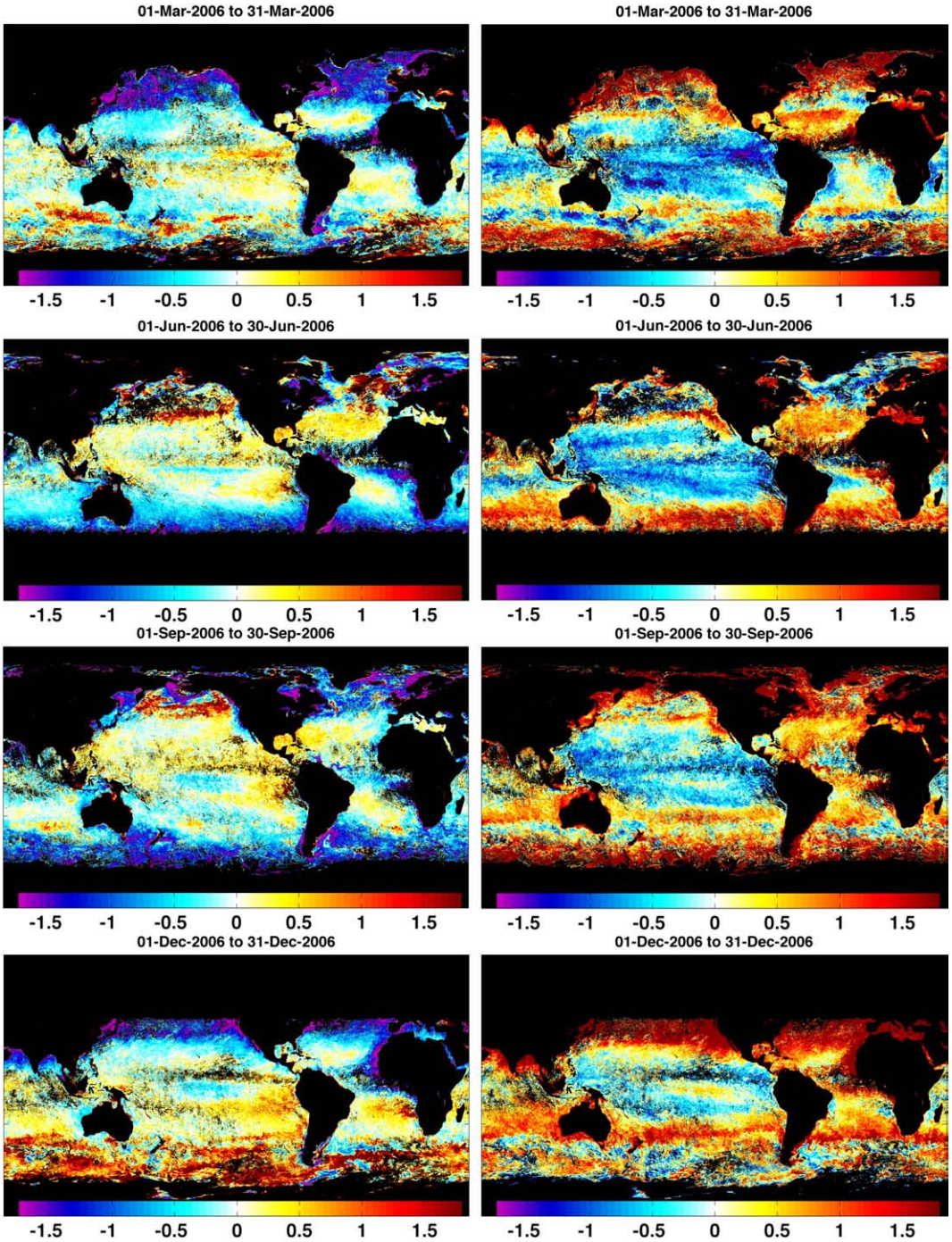

Fig. 6. Global distribution of normalized ocean color anomalies (AS^N). Left and right columns show the backscattering and CDM anomalies, respectively (warmer colors represent higher anomalies). From top to bottom, monthly images for March, June, September and December for the year 2006.

et al., 2003). In addition, the very high proportion of CDM absorption to total absorption has been observed to have a strong effect on ocean color algorithms (Matsuoka et al., 2007). However, while consistent with in situ observations, these anomalies at high latitudes must be verified due to the potential for artifacts in the imagery (see Section 3.4.3).

3.4.2. Evaluation of the CDM anomaly

An evaluation of the CDM anomaly ($^{412}_{488}AS^N$) using data from BBOP in the Sargasso Sea is shown in Fig. 8. This time series is influenced by an annual spring bloom occurring at the end of January or beginning of February, which is represented by the yearly increase in the algal absorption coefficient (Fig. 8A). The CDOM absorption coefficient, by contrast, shows less variability remaining fairly constant throughout the year (Fig. 8A); the first two years of the time series display the most variability in CDOM absorption. The ratio of CDOM to phytoplankton absorption shows a minimum at the beginning of each year and a maximum during the summer (Fig. 8B, this is most clearly shown when in situ sampling was more frequent during the first two years). As shown in Fig. 8C, variability in this ratio is clearly reflected in the 340 to 443 nm spectral ratio of the diffuse attenuation coefficient, once the diffuse attenuation coefficient of pure water has been subtracted (Morel and Maritorena, 2001). While satellite measurements of the CDM anomaly ($^{412}_{488}AS^N$ in

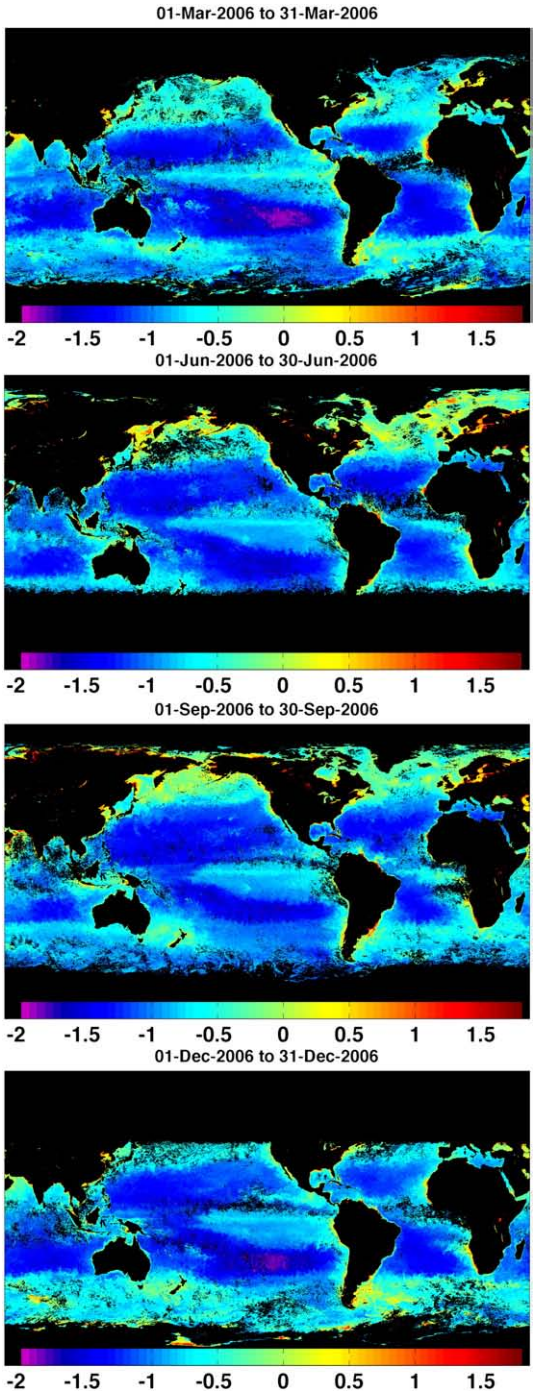

Fig. 7. Global distribution of $\log_{10}$[Chl]. From top to bottom, monthly images for March, June, September and December for the year 2006 (warmer colors represent higher [Chl]).

Fig. 8D) show more variability than in situ measurements, the minimum values (consistent with an interpretation of a higher phytoplankton concentration for a given CDOM absorption) always occur in January and February (except in 2006 when a few data points are missing) when the values are negative. For most of the year ${}^{412}_{488}AS^N$ is positive suggesting a greater amount of CDM for a given [Chl] than the average (${}^{412}_{488}AS^N \equiv 0$) for the global ocean at this location. This trend of negative ${}^{412}_{488}AS^N$ occurring during seasonal algal blooms is also observed in the CDM anomaly maps shown in Fig. 6. (right column), particularly in the northwest Mediterranean Sea (March), North Atlantic Ocean (June) and the Pacific sector of the Southern Ocean (December).

3.4.3. Match-up database

The Aqua match-up database allows us to examine whether ocean color anomalies explain any of the discrepancies observed between algal biomass estimated using ocean color algorithms and biomass measured in situ and if so, whether the discrepancies are consistent with our interpretation of the anomalies. In Fig. 9 (panels A to D), the CDM anomalies are shown to be associated with biases in the estimated [Chl] using the standard MODIS algorithm, OC3M (O'Reilly et al., 2000). We show that "high CDM" waters (panel C) are generally associated with higher estimates of [Chl], whereas "low CDM" waters (panel D) are associated with lower estimates of [Chl]. This result is to be expected as the increased absorption in the blue wavelengths due to CDM has the same effect on the blue-to-green ratio as phytoplankton absorption. As such, the relative concentration of CDM to [Chl] differs from the mean trend upon which the empirical algorithm is based and there should be some error in the estimates. There is no bias for "average waters" (Fig. 9B). Therefore, because the CDM anomaly identifies important variability in ocean color that can be used for identifying "high CDM" and "low CDM" waters, this information can serve as the basis for simple corrections to global empirical algorithms for estimating [Chl] in Case 1 waters. However, due to the limited match-up dataset available for Aqua, it would be premature to propose a correction scheme at this time.

Beyond potential biases in the method, there can be problems in the remote sensing data itself. Despite the fact that a shorter time series is available as compared to SeaWiFS, Level-3 data from the MODIS-Aqua sensor were chosen for this study because they are the least affected by residual sun glint (Morel and Gentili, 2008). Accordingly, we do not see obvious effects of sun glint in the nLw(551) images. As for any global analysis, biases in the remotely sensed data, for example due to errors in the atmospheric correction or very low signal to noise ratio in some bands, will appear as local or regional biases in our estimates of the anomalies. For example, the strong CDM anomaly observed poleward of ~45° in both the northern and southern hemispheres (Fig. 6, right column) during their respective winter seasons requires further investigation especially because of the conspicuously strong variation of this feature with the time of year (i.e. solar zenith angle; not shown).

In Fig. 9, we show that for the match-up database there is no bias in the estimated [Chl] associated with the b_{bp} anomaly for the "average" or "low" backscattering waters (Fig. 9F and H). However, for waters with "high b_{bp}" (Fig. 9G), [Chl] tends to be overestimated. This finding is in agreement with Balch et al. (2005) and Claustre et al. (2002) who found that higher scattering waters tend to be greener. A simple correction scheme for improving estimates of [Chl] using global empirical algorithms could also be developed for high scattering waters using the b_{bp} anomaly data.

It is generally recognized that a complete separation of the effects of CDM and backscattering on the optical characteristics of water can be difficult to achieve (e.g. Roesler and Boss, 2003). We have shown that waters with high CDM anomalies are associated with overestimates of [Chl] using band ratio algorithms such as OC3M. Our method may also include some error because it uses MBR as a reference. In waters with high CDM, the CDM may have a significant influence on MBR which could cause it to depart strongly from the global mean optical relationships, thus influencing the value of $n\hat{L}w_{551}$ and hence AS_{551} upon which the b_{bp} anomaly is based. The expected tendency would be for b_{bp} anomalies to be underestimated in waters with high CDM anomalies resulting in a negative correlation between the two anomalies. For example, this may be the case in the high northern latitudes which at times show high CDM and low b_{bp} anomalies. For the world ocean, there is a small negative covariation between the two anomalies (correlation coefficient, $r = -0.40$, see Fig. 10); it is difficult to know whether this reflects real in situ covariation or if it is a bias in the method. Indeed,

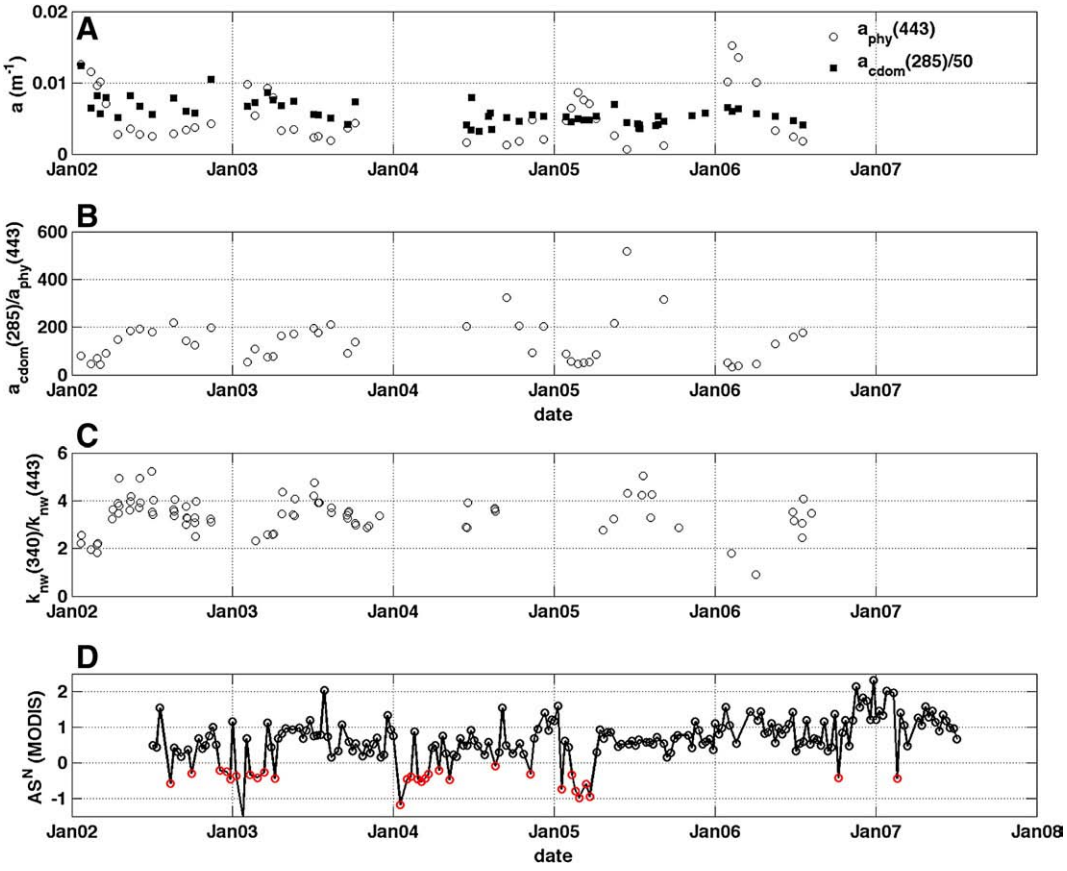

Fig. 8. Data from BBOP. A) The absorption coefficients of phytoplankton at 443 nm and CDOM at 285 nm (both in units of m^{-1}). Note that the CDOM absorption coefficients are divided by 50 to fit the same scale. B) Ratio of CDOM to phytoplankton absorption from panel A. C) Ratio of the diffuse attenuation coefficient at 340 and 443 nm, with the diffuse attenuation coefficient of pure water subtracted D) CDM anomaly, ${}^{412}_{488}AS^N$, from MODIS Aqua data taken at the BATS station. Red circles highlight negative values.

another potential bias in the estimation of the CDM anomaly could be the effect of variability in the b_{bp} amplitude on ${}^{412}_{488}R_{Lw} = nLw_{412}/nLw_{488}$. However, as shown in Fig. 4B a factor of two variability in the b_{bp} amplitude has a rather small effect on ${}^{412}_{488}R_{Lw}$ as compared to variability in CDM (Fig. 4D). Corrections for these second order effects are likely possible using iterative procedures, but are beyond the scope of the present study.

3.4.4. Algorithms for b_{bp} and CDM

In this section we are interested in exploiting the anomalies described above and developing algorithms for b_{bp} and CDM usable for remote sensing. The algorithms that are presented in this section are also provided in a summarized fashion in Appendix 3 to allow an easier implementation.

Empirical algorithms for Case 1 waters describe the mean relationship between the variable of interest (e.g. [Chl] or $K_d(490)$) and spectral ratios of nLw. There is usually significant variability around the mean relationship, but it is most often ignored while developing a functional description of the mean trend. In this study, we focus specifically on two sources of variability: the amount of backscattering and the proportion of CDM to chlorophyll, both for a given [Chl] (where [Chl] is represented by the MBR). It should thus be possible, if average relationships between backscattering and [Chl] (i.e. $b_{bp}(551) = f_1([Chl])$) and between CDM and [Chl] (i.e. $a_{CDM}(412) = g_1([Chl])$) exist, to refine these relationships based on the work described herein. In essence, we want to write

$$b_{bp}(551) = f_1([Chl]) + f_2(AS_{551}) \tag{4}$$

$$a_{CDM}(412) = g_1([Chl]) + g_2\left({}^{412}_{488}AS\right) \tag{5}$$

Note that the second term on the right-hand-side of these equations (f_2 and g_2) may or may not depend on [Chl] (or equivalently on MBR) depending on how the radiance anomaly must be transformed to be expressed in terms of anomalies in $b_{bp}(551)$ or $a_{CDM}(412)$.

We begin by describing an approach for the b_{bp} algorithm (Eq. (4)). Morel and Maritorena (2001) have established the following relationship for b_{bp} vs. [Chl],

$$b_{bp}^{MM01}(551) = 0.347[Chl]^{0.766}\left(\frac{551}{660}\right)^{\upsilon}\{0.002 + 0.01\ [0.50\text{–}0.25log_{10}([Chl])]\}, \tag{6}$$

where the varying exponent υ is expressed as

$$\begin{array}{ll} \upsilon = (1/2)(\log_{10}[Chl]-0.3), & 0.02 < [Chl] < 2mg^{-3}, \\ \upsilon = 0, & [Chl] > 2mg^{-3} \end{array} \tag{7}$$

This relationship was shown to be robust when tested against in situ data (Huot et al., 2008), and can be used in place of $f_1([Chl])$ for the estimation of global b_{bp} values. To use anomalies in an algorithm they need to be "calibrated" in terms of the desired variable, which in this case is $b_{bp}(551)$. We will use a simple approach by considering that $f_2(AS_{551})$ can be described as

$$f_2(AS_{551}) = \frac{AS_{551}\ (f_1([Chl]) + b_{bw})}{n\hat{L}w_{551}} = \frac{AS_{551}\ \left(b_{bp}^{MM01} + b_{bw}\right)}{n\hat{L}w_{551}}, \tag{8}$$

where $n\hat{L}w_{551}$ is the value of our fit for nLw vs. MBR (e.g. Fig. 1). Eq. (8) can be thought of as a relative departure from the mean relationship (i.e. the ratio "$AS_{551}/n\hat{L}w_{551}$") times the total backscattering coefficient (i.e. "$(b_{bp}^{MM01} + b_{bw})$"). This description thus assumes that all

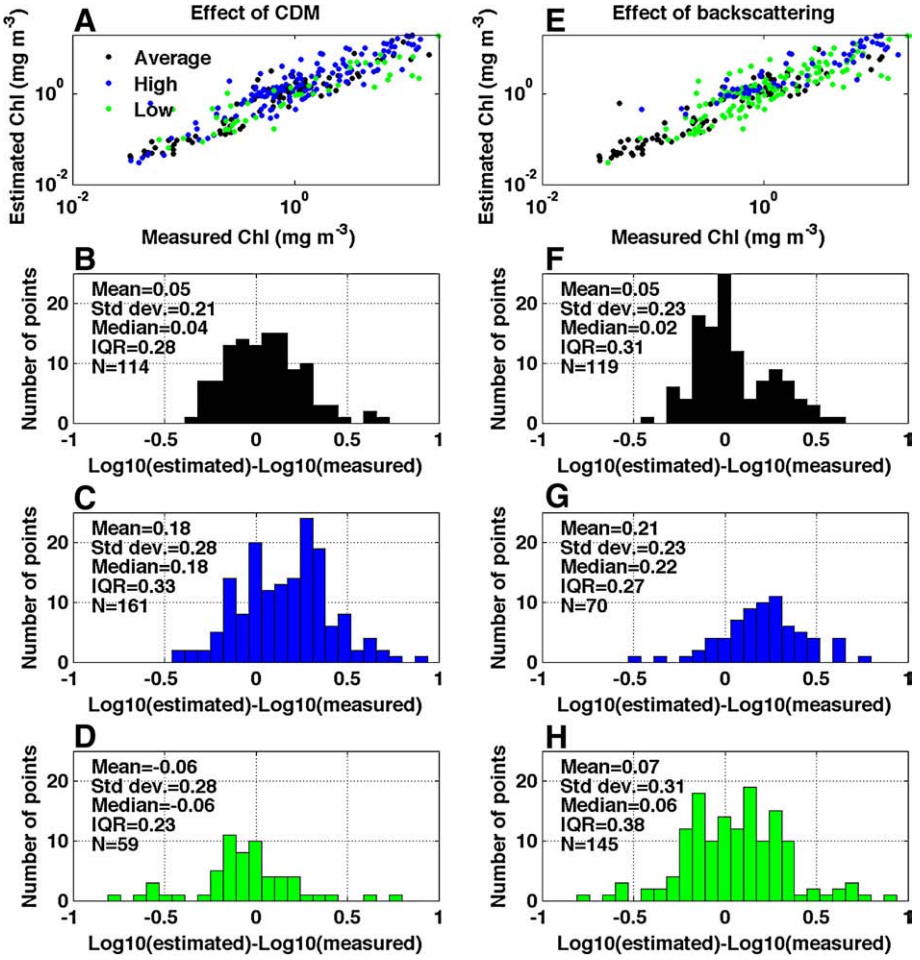

Fig. 9. Comparison with the Aqua match-up database of the measured and estimated [Chl] using the OC3M algorithm. The left and right columns examine respectively the effects of the CDM and b_{bp} anomalies on the retrieval of [Chl]. A) We use the normalized CDM anomaly ($^{412}_{488}AS^{N}$) to separate the points of the Aqua match-up dataset according to their a_{CDM} relative to [Chl]. Values >0.68 were considered "high CDM" (blue dots), <−0.68 "low CDM" (green dots), and between −0.68 and 0.68 "average" (black dots) (in other words, points within one standard deviation of the global mean are considered average). B) A histogram of the difference between [Chl] estimated by OC3M and [Chl] measured in situ is presented for the "average" waters. Similarly, histograms for (C) "high CDM" waters and (D) "low CDM" waters are presented. E) We use the normalized b_{bp} anomaly (AS^{N}_{551}) to separate the points of the Aqua match-up dataset according to their b_{bp} anomalies. The same numerical criteria as in panel A are used to separate high, average and low backscattering points. F) A histogram of the difference between [Chl] estimated by OC3M and [Chl] measured in situ is presented for "average" waters. Similarly, histograms for (G) "high backscattering" waters and (H) "low backscattering" waters are presented.

variability in AS_{551} originates from variability in b_{bp} (i.e. that the absorption anomalies are small at 551 nm; see Appendix 2 for a complete derivation of Eq. (8) and Appendix 3 for a step-by-step description of the algorithm).

Due to the limited in situ data available for b_{bp}, it is difficult to validate b_{bp} estimates retrieved from remote sensing algorithms. We will use the BIOSOPE dataset to qualitatively assess our algorithm (Fig. 11 and 12). Fig. 11A and B (gray points) show the results obtained by two currently used ocean color algorithms, the Garver–Siegel–Maritorena model (GSM, Garver and Siegel, 1997; Maritorena et al., 2002) and the Quasi-Analytical Algorithm (QAA, Lee et al., 2002). In all panels of Fig. 11, a black line is used to represent the best fit between the backscattering coefficient and [Chl] at either 443 (for GSM and QAA) or 551 nm (for this study) as measured in situ along the BIOSOPE transect. In situ measurements are represented as colored points in Fig. 11B for 443 nm (see Huot et al. (2008) for details and other wavelengths). The magenta line represents the b_{bp}^{MM01} model (Eqs. (6) and (7)) of backscattering by particles, that is the "f_1([Chl])" term when the OC3M chlorophyll algorithm is used as input. The scatter around this line in panel C is due to the $f_2(AS_{551})$ term. Both the GSM and QAA models tend to overestimate b_{bp} in the low [Chl] waters of the South Pacific Gyre whereas their accuracy tends to improve at [Chl] above ~0.1 mg m^{-3}. Our algorithm performs well across the entire range of [Chl] despite a slight (~30%) underestimate above [Chl] ~0.4 mg m^{-3}. In this region, AS_{551} is strongly negative (data not shown) and lowers the prediction. It is difficult to say which aspect of our algorithm causes these departures, whether the mean trend $b_{bp}^{MM01}(551)$ is not representative of the average conditions at high [Chl] in the world ocean (but see Huot et al., 2008 which suggests otherwise), AS is incorrectly retrieved in this region due to biases in ocean color (e.g. atmospheric corrections), or if our algorithm simply fails at high [Chl]. It is important to remember that this algorithm relies on the mean trend (in this case taken from Morel and Maritorena, 2001) being representative of the average oceanic trend (i.e. if $b_{bp}^{MM01}(551)=b_{bp}$ in Appendix 1); any departure from the mean oceanic trend will show up as a bias in the algorithm. In any case, the upwelling conditions of these high [Chl] waters resulted in highly dynamic conditions for which a mean trend is not very meaningful, as opposed to the more quiescent Gyre and equatorial waters sampled at lower [Chl] (<~0.4 mg m^{-3}).

In Fig. 11, the retrieval of the mean b_{bp} trend shows essentially that the mean relationship described by Morel and Maritorena (2001) is representative of the BIOSOPE data (compare best fit and b_{bp}^{MM01}. However, this does not mean that the dispersion around the trend is

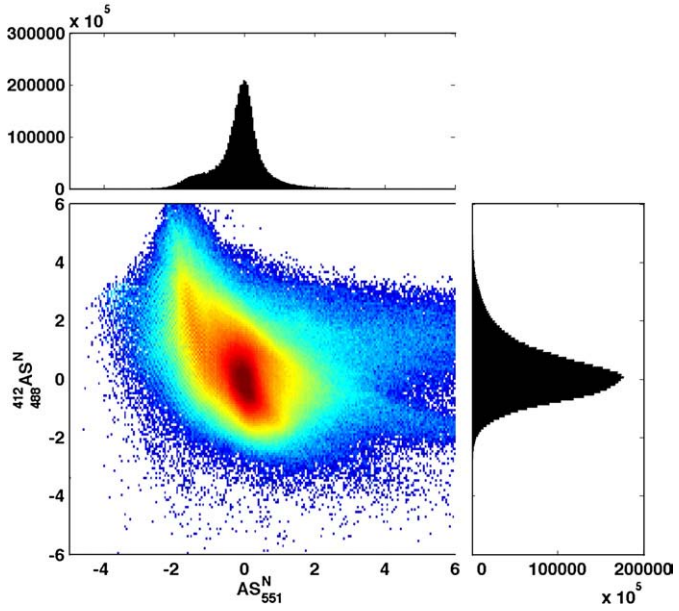

Fig. 10. Covariation between the backscattering (AS^N_{551}) and CDM anomalies ($^{412}_{488}AS^N$). The main panel shows a density plot of the points as a function of the two anomalies. The correlation coefficient is $r=-0.40$. The panels on the top (AS^N_{551}) and right-hand-side ($^{412}_{488}AS^N$) are histograms showing the number of points as a function of the anomalies, roughly 95% of the points fall within values of ±2.

correct as this originates from the f_2(AS) term. The anomalies around the mean trend are shown in Fig. 12 and are used to evaluate the f_2(AS) component of the algorithm. In this figure, differences between the measured b_{bp} and the mean trend with respect to [Chl] are shown as a function of longitude along the BIOSOPE cruise track both for our satellite-based estimates of b_{bp} (Fig. 12A) and the in situ measurements (Fig. 12B). In the case of the satellite data, the $b_{bp}^{MM01}(551)$ line is subtracted from the mean trend thus leaving the $f_2(AS)$ term while for the in situ data a power law fit to the data is used (given in Huot et al., 2008). A fifth order polynomial fit to these anomalies is drawn on both panels to illustrate the trend with longitude observed in these anomalies. This analysis demonstrates the strong similarity between two independent datasets; in both cases, the optical conditions in the center of the South Pacific Gyre (longitude −95 to −135°W) show distinctly lower b_{bp} anomalies than the surrounding waters.

The seasonal progression of global b_{bp} as retrieved with this algorithm is presented in Fig. 14 (left column). The overall spatial distribution is generally consistent with the five-year mean field for b_{bp} produced by the GSM model (see Fig. 2D in Siegel et al., 2005); however, our values in the low chlorophyll regions are considerably lower than their estimates, consistent with the differences observed in Fig. 11. In Fig. 15 (left panel), the dispersion in $b_{bp}(551)$ is shown to be evenly distributed around a continually increasing [Chl]. Note the relatively wide scatter, roughly a factor of 10 around the trend line.

An algorithm for CDM is not as easily formulated and we cannot provide an analytical description of g_2(AS); however, we present here a tentative algorithm for completeness. The mean relationship between [Chl] and CDM has recently been described for the BIOSOPE region by Morel et al. (2007a). Even more recently, Morel (in press) found the following relationship (M08 on Fig. 13A) from an analysis of Case 1 waters:

$$a^{fit}_{CDOM}([Chl],\lambda) = 0.0649[Chl]^{0.63}\exp(-s_{[chl]}[\lambda-400]) \quad (9)$$

where we used $s_{[chl]}=0.018$ nm^{-1} to transfer from 400 to 412 nm for this study. From the IOCCG dataset, by restricting the fit to $-0.2<{}^{412}_{488}AS<0.2$, and thus removing outlier Case 2 waters, we find the following relationship (best fit on Fig. 13A):

$$a^{fit}_{CDOM}([Chl],\lambda) = 0.117[Chl]^{0.774} \quad (9a)$$

To describe a simple algorithm for CDM absorption, we can use either Eqs. (9) or (9a) in place of $g_1(AS)$ in Eq. (5) and add

$$g_2(AS) = -\gamma\frac{{}^{412}_{488}AS\ (g_1([Chl]) + a_w(412))}{{}^{412}_{488}\hat{R}_{Lw}} = -\gamma\frac{{}^{412}_{488}AS\ \left(a^{fit}_{CDOM}([Chl],412) + a_w(412)\right)}{{}^{412}_{488}\hat{R}_{Lw}} \quad (10)$$

where γ is a proportionality constant (see below) and $a_w(412)$= 0.00452 m^{-1} is the absorption coefficient for pure water (Pope and Fry, 1997) at 412 nm. Note that the negative sign in front of the right hand side of Eq. (10) accounts for the fact that an increase in CDM leads to negative ${}^{412}_{488}AS$. This algorithm should be refined as more in situ data becomes available. In particular, more work will be required to constrain the value of γ. A step-by-step description is provided in Appendix 3.

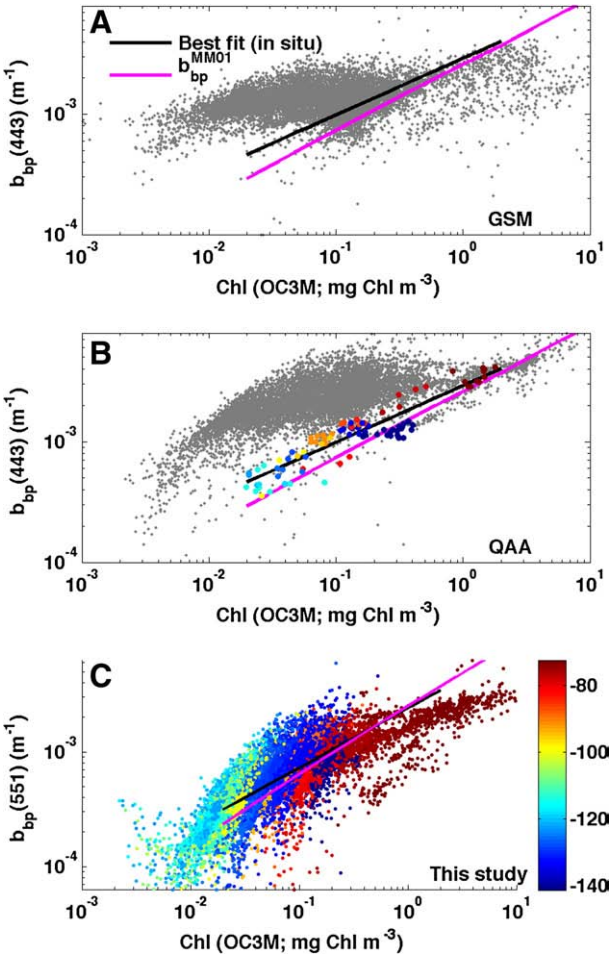

Fig. 11. Results from three b_{bp} algorithms for the BIOSOPE cruise track. A) Results from the GSM algorithm (Maritorena et al., 2002). B) Results from the QAA model (Lee et al., 2002). C) Results from this study based on nLw(551) anomalies (points are colored according to longitude). For all figures, the black line represents the best fit to in situ b_{bp} measurements made during the BIOSOPE cruise (at 443 nm for panels A and B, and 551 nm for panel C), and the magenta line is the representation of backscattering (Eqs. (6) and (7), applied at 443 nm instead of 551 nm for panels A and B). In panel B, the colored points are in situ data at 443 nm measured with a Hydroscat instrument and used to calculate the best fit line (matches the color bar shown in panel C). Panels A and B are modified from the work of Huot et al. (2008).

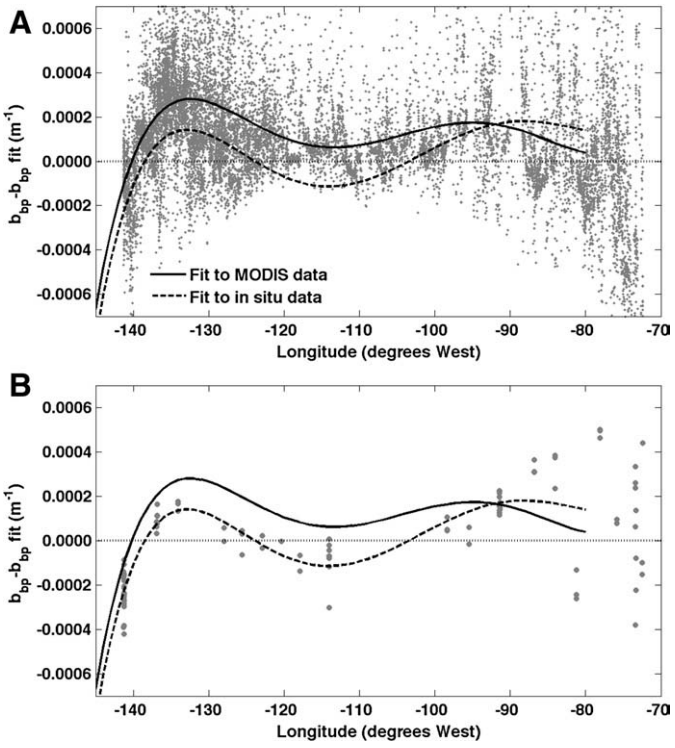

Fig. 12. Differences between b_{bp} and the best fit line as a function of longitude along the BIOSOPE cruise track. A) Difference of b_{bp} retrieved from our algorithm and b_{bp}^{MM01} (satellite data in Fig. 11C minus results of Eqs. (6) and (7) when the OC3M algorithm is used for [Chl]). B) Same as panel A except that the data are only from the Hydroscat instrument (analogous to subtracting the black line in Fig. 11B from the colored points, but for the data from the waveband at 550 nm). The best fit lines are fifth order polynomials to the data points in each panel to illustrate the trends in the anomalies. The best fit for the lower panel (Hydroscat data) is also shown in the upper panel, and vice versa, for comparison.

A first test of the suitability of the algorithm is presented in Fig. 13. We present this analysis as a proof of concept rather than as a rigorous validation or evaluation of the algorithm because of the coastal bias in the IOCCG dataset. This limitation is compounded by the fact that we lose the MODIS-specificity of the algorithm (e.g. inherent corrections for potential calibration and atmospheric correction biases) by using an in situ dataset and not a "match-up dataset". In panel A, the covariation between [Chl] measured in situ and $a_{CDM}(412)$ is presented along with the lines for Eqs. (9) and (9a). We note here that because of the mostly coastal nature of the dataset, the departure of Eq. (9) from the dataset is not surprising. For this "proof of concept", we will use the best fit (Eq. (9a)) for the rest of the analysis of the IOCCG dataset, but will keep Eq. (9) for the global application of the algorithm. The second panel shows the relationship between the measured $a_{CDM}(412)$ and the g_1 ([Chl]) term of the algorithm when [Chl] is estimated using the OC3M algorithm applied to the in situ nLw dataset. We observe that the blue and orange/red points corresponding respectively to low and high $\frac{412}{488}AS$ (high and low CDM to [Chl] ratios) are generally above and below the 1:1 line. The third panel shows the complete algorithm with the $g_2(\frac{412}{488}AS)$ term added. A value of $\gamma=0.4$ was used for $g_2(\frac{412}{488}AS)$ and was obtained by optimizing the r^2 and RMSE values in this panel. We observe an overall tightening of the relationship (reflected in the r^2 and RMSE when compared with the other panels).

While the comparison is unfair, because we used the dataset to fine-tune the algorithm (by changing γ) in this proof of concept, we note that the RMSE obtained (which includes all 656 datapoints) is better than those obtained by all semi-analytical models tested on the same dataset (IOCCG, 2006). The RMSE using only the g_1 term compares well with the best semi-analytical models. At the very least, this shows that there is information in the reflectance spectra that is not fully exploited by semi-analytical models.

For this preliminary analysis, we present the global distribution obtained with this algorithm in Fig. 14 (right column). We remind the reader that Eq. (9) has been used here instead of Eq. (9a) (Eq. (9a) was used in Fig. 13) as it was developed for Case 1 waters. For this analysis we have set the γ constant in the $g_2(\frac{412}{488}AS)$ term to 1. This value provides a dispersion around the mean trend that appears "more reasonable" than using $\gamma=0.4$ which provides very little variability (which could have arisen due to the coastal nature of the dataset). These distributions are consistent with the findings of other remote sensing algorithms (Siegel

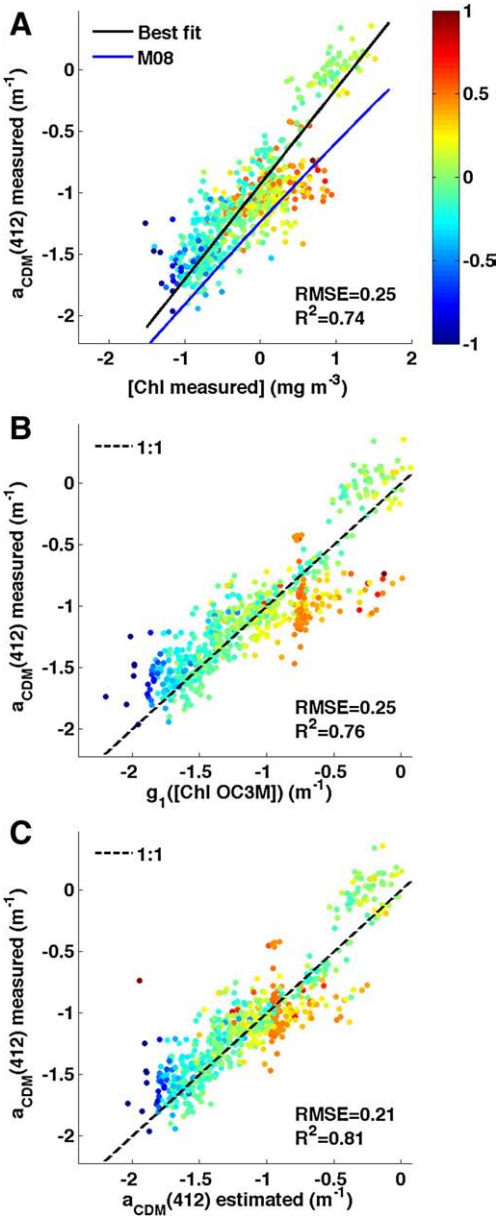

Fig. 13. Evaluation of the CDM algorithm with the IOCCG dataset. A) Measured $a_{CDM}(412)$ as a function of in situ [Chl] measured by HPLC or fluorometrically. The black line is the best fit to the data for which $-0.2<\frac{412}{488}AS<0.2$ (Eq. (9a)). The blue line corresponds to Eq. (9). For all three panels, the points are colored according to $\frac{412}{488}AS$. The RMSE and R^2 given in this panel are for the best fit model. B) Measured $a_{CDM}(412)$ as a function of g_1[Chl OC3M]). Here, g_1[Chl OC3M]) is evaluated by using the [Chl] estimated from the MODIS OC3M algorithm applied on the nLw(λ) measured in situ to evaluate the best fit equation of panel A (Eq. (9a)). C) Measured $a_{CDM}(412)$ as a function of the estimated $a_{CDM}(412)$ using the CDM algorithm (Eqs. (5), (9a) and (10)).

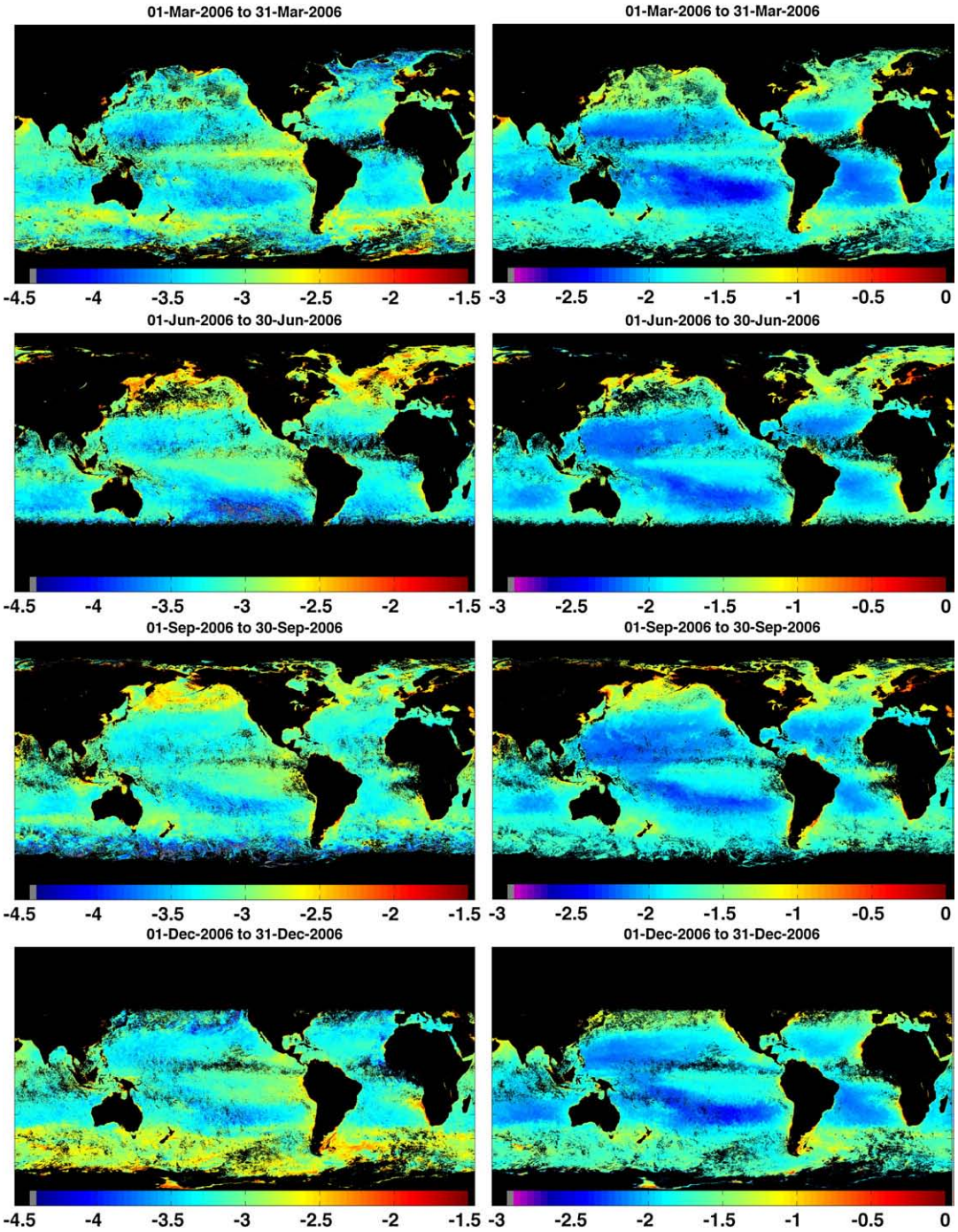

Fig. 14. Global estimates of IOPs. Left column shows b_{bp} ($\log_{10}[b_{bp}(551)]$; $b_{bp}(551)$ in m^{-1}). Right column shows a_{CDM} ($\log_{10}[a_{CDM}(412)]$; $a_{CDM}(412)$ in m^{-1}). From top to bottom, monthly images for March, June, September and December for the year 2006.

et al., 2002, 2005) as well as our current knowledge of CDM distributions. We estimate high CDM values in coastal waters, in the Arctic Ocean, and in the plumes from major rivers such as the Orinoco and Amazon; however, the absolute values of these results remain tentative. Fig. 15 (right panel) illustrates the relationship between $a_{CDM}(412)$ and [Chl] for the month of March 2006 (some variability arises through the year, not shown). It clearly shows the effect of the mean trend with [Chl] on which is superimposed a fair amount of variability. Clearly, the trend with [Chl] has the strongest effects on this algorithm as reflected in the maps in Fig. 14. The dispersion is, however, directly proportional to the constant used in the g_2 term (i.e. $\gamma=1$), and as such remains unverified.

3.4.5. Algal groups and IOPs

Our analysis suggests that the strongest influences on ocean color anomalies are particulate backscattering and the proportion of CDM to chlorophyll. Alvain and colleagues (2005, 2008) also calculate similar anomalies but their objective is to show that spectral changes in nLw are related to phytoplankton accessory pigments and then use the anomalies to identify phytoplankton groups using ocean color. They do not attempt to identify the sources of the anomalies. In general, our statistical and modeling studies (Figs. 3 and 4) suggest that the differences amongst their criteria (their Table 5 in the 2005 publication; compare also Fig. 2 in Alvain et al. 2008, with Figs. 3C and

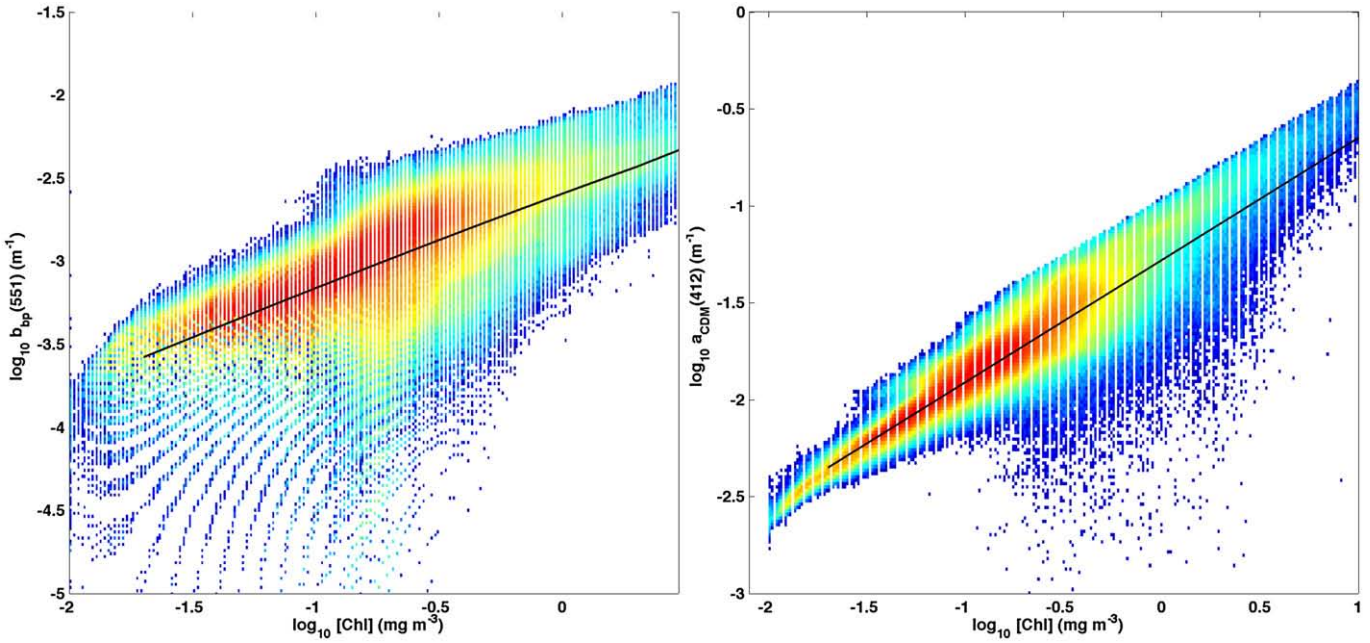

Fig. 15. Estimated relationships for the backscattering coefficient at 551 nm (left panel) and CDM absorption coefficient at 412 nm (right panel) as a function of [Chl] for the month of March 2006. Warmer colors indicate a higher density of points. The black lines correspond to the respective Case 1 representations for $b_{bp}(551)$ and $a_{CDM}(412)$ (Eqs. (6), (7) and (9)).

4A herein) are primarily due to backscattering anomalies, and to a lesser extent to CDM. For example, their criteria for classifying nLw anomalies as diatoms appears to be a description of waters with relatively high backscattering relative to the global mean. The criterion for their haptophyte category (renamed nanoeukaryotes in the 2008 paper) is a description of waters with lower backscattering. Consequently, we note that our maps for the b_{bp} anomaly (Fig. 6, left column) are almost identical to their global maps of phytoplankton assemblages (Fig. 6 in Alvain et al., 2005). Our interpretation in terms of backscattering and CDM may also explain the biases arising in empirical ocean color algorithms (Alvain et al., 2006) when species are used as classifiers. A possible reconciliation of the two interpretations could be imagined under two hypothetical scenarios: 1) if specific phytoplankton groups influence or dictate the bio-optical conditions (mostly backscattering), or 2) if the groups are associated with – but do not cause – specific optical conditions. Both hypotheses remain untested and are difficult to test with current datasets.

4. Summary and conclusion

For Case 1 waters, unique relationships have been defined between [Chl] and ocean color (O'Reilly et al., 1998; Morel and Maritorena 2001) as well as other optical properties (Bricaud et al., 1995; Loisel and Morel 1998; Morel and Loisel 1998). These empirical models were developed from extensive measurements; however, it is well recognized that dispersion occurs around the mean trends for all of these relationships. We identify in this paper the main sources of second order variability in ocean color and provide a method for observing their global distribution.

According to our study, the two principal types of second order variability viewed from space are: 1) nLw anomalies due to particulate backscattering; and 2) nLw band ratio anomalies associated with the proportion of CDM relative to phytoplankton. We show that strong seasonal and regional trends are associated with these anomalies, which can lead to systematic biases in operational algorithms for estimating [Chl]. Several important advantages of using an empirical approach are that the effects of Raman scattering, absorption and backscatter by pure water, and optical constants (e.g. f/Q) are inherently accounted for, thus avoiding the use of various approximations. In addition, our empirical approach is much less sensitive to problems of sensor calibration compared to semi-analytical approaches that use nLw because we use differences from the global mean nLw, which are themselves influenced (or biased) by these same calibration errors. The method also avoids the effects of any systematic errors in the satellite data with [Chl] or MBR because these biases are subtracted when the global mean of nLw is removed. This empirical method for observing ocean color anomalies is an independent approach suitable for comparisons with semi-analytical models. Further identification of the oceanographic and biogeochemical sources of the anomalies is complex and beyond the scope of the present study.

By focusing on the origin and global distribution of ocean color anomalies, we develop a useful observational tool for characterizing the marine optical environment. This is also an essential first step for hypotheses concerning secondary correlations between ocean color anomalies and other properties, such as phytoplankton groups. Regional studies that allow intensive in situ validation are a promising avenue for further testing the use of second order variability in ocean color remote sensing.

Acknowledgements

We gratefully thank David Antoine, Annick Bricaud and André Morel for their insightful comments and suggestions during the preparation of the manuscript. We also thank Heidi Sosik for her thorough and helpful review as well as one anonymous reviewer who helped to improve the paper.

Appendix 1. Modeling study

Our standard representation of Case 1 waters is based on the model described by Morel and Maritorena (2001), Morel and Gentili (2004), and Morel et al. (2007b). As shown schematically in Table A1, the first step is to develop an empirical chlorophyll algorithm that is consistent with this model. To achieve this aim, a series of [Chl] from 0.03 to 3 mg m^{-3} is inserted into the model described by Morel and Maritorena (2001), using the new parameterization for

Table A1
Steps to obtain the simulated anomaly spectra

Step 1: An empirical algorithm for [Chl] vs. MBR that is consistent with MM01v07.
[Chl] series ⇒ MM01v07 → $a_{nw}(\lambda)$, $b_{bp}(\lambda)$ ⇒ Hydrolight → series of nLw(λ)
functional representation of max[nLw(443),nLw(488)]/nLw(551) vs. [Chl] → HOC3
Step 2: Sensitivity analysis conducted using Hydrolight by varying the standard IOPs of MM01v07 to obtain anomalous nLw spectra.

$$[\mathrm{Chl}] = 0.03, 0.12, 0.21, 0.5 \Rightarrow \mathrm{MM01v07}$$
$$\downarrow$$
$$a_{nw}(\lambda), b_{bp}(\lambda) \text{ modified to } \underbrace{\begin{cases} a_{nw}(\lambda) \pm a_{CDM}(\lambda),\ b_{bp}(\lambda) \\ a_{nw}(\lambda),\ b_{bp}(\lambda) \pm \Delta b_{bp} \\ a_{nw}(\lambda),\ b_{bp}(\lambda) \pm \mathrm{slope} \end{cases}}_{\text{Hydrolight} \rightarrow 8\ nLw_{var}(\lambda)}$$

Step 3: Obtaining nLw spectra for the [Chl] (or MBR) that would be retrieved from an empirical chlorophyll algorithm using the anomaly spectra.
8 $nLw_{var}(\lambda)$ ⇒ HOC3 → [Chl] ⇒ MM01v07 → $a_{nw}(\lambda)$, $b_{bp}(\lambda)$ ⇒ Hydrolight → 8 $nLw_{std}(\lambda)$
Step 4: Obtaining simulated anomaly spectra, AS^{sim}, equivalent to those derived from satellite data.
$AS_k^{sim} = nLw_{var} - nLw_{std}$
$_i^jAS^{sim} = {}_i^jR_{var} - {}_i^jR_{std}$ where $_i^jR$ is a spectral ratio of nLw

The single arrow (→) indicates an "output". The double arrow (⇒) means "inserted into". See text for further details.

$K_d(\lambda)$ vs. [Chl] from Morel et al. 2007a (we refer to this model as MM01v07). This provides us with total non-water absorption (a_{nw}) and backscattering (b_{bp}) coefficients (i.e. the inherent optical properties or IOPs). These IOPs are then inserted into a radiative transfer model (Hydrolight, Sequoia Scientific, Inc.) to calculate nLw. The phase functions used are parameterized as in Morel et al. (2002). A functional representation of this set of [Chl] vs. nLw ratios constitutes our "standard chlorophyll algorithm" (herein referred to as the "Hydrolight Ocean Color three-band algorithm" or HOC3). While HOC3 differs slightly from the algorithms proposed in Morel et al. (2007b), it allows us to obtain the [Chl] that would be returned by an empirical algorithm for any MBR (e.g. similar to OC3M). We can thus create an nLw spectrum using a radiative transfer model with any set of IOPs and obtain its [Chl]. This [Chl] can in turn be input into MM01v07 to obtain an nLw spectrum that is consistent with the natural covariation of IOPs included within MM01v07. This step provides two spectra with the same MBR but having otherwise different shapes.

In the second step shown in Table A1, we perform a simple sensitivity analysis by varying the input IOPs for a series of Hydrolight simulations at four selected [Chl] (i.e. 0.03, 0.12, 0.21, 0.5 mg m^{-3}) thus modifying the standard IOPs that were used in MM01v07. We design three cases, each with two sub-cases: 1) to examine the effect of CDM absorption, we modify $a_{nw}(\lambda)$ from MM01v07 by adding and subtracting $0.5*a_{nw}(400)\exp(-0.017*(\lambda-400))$; 2) to examine the effect of the backscattering amplitude, we alter b_{bp} by multiplying and dividing the absolute value by 2; and 3) to examine the effect of the backscattering shape, we change the slope parameter for backscattering (v in MM01v07) by adding and subtracting 0.5 to this value. These Hydrolight simulations provide eight new spectra, called nLw_{var} (the subscript "var" stands for variations), which have slightly different MBR (referred to as MBR_{var}) than those obtained using the standard MM01v07 model. Additionally and in a similar fashion (not shown in Table A1), to examine the effect of varying CDM slopes on the anomalies, the same simulations as above for variable CDM are run but with the slope of the CDM absorption changing from −0.017 to successively: −0.008, −0.011, −0.14 and −0.20.

We now need an nLw reference spectrum with the same MBR as MBR_{var}. As shown in the third step in Table A1, to obtain our reference nLw, which is consistent with MM01v07, we calculate [Chl] from MBR_{var} using HOC3, insert this [Chl] into MM01v07, and retrieve new IOPs that are then inserted into Hydrolight to calculate the "standard" spectrum (nLw_{std}) with an MBR equal to MBR_{var}. In the fourth step, we calculate the difference between nLw_{var} (akin to nLw in satellite imagery) and nLw_{std} (akin to the statistical average relationships computed for $\hat{nLw}$ vs. MBR). In this way, the variable AS_k^{sim} simulates the effect of a diversity of IOPs on the nLw anomalies associated with our clusters. Similarly, $_i^jAS^{sim}$ is calculated to simulate spectral ratios of nLw. Note that AS^{sim} cannot be normalized in a meaningful way as was done in the case of AS to obtain AS^N, these simulations thus need to be compared directly with unnormalized AS spectra.

Appendix 2. Verifying Eqs. (4) and (8)

The particulate backscattering (b_{bp}) and non-water absorption coefficients (a_{nw}) can be expressed as the sum of a mean value ($\overline{b}_{bp}$, $\overline{a}_{nw}$) at a given MBR and a deviation from that mean value (Δb_{bp}, Δa_{nw},)

$$b_{bp} = \overline{b}_{bp} + \Delta b_{bp}$$

$$a_{nw} = \overline{a}_{nw} + \Delta a_{nw}$$

Note that, here, all IOPs are at 551 nm. For a given MBR, we can thus express nLw_{551} as

$$nLw_{551} = \frac{F_o \Re f}{Q} \frac{b_{bp} + b_{bw}}{a_{nw} + a_w} = C_1 \frac{b_{bp} + b_{bw}}{a_{nw} + a_w} = C_1 \frac{\overline{b}_{bp} + \Delta b_{bp} + b_{bw}}{\overline{a}_{nw} + \Delta a_{nw} + a_w}, \quad \text{(A2)}$$

where $C_1 = F_o \Re f / Q$ are optical constants (see Morel et al., 2002 for definitions) and b_{bw} and a_w are respectively the backscattering and absorption coefficients for water. In a similar manner, we can express the result of the fit of nLw_{551} vs. MBR as

$$\hat{nLw}_{551} = \frac{F_o \hat{\Re} \hat{f}}{\hat{Q}} \frac{\overline{b}_{bp} + b_{bw}}{\overline{a}_{nw} + a_w} = C_2 \frac{\overline{b}_{bp} + b_{bw}}{\overline{a}_{nw} + a_w} \quad \text{(A3)}$$

The anomaly spectrum for backscattering (AS_{551}) is equal to:

$$AS_{551} = nLw_{551} - \hat{nLw}_{551} \quad \text{(A4)}$$

Assuming that $C_1 = C_2$ and that Δa_{nw} is negligible relative to $\overline{a}_{nw} + a_w$ and substituting Eqs. (A2) and (A3) into Eq. (A4) we find:

$$AS_{551} = C_2 \frac{\Delta b_{bp}}{\overline{a}_{nw} + a_w} \quad \text{(A5)}$$

By, substituting Eqs. (A3) and (A5) into Eq. (8) and then into Eq. (4). we find:

$$b_{bp} = f_1([\mathrm{Chl}]) + \frac{\Delta b_{bp}}{\overline{b}_{bp} + b_{bw}} \left(b_{bp}^{MM01} + b_{bw} \right)$$

Finally, if we assume that $f_1([\mathrm{Chl}]) = b_{bp}^{MM01}$ and $b_{bp}^{MM01} = \overline{b}_{bp}$ and simplify, we obtain:

$$b_{bp} = b_{bp}^{MM01} + \Delta b_{bp}.$$

Consequently, within the limit of our assumptions, Eqs. (4) and (8) indeed estimate b_{bp}, including the difference from the average relationship.

Appendix 3. Step by step algorithms for $b_{bp}(551)$ and $a_{CDM}(412)$

Algorithm for $b_{bp}(551)$

1. Calculate MBR as $\mathrm{MBR} = \max({}_{551}^{443}R_{Lw}, {}_{551}^{448}R_{Lw})$.
2. Compute AS_{551} by subtracting from each nLw_{551} the oceanic average $\hat{nLw}_{551}$ corresponding to its MBR using the following sixth order polynomial (i.e. $AS_{551} = nLw_{551} - \hat{nLw}_{551}$):

$$\hat{nLw}_{551} = 0.7504 - 0.2779\mathrm{MBR} + 0.07443\mathrm{MBR}^2 - 0.01084\mathrm{MBR}^3 + 8.744 \times 10^{-4}\mathrm{MBR}^4 - 3.609 \times 10^{-5}\mathrm{MBR}^5 + 5.702 \times 10^{-7}\mathrm{MBR}^6$$

Note that this relationship is specific to MODIS Aqua, for application to another ocean color sensor it would be preferable to derive a new relationship.

3. Compute [Chl] using the latest version of OC3M (or another empirical band ratio algorithm of your choice).

4. Compute $b_{bp}(551)$ using Eq. (4) by replacing the term $f_1([\mathrm{Chl}])$ by the right side of Eq. (6) and $f_2(AS_{551})$ by the rightmost version of Eq. (8).

Algorithm for $a_{CDM}(412)$

1. Calculate MBR as $\mathrm{MBR}=\max({}^{443}_{551}R_{Lw}, {}^{448}_{551}R_{Lw})$.

2. Compute ${}^{412}_{488}AS$ by subtracting from each ${}^{412}_{488}R_{Lw}$ the oceanic average ${}^{412}_{488}\hat{R}_{Lw}$ (see Section 2.1.1 for definitions) corresponding to its MBR using the following fifth order polynomial (i.e. ${}^{412}_{488}AS={}^{412}_{488}R_{Lw}-{}^{412}_{488}\hat{R}_{Lw}$):

$$ {}^{412}_{488}\hat{R}_{Lw} = 0.1271 + 0.5300\mathrm{MBR} - 0.09602\mathrm{MBR}^2 + 0.01111\mathrm{MBR}^3 - 4.620 \times 10^{-4}\mathrm{MBR}^4 + 5.774 \times 10^{-5}\mathrm{MBR}^5 $$

Note that this relationship is specific to MODIS Aqua, for application to another ocean color sensor it would be preferable to derive a new relationship.

3. Compute [Chl] using the latest version of OC3M (or another empirical band ratio algorithm of your choice).

4. Compute $a_{CDM}(412)$ using Eq. (5) by replacing the term $g_1([\mathrm{Chl}])$ by the right-hand-side side of Eq. (9) and $g_2({}^{412}_{488}AS)$ by the last version of Eq. (10) and using $\gamma=1$.

References

Alvain, S., Moulin, C., Dandonneau, Y., & Breon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Research Part I-Oceanographic Research Papers*, *52*, 1989–2004.

Alvain, S., Moulin, C., Dandonneau, Y., & Loisel, H. (2008). Seasonal distribution and succession of dominat phytoplankton groups in the global ocean: A satellite view. *Global Biogeochemical Cycles*, *22*. doi:10.1029/2007GB003154

Alvain, S., Moulin, C., Dandonneau, Y., Loisel, H., & Breon, F. M. (2006). A species-dependent bio-optical model of case I waters for global ocean color processing. *Deep-Sea Research Part I-Oceanographic Research Papers*, *53*, 917–925.

Antoine, D., Morel, A., Gordon, H. R., Banzon, V. F., & Evans, R. H. (2005). Bridging ocean color observations of the 1980s and 2000s in search of long-term trends. *Journal of Geophysical Research-Oceans*, *110*. doi:10.1029/2004JC002620

Arrigo, K. R., Robinson, D. H., Worthen, D. L., Schieber, B., & Lizotte, M. P. (1998). Bio-optical properties of the southwestern Ross Sea. *Journal of Geophysical Research-Oceans*, *103*, 21683–21695.

Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing Environment*, *102*, 12–23.

Balch, W. M., Gordon, H. R., Bowler, B. C., Drapeau, D. T., & Booth, E. S. (2005). Calcium carbonate measurements in the surface global ocean based on Moderate-Resolution Imaging Spectroradiometer data. *Journal of Geophysical Research-Oceans*, *110*.

Balch, W. M., Kilpatrick, K. A., & Trees, C. C. (1996). The 1991 coccolithophore bloom in the Central North Atlantic. 1. Optical properties and factors affecting their distribution. *Limnology and Oceanography*, *41*, 1669–1683.

Bricaud, A., Babin, M., Morel, A., & Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. *Journal of Geophysical Research*, *100*, 13321–13332.

Bricaud, A., Bédhomme, A. -L., & Morel, A. (1988). Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation. *Journal of Plankton Research*, *10*, 851–873.

Bricaud, A., Claustre, H., Ras, J., & Oubelkheir, K. (2004). Natural variability of phytoplankton absorption in oceanic waters: influence of the size structure of algal populations. *Journal of Geophysical Research-Oceans*, *109*, C11010. doi:10.1029/2004JC002419

Bricaud, A., Morel, A., Babin, M., Allali, K., & Claustre, H. (1998). Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: analysis and implications for bio-optical models. *Journal of Geophysical Research-Oceans*, *103*, 31033–31044.

Carder, K. L., Chen, F. R., Lee, Z. P., Hawes, S. K., & Kamykowski, D. (1999). Semianalytic moderate-resolution imaging spectrometer algorithm for chlorophyll a concentration and absorption with bio-optical domains based on nitrate-depletion temperatures. *Journal of Geophysical Research*, *104*, 5403–5421.

Carder, K. L., Hawes, S. K., Baker, K. A., Smith, R. C., Steward, R. G., & Mitchell, B. G. (1991). Reflectance model for quantifying chlorophyll-a in the presence of productivity degradation products. *Journal of Geophysical Research-Oceans*, *96*, 20599–20611.

Claustre, H., Morel, A., Hooker, S. B., Babin, M., Antoine, D., Oubelkheir, K., Bricaud, A., Leblanc, K., Queguiner, B., & Maritorena, S. (2002). Is desert dust making oligotrophic waters greener? *Geophysical Research Letters*, *29*.

Cota, G. F., Harrison, W. G., Platt, T., Sathyendranath, S., & Stuart, V. (2003). Bio-optical properties of the Labrador Sea. *Journal of Geophysical Research-Oceans*, *108*.

Dierssen, H. M., & Smith, R. C. (2000). Bio-optical properties and remote sensing ocean color algorithms for Antarctic Peninsula waters. *Journal of Geophysical Research-Oceans*, *105*, 26301–26312.

Garver, S. A., & Siegel, D. A. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation 1. Time series from the Sargasso Sea. *Journal of Geophysical Research*, *102*, 18607–18625.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., & Clark, D. K. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research*, *93*, 10909–10924.

Gordon, H. R., & Morel, A. (1983). *Remote assessment of ocean color for interpretation of satellite visible imagery: A review*. Heidelberg, Germany: Springer Verlag.

Gordon, H. R., & Wang, M. H. (1994). Retrieval of water-leaving radiance and aerosol optical-thickness over the oceans with SeaWiFS — a preliminary algorithm. *Applied Optics*, *33*, 443–452.

Hoge, F. E., & Lyon, P. E. (1996). Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: an analysis of model and radiance measurement errors. *Journal of Geophysical Research-Oceans*, *101*, 16631–16648.

Huot, Y., Morel, A., Twardowski, M. S., Stramski, D., & Reynolds, R. A. (2008). Particle optical backscattering along a chlorophyll gradient in the upper layer of the eastern South Pacific Ocean. *Biogeosciences*, *5*, 495–507.

IOCCG (2000). Remote sensing of ocean colour in coastal, and other optically-complex, waters. In S. Sathyendrath (Ed.), *Reports of the International Ocean-Colour Coordinating Group* Dartmouth: IOCCG.

IOCCG (2006). Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and applications. In Z. P. Lee (Ed.), *Reports of the International Ocean-Colour Coordinating Group* (pp. 126). Dartmouth: IOCCG.

Lee, Z. P., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*, 5755–5772.

Lee, Z. P., & Hu, C. (2006). Global distribution of Case-1 waters: an analysis from SeaWiFS measurements. *Remote Sensing Environment*, *101*, 270–276.

Loisel, H., & Morel, A. (1998). Light scattering and chlorophyll concentration in case 1 waters: a reexamination. *Limnology and Oceanography*, *43*, 847–858.

Loisel, H., Nicolas, J., Deschamps, P., & Frouin, R. (2002). Seasonal and inter-annual variability of particulate organic matter in the global ocean. *Geophysical Research Letters*, *29*, 2196. doi:10.1029/2002GL015948

Loisel, H., Nicolas, J. M., Sciandra, A., Stramski, D., & Poteau, A. (2006). Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *Journal of Geophysical Research-Oceans*, *111*, C09024. doi:10.1029/2005JC003367

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297). Berkeley: University of California Press.

Maritorena, S., Siegel, D. A., & Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, *41*, 2705–2714.

Matsuoka, A., Huot, Y., Shimada, K., & Saitoh, S. -I. (2007). Bio-optical characteristics of the Western Arctic Ocean: implications for ocean color algorithms. *Canadian Journal of Remote Sensing*, *33*, 503–518.

McClain, C. R., Feldman, G. C., & Hooker, S. B. (2004). An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. *Deep Sea Research, II*, *51*, 5–42.

McClain, C. R., Signorini, S. R., Tai, K. -S., Arrigo, K., & Murtugudde, R. (1998). *An ecosystem model for the simulation of physical and biological oceanic processes — IDAPAK User's guide and applications*. Greenbelt: NASA.

Mitchell, B. G. (1992). Predictive biooptical relationships for polar oceans and marginal ice zones. *Journal of Marine Systems*, *3*, 91–105.

Mitchell, B. G., & Holm-hansen, O. (1991). Biooptical properties of Antarctic Peninsula waters — differentiation from temperate ocean models. *Deep-Sea Research Part A-Oceanographic Research Papers*, *38*, 1009–1028.

Morel, A. (1980). In-water and remote measurements of ocean color. *Boundary Layer Meteorology*, *18*, 177–201.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (case 1 waters). *Journal of Geophysical Research-Oceans*, *93*, 10749–10768.

Morel, A. (in press). Are the empirical laws describing the bio-optical properties of Case 1 waters consistent and internally compatible? Journal of Geophysical Research-Oceans.

Morel, A., Antoine, D., & Gentili, B. (2002). Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. *Applied Optics*, *41*, 6289–6306.

Morel, A., Claustre, H., Antoine, D., & Gentili, B. (2007a). Natural variability of bio-optical properties in Case 1 waters: attenuation and reflectance within the visible and near-UV spectral domains, as observed in South Pacific and Mediterranean waters. *Biogeosciences Discussions*, *4*, 2147–2178.

Morel, A., & Gentili, B. (2004). Radiation transport within oceanic (case 1) water. *Journal of Geophysical Research*, *109*, C06008. doi:10.1029/02003JC002259

Morel, A., & Gentili, B. (2008). Practical application of the "turbid water" flag in ocean color imagery: interference with sun-glint contaminated pixels in the open ocean. *Remote Sensing of Environment*, *112*, 934–938.

Morel, A., Huot, Y., Gentili, B., Werdell, J. P., Hooker, S. B., & Franz, B. (2007b). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, *111*, 69–88.

Morel, A., & Loisel, H. (1998). Apparent optical properties of oceanic water: dependence on the molecular scattering contribution. *Applied Optics*, *37*, 4765–4776.

Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: a re-appraisal. *Journal of Geophysical Research-Oceans*, *106*, 7163–7180.

Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, *22*, 709–722.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., & McClain, C. R. (1998). Ocean color chlorophyll algorithm for SeaWiFS. *Journal of Geophysical Research-Oceans*, *103*, 24937–24953.

O'Reilly, J. E., Maritorena, S., Siegel, D. A., O'Brien, M. C., Toole, D., Mitchell, B. G., Kahru, M., Chavez, F. P., Strutton, P., Cota, G., Hooker, S. B., McClain, C. R., Carder, K. L., Muller-Karger, F., Harding, L., Magnuson, A., Phinney, D., Moore, G. F., Aiken, J., Arrigo, K. R., Letelier, R., & Culver, M. (2000). Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. In S. B. Hooker & E.R. Firestone (Eds.), *SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3*, *vol. 11*. (pp. 9–23)Greenbelt, Maryland: NASA, Goddard Space Flight Center.

Patt, F. S., Barnes, R. A., Eplee, R. E., Jr., Franz, B. A., Robinson, W. D., Feldman, G. C., Bailey, S. W., Gales, J., Werdell, P. J., Wang, M., Frouin, R., Stumpf, R. P., Arnone, R. A., Gould, R. W., Jr., Martinolich, P. M., Ransibrahmanakul, V., O'Reilly, J. E., & Yoder, J. (2003). Algorithm updates for the fourth SeaWiFS reprocessing. *NASA. Tech. Memo.*, *vol. 206892* (pp. 74). Greenbelt: NASA Goddard Space Flight Center.

Pope, R. M., & Fry, E. S. (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. *Applied Optics*, *36*, 8710–8723.

Reynolds, R. A., Stramski, D., & Mitchell, B. G. (2001). A chlorophyll-dependent semianalytical reflectance model derived from field measurements of absorption and backscattering coefficients within the Southern Ocean. *Journal of Geophysical Research-Oceans*, *106*, 7125–7138.

Roesler, C. S., & Boss, E. (2003). Spectral beam attenuation coefficient retrieved from ocean color inversion. *Geophysical Research Letters*, *30*, 1468. doi:10.1029/2002GL016185

Roesler, C. S., & Perry, M. J. (1995). In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance. *Journal of Geophysical Research*, *100*, 13279–13294.

Sathyendranath, S., Cota, G., Stuart, V., Maass, H., & Platt, T. (2001). Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. *International Journal of Remote Sensing*, *22*, 249–273.

Sathyendranath, S., Prieur, L., & Morel, A. (1989). A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters. *International Journal of Remote Sensing*, *10*, 1373–1394.

Siegel, D. A., Maritorena, S., & Nelson, N. B. (2005). Independence and interdependencies among global ocean color properties: reassessing the bio-optical assumption. *Journal of Geophysical Research-Oceans*, *110*.

Siegel, D. A., Maritorena, S., Nelson, N. B., Hansell, D. A., & Lorenzi-Kayser, M. (2002). Global distribution and dynamics of colored dissolved and detrital organic materials. *Journal of Geophysical Research*, *107*, 3228.

Snyder, W. A., Arnone, R. A., Davis, C. O., Goode, W., Gould, R. W., Ladner, S., Lamela, G., Rhea, W. J., Stavn, R., Sydor, M., & Weidemann, A. (2008). Optical scattering and backscattering by organic and inorganic particulates in U.S. coastal waters. *Applied Optics*, *47*, 666–677.

Stramska, M., Stramski, D., Hapter, R., Kaczmarek, S., & Ston, J. (2003). Bio-optical relationships and ocean color algorithms for the north polar region of the Atlantic. *Journal of Geophysical Research-Oceans*, *108*.

Stramski, D., Bricaud, A., & Morel, A. (2001). Modeling the inherent optical properties of the ocean based on the detailed composition of the planktonic community. *Applied Optics*, *40*, 2929–2945.

Stramski, D., Reynolds, R. A., Kahru, M., & Mitchell, B. G. (1999). Estimation of particulate organic carbon in the ocean from remote sensing. *Science*, *285*, 239–242.

Stramski, D., Reynolds, R. A., Babin, M., Kaczmarek, S., Lewis, M. R., Röttgers, R., Sciandra, A., Stramska, M., Twardowski, M. S., Franz, B. A., & Claustre, H. (2007). Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. *Biogeosciences*, *5*, 171–201.