JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. C11, PAGES 24,937–24,953, OCTOBER 15, 1998

Ocean color chlorophyll algorithms for SeaWiFS

John E. O'Reilly,[1] Stéphane Maritorena,[2] B. Greg Mitchell,[3] David A. Siegel,[4] Kendall L. Carder,[5] Sara A. Garver,[6] Mati Kahru,[3] and Charles McClain[7]

Abstract. A large data set containing coincident in situ chlorophyll and remote sensing reflectance measurements was used to evaluate the accuracy, precision, and suitability of a wide variety of ocean color chlorophyll algorithms for use by SeaWiFS (Sea-viewing Wide Field-of-view Sensor). The radiance-chlorophyll data were assembled from various sources during the SeaWiFS Bio-optical Algorithm Mini-Workshop (SeaBAM) and is composed of 919 stations encompassing chlorophyll concentrations between 0.019 and 32.79 $\mu g\ L^{-1}$. Most of the observations are from Case I nonpolar waters, and ~20 observations are from more turbid coastal waters. A variety of statistical and graphical criteria were used to evaluate the performances of 2 semianalytic and 15 empirical chlorophyll/pigment algorithms subjected to the SeaBAM data. The empirical algorithms generally performed better than the semianalytic. Cubic polynomial formulations were generally superior to other kinds of equations. Empirical algorithms with increasing complexity (number of coefficients and wavebands), were calibrated to the SeaBAM data, and evaluated to illustrate the relative merits of different formulations. The ocean chlorophyll 2 algorithm (OC2), a modified cubic polynomial (MCP) function which uses Rrs490/Rrs555, well simulates the sigmoidal pattern evident between log-transformed radiance ratios and chlorophyll, and has been chosen as the at-launch SeaWiFS operational chlorophyll *a* algorithm. Improved performance was obtained using the ocean chlorophyll 4 algorithm (OC4), a four-band (443, 490, 510, 555 nm), maximum band ratio formulation. This maximum band ratio (MBR) is a new approach in empirical ocean color algorithms and has the potential advantage of maintaining the highest possible satellite sensor signal:noise ratio over a 3-orders-of-magnitude range in chlorophyll concentration.

1. Introduction

The influence of phytoplankton on the color of seawater has been studied for several decades. It is well understood that chlorophyll *a*, the primary photosynthetic pigment in phytoplankton, absorbs relatively more blue and red light than green, and the spectrum of backscattered sunlight or color of ocean water progressively shifts from deep blue to green as the concentration of phytoplankton increases [e.g., *Yentsch*, 1960]. Following successful high-altitude aircraft studies relating ocean color to chlorophyll concentration [*Clark et al.*, 1970; *Hovis*, 1981], satellite ocean color research began in the late 1970s with the coastal zone color scanner (CZCS) aboard the Nimbus 7 satellite which acquired data from October 1978 to June 1986 [*Evans and Gordon*, 1994; *Acker*, 1994]. Because phytoplankton are the major contributor to ocean color in offshore water, the passive remote measurements of the CZCS over the oceans were successfully used to quantify in situ phytoplankton chlorophyll concentrations. The CZCS ocean color data profoundly enriched our understanding of the global distribution of phytoplankton by providing a synoptic, spatially and temporally cohesive picture of phytoplankton biomass variability only partially resolved by previous shipboard sampling [*Yoder et al.*, 1988; *Feldman et al.*, 1989; *Aiken et al.*, 1992; *McClain*, 1993; *Yoder et al.*, 1993; *Mitchell*, 1994]. Moreover, combining CZCS data with shipboard data and other satellite measurements, such as sea surface temperature from the advanced very high resolution radiometer (AVHRR), provided insights into linkages between physical and biological oceanographic properties [e.g., *Sathyendranath et al.*, 1991; *Denman and Abbott*, 1994] and permitted satellite-based estimates of regional and global phytoplankton primary production [e.g., *Smith et al.*, 1982; *Campbell and O'Reilly*, 1988; *Platt et al.*, 1991; *Longhurst et al.*, 1995; *Antoine et al.*, 1996; *Behrenfeld and Falkowski*, 1997]. Satellite ocean color data provide the only practical means for monitoring the spatial and seasonal variations of near-surface phytoplankton, information essential for the study of oceanic primary production, global carbon and other biogeochemical cycles, as well as fisheries research.

More than a decade after the end of the pioneer CZCS mission, a new generation of ocean color sensors is emerging (Table 1). These new sensors have more wave bands and higher precision and are designed to avoid some of the limitations of the CZCS [*Hooker et al.*, 1993]. Along with improved sensors, improvements in bio-optical algorithms are required for making accurate estimates of chlorophyll *a* from satellite radiance data. Such improvements are expected to enhance the accuracy of global ocean phytoplankton biomass assessments.

Since the 1970s, a variety of bio-optical algorithms have been

[1]NOAA National Marine Fisheries Service, Narragansett, Rhode Island.
[2]Universities Space Research Association, NASA Goddard Space Flight Center, Greenbelt, Maryland.
[3]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California.
[4]Department of Geography, University of California, Santa Barbara.
[5]University of South Florida, St. Petersburg.
[6]California State Polytechnic University, Pomona.
[7]NASA Goddard Space Flight Center, Greenbelt, Maryland.

Paper number 98JC02160.
0148-0227/98/98JC-02160$09.00

Table 1. Center Wave Bands for Historical and Contemporary Ocean Color Sensors (400–700 nm Range)

Band, nm	CZCS	SeaWiFS	OCTS	POLDER	MODIS	MERIS
412		1	1		8	1
443	1	2	2	1	9	2
490		3	3	2	10	3
510		4				4
520	2		4			
530					11	
550	3				12	
555		5				
560						5
565			5	3		
620						6
665			6		13	7
670	4	6				
678					14	
682						8

See notation section for acronym definitions.

developed to estimate chlorophyll *a* (C) or chlorophyll *a* + phaeopigments ([C + P]) concentration from ocean radiance data. Most of these are empirical equations derived by statistical regression of radiance versus chlorophyll. Advances in various theoretical studies and new parameterizations of some optical properties have yielded better knowledge of the marine light field and have provided new tools for modeling ocean color [e.g., *Gordon et al.*, 1988; *Morel*, 1988; *Sathyendranath et al.*, 1989; *Bricaud et al.*, 1995]. The emergence and development of semianalytic (or semiempirical) ocean color algorithms largely result from these improvements in understanding the relationship between remote sensing reflectance (Rrs) and backscattering to absorption ratio [*Morel and Prieur*, 1977; *Carder et al.*, 1986]. Semianalytic algorithms use analytical, optical, Rrs models that can be inverted to derive chlorophyll, absorption coefficients of other optically active components in the water, such as gelbstoff, or the backscattering coefficient b_b. Empiricism is involved in the parameterization of several terms used in these models (e.g., backscattering, chl *a*-specific absorption coefficient, spectral shapes of detrital absorption). This admixture of theory and empiricism is the reason the term semianalytic has been applied [e.g., *Gordon and Morel*, 1983] to describe such algorithms.

Despite these advances, the development and evaluation of the accuracy and precision of ocean color chlorophyll algorithms has been impeded by the limited number and geographic distribution of simultaneous in situ radiance and chlorophyll data and the even smaller number of in situ measurements coincident with satellite data [*Gordon et al.*, 1983; *Balch et al.*, 1992]. For example, the empirical algorithm, widely applied in the processing of the global CZCS data set [*Gordon et al.*, 1983; *Feldman et al.*, 1989; *Evans and Gordon*, 1994], was derived from the Nimbus Experiment Team radiance-chlorophyll data set [*Acker*, 1994] which contains less than 60 stations.

In January 1997, NASA convened a small working group (SeaWiFS Bio-optical Algorithm Mini-Workshop; hereinafter referred to as SeaBAM) whose primary goal was the identification of chlorophyll *a* (C) and chlorophyll *a* + phaeopigments ([C + P]) algorithms suitable for operational use by SeaWiFS [*Firestone and Hooker*, 1998]. Such algorithms are expected to encompass accurately a large diversity of bio-optical conditions since they will be used routinely to process data at the global scale. To achieve this goal, a large, globally representative evaluation data set, the SeaBAM data, was compiled from various sources, and criteria for objective evaluation of algorithms were developed. The SeaBAM activity also provided an opportunity to evaluate and compare chlorophyll *a* and [C + P] algorithms from past (CZCS), current (OCTS, POLDER), and near-future (MODIS) sensors. Such an evaluation would also provide useful information for the recently initiated ocean color satellite intercomparison studies such as SIMBIOS (Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies) [*Esaias et al.*, 1995].

In this paper we report on the results of the evaluation of 17 algorithms tested using the SeaBAM data. The composition and characteristics of the SeaBAM data and criteria used in the evaluation of these algorithms are described. The relative merits of various algorithm formulations, tuned to SeaBAM data, are also presented, and their suitability for operational use by SeaWiFS and compatibility with past ocean color data are discussed.

2. Algorithms

Two semianalytic models and 15 empirical equations were evaluated (Table 2). Some of the algorithms require Rrs, and others require normalized water-leaving radiance Lwn. The equations presented in this paper reflect the versions of the algorithms in April 1997 at the conclusion of the SeaBAM activity [*Firestone and Hooker*, 1998], except for the global versions of the Carder and Garver-Siegel models which were parameterized and evaluated following this workshop. This diverse collection of algorithms is briefly described below.

2.1. Semianalytic Models

The Carder model (K. L. Carder et al., Semianalytic MODIS algorithms for chlorophyll and absorption with bio-optical domains based on nitrate depletion temperatures, submitted to *Journal of Geophysical Research*, 1998), is a semianalytic algorithm based on the $b_b/(a + b_b)$ to Rrs relationship [*Gordon et al.*, 1988]. It uses the Rrs at four SeaWiFS wavelengths to derive the absorption coefficient of phytoplankton at 675 nm, $a_{ph}(675)$, and the absorption coefficient of colored dissolved organic matter (CDOM) at 400 nm, $a_g(400)$. Chlorophyll *a* concentration is then calculated from an empirical relationship between $a_{ph}(675)$ and chlorophyll *a*. A default, two-wavelength empirical algorithm (Rrs490/Rrs555) is used when $a_{ph}(675)$ is outside a predetermined search range. Two versions of the Carder model were evaluated: an initial version parameterized for subtropical, unpackaged pigment data and a second version parameterized for more packaged pigments and global application [*Carder et al.*, 1998].

The Garver/Siegel model [*Garver and Siegel*, 1997] is a semianalytic algorithm based on the quadratic form of the $b_b/(a + b_b)$ to Rrs relationship. The model uses predefined shapes for specific absorption and backscattering coefficients to derive, through a nonlinear statistical method, the chlorophyll *a* concentration, the absorption coefficient due to phytoplankton at 441 nm, $a_{ph}(441)$, the absorption coefficient due to other particulate and dissolved matter at 441 nm, $a_{dm}(441)$, and the backscattering coefficient of particles at the same wavelength, $b_{bp}(441)$. The model was initially tested against data from the

Table 2. Empirical Algorithms

Algorithm	Type	Result Equation(s)	Band Ratio (R), Coefficients (a)	Reference
Global processing (GPs)	power	$C_{13} = 10^{(a0+a1*R1)}$ $C_{23} = 10^{(a2+a3*R2)}$ [C + P] = C_{13}; if C_{13} and $C_{23} > 1.5$ μg L^{-1} then [C + P] = C_{23}	$R1$ = log(Lwn443/Lwn550) $R2$ = log(Lwn520/Lwn550) a = [0.053, −1.705, 0.522, −2.440]	1
Clark three-band (C3b)	power	$[C + P] = 10^{(a0+a1*R)}$	R = log((Lwn443 + Lwn520)/Lwn550) a = [0.745, −2.252]	2
Aiken-C	hyperbolic + power	$C_{21} = \exp(a0 + a1*\ln(R))$ $C_{23} = (R + a2)/(a3 + a4*R)$ $C = C_{21}$; if $C < 2.0$ μg L^{-1} then $C = C_{23}$	R = Lwn490/Lwn555 a = [0.464, −1.989, −5.29, 0.719, −4.23]	3
Aiken-P	hyperbolic + power	$C_{22} = \exp(a0 + a1*\ln(R))$ $C_{24} = (R + a2)/(a3 + a4*R)$ [C + P] = C_{22}; if [C + P] < 2.0 μg L^{-1} then [C + P] = C_{24}	R = Lwn490/Lwn555 a = [0.696, −2.085, −5.29, 0.592, −3.48]	3
OCTS-C	power	$C = 10^{(a0+a1*R)}$	R = log((Lwn520 + Lwn565)/Lwn490) a = [−0.55006, 3.497]	4
OCTS-P	multiple regression	$[C + P] = 10^{(a0+a1*R1+a2*R2)}$	$R1$ = log(Lwn443/Lwn520) $R2$ = log(Lwn490/Lwn520) a = [0.19535, −2.079, −3.497]	5
POLDER	cubic	$C = 10^{(a0+a1*R+a2*R^2+a3*R^3)}$	R = log(Rrs443/Rrs565) a = [0.438, −2.114, 0.916, −0.851]	6
CalCOFI two-band linear	power	$C = 10^{(a0+a1*R)}$	R = log(Rrs490/Rrs555) a = [0.444, −2.431]	7
CalCOFI two-band cubic	cubic	$C = 10^{(a0+a1*R+a2*R^2+a3*R^3)}$	R = log(Rrs490/Rrs555) a = [0.450, −2.860, 0.996, −0.3674]	7
CalCOFI three-band	multiple regression	$C = \exp(a0 + a1*R1 + a2*R2)$	$R1$ = ln(Rrs490/Rrs555) $R2$ = ln(Rrs510/Rrs555) a = [1.025, −1.622, −1.238]	7
CalCOFI four-band	multiple regression	$C = \exp(a0 + a1*R1 + a2*R2)$	$R1$ = ln(Rrs443/Rrs555) $R2$ = ln(Rrs412/Rrs510) a = [0.753, −2.583, 1.389]	7
Morel-1	power	$C = 10^{(a0+a1*R)}$	R = log(Rrs443/Rrs555) a = [0.2492, −1.768]	8
Morel-2	power	$C = \exp(a0 + a1*R)$	R = ln(Rrs490/Rrs555) a = [1.077835, −2.542605]	9
Morel-3	cubic	$C = 10^{(a0+a1*R+a2*R^2+a3*R^3)}$	R = log(Rrs443/Rrs555) a = [0.20766, −1.82878, 0.75885, −0.73979]	9
Morel-4	cubic	$C = 10^{(a0+a1*R+a2*R^2+a3*R^3)}$	R = log(Rrs490/Rrs555) a = [1.03117, −2.40134, 0.3219897, −0.291066]	9

References: 1, *Evans and Gordon* [1994]; 2, *Muller-Karger et al.* [1990]; D. Clark; *McClain and Yeh* [1994]; 3, *Aiken et al.* [1995]; 4, Science on the GLI Mission, p. 16; Ocean Optics XIII, Halifax, October 1996; 5, Ocean Optics XIII, Halifax, October 1996; personal communication to C. McClain, NASA; 6, A. Bricaud, personal communication to S. Maritorena; 7, *Mitchell and Kahru* [1998]; 8, Ocean Optics XIII, Halifax, October 1996; 9, A. Morel, personal communication to S. Maritorena.

Sargasso Sea [*Garver and Siegel*, 1997]. Recent developments included the use of $a_w(\lambda)$ values from *Pope and Fry* [1997] instead of *Smith and Baker*'s [1981], the chlorophyll-specific phytoplankton absorption spectra of *Morel* [1988] instead of that from *Bricaud et al.* [1995], and a different value for the exponential decay constant of the detrital and dissolved absorption. Details of adjustments made to the model for the SeaBAM intercomparisons are provided by *Garver* [1997].

2.2. Empirical Models

Most CZCS-pigment estimates have been made using the global processing switching (GPs) algorithm [*Gordon et al.*, 1983; *Feldman et al.*, 1989; *Evans and Gordon*, 1994] which uses Lwn443/Lwn550 at concentrations below ~1.5 μg L^{-1} and switches to Lwn520/Lwn550 above 1.5 μg L^{-1}, when the former band ratio gets too low (Table 2). The Clark three-band (C3b) [*Muller-Karger et al.*, 1990] uses the same bands as the GPs but avoids band switching by summing the 443 and 520 channels, thereby compensating for the weakness of the 443 nm band at high pigment concentrations. The Aiken hyperbolic models estimate C and [C + P] by the combination of a hyperbolic function up to 2 μg L^{-1} with a power function at higher concentrations [*Aiken et al.*, 1995]. The OCTS-C model is a power-law formulation which uses the sum of Lwn520 and Lwn565 over Lwn490 to estimate C, whereas the OCTS [C + P] model (OCTS-P) uses two-band ratios, Lwn443/Lwn520 and Lwn490/Lwn520, in a multiple regression function. The POLDER algorithm is considered empirical because it is based on a simple equation relating C to a band ratio, although the equation was actually derived from the use of a modified version of the semianalytic model of *Morel* [1988], parameterized for absorption instead of diffuse attenuation coefficient (A. Bricaud, personal communication, 1997).

The CalCOFI algorithms are derived from CalCOFI data [*Mitchell and Kahru*, 1998]. The CalCOFI two-band relates C to Rrs490/Rrs555 using a power equation. The CalCOFI two-band cubic is a third-order polynomial equation using Rrs490/Rrs555. The CalCOFI three-band, a multiple regression equation, has similarities with the OCTS-P algorithm and uses the Rrs490/Rrs555 and Rrs510/Rrs555 band ratios. The functional form of the CalCOFI four-band equation is similar to CalCOFI three-band except that it uses Rrs443/Rrs555 and Rrs412/Rrs510 (Table 2). The Morel-1 equation was presented at the Ocean Optics XIII meeting [*Morel*, 1997] and relates C to Rrs443/Rrs555 using a power equation (Table 2). Morel-2 is

Table 3. Data Sources and Characteristics of SeaBAM Data Set

Data Set	Provider/PI	Location	Date	n	f_{chla}	f_{phaeo}	h_{chla}	h_{phaeo}	Wavelength
BBOP92-93	D. Siegel	Sargasso Sea	monthly, 1992–1993	72	72	72	72		410, 441, 488, 520, 565, 665
BBOP94-95	D. Siegel	Sargasso Sea	monthly, 1994–1995	67	61	61	67		410, 441, 488, 510, 555, 665
WOCE	J. Marra	50°S–13°N, 88°–91°W 10°S–30°N, 18°–37°W	March 1993 April 1994	70	70				410, 441, 488, 520, 565, 665
EQPAC	C. Davis	0, 140°W	March and Sept. 1992	126			126		410, 441, 488, 520, 550, 683
NABE	C. Trees	46°–59°N, 17°–20°W	May 1989	72			72		412, 441, 488, 521, 550 ···
NABE	C. Davis	46°N, 19°W	April 1989	40			40		410, 441, 488, 520, 550, 683
CARDER	K. Carder	North Atlantic Pacific Gulf Mexico Arabian Sea	Aug. 1991 July 1992 April 1993 Nov. 1994 and June 1995	87	87				412, 443, 490, 510, 555, 670
CALCOFI	G. Mitchell	California Current	quarterly, Aug. 1993 to July 1996	303	303	303			412, 443, 490, 510, 555, 665
MOCE1	D. Clark	Monterey Bay	Sept. 1992	8	8	0	8		412, 443, 490, 510, 555 ···
MOCE2	D. Clark	Gulf California	April 1993	5	5	5	5		412, 443, 490, 510, 555 ···
North Sea	R. Doerffer	55°–52°N, 0°–8°E	July 1994	10			10		412, 443, 490, 510, 555, 670
Chesapeake Bay	L. Harding	~37°N, 75°W	April and July 1995	9			9	9	412, 443, 490, 510, 555, 671
Canadian Arctic	G. Cota	~74.38°N, 95°W	August 1996	8	8	7			412, 443, 490, 509, 555, 665
AMT	G. Moore	50°N–50°S	Sept. 1995 and April 1996	42	42		33		412, 443, 490, 510, 555 ···
Total				919	656	448	442	9	

f_{chla}: fluorometric chlorophyll a; f_{phaeo}: fluorometric phaeophytin a; h_{chla}: HPLC chlorophyll a; h_{phaeo}: HPLC phaeophytin a.

similar to Morel-1 but uses Rrs490/Rrs555. Morel-3 and -4 are other examples of a cubic polynomial with Rrs443/Rrs555 and Rrs490/Rrs555, respectively, and were derived from in situ measurements and an updated version of the *Morel* [1988] semianalytic model (A. Morel, personal communication, 1997).

3. SeaBAM Data Set

To evaluate the performance of chlorophyll and [C + P] algorithms to be used at global scale with SeaWiFS data, an in situ data set was needed to compare with results predicted by the various models. Such an evaluation data set should, ideally, meet the following requirements: (1) contain Rrs or Lwn at or close to the SeaWiFS visible wavelengths; (2) have the in situ chlorophyll a concentrations associated with the stations from which Rrs or Lwn were available or derivable, (3) encompass the widest possible chlorophyll a concentration range, (4) contain data from the widest possible variety of bio-optical provinces, (5) not contain data used for the development of the algorithms under evaluation, and (6) be the same for all algorithms under evaluation.

The constraints imposed by requirements 1 and 5 resulted in an evaluation data set too small to ensure significance or generality to the evaluation results. Therefore a data set was created by merging the data used by the various SeaBAM participants [*Firestone and Hooker*, 1998] as well as other data available in the NASA SeaBASS bio-optical archive [*Hooker et al.*, 1994].

3.1. Radiometric Data: Sources, Processing, and Quality Control

The Rrs in the SeaBAM data originated from various sources and were derived in different ways depending on which investigator processed the data (Table 3). The BBOP and JGOFS data (WOCE, EQPAC, and NABE) were assembled and processed at University of California at Santa Barbara and represent almost one half of the SeaBAM data [*Garver*, 1997]. The Carder data set is composed of above-water measurements [*Carder and Steward*, 1985; *Lee et al.*, 1994] collected during various cruises and locations. The "optimization method" [*Lee et al.*, 1996] was used on high-chlorophyll data where water-leaving radiance in the infrared was not zero, while the "quick and easy" method [*Lee et al.*, 1996] was used on the remaining Carder data. The CalCOFI data set contains more than 300 stations, which were processed at Scripps Institution of Oceanography [*Mitchell and Kahru*, 1998]. Atlantic meridional transect (AMT) data were provided directly to us by Plymouth Marine Laboratory as Rrs. The North Sea, MOCE, Chesapeake Bay, and Canadian Arctic measurements were also extracted from the SeaBASS archive, and the latter three Rrs sets were processed at Goddard Space Flight Center (GSFC). The North Sea and Chesapeake Bay data were collected in waters with high total suspended matter loads (>1 mg/L) and should be considered Case II. A few other stations may also be Case II, but such a classification is difficult to conduct from the sole basis of the radiometric data.

All data were processed in ways compatible with the SeaWiFS protocols [*Mueller and Austin*, 1995]. However, data were not corrected for instrument self-shading [*Gordon and Ding*, 1992] at high C concentration because some of the information required to perform these corrections (i.e., absorption coefficient of the medium, radius of the instrument, Sun zenith angle, and ratio between diffuse and direct Sun irradiance) was not available. The various processing methodologies used for in-water data resulted in Rrs expressed as $L_w/E_d(0+)$, $L_u(0-)/E_d(0-)$ or $L_u(0-)/E_d(0+)$, where L_w is the water-leaving radiance, L_u is the upwelling radiance, E_d is the downwelling irradiance, and 0+ or 0− indicates measurements just above or just beneath the sea surface, respectively. For consistency, all data were converted to Rrs = $L_w/E_d(0+)$ using $L_w = 0.54L_u(0-)$, and $E_d(0-) = 0.96E_d(0+)$, where 0.54 is a mean coefficient summarizing the effect of internal reflection of the upwelling flux during transmission through the interface, and 0.96 accounts for the loss of the

downwelling flux by reflection at the air-sea interface [*Austin*, 1974; *Gordon et al.*, 1988; *Morel and Antoine*, 1994]. Both coefficients assume low solar zenith angle and calm sea surface.

The SeaBAM data set includes ~10% of Rrs data calculated from above-surface measurements (i.e., Carder and North Sea data), while all other data were derived from in-water measurements. The protocols for above-surface measurements are still under development, and the agreement between above-surface and in-water measurements is extremely dependent upon experimental and environmental conditions. However, on the basis of the tests described below, the Rrs from above-surface measurements did not differ significantly from the other data.

Radiance band ratios versus C and single bands versus C were plotted for the various data sets in order to identify outliers. These plots were very useful in revealing data with errors and in determining which data could be corrected and which should be removed from the SeaBAM data. A second quality control measure used preliminary results from eight chlorophyll or [C + P] algorithms to identify stations with one or more "anomalous" radiance values. These models (CalCOFI two-band linear; CalCOFI four-band; GPs, Morel-1; Morel-3; OCTS-C; Carder; Garver/Siegel) were chosen because their outputs showed a good linearity in log-log space (minimal curvilinearity) and/or they use three or more wavelengths (see Table 2). A station was considered an outlier when the ratio of the modeled chlorophyll *a* concentration to the in situ concentration exceeded 5:1 or was less than 1:5 for two or more models. These rejection criteria were chosen not to be too restrictive in order to eliminate only extreme stations. After the elimination of 54 suspect stations, 919 stations remained in the final SeaBAM data set (Table 3).

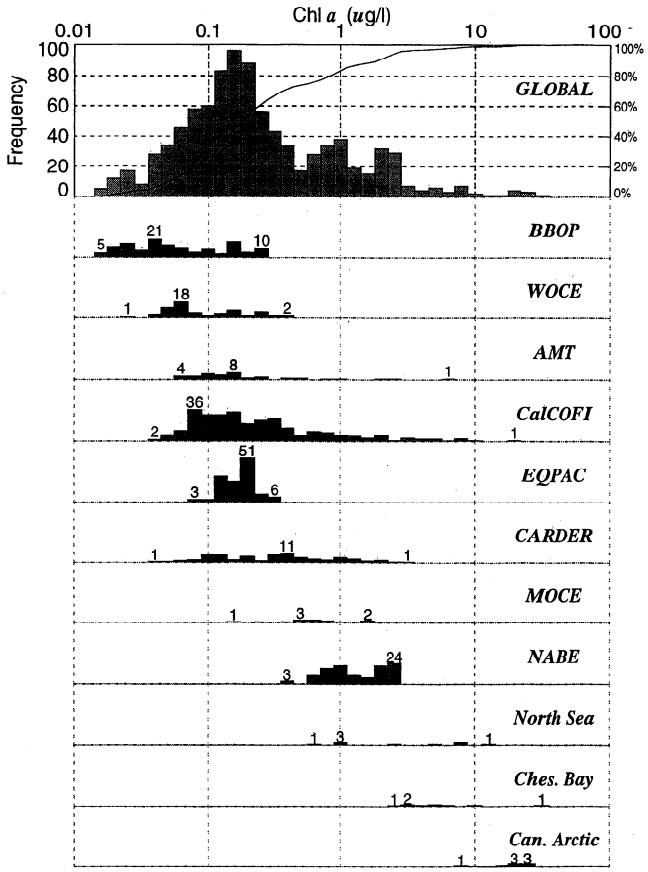

Figure 1. Frequency distribution of Chl *a* concentration in the SeaBAM data set (top panel) and in the 11 subsets.

3.2. In Situ Chlorophyll *a* Data

The SeaBAM data set is comprised of fluorometric [*Yentsch and Menzel*, 1963; *Holm-Hansen et al.*, 1965] and/or high-performance liquid chromatography (HPLC) measurements of chlorophyll *a* and phaeophytin *a* (Table 3). For the purpose of evaluating the suitability of models for global application, where chlorophyll concentrations range over 3 orders of magnitude, the largest and most representative data set is required. It was not possible to stratify the SeaBAM data by chlorophyll method and still achieve global coverage. Therefore HPLC and fluorometric measurements were merged to form the chlorophyll *a* evaluation data set. HPLC C was preferentially used, when available, because this method is considered more precise than fluorometric methods and fluorometric C was used for subsets which do not contain HPLC data.

The distribution of the chlorophyll *a* data in the SeaBAM data and its subsets are illustrated in Figure 1. The concentrations range between 0.019 and 32.79 $\mu g\ L^{-1}$ with a geometric mean of 0.27 $\mu g\ L^{-1}$, somewhat higher than the global ocean mean (0.19 $\mu g\ L^{-1}$) reported by *Antoine et al.* [1996]. The BBOP, WOCE, and EQPAC data sets represent most of the data at the low end of the concentration range. The Carder, CalCOFI, and AMT data cover a relatively wide concentration range, from oligotrophic to eutrophic waters. NABE, North Sea, Chesapeake Bay, and Canadian Arctic data are mostly at concentrations over 1 $\mu g\ L^{-1}$. If 0.1 and 1 $\mu g\ L^{-1}$ are taken as approximate limits between oligotrophic and mesotrophic waters and between mesotrophic and eutrophic waters [*Antoine et al.*, 1996], then the SeaBAM data are approximately composed of 23, 59, and 18% of oligotrophic, mesotrophic, and eutrophic stations, respectively. Comparing these proportions with those reported by *Antoine et al.* [1996] for the world ocean (oligotrophic: 55.8%; mesotrophic: 41.8%; and eutrophic: 2.4%) reveals that the SeaBAM data tend to be overrepresented by mesotrophic and eutrophic waters. Alternatively, for algorithm evaluation purposes, where a more uniform distribution over the concentration range may be desirable, the data set has a relative insufficiency of concentrations exceeding 7–8 $\mu g\ L^{-1}$.

3.3. In Situ [C + P]: Data and Estimates

The [C + P] algorithms for SeaWiFS are needed to permit comparisons with historical CZCS data. Most algorithms we evaluated estimate C, but several were designed to estimate [C + P] (Table 2). Only 448 SeaBAM stations had fluorometric measurements of both chlorophyll *a* and phaeophytin *a* (Table 3). Since equitable algorithm comparisons require data sets with similar characteristics and number of stations, a statistical relationship between C and [C + P] was investigated to allow reasonable estimates of [C + P] for those stations where only chlorophyll *a* was available. A set of 2329 stations sampled from 1978 to 1996, including those available in the SeaBAM data with fluorometric measurements of both Chl *a* and phacopigment in the 0–10 m depth range, was extracted from the SeaBASS "Historical Pigment Database" [*Hooker et al.*, 1994]. The data were mostly collected off the U.S. coasts and the Atlantic and Pacific Oceans by various investigators. Stations with [phaeo] > [Chl *a*] (n = 67) were considered extreme cases and removed. A type II (RMA) regression on log-

Table 4. Mean Extraterrestrial Solar Irradiance

Wavelength, nm	$\overline{F_o}$ $\mu W\ cm^{-2}\ nm^{-1}$
*SeaWiFS Bands**	
412	170.7943
443	189.4438
490	193.6842
510	188.3675
555	185.3973
670	153.3877
765	122.5128
865	99.0214
OCTS Bands†*	
412	170.96
443	188.17
490	194.59
520	185.74
565	184.49
670	153.12
765	122.61
865	98.55

*H. Gordon, personal communication, 1998.
†Advance Earth Observing Satellite (ADEOS). OCTS Data Processing Algorithms Description version 2.01, NASDA, June 1997.

transformed data yielded the following equation (n = 2262; R^2 = 0.993):

$$[C + P] = 1.34*C^{0.983} \quad (1)$$

It is acknowledged that C and [C + P] cannot be considered as completely independent variables in the above regression. The major benefit of this approach is to allow the derivation of a chlorophyll/[C + P] relationship from a more stable basis than one based on chlorophyll versus phaeopigment. Because (1) yielded C/[C + P] ratios comparable to other reported ratios [e.g., *Smith and Baker*, 1978; *Morel and Berthon*, 1989; *Balch et al.*, 1992], it was used to estimate the [C + P] concentration for stations where the complete fluorometric information was missing (n = 471).

3.4. Radiometric Data Adjustments

For algorithms that required normalized water-leaving radiances as input, Rrs was multiplied by the mean extraterrestrial solar irradiance [*Neckel and Labs*, 1984] weighted by the spectral response of the relevant sensor bands (see Table 4). In addition, because the wavelengths required by the various algorithms (Table 2) did not always match those available in the SeaBAM data (Table 3), several radiometric adjustments were applied to some data sets. These adjustments were aimed to enhance the consistency of the algorithm comparison by testing all algorithms using the full dynamic range of available radiance data (n = 919).

For the first three shorter wavelengths the maximum difference between data and bands required by the various algorithms is 2 nm and was considered negligible. The major differences occur for algorithms that require either 510 or 520 nm data, while SeaBAM contains a mixture of these pairs of measurements. Similar mismatches exist between algorithms using either 550, 555, or 565 nm data and SeaBAM data which are comprised of a mix of these three wavelengths.

3.4.1. 565–555 nm. Even though chlorophyll absorbs light weakly in the 550–565 nm region and the Rrs spectrum is relatively insensitive to changes in C concentration at these wavelengths, 555 and 565 nm data are not interchangeable. This is particularly evident at low C concentrations, where for example, substituting Rrs565 for Rrs555 in a band ratio would give anomalously higher reflectance ratios than those expected for the clearest waters, based on backscattering data derived from *Morel* [1974] and recent absorption coefficients for pure seawater from *Pope and Fry* [1997].

The feasibility of estimating Rrs555 from measurements of Rrs565 was explored using BBOP94-95 data which have concurrent measurements at both wavelengths. A strong linear relationship was found (n = 78; R^2 = 0.975):

$$Rrs555 = 1.0628*Rrs565 + 0.0002 \quad (2)$$

Equation (2) was therefore applied to the BBOP92-93 Rrs565 data, to generate proxy estimates of Rrs555, and to the WOCE data set, which has a narrow range of low C concentrations very similar to the BBOP data (Figure 1).

3.4.2. Other radiometric adjustments. Two algorithms, OCTS-C and POLDER, use radiance data at 565 nm instead of 555 nm (Table 2). It is inappropriate to invert (2) to generate proxy Rrs565 from Rrs555 data because its applicability is restricted to the low C concentrations (<0.4 $\mu g\ L^{-1}$) used in its derivation. Another approach was thus used to convert between 555 and 565 nm and between 510 and 520 nm data. It uses the reflectance ratios predicted at any given C concentration by the semianalytic model of *Morel* [1988] adapted with *Pope and Fry* [1997] a_w data. These predicted reflectance ratios are shown in Figure 2 along with those predicted by the semianalytic model of *Gordon et al.* [1988]. According to these models, the Rrs555/Rrs565 reflectance ratio decreases from ~1.13 to ~0.97 as C increases from 0.015 to 7 $\mu g\ L^{-1}$, whereas for the same concentration range, Rrs510/Rrs520 varies from ~1.32 to ~0.95. Contrary to the 555–565 nm region, the 510–520 nm domain is highly influenced by pigment absorption and is thus more variable. The Rrs510/Rrs520 versus C and Rrs555/

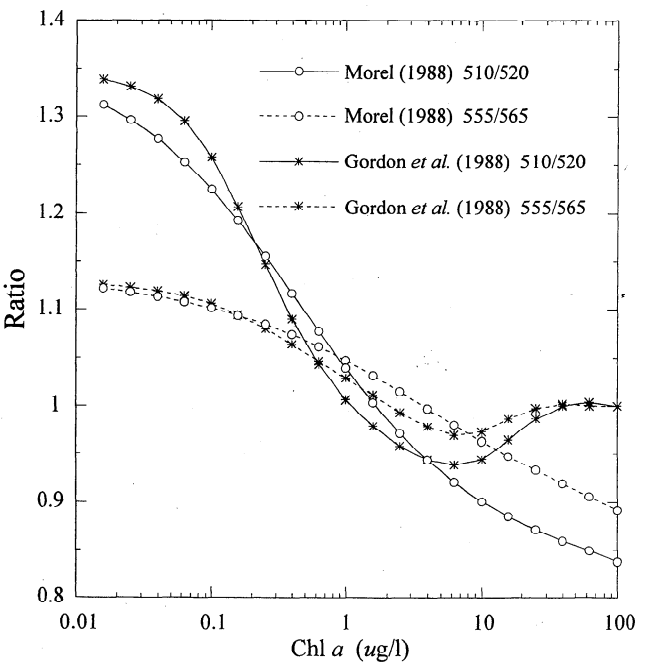

Figure 2. Ratios of Rrs510/Rrs520 and Rrs555/Rrs565 predicted from semianalytic models of *Morel* [1988] and *Gordon et al.* [1988] as a function of [Chl *a*].

Rrs565 versus C relationships derived from Morel's model were used as a basis for estimating Rrs at one wavelength from the other as a function of in situ C concentration.

3.5. Final Rrs-Chlorophyll Data Set

The Rrs490/Rrs555 ratio versus in situ C for the SeaBAM data and for its subsets is shown in Figure 3. The figure illustrates the dynamic range associated with each subset and of the combined data set. The dispersion (orthogonal to main axis) of the data is another informative feature of these plots. The CalCOFI data set, for instance, appears very coherent, whereas some other data sets are slightly noisier. Although SeaBAM data originate from various investigators and were processed differently by different people, the variability in the radiometric data is reasonably limited. At high C concentration the dispersion of radiance ratios in the SeaBAM data increases, mostly because of the presence of Case II waters. The shape of the scatterplot for the SeaBAM data is clearly sigmoid (in log-log space) as predicted by the *Morel* [1988] and *Gordon et al.* [1988] models. This trend, though less marked, is seen in the CalCOFI data. At lowest C concentrations the highest Rrs490/Rrs555 ratios are slightly lower than the theoretical limit (~6.66) for clear natural waters.

4. Algorithm Evaluation Criteria

A variety of statistical and graphical criteria were used to evaluate agreement between C, estimated by the various models, and in situ C (Table 5). Statistical comparisons between in situ C and model C were based on log-transformed data, in part, to encompass the several-orders-of-magnitude variation in C, and because log-transformed C was more normally distributed than untransformed data (Figure 1) [also see *Campbell*, 1995a]. The slope and intercept of the linear equation relating model to in situ data was computed using a type II (reduced major axis) functional regression model which is considered the appropriate model when the assignments to x or y axes are arbitrary and when substantial variance is expected in both variables [*Ricker*, 1973; *Laws and Archie*, 1981; *Press and Teukolsky*, 1992]. Statistics (Table 5) such as regression slope and intercept, coefficient of determination (R^2) and root-mean-square error (RMS) provide a numerical index of model performance but may not indicate nonlinear trends or other important features in the relationship between model and in situ data. Several complementary graphical portrayals were constructed to illustrate these features, including standard scatter, relative frequency, and empirical quantile-quantile plots [*Chambers et al.*, 1983], hereinafter abbreviated as q-q. The q-q plot, when both sets have the same number of observations, as is our case, is simply a plot of the model data sorted (ascending order) against the sorted in situ data.

Table 5. Criteria for Model Evaluation (Log-Transformed Data)

	Evaluation
	Statistical
Regression slope	1 ± 0.01
Regression intercept	0 ± 0.01
Bias	0 ± 0.01
R^2	>0.9
RMS	<0.185
Negative estimates	none
	Graphical
Scatter	linear distribution; few outliers (model:in situ, <5:1 and >1:5)
Quantile-quantile	linear; data overlap the 1:1 line; no discontinuities
Relative frequency	congruency with in situ data

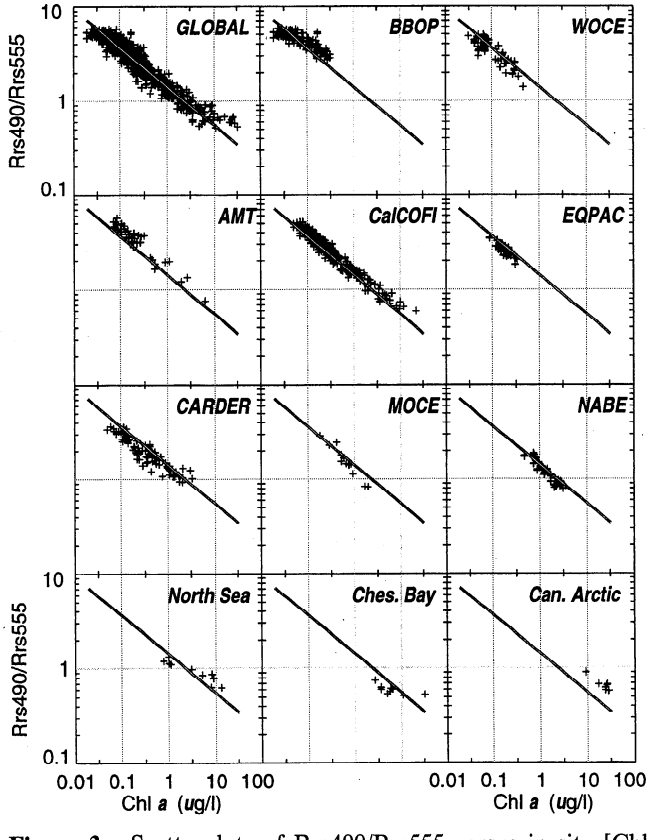

Figure 3. Scatterplots of Rrs490/Rrs555 versus in situ [Chl *a*] for the SeaBAM data set (top left panel) and in the 11 subsets. The RMA regression line for the data set is repeated in each of the 11 subset plots as a reference.

5. Algorithm Evaluation Results

The statistical results of the algorithm evaluation are presented in Table 6. Graphical results for 12 algorithms, selected because they represent a particular functional form or satellite sensor, are illustrated in Figure 4.

In general, all algorithms performed reasonably well, at least in part of the whole concentration range. It is noteworthy that two [C + P] algorithms used extensively to process CZCS data, the GPs and Clark three-band, tended to underestimate in situ [C + P] (Table 6 and Figure 4). Relative frequency distribution plots reveal that modes for these models are close to the in situ [C + P] mode but that both models overestimate the frequency of low concentrations. A discontinuity induced by the equation switch in the GPs (Table 2) is evident in the q-q plot. The statistical artifacts in CZCS [C + P] retrievals resulting from the GPs algorithm switch (e.g., bimodality, frequency discontinuity) have been reported elsewhere [*Muller-Karger et al.*, 1990; *Denman and Abbott*, 1988; *Campbell*, 1995b].

A discontinuity is also observed with the Aiken-C algorithm which switches from the hyperbolic to the power equation at 2 $\mu g\ L^{-1}$, and there is a marked curvature evident in scatter and

Table 6. Summary of Statistical Results of Algorithm Evaluations

Rank	Algorithm	N	Intercept	Slope	Rsq	RMS	Bias	Nonlinear	Discontinuity
1	Morel-1	919	0.038	0.975	0.917	0.179	0.052	X	
2	Morel-3	919	0.040	0.970	0.915	0.183	0.058		
3	CalCOFI 2-band cubic	919	0.072	0.980	0.918	0.190	0.083		
4	OCTS-C	919	0.054	1.148	0.933	0.190	−0.030	X	
5	Carder (global)	919	−0.033	0.990	0.876	0.213	−0.027		
6	CalCOFI 2-band linear	919	0.074	0.991	0.915	0.192	0.079	X	
7	Morel-2	919	0.081	1.037	0.915	0.190	0.060		
8	CalCOFI 3-band	919	0.062	0.939	0.908	0.205	0.097		
9	Morel-4	919	0.102	1.059	0.907	0.204	0.069		
10	Siegel-Garver (global)	919	−0.012	0.928	0.734	0.311	0.029		
11	GPs	919	−0.239	1.004	0.923	0.292	−0.241		X
12	CalCOFI 4-band	919	0.073	0.934	0.900	0.218	0.110		
13	POLDER	919	0.215	1.190	0.921	0.241	0.107		
14	Carder (subtropical)	919	−0.128	1.073	0.872	0.284	−0.169		
15	Aiken-C	877*	−0.094	1.083	0.774	0.330	−0.139	X	X
16	Aiken-P	877*	−0.120	1.118	0.787	0.339	−0.168	X	X
17	Clark 3-band	919	−0.306	0.913	0.905	0.323	−0.267		
18	Siegel-Garver (BBOP)	919	0.141	0.776	0.896	0.345	0.269		
19	OCTS-P	919	−0.345	1.750	0.913	0.842	−0.680	X	

*Forty-two negative estimates.

q-q plots. *Aiken et al.* [1995] noted that an insufficient number of high chlorophyll stations in their data set prevented the fitting of a single hyperbolic equation to the entire set. Note also that this model generated negative concentrations for 42 stations because it uses a "clear water" Rrs490/Rrs555 limit (5.29), which is below that observed for these stations.

Considering the two semianalytic algorithms, the Carder algorithm yielded better overall agreement with in situ C than the Garver/Siegel model. As described earlier, the philosophy of model inversion is very different in these models despite the fact they are both based on the same formulation linking Rrs to absorption and backscattering. The parameterization in these two models is also very different. For instance, some of the coefficients in the Carder model were based on specific tuning to in situ data, whereas the Garver/Siegel model uses parameters available in the literature without any particular tuning. Compared with the SeaBAM data, the Garver/Siegel model underestimated the lowest concentrations as well as concentrations above 1 $\mu g\ L^{-1}$. The Carder subtropical model (Table 6) performed well in the midconcentration range, while it generally underestimated concentrations at the low and high ends of the range. In sharp contrast the Carder global model performed as well as several of the best empirical models (see Table 6). For both semianalytic models, the outliers in the 1–10 $\mu g\ L^{-1}$ concentration range are Case II stations.

A residual sigmoid pattern in the q-q plots is a recurrent feature in several algorithms (e.g., GPs, Clark three-band, CalCOFI three-band, CalCOFI four-band, Morel-1, OCTS-C). The Morel-1 and Morel-3 algorithms use Rrs443/Rrs555, but it is clear from the q-q plots that Morel-3, the cubic polynomial, agrees better with the 1:1 line. The relative frequency plot also shows a better agreement between modeled and actual data for the Morel-3 algorithm. Another interesting aspect also well illustrated by the Morel-1 and -3 algorithms is the lower dispersion of the data at low C concentrations, typical of algorithms using the 443/555 ratio.

The OCTS and POLDER are the operational algorithms used for the 9 months of data collected by ADEOS. The POLDER algorithm gives reasonable estimates at concentrations under 0.4 $\mu g\ L^{-1}$, but it overestimates higher concentrations. The OCTS-C algorithm exhibits some curvature along the concentration range, and its slope departs significantly from 1, but the negative bias is small (Table 6), and the algorithm behaves well globally.

Among the CalCOFI algorithms, the best results are obtained with the two-band cubic polynomial function. The algorithm performs well at all concentrations except at the lower end where it overestimates concentrations. It is noteworthy that in this particular case, an increasing complexity in the formulations (three and four bands used in quadratic functions) did not increase overall performance of these algorithms.

The algorithms are ranked according to their overall statistical performance in Table 6. Evidence of model discontinuities or nonlinearity is also summarized in this table. For each statistical parameter, the algorithms were ranked (slope closest to 1, intercept and bias closest to 0, highest R^2), and these scores were summed to yield the overall final rank. While this ordination scheme arbitrarily gives the same weight to each statistical parameter, it nevertheless does indicate the overall performances of the various models. Features such as discontinuities, curvatures, and mismatches in relative frequency with in situ were not used in the ranking. Consideration of these aspects would reduce the rank of some algorithms which are highly ranked. For example, the Morel-1, CalCOFI two-band, and CalCOFI cubic and OCTS-C are ranked higher than other empirical algorithms, but the graphical results show nonlinear trends for several of them. In summary, empirical equations generally performed better than the semianalytic algorithms and when considering both statistical and graphical criteria, those using cubic polynomial formulations, such as Morel-3 and CalCOFI cubic, performed best.

6. Analysis of Functional Forms

The empirical models tested above vary in their formulation and complexity and collectively represent approaches of the last two decades. They use single or multiple band ratios and different formulations: power function, multiple regression, hyperbolic, second-order and third-order polynomials, and most use log-transformed data. Since these algorithms were developed and tuned using different data sets, it is difficult to

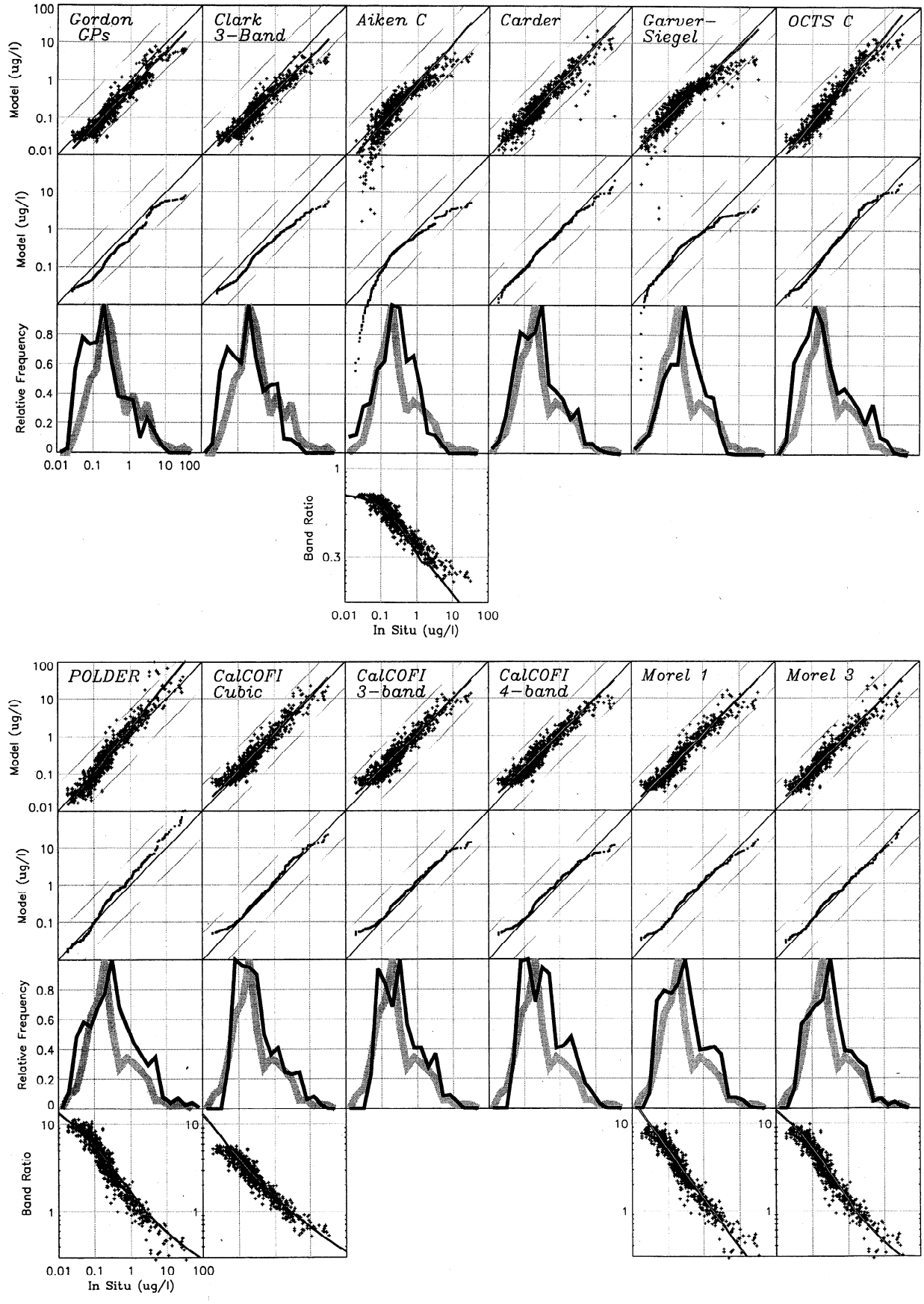

Figure 4. Comparisons between model and in situ data: (top) Gordon GPS, Clark three-band, Aiken-C, Carder (global), Garver/Siegel (global), and OCTS-C models; (bottom) POLDER, CalCOFI cubic, CalCOFI three-band, CalCOFI four-band, Morel-1, and Morel-3 models. From top to bottom: Scatterplots; quantile-quantile plots; relative frequency of model (thin black line) and in situ (thick faint line); band ratio versus in situ C for two-band ratio algorithms (pluses) and band ratio versus model (curve). Note that the axes for each row of figures are shown in column 1. Also shown are lines indicating model:in situ ratios of 1:5 and 5:1.

Table 7. Ocean Chlorophyll Algorithms (Empirical Algorithms Tuned to the SeaBAM Data)

Algorithm	Functional Form	Band Ratio (R)	Coefficient				
			$a0$	$a1$	$a2$	$a3$	$a4$
OC1a	power	R = log(Rrs490/Rrs555)	0.3734	−2.4529			
OC1b	geometric	R = log(Rrs490/Rrs555)	0.3636	−2.3500	−0.0100		
OC1c	quad. poly.	R = log(Rrs490/Rrs555)	0.3920	−2.8550	0.6580		
OC1d	cubic poly.	R = log(Rrs490/Rrs555)	0.3335	−2.9164	2.4686	−2.5195	
OC2a	MCP	R = log(Rrs412/Rrs555)	0.2457	−1.7620	0.2830	0.1035	−0.0388
OC2b	MCP	R = log(Rrs443/Rrs555)	0.1909	−1.9961	1.3020	−0.5091	−0.0815
OC2	MCP	R = log(Rrs490/Rrs555)	0.3410	−3.0010	2.8110	−2.0410	−0.0400
OC2d	MCP	R = log(Rrs510/Rrs555)	0.4487	−4.3665	2.7130	−0.2698	−0.0821
OC2e	MCP	R = log(Rrs520/Rrs555)	0.5072	−6.2432	2.7787	3.3845	−0.0413
OC3d	MCP	R = log((Rrs443 > Rrs490)/Rrs555)	0.3483	−2.9959	2.9873	−1.4813	−0.0597
OC3e	MCP	R = log((Rrs443 > Rrs520)/Rrs555)	0.5179	4.7478	6.7321	−4.1287	−0.0121
OC4	MCP	R = log((Rrs443 > Rrs490 > Rrs510)/Rrs555)	0.4708	−3.8469	4.5338	−2.4434	−0.0414

Formulations: Power, $C = 10^{(a0+a1*R)}$; geometric, $C = 10^{(a0+a1*R)} + a2$; quadratic polynomial, $C = 10^{(a0+a1*R+a2*R^2)}$; cubic polynomial, $C = 10^{(a0+a1*R+a2*R^2+a3*R^3)}$; MCP, $C = 10^{(a0+a1*R+a2*R^2+a3*R^3)} + a4$.

use their results to determine which formulations have the highest potential merit. To investigate this aspect, several simple empirical formulations were "tuned" to the SeaBAM data to achieve a slope of 1.000 and an intercept of 0.000, maximum R^2, and minimum RMS.

The power equation has been widely used to relate radiance ratios to C, in part, due to the relative ease of derivation of model parameters using a simple linear regression of log-transformed data [e.g., *Clark*, 1981; *Smith and Baker*, 1982; *Gordon et al.*, 1983; *Mitchell and Holm-Hansen*, 1991]. Despite its simplicity, the power model (Table 7) captures a large fraction of the variation in radiance band ratios and the 3-orders-of-magnitude variation in C (Figure 5). It is clear that the power model fits the most frequent, central data but does not fit the extremes well. There is a significant overall sigmoid pattern evident in the scatter and q-q plots. This residual curvature results from the inability of power equations to capture the inherent sigmoid relationship between commonly used band ratios and in situ C in log-log space. Additionally, the relative frequency distribution of model C is much broader than in situ.

Geometric models have been used only rarely. One example was reported by *Hojerslev* [1981]. Compared with the power equation, the geometric model agrees better with in situ C at low concentrations, but neither captures the inherent sigmoid pattern evident in plots of band ratios versus C which consequently results in a "residual" sigmoid trend in the plot of model C versus in situ C (Figure 5).

A quadratic polynomial (second order) achieved a better match with highest in situ C than the power model but a relatively poorer match with lowest C values (Figure 5). The POLDER, CalCOFI two-band cubic, Morel-3, and Morel-4 algorithms use a cubic polynomial (third-order) equation and generally performed better than most other algorithms tested. Our fit of a cubic polynomial equation to SeaBAM Rrs490/Rrs555 and C data matches the mode of the in situ distribution (relative frequency plots) better than the power, geometric, or quadratic models, but like these other models, it does not simulate well low C concentrations, and the frequency distribution around the mode is still too broad relative to that of in situ C (Figure 5). In the case of the SeaBAM data, no major improvements are observed when coefficients are derived for higher-order polynomials (i.e., order >3).

Several other conventional two-band and three-band ratio algorithms were explored. A simple hyperbolic equation did not fit the SeaBAM data well (not shown). This might be expected because this model assumes symmetry, while the tails (asymptotes) of the Rrs490/Rrs555 versus C joint distribution are not symmetrical. Some band combinations in cubic polynomials (e.g., (Rrs443 + Rrs490)/Rrs555, or (Rrs443 + Rrs490 + Rrs510)/Rrs555, not shown) have interesting potential but did not yield to major improvements over the use of a simple Rrs490/Rrs555 band ratio.

7. Discussion

7.1. SeaBAM Data Set

This analysis is based on the largest in situ data set ever assembled (to our knowledge) for ocean color algorithm studies. However, our algorithm evaluations are only as good as the data set itself. The data quality control procedures were designed to identify extreme outliers and erroneous data, but the rejection criteria were deliberately not too severe, so some of the remaining stations may still depart slightly from the general trend. Differences in data acquisition methodologies (e.g., above-surface versus in-water measurements), radiometer designs, calibrations, data processing, and environmental factors (sea and sky state) are probably responsible for part of the variability observed. If radiometric data were measured and processed in a more consistent manner, the dispersion within some of the subsets might be reduced [see *Siegel et al.*, 1995], but it is not certain that dispersion (orthogonal to axial trends; see Figure 3) in the SeaBAM data would decrease, because the data set would still reflect the real inherent variability in bio-optical properties of the waters represented by the various subsets.

The SeaBAM chlorophyll data set was formed by merging HPLC and fluorometric measurements of chlorophyll to encompass the largest possible range of data and bio-optical provinces. The consequences of using either fluorometric or HPLC chlorophyll are difficult to assess mostly because limited information about the relationship between HPLC and fluorometric data is available in the literature [e.g., *Trees et al.*, 1985; *Bricaud et al.*, 1995]. Moreover, the equivalence of these two measurements may vary with season, location, depth and concentration range of data, as well as the way pigments were separated and the kind of statistical analyses performed. Fluorometric concentrations are, usually, higher than HPLC esti-

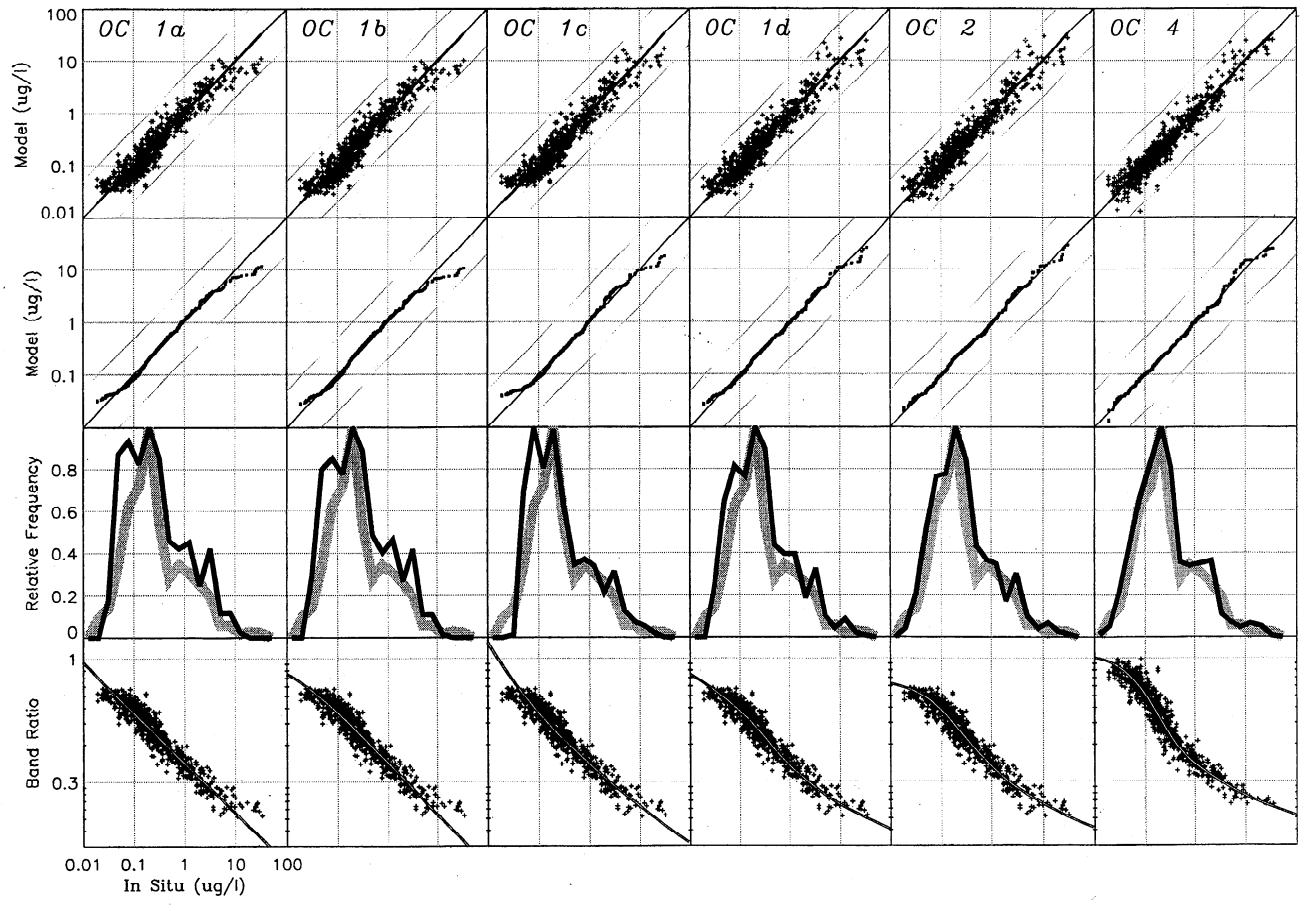

Figure 5. Comparisons between models tuned to the SeaBAM data set and in situ data: OC1a (power), OC1b (geometric), OC1c (quadratic polynomial), OC1d (cubic polynomial), OC2 (modified cubic polynomial), and OC4 (maximum band ratio) algorithms. See Figure 4 caption for additional details.

mations (e.g., in the SeaBAM data), but it is unrealistic to consider a unique relationship between the two methods to convert data from various dates and regions. Given these uncertainties, it seemed prudent not to apply any fluorometric/HPLC conversion scheme to the SeaBAM data. Additional independent analyses are required which evaluate the comparability of fluorometric and HPLC chlorophyll *a* and the impact of blending measurements from these two methods on algorithm validation and calibration.

Whether the SeaBAM data adequately represent global bio-optical variability is another important question. Several major bio-optical provinces of the world ocean are poorly or not represented; the data set contains very little data from polar regions, none from Antarctica, and an insufficient number of C observations above 8 and below 0.03 μg L^{-1}. Future expansion of the SeaBAM data set should not be considered as a simple increase in the number of stations available but rather as an increase in the diversity of bio-optical conditions it encompasses. Since radiometric bands separated by 10 nm (i.e., 510 versus 520 nm, or 555 versus 565 nm) are not interchangeable, radiometric adjustments such as those we performed are, in the present state of the data set, necessary to ensure that algorithm comparisons and analyses are equitably based on the complete set of 919 stations. Additions to the SeaBAM data set should therefore take into consideration the increasing need for intercomparison and evaluation of data or algorithms related to various satellite sensors (e.g., CZCS, OCTS, POLDER, SeaWiFS, MODIS). This underscores the importance of acquiring in situ observations which include all wavelengths involved (Table 1). Correction for instrument self-shading at high C is also advisable, although recent results from *Kahru and Mitchell* [1998a] suggest that the correction for the Lwn490/Lwn555 band ratio is probably small (e.g., mean of <10% for 11 stations, with C between 10 and 32.5 μg L^{-1}).

7.2. Empirical Versus Semianalytical Algorithms

The semianalytic algorithms evaluated yielded results inferior to those from several empirical expressions. Yet semianalytic algorithms are potentially much more useful because they allow the derivation of other in-water, optically active constituents besides C, such as nonchlorophyllous absorption, CDOM, or backscattering. In their present state, semianalytic algorithms are handicapped by their design and parameterization. Because they may employ four or more radiance bands, semianalytic algorithms also require more consistent data sets with high spectral fidelity in order to perform as well as or better than simple two-band empirical algorithms. Some simplifying assumptions, not always true in the world ocean, are used to limit the number of unknowns in semiempirical models. These assumptions often result in making constant some parameters which actually vary in the ocean (e.g., the slope in detrital absorption). Other limitations come from the parameterization of some terms used in the Rrs = $f(b_b/(a + b_b))$ equation. Most of these are wavelength dependent but some

Table 8. Statistical Results for Chlorophyll Algorithms Tuned to SeaBAM Data

Algorithm	Type-Wavelengths, nm	Rsq	RMS
OC1a	power-490, 555	0.915	0.175
OC1b	geometric-490, 555	0.912	0.178
OC1c	quadratic-490, 555	0.917	0.173
OC1d	cubic-490, 555	0.918	0.172
OC2a	MCP-412, 555	0.892	0.197
OC2b	MCP-443, 555	0.916	0.173
OC2	MCP-490, 555	0.918	0.172
OC2d	MCP-510, 555	0.849	0.235
OC2e	MCP-520, 555	0.744	0.311
OC3d	MCP, MBR-443, 490, 555	0.928	0.161
OC3e	MCP, MBR-443, 520, 555	0.921	0.169
OC4	MCP, MBR-443, 490, 510, 555	0.932	0.156

For all algorithms, N = 919, intercept = 0.000, slope = 1.000, and bias = 0.000. MCP, modified cubic polynomial; MBR, maximum band ratio.

terms, such as specific absorption coefficient of phytoplankton a^*_{ph}, are dependent on phytoplankton concentration or phytoplankton community structure, which depend on the trophic status of the *waters* [*Mitchell and Kiefer*, 1988; *Bricaud and Stramski*, 1990; *Cleveland*, 1995; *Sosik and Mitchell*, 1995]. This suggests it may be necessary to use a parameterization which takes biological/ecological variability into account [e.g., *Cleveland*, 1995; *Bricaud et al.*, 1995]. Other parameterization weaknesses exist at high C concentrations [see *Morel*, 1997] which, in the case of the Carder algorithm, are partially canceled by the use of an empirical default algorithm. The improved results obtained with the "global" version of both semianalytic models (Table 6) are encouraging examples showing that when appropriately parameterized, these algorithms can achieve global estimates which are reasonably accurate. Some of these parameterization problems may cancel at local or regional scales where variability introduced by biological processes may be more limited. At present the predictive skill of the semianalytic algorithms remains, however, inferior to that of empirical algorithms when applied to widely varying bio-optical provinces. Until such domains are well sorted and understood, the most conservative approach for acquiring a global satellite chlorophyll data set is the use of a globally tuned empirical algorithm.

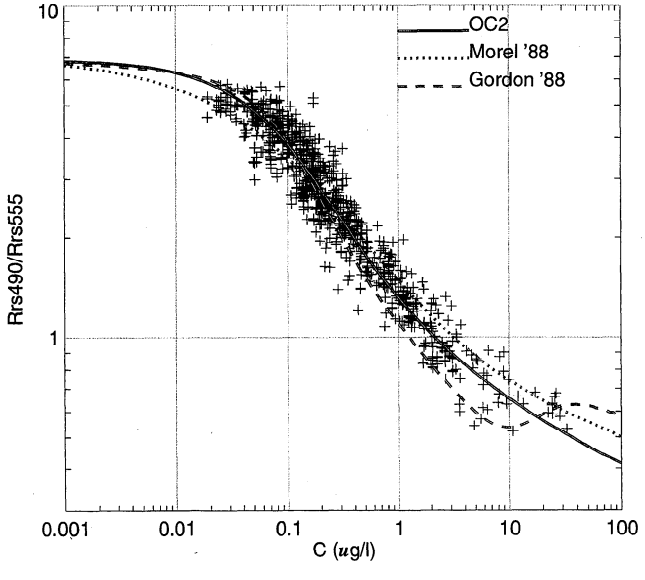

Figure 6. Relationship between chlorophyll and Rrs490/Rrs555 for the ocean chlorophyll 2 empirical algorithm (solid curve) and *Gordon et al. [1988] (dashed curve) and Morel* [1988] (dotted curve) semianalytic models. In situ Rrs490/Rrs555 versus in situ C (pluses). (The a_w coefficients used in semianalytic models are from *Pope and Fry* [1997]; b_{bw} coefficients are from *Morel* [1974].

7.3. New Formulations for Empirical Algorithms

Most algorithms presented here perform reasonably well, and several may be used with satisfactory results for particular regions or concentration ranges. Those performing best over the whole C concentration range are based on cubic polynomials. However, none of them fit the low C values very well.

7.3.1. Ocean chlorophyll 2 (OC2), modified cubic polynomial. In our attempts to fit a cubic polynomial to the SeaBAM data, we were also not able to capture the lowest C-highest ratios of Rrs490/Rrs555 without compromising the fit elsewhere. Inspection of scatterplots and q-q plots suggested that the cubic polynomial formula required an additional correction term (coefficient) which influences the shape of the curve at low C but has little effect at higher concentrations. Such a "modified cubic polynomial" (MCP) formula (Table 7), named "ocean chlorophyll 2" (OC2), yielded very good statistical results when tuned to the SeaBAM data (R^2 = 0.918; RMS = 0.172 (Table 8)). Algorithm tuning involved determination of MCP coefficients using iterative minimization routines (IDL, Research System Incorporated) to achieve a slope of 1.000 intercept of 0.000, minimum RMS of q-q, and maximum R^2 between model and observed chlorophyll data.

The agreement between the OC2 model and the in situ data throughout the range of C is excellent, and the relative frequency distributions of model and in situ C are highly congruent (Figure 5). The OC2 model captures the inherent sigmoid relationship between in situ band ratio and C, evident in semianalytic models such as *Gordon et al.* [1988] and *Morel* [1988] (Figure 6). Our empirical fit (OC2) suggests that the sigmoid relationship is asymmetric, with steeper curvature at low C than at high C concentrations (but more observations with C above 20 μg L^{-1} are required to confirm this). At low C the OC2 asymptotically approaches the expected clear water value (~6.6, Table 9) for the Rrs490/Rrs555 radiance ratio.

7.3.2. Ocean chlorophyll 4 (OC4), maximum band ratio algorithm. In the algorithm evaluation the global processing switching (GPs) algorithm yielded one of the highest coefficients of determination (R^2 = 0.927) between in situ and

Table 9. Clear Water Values for Band Ratios With Rrs555 in Denominator As Predicted by Model Using $b_b/a + b_b$ Formulation and Model Based on b_b/a Relationship (With a_w Values From *Pope and Fry* [1997] and b_{bw} From *Morel* [1974]

Rrs Band Ratio	Rrs = $f(b_b/(a + b_b))$	Rrs = $f(b_b/a)$	MCP Algorithm
412/555	27.61	47.03	28.52
443/555	16.53	21.78	11.91
490/555	6.13	6.66	6.80
510/555	2.52	2.58	3.12
520/555	1.87	1.89	2.42

Clear water values predicted by the MCP equations derived from the SeaBAM data set (Table 7) are also indicated.

model. The GPs achieves this by switching from a 443/550 to a 520/550 band ratio, thereby avoiding the relatively lower and noisier 443/555 ratios when C exceeds ~1.5 μg L^{-1}. The GPs algorithm follows the well-known shift of the maximum of Rrs spectra toward higher wavelengths with increasing C. The strategy behind the GPs is sound and insightful, but the switching between power equations leads to the artifacts described above.

With this in mind, a strategy was devised to maximize model precision over the entire chlorophyll concentration range. The functional form of this algorithm, named ocean chlorophyll 4 (OC4), is a modified cubic polynomial relating a band ratio to C (Table 7). The significant departure from previous band ratio algorithms is that the band ratio is determined by whichever ratio, Rrs443/Rrs555, or Rrs490/Rrs555, or Rrs510/Rrs555, is greatest. Thus the OC4 maximum band ratio (MBR) model uses three-band ratios but only a single set of coefficients in a single MCP equation. Similar MBR models for three-band combinations are shown in Table 7.

After tuning to the SeaBAM data set, the OC4 model yields an R^2 of 0.932 and RMS of 0.156 (Figure 5, Table 8). Of the three-band ratios considered, Rrs443/Rrs555 was maximal from lowest C to values of ~0.3 μg L^{-1}; Rrs490/Rrs555 generally dominated between 0.3 and ~1.5 μg L^{-1}; and Rrs510/Rrs555 dominated when C exceeded ~1.5 μg L^{-1} (Figure 7 (top)). Note that the ranges of dominant band ratios overlap by ~10–30%, so there is a smooth transition from Rrs443/Rrs555 to Rrs490/Rrs555 to Rrs510/Rrs555 with decreasing band ratio (Figure 7 (bottom)). This overlap is a desirable property because it implies that discontinuities in frequency distributions of C estimated with OC4 are unlikely.

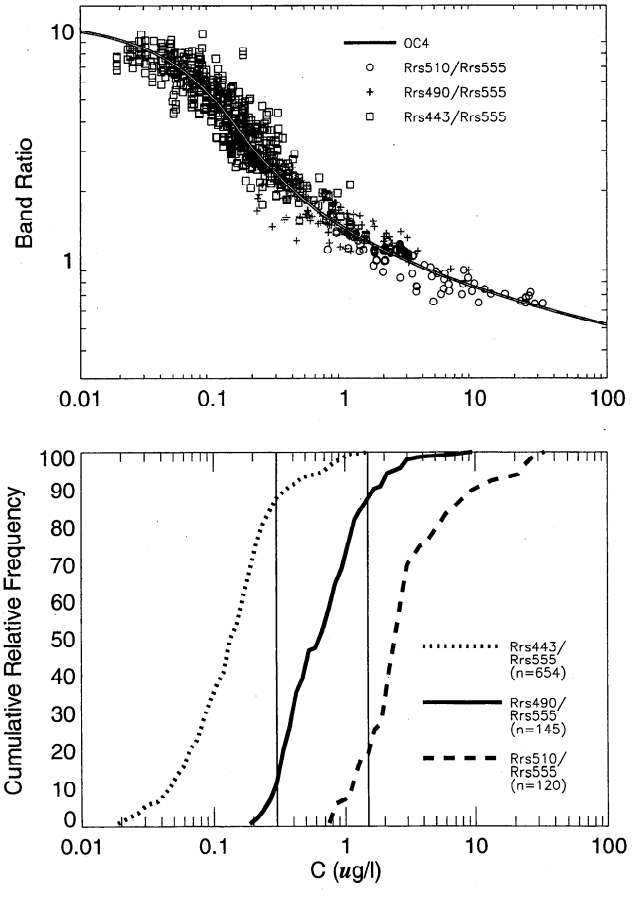

Figure 7. Ocean chlorophyll 4 algorithm. (top) In situ band ratio versus C. OC4 model is represented by curved line. The in situ data are represented by symbols indicating dominant band ratio. (bottom) Cumulative relative frequency distribution of maximum band ratios showing regions of dominance overlap between Rrs443 and Rrs490 and between Rrs490 and Rrs510 (vertical lines at 0.3 and 1.5 μg L^{-1}, respectively).

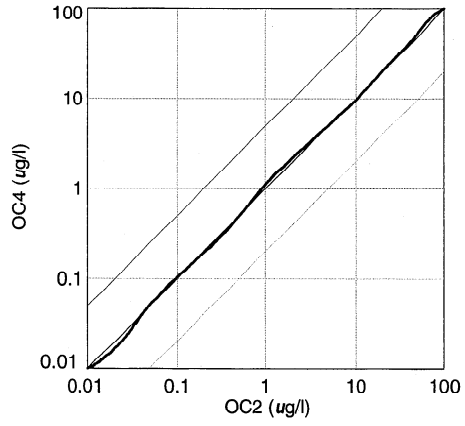

Figure 8. Comparisons between OC4 and OC2. Quantile-quantile plot using simulated radiance ratios with random noise added.

Possible discontinuities in the OC4 model were investigated by subjecting OC4 to a large, continuously varying population of simulated radiance ratios with random noise added. To simulate all band ratios involved in the OC4 model, a MCP equation was derived for each (Table 7), similar to the OC2 model. By inverting these equations it was then possible to generate values for all radiance ratios for any given chlorophyll concentration. Random noise was introduced as a function of C concentration and wavelength [*André and Morel*, 1991] so that several realistic band ratio combinations could be generated at each C concentration. OC2 served as a reference since it does not generate any discontinuities for C between 0.001 and 100 μg L^{-1}. If discontinuities were present, we would expect them to appear near C values of 0.3 μg L^{-1} and 1.5 μg L^{-1}, the general regions where dominant band ratios shift (Figure 7 (bottom)). The q-q comparison shown in Figure 8 indicates that no discontinuities were observed. While further tests are needed, these results indicate that discontinuities do not result from maximum band ratio models such as OC4 when ample overlap exists between adjacent dominant band ratios, as is the case for the SeaWiFS bands.

This maximum band ratio model is a new approach in empirical ocean color algorithms. It has the potential advantage of maintaining the highest possible satellite sensor signal:noise reflectance ratio over a broad range of C concentrations. This aspect is important for passive ocean color sensors aboard satellites since normalized water-leaving radiances retrieved for the 443 nm band, after atmospheric correction, may be quite low or below the sensor detection threshold in chlorophyll-rich coastal water or offshore phytoplankton blooms [*Gordon*, 1987]. The MBR model may also be a useful approach with sensors having many radiance bands (e.g., MODIS or hyperspectral data). MBR models such as OC4 might also be useful to define operationally three ocean realms with respect to trophic status: oligotrophic (<0.3 μg L^{-1}), mesotrophic (0.3–1.5 μg L^{-1}), and eutrophic (>1.5 μg L^{-1}), depending on

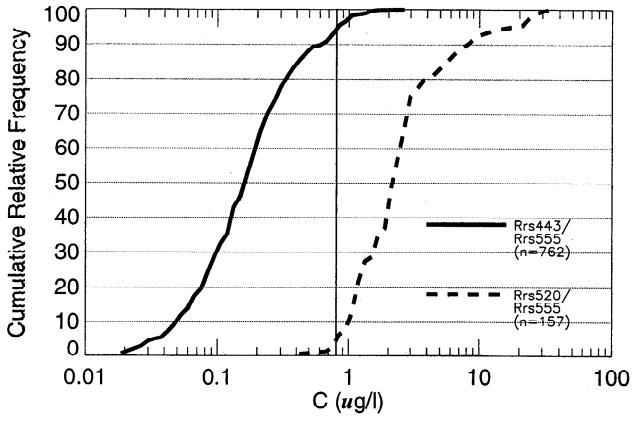

Figure 9. Ocean chlorophyll 3e algorithm. Cumulative relative frequency distribution of maximum band ratios showing regions of dominance overlap between Rrs443 and Rrs520 (vertical line at 0.8 μg L^{-1}).

whether radiances in the 443, 490, or 510 nm bands dominate. The scheme parallels that described by *Antoine et al.* [1996] in a study based on CZCS data.

7.4. Comparisons Between Contemporary and Historical Ocean Color Data

There is keen interest in the ocean color community in assessing long-term (decadal) changes in phytoplankton biomass, and primary productivity in the world oceans, and particular interest in investigating the potential influences of climate change on the magnitude or redistribution of oceanic productivity. The CZCS mission, from 1978 to 1986 [*Yoder et al.*, 1988; *Hooker et al.*, 1992; *McClain*, 1993], established a baseline which may be compared with time series developing or anticipated from contemporary ocean color sensors such as SeaWiFS, OCTS, and MODIS. The comparability between ocean color products embraces various aspects such as sensor characteristics and the way data should be processed and atmospherically corrected. Accurate comparisons also require that similar products be compared (i.e., chlorophyll *a* or [C + P]). The primary CZCS product was [C + P], defined as the sum of chlorophyll *a* + phaeophytin *a* as measured by the fluorometric method. However, since fluorometric phaeophytin determinations are often contaminated by the presence of other forms of chlorophyll [*Lorenzen and Jeffrey*, 1980; *Trees et al.*, 1985; *Vernet and Lorenzen*, 1987], the definition of [C + P] is confounded and imprecise. Also, improvements in HPLC technique and new equipment and methods for fluorometric measurements of chlorophyll *a* [e.g., *Welschmeyer*, 1994] will make it increasingly difficult to compare new "sea truth" chlorophyll data with earlier fluorometric [C + P] estimates used as sea truth for the CZCS mission. For these reasons it appears preferable that comparisons between past and new ocean color data be based on chlorophyll *a*. This implies a need to reprocess the CZCS data using a suitable chlorophyll algorithm.

Because the SeaBAM data set was designed primarily for SeaWiFS, its radiometric composition is less suited for developing a CZCS chlorophyll algorithm. The only CZCS chlorophyll algorithm which can be derived without radiometric adjustments of the SeaBAM data must use 443 and 555 nm (accepting the practical equivalence of Rrs550 and Rrs555). CZCS algorithms employing radiance data from 443, 520, and 550 nm bands using a maximum band ratio approach, such as used in OC4, might be suitable for CZCS reprocessing. For instance, the ocean chlorophyll 3e (OC3e), a three-band algorithm (Table 7), gives better agreement with in situ C than OC2b (Table 8). Even though Rrs490 is not used in OC3e, there appears to be sufficient overlap (~10%) to ensure a smooth transition between dominance by Rrs443/Rrs555 and Rrs520/Rrs555 at low and high concentrations of C, respectively (Figure 9). These results, although promising, must be considered preliminary because in 85 of the 157 instances where Rrs520/Rrs555 was the dominant band ratio (Figure 9), Rrs520 data were estimated from the adjacent Rrs510 band, as described in section 3.4.2.

8. Conclusion

The major focus of this paper was the identification of an algorithm which would allow estimates of in situ C concentrations from SeaWiFS data with the highest possible accuracy and precision over a wide range of bio-optical conditions. While several SeaWiFS-compatible algorithms performed well, the SeaBAM participants [*Firestone and Hooker*, 1998] recommended OC2 as the at-launch SeaWiFS operational chlorophyll *a* algorithm for several reasons. The potential robustness of an algorithm tuned to a large and quality-controlled data set is a major reason. The simple and reversible functional form used by OC2, as well as its statistical and graphical results, were considered superior to other formulations evaluated. Its use of the 490 nm band allows reliable chlorophyll estimates over a wide range of concentrations; Statistical results using the 490/555 band ratio were superior to any other two-band combination (Table 8), as was also reported by *Aiken et al.* [1995]. Despite their evident benefits at low and midconcentrations, two-band algorithms based on the 443/555 ratio were rejected because they gave less precise estimates at high chlorophyll concentrations and because there existed uncertainties associated with atmospheric corrections of sensor radiances at low wavelengths [e.g., *Gordon*, 1987]. While the 443 nm band is nearer the chlorophyll absorption peak than the 490 band and should therefore be more responsive to variation in chlorophyll *a* concentration, the 443 band is also more likely to be influenced by CDOM absorption which decreases exponentially with increasing wavelength [*Bricaud et al.*, 1981; *Roesler et al.*, 1989]. It must also be kept in mind that the 490 nm band works well in band ratio algorithms such as OC2 because of the usually strong correlation between chlorophyll *a*, accessory pigments such as carotenoids, and other covarying substances influencing absorption and reflectance at 490 nm [*Yentsch*, 1960; *Aiken et al.*, 1995]. This covariance of in-water optical properties influencing Rrs at 443 and 490 nm is also implied by their high correlation (R^2 = 0.92, log-transformed data) in the SeaBAM data. Reduced accuracy should therefore be expected from OC2 when phytoplankton pigment composition, pigment-packaging, or bio-optical conditions deviate markedly from those embodied in SeaBAM data [e.g., *Kahru and Mitchell*, 1998b]. Additional drawbacks in using OC2 are that it must be validated using new independent observations and that it, or any algorithm derived from the SeaBAM, will inherit the limitations of the SeaBAM data. While results for OC4 were superior to OC2, OC4 is not so suitable as the initial operational algorithm for SeaWiFS because its use would require accurate atmospheric corrections and on-orbit calibrations in four bands (instead of two), and

this can be assessed only after the collection of sufficient data to validate and fine tune the sensor calibrations.

The SeaBAM data, collected in many oceanic provinces, are reasonably coherent, implying that OC2, derived from these data, should perform reasonably well for tropical, subtropical, and temperate waters. It is unlikely, however, that any single equation would equitably render all bio-optical diversity present in the world ocean. For example, while still not well understood, evidence is accumulating that polar waters have bio-optical properties which differ from lower-latitude oceanic Case I waters [*Mitchell and Holm-Hansen*, 1991; *Mitchell*, 1992; *Cota*, 1997]. Moreover, CDOM, "package effects" [*Kirk*, 1975; *Morel and Bricaud*, 1981] and phytoplankton species composition are important factors ruling light-related phenomena in Antarctica [*Mitchell and Holm-Hansen*, 1991; *DiTullio and Smith*, 1996; *Claustre et al.*, 1997]. In the context of results presented by *Mitchell* [1992], any algorithm derived from low and midlatitude Oceanic Case I data might be expected to underestimate C concentration in Antarctic waters. On the basis of some recent optical measurements collected in the Ross Sea, the converse trend may also be observed [*Schieber*, 1998]. Coccolithophorid blooms [*Balch et al.*, 1991; *Ackleson et al.*, 1994; *Brown and Yoder*, 1994] and perhaps cyanobacteria blooms, are other examples of situations where algorithms derived from Case I waters dominated by other species are likely to fail. Case II waters present additional complications and challenges as they can be of different types depending on whether they are dominated by CDOM, nonchlorophyllous particles, or a variable mix of both [*Carder et al.*, 1989; see also *Siegel and Michaels*, 1996 about CDOM in Case I waters]. Specific algorithms or different parameterizations will be required to handle these kinds of situations or regions.

Algorithms designed for use at global scales are likely to be less accurate at local and regional scales than those developed for particular regions or bio-optical conditions, and vice versa. For local and regional studies the use of a regionally parameterized algorithm makes sense, but for the purpose of producing chlorophyll maps at global scale, it would be very difficult to use a patchwork of regional algorithms because of consistency and transition issues between algorithms, and the fundamental lack of data in many regions. The calibration and validation activities planned for SeaWiFS will allow extensive testing of OC2 and other algorithms and will determine whether algorithm modifications or adjustments are required to achieve the most reliable global estimates of phytoplankton biomass.

Notation

$a(\lambda)$	absorption coefficient, m^{-1}.
$a_{dm}(\lambda)$	absorption coefficient for particulate and dissolved matter, m^{-1}.
$a_g(\lambda)$	absorption coefficient for CDOM, m^{-1}.
$a_{ph}(\lambda)$	absorption coefficient for phytoplankton, m^{-1}.
$a_w(\lambda)$	absorption coefficient of pure seawater, m^{-1}.
$b_b(\lambda)$	backscattering coefficient, m^{-1}.
$b_{bp}(\lambda)$	backscattering coefficient for particles, m^{-1}.
$b_{bw}(\lambda)$	backscattering coefficient of pure seawater, m^{-1}.
C	chlorophyll *a* concentration, $\mu g\ L^{-1}$.
$E_d(z, \lambda)$	downwelling irradiance at depth z and wavelength λ, $\mu W\ cm^{-2}\ nm^{-1}$.
$\overline{Fo}(\lambda)$	mean extraterrestrial solar irradiance, $\mu W\ cm^{-2}\ nm^{-1}$.
$L_u(z, \lambda)$	upwelling radiance at depth z and wavelength λ, $mW\ cm^{-2}\ nm^{-1}\ sr^{-1}$.
$L_w(\lambda)$	water-leaving radiance, $\mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$.
Lwn(λ)	normalized water-leaving radiance, $mW\ cm^{-2}\ nm^{-1}\ sr^{-1}$.
P	phaeophytin concentration, $\mu g\ L^{-1}$.
Rrs	remote sensing reflectance, sr^{-1}.

Cruises and field studies

AMT	Atlantic meridional transect.
BBOP	Bermuda BioOptics Project.
CalCOFI	California Cooperative Oceanic Fisheries Investigation.
EQPAC	equatorial Pacific.
JGOFS	Joint Global Ocean Flux Study.
MOCE	Marine Optical Characterization Experiment.
NABE	North Atlantic Bloom Experiment.
WOCE	World Ocean Circulation Experiment.

Satellite ocean color sensors

CZCS	coastal zone color scanner.
MERIS	medium-resolution imaging spectrometer.
MODIS	moderate-resolution imaging spectroradiometer.
OCTS	ocean color temperature sensor.
POLDER	Polarization and Directionality of Earth Reflectances.
SeaWiFS	Sea-viewing Wide Field-of-view Sensor.

Acknowledgments. We are extremely grateful to S. Hawes, J. Landry, G. Moore, J. Mueller, B. Schieber, and J. Yoder for their assistance. The authors thank J. Aiken, G. Moore, and S. Hooker for providing the AMT data, and anonymous referees for their helpful comments. This work was supported by various grants and contracts from the NASA Ocean Biogeochemistry Program.

References

Ackleson, S. G., W. M. Balch, and P. M. Holligan, Response of water-leaving radiance to particulate calcite and chlorophyll *a* concentrations: A model for Gulf of Maine coccolithophore blooms, *J. Geophys. Res.*, *99*, 7483–7499, 1994.

Acker, J. G., The heritage of SeaWiFS: A retrospective on the CZCS NIMBUS Experiment Team (NET) Program, *NASA Tech. Memo. 104566*, vol. 21, edited by S. B. Hooker and E. R. Firestone, 44 pp., NASA Goddard Space Flight Cent., Greenbelt, Md., 1994.

Aiken, J., G. F. Moore, and P. M. Holligan, Remote sensing of oceanic biology in relation to global climate change, *J. Phycol.*, *28*, 579–590, 1992.

Aiken, J., G. F. Moore, C. C. Trees, S. B. Hooker, and D. K. Clark, The SeaWiFS CZCS-type pigment algorithm, *NASA Tech. Memo. 104566*, vol. 29, 34 pp., 1995.

André, J.-M., and A. Morel, Atmospheric corrections and interpretation of marine radiances in CZCS imagery, revisited, *Oceanol. Acta*, *14*, 3–22, 1991.

Antoine, D., J.-M. André, and A. Morel, Oceanic primary production, 2, Estimation at global scale from satellite (CZCS) chlorophyll, *Global Biochem. Cycles*, *10*(1), 57–69, 1996.

Austin, R. W., The remote sensing of spectral radiance from below the ocean surface, in *Optical Aspects of Oceanography*, edited by N. G. Jerlov and E. S. Nielsen, pp. 317–344, Academic, San Diego, Calif., 1974.

Balch, W. R., P. M. Holligan, S. G. Ackleson, and K. J. Voss, Biological and optical properties of mesoscale coccolithophore blooms in the Gulf of Maine, *Limnol. Oceanogr.*, *36*, 629–643, 1991.

Balch, W., R. Evans, J. Brown, G. Feldman, C. McClain, and W. Esaias, The remote sensing of primary productivity, Use of a new data compilation to test satellite algorithms, *J. Geophys. Res.*, *97*, 2279–2293, 1992.

Behrenfeld, M. J., and P. G. Falkowski, Photosynthetic rates derived

from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, *42*(1), 1–20, 1997.

Bricaud, A., and D. Stramski, Spectral absorption coefficients of living phytoplankton and nonalgal biogenous matter: A comparison between the Peru upwelling and the Sargasso Sea, *Limnol. Oceanogr.*, *35*(3), 562–582, 1990.

Bricaud, A., A. Morel, and L. Prieur, Absorption by dissolved organic matter of the sea (yellow substances) in the UV and visible domains, *Limnol. Oceanogr.*, *26*, 43–53, 1981.

Bricaud, A., M. Babin, A. Morel, and H. Claustre, Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization, *J. Geophys. Res.*, *100*, 13,321–13,332, 1995.

Brown, C. W., and J. A. Yoder, Coccolithophorid blooms in the global ocean, *J. Geophys. Res.*, *99*, 7467–7482, 1994.

Campbell, J. W., The lognormal distribution as a model for bio-optical variability in the sea, *J. Geophys. Res.*, *100*, 13,237–13,254, 1995a.

Campbell, J. W., Level-3 SeaWiFS data products: Spatial and temporal binning algorithms, *NASA Tech. Memo. 104566*, vol. 32, 73 pp. plus color plates, 1995b.

Campbell, J. W., and J. E. O'Reilly, Role of satellites in estimating primary productivity on the northwest Atlantic continental shelf, *Cont. Shelf Res.*, *8*(2), 179–204, 1988.

Carder, K. L., and R. G. Steward, A remote-sensing reflectance model of a red tide dinoflagellate off West Florida, *Limnol. Oceanogr.*, *30*, 286–298, 1985.

Carder, K. L., R. G. Steward, J. H. Paul, and G. A. Vargo, Relationships between chlorophyll and ocean color constituents as they affect remote-sensing reflectance models, *Limnol. Oceanogr.*, *31*(2), 403–413, 1986.

Carder, K. L., R. G. Steward, G. R. Harvey, and P. B. Ortner, Marine humic and fulvic acids: Their effects on remote sensing of ocean chlorophyll, *Limnol. Oceanogr.*, *34*(1), 68–81, 1989.

Chambers, J. M., W. S. Cleveland, B. Kleiner, and P. A. Tukey, *Graphical Methods for Data Analysis*, 395 pp., Wadsworth, Belmont, Calif., 1983.

Clark, D. W., Phytoplankton pigment algorithms for the Nimbus-7 CZCS, in *Oceanography From Space*, edited by J. F. R. Gower, pp. 227–237, Plenum, New York, 1981.

Clark, G. L., G. C. Ewing, and C. J. Lorenzen, Spectra of backscattered light from the sea obtained from aircraft as a measure of chlorophyll concentration, *Science*, *167*, 1119–1121, 1970.

Claustre, H., M. A. Moline, and B. B. Prézelin, Sources of variability in the column photosynthetic cross section for Antarctic Coastal Waters, *J. Geophys. Res.*, *102*, 25,047–25,060, 1997.

Cleveland, J. S., Regional models for phytoplankton absorption as a function of chlorophyll a concentration, *J. Geophys. Res.*, *100*, 13,333–13,344, 1995.

Cota, G. F., Bio-optical properties of Arctic Coastal Waters, paper presented at the ASLO Meeting on Aquatic Science, Am. Soc. of Limnol. and Oceanogr., Santa Fe, N. M., Feb. 10–14, 1997.

Denman, K. L., and M. R. Abbott, Time evolution of surface chlorophyll patterns from cross-spectrum analysis of satellite color images, *J. Geophys. Res.*, *93*, 6789–6798, 1988.

Denman, K. L., and M. R. Abbott, Timescales of pattern evolution from cross-spectrum analysis of advanced very high resolution radiometer and coastal zone color scanner imagery, *J. Geophys. Res.*, *99*, 7433–7442, 1994.

DiTullio, G. R., and W. O. Smith Jr., Spatial patterns in phytoplankton biomass and pigment distributions in the Ross Sea, *J. Geophys. Res.*, *101*, 18,467–18,477, 1996.

Esaias, W. E., G. C. Feldman, R. Frouin, W. W. Gregg, S. B. Hooker, and C. R. McClain, Goddard Space Flight Center Ocean Color Group, 1995, in *Ocean Color Multisensor Data Evaluation and Utilization Plan*, 72 pp., NASA, 1995.

Evans, R. H., and H. R. Gordon, Coastal zone color scanner "system calibration": A retrospective examination, *J. Geophys. Res.*, *99*, 7293–7307, 1994.

Feldman, G., et al., Ocean Color: Availability of the global data set, *Eos*, *70*, 634–641, 1989.

Firestone, E. R., and S. B. Hooker (Eds.), SeaWiFS prelaunch technical report series final cumulative index, vol. 43. *NASA Tech. Memo., TM-1998-104566* 4–8, 1998.

Garver, S. A., Variability in ocean color observations and their use in the study of upper open ocean ecosystem dynamics, 118 pp., Ph.D. thesis, Dep. Phys. Geogr., Univ. of Calif., Santa Barbara, 1997.

Garver, S. A., and D. A. Siegel, Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation, 1, Time series from the Sargasso Sea, *J. Geophys. Res.*, *102*, 18,607–18,625, 1997.

Gordon, H. R., Calibration requirements and methodology for remote sensors viewing the ocean in the visible, *Remote Sens. Environ.*, *22*, 103–126, 1987.

Gordon, H. R., and K. Ding, Self-shading of in-water optical instruments, *Limnol. Oceanogr.*, *37*(3), 491–500, 1992.

Gordon, H. R., and A. Morel, Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery, A review, in *Lecture Notes on Coastal and Estuarine Studies*, 113 pp., Springer-Verlag, New York, 1983.

Gordon, H. R., D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans, and W. W. Broenkow, Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates, *Appl. Opt.*, *22*, 20–36, 1983.

Gordon, H. R., O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, A semianalytic radiance model of ocean color, *J. Geophys. Res.*, *93*, 10,909–10,924, 1988.

Hojerslev, N., Assessement of some suggested algorithms for sea colour and surface chlorophyll, in *Oceanography From Space*, edited by J. F. R. Gower, pp. 347–353, Plenum, New York, 1981.

Holm-Hansen, O., C. J. Lorenzen, R. W. Holmes, and J. D. H. Strickland, Fluorometric determination of chlorophyll, *J. Cons. Cons. Int. Explor. Mer.*, *30*(1), 3–15, 1965.

Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, An overview of SeaWiFS and ocean color, *NASA Tech. Memo. 104566*, vol. 1, 24 pp., 1992.

Hooker, S. B., C. R. McClain, and A. Holmes, Ocean Color Imaging: CZCS to SeaWiFS, *Mar. Technol. Soc.*, *27*(1), 2–15, 1993.

Hooker, S. B., C. R. McClain, J. K. Firestone, T. L. Westphal, E.-N. Yeh, and Y. Ge, The SeaWiFS bio-optical archive and storage system (SeaBASS), 1, *NASA Tech. Memo. 104566*, vol. 20, 40 pp., 1994.

Hovis, W. A., The NIMBUS-7 Coastal Zone Color Scanner (CZCS) program, *Oceanography From Space*, edited by J. F. R. Gower, pp. 213–225, Plenum, New York, 1981.

Kahru, M., and B. G. Mitchell, Evaluation of instrument self-shading and environmental errors on ocean color algorithms, in *Ocean Optics XIV, Proc. SPIE Int. Soc. Opt. Eng.*, in press, 1998a.

Kahru, M., and B. G. Mitchell, Spectral reflectance and absorption of a massive red tide off southern California, *J. Geophys. Res.*, in press, 1998b.

Kirk, J. T. O., A theoretical analysis of the contribution of algal cells to the attenuation of light within waters, II, Spherical cells, *New Phytol.*, *75*, 21–36, 1975.

Laws, E. A., and J. W. Archie, Appropriate use of regression analysis in marine biology, *Mar. Biol.*, *65*, 13–16, 1981.

Lee, Z., K. L. Carder, S. K. Hawes, R. G. Steward, T. G. Peacock, and C. O. Davis, Model for the interpretation of hyperspectral remote-sensing reflectance, *Appl. Opt.*, *33*(24), 5721–5732, 1994.

Lee, Z. P., K. L. Carder, S. K. Hawes, R. G. Steward, T. G. Peacock, and C. O. Davis, Remote-sensing reflectance and inherent optical properties of oceanic waters derived from above water measurements, *Proc. SPIE*, *2963*, 160–166, 1996.

Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill, An estimate of global primary production in the ocean from satellite radiometer data, *J. Plank. Res.*, *17*(6), 1245–1271, 1995.

Lorenzen, C. J., and S. W. Jeffrey, Determination of chlorophyll in seawater, *UNESCO Tech. Pap., Mar. Sci.*, *35*, 20, 1980.

McClain, C. R., Review of major CZCS applications: U.S. case studies, *Ocean Colour: Theory and Applications in a Decade of CZCS Experience*, edited by V. Barale and P. M. Schlittenhardt, pp. 167–188, ECSC, EEC, EAEC, Brussels and Luxembourg, 1993.

McClain, C. R., and E.-N. Yeh, CZCS bio-optical algorithm comparison, in *Case Studies for SeaWiFS Calibration and Validation*, 1, *NASA Tech. Memo. 104566*, vol. 13, pp. 3–8, 1994.

Mitchell, B. G., Predictive bio-optical relationships for polar oceans and marginal ice zones, *J. Mar. Syst.*, *3*, 91–105, 1992.

Mitchell, B. G., Coastal zone color scanner retrospective, *J. Geophys. Res.*, *99*, 7291–7292, 1994.

Mitchell, B. G., and O. Holm-Hansen, Bio-optical properties of Antarctic Peninsula waters: Differentiation from temperate ocean models, *Deep Sea Res.*, *38*(89), 1009–1028, 1991.

Mitchell, B. G., and M. Kahru, Algorithms for SeaWiFS developed

with the CalCOFI data set, *CalCOFI Rep. 39*, 26 pp., Calif. Coop. Oceanic Fish. Invest. Rep., LaJolla, Calif., 1998.

Mitchell, B. G., and D. A. Kiefer, Variability in pigment specific particulate fluorescence and absorption spectra in the northeastern Pacific Ocean, *Deep Sea Res., 35*(5), 665–689, 1988.

Morel, A., Optical properties of pure water and pure seawater, in *Optical Aspects of Oceanography*, edited by N. G. Jerlov and E. Steemann Nielsen, pp. 1–24, Academic, San Diego, Calif., 1974.

Morel, A., Optical modeling of the upper ocean in relation to its biogenous matter content (Case I waters), *J. Geophys. Res., 93*, 10,749–10,768, 1988.

Morel, A., Optical properties of oceanic Case 1 waters, revisited, in *Ocean Optics XIII*, edited by S. G. Ackleson and R. Frouin, *Proc. SPIE, 2963*, 108–114, 1997.

Morel, A., and D. Antoine, Heating rate within the upper ocean in relation to its bio-optical state, *J. Phys. Oceanogr., 24*(7), 1652–1665, 1994.

Morel, A., and J.-F. Berthon, Surface pigments, algal biomass profiles, and potential production of the eutrophic layer: Relationships reinvestigated in view of remote-sensing applications, *Limnol. Oceanogr., 34*(8), 1545–1562, 1989.

Morel, A., and A. Bricaud, Theoretical results concerning light absorption in a discrete medium and application to specific absorption of phytoplankton, *Deep Sea Res., 28*, 1375–1393, 1981.

Morel, A., and L. Prieur, Analysis of variations in ocean color, *Limnol. Oceanogr., 22*(4), 709–722, 1977.

Mueller, J. L., and R. W. Austin, Ocean optics protocols for SeaWiFS validation, Revision 1, *NASA Tech. Memo, 104566*, vol. 25, 67 pp., 1995.

Muller-Karger, F. E., C. R. McClain, R. N. Sambrotto, and G. C. Ray, A comparison of ship and coastal zone color scanner mapped distribution of phytoplankton in the southeastern Bering Sea, *J. Geophys. Res., 95*, 11,483–11,499, 1990.

Neckel, H., and D. Labs, The solar radiation between 3300 and 12500 Å, *Solar Phys., 90*, 205–258, 1984.

Platt, T., C. Caverhill, and S. Sathyendranath, Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic, *J. Geophys. Res., 96*, 15,147–15,159, 1991.

Pope, R. M., and E. S. Fry, Absorption spectrum (380–700 nm) of pure water, II, Integrating cavity measurements, *Appl. Opt., 36*(33) 8710–8723, 1997.

Press, W. H., and S. A. Teukolsky, Fitting straight line data with errors in both coordinates, *Comput. Phys., 6*, 274–276, 1992.

Ricker, W. E., Linear regressions in fishery research, *J. Fish. Res. Board Can., 30*, 409–434, 1973.

Roesler, C. S., M. J. Perry, and K. L. Carder, Modeling in situ phytoplankton absorption from total absorption spectra in productive inland waters, *Limnol. Oceanogr., 34*, 1510–1523, 1989.

Sathyendranath, S., L. Prieur, and A. Morel, A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters, *Int. J. Remote Sens., 10*, 1373–1394, 1989.

Sathyendranath, S., T. Platt, E. P. W. Horne, W. G. Harrison, O. Ulloa, R. Outerbridge, and N. Hoepffner, Estimation of new production in the ocean by compound remote sensing, *Nature, 353*, 129–133, 1991.

Schieber, B. D., Ross Sea variability and performance of Case I ocean color algorithms, M.S. thesis, 96 pp., Mar. Estuarine and Environ. Stud., Univ. of Maryland, College Park, 1998.

Siegel, D. A., and A. F. Michaels, Quantification of non-algal light attenuation in the Sargasso Sea: Implications for biogeochemistry and remote sensing, *Deep Sea Res., 43*(2–3), 321–345, 1996.

Siegel, D. A., M. C. O'Brien, J. C. Sorensen, D. A. Konnoff, E. A. Brody, J. L. Mueller, C. O. Davis, W. J. Rhea, and S. B. Hooker, Results of the SeaWiFS Data Analysis Round-Robin (DARR-94), July 1994, *NASA Tech. Memo. 104566*, vol. 26, 58 pp., 1995.

Smith, R. C., and K. S. Baker, The bio-optical state of ocean waters and remote sensing, *Limnol. Oceanogr., 23*, 247–249, 1978.

Smith, R. C., and K. S. Baker, Optical properties of the clearest natural waters (200–800 nm), *Appl. Opt., 20*(2), 177–184, 1981.

Smith, R. C., and K. S. Baker, Oceanic chlorophyll concentrations as determined by satellite (Nimbus-7 coastal zone color scanner), *Mar. Biol., 66*, 269–279, 1982.

Smith, R. C., R. W. Eppley, and K. S. Baker, Correlation of primary production as measured aboard ship in Southern California coastal waters and as estimated from satellite chlorophyll images, *Mar. Biol., 66*, 281–288, 1982.

Sosik, H. M., and B. G. Mitchell, Light absorption by phytoplankton, photosynthetic pigments and detritus in the California Current System, *Deep Sea Res., 42*(10), 1717–1748, 1995.

Trees, C. C., M. C. Kennicutt, and J. M. Brooks, Errors associated with the standard fluorimetric determination of chlorophylls and phaeopigments, *Mar. Chem., 17*, 1–12, 1985.

Vernet, M., and C. J. Lorenzen, The presence of chlorophyll b and the estimation of phaeopigments in marine phytoplankton, *J. Plank. Res., 9*(2), 255–265, 1987.

Welschmeyer, N. A., Fluorometric analysis of chlorophyll *a* in the presence of chlorophyll *b* and phaeopigments, *Limnol. Oceanogr., 39*(8), 1985–1992, 1994.

Yentsch, C. S., The influence of phytoplankton pigments on the colour of seawater, *Deep Sea Res., 7*, 1–9, 1960.

Yentsch, C. S., and D. W. Menzel, A method for the determination of phytoplankton, chlorophyll, and phaeophytin by fluorescence, *Deep Sea Res., 10*, 221–231, 1963.

Yoder, J. A., W. E. Esaias, G. C. Feldman, and C. R. McClain, Satellite ocean color status report, *Oceanogr. Mag., 1*, 18–20, 35, 1988.

Yoder, J. A., C. R. McClain, G. C. Feldman, and W. E. Esaias, Annual cycles of phytoplankton chlorophyll concentrations in the global ocean: A satellite view, *Global Biochem. Cycles, 7*, 181–193, 1993.

K. L. Carder, University of South Florida, St. Petersburg, FL 33701.

S. A. Garver, California State Polytechnic University, Pomona, CA 91768.

M. Kahru and B. G. Mitchell, Scripps Institution of Oceanography, UCSD, La Jolla, CA 92093.

S. Maritorena, Universities Space Research Association, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

C. McClain, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

J. E. O'Reilly, NOAA National Marine Fisheries Service, Northeast Fisheries Science Center, 28 Tarzwell Drive, Narragansett, RI 02882. (e-mail: oreilly@fish1.gso.uri.edu)

D. A. Siegel, Department of Geography, University of California, Santa Barbara, CA 93106.

(Received October 28, 1997; revised June 3, 1998; accepted June 24, 1998.)