Chapter 2

Ocean Color Chlorophyll *a* Algorithms for SeaWiFS, OC2 and OC4: Version 4

John E. O'Reilly
NOAA, National Marine Fisheries Service, Narragansett, Rhode Island

Stephane Maritorena, David A. Siegel, Margaret C. O'Brien, and Dierdre Toole
ICESS/University of California Santa Barbara, Santa Barbara, California

B. Greg Mitchell and Mati Kahru
Scripps Institution of Oceanography, University of California San Diego, California

Francisco P. Chavez and P. Strutton
Monterey Bay Aquarium Research Institute, Moss Landing, California

Glenn F. Cota
Old Dominion University, Norfolk, Virginia

Stanford B. Hooker and Charles R. McClain
NASA Goddard Space Flight Center, Greenbelt, Maryland

Kendall L. Carder and Frank Muller-Karger
University of South Florida, St. Petersburg, Florida

Larry Harding and Andrea Magnuson
Horn Point Laboratory, University of Maryland, Cambridge, Maryland

David Phinney
Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine

Gerald F. Moore and James Aiken
Plymouth Marine Laboratory, Plymouth, United Kingdom

Kevin R. Arrigo
Department of Geophysics, Stanford University, Stanford, California

Ricardo Letelier
College of Oceanic and Atmospheric Sciences, Oregon State University

Mary Culver
NOAA, Coastal Services Center, Charleston, South Carolina

Abstract

This chapter describes the revisions (version 4) to the ocean chlorophyll two- and four-band algorithms as well as the very large *in situ* data set used to update these algorithms for use in the third reprocessing of SeaWiFS data. The *in situ* data set is substantially larger ($N = 2{,}853$) than was used to develop earlier versions of OC2 and OC4, includes samples from a greater variety of bio-optical provinces, and better represents oligotrophic and eutrophic waters. The correlation between chlorophyll *a* concentration, C_a, estimated using OC4 and *in situ* C_a ($\tilde{C}_a$) estimated from fluorometric and HPLC analyses was slightly higher than that for OC2. Also, OC4 would be expected to perform better than OC2, when applied to satellite-derived, water-leaving radiances retrieved from oligotrophic and eutrophic areas. Variations of the OC4 algorithm are provided for other ocean color sensors to facilitate comparisons with SeaWiFS.

2.1 INTRODUCTION

The accuracy, precision, and utility of an empirical ocean color algorithm for estimating global chlorophyll *a* distributions depends on the characteristics of the algorithm and the *in situ* observations used to develop it. The empirical pigment and chlorophyll algorithm widely used in the processing of the global CZCS data set was developed using fewer than 60 *in situ* radiance and chlorophyll *a* pigment observations (Evans and Gordon 1994). Since the CZCS period, a number of investigators have measured *in situ* remote sensing reflectance, $\tilde{R}_{rs}(\lambda)$, and *in situ* chlorophyll *a* concentration, $\tilde{C}_a$, from a variety of oceanic provinces. In 1997, the SeaBAM group (Firestone and Hooker 1998) assembled a large $\tilde{R}_{rs}(\lambda)$ and $\tilde{C}_a$ data set containing 919 observations. This data set was used to evaluate the statistical performance of chlorophyll *a* algorithms and to develop the ocean chlorophyll 2-band (OC2) and ocean chlorophyll 4-band (OC4) algorithms (O'Reilly et al. 1998).

OC2 predicts C_a from the $R_{rs}(490)/R_{rs}(555)$ band ratio using a modified cubic polynomial formulation. Hereafter, the R_{rs} ratio constructed from band A divided by band B is indicated by R_B^A, i.e., the $R_{rs}(490)/R_{rs}(555)$ band ratio is represented by R_{555}^{490}. OC4 also relates band ratios to chlorophyll *a* with a single polynomial function, but it uses the maximum band ratio (MBR) determined as the greater of the R_{555}^{443}, R_{555}^{490}, or R_{555}^{510} values. OC2 was employed as the standard chlorophyll *a* algorithm by the SeaWiFS project following the launch of SeaWiFS in September, 1997. Although the statistical characteristics of OC4 were superior to those of OC2, the SeaBAM group recommended using the simpler 2-band OC2 at launch.

With the goal of improving estimates in chlorophyll-rich waters, OC2 was revised (version 2) based on an expanded data set of 1,174 *in situ* observations (Maritorena and O'Reilly, 2000) and applied by the SeaWiFS project in the second data reprocessing (McClain 2000). Additional *in situ* data have become available as the result of new programs (e.g., SIMBIOS) and the continuation and expansion of ongoing field campaigns. These new data increase the variety of bio-optical provinces represented in the original data set and fill in regions of the $R_{rs}(\lambda)$ and C_a domain which were not previously well represented. Also, results from over 2.5 years of SeaWiFS data are now available to assess the overall performance of the SeaWiFS instrument and identify areas where improvements are needed in the processing of satellite ocean color data (McClain 2000).

An update to the OC2 and OC4 chlorophyll algorithms for SeaWiFS are presented here along with a description of the major features of the very large *in situ* data set used to refine these models, and a comparison of the updated algorithms with earlier versions. MBR chlorophyll algorithms for several other satellite ocean color sensors are also provided to facilitate intercomparisons with SeaWiFS.

2.2 THE *IN SITU* DATA SET

A very large data set of $\tilde{R}_{rs}(\lambda)$ and $\tilde{C}_a$ measurements were assembled for the purpose of updating ocean color chlorophyll algorithms for SeaWiFS calibration and validation activities. The data sets and the principal investigators responsible for collecting the data are provided in Table 2. Table 3 gives the location and acquisition time periods of the data, along with an indication of the number of observations, how the chlorophyll *a* concentration was determined (fluorometry or HPLC), and how the radiometric observations were made (above- or in-water). The wavelengths of the latter are presented in Table 4.

The data set has a total of 2,853 *in situ* observations. It is the largest ever assembled for algorithm refinement, and represents a large diversity of bio-optical provinces. The C_a data are derived from a mixture of HPLC and fluorometric measurements from surface samples: 28% and 72% of the data, respectively (Table 3). The C_a values range from 0.008–90 $\mathrm{mg\,m^{-3}}$. The relative frequency distribution of C_a has a primary and secondary peak at 0.2 $\mathrm{mg\,m^{-3}}$ and approximately 1 $\mathrm{mg\,m^{-3}}$, respectively (Fig. 1). Oceanic regions with C_a between 0.08–3 $\mathrm{mg\,m^{-3}}$ are relatively well represented. There are 238 observations of C_a exceeding 5 $\mathrm{mg\,m^{-3}}$ and 116 observations with C_a less than 0.05 $\mathrm{mg\,m^{-3}}$. A comparison of the C_a frequency distribution with those from previous versions (O'Reilly et al. 1998 and Maritorena and O'Reilly 2000) shows that oligotrophic and eutrophic waters are relatively better represented in the current data set. The present data set also has a more equitable distribution over a broader range of C_a (i.e., 0.08–3 $\mathrm{mg\,m^{-3}}$).

Measurements of $R_{rs}(\lambda)$ were made using both above- and in-water radiometers: 88% and 12% of the data, respectively (Table 3). In several subsets, multiple R_{rs} measurements were taken at stations where only a single C_a measurement was made. For these subsets (BBOP9293, WOCE, EQPAC, NABE, Goa97, Ber96, Ber95, Lab98, Lab97, Lab96, Res96, Res952, Res94), the median R_{rs} value was paired with the solitary C_a observation and added to the data set.

Except in a limited number of circumstances, band ratios determined from the median R_{rs} values agreed well with the individual band ratios. Several subsets, however, required adjustments to the $\tilde{R}_{rs}(\lambda)$ values to conform with the SeaWiFS band set. The $\tilde{R}_{rs}(555)$ value was estimated from the $\tilde{R}_{rs}(565)$ measurement for the BBOP9293 and WOCE data using an equation derived from concurrent measurements of $\tilde{R}_{rs}(555)$ and $\tilde{R}_{rs}(565)$ from 1994–1995 BBOP surveys (equation 2 from O'Reilly et al. 1998). The $\tilde{R}_{rs}(555)$ value for the CBAY-MAB subset was computed by averaging the $\tilde{R}_{rs}(550)$ and $\tilde{R}_{rs}(560)$ values. The $\tilde{R}_{rs}(510)$ value was estimated from the $\tilde{R}_{rs}(520)$ values for the WOCE, EQPAC, NABE, and BBOP9293 data sets using the following conversion equation based on Morel and

Table 2. The data sets and the investigators responsible for the data collection activity.

No.	*Data Set*	*Investigators*
1	Roaverrs 97	Arrigo, K.
2	CARDER	Carder, K.
3	CARDER	Carder, K.
4	CARDER	Carder, K.
5	CARDER	Carder, K.
6	MF0796	Carder, K.
7	TOTO	Carder, K.
8	CoBOP	Carder, K., J. Patch
9	EcoHAB	Carder, K., J. Patch
10	Global	Chavez, F.
11	MBARI EqPac	Chavez, F., P. Strutton
12	MOCE-1	Clark,D.
13	MOCE-2	Clark,D.
14	MOCE-4	Clark,D., C. Trees
15	Goa97	Cota, G.
16	Ber95	Cota, G., Saitoh
17	Ber96	Cota, G., Saitoh
18	Lab96	Cota, G., G. Harrison
19	Lab97	Cota, G., G. Harrison
20	Res94	Cota, G.
21	Res95-2	Cota, G.
22	Res96	Cota, G.
23	Res98	Cota, G.
24	CSC	Culver, M., A. Subramaniam
25	CSC	Culver, M., A. Subramaniam
26	CSC	Culver, M., A. Subramaniam
27	EqPac	Davis, C.
28	NABE	Davis, C.
29	CB-MAB	Harding, L., A. Magnuson
30	AMT-1	Hooker, S., G. Moore
31	AMT-2	G.Moore, S. Hooker
32	AMT-3	Hooker, S., J. Aiken, S. Maritorena
33	AMT-4	Hooker,S.,S.Maritorena
34	AMT-5	Hooker, S., S. Maritorena
35	AMT-6B	Moore, G., S. Hooker, S. Maritorena
36	AMT-6	Hooker, S., S. Maritorena
37	AMT-7	Hooker, S., S. Maritorena
38	AMT-8	Hooker, S., S. Maritorena
39	HOT	Letelier, R., R. Bidigare, D. Karl
40	WOCE	Marra, J.
41	WOCE	Marra, J.
42	CalCOFI	Mitchell, G., M. Kahru
43	CalCOFI	Mitchell, G., M. Kahru
44	RED9503	Mitchell, G., M. Kahru
45	AI9901	Mitchell, G., M. Kahru
46	JES9906	Mitchell, G., M. Kahru
47	CARIACO	Muller-Karger, F., R. Varela, J. Akl, A. Rondon, G. Arias
48	NEGOM	Muller-Karger, F., C. Hu, D. Biggs, B. Nababan, D. Nadeau, J. Vanderbloemen
49	ORINOCO	Muller-Karger, F., R. Varela, J. Akl, A. Rondon, G. Arias
50	GOM	Phinney, D.
51	Arabian Sea	Phinney, D.
52	FL-Cuba	Phinney, D.
53	BBOP 9293	Siegel, D., M. O'Brien, N. Nelson, T. Michaels

Table 2. (cont.) The data sets and the investigators responsible for the data collection activity.

No.	*Data Set*	*Investigators*
54	BBOP 9499	Siegel, D., M. O'Brien, N. Nelson, T. Michaels
55	Plumes & Blooms	Siegel, D., D. Toole, L. Mertes, R. Smith, L. Washburn, M. Brzezinski
56	NABE	Trees, C.
57	CoASTS	Zibordi, G.

Table 3. Data sources, locations, and acquisition dates (summarized by the three-letter month abbreviation and the two-digit year) of the global data set. N is the number of samples, the C_a column indicates the method(s) used for chlorophyll a determination (F for fluorometry and H for HPLC), and the R_{rs} column indicates the type of radiometric used for measuring remote sensing reflectance (A for above water and B for below water).

No.	*Data Set*	*Location*	*Dates*	N	C_a	R_{rs}
1	Roaverrs 97	Ross Sea	Dec97–Jan98	73	H	B
2	CARDER	North Atlantic	Aug91	87	F	A
3	CARDER	Pacific	Jul92		F	A
4	CARDER	Gulf of Mexico	Apr93		F	A
5	CARDER	Arabian Sea	Nov94, Jun95		F	A
6	MF0796	Bering Sea	Apr96	22	F	A
7	TOTO	Bahamas	Apr98, Apr99	26	F	A
8	CoBOP	Bahamas	May98, May–Jun99	43	F	A
9	EcoHAB	W. Florida Shelf	Mar99–Mar00 (6 Surveys)	57	F	A
10	Global	Global	Nov93–Jul98 (18 Surveys)	284	F	B
11	MBARI EqPac	Equatorial Pacific	Oct97–Nov99 (6 Surveys)	89	F	B
12	MOCE-1	Monterey Bay	Sep92	8	H	B
13	MOCE-2	Gulf of California	Apr93	5	H	B
14	MOCE-4	Hawaii	Jan–Feb98	20	F	B
15	Goa97	Gulf of Alaska	Oct97	10	F	B
16	Ber95	Bering Sea	Jul95	17	F	B
17	Ber96	Bering Sea	Jul96	16	F	B
18	Lab96	Labrador Sea	Oct–Nov96	68	F	B
19	Lab97	Labrador Sea	May–Jun97	71	F	B
20	Res94	Resolute	Aug94	9	F	B
21	Res95-2	Resolute	Aug95	14	F	B
22	Res96	Resolute	Aug96	11	F	B
23	Res98	Resolute	Aug98	91	F	B
24	CSC	Onslow Bay and S. MAB	May97	12	F	B
25	CSC	S. Mid-Atlantic Bight	Sep97, Nov97, Apr98, Feb99	45	F	B
26	CSC	Gulf of Mexico	Apr99	6	F	B
27	EqPac	0°N,140°W	Mar92, Sep92	36	H	B
28	NABE	46°N,19°W	Apr89	6	H	B
29	CB-MAB	Chesapeake Bay and MAB	Apr96–Oct98 (9 Surveys)	197	H	B
30	AMT-1	E. North Atlantic and W. South Atlantic	Sep–Oct95	23	F	B
31	AMT-2	E. North Atlantic and W. South Atlantic	Apr–May96	19	F	B
32	AMT-3	E. North Atlantic and W. South Atlantic	Sep–Oct96	20	H	B
33	AMT-4	E. North Atlantic and W. South Atlantic	Apr–May97	21	H	B
34	AMT-5	E. North Atlantic and W. South Atlantic	Sep–Oct97	45	H	B
35	AMT-6B	E. North Atlantic and W. South Atlantic	Apr–May98	62	H	B
36	AMT-6	E. North Atlantic and E. South Atlantic	May–Jun98	35	H	B
37	AMT-7	E. North Atlantic and W. South Atlantic	Sep–Oct98	52	H	B
38	AMT-8	E. North Atlantic and W. South Atlantic	May–Jun99	46	H	B
39	HOT	N. Pacific Subtropical Gyre (ALOHA)	Feb98–May99	50	H,F	B
40	WOCE	50°S–13°N,88–91°W	Mar–Apr93	15	F	B
41	WOCE	10°S–30°N,18–37°W	Apr–May94	27	F	B
42	CalCOFI	California Coast	93–97 (16 Surveys)	299	F	B

Table 3. (cont.) The data sources, locations, and acquisition dates of the global data set.

No.	*Data Set*	*Location*	*Dates*	*N*	C_a	R_{rs}
43	CalCOFI	California Coast	97–99 (6 Surveys)	100	F	B
44	RED9503	California Coast (Red Tide)	Mar95	9	F	B
45	AI9901	Subtrop. Atlantic, Indian Ocean	Jan–Mar99	36	F	B
46	JES9906	E. Japan Sea	Jun–Jul99	37	F	B
47	CARIACO	Cariaco Basin	May96–Aug99	35	F	A
48	NEGOM	NE Gulf of Mexico	Jul–Aug98	13	F	A
49	ORINOCO	Orinoco Delta, Paria Gulf, Orinoco Plume	Jun98, Oct98, Feb99, Oct99	48	F	A
50	GOM	Gulf of Maine	Mar95–Apr99 (11 Surveys)	92	F	C
51	Arabian Sea	Arabian Sea	Jul95, Sep95, Oct95	15	F	C
52	FL-Cuba	Florida–Cuba	Mar99	13	F	C
53	BBOP 9293	Sargasso Sea (BATS)	92–93	30	H	B
54	BBOP 9499	Sargasso Sea (BATS)	Jan94–Aug99	83	H	B
55	Plumes & Blooms	Santa Barbara Channel	Aug96–June99	251	F	B
56	NABE	46–59°N,17–20°W	May89	19	H	B
57	CoASTS	N. Adriatic Sea	Sep97–Jan98	35	H	B

Table 4. The wavelengths of the radiometer data.

No.	*Data Set*	*Nominal Center Wavelengths* [nm]																		
1	Roaverrs 97		412	443		490	510				555					655				
2	CARDER		412	443		490	510				555							670		
3	CARDER		412	443		490	510				555							670		
4	CARDER		412	443		490	510				555							670		
5	CARDER		412	443		490	510				555							670		
6	MF0796		412	443		490	510				555							670		
7	TOTO		412	443		490	510				555							670		
8	CoBOP		412	443		490	510				555							670		
9	EcoHAB		412	443		490	510				555							670		
10	Global		412	443		490	510				555					656	665			
11	MBARI EqPac		412	443		490	510				555							670		
12	MOCE-1		412	443		490	510				555									
13	MOCE-2		412	443		490	510				555									
14	MOCE-4		412	443		490	510				555							670		
15	Goa97		412	443		490	510	520	532		555		565		619		665		683	700
16	Ber95		412	443		490	510				555						665		683	
17	Ber96		412	443		490	510	520	532		555		565		619		665		683	700
18	Lab96		412	443		490	510	520	532		555		565		619		665		683	700
19	Lab97		412	443		490	510	520	532		555		565		619		665		683	700
20	Res94		412	443		490	510				555						665		683	
21	Res95-2		412	443		490	510				555						665		683	
22	Res96		412	443		490	510	520	532		555		565		619		665		683	700
23	Res98		412	443		490	510	520	532		555		565		619		665		683	700
24	CSC	380	412	443		490	510				555								683	
25	CSC	380	412	443		490	510				555								683	
26	CSC	380	412	443		490	510				555								683	
27	EqPac		410	441		488		520		550									683	
28	NABE		410	441		488		520		550									683	
29	CB-MAB		412	443	455	490	510		532	550		560		589	625			671	683	700
30	AMT-1		412	443		490	510				555						665			
31	AMT-2		412	443		490	510				555						665			
32	AMT-3		412	443		490	510				555						665			
33	AMT-4		412	443		490	510				555						665			
34	AMT-5		412	443		490	510				555						665			

Table 4. The wavelengths of the instruments.

No.	*Data Set*	*Nominal Center Wavelengths* [nm]																			
35	AMT-6B				412	443			490	510				555					665		
36	AMT-6				412	443			490	510				555					665		
37	AMT-7				412	443			490	510				555					665		
38	AMT-8				412	443			490	510				555					665		
39	HOT				412	443			490	510				555						670	
40	WOCE				410	441			488		520				565				665		
41	WOCE				410	441			488		520				565				665		
42	CalCOFI	340	380	395	412	443	455		490	510		532		555		570			665		
43	CalCOFI				412	443			490	510				555					665		
44	RED9503	340	380	395	412	443	455		490	510		532		555		570			665		
45	AI9901				412	443			490	510				555					665		
46	JES9906				412	443			490	510				555					665		
47	CARIACO				412	443			490	510				555				656			
48	NEGOM				412	443			490	510				555						670	
49	ORINOCO				410	443			490	510				555						670	
50	GOM				412	443			490	510				555					665		
51	Arabian Sea				412	443			490	510				555					665		
52	FL-Cuba				412	443			490	510				555					665		
53	BBOP 9293				410	441			488		520				565				665		
54	BBOP 9499				410	441		465	488	510	520			555	565	589	625		665		683
55	Plumes & Blooms				412	443			490	510				555				656			
56	NABE				412	441			488		521		550								
57	CoASTS				412	443			490	510				555				655			683

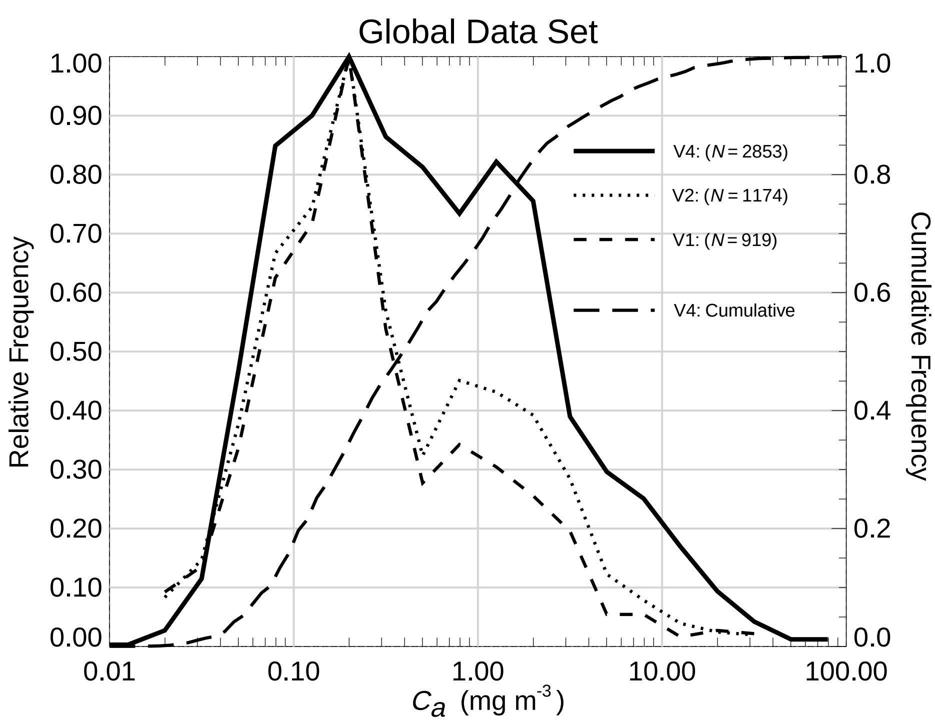

Fig. 1. The relative frequency distribution of C_a concentration in the *in situ* data used to develop versions 4 and earlier versions of the ocean chlorophyll algorithms (V1 is version 1, V2 is version 2, and V4 is version 4). The version 3 data set, an intermediate test set, is not described here). Relative frequency is the observed frequency normalized to the maximum frequency.

Maritorena (2000):

$$R_{\mathrm{rs}}(510) = R_{\mathrm{rs}}(520)\Big[1.0605321 - 0.1721619\,\gamma_a + 0.0295192\,\gamma_a^2 + 0.0150622\,\gamma_a^3 - 0.004133924\,\gamma_a^4\Big] \qquad (3)$$

where $\gamma_a = \log(C_a)$.

The Chesapeake Bay and Mid-Atlantic Bight (CB-MAB) $\tilde{R}_{\mathrm{rs}}(\lambda)$ measurements were corrected for the influence of radiometer self-shading (Gordon and Ding 1992, and Zibordi and Ferrari 1995) using equations provided by G. Zibordi. Corrections for radiometer shading by the Acqua Alta Oceanographic Tower were also applied to the CoASTS $\tilde{R}_{\mathrm{rs}}(\lambda)$ data (Zibordi et al. 1999).

Interpolated estimates of R_{rs} were also generated for non-SeaWiFS wavelengths which were not consistently present in the global data set to develop chlorophyll algorithms similar to OC4 for use by other ocean color sensors. The interpolation–extrapolation method consisted of two steps. A cubic spline interpolation method (using IDL, version 5.3) and four measured adjacent R_{rs} values were used to derive the interpolated R_{rs} estimate ($\hat{R}_{\mathrm{rs}}$). The interpolated values were then regressed against those measured R_{rs} values present in the global data set and the resulting regresssion equation (Table 5) was applied in the second step to remove bias in the interpolated values. This scheme resulted in good agreement between interpolated and measured R_{rs} over a wide range of chlorophyll concentration (Fig. 2).

The characteristics of the R_{rs} data most relevant to bio-optical algorithms are illustrated in Fig. 3. An important feature revealed by these plots is the dispersion of the data (variability orthogonal to the major axis of the data). A pattern common to these plots is the progressive increase in dispersion with increasing chlorophyll concentration and decreasing band ratio. This is most evident in the plots of R^{412}_{555} and R^{443}_{555} versus C_a.

Considering only the degree of scatter evident in these plots, the R^{443}_{555} provide the most precise (lowest dispersion) C_a estimates at concentrations approximately less than $0.4\,\mathrm{mg\,m^{-3}}$, whereas, the R^{510}_{555} and R^{490}_{555} band ratios would provide relatively more precise estimates of C_a in chlorophyll-rich waters. Over the entire data domain, R^{490}_{555} yields the highest correlation with C_a, $R^2 = 0.862$ (Fig. 3), followed by R^{443}_{555}, $R^2 = 0.847$. It must be kept in mind, however, that R^2 is an index of the degree of linear association and a simple linear model is generally not the best model to describe the band ratio C_a relationships over the entire range of the data.

2.3 OC2 AND OC4

The R_{rs} and C_a data (N=2,853) were used to revise the OC2 and OC4 C_a algorithms. Four observations, with $\tilde{C}_a$ greater than $64\,\mathrm{mg\,m^{-3}}$, were widely scattered in plots of band ratios versus C_a and were not used. A test version of the OC4 MBR model revealed 45 observations had $\log(C_a)/\log(\tilde{C}_a)$ values exceeding three standard deviations, so these data were also discarded. The final model coefficients were derived using the remaining 2,804 R_{rs} and $\tilde{C}_a$ combinations. Algorithm refinement involved the determination of model coefficients using iterative minimization routines (IDL, Research System Incorporated) to achieve a slope of 1.000, an intercept of 0.000, minimum RMS error, and maximum R^2 between model and measured $\tilde{C}_a$ concentration. The first version of OC4 (O'Reilly et al. 1998) was formulated as a modified cubic polynomial (i.e., a third order polynomial plus an extra coefficient), however, the current version of OC4 uses a fourth order polynomial (five coefficients), because this yielded better statistical agreement between model and $\tilde{C}_a$ than a MCP formulation. A MCP equation was used to refine OC2 to the same set of values (N=2,804) used to update OC4.

The fourth order polynomial equation for OC4 version 4, hereafter referred to as OC4v4, is:

$$C_a = 10.0^{\left(0.366 - 3.067R_4 + 1.930R_4^2 + 0.649R_4^3 - 1.532R_4^4\right)} \qquad (4)$$

where $R_4 = \log_{10}\left(R^{443}_{555} > R^{490}_{555} > R^{510}_{555}\right)$. The modified cubic polynomial equation for OC2 version 4, hereafter referred to as OC2v4, is:

$$C_a = 10.0^{\left(0.319 - 2.336R_2 + 0.879R_2^2 - 0.135R_2^3\right)} - 0.071 \qquad (5)$$

where $R_2 = \log_{10}\left(R^{490}_{555}\right)$.

The statistical and graphical characteristics of these two algorithms are illustrated in Figs. 4 and 5. The R^2 value between $\tilde{C}_a$ and (model) C_a is slightly higher with OC4 (0.892) than OC2 (0.883). Both models yield a relative frequency distribution that is approximately congruent with the $\tilde{C}_a$ distribution. The OC2 and OC4 models are extrapolated to a C_a value of 0.001, well below the lowest concentration ($0.008\,\mathrm{mg\,m^{-3}}$) present in the *in situ* data (Figs. 4e and 5e). If clear (clearest) water is operationally defined as $C_a = 0.001\,\mathrm{mg\,m^{-3}}$, then the clear water reflectance ratio (R^{443}_{555}) predicted by OC4 is within the theoretical range given in Table 6, whereas the extrapolated clear water R^{490}_{555} reflectance ratio for OC2 is greater than the theoretical clear water estimates.

Since the OC2v4 and OC4v4 algorithms were tuned to the same data set, their C_a estimates should be very highly correlated and internally consistent, with a slope of 1 and an intercept of 0. This is illustrated in Fig. 6. The reduced scatter (orthogonal to the 1:1 line), centered

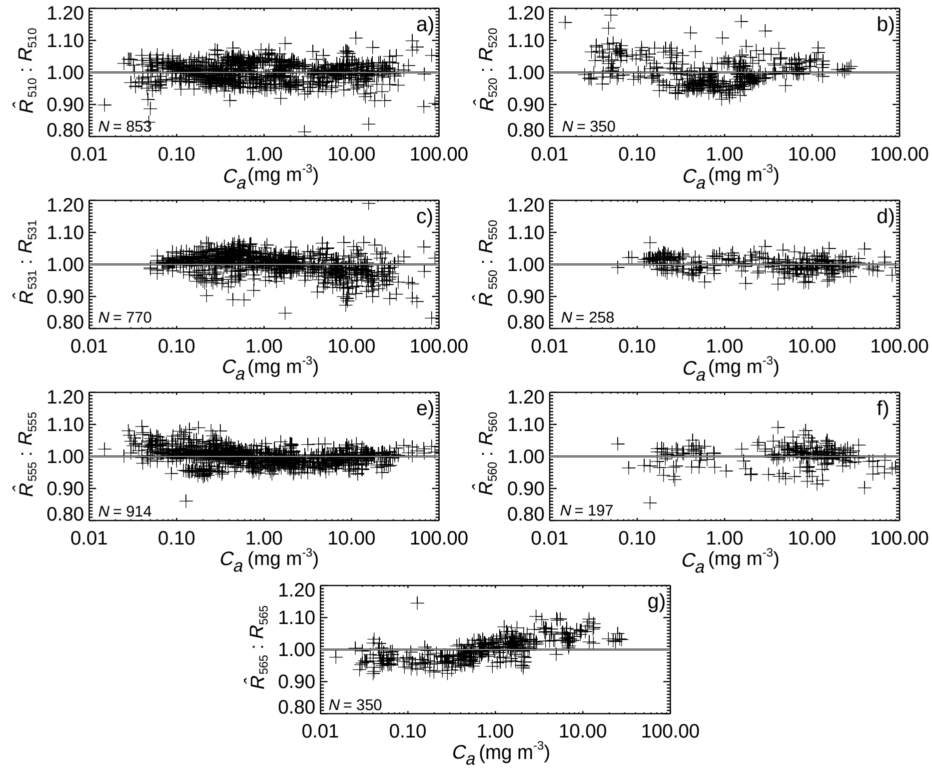

Fig. 2. Ratio of R_{rs} based on interpolated R_{rs} ($\hat{R}$) to measured R_{rs} (R) versus chlorophyll concentration (C_a): a) $\hat{R}_{510}$:R_{510}; b) $\hat{R}_{520}$:R_{520}; c) $\hat{R}_{531}$:R_{531}; d) $\hat{R}_{550}$:R_{550}; e) $\hat{R}_{555}$:R_{555}; f) $\hat{R}_{560}$:R_{560}; and g) $\hat{R}_{565}$:R_{565}.

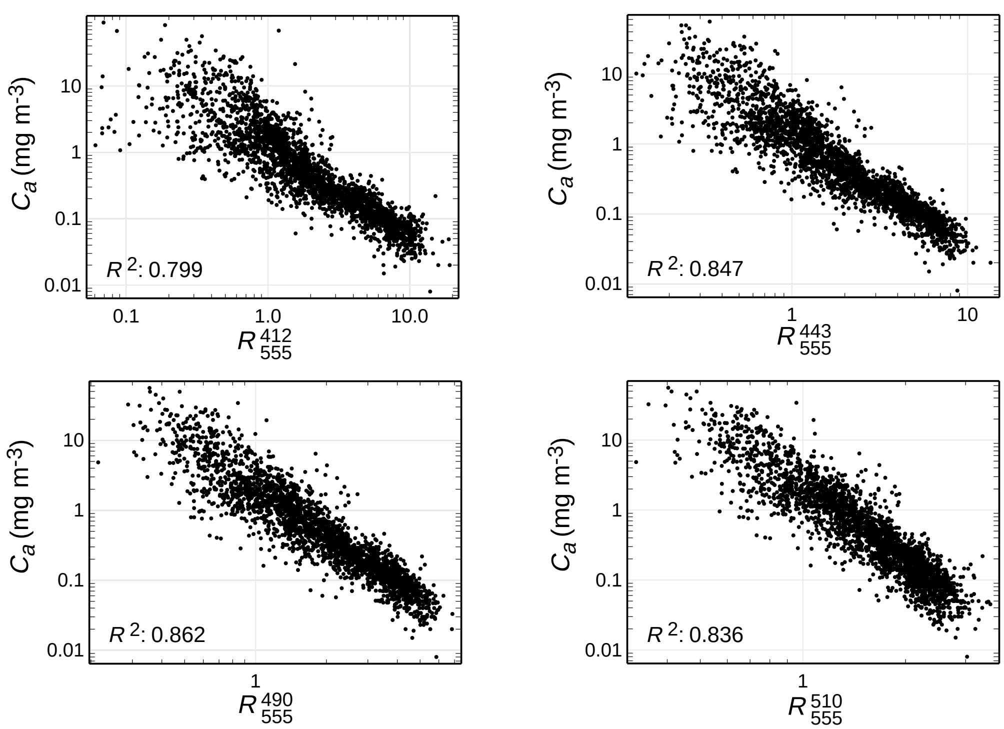

Fig. 3. The relationship between R^{412}_{555}, R^{443}_{555}, R^{490}_{555} and R^{510}_{555} band ratios and chlorophyll concentrations less than $64\,\mathrm{mg\,m^{-3}}$ ($N = 2849$, except for R^{412}_{555} where $N = 2813$).

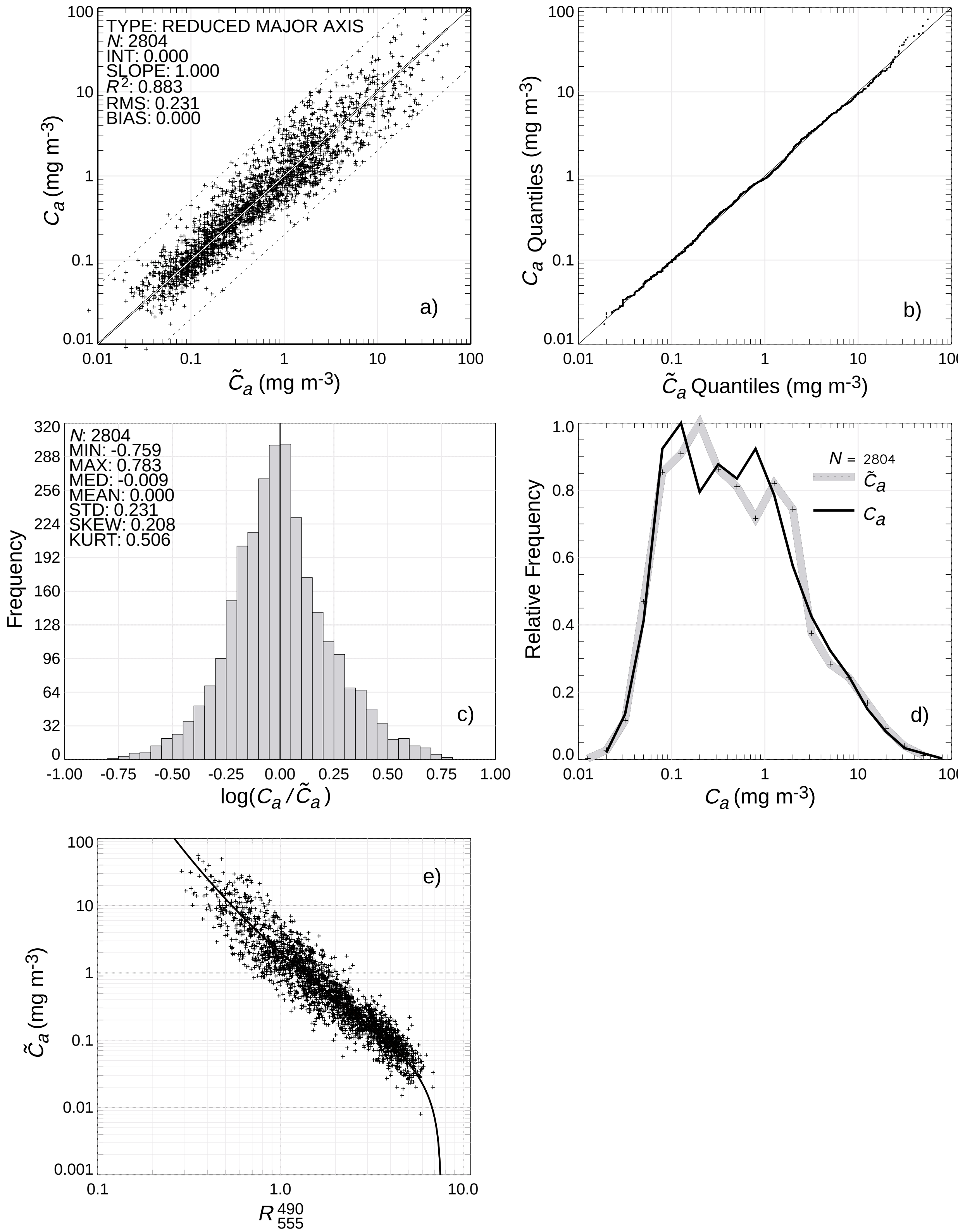

Fig. 4. Comparisons between OC2v4 modeled values (C_a) and *in situ* data ($\tilde{C}_a$): **a)** Scatterplot of C_a versus $\tilde{C}_a$; **b)** Quantile–quantile plot of C_a versus $\tilde{C}_a$; **c)** Frequency distribution of $\log(C_a/\tilde{C}_a)$; **d)** Relative frequency of C_a (thin solid curve) and $\tilde{C}_a$; **e)** R^{490}_{555} versus $\tilde{C}_a$. Also shown is the OC2v4 model (solid curve).

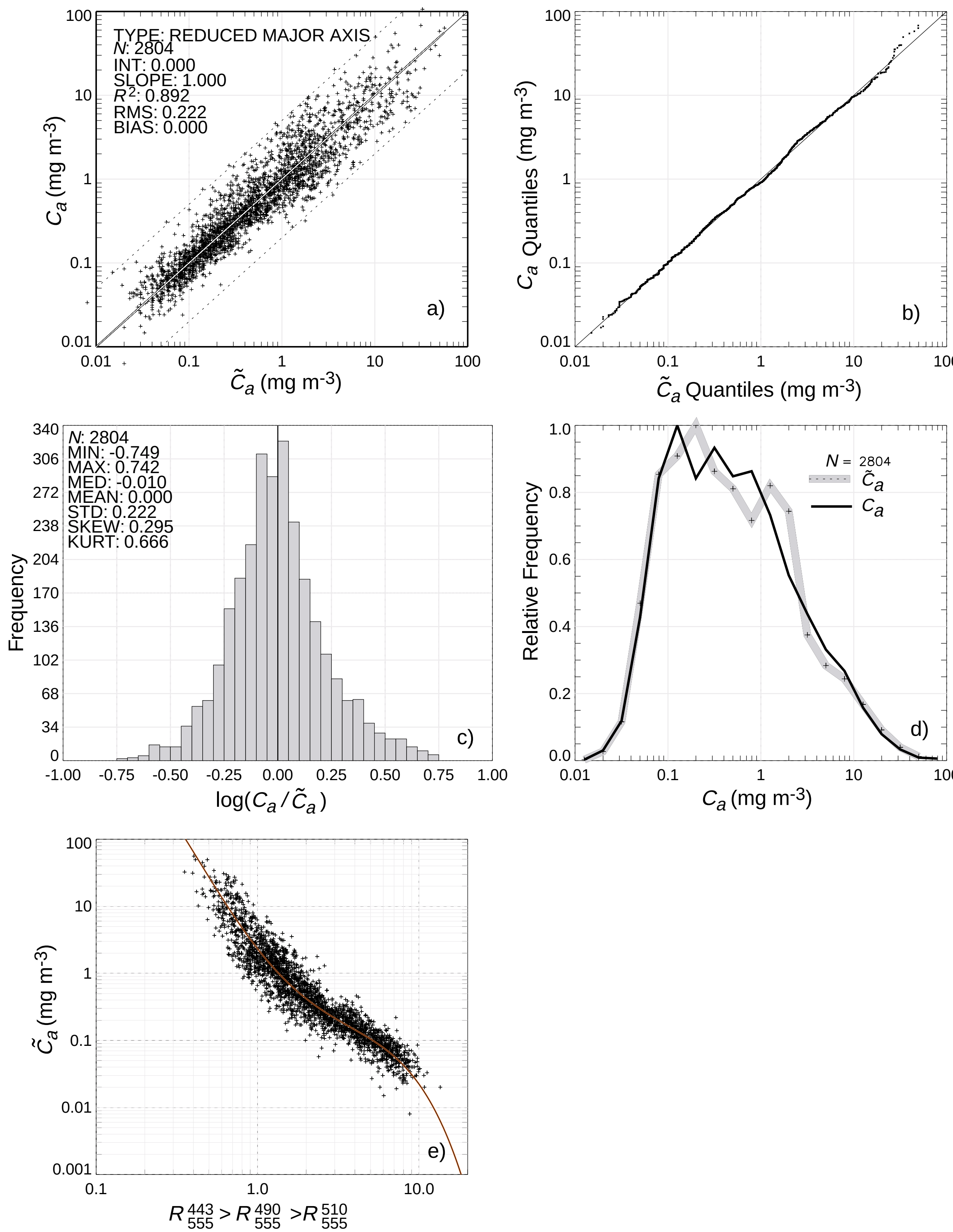

Fig. 5. Comparisons between OC4v4 modeled values (C_a) and *in situ* data ($\tilde{C}_a$): **a)** Scatterplot of C_a versus $\tilde{C}_a$; **b)** Quantile–quantile plot of C_a versus $\tilde{C}_a$; **c)** Frequency distribution of $\log(C_a/\tilde{C}_a)$; **d)** Relative frequency of C_a (thin solid curve) and $\tilde{C}_a$; **e)** R_{555}^{490} versus $\tilde{C}_a$. Also shown is the OC4v4 model (solid curve).

Table 5. Regression statistics (reduced major axis) for the linear relationship between log (measured R_{rs}) and log (interpolated R_{rs}), where m is the slope and b is the intercept.

R_{rs}	N	R^2	m	b
510	853	0.995	0.9948	0.00299
520	350	0.990	1.0328	0.06280
531	770	0.995	0.9614	−0.1005
550	258	0.999	0.9827	−0.0425
555	914	0.998	1.0032	0.01141
560	197	0.998	1.0178	0.02361
565	350	0.989	1.0487	0.11512

Table 6. Comparison between theoretical and extrapolated clear water reflectance ratios using OC2 and OC4 algorithms, where a is the absorption per meter, b_b is the backward scattering coefficient per meter, and f is the function (unspecified). The theoretical reflectance ratios are based on the absorption and backscattering values from Pope and Fry (1977) and Morel (1974).

R_{rs} Band Ratio	$R_{\mathrm{rs}} = f\frac{b_b}{a+b_b}$	$R_{\mathrm{rs}} = f\frac{b_b}{a}$	Algorithm
443:555	16.53	21.78	18.21 (OC4)
490:555	6.13	6.66	7.502 (OC2)

Table 7. Maximum band ratio algorithms for the SeaWiFS, CZCS, OCTS, MODIS and MERIS sensors.

Sensor	*Name*	*Equation*
SeaWiFS	OC4	$C_a = 10.0^{(0.366 - 3.067R_4 + 1.930R_4^2 + 0.649R_4^3 - 1.532R_4^4)}$ where $R_4 = \log_{10}\left(R_{555}^{443} > R_{555}^{490} > R_{555}^{510}\right)$
MODIS	OC4M	$C_a = 10.0^{(0.366 - 3.067R_{4\mathrm{M}} + 1.930R_{4\mathrm{M}}^2 + 0.649R_{4\mathrm{M}}^3 - 1.532R_{4\mathrm{M}}^4)}$ where $R_{4\mathrm{M}} = \log_{10}\left(R_{550}^{443} > R_{550}^{490} > R_{550}^{530}\right)$
OCTS	OC3O	$C_a = 10.0^{(0.366 - 3.067R_{3\mathrm{O}} + 1.930R_{3\mathrm{O}}^2 + 0.649R_{3\mathrm{O}}^3 - 1.532R_{3\mathrm{O}}^4)}$ where $R_{3\mathrm{O}} = \log_{10}\left(R_{565}^{443} > R_{565}^{490} > R_{565}^{520}\right)$
CZCS	OC3C	$C_a = 10.0^{(0.366 - 3.067R_{3\mathrm{C}} + 1.930R_{3\mathrm{C}}^2 + 0.649R_{3\mathrm{C}}^3 - 1.532R_{3\mathrm{C}}^4)}$ where $R_{3\mathrm{C}} = \log_{10}\left(R_{550}^{443} > R_{550}^{520}\right)$
MERIS	OC4E	$C_a = 10.0^{(0.366 - 3.067R_{4\mathrm{E}} + 1.930R_{4\mathrm{E}}^2 + 0.649R_{4\mathrm{E}}^3 - 1.532R_{4\mathrm{E}}^4)}$ where $R_{4\mathrm{E}} = \log_{10}\left(R_{560}^{443} > R_{560}^{490} > R_{560}^{510}\right)$

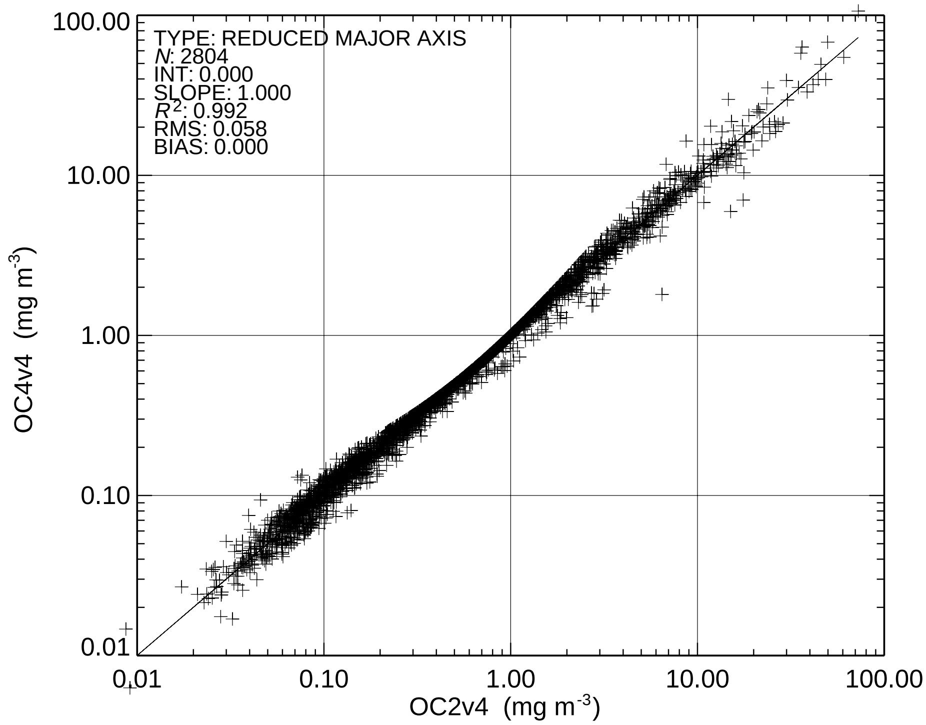

Fig. 6. Comparisons of C_a from OC2 and OC4 when using $\tilde{R}_{\mathrm{rs}}$ from the *in situ* data set.

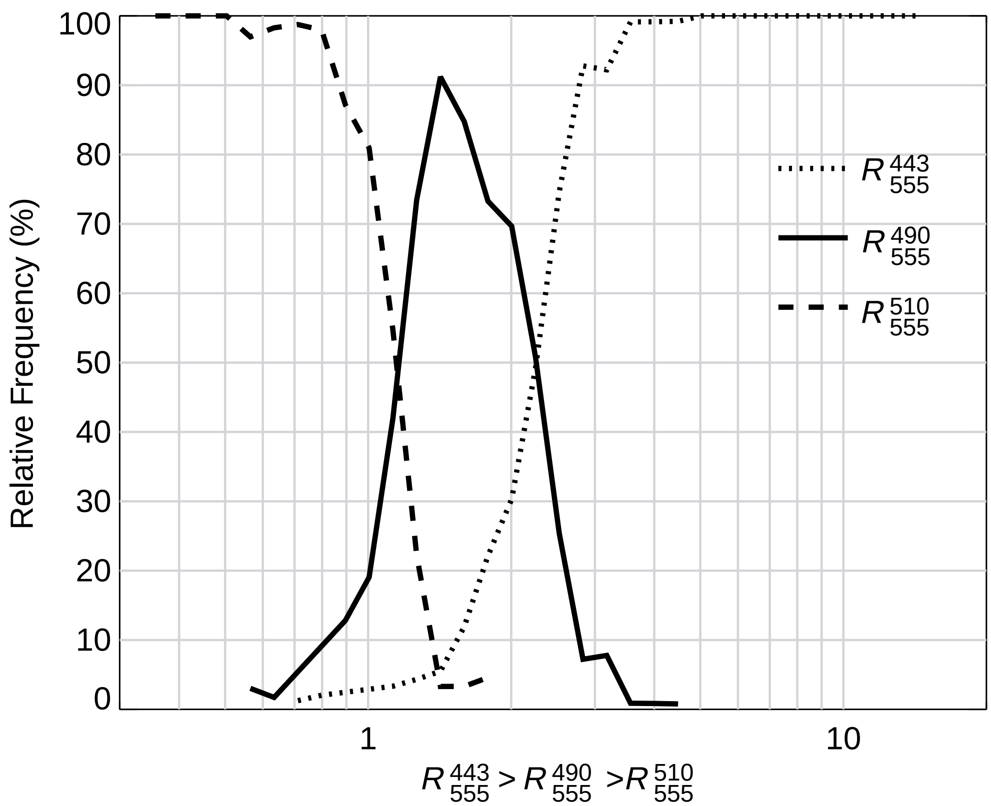

Fig. 7. The relative frequency of band ratios used in the OC4 model versus the maximum band ratio.

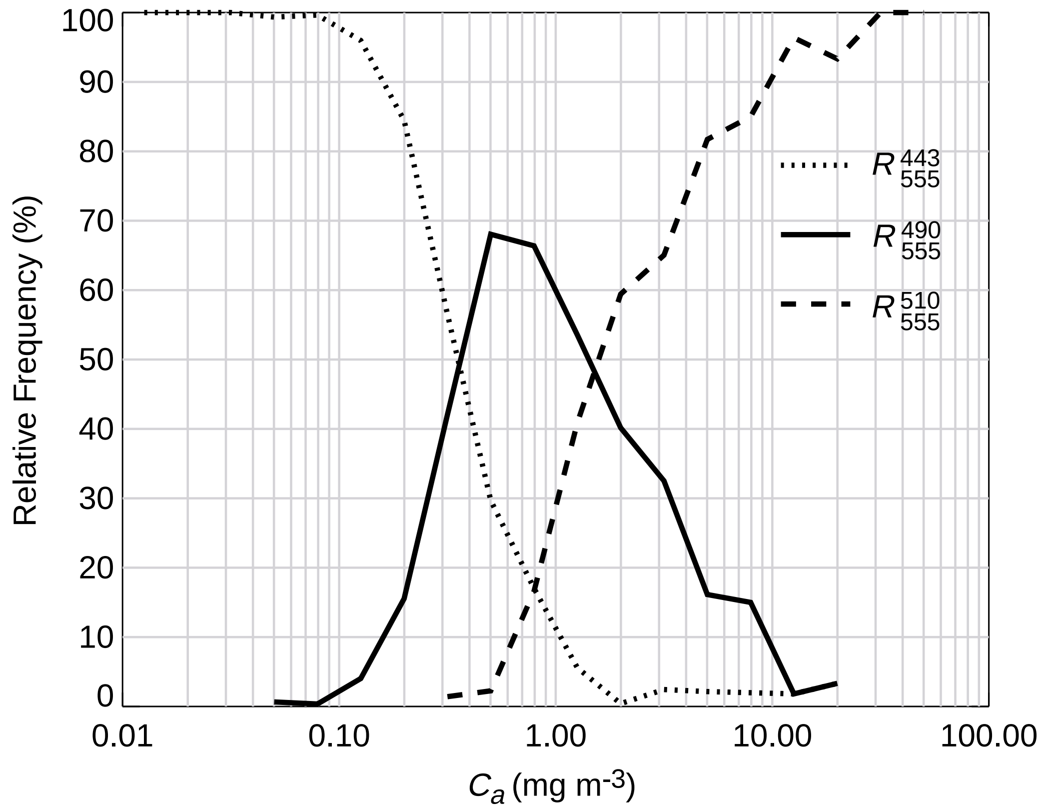

Fig. 8. The relative frequency of band ratios used in the OC4 model versus chlorophyll concentration.

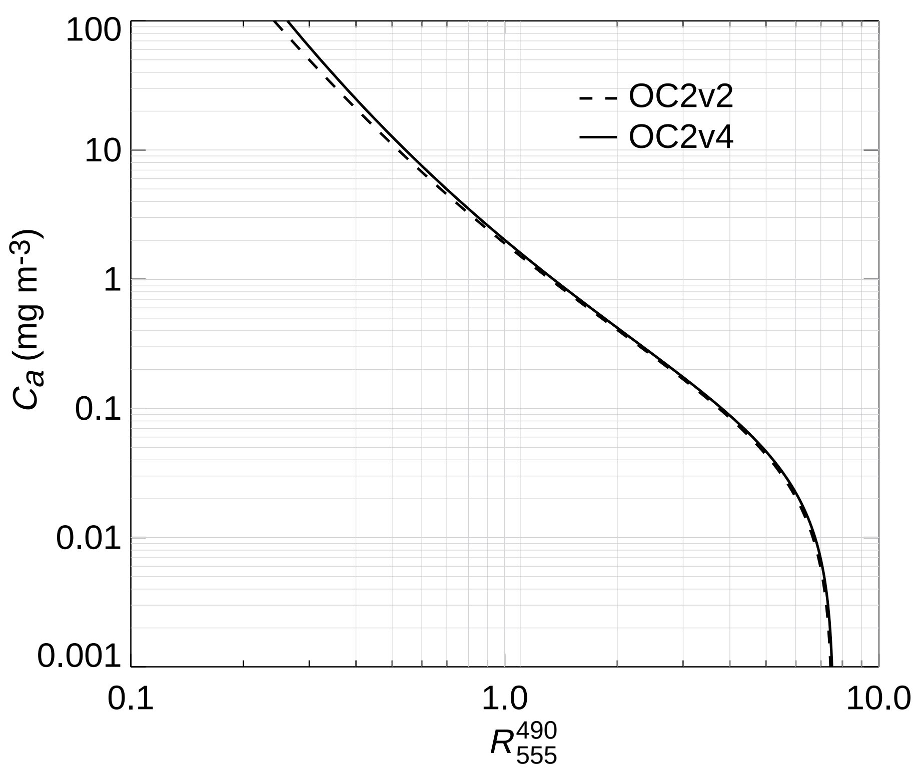

Fig. 9. Comparison of C_a estimates from OC2v4 with OC2v2.

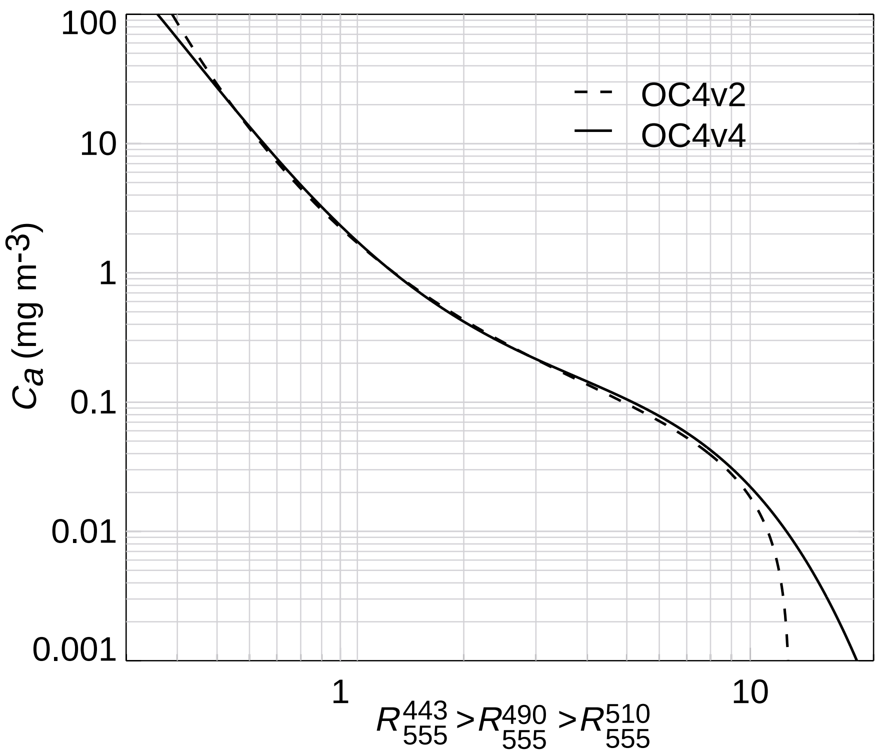

Fig. 10. Comparison of C_a estimates from OC4v4 and OC4v2 models.

at about $1\,\mathrm{mg\,m^{-3}}$, indicates the region where both algorithms use the 490 nm band.

Additional noteworthy characteristics of OC4 are illustrated in Figs. 7 and 8. The R^{443}_{555} ratio dominates (50%) at MBRs above approximately 2.2, R^{490}_{555} between 2.2 and 1.1, and R^{510}_{555} at MBRs below 1.1 (Fig. 7). With respect to chlorophyll concentration, the R^{443}_{555} ratio dominates (50%) when C_a is below approximately $0.33\,\mathrm{mg\,m^{-3}}$, R^{490}_{555} for C_a between 0.33–$1.4\,\mathrm{mg\,m^{-3}}$, and R^{510}_{555} when C_a exceeds approximately $1.4\,\mathrm{mg\,m^{-3}}$ (Fig. 8).

Relative to OC2v2, OC2v4 predicts slightly higher C_a above concentrations of $3\,\mathrm{mg\,m^{-3}}$ (Fig. 9), while OC4v4 generates slightly lower C_a estimates at very high concentrations (Fig. 10). At C_a below $0.03\,\mathrm{mg\,m^{-3}}$, OC2v4 estimates are very similar to OC2v2, while OC4v4 estimates are slightly higher than those from OC4v2, particularly so when C_a is below $0.01\,\mathrm{mg\,m^{-3}}$. (Version 3 equations were preliminary and provided to the SeaWiFS Project for testing and evaluation and are not described here.)

There is considerable interest and benefit from comparing and merging data from various ocean color sensors (Gregg and Woodward 1998). This is one of the major objectives of SIMBIOS (McClain and Fargion 1999). In the particular case of ocean color data merging, one methodological issue to be resolved is how data from satellite sensors having different center band wavelengths can be merged to generate seamless maps of chlorophyll *a* distribution. Among several possible approaches, one is to develop internally consistent, sensor-specific variations of empirical chlorophyll *a* algorithms tuned to the same data set. This implies a comprehensive suite of *in situ* measurements at wavelengths matching the various satellite spectrometers or perhaps hyperspectral *in situ* data. To facilitate comparisons with SeaWiFS chlorophyll *a*, MBR algorithms for several ocean color sensors are presented in Table 7. These algorithms must be considered as an approximation, because the *in situ* data set is biased to SeaWiFS channels and a number of radiometric adjustments were made to the $R_{\mathrm{rs}}(\lambda)$ data to compensate for wavelength differences among the sensors (Table 4).

2.4 CONCLUSIONS

A large data set of $\tilde{R}_{\mathrm{rs}}$ and $\tilde{C}_a$ measurements was compiled and used to update the OC2 and OC4 bio-optical chlorophyll *a* algorithms. The present data set is substantially larger (N=2,853) than that used to develop the version 2 algorithms (N=1,174), includes samples from a greater variety of bio-optical provinces, and better represents oligotrophic and eutrophic waters.

Over the 4-decade range in chlorophyll *a* concentration encompassed in the data set (0.008–$90\,\mathrm{mg\,m^{-3}}$), the R^{490}_{555} band ratio is the best overall single band ratio index of chlorophyll *a* concentration. In oligotrophic waters, however, the R^{443}_{555} ratio yields the best correlation with C_a and lowest RMS error, while in waters with chlorophyll concentrations exceeding approximately $3\,\mathrm{mg\,m^{-3}}$, the R^{510}_{555} ratio is the best-correlated index. OC4 takes advantage of this band-related shift in precision, and the well-known shift of the maximum of $R_{\mathrm{rs}}(\lambda)$ spectra towards higher wavelengths with increasing C_a. Dispersion between the OC2 model and $\tilde{C}_a$ tended to increase with increasing chlorophyll concentrations above $1\,\mathrm{mg\,m^{-3}}$, whereas dispersion using OC4 remained relatively low and uniform throughout the range of *in situ* data. Consequently, OC4 yields a slightly higher R^2 and lower RMS error than OC2.

Statistical comparisons of algorithm performance with respect to *in situ* data, however, provide only partial information about their performance when applied to satellite-derived water-leaving radiances. Operationally, OC4 would be expected to generate more accurate C_a estimates than OC2 for several reasons. In oligotrophic water, OC4 would be expected to provide more accurate C_a estimates than OC2, because the signal-to-noise ratio (SNR) is greater in the 443 nm band than the 490 nm band. In eutrophic waters, strong absorption in the blue region of the spectrum results in lower SNR for water-leaving radiances retrieved in the 412 nm and 443 nm bands relative to the 490 nm and 510 nm bands. Furthermore, the influence of the atmospheric correction scheme on the accuracy of derived water-leaving radiances used in band-ratio algorithms must be considered. The SeaWiFS atmospheric correction algorithm (Gordon and Wang 1994 and Wang 2000) uses the near infrared bands (765 and 865 nm) to characterize aerosol optical properties and estimates aerosol contribution to total radiance in the visible spectrum by extrapolation. The 510 nm band, being closer to the near infrared bands, is less prone to extrapolation errors than the 490 nm and 443 nm bands. In chlorophyll-rich water, therefore, OC4 would be expected to provide more accurate estimates of C_a than OC2.

The present version of the $\tilde{R}_{\mathrm{rs}}(\lambda)$ and $\tilde{C}_a$ data set represents a significant improvement in size, quality, and bio-optical diversity when compared with earlier versions, but it still lacks observations from the clearest oceanic waters. These observations are required to resolve the asymptotic relationship expected between $R_{\mathrm{rs}}(\lambda)$ and C_a as chlorophyll a concentration diminishes below $0.01\,\mathrm{mg\,m^{-3}}$, and reflectance band ratios approach the theoretical values for pure sea water. They are also needed to determine if the OC2 and OC4 extrapolations beyond the lowest $\tilde{C}_a$ are accurate. Given the spatially and temporally comprehensive time series achieved by the SeaWiFS mission, these clearest water regions and optimal sampling times may now be easily identified and targeted for special shipboard surveys. Although clearest waters encompass a relatively small fraction of the global ocean, these and highly eutrophic areas represent bio-optical and ecological extremes and changes in their magnitude or areal distribution may provide very sensitive indicators of global change.

Acknowledgments

We would like to acknowledge the following individuals for their significant contribution of *in situ* data and ideas: J. Marra, C Davis, D. Clark, G. Zibordi, C. Trees, R. Bidigare, D. Karl, J. Patch, R. Varela, J. Akl, C. Hu, A. Subramaniam, N. Nelson, T. Michaels, R. Smith, and A. Morel.