Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere

D. A. Siegel,[1] S. Maritorena,[1] N. B. Nelson,[1] M. J. Behrenfeld,[2] and C. R. McClain[3]

Received 4 August 2005; revised 19 September 2005; accepted 21 September 2005; published 27 October 2005.

[1] Satellite ocean color data enable the global assessment of the ocean biosphere through determinations of chlorophyll concentrations. However, ocean color is not a function of chlorophyll alone. We assess differences between two ocean color models with nearly identical validation statistics. The resulting chlorophyll retrievals show systematic differences which are consistent with each model's ability to account for the absorption of light by colored dissolved organic materials. These differences are often large and approach 100% poleward of 40° latitude. We conclude that the discrepancies are due to fundamental differences in model assumptions and their empirical tuning using geographically limited, in situ data. This source of uncertainty is important as the choice of ocean color model alters modeled rates of global net primary production by more than 30%. The ultimate resolution of this issue requires continued improvements in remote sensing algorithms and validation data as well as satellite technology. **Citation:** Siegel, D. A., S. Maritorena, N. B. Nelson, M. J. Behrenfeld, and C. R. McClain (2005), Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere, *Geophys. Res. Lett.*, *32*, L20605, doi:10.1029/2005GL024310.

[2] The global assessment of phytoplankton biomass and its variations in time and space is essential for the long-term evaluation of ocean ecosystem health and for understanding changes in the ocean carbon cycle [*Field et al.*, 1998; *Gregg and Conkright*, 2002; *Fasham*, 2003]. The sheer size of the ocean and the costs associated with its in situ sampling have led to the deployment of satellite ocean color missions [*IOCCG*, 1999; *McClain et al.*, 2004]. These global determinations of the upper ocean chlorophyll distribution have produced the first consistent views of the space/time dynamics of the ocean biosphere [*Yoder et al.*, 1993; *Longhurst*, 1995; *Behrenfeld et al.*, 2001]. However, satellite ocean color data are produced through a complex procedure which accounts for atmospheric, surface and in-water effects to produce useful products like the chlorophyll *a* concentration [*Gordon and Morel*, 1983; *McClain et al.*, 2004]. Some of the models used have roots in first principles while others are empirical and are constructed by statistically modeling field observations. A critical part of this procedure is the bio-optical model which relates a measure of ocean color, the water-leaving radiance spectrum, to an in-water constituent, such as the chlorophyll concentration [*Gordon and Morel*, 1983; *O'Reilly et al.*, 1998]. These models are developed and validated using limited in situ data which do not span the full range of oceanic conditions [*Claustre and Maritorena*, 2003]. Hence, this data limitation creates a potential for significant biases in remote sensing products with important implications.

[3] A five year time series of monthly satellite ocean color observations from the Sea-viewing Wide-Field of view Sensor (SeaWiFS) [*McClain et al.*, 2004] is used to determine surface chlorophyll concentrations (Chl) using the operational empirical bio-optical algorithm (OC4v4) [*O'Reilly et al.*, 1998, 2000] and a semi-analytical algorithm (GSM) [*Maritorena et al.*, 2002; *Maritorena and Siegel*, 2005]. Both algorithms have been developed using the best available data set of biological (chlorophyll concentrations) and optical (water-leaving radiance spectra) properties. The OC4v4 algorithm is a polynomial relationship of water-leaving radiance ratios numerically fit to global chlorophyll observations [*O'Reilly et al.*, 1998]. It assumes that the major optically-active components in the surface ocean covary with Chl in a consistent manner globally. In contrast, the GSM algorithm considers that Chl, colored dissolved and detrital organic materials (CDM) and particulate abundances each independently affect ocean color and these properties are retrieved simultaneously from a water-leaving radiance spectrum [*Maritorena et al.*, 2002; *Siegel et al.*, 2002, 2005]. Values of the parameters used in the GSM model are derived using a very similar data set to the one used for developing the OC4v4 algorithm, but also includes CDM and the optical backscattering due to particulates [*Maritorena et al.*, 2002; *Maritorena and Siegel*, 2005]. Both algorithms are good Chl predictors as demonstrated using a match-up data set of SeaWiFS imagery and coincident in situ observations (Table 1). When water depths are greater than 1000 m (chosen to reflect open ocean conditions), the performance of the two bio-optical algorithms is indistinguishable (Table 1).

[4] However, comparison of the two global Chl climatologies shows large qualitative and quantitative differences (Figure 1). Normalized percentage differences (ΔChl) exceed 50% over large expanses of the ocean where retrievals found using the empirical algorithm (OC4v4) are greater than the semi-analytical algorithm (GSM). Large differences are seen poleward of 40° latitude, particularly in the northern hemisphere where they approach 100%. On the other hand, GSM algorithm Chl values are greater than the OC4v4 Chl retrievals by as much as 50% in the clear waters of the subtropical gyres. These differences are of the same size as errors in Chl retrievals reported from the

[1]Institute for Computational Earth System Science, University of California, Santa Barbara, California, USA.
[2]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA.
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Copyright 2005 by the American Geophysical Union.
0094-8276/05/2005GL024310

Table 1. Validation Statistics for Chlorophyll Concentration Retrievals[a]

	OC4v4 vs. in Situ	GSM vs. in Situ	OC4v4 vs. in Situ (Z > 1000 m)	GSM vs. in Situ (Z > 1000 m)
N	1378	979	344	324
R^2	0.757	0.689	0.706	0.823
Slope	0.947	0.876	0.951	0.815
Intercept	−0.015	−0.244	−0.161	−0.156
RMS	0.290	0.381	0.175	0.259
BIAS	−0.011	−0.216	−0.047	−0.148

[a]This is an update of the validation table by *Siegel et al.* [2005] using a more extensive data set. No interpretative differences were found in comparison of the two statistical summaries of validation performance. Field-satellite data match-ups were constructed using a 9 pixel median value from available SeaWiFS imagery with a time difference of 3 hours or less (see seabass.gsfc.nasa.gov/matchup_results.html). Values were not considered in the statistical comparison if the satellite-sensed water-leaving radiance at 555 nm > 1.3 μW cm^{-2} nm^{-1} sr^{-1} to avoid extremely turbid coastal waters [*Otero and Siegel*, 2004]. Further, GSM retrievals were not considered if the satellite-sensed water-leaving radiance at 412 nm was <0.17 μW cm^{-2} nm^{-1} sr^{-1} to eliminate obvious problems with the atmospheric correction procedures [*Maritorena and Siegel*, 2005]. Water depth was determined by comparing each in-situ measurement location with the ETOPO2 digital bathymetry map. Statistical quantities are calculated on log-transformed data following the procedures by *O'Reilly et al.* [1998].

previous generation of satellite ocean color observations from the Southern Ocean [*Mitchell and Holm-Hansen*, 1991; *Sullivan et al.*, 1993] but extend over larger regions of the oceans than just the Southern Ocean (Figure 1).

[5] The spatial patterns of ΔChl and the colored detrital material (CDM) distribution suggest a central role for CDM in creating the observed differences (Figures 1c and 1d). Regions with high average CDM retrievals correspond to regions where the OC4v4 algorithm retrieves higher Chl values than does the GSM algorithm. This can also be seen in the strong correspondence observed between ΔChl and CDM for the entire 5 year data set (Figure 2). The cause for the large differences in global Chl climatologies appears to lie in differences in the underlying assumptions used in the two models. Here, the semi-analytical model (GSM) is able to account for absorption of light by CDM independently while the empirical model (OC4v4) assumes that CDM covaries in a consistent way with Chl.

[6] For nearly all of the ocean, the CDM signal is driven by changes in the colored dissolved organic material content (CDOM) [*Siegel et al.*, 2002]. High quality, open ocean CDOM observations are even rarer than Chl observations [*Nelson and Siegel*, 2002]. That said, the GSM algorithm performs well for predicting CDOM [*Siegel et al.*, 2005] as demonstrated using match-up data of satellite and field observations (R^2 = 0.61; N = 112) and from meridional transect observations from the North Atlantic Ocean (R^2 = 0.65; N = 111). Thus, the correspondence between satellite determinations of CDM and in situ CDOM observations and between ΔChl and CDM signals all suggest that the varying CDOM contribution is not properly accounted for in the OC4v4 algorithm [*Siegel et al.*, 2005].

[7] Other processes could conceivably create the observed discrepancies though it is hard to make a convincing argument. For example, land-sea interactions are not driving the observed differences as the expected patterns from riverine inputs are largely inconsistent with the observed ΔChl distribution (Figure 1c) [see *Siegel et al.*, 2002]. Further, it is also unlikely that the observed differences are an artifact of the procedures used to correct the satellite

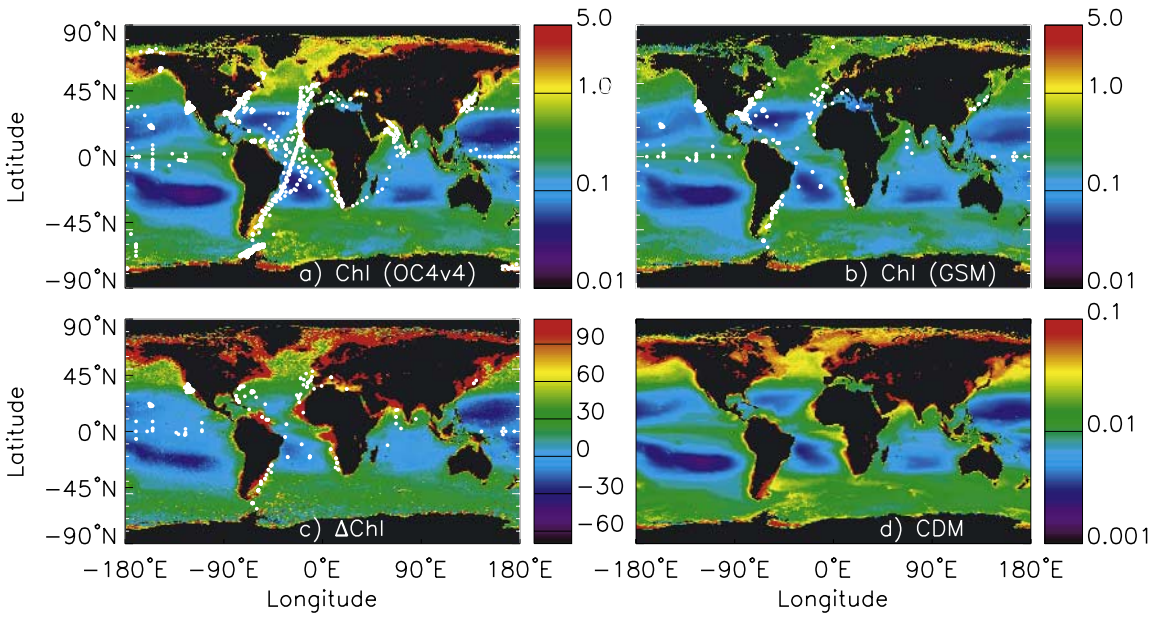

Figure 1. Global climatology of (a) Chl(OC4v4) (mg m^{-3}), (b) Chl(GSM) (mg m^{-3}), (c) the normalized percentage difference between them (ΔChl = 100*(Chl(OC4v4) − Chl(GSM))/Chl(GSM); units %) and (d) the CDM distribution (m^{-1}). These climatologies are constructed from simple averages from the available SeaWiFS data set (October 1997 to December 2004) averaged to 1/3° spatial resolution. Also shown in Figure 1a are the locations of in situ observations from the NOMAD data set (white dots) [*Werdell and Bailey*, 2005] (see Figure 3a), in Figure 1b all of the locations of the in situ data set used to validate the ocean color products in Table 1 and in Figure 1c those validation data locations in Table 1 where water depths are greater than 1000 m. Note that wide expanses of the open ocean contribute little to bio-optical databases.

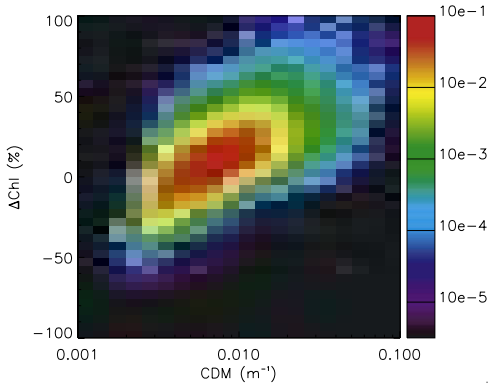

Figure 2. Two-dimensional probability density distributions of the co-occurrence ΔChl and CDM from the SeaWiFS 5 year, global climatology. An increase in the CDM retrieval leads to a positive bias or an increase in the retrieved value of ΔChl between the two bio-optical models. Correlation analysis using all the data used to construct this figure (N = 1,876,368) results in a significant positive slope of 49.2 %/m^{-1} with a regression coefficient (r value) of 0.384. Values are probability of occurrence scaled so that the integral over all parameter space is equal to one. The color scale is log10 transformed.

signals for the atmospheric path as there is no correspondence in spatial patterns between ΔChl and retrieved aerosol property indices in either space or time (data not shown). Changes in phytoplankton community structure or photoadaptation, which alter phytoplankton light absorption per unit chlorophyll in response to light and other environmental stresses [*Bricaud et al.*, 1998; *Cota et al.*, 2004], are also not likely to create the observed differences. The effects of photoadaptation and community structure shifts on light absorption per unit chlorophyll are typically modeled using non-linear power-law relationships with Chl [*Bricaud et al.*, 1998; *Carder et al.*, 2004]. Hence, the global relationship between phytoplankton absorption and Chl will be accounted for in the empirical fittings of the two bio-optical models. Residual differences may still occur, although the observed, large-scale patterns of ΔChl (Figure 1c) and its relationship to the derived CDM distribution (Figure 1d) seem to preclude this possibility.

[8] We conclude that the observed differences in the two Chl retrievals are built into the models themselves. A comparison of clear water, in situ observations (in situ Chl < 0.25 mg m^{-3} defined by [*Gordon and Clark*, 1981]) shows differences between the two models (ΔChl) that increase as a function of CDM (Figure 3a; data from *Werdell and Bailey* [2005]). This trend is also seen as a function of in situ spectrophotometric observations of CDM (Figure 3b; data from *Siegel et al.* [2005]). Type II regression statistics show weak yet significant positive relationships between the ΔChl and either the algorithm produced CDM value or the in situ observed CDOM. The correspondence between these figures produced with in situ observations and the same figure constructed with satellite data (Figure 2) is striking supporting our argument that the observed differences is built into the algorithms. In turn, these differences are created by the limitations of the data sets used to parameterize the two bio-optical algorithms. Relatively few of the in situ data locations used to develop the two algorithms have come from locations where ΔChl values are large (see the locations of the NOMAD data in the upper left panel of Figure 1).

[9] Empirical modeling requires that model performance is optimized through comparison to a development data set. Performance can, however, be significantly degraded outside of the range of applicability of the development data set [*Davis*, 1977]. It is for this reason that an evolution toward mechanistic relationships is preferable. Here, we apply a semi-analytical, bio-optical model (GSM) which accounts for the independence among open ocean optical properties [*Siegel et al.*, 2002]. Comparison of the two bio-optical models indicates that there may be a serious bias in our interpretations of satellite ocean color data. This issue revolves around the contribution that colored dissolved organic materials make to ocean color variations [*Siegel et al.*, 2005]. Empirical algorithms, like OC4v4, assume that the CDM to Chl relationship is a fixed function of Chl for all ocean waters. The advantage of a semi-analytical approach is that it can accommodate the independent and highly variable contributions of CDM and chlorophyll to ocean optical properties.

[10] Implications of our results go well beyond the simple quantification of light absorption by two components of the upper ocean. Perhaps most importantly, these discrepancies between models seriously impact assessments of ocean net primary production (NPP) and quantifying global ocean carbon cycling. For example, the Vertically Generalized Productivity Model [*Behrenfeld and Falkowski*, 1997; *Behrenfeld et al.*, 2001] applied using OC4v4 Chl yields a global net primary production of 53.5 Gt C/y, whereas if the GSM Chl field is used it gives a value of 37.1 Gt C/y. This difference of 16.4 Gt C/y is particularly worrisome when compared to the magnitude of natural interannual variations in NPP, such as the 5 Gt C/y change associated with the 1997–1999 El Nino to La Nina transition [*Behrenfeld et al.*,

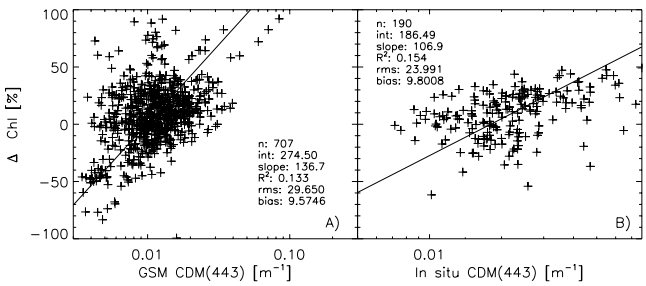

Figure 3. Scatterplot of ΔChl vs. CDM derived from in situ databases for open ocean conditions where (a) in situ radiometric determinations from the NOMAD data set [*Werdell and Bailey*, 2005] are used to estimate CDM using the GSM model [*Maritorena et al.*, 2002] and (b) spectrophotometric observations [*Siegel et al.*, 2005] are used to determine CDM. In situ chlorophyll concentrations of less than 0.25 mg m^{-3} are used to discriminate clear water conditions in the data sets [*Gordon and Clark*, 1981]. In situ CDM is defined here as the sum of detrital particulate and CDOM absorption at 440 nm taken from spectrophotometric data. A type II regression is performed in each plot which demonstrates that the value of ΔChl increases with increasing CDM value for both cases.

2001]. The detection of longer-term temporal change in operational chlorophyll records [*Gregg and Conkright*, 2002] may also be easily confounded by details of the bio-optical models identified here. For example if CDM varies independently from Chl, then changes in global biospheric processes may be incorrectly attributed [*Siegel et al.*, 2005]. Finally, a path for retrieving information on growth rates from space has recently been developed that is based on phytoplankton chlorophyll-to-carbon ratios derived from GSM absorption and backscattering products [*Behrenfeld et al.*, 2005]. Clearly, any uncertainty in the attribution of absorption to CDM versus chlorophyll will impact interpretations and calculations made from satellite ocean color imagery.

[11] The present work points out how differences in algorithms with near identical validation characteristics (Table 1) can produce important differences when applied to the global ocean. This will have an important bearing on the remote assessment of ecosystem functioning and carbon cycling for the world's oceans using satellite ocean color data products. The resolution of this issue will require continued improvements in remote sensing algorithms, the in situ data sets used in validating these models and the ocean viewing instrumentation deployed in space. This work clearly suggests that remote sensing algorithms must evolve from empirical toward mechanistic approaches so that confounding influences of independent ocean optical properties can be diagnosed successfully. This path must also emphasize the importance of in situ sampling covering the entire parameter range that the ocean provides, a limitation that currently reflects the fact that much of the ocean remains largely unexplored [*Claustre and Maritorena*, 2003]. These improvements in turn must be coupled to the development of new satellite-based technologies capable of accurately separating signals from the dominant in-water constituents. It is along this path where we will reduce the uncertainty in our remote assessments of the ocean biosphere due to competing ocean optical constituents.

[12] **Acknowledgments.** The authors would like to acknowledge the support of the NASA Ocean Biology and Biogeochemistry program and the National Science Foundation. Manuela Lorenzi-Kayser and Robert O'Malley performed the satellite data analyses and David Court, Sean Bailey and Jeremy Werdell prepared the satellite match-up and in situ data bases. Comments from the anonymous reviewers as well as Scott Doney and Tiho Kostadinov were useful in preparing the final manuscript. The authors would also like to thank the many scientists who have shared their in situ data in public databases, such as SeaBASS, which has made this work possible.

References

Behrenfeld, M. J., and P. G. Falkowski (1997), Photosynthetic rates derived from satellite-based chlorophyll concentrations, *Limnol. Oceanogr.*, *42*, 1–20.

Behrenfeld, M. J., et al. (2001), Biospheric primary production during an ENSO transition, *Science*, *291*, 2594–2597.

Behrenfeld, M. J., E. Boss, D. A. Siegel, and D. M. Shea (2005), Carbon-based ocean productivity and phytoplankton physiology from space, *Global Biogeochem. Cycles*, *19*, GB1006, doi:10.1029/2004GB002299.

Bricaud, A., A. Morel, M. Babin, K. Allali, and H. Claustre (1998), Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models, *J. Geophys. Res.*, *103*, 31,033–31,044.

Carder, K. L., F. R. Chen, J. P. Cannizzaro, J. W. Campbell, and B. G. Mitchell (2004), Performance of the MODIS semi-analytical ocean color algorithm for chlorophyll-*a*, *Adv. Space Res.*, *33*, 1152–1159.

Claustre, H., and S. Maritorena (2003), The many shades of ocean blue, *Science*, *302*, 1514–1515.

Cota, G. F., J. Wang, and J. C. Comiso (2004), Transformation of global satellite chlorophyll retrievals with a regionally tuned algorithm, *Remote Sens. Environ.*, *90*, 373–377.

Davis, R. E. (1977), Techniques for statistical analysis and prediction of geophysical fluid systems, *J. Geophys. Astrophys. Fluid Dyn.*, *8*, 245–277.

Fasham, M. J. R. (Ed.) (2003), *Ocean Biogeochemistry: The Role of the Ocean Carbon Cycle in Global Change*, *Global Change—IGBP Ser.*, vol. XVIII, 297 pp., Springer, New York.

Field, C. B., M. J. Behrenfeld, J. T. Randerson, and P. G. Falkowski (1998), Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, *281*, 237–240.

Gordon, H. R., and D. K. Clark (1981), Clear water radiances for atmospheric correction of Coastal Zone Color Scanner imagery, *Appl. Opt.*, *20*, 4175–4180.

Gordon, H. R., and A. Y. Morel (1983), *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*, 114 pp., Springer, New York.

Gregg, W. W., and M. E. Conkright (2002), Decadal changes in global ocean chlorophyll, *Geophys. Res. Lett.*, *29*(15), 1730, doi:10.1029/2002GL014689.

International Ocean Colour Coordinating Group (IOCCG) (1999), *Status and Plans for Missions*, edited by J. A. Yoder, Dartmouth, N. S., Canada.

Longhurst, A. (1995), Seasonal cycles of pelagic production and consumption, *Prog. Oceanogr.*, *36*, 77–167.

Maritorena, S., and D. A. Siegel (2005), Consistent merging of satellite ocean color data using a semi-analytical model, *Remote Sens. Environ.*, *94*, 429–440.

Maritorena, S., D. A. Siegel, and A. R. Peterson (2002), Optimal tuning of a semi-analytical model for global applications, *Applied Opt.*, *41*, 2705–2714.

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

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

Nelson, N. B., and D. A. Siegel (2002), Chromophoric DOM in the open ocean, in *Biogeochemistry of Marine Dissolved Organic Matter*, edited by D. A. Hansell and C. A. Carlson, pp. 547–578, Elsevier, New York.

O'Reilly, J., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, M. Kahru, S. A. Garver, and C. R. McClain (1998), Ocean color algorithms for SeaWiFS, *J. Geophys. Res.*, *103*, 24,937–24,953.

O'Reilly, J. E., et al. (2000), Ocean color chlorophyll-*a* algorithms for SeaWiFS, OC2 and OC4, in *SeaWiFS Postlaunch Calibration and Validation Analyses: Part 3*, *SeaWiFS Postlaunch Tech. Rep. Ser.*, vol. 11, version 4, edited by S. B. Hooker and E. R. Firestone, pp. 9–23, NASA Goddard Space Flight Cent., Greenbelt, Md.

Otero, M. P., and D. A. Siegel (2004), Spatial and temporal characteristics of sediment plumes and phytoplankton blooms in the Santa Barbara Channel, *Deep Sea Res., Part II*, *51*, 1129–1149.

Siegel, D. A., S. Maritorena, N. B. Nelson, D. A. Hansell, and M. Lorenzi-Kayser (2002), Global ocean distribution and dynamics of colored dissolved and detrital organic materials, *J. Geophys. Res.*, *107*(C12), 3228, doi:10.1029/2001JC000965.

Siegel, D. A., S. Maritorena, N. B. Nelson, and M. J. Behrenfeld (2005), Independence and interdependencies of global ocean color properties: Reassessing the bio-optical assumption, *J. Geophys. Res.*, *110*, C07011, doi:10.1029/2004JC002527.

Sullivan, C. W., K. R. Arrigo, C. R. McClain, J. C. Comiso, and J. K. Firestone (1993), Distributions of phytoplankton blooms in the Southern Ocean, *Science*, *262*, 1832–1837.

Werdell, P. J., and S. W. Bailey (2005), An improved bio-optical data set for ocean color algorithm development and satellite data product validation, *Remote Sens. Environ.*, *98*, 122–140.

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

M. J. Behrenfeld, Department of Botany and Plant Pathology, Cordley Hall 2082, Oregon State University, Corvallis, OR 97331–2902, USA.

C. R. McClain, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA.

S. Maritorena, N. B. Nelson, and D. A. Siegel, Institute for Computational Earth System Science, University of California, Santa Barbara, 6844 Ellison Hall, Santa Barbara, CA 93106–3060, USA. (davey@icess.ucsb.edu)