

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

D. A. Siegel,¹ S. Maritorea,¹ N. B. Nelson,¹ M. J. Behrenfeld,² and C. R. McClain³

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. Maritorea, 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 Maritorea, 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) [Maritorea *et al.*, 2002; Maritorea 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 [Maritorea *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 [Maritorea *et al.*, 2002; Maritorea 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

¹Institute for Computational Earth System Science, University of California, Santa Barbara, California, USA.

²Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA.

³NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

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

^aThis 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 $\mu\text{W cm}^{-2} \text{nm}^{-1} \text{sr}^{-1}$ 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 $\mu\text{W cm}^{-2} \text{nm}^{-1} \text{sr}^{-1}$ 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 CDOM 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 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).

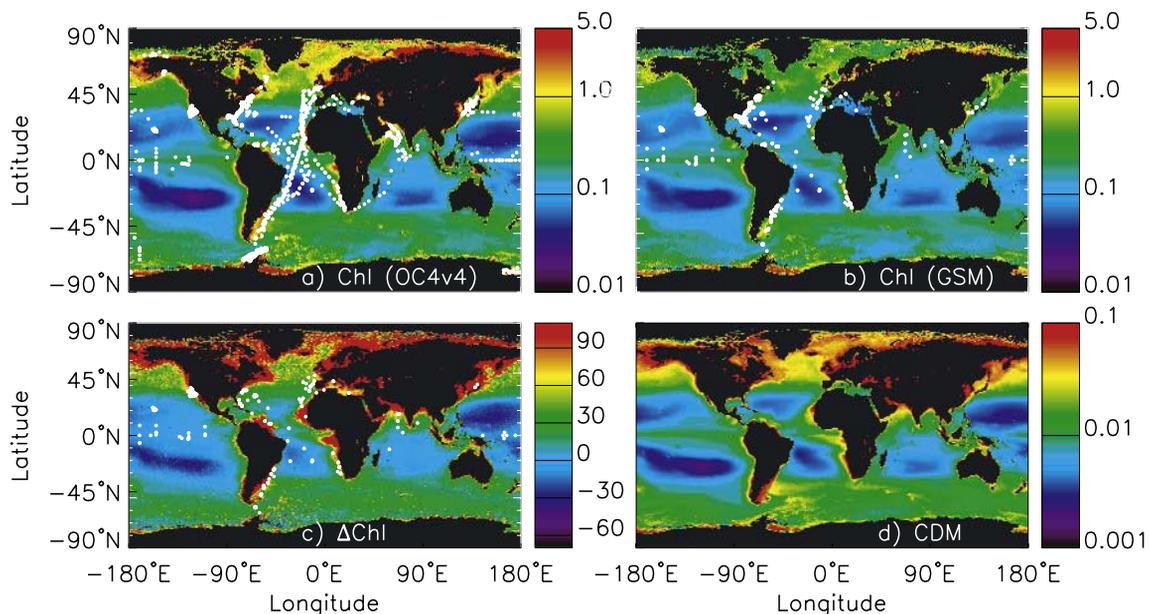


Figure 1. Global climatology of (a) Chl(OC4v4) (mg m^{-3}), (b) Chl(GSM) (mg m^{-3}), (c) the normalized percentage difference between them ($\Delta\text{Chl} = 100 \cdot (\text{Chl}(\text{OC4v4}) - \text{Chl}(\text{GSM})) / \text{Chl}(\text{GSM})$; units %) and (d) the CDM distribution (m^{-1}). These climatologies are constructed from simple averages from the available SeaWiFS data set (October 1997 to December 2004) averaged to $1/3^\circ$ 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.