Remote Sensing of Environment 113 (2009) 1319–1330

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua

P. Jeremy Werdell [a,*], Sean W. Bailey [a], Bryan A. Franz [a], Lawrence W. Harding Jr. [b], Gene C. Feldman [a], Charles R. McClain [a]

[a] NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
[b] Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, Maryland 21613, USA

ARTICLE INFO

Article history:
Received 10 September 2008
Received in revised form 18 February 2009
Accepted 21 February 2009

Keywords:
Chlorophyll
Remote sensing
Ocean color
Water quality
SeaWiFS
MODIS
Chesapeake Bay

ABSTRACT

Concentrations of the phytoplankton pigment chlorophyll-a (C_a) provide indicators of nutrient over-enrichment that has negatively affected Chesapeake Bay, U.S.A. C_a time-series from the National Aeronautics and Space Administration (NASA) Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer aboard the Aqua spacecraft (MODIS-Aqua) provide observations on temporal and spatial scales that far exceed current field and aircraft sampling strategies. These sensors provide consistent, frequent, and high density data to potentially complement ongoing Bay monitoring activities. We used the *in situ* Water Quality Monitoring Data set of the Chesapeake Bay Program to evaluate decade-long time-series of SeaWiFS and MODIS-Aqua C_a retrievals in the Bay. The accuracy of the retrievals generally degraded with increasing latitude as the optical complexity increases northward. C_a derived using empirical ("band ratio") algorithms overestimated *in situ* measurements by 10–50 and 40–100% for SeaWiFS and MODIS-Aqua, respectively, but with limited variability. C_a derived using spectral-matching algorithms showed less bias for both sensors, but with significant variability and sensitivity to radiometric errors. Regionally-tuned empirical algorithms performed best throughout the Bay, offering a combination of reasonable accuracy and high spatial coverage. The radiometric spectral resolution used as input to the algorithms strongly influenced the quality of C_a retrievals from both sensors. These results establish a baseline quantification of algorithm and sensor performance in a variable and stressed ecosystem against which novel approaches might be compared.
© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The National Aeronautics and Space Administration (NASA) Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer aboard the Aqua spacecraft (MODIS-Aqua) provide an opportunity to study the marine biosphere on spatial and temporal scales unattainable by conventional sampling methods. These ocean color satellites measure the upwelling radiance emitted from the top of the Earth's atmosphere at discrete visible and infrared wavelengths. Atmospheric correction algorithms (Gordon & Wang, 1994) are used to remove the contribution of the atmosphere from the total signal and produce estimates of spectral radiance exiting the water mass. These water-leaving radiances ($L_w(\lambda)$; $\mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$) are used to retrieve additional geophysical properties via the application of bio-optical algorithms. Typically, such algorithms adopt one of two forms: (1) empirical expressions that are derived by statistical correlation of coincident *in situ* radiometry and a biogeochemical product of interest (e.g., O'Reilly et al., 1998); and (2) semi-analytical expressions wherein a simplified form of the radiative transfer equation is inverted to retrieve the product (e.g., Maritorena et al., 2002).Many examples exist for both forms, all of which exhibit individual strengths and weaknesses for coastal and estuarine remote sensing.

Chesapeake Bay is one of the largest and historically most productive estuaries in North America. An extensive watershed contributes an annual average of 2300 $m^3\ s^{-1}$ freshwater flow to the Bay with accompanying dissolved and particulate matter, including nutrients and sediments. Unfortunately, the Chesapeake Bay suffers from excessive anthropogenic nutrient loading (Fisher et al., 2006; Malone, 1992) and, consequently, intense seasonal anoxia (Hagy et al., 2004; Kemp et al., 2005). As such, considerable effort has focused on the relationship of freshwater flow and nutrient input to water quality, expressed as elevated algal biomass (Malone et al., 1996) and primary productivity (Harding et al., 2002). Algal biomass, commonly quantified as the concentration of the phytoplankton pigment chlorophyll-a (C_a; mg m^{-3}), is an important indicator of eutrophication in marine ecosystems (Smith, 2006).

Long-term (e.g., decadal) time-series of *in situ* measurements currently provide the main data source for evaluating water quality. The Chesapeake Bay Program (CBP), a cooperative effort between the federal government (e.g., U.S. Environmental Protection Agency) and state and local governments within the Chesapeake Bay watershed, initiated a

* Corresponding author. Science Systems and Applications Inc., Lanham, Maryland 20706, USA.
E-mail address: jeremy.werdell@nasa.gov (P.J. Werdell).

0034-4257/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2009.02.012

Table 1
Coefficients for the OC version 5 algorithms (O'Reilly, personal communication).

	λ_b	λ_g	c_0	c_1	c_2	c_3	c_4
OC4	443,490,510	555	0.3080	−3.0882	3.0440	−1.2013	−0.7992
OC3S[a]	443,490	555	0.2409	−2.4768	1.5296	0.1061	−1.1077
OC3M[b]	443,488	551	0.2254	−2.6354	1.8071	0.0063	−1.2931

[a] For SeaWiFS.
[b] For MODIS-Aqua.

Water Quality Monitoring Program in 1984 to facilitate watershed restoration efforts (Chesapeake Bay Program, 1993). The CBP selected 49 stations in the mainstem Bay to be sampled approximately monthly to measure 19 hydrographic water quality parameters. Unfortunately, the time-series remains temporally and spatially discontinuous and susceptible to unfavorable sampling conditions (e.g., poor weather) that may interrupt *in situ* data collection. Data products from SeaWiFS and MODIS-Aqua have an advantage over *in situ* and aircraft measurements in this regard given their consistent, frequent, and high density sampling of Chesapeake Bay (limited, to a first order, only by cloud cover).

The CBP Water Quality Monitoring Data set provides a unique opportunity to compare decade-long time-series of satellite and *in situ* C_a observations of a highly variable and stressed ecosystem. A time-series of this length permits separation of sampling variability from geophysical change and facilitates understanding of how the ecosystem responds to stresses. Space borne sensors with the ability to capture phytoplankton dynamics and hydrographic phenomena in such a geographically confined, dynamic environment provide novel data sources for ecosystem management. To this end, we present a case study that evaluated SeaWiFS and MODIS-Aqua time-series of C_a in Chesapeake Bay using the CBP Water Quality Monitoring Data set. Harding et al. (2005) and Signorini et al. (2005) validated C_a retrievals from SeaWiFS in the Bay using temporally limited *in situ* data sets. To our knowledge, similar analyses have yet to be executed for MODIS-Aqua or for the full decade-long SeaWiFS time-series.

We describe the spatial and temporal variability of SeaWiFS and MODIS-Aqua C_a estimates in the Chesapeake Bay with the goal of evaluating the performance of two bio-optical algorithms, the congruity of the two sensors, and the applicability of these data for water quality monitoring. Our four step approach begins with a brief review of the C_a algorithms (one for each algorithmic form), followed by an explanation of the satellite and *in situ* data acquisition and preparation. We then quantify the long-term accuracies of the remote-sensing algorithms using *in situ* data as ground truth. Finally, we assess the consistency of C_a estimated from SeaWiFS and MODIS-Aqua for the Bay. We used standard (operational) atmospheric correction and bio-optical algorithms for this study to establish baseline quantification of algorithm performance and sensor congruence. In doing so, we illustrate the sensitivity of the approaches to sensor spectral resolution and the benefits of regional algorithm parameterization.

2. Methods

2.1. Bio-optical algorithms

The types and relative concentrations of optically active constituents in the water column (e.g., algae and chromophoric dissolved organic matter) largely determine the spectral shape and magnitude of $L_w(\lambda)$. Given the optical complexity of Chesapeake Bay and its spatial and temporal variability, considerable effort has focused on developing approaches to relate $L_w(\lambda)$ to the biogeochemistry of the Bay (Gitelson et al., 2007; Magnuson et al., 2004; Tzortziou et al., 2007; Zawada et al., 2007). A survey of all relevant algorithms directed at this problem is well beyond the scope of this paper. Rather, we focus on two prominent C_a algorithms currently used in the operational processing of SeaWiFS and MODIS-Aqua (McClain et al., 2006). Our goal is to define the state-of-the-art regarding the use of these widely-accessible algorithms for water quality monitoring in Chesapeake Bay, which will in turn provide a foundation for evaluating novel or expanded algorithm approaches.

2.1.1. Ocean Chlorophyll algorithm

The first algorithm under consideration is the empirical Ocean Chlorophyll (OC) form described in O'Reilly et al. (1998), which currently provides the operational C_a products for SeaWiFS and MODIS-Aqua. This form describes the polynomial best fit that relates log-transformed C_a to a log-transformed ratio of remote-sensing reflectances (R_{rs}; the ratio of L_w to downwelling surface irradiance; sr^{-1}):

$$log_{10}(C_a) = c_0 + \sum_{i=1}^{N} c_i log_{10}\left(\frac{R_{rs}(\lambda_b)}{R_{rs}(\lambda_g)}\right) \quad (1)$$

where the coefficients c_0 and c_i are listed in Table 1. The standard SeaWiFS algorithm, OC4, uses four $R_{rs}(\lambda)$, where λ_b is the greatest of $R_{rs}(443)$, $R_{rs}(490)$, and $R_{rs}(510)$ and λ_g is $R_{rs}(555)$. The shift from 443 to 490 nm and from 490 to 510 nm generally occurs near C_a of 0.5 and 2 mg m^{-3}, respectively. The MODIS-Aqua version, OC3, uses three $R_{rs}(\lambda)$, where λ_b is the greater of $R_{rs}(443)$, $R_{rs}(488)$ and λ_g is $R_{rs}(551)$. A three-wavelength version also exists for SeaWiFS, but will not be highlighted in this analysis until section 4.3. The polynomials were derived using the globally-distributed NASA bio-optical Marine Algorithm Data set (O'Reilly, personal communication; Werdell & Bailey, 2005). As such, each OC algorithm describes a mean trend over a wide C_a range and not specific bio-optical relationships at regional scales or within narrow C_a ranges (Brown et al., 2008). Data collected in the mainstem Chesapeake Bay account for 9% of all NOMAD observations and 27% of those from eutrophic waters (i.e., C_a>1 mg m^{-3}) (Fig. 1). Regionally-tuned variants exist, as will be discussed in section 4.4, but we first evaluate standard SeaWiFS and MODIS-Aqua algorithms to develop a baseline against which Bay-specific versions can be compared.

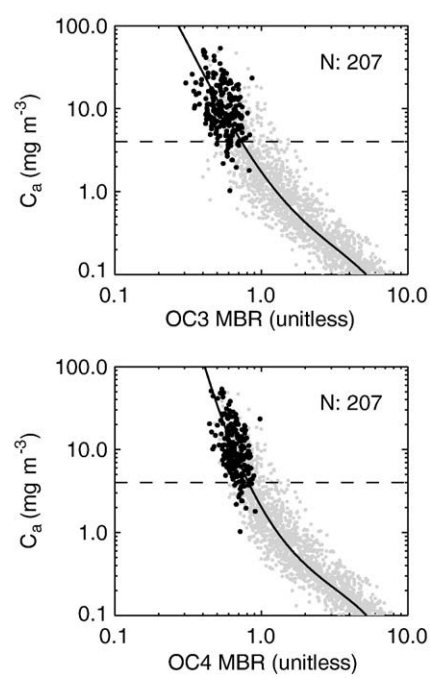

Fig. 1. *In situ* R_{rs} maximum band ratios (MBR) versus C_a for the OC3 and OC4 algorithms. The gray circles show the NOMAD data used to derive the polynomial coefficients for each expression (N=2389). The black circles indicate stations located within the Chesapeake Bay, with the number of stations also reported. The solid lines were generated using Eq. (1) and Table 1. Dashed horizontal lines demark C_a=4 mg m^{-3}.

2.1.2. Garver–Siegel–Maritorena algorithm

The second algorithm is a modified version of the semi-analytical Garver–Siegel–Maritorena model (Maritorena et al., 2002) specifically parameterized for Chesapeake Bay by Magnuson et al. (2004). This approach relates spectral inherent optical properties to $R_{rs}(\lambda)$ via a polynomial expansion of the spectrally dependent ratio $b_b\,(a+b_b)^{-1}$, where $a(\lambda)$ (m^{-1}) and $b_b(\lambda)$ (m^{-1}) are the marine absorption and backscattering coefficients, respectively (Gordon et al., 1975). These spectral coefficients are commonly partitioned into components via $a=a_w+a_\phi+a_{dg}$ and $b_b=b_{bw}+b_{bp}$, where the subscripts w, ϕ, dg, and p indicate water (known), phytoplankton, non-algal particles (NAP) + chromophoric dissolved organic matter (CDOM), and total particles, respectively. Note that the a_{dg} combination cannot currently be decomposed into its two components using remote-sensing methods.

In GSM, the components are further expanded into:

$$a_\phi(\lambda) = C_a a_/^*(\lambda), \tag{2}$$

$$a_{dg}(\lambda) = a_{dg}(443)exp[-S(\lambda - 443)], \tag{3}$$

$$b_{bp}(\lambda) = b_{bp}(443)(\lambda/443)^{-\eta}, \tag{4}$$

where a_ϕ^* is the C_a-specific absorption coefficient ($m^2\ mg^{-1}$), S is the spectral decay constant for NAP + CDOM absorption (unitless), and η is the power-law exponent for the particulate backscattering coefficient (unitless). In GSM, these three terms are assigned constant values. Using $R_{rs}(\lambda)$ as input, the Levenberg–Marquardt nonlinear least-squares procedure is employed to solve for the remaining unknown terms, namely C_a, $a_{dg}(443)$, and $b_{bp}(443)$. Failure of the model parameterization (Eqs. (2)–(4)) to capture natural biogeophysical variability and ambiguity (nonuniqueness) in the least-squares fit both contribute to uncertainty in the derived products (Defoin-Patel & Chami, 2007). Magnuson et al. (2004) derived seasonally and geographically varying constants for Eqs. (2)–(4) using *in situ* data from Chesapeake Bay (see their Table 5). We evaluate only this regionally-parameterized version as improvements realized with Bay-specific tuning of GSM have been independently verified (Signorini et al., 2005; Werdell et al., 2007).

2.2. Data acquisition

2.2.1. Satellite data

We acquired ~6200 SeaWiFS and ~3000 MODIS-Aqua spatially-extracted Level-1A files containing all or part of the Bay from the NASA Ocean Biology Processing Group (McClain et al., 2006). These files represent a significant time-series for both sensors, spanning September 1997 to March 2007 for SeaWiFS and June 2002 to March 2007 for MODIS-Aqua, at ~1 km^2 spatial resolution at nadir. We generated Level-2 C_a files using the OBPG processing software MSl12 (Franz et al., 2005) configured for SeaWiFS Reprocessing 5.2 and MODIS-Aqua Reprocessing 1.1, which includes the (Gordon & Wang, 1994) atmospheric correction approach, plus corrections for near-infrared water-leaving radiances, bi-directional reflectance, and spectral band-pass effects (Morel et al., 2002; Patt et al., 2003). GSM processing used all visible SeaWiFS (412, 443, 490, 510, 555, and 670 nm) and MODIS-Aqua (412, 443, 488, 531, 551, 667, and 678 nm) wavelengths. The operational pixel-masking scheme for each sensor was adopted, with the exception that pixels with stray light contamination were retained (i.e., to avoid excluding near-shore waters where light reflected off land occasionally enters the field of view of the satellite). Quality control metrics described in Werdell et al. (2007) were applied to both data sets to ensure that only the most reliable data were retained for analysis. This included removing scenes with satellite zenith angles greater than 54° or fewer than 25% cloud-free marine pixels. In the end, approximately 9 days of data per month were available for each sensor.

2.2.2. In situ data

We acquired ~15,750 discrete fluorometric and spectrophotometric C_a samples from the CBP Water Quality Monitoring Data set (Chesapeake Bay Program, 1993) (Fig. 2). These data were supplemented with an additional ~2300 independent fluorometric C_a samples collected as part of the NASA Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) and NSF Land Margin Ecosystem Research (LMER) Trophic Interactions in Estuarine Systems (TIES) programs (Harding & Magnuson, 2003; Harding et al., 2005). We considered only near-surface samples (depths ≤ 1 m) for comparison with the satellite C_a retrievals as the Bay typically has shallow optical depths. Replicate samples were averaged. Additional details regarding the post-collection treatment of *in situ* data, including quality control metrics, are provided in Werdell and Bailey (2005). Of the ~18,050 *in situ* measurements, 8 and 52 stations, respectively, had C_a < 0.3 and < 1 mg m^{-3}, and 24 stations had values > 100 mg m^{-3}. The lowest and median *in situ* C_a were 0.17 and 7.7 mg m^{-3}, respectively. Both *in situ* and derived C_a exclude contributions by phaeopigments in this analysis.

2.3. Data analysis

2.3.1. Study area and data stratification

The three largest tributaries, the Susquehanna, Potomac, and James Rivers, contribute over 85% of the total riverine input to the Chesapeake Bay. The Susquehanna River at the head of the estuary contributes over

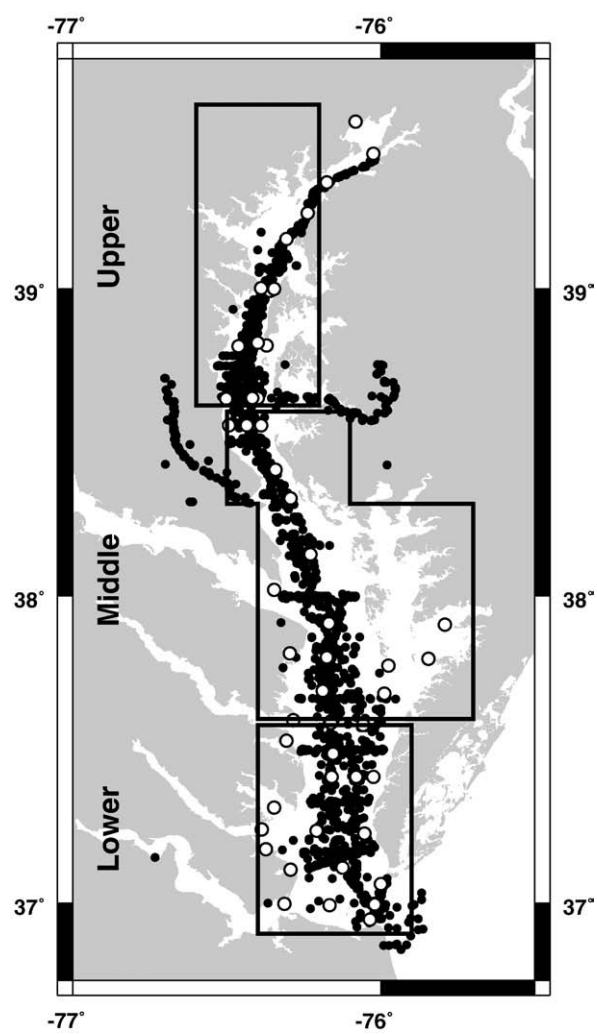

Fig. 2. Locations of *in situ* C_a within the Chesapeake Bay and boxes denoting our adopted regional designations. The CBP and SIMBIOS/LMER-TIES data sets are indicated by empty and filled circles, respectively.

half of the total freshwater and inorganic nutrient inputs (Fisher et al., 1988). Our study period (September 1997 through March 2007) encompasses three wet years (1998, 2003, and 2004; upper quartile compared to mean long-term inflow) and three dry years (1999, 2001, and 2002; lower quartile),with the remaining four considered normal (U.S. Geological Survey, 2007). Despite their designation as normal years, 2000 and 2005 had low and high inflows, respectively, suggesting that 1999–2002 to be appreciably dry compared to 2003–2005 (Acker et al., 2005).

Bio-optical properties of the Bay are predominantly controlled by the annual cycle of phytoplankton (Marshall et al., 2006) confounded by spatially and temporally varying sources of terrestrially-derived dissolved and particulate matter (Gallegos et al., 2005). The winter–spring freshet from the Susquehanna River largely regulates the timing and magnitude of the spring diatom bloom (Adolf et al., 2006; Miller et al., 2006), which is followed by a summer maximum of primary productivity and increased relative abundance of picoplankton and flagellates (Harding et al., 2002). Seasonal changes in $a_{\phi}(\lambda)$ accompany this shift in phytoplankton species. Concentrations of CDOM and NAP generally correlate with riverine discharge, thereby creating latitudinal gradients in $a_{dg}(\lambda)$ and S, although this pattern can be obscured by frontal features, tidal cycles, estuarine circulation, and sediment resuspension (Harding et al., 2005; Zawada et al., 2007). The relative concentrations and size distributions of biogenic and mineral particles, which seldom directly covary, determine $b_{bp}(\lambda)$ and η. Gallegos et al. (2005), for example, observed late summer increases in η in a sub-estuary of the Bay that were consistent with seasonal reductions in phytoplankton size class and concentrations of suspended solids (Zawada et al., 2007).

We stratified the C_a data spatially and temporally to facilitate consideration of scales of biogeochemical variability when interpreting our results. Seasons were defined using the day of year of data collection, with days 80, 172, 266, and 355 defining the transitions of winter–spring, spring–summer, summer–fall, and fall–winter (roughly following Northern Hemisphere equinoxes and solstices). We adopted the regional stratification of Magnuson et al. (2004), who defined the boundaries between the Lower–Middle and Middle–Upper Bays to be latitudes 37.6° and 38.6°N, respectively (Fig. 2). Note that our spatial stratification follows a latitudinal salinity (Sal) gradient, with the Upper zone largely oligohaline (Sal ≤ 10), the Middle mesohaline (10 < Sal ≤ 20), and Lower polyhaline (Sal > 20). We excluded dates when satellite sample sizes for a given region were less than 200 valid (unmasked) marine pixels. For both sensors, the latter eliminated <4% of all dates for the Lower and Middle Bays, but excluded an additional 1–3 days per month for the Upper Bay.

2.3.2. Analysis tools

We employed three methods to evaluate the C_a retrievals from SeaWiFS and MODIS-Aqua using the *in situ* measurements. The accuracy of the C_a retrievals was best quantified using this combination of approaches for validation, as each method had strengths and weaknesses. First, we statistically compared coincident Level-2 satellite and *in situ* C_a using the OBPG satellite data product validation system (Bailey & Werdell, 2006). For this "match-up" analysis, we retained the default (globally-parameterized) OBPG configuration, namely: (1) temporal coincidence was defined as ±3 h; (2) satellite C_a were the filtered median (via the semi-interquartile range) of all unmasked pixels within a 5 × 5 pixel box centered on the *in situ* target; and (3) satellite C_a were excluded when more than 50% of marine pixels within this box were masked or when the coefficient of variation of the valid marine pixels exceeded 0.15. While ecological patchiness might dictate a smaller box, uncertainty about the median derived from valid pixels within this box increases in accordance with the decreased sample size. Harding et al. (2005) derived analogous results using 3 × 3 and 5 × 5 pixel boxes and found reducing the temporal threshold from same-day to ±3 h significantly improved the quality of C_a match-ups. Tidal and estuarine circulation might dictate further reduction of the temporal threshold. The corresponding reduction in the number of match-ups proved too costly for our analysis, however, as reducing the threshold from ±3 h to ±30 min resulted in ~65% fewer match-ups for both sensors. Sensor digitization and algorithm noise precluded the use of a single pixel in such analyses (Hu et al., 2001).

Next, we generated regional histograms and time-series of satellite-derived and *in situ* C_a as described in Werdell et al. (2007). For the histogram analyses, we calculated: (1) relative percent differences (RPD) between the satellite and *in situ* distribution medians to quantify relative biases in the C_a estimates; and (2) percent differences in the semi-interquartile ranges (SPD) of the distributions to quantify relative differences in distribution widths (here, the semi-interquartile range is the range covered by values of $\ln(C_a)$ such that 50% of $\ln(C_a)$ values occur with equal probability on either side of the median). The reported percent differences are relative to the *in situ* measurement, for example, $100\% \cdot (C_a^{satellite}/C_a^{in\ situ} - 1)$. We prepared the time-series by calculating the monthly geometric means of all available (unmasked) data. This averaging improved the clarity of the time-series and eliminated anomalous C_a retrievals. Alternate filtering and statistical smoothing techniques produced slight quantitative differences, but revealed similar spatial and temporal patterns. For Level-2 match-ups and time-series, only retrievals within $0.001 \le C_a \le 100$ mg m^{-3} were considered, as this is the effective operational range of the OC and GSM algorithms.

3. Results

3.1. Level-2 match-ups

The comparison of coincident satellite and *in situ* observations provided estimates of the accuracy and precision of the SeaWiFS and MODIS-Aqua data products in Chesapeake Bay. OC-derived C_a for the Bay were positively biased, as indicated by the median satellite-to-*in situ* ratios of 1.30 and 1.69, respectively (Fig. 3). For comparison, Harding et al. (2005) and Signorini et al. (2005) reported ratios of 1.97 and 1.24, respectively, for SeaWiFS (using shorter time-series and previous SeaWiFS Reprocessing configurations). The elevated retrievals persisted in all regions of the mainstem Bay, but were most pronounced at the lowest C_a. As illustrated in Fig. 1, NOMAD Chesapeake Bay stations with C_a<4 mg m^{-3} are not well represented by either OC3 or OC4 (17 stations), whereas both algorithms clearly bisect those stations with higher C_a. When only C_a>4 mg m^{-3} were considered, SeaWiFS match-up ratios and median absolute percent differences (APD) reduced to 1.03 and 35.3%, respectively. For MODIS-Aqua, however, this reduction in dynamic range reduced the ratio and APD to only 1.65 and 65.1%, suggesting that either inadequate sampling (41 stations compared to 252 for SeaWiFS) or that inaccurate reflectance ratios were used as inputs to OC3. In contrast, GSM-derived C_a maintained nominal biases for both sensors, with ratios deviating from unity by ~10%. Despite its improved match-up accuracy, however, GSM returned 40% fewer stations for both sensors with C_a between 0.001 and 100 mg m^{-3}. The APD for GSM improved upon that for OC by 29% for MODIS-Aqua, but degraded by 9% for SeaWiFS. Unfortunately, match-ups for both sensors were acquired under a limited range of environmental conditions (e.g., varied water turbidity and solar and satellite geometries), which limited our interpretation of these results and reinforced the need to consider these statistics in combination with those from the frequency distributions and time-series.

3.2. Data distributions

Comparisons of frequency distributions identified relative biases and differences in dynamic ranges among the satellite and *in situ* data sets (Table 2). Here, coincidence is defined only through season and region, which differs from the Level-2 validation approach presented

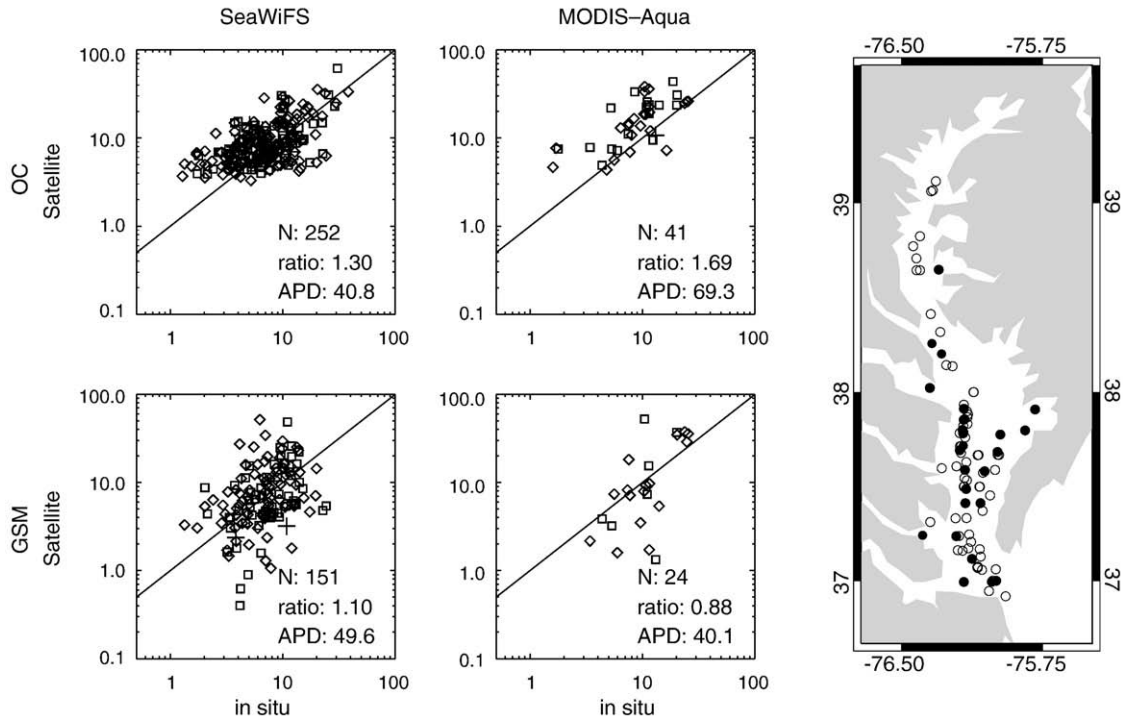

Fig. 3. SeaWiFS and MODIS-Aqua satellite-to-*in situ* C_a validation results for the OC and GSM algorithms. *N*, ratio, and APD indicate the sample size, median satellite-to-*in situ* ratio, and median absolute percent difference, respectively. The APD for each match-up is calculated via 100% $|C_a^{satellite}/C_a^{in\ situ} - 1|$. The crosses, empty squares, and empty diamonds denote Upper, Middle, and Lower Bay stations, respectively. The solid line demarks a 1:1 relationship. In the map, empty and filled circles show the locations of SeaWiFS and MODIS-Aqua match-ups, respectively.

in section 3.1. *In situ* C_a in the Bay were lognormally distributed over a broad dynamic range (~1 to 50 mg m^{-3}) (Fig. 4). SeaWiFS and MODIS-Aqua C_a exhibited similarly shaped lognormal distributions, particularly those derived using the OC algorithms, despite an increased number of observations relative to the sample size of the *in situ* data. The similar widths of the OC3- and OC4-derived C_a suggest the breadth of their dynamic ranges to be consistent with the *in situ* C_a, however, positive biases appeared in the Middle region and intensified in the Upper region. Few C_a<3 mg m^{-3} were retrieved, similar to the results in section 3.1. In contrast, GSM returned an increased number of low (<1 mg m^{-3}) and high (>30 mg m^{-3}) C_a, effectively flattening the histogram relative to the *in situ* measurements. For both sensors, GSM returned ~41, 38, and 86% fewer viable pixels on average than the OC algorithms in the Lower, Middle, and Upper Bay, respectively.

Differences between the remotely-sensed and *in situ* C_a frequency distributions showed strong latitudinal dependencies. SeaWiFS and MODIS-Aqua C_a from both algorithms showed marginal long-term biases compared to *in situ* data in the Lower Bay, at least with regard to their modes (Fig. 4). Full-year (not stratified by season) SeaWiFS C_a distributions from OC4 also aligned well with the *in situ* data in the Middle Bay, whereas the full-year MODIS-Aqua C_a from OC3 showed moderate, positive departures. Full-year SeaWiFS C_a distributions from GSM remained similar for all Bay regions, while those for MODIS-Aqua did so for the Lower and Middle regions. The RPD and SPD for both sensors reached their respective maxima in the Upper Bay, independent of season, emphasizing the inability of the algorithms to adequately account for non-algal contributions to $R_{rs}(\lambda)$ in optically complex locations (Table 2). Interestingly, MODIS-Aqua retrievals from both algorithms in the Upper Bay far exceeded those of SeaWiFS in this region (by factors of two or more). At higher latitudes, GSM from both sensors generally reported lower RPD relative to the OC algorithms, but higher SPD (broad, flat histograms whose centers more closely matched the *in situ* distributions).

3.3. Time-series

Comparison of time-series identified relative biases and seasonal differences amongst our coincident data sets. In the Lower and Middle Bays, moderate biases persisted in monthly comparisons of *in situ* C_a with the retrievals from OC (Fig. 5) and GSM (Fig. 6), although satellite estimates largely fell within one standard deviation of *in situ* averages. Both algorithms exhibited seasonal patterns in these zones (e.g., distinct winter–spring blooms) and paralleled the interannual variability of *in situ*

Table 2
Percent differences of satellite and *in situ* C_a distribution medians (RPD) and semi-interquartile ranges (SPD).

		SeaWiFS				MODIS-Aqua			
		OC4		GSM		OC3		GSM	
		RPD	SPD	RPD	SPD	RPD	SPD	RPD	SPD
Upper Bay	Spring	76.6	22.3	58.5	94.7	166.2	5.7	161.9	171.3
	Summer	62.2	122.3	−31.1	90.9	124.2	44.0	64.5	218.0
	Fall	74.1	54.8	17.2	182.4	187.8	27.7	227.4	208.1
	Winter	81.5	46.0	29.6	172.5	185.6	28.3	124.4	172.6
	All	53.9	44.1	3.6	132.1	138.4	16.9	124.1	174.2
Middle Bay	Spring	45.2	6.2	55.7	5.9	99.2	−10.7	54.2	44.9
	Summer	24.1	40.0	−28.8	33.8	66.8	26.2	−39.3	44.2
	Fall	22.1	5.7	−2.3	88.4	96.1	16.5	−3.5	107.6
	Winter	55.5	38.6	3.2	157.3	143.7	10.4	−4.3	157.5
	All	28.9	26.1	−0.1	82.3	93.1	13.5	−7.8	104.6
Lower Bay	Spring	21.7	−15.1	56.0	−2.1	67.6	−7.4	47.3	27.0
	Summer	1.5	9.1	−33.6	10.1	24.7	30.5	−45.0	54.5
	Fall	−13.9	−34.7	−26.9	30.2	17.6	−14.0	−22.0	69.7
	Winter	22.2	25.6	−7.7	194.0	120.5	23.1	−4.6	186.7
	All	1.7	−5.3	−7.6	54.7	47.1	10.6	−10.3	88.6

RPD were calculated as 100% · $(C_a^{satellite}/C_a^{in\ situ} - 1)$. Use of the median C_a to calculate RPD is prudent given the unimodal symmetry of the lognormally transformed data and the proximity of these values to the distribution modes. Semi-interquartile ranges (SIQR) were computed for $\ln(C_a^{satellite})$ and $\ln(C_a^{in\ situ})$. SPD were calculated as 100% · $[\exp(\mathrm{SIQR}^{satellite} - \mathrm{SIQR}^{in\ situ}) - 1]$.

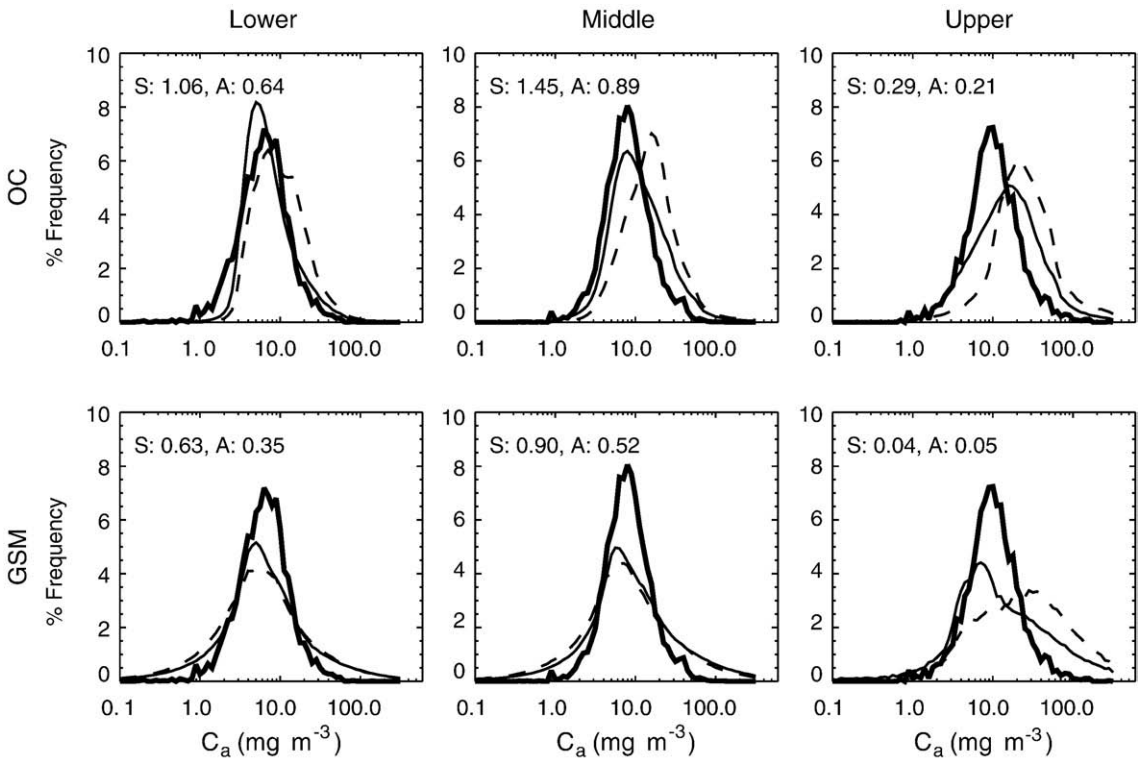

Fig. 4. *In situ* C_a distributions (thick lines) in the Lower, Middle, and Upper Bays compared to OC and GSM retrievals for SeaWiFS (thin solid) and MODIS-Aqua (thin dashed). OC4 was used for SeaWiFS and OC3 for MODIS-Aqua. Samples sizes (in million pixels) for satellite retrievals are provided in each panel, with SeaWiFS indicated by *S* and MODIS-Aqua indicated by *A*. *In situ* sample sizes are 7204, 5814, and 3660 for the Lower, Middle, and Upper Bay, respectively. Data from all four seasons are included.

C_a. The algorithms also captured the transition from dry conditions (1999–2002) to wet conditions (2003–2005). For our study period, the Bay experienced enhanced winter–spring blooms, stronger and longer-lived summer blooms, and increased long-term median *in situ* C_a (from ~6 to 10 mg m^{-3}, ~8 to 11 mg m^{-3}, and ~10 to 12 mg m^{-3} in the Lower, Middle, and Upper Bays, respectively) as it transitioned into a series of wet years. Unfortunately, satellite retrievals using OC and GSM in the Upper Bay did not display comparable seasonal patterns to *in situ* C_a.

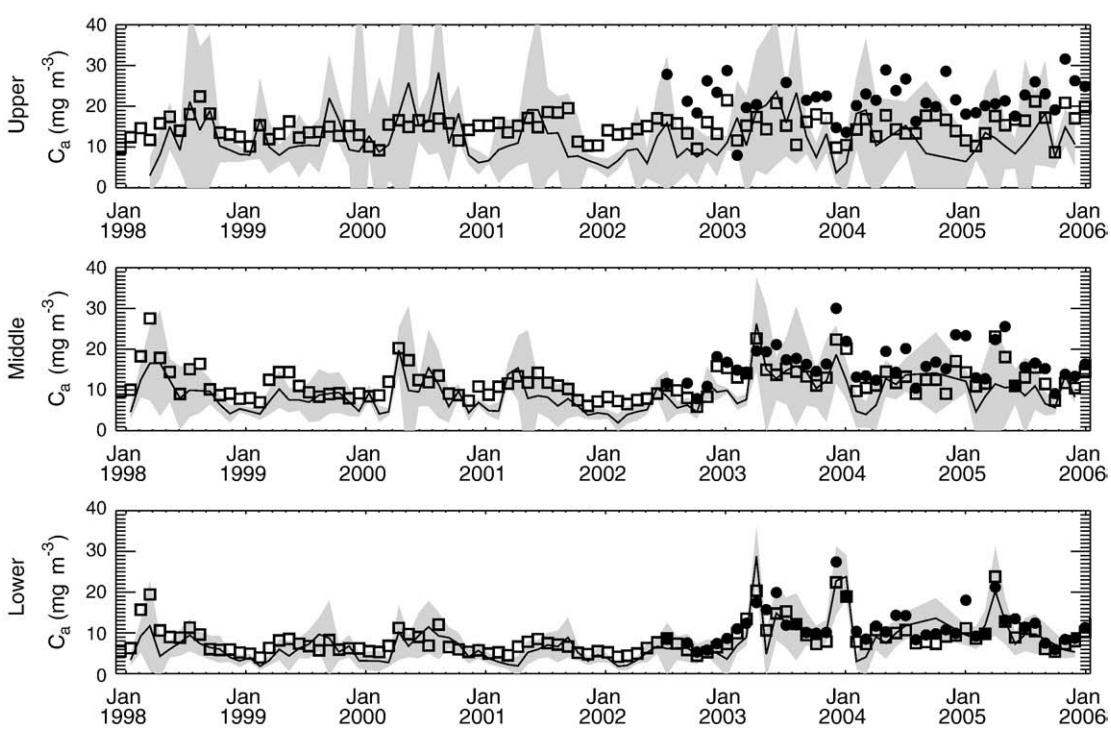

Fig. 5. Monthly averages of *in situ* C_a (thin lines) in the Upper, Middle, and Lower Bays compared to OC retrievals for SeaWiFS (OC4; empty squares) and MODIS-Aqua (OC3; filled circles). The grey shaded area represents one standard deviation about the *in situ* averages. The RPD reported in the text was calculated as $100\% \cdot (\mathrm{median}(C_a^{satellite}/C_a^{in\ situ}) - 1)$ using each monthly satellite and *in situ* pair.

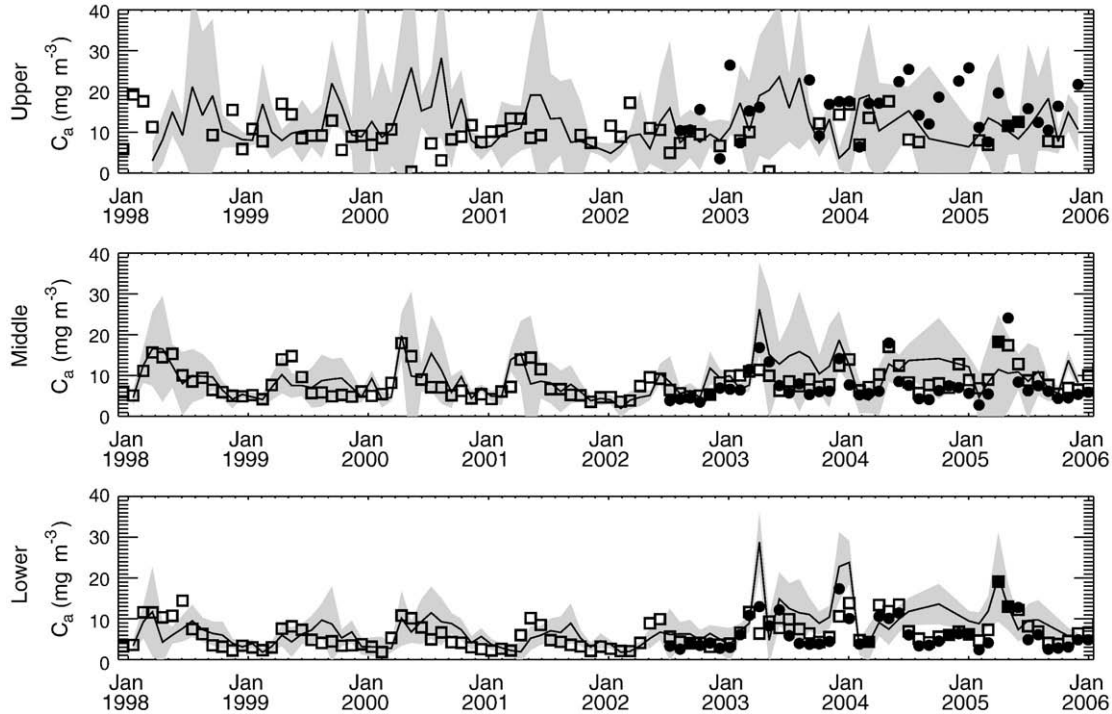

Fig. 6. Monthly averages of *in situ* C_a (thin lines) in the Upper, Middle, and Lower Bays compared to GSM retrievals for SeaWiFS (empty squares) and MODIS-Aqua (filled circles). The grey shaded area represents on standard deviation about the *in situ* averages.

Consistent with results in sections 3.1 and 3.2, the GSM sample sizes for both sensors' time-series where retrieved C_a fell between 0.001 and 100 mg m^{-3} trailed those for the OC algorithms by ~50%. The persistent differences in sample sizes for the algorithms will be reviewed in section 4.1.

SeaWiFS OC4 retrievals consistently overestimated *in situ* C_a by 30 and 48% for the Lower and Middle Bays, respectively, in dry years (Fig. 5). These RPD fell to 11 and 23% in wet years. Biases in Upper Bay retrievals were less responsive to inflow variability, averaging 39% in dry years and 29% in wet years. Analyses of streamflow effects were not performed for MODIS-Aqua, which began data collection in mid-June 2002 and has observed predominantly wet conditions. MODIS-Aqua OC3retrievals overestimated *in situ* C_a by 20, 51, and 97% in the Lower, Middle, and Upper Bays, respectively. Over comparable time spans, the MODIS-Aqua RPD compare well with those for SeaWiFS in the Lower Bay, reasonably well in the Middle Bay (possibly degrading because of several outliers in 2004), and poorly in the Upper Bay. The monthly MODIS-Aqua C_a from OC3 show little seasonality (not unlike SeaWiFS) in the Upper Bay, with values exceeding the *in situ* measurements twofold.

Unlike the consistent positive bias of the OC4-derived C_a, SeaWiFS GSM retrievals overestimated *in situ* C_a in the winter–spring, but underestimated *in situ* C_a in the summer–fall (Fig. 6). For our full study period, RPD increased from −36 and −21% during summer–fall to 16 and 17% during winter–spring in the Lower and Middle Bays, respectively. Further, C_a retrievals from GSM responded marginally to increased streamflow, with RPD shifting from −15 and −6% for dry conditions to −20 and −1% for wet conditions in the Lower and Middle Bays, respectively. Clearly, negative biases that emerged between each winter–spring bloom controlled the magnitudes of the long-term offsets. As was the case for the OC retrievals, reasonable C_a were occasionally achieved in the Upper Bay, but captured little seasonal variability. MODIS-Aqua failed to improve upon the long-term biases of SeaWiFS for similar time ranges, and reported larger seasonal amplitudes (i.e., greater overestimates of *in situ* C_a in winter–spring and increased underestimates in summer–fall). For the Lower and Middle Bays, MODIS-Aqua reported long-term RPD of −30 and −32%, but summer–fall and winter–spring RPD of −48 and −38% and −2 and −17%, respectively. MODIS-Aqua produced highly variable GSM-derived C_a in the Upper Bay (Fig. 6).

4. Discussion

4.1. Algorithm performance

The Level-2 match-ups, frequency distributions, and time-series convey qualitative and quantitative information in different ways, both visually and statistically. Summarizing section 3, OC-derived C_a from SeaWiFS overestimated *in situ* observations on the order of 10, 30, and 50% in the Lower, Middle, and Upper Bays, respectively, although these statistics varied strongly by season. Likewise, OC-derived C_a from MODIS-Aqua generally exceeded *in situ* observations by 40, 70, and 100% in these regions. GSM-derived C_a from SeaWiFS matched *in situ* observations within ±25% in the Lower and Middle Bays, although these statistics also varied strongly by season. GSM-derived C_a from MODIS-Aqua largely fell within the same range, but showed greater negative biases. For both sensors, broad dynamic ranges confounded the interpretation of GSM retrievals of C_a in the Upper Bay. With regard to water quality monitoring, OC algorithms returned C_a with realistic variability, as indicated by the dynamic range of the retrievals, but with spatially dependent, positive biases. The effectiveness of GSM for water monitoring is more difficult to assess given its increased variability. Both algorithms showed bias shifts as the seasons progressed, with zonal dependence in the direction of the shifts, which indicates a temporally-dependent ability to capture seasonal phenomena (Table 2). Interestingly, lowest RPD from both algorithms were often accompanied by the highest SPD, emphasizing the need to consider both statistics.

OC-derived C_a consistently exceeded *in situ* observations, but were constrained to a narrow dynamic range, and thereby showed less variability in magnitude compared to the GSM retrievals. Strong positive biases were evident in the Upper Bay and lower C_a values (<4 mg m^{-3}) were under-represented compared to extensive *in situ* observations. We largely attribute the improved agreement with ground truth for years with increased streamflow to the increased

phytoplankton biomass in the wet years. This increased biomass pushed C_a into the range where the OC algorithms performed best in the Bay, namely $\sim 5 \leq C_a \leq 20$ mg m^{-3}. The OC algorithms showed the least bias in the summer, possibly owing to 45% of Chesapeake Bay stations in NOMAD (the tuning dataset) having been visited between June and August, although they also report the largest SPD in this season.

GSM-derived C_a showed less bias relative to OC values in the Lower and Middle Bays, but increased overall variability, as indicated by scatter in the Level-2 match-ups (Fig. 3), widths of the data distributions (Fig. 4), and seasonal amplitudes of the time-series (Fig. 6). The similar C_a magnitudes retrieved in dry and wet conditions suggest GSM to be insensitive to the effects of changes in streamflow. For example, GSM moderately underestimated C_a in dry summers, but significantly underestimated C_a in wet summers, indicating imperfect algorithm parameterization for higher inflow conditions. The assigned $a^*_\phi(\lambda)$ and S, for example, may not have perfectly captured the interannual variabilities of the annual phytoplankton cycle (Adolf et al., 2006; Marshall et al., 2006) and inflow of riverine humus and sediments, respectively. Magnuson et al. (2004) developed their regional GSM parameterization using data collected between 1996 and 2002, encompassing both dry and wet conditions. In general, the wide C_a distributions from GSM illustrate the challenge of algorithm parameterization in highly productive and stratified waters. In all three Bay regions, GSM retrievals overestimated C_a in the spring and underestimated C_a in the summer. GSM retrievals were least biased in the fall and winter, but had elevated SPD in these seasons (common S and $a^*_\phi(\lambda)$ were used for both, as Magnuson et al. (2004) did not collect winter data).

The empirical algorithms returned a greater number of valid Level-2 pixels than the semi-analytical approach. Recall, we considered only retrievals within $0.001 < C_a \leq 100$ mg m^{-3} to be valid, and clouds equally confounded both algorithmic forms. GSM and other "spectral-matching" algorithms fail in the presence of negative $L_w(\lambda)$, which commonly occur in the presence of absorbing aerosols that cannot be detected by the atmospheric correction process (Werdell et al., 2007). This misinterpretation of aerosol type over the Bay leads to underestimates of blue L_w, where the signals are already low because of high a_{dg} (and a_ϕ to a lesser degree) in that spectral region (Holben et al., 2001). Note, low red L_w signals cause analogous inversion problems in the open ocean. With regard to Level-2 spatial coverage in the Bay, OC algorithms outperformed GSM as they require only wavelengths between ~490 and 555 nm to return reasonable C_a. Note, however, that the spatial and temporal aggregation of the Level-2 data into Level-3 monthly averages equalizes the spatial coverage of both algorithm approaches. For a case study in April 2006, Werdell et al. (2007) reported the percentage of valid C_a retrieved in the Bay increased from 45% at Level-2 to 65 and 75% at Level-3 at 2 and 4 km spatial resolution, respectively.

4.2. Satellite performance

SeaWiFS-derived C_a for both algorithm forms most closely matched the *in situ* measurements. In general, the magnitude of MODIS-Aqua retrievals exceeded those from SeaWiFS (more so with increasing latitude) for the years 2002 to 2006. RPD of the year-long data distributions showed MODIS-Aqua C_a from the OC algorithms were 45, 50, and 55% higher than those for SeaWiFS in the Lower, Middle, and Upper Bays, respectively, although SPD differed by only ±15% on average. The time-series biases were less severe, reporting C_a increases of only 15, 19, and 43% in the three regions (also calculated relative to SeaWiFS). GSM retrievals for MODIS-Aqua more closely matched those from SeaWiFS, with the exception of the Upper Bay (Fig. 4.). RPD of the year-long data distributions showed GSM-derived C_a from MODIS-Aqua fell below those from SeaWiFS by only −3 and −8% in the Lower and Middle Bays, but exceeded SeaWiFS by 116% in the Upper Bay, with SPD differing by 17% on average. The time-series biases were more consistent, reporting values of −20% for all three regions.

Differences in C_a from SeaWiFS and MODIS-Aqua trace to differences in their respective radiometry and to differences in the satellite-specific bio-optical algorithms. In effect, we assessed the tolerance of each C_a algorithm to regional errors in $L_w(\lambda)$ and evaluated the full sensor system (instrument calibration plus atmospheric correction algorithms plus bio-optical algorithms). We explore differences in SeaWiFS and MODIS-Aqua radiometry and algorithms in this section and the subsequent section, respectively. SeaWiFS and MODIS-Aqua radiances maintain long-term relative biases, the magnitudes of which vary by wavelength, season, and trophic level. For example, while radiative transfer theory and field observations suggest that $L_w(551)$ exceeds $L_w(555)$ by several percent in most waters, the long-term ratio of MODIS-Aqua $L_w(551)$ to SeaWiFS $L_w(555)$ is ~0.95 (Franz et al., 2005). Reducing the R_{rs} maximum band ratio used as input into OC3 by 10% (if, e.g., MODIS-Aqua $L_w(551)$ are 10% low) raises the derived C_a by ~35% in eutrophic waters (Fig. 1). A spectral-matching algorithm requires more accurate input $R_{rs}(\lambda)$ than its empirical counterparts to retrieve reasonable C_a. We have focused only on errors associated with GSM parameterization and difficulties with spatial and temporal transitions within this parameterization. The sensitivity of C_a from GSM to errors in the $R_{rs}(\lambda)$ used as input has yet to be explored for Chesapeake Bay. A full discussion of SeaWiFS and MODIS-Aqua radiometric differences exceeds the scope of this work; we simply reiterate that differences in radiometry lead to significant differences in derived C_a.

The evaluation of satellite-derived $R_{rs}(\lambda)$ in the Bay would provide an excellent precursor to this paper, but the scarcity of high-quality radiometry precludes the execution of similarly robust spatial and temporal studies (radiometric measurements are not made as part of the CBP Water Quality Monitoring program). Harding et al. (2005), Signorini et al. (2005), and Werdell et al. (2007) reported acceptable $L_w(\lambda)$ Level-2 match-ups for SeaWiFS in the lower half of the Bay using smaller data sets. To our knowledge, similar analyses have yet to be conducted for MODIS-Aqua. To first order, C_a retrievals depend largely on the efficacy of the atmospheric correction process, which may be problematic in coastal environments. For example, the corrections for near-infrared water-leaving radiances, bi-directional reflectance, and spectral band-pass effects all rely on bio-optical models that can perform poorly in optically complex waters and do not permanently represent all geographic zones (Morel et al., 2002; Patt et al., 2003). Adjacency effects on the satellite radiometry become prominent where the Bay narrows, requiring an additional correction (applied in this analysis) to remove "stray" light reflected by neighboring land (Barnes et al., 1995; Santer & Schmechtig, 2000). While recent studies suggested improved methods for estimating aerosol radiances over coastal or turbid waters, few have been validated comprehensively against ground truth (Ahmad et al., 2007; Kuchinke et al., 2009; Wang & Shi, 2005). Conceivably, the use of regionally specific aerosol models will improve $R_{rs}(\lambda)$ retrievals if, for example, they can be developed using *in situ* measurements of atmospheric properties over the Bay (Holben et al., 2001).

4.3. Spectral considerations

The two OC algorithms, OC3 and OC4, originated from a common *in situ* data set, but inherent differences remain, particularly at high C_a (Morel et al., 2007). When C_a rises above ~2 mg m^{-3}, λ_b sequences from 490 to 510 nm for OC4, but remains at 490 nm for OC3, resulting in a different, less variable, data series against which to regress C_a (Fig. 1). To quantify the differences in performance of OC3 and OC4 in the absence of satellite-to-satellite radiometric biases, we reprocessed the full SeaWiFS time-series using a SeaWiFS-specific version of OC3 (Table 1). The full-year SeaWiFS C_a from OC3 (Fig. 7) behaved similarly to those for MODIS-Aqua (Fig. 4), with biases increasing

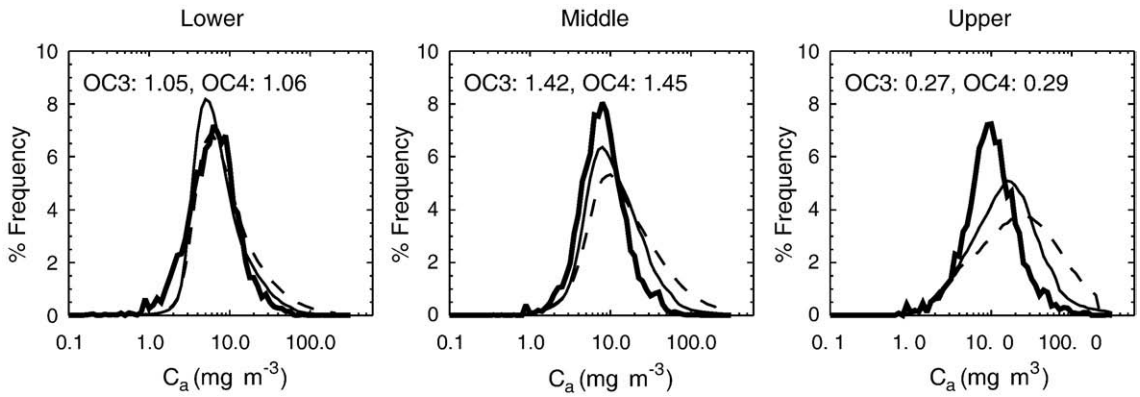

Fig. 7. *In situ* C_a distributions (thick lines) in the Lower, Middle, and Upper Bays compared to SeaWiFS retrievals from OC4 (thin solid) and OC3 (thin dashed) algorithms. Samples sizes (in million pixels) for satellite retrievals are provided in each panel. *in situ* sample sizes are as reported for Fig. 4. Data from all four seasons are included.

strongly with increasing latitude. RPD of the frequency distributions showed C_a from OC3 were 19, 25, and 37% higher than those for OC4 in the Lower, Middle, and Upper Bays, respectively, with SPD increasing accordingly by 20, 26, and 48%. Similarly, biases in the C_a time-series (data not shown) increased by ~20% for all three zones. For comparison, the SeaWiFS OC3 time-series showed little bias relative to the MODIS-Aqua time-series, reporting differences of only −7, −3, and 19% for the Lower, Middle, and Upper Bays, respectively. The latter suggests that a sizeable portion of the SeaWiFS and MODIS-Aqua differences stem from algorithm differences. OC4 outperformed OC3 in the mainstem Bay, the difference being the inclusion of a useful "green" band (510 nm). The proximity of the 531 nm channel on MODIS-Aqua to its 551 nm channel precludes its viable use in an OC algorithm as the two overly covary.

Given that negative $R_{rs}(\lambda)$ prevent GSM from returning reasonable C_a, we further examined whether ignoring portions of the spectra increased the volume of retrievals without sacrificing overall data quality. Using Level-2 match-up data, were calculated SeaWiFS C_a for three reduced-wavelength scenarios (five input $R_{rs}(\lambda)$ rather than six) for comparison with the standard results in section 3.1. We sequentially removed one blue (412 nm), green (510 nm), and red (670 nm) band to evaluate the importance of each spectral region. Spectral-matching algorithms require the assignment of a unique spectral shape to each water column constituent of interest (Eqs. (2)–(4)). Our interpretation of these match-ups, therefore, relies largely on differences in spectral shapes of $a_{dg}(\lambda)$ and $a_\phi(\lambda)$. We adopted spectrally flat $b_{bp}(\lambda)$ from Magnuson et al. (2004), which minimizes its role in this analysis. From 400 to 700 nm, the predominant differences in $a_{dg}(\lambda)$ and $a_\phi(\lambda)$ are depressed algal absorption from 400 to 440 nm and enhanced algal absorption from 600 to 700 nm. $R_{rs}(412)$ and $R_{rs}(670)$ are critical for distinguishing between these types of absorbing material. From 490 to 550 nm, the $a_\phi(\lambda)$ proposed by Magnuson et al. (2004) are typically less concave than $a_{dg}(\lambda)$ and have increased spectral slopes. Note that minor absorption peaks in this limited spectral range are not captured by the SeaWiFS spectral suite.

Reducing the number of input $R_{rs}(\lambda)$ for GSM did not sufficiently improve its spatial coverage to warrant the accompanying increase in C_a variability. The loss of any band confounded the ability of GSM to distinguish C_a from NAP + CDOM (Fig. 8). This was most obvious when we ignored $R_{rs}(412)$ and $R_{rs}(510)$, as the sample sizes (N) were unchanged relative to the full-spectrum results, but positive biases appeared (ratios of 1.39 and 1.22, respectively) and APD increased significantly (~50 and 20%). Note the expanded dynamic range of GSM-derived C_a for both cases. GSM ineffectively distinguished between C_a and NAP + CDOM with $R_{rs}(412)$ absent as the inversion no longer made use of the unique spectral differences between $a_{dg}(412)$ and $a_\phi(412)$. The match-up variability increased, but the retrieved C_a remained in a reasonable range (0.001 to 100 mg m^{-3}) because of the dominant contribution of $a_\phi(670)$ to $R_{rs}(670)$, which effectively constrained the inversion system. An additional 14 stations returned valid C_a, however, the median satellite-to-*in situ* ratio for only those stations with negative $R_{rs}(412)$ was 3.7. GSM also imperfectly distinguished between C_a and NAP + CDOM with $R_{rs}(510)$ absent as the inversion incorrectly assumed a linear slope between $R_{rs}(490)$ and $R_{rs}(555)$. This portion of the visible spectrum is typically concave, and the magnitude of the inward curve is determined by the relative contributions of C_a and NAP + CDOM. In

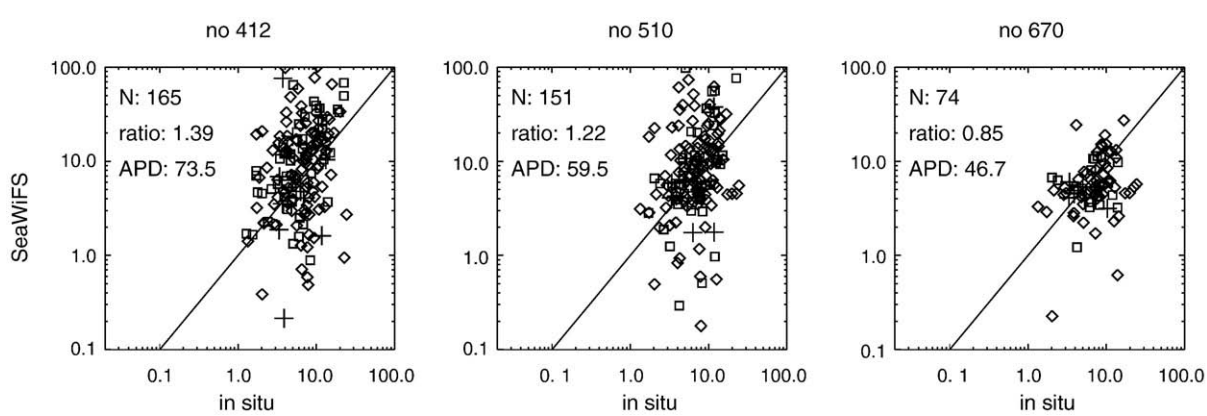

Fig. 8. SeaWiFS satellite-to-*in situ* C_a validation results for the GSM algorithm for three varied $R_{rs}(\lambda)$ spectral resolutions. From left to right, the wavelength suites used are 443, 490, 510, 555, and 670 nm; 412, 443, 490, 555, and 670 nm; and 412, 443, 490, 510, and 555 nm. Definitions for the validation statistics and symbols are as in Fig. 3. The solid line demarks a 1:1 relationship.

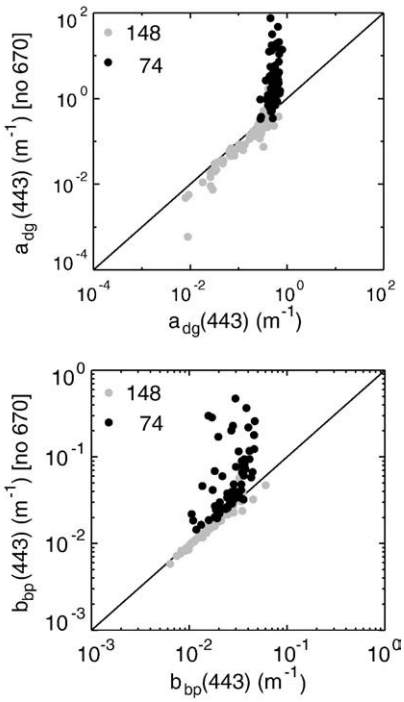

Fig. 9. Comparison of SeaWiFS GSM-derived $a_{dg}(443)$ and $b_{bp}(443)$ calculated using five and six visible $R_{rs}(\lambda)$ as input (412, 443, 490, 510, and 555 nm versus 412, 443, 490, 510, 555, and 670 nm, respectively). Gray circles indicate retrievals where derived C_a fall between 0.001 and 100 mg m^{-3}. Black circles indicate retrievals where derived C_a fall below 0.001 mg m^{-3}. Sample sizes for each are reported.

contrast, GSM produced favorable results with $R_{rs}(670)$ absent, yet N was reduced by half. With the red portion of the spectrum now unconstrained (i.e., the inversion no longer exploits the dominant contribution of $a_{\phi}(670)$ to $R_{rs}(670)$), the magnitude of the retrieved a_{dg} (443) increased to unreasonable levels (Fig. 9). The retrieved C_a fell well below 0.001 mg m^{-3} when GSM returned $a_{dg}(443)$ in excess of ~0.5 m^{-1}, resulting in a strongly bimodal distribution. Further analysis of the choice of inversion method (e.g., Levenberg–Marquardt) is beyond the scope of this work.

4.4. Regional algorithms

OC-derived C_a from SeaWiFS and MODIS-Aqua showed positive, latitudinal biases relative to *in situ* measurements and may not yet be sufficiently accurate for Bay water quality monitoring. Recall, however, we used globally-parameterized OC coefficients (Table 1), whereas several options for regional tuning remain to be exploited. Regression coefficients derived using a Bay-only data set provide the most direct approach. For example, Werdell et al. (2007) presented a Bay-specific version of OC3 and reported MODIS-Aqua Level-2 match-up APD and median satellite-to-*in situ* ratios decreased from 70% and 1.70 to 35% and 1.17, respectively. Given that several other research groups have also used such an approach to retrieve reasonable C_a (e.g., Old Dominion University and NOAA CoastWatch–East Coast Node), we explored two alternative mechanisms for regional OC algorithm tuning.

Hyde et al. (2007) proposed an empirical correction for SeaWiFS OC retrievals in Massachusetts Bay based on type II linear regression (their Eq. (3)):

$$OC_{corr} = 10^{\wedge}\left(\frac{log_{10}(OC) - \alpha}{\beta}\right) \quad (5)$$

where α and β are the regression slope and intercept, respectively. We explored the effectiveness of such an approach for Chesapeake Bay using Level-2 match-up data (Fig. 3). The resulting α and β were 0.249 and 0.844 for SeaWiFS and 0.264 and 0.960 for MODIS-Aqua. Interestingly, the SeaWiFS values closely resemble the 0.248 and 0.832 reported by Hyde et al. (2007) for Massachusetts Bay. RPD of the year-long data distributions showed SeaWiFS C_a from OC_{corr} reduced to −1and 17% in the Middle and Upper Bays, but increased to −22% in the Lower Bay (Fig. 10). For MODIS-Aqua, RPD reduced to −4, 26, and

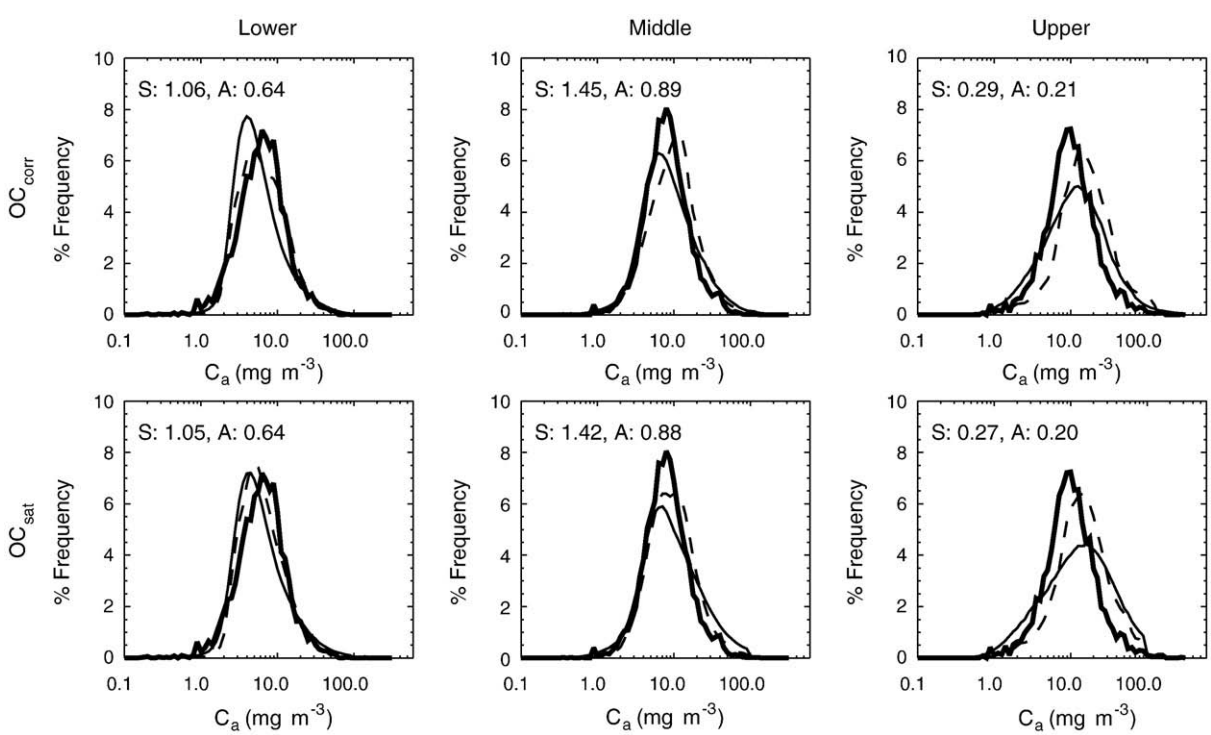

Fig. 10. *In situ* C_a distributions (thick lines) in the Lower, Middle, and Upper Bays compared to OC_{corr} and OC_{sat} retrievals for SeaWiFS (thin solid) and MODIS-Aqua (thin dashed). Samples sizes (in million pixels) for satellite retrievals are provided in each panel, with SeaWiFS indicated by *S* and MODIS-Aqua indicated by *A*. *In situ* sample sizes are as reported for Fig. 4. Data from all four seasons are included.

52% in the Lower, Middle, and Upper Bays, respectively. For both sensors, the widths of the OC_{corr} distributions matched those of the *in situ* observations almost exactly. The time-series biases had similar trends, with values of −9, 11, and 10% in the three zones for SeaWiFS and −20, 6, and 38% for MODIS-Aqua (data not shown). Level-2 match-up results for OC_{corr} did not provide independent verification of this approach, however, as we used these data in the derivation of OC_{corr}.

Next, we derived a series of regression coefficients (OC_{sat}) using the *in situ* C_a and satellite $R_{rs}(\lambda)$ from Level-2 match-up data. Using satellite radiometry in lieu of *in situ* measurements provides some advantage, as the bio-optical algorithm now directly addresses biases from the atmospheric correction process. Given the limited MODIS-Aqua sample size, we combined data from the two sensors and focused solely on OC3. The limited dynamic range of $\log_{10}(C_a)$ in this data set led us to abandon the polynomial expansion in favor of type II linear regression to derive an intercept (c_0 in Eq. (1)) and slope (c_1) of 0.0667 and −2.874, respectively. RPD of the year-long data distributions showed SeaWiFS C_a from OC_{sat} reduced to 6 and 30% in the Middle and Upper Bays, but increased to −17% in the Lower Bay (Fig. 10). MODIS-Aqua RPD reduced to −8, 16, and 44% in the Lower, Middle, and Upper Bays, respectively. Again, the widths of the OC_{sat} distributions for both sensors matched those of the *in situ* observations almost exactly. The time-series biases had similar trends with values of −6, 13, and 16% in the three zones for SeaWiFS and −22, 2, and 37% for MODIS-Aqua (data not shown). As we observed for OC_{corr}, the Level-2 match-up results for OC_{sat} did not provide independent verification of this approach, as we used these data to develop OC_{sat}.

Both OC_{corr} and OC_{sat} showed improved C_a retrievals compared to their global counterparts (Fig. 10). For both sensors, C_a retrievals closely matched *in situ* measurements in the Middle Bay. Some positive biases remained in the Upper Bay retrievals, but RPD of the year-long distributions were reduced to one-third or one-half of their original values. In the Lower Bay, RPD for OC_{corr} and OC_{sat} showed slight negative biases in both the time-series (more so for MODIS-Aqua)and data distributions (more so for SeaWiFS), although both had identical dynamic ranges with *in situ* C_a. SeaWiFS retrievals from both regional algorithms underestimated C_a in the Lower Bay more severely during wet years, possibly owing to 81% of the SeaWiFS Level-2 match-up stations having been visited pre-2003 (biases of −2 and −14% for dry and wet years for OC_{corr} and 6 and −17% for dry and wet years for OC_{sat}). Conceivably, differences in satellite and *in situ* sampling and sample sizes contribute most significantly to these minor biases, rather than the bio-optical algorithms and satellite radiometry used as input. Wherever variabilities of OC_{corr} and OC_{sat} did not improve upon those for the global algorithms, the removal of biases in the C_a retrievals argues in favor of such approaches for satellite-based Bay water quality monitoring. We expect the results to improve further with additional algorithm refinements as new data become available to support seasonally-dependent versions of OC_{corr} and OC_{sat}.

Alternative algorithmic approaches exist with the potential to improve satellite C_a retrievals in the Bay. For example, an emerging body of work encourages R_{rs} band ratio algorithms to include red and near-infrared wavelengths (Gitelson et al., 2007; Tzortziou et al. 2007). Our understanding of the behavior of marine absorption and backscattering properties within the Bay has also increased (Gallegos et al., 2005; Tzortziou et al., 2006; Zawada et al., 2007). In the same manner, the bio-optical data described by Magnuson et al. (2004) have expanded sufficiently to warrant the reparameterization of GSM. This might include novel mechanisms for transitioning between spatial and temporal boundaries and accounting for varied streamflow. Naturally, the latter would require ancillary geophysical inputs or be more suitable for retrospective processing, as operational ocean color approaches currently lack predictive capabilities. Mechanisms for transitioning between trophic levels or bioregimes are emerging, which show promise for improved bio-optical model parameterization if they can be developed on regional levels (Devred et al., 2007; Moore et al., 2001). Improved spatial resolution is also desirable, and the land and atmosphere channels on MODIS-Aqua provide a possible avenue for decreasing the satellite footprint (Franz et al., 2006). Given the spectral limitations of the higher resolution visible bands (469 and 555 nm at 500 m resolution and 645 nm at 250 m resolution) and their reduced signal-to-noise ratios, sacrifices in algorithm quality accompany their use (although highly scattering coastal waters often have increased radiometric signals relative to the open ocean in green and red spectral bands).

5. Conclusions

We have presented the regional and seasonal variability of C_a from SeaWiFS and MODIS-Aqua in the Chesapeake Bay. Specifically, we compared satellite retrievals from two common approaches with *in situ* measurements to establish baseline quantification of algorithm and sensor performance against which novel approaches might be compared. C_a from SeaWiFS and MODIS-Aqua were reasonable in the Lower and Middle Bays, but the accuracy of the retrievals degraded with increasing latitude as the marine optical complexity increased northward. Retrieving high-quality satellite radiances in the Upper Bay was increasingly difficult because of the narrowing shape of the estuary, the increased turbidity, and the proximity to sources of urban aerosols that confounded the atmospheric correction process. We found advantages of the OC algorithms to include superior spatial coverage and decreased sensitivity to errors in the input radiometry. Advantages of the GSM algorithm as parameterized by Magnuson et al. (2004) include its ability to simultaneously retrieve additional marine optical properties and its spatially and temporally transitioning parameterization. Overall, the regionally-tuned OC_{corr} and OC_{sat} algorithms performed best throughout the mainstem Bay, offering the combination of reasonable accuracy and high spatial coverage. While GSM produced good results under many conditions, it remained highly sensitive to input $R_{rs}(\lambda)$ and its internal parameterization (e.g., S, η, and $a^*_\phi(\lambda)$), departures from which (through the natural biogeochemical variability of the Bay) increased its overall variability. The OC algorithms most naturally represented "mean" Bay conditions and better constrained the retrieved C_a dynamic range. We demonstrated the value of "green" spectral resolution (e.g., 510 nm) for coastal ocean color applications, which provides a cautionary tale for future satellite instruments whose proposed data processing includes similar empirical or spectral-matching algorithms. Moreover, we found that reducing the spectral resolution of either algorithm was disadvantageous for water quality monitoring in the Chesapeake Bay.

The CBP Water Quality Monitoring Data set provided a unique opportunity to compare decade-long time-series of satellite and *in situ* C_a observations. While the designation of an uncertainty budget for SeaWiFS- and MODIS-Aqua-derived C_a would provide a desirable product of this paper, we were hesitant to assign a restrictive one, given the statistical incompatibility of our varied analyses. The Level-2 match-ups, data distributions, and time-series all described departures of the satellite retrievals from ground truth estimates, but exact quantities and magnitudes varied based on the spatial and temporal binning applied within each analysis. Nevertheless, all three analyses converged on the direction of the differences (e.g., positive or negative biases) and their general magnitudes. The coverage needs and accuracy requirements of each data manager will dictate whether satellite-derived C_a are of sufficient quality for use in water quality monitoring of Chesapeake Bay. These obligations will further determine the optimal geophysical algorithms, processing approaches, and binning strategies (e.g., Level-2 or Level-3) to be applied. We propose that the spatially and temporally rich data streams from ocean color satellites have potential to improve both our scientific understanding of coastal processes and our ability to monitor Chesapeake Bay marine resources, provided their uncertainties can be rigorously quantified.

Acknowledgements

We thank Michael Mallonee, David Jasinski, and Mark Trice for the assistance with the CBP Water Quality Monitoring data. We also thank Shawna Karlson, Kevin Sellner, Chris Kinkade, Eric Stengel, Michael Ondrusek, Chris Brown, and our anonymous reviewers for the valuable comments at various stages of this project. Support for this work was provided through the NASA MODIS Science Team.

References

Acker, J. G., Harding, L. W., Jr., Leptoukh, G., Zhu, T., & Shen, S. (2005). Remotely-sensed chl *a* at the Chesapeake Bay mouth is correlated with annual freshwater flow to Chesapeake Bay. *Geophysical Research Letters*, *32*, L05601. doi:10.1029/2004GL021852

Adolf, J. E., Yeager, C. L., Miller, W. D., Mallonee, M. E., & Harding, L. W., Jr. (2006). Environmental forcing of phytoplankton floral composition, biomass, and primary productivity in Chesapeake Bay, USA. *Estuarine, Coastal and Shelf Science*, *67*, 108–122.

Ahmad, Z., McClain, C. R., Herman, J. R., Franz, B. A., Kwiatkowska, E. J., Robinson, W. D., et al. (2007). Atmospheric correction for NO_2 absorption in retrieving water-leaving reflectances from the SeaWiFS and MODIS measurements. *Applied Optics*, *46*, 6504–6512.

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

Barnes, R. A., Holmes, A. W., & Esaias, W. E. (1995). *Stray light in the SeaWiFS radiometer. NASA Tech. Memo. 104566, Vol. 31.* (pp. 76). Greenbelt: NASA Goddard Space Flight Center.

Brown, C., Huot, Y., Werdell, P. J., Gentili, B., & Claustre, H. (2008). The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. *Remote Sensing of Environment*, *112*, 4186–4203.

Chesapeake Bay Program (1993). Guide to using Chesapeake Bay program water quality monitoring data. *CBP/TRS 78/92* (pp. 127). Annapolis: Chesapeake Bay Program.

Defoin-Patel, M., & Chami, M. (2007). How ambiguous is the inverse problem of ocean color in coastal waters? *Journal of Geophysical Research*, *112*, C03004. doi:10.1029/2006JC003847

Devred, E., Sathyendranath, S., & Platt, T. (2007). Delineation of ecological provinces using ocean colour radiometry. *Marine Ecology. Progress Series*, *346*, 1–13.

Fisher, T. R., Hagy, J. D., Boynton, W. R., & Williams, M. R. (2006). Cultural eutrophication in the Choptank and Patuxent estuaries of Chesapeake Bay. *Limnology and Oceanography*, *51*, 435–447.

Fisher, T. R., Harding, L. W., Jr., Stanley, D. W., & Ward, L. G. (1988). Phytoplankton, nutrients, and turbidity in the Chesapeake, Delaware, and Hudson estuaries. *Estuarine, Coastal and Shelf Science*, *27*, 61–93.

Franz, B. A., Werdell, P. J., Meister, G., Bailey, S. W., Eplee, Jr., R. E., Feldman, G. C., et al. (2005). The continuity of ocean color measurements from SeaWiFS to MODIS. *Proceedings SPIE 5882*. doi:10.1117/12.620069

Franz, B. A., Werdell, P. J., Meister, G., Kwiatkowska, E. J., Bailey, S. W., Ahmad, Z., et al. (2006). MODIS land bands for ocean remote sensing applications. *Proceedings Ocean Optics, Vol. XVIII*. Canada: Montreal.

Gallegos, C. L., Jordon, T. E., Hines, A. H., & Weller, D. E. (2005). Temporal variability of optical properties in a shallow, eutrophic estuary: Seasonal and interannual variability. *Estuarine, Coastal and Shelf Science*, *64*, 156–170.

Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*, 464–472.

Gordon, H. R., Brown, O. B., & Jacobs, M. M. (1975). Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics*, *14*, 417–427.

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

Hagy, J. D., Boynton, W. R., Keefe, C. W., & Wood, K. V. (2004). Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries*, *27*, 634–658.

Harding, L. W., Jr., & Magnuson, A. (2003). Bio-optical and remote sensing observations in Chesapeake Bay. In G. S. Fargion & C. R. McClain (Eds.), *SIMBIOS Project 2003 Annual Report. NASA Tech. Memo. 212251* (pp. 84–97). Greenbelt: NASA Goddard Space Flight Center.

Harding, L. W., Jr., Magnuson, A., & Mallonee, M. E. (2005). SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic Bight. *Estuarine, Coastal and Shelf Science*, *62*, 75–94.

Harding, L. W., Jr., Mallonee, M. E., & Perry, E. S. (2002). Toward a predictive understanding of primary productivity in a temperate, partially stratified estuary. *Estuarine, Coastal and Shelf Science*, *55*, 437–463.

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. E., Slutsker, I., Abuhassan, N., et al. (2001). An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *Journal of Geophysical Research*, *106*, 12,067–12,097.

Hu, C., Carder, K. L., & Muller-Karger, F. E. (2001). How precise are SeaWiFS ocean color estimates? Implications of digitization–noise errors. *Remote Sensing of Environment*, *76*, 239–249.

Hyde, K. J. W., O'Reilly, J. E., & Oviatt, C. A. (2007). Validation of SeaWiFS chlorophyll a in Massachussetts Bay. *Continental Shelf Research*, *27*, 1677–1691.

Kemp, W. M., Boynton, W. R., Adolf, J. E., Boesch, D. F., Boicourt, W. C., Brush, G., et al. (2005). Eutrophication of Chesapeake Bay: Historical trends and ecological implications. *Marine Ecology. Progress Series*, *303*, 1–29.

Kuchinke, C. P., Gordon, H. R., Harding, L. W., Jr., & Voss, K. J. (2009). Spectral optimization for constituent retrieval in Case 2 waters II: Validation study in the Chesapeake Bay. *Remote Sensing of Environment*, *113*, 610–621.

Magnuson, A., Harding, L. W., Jr., Mallonee, M. E., & Adolf, J. (2004). Bio-optical model for Chesapeake Bay and the Middle Atlantic Bight. *Estuarine, Coastal and Shelf Science*, *61*, 403–424.

Malone, T. C. (1992). Effects of water column processes on dissolved oxygen: Nutrients, phytoplankton, and zooplankton. In D. Smith, M. Leffler, & G. Mackiernan (Eds.), *Oxygen dynamics in the Chesapeake Bay: A synthesis of research* (pp. 61–112). College Park: University of Maryland Sea Grant College Publications.

Malone, T. C., Conley, D. J., Fisher, T. R., Glibert, P. M., Harding, L. W., Jr., & Sellner, K. G. (1996). Scales of nutrient-limited phytoplankton productivity in Chesapeake Bay. *Estuaries*, *19*, 371–385.

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

Marshall, H. G., Lacouture, R. V., Buchanan, C., & Johnson, J. M. (2006). Phytoplankton assemblages associated with water quality and salinity regions in Chesapeake Bay, USA. *Estuarine, Coastal and Shelf Science*, *69*, 10–18.

McClain, C. R., Hooker, S. B., Feldman, G. C., & Bontempi, P. (2006). Satellite data for ocean biology, biogeochemistry, and climate research. *EOS Transactions AGU*, *87*, 337.

Miller, W. D., Kimmel, D. G., & Harding, L. W., Jr. (2006). Predicting spring discharge of the Susquehanna River from a winter synoptic climatology for the eastern United States. *Water Resources Research*, *42*, W05414. doi:10.1029/2005WR004270

Moore, T. S., Campbell, J. W., & Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions Geoscience and Remote Sensing*, *39*, 1764–1776.

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

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

O'Reilly, J.E., personal communication. Northeast Fisheries Center, Narragansett Laboratory, NOAA National Marine Fisheries Service, Narragansett, Rhode Island 02882.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, *103*, 24,937–24,953.

Patt, F. S., Barnes, R. A., Eplee, R. E., Jr., Franz, B. A., Robinson, W. D., Feldman, G. C., et al. (2003). Algorithm updates for the fourth SeaWiFS data reprocessing. *NASA Tech. Memo. 206892, Vol. 22*. (pp. 74). Greenbelt: NASA Goddard Space Flight Center.

Santer, R., & Schmechtig, C. (2000). Adjacency effects on water surfaces: Primary scattering approximation and sensitivity study. *Applied Optics*, *39*, 361–375.

Signorini, S. R., McClain, C. R., Mannino, A., & Bailey, S. W. (2005). Report on ocean color and carbon study for the South Atlantic Bight and Chesapeake Bay regions. *NASA Tech. Memo. 212787* (pp. 45). Greenbelt: NASA Goddard Space Flight Center.

Smith, V. H. (2006). Responses of estuarine and coastal marine phytoplankton to nitrogen and phosphorus enrichment. *Limnology and Oceanography*, *51*, 377–384.

Tzortziou, M., Herman, J. R., Gallegos, C. L., Neale, P. J., Subramaniam, A., Harding, L. W., Jr., et al. (2006). Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure. *Estuarine, Coastal and Shelf Science*, *68*, 348–362.

Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., & Harding, L. W., Jr. (2007). Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *72*, 16–32.

U.S. Geological Survey (2007). *Estimated streamflow entering Chesapeake Bay.* http://md.water.usgs.gov/monthly/bay.html

Wang, M., & Shi, W. (2005). Estimation of ocean contribution at the MODIS near-infrared wavelengths along the east coast of the U.S.: Two case studies. *Geophysical Research Letters*, *32*, L13606. doi:10.1029/2005GL022917

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

Werdell, P. J., Franz, B. A., Bailey, S. W., Harding, L. W., Jr., & Feldman, G. C. (2007). Approach for the long-term spatial and temporal evaluation of ocean color satellite data products in a coastal environment. *Proceedings SPIE 6680*. doi:10.1117/12.732489

Zawada, D. G., Hu, C., Clayton, T., Chen, Z., Brock, J. C., & Muller-Karger, F. E. (2007). Remote sensing of particle backscattering in Chesapeake Bay: A 6-year SeaWiFS retrospective view. *Estuarine, Coastal and Shelf Science*, *73*, 792–806.