Remote Sens. **2014**, *6*, 8524-8540; doi:10.3390/rs6098524

OPEN ACCESS

remote sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Evaluation of Satellite Retrievals of Ocean Chlorophyll-a in the California Current

Mati Kahru [1,*], Raphael M. Kudela [2], Clarissa R. Anderson [2], Marlenne Manzano-Sarabia [3] and B. Greg Mitchell [1]

[1] Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 95064, USA; E-Mail: gmitchell@ucsd.edu

[2] Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA 95064, USA; E-Mails: kudela@ucsc.edu (R.M.K.); clrander@ucsc.edu (C.R.A.)

[3] Facultad de Ciencias del Mar, Universidad Autónoma de Sinaloa, Mazatlán, Sinaloa, Mexico; E-Mail: mmanzano@uas.edu.mx

* Author to whom correspondence should be addressed; Mati Kahru; E-Mail:mkahru@ucsd.edu; Tel.: +1-858-534-8947; Fax: +1-858-822-0562.

Received: 14 July 2014; in revised form: 28 August 2014 / Accepted: 3 September 2014 / Published: 11 September 2014

Abstract: Retrievals of ocean surface chlorophyll-a concentration (Chla) by multiple ocean color satellite sensors (SeaWiFS, MODIS-Terra, MODIS-Aqua, MERIS, VIIRS) using standard algorithms were evaluated in the California Current using a large archive of *in situ* measurements. Over the full range of *in situ* Chla, all sensors produced a coefficient of determination (R^2) between 0.79 and 0.88 and a median absolute percent error (MdAPE) between 21% and 27%. However, at *in situ* Chla > 1 mg m^{-3}, only products from MERIS (both the ESA produced *algal_1* and NASA produced *chlor_a*) maintained reasonable accuracy (R^2 from 0.74 to 0.52 and MdAPE from 23% to 31%, respectively), while the other sensors had R^2 below 0.5 and MdAPE higher than 36%. We show that the low accuracy at medium and high Chla is caused by the poor retrieval of remote sensing reflectance.

Keywords: ocean color; phytoplankton; chlorophyll; California Current; remote sensing

1. Introduction

The California Current (CC) has been a test bed for ocean color algorithm development for decades (e.g., [1–5]), and standard NASA empirical ocean color algorithms [6,7] were originally parameterized with datasets with approximately a third of all *in situ* measurements of radiometry and chlorophyll-a concentration (Chla, mg m^{-3}) from the CC [8]. While the proportion of CC data is smaller in the current NOMAD (NASA bio-Optical Marine Algorithm Dataset) version 2 dataset that is used to derive the coefficients of standard ocean color (OC) Chla algorithms (version 6) (http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/), it is still significant, accounting for 11.7% of all observations and 10.7% of Chla values > 1 mg m^{-3}. There is therefore an assumption that standard ocean color algorithms work well in the CC and have no significant bias. Yet, despite progress with data reprocessing and vicarious calibration of satellite ocean color radiometry, match-ups of Chla with *in situ* data in the CC consistently show significant systematic error at medium and high Chla [5]. Here, we evaluate how well Chla is retrieved by standard algorithms of all major ocean color satellite sensors (SeaWiFS, MODIS-Terra, MODIS-Aqua, MERIS and VIIRS) in comparison to a large archive of *in situ* Chla measurements from the CC region. We also compare some algorithms that are specifically designed for high Chla waters. The response of the California Current ecosystems to global and regional forcings is an area of active study, and trends of increasing phytoplankton biomass have been detected based on both *in situ* [9,10] and satellite [5,11] data. However, satellite detection of trends may be impacted if the estimates are biased, potentially not detecting real trends or falsely identifying trends due to biased satellite retrievals.

2. Data and Methods

2.1. In Situ Chla Data

The sources of *in situ Chla* data are listed in Table 1. Over half of the *in situ* Chla data were collected by the California Cooperative Oceanic Fisheries Investigations (CalCOFI). Quarterly cruises have been conducted on a regular grid of stations as far as 600 km offshore by CalCOFI [12]. A related California Coastal Ecosystem-Long Term Ecological Research (CCE-LTER) program that carries out cruises with flexible ground coverage was the second largest data contributor. The total number of near-surface Chla samples that were used to validate satellite data from 1997–2013 was about 7500 (Table 1), *i.e.*, about 440 stations per year.

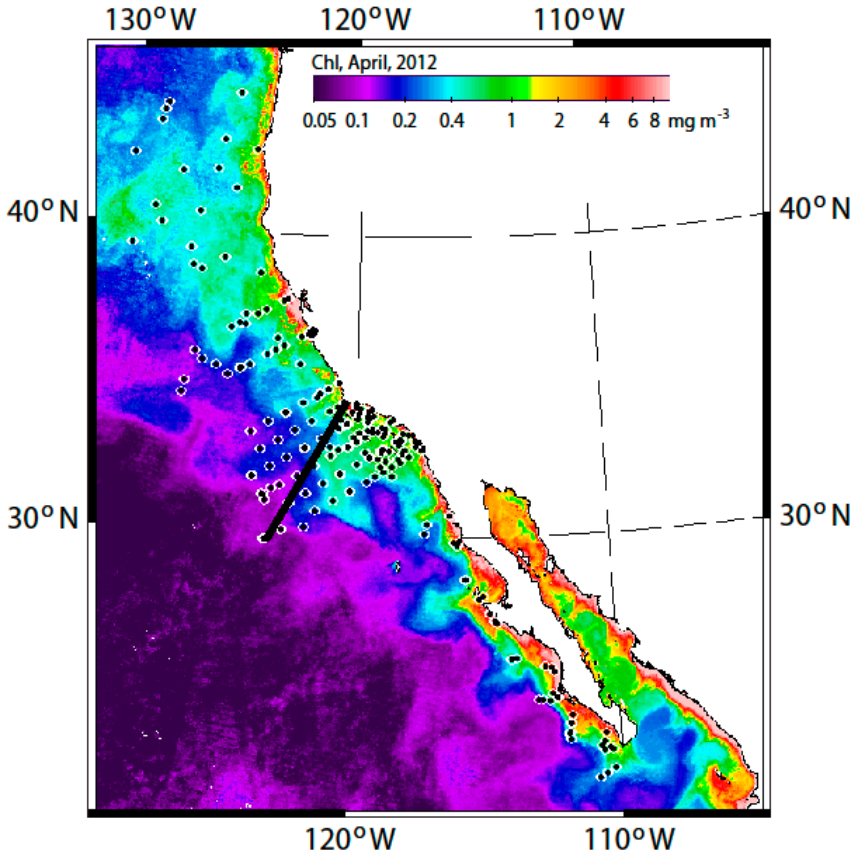
For each station, the sample nearest to the surface (typically 1–10 m) was used. This included only the high-quality datasets that were far enough from the coast to provide at least five valid satellite pixels in the 3 × 3-pixel window centered at the *in situ* sample location. Datasets acquired too near the coast were excluded, since they are typically affected by coastal runoff and river plumes, as well as by land adjacency effects. The excluded data include the optically complex (Case 2) waters in the Plumes and Blooms study in the Santa Barbara Channel [4] and several projects in the Monterey Bay (e.g. [13,14]). A few other datasets were excluded due to questionable calibration accuracy and sample collection procedures. Most of the *in situ* Chla (mg m^{-3}) samples were processed with the standard fluorometric method [15]. Fluorometric Chla values were replaced with total Chla when measurements with the more accurate HPLC method were available. A comparison of the Chla measurements made

with the fluorometric and HPLC methods showed very high correlation, with a slope close to one and an intercept not distinguishable from zero (e.g., Figure 5 in [8]). A typical spatial pattern for Chla in the CC region and the station locations for match-ups with MODIS-Aqua satellite measurements are shown in Figure 1.

Table 1. Sources of *in situ* surface Chla data and the corresponding number of stations.

Data Source	1997–2013
CalCOFI, California Cooperative Oceanic Fisheries Investigations (1997–2013)	4602
CCE-LTER, California Current Ecosystem Long-Term Ecological Research (2006–2013)	818
Oregon, California, Washington Line-transect and Ecosystem (ORCAWALE) Survey, NOAA Southwest Fisheries Science Center, 2008	405
Line 60, CalCOFI Line 60 by the Monterey Bay Aquarium Research Institute, 1997–2008	123
Line 67, CalCOFI Line 67 by the Monterey Bay Aquarium Research Institute, 1997–2009	556
Delphinus, NOAA SWFSC survey of the *Delphinus* species, 2009	465
CIMT, NOAA Center for Integrated Marine Technology, 2002–2006	385
PaCOOS, Pacific Coastal Ocean Observing System, 2005–2007	119
TOTAL	7473

Figure 1. Locations of the MODIS-Aqua Chla match-ups (black dots with white circles) within 3 h time difference overlaid on the April 2012, Chla composite. The black line shows the location of the 5 km-wide strip from the coast to offshore along which satellite-to-satellite match-ups were assembled.



2.2. Match-Ups between Satellite and In Situ Data

We used data from five different satellite sensors: SeaWiFS, MODIS-Terra (MODIST), MERIS, MODIS-Aqua (MODISA) and VIIRS. For MERIS, we used two different estimates of Chla from the same reduced resolution (RR) data: the standard ESA processed *algal_1* product (designated here as MERISRR) and the NASA processed *chlor_a* product (MERISNASARR). All satellite data were acquired at Level 2 (*i.e.*, processed to surface quantities, but unmapped) with approximately 1-km ground resolution.

SeaWiFS (1997–2010, version 2010.0), MODIST (2000–2014, version 2010.0) MODISA (2002–2014, version 2013.0), MERISNASARR (2002–2012, version 2012.1.1) and VIIRS processed by NASA (2012–2014, version 2013.1.1) were obtained from NASA's Ocean Color web (http://oceancolor.gsfc.nasa.gov/). Level-2 MERIS RR data processed by ESA (2002–2012, third reprocessing) were downloaded from ESA's MERIS Catalogue and Inventory (http://merci-srv.eo.esa.int/merci/welcome.do).

The validation of satellite products using quasi-simultaneous and spatially-collocated measurements (match-ups) of satellite and *in situ* data followed the general procedures of previous studies (e.g. [2,5,16,17]). For each Level-2 pixel, we used the corresponding Level-2 flags. For NASA processed data, the following flags made a pixel invalid: ATMFAIL, LAND, HISATZEN, CLDICE, CHLFAIL, SEAICE, NAVFAIL and HIPOL (see http://oceancolor.gsfc.nasa.gov/VALIDATION/flags.html for an explanation of the flags). In contrast with our previous study [5], the flag PRODFAIL, which indicated failure of any of the derived products, was not used, because it removes many valid chlorophyll retrievals due to failure of other products not relevant to the analysis, such as fluorescence line height (FLH). For MERISRR, the following flags made a pixel invalid: LOW_SUN, HIGH_GLINT, ICE_HAZE, SUSPECT, COASTLINE, PCD_19, PCD_18, PCD_17, PCD_16, PCD_15, PCD_14, PCD_1_13, CLOUD and LAND (https://earth.esa.int/instruments/meris/data-app/level2.html). All variables in Level-2 files were extracted from a 3 × 3-pixel window centered at the pixel nearest to the *in situ* sample. For statistical analysis, we accepted only those match-ups (at least five valid pixels (out of nine)). The maximum temporal difference between satellite and *in situ* measurements was set at three hours, but was relaxed to six hours for some tests with VIIRS in order to increase the number of match-ups. Satellite match-ups with a high range of variability within the 3 × 3-pixel window were excluded if $(Max - Min)/Min > 1$ for the standard Chla variable (*chlor_a* for NASA products or *algal_1* for ESA products). These match-ups were typically located near cloud edges and were deemed unreliable. The arithmetic mean Chla value of all valid pixels within the 3 × 3-pixel window was used as the satellite retrieval. The spatial distribution of MODISA match-ups with *in situ* measurements of Chla is shown in Figure 1.

2.3. Sensor Comparison

In order to evaluate the errors of the satellite remote sensing reflectance, $Rrs(\lambda)$, we compared *Rrs* data from spatially and temporally overlapping satellite sensors. We created daily satellite datasets for each sensor by mapping Level-2 *Rrs* to a standard map in Albers conic equal area projection, with each pixel being approximately 1 km^2. We then found match-ups between these mapped datasets for the

same pixel and the same day. Differences of a few hours in the timing between different satellite sensors are unavoidable, as the SeaWiFS overpass time nominally occurred at local noon, but drifted later towards the afternoon, while the MERIS overpass was approximately 10 AM, the MODIS-Terra overpass time approximately 10:30 AM, and the MODIS-Aqua and VIIRS overpasses at approximately 1:30 PM. As satellite sensors generate large amounts of overlapping data, we picked a 5 km-wide transect from the coast to offshore (Figure 1) that covered a range of environments from the high Chla coastal upwelling band to the oligotrophic low Chla offshore waters. Satellite-to-satellite match-ups were then picked along that transect. We chose the first 99 days of 2012 for a comparison of MODIST, MODISA, MERIS and VIIRS data. A comparison with SeaWiFS was performed during the first 99 days of 2004, as SeaWiFS data were not available in 2012. A comparison of sensors using a full year (not shown) provided similar results. The shorter, 99-day interval was chosen to simplify analysis and graphical representation of the results.

2.4. Statistical Estimates of Model Performance

We used several statistical measures to assess the performance of satellite products against *in situ* observations (satellite to *in situ* match-ups) and between different satellite sensors (*i.e.*, inter-sensor match-ups). For satellite to *in situ* match-ups, O_i is the i-th observation of an *in situ* variable and P_i is the corresponding predicted satellite variable. For sensor match-ups, the choice of the observed *versus* predicted variable is arbitrary, but we used MODISA as the common variable when comparing with other sensor values. We selected MODISA as the common sensor against which the other sensors were compared, as it is the only one to overlap temporally with all of the other sensors considered here, has a good calibration history and is currently operational. The coefficient of determination (R^2) on $\log_{10}$ transformed variables was used as a measure of covariance that captures the proportion of variance in one variable that can be predicted from another. As an estimate of scatter, we used the median absolute percentage error (MdAPE), which was calculated as MdAPE = 100 × median ($|(P_i - O_i)/O_i|$). For compatibility with some earlier studies, we also used the root mean square (RMS) error of $\log_{10}$ transformed variables. Log-transformation was needed, as the distribution of Chla is close to lognormal (e.g., Figure 2). As an estimate of general bias (e.g., too high or too low), we used the median relative percentage error (MRPE), which was calculated as MRPE = 100 × median ($(P_i - O_i)/O_i$). Both MdAPE and MRPE were calculated for P_i and O_i in natural (*i.e.*, not $\log_{10}$ transformed) units. We also include the slope of the reduced major axis (RMA) regression, calculated for $\log_{10}$ transformed variables.

3. Results

In order to find the distribution characteristics of both *in situ* and satellite Chla in the match-up datasets before eliminating any of the datasets, we constructed histograms of the match-up within five days. Because of the approximately lognormal distribution of Chla, both axes are logarithmic (Figure 2).

To a first approximation, the histograms of *in situ* and satellite data in over 4500 match-ups for MODIS-Aqua are quite similar. While some satellite retrievals were below the minimum measured *in situ* Chla (0.02 mg m^{-3}), their numbers were very low (note the logarithmic scale of Figure 2). A

more significant difference is present at the high Chla end. *In situ* match-up values peak at 47 mg m^{-3}, but MODIS-Aqua match-up values reach 249 mg m^{-3} (the mean of the valid pixels).

Satellite to *in situ* match-ups of Chla over the full range of *in situ* Chla (Figure 3) within a 3-h time difference show the highest coefficient of determination for the two MERIS products (R^2 = 0.88 for MERISRR and 0.82 for MERISNASARR) and the lowest for MODIST (R^2 = 0.79), VIIRS (R^2 = 0.80) and SeaWiFS (R^2 = 0.81) (Table 2 and Figure 4A).

Figure 2. Histograms of *in situ* Chla and MODISA-derived Chla in match-ups with up to a five-day time difference. The full range is 0.02–47 mg m^{-3} for *in situ* and 0.01–249 mg m^{-3} for the satellite retrievals (the mean of the valid pixels). The cumulative histograms are very close and practically overlap with each other.

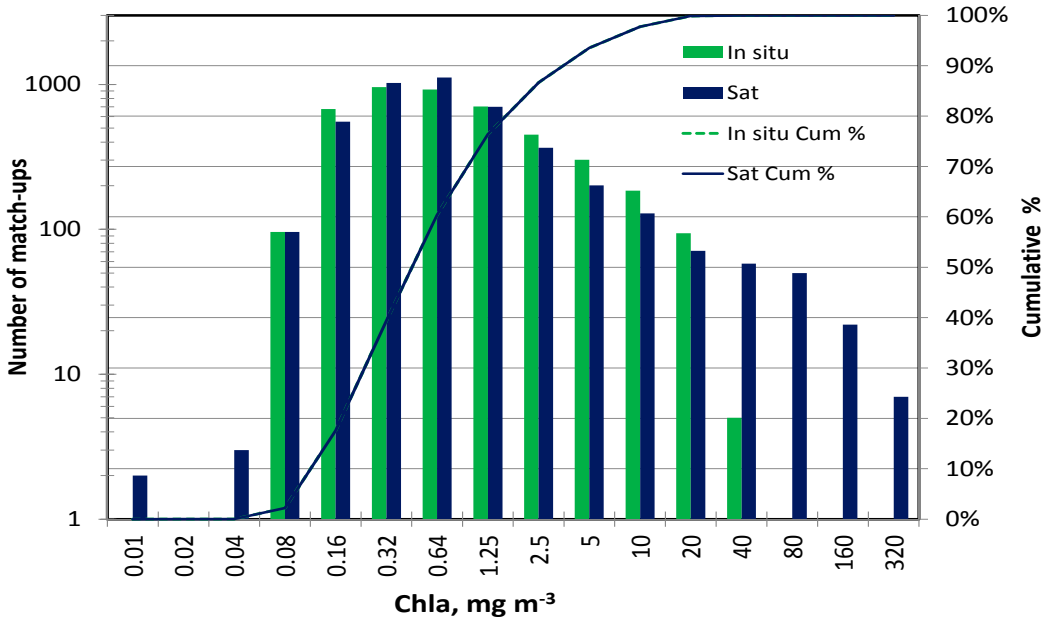


Table 2. Statistics for all match-ups with up to a 3-h time difference (also 6 h for VIIRS) and at least five valid pixels. N = number of match-ups, R^2 = coefficient of determination, MdAPE = median absolute percent error, Bias = median relative percent error, RMS = Root Mean Square error, RmaSlope = slope of the reduced major axis linear regression. MERISRR, standard ESA processed *algal_1* product; MERISNASARR, the NASA processed *chlor_a* product.

Sensor	N	R^2	MdAPE	Bias (MRPE)	RMS	RmaSlope
SeaWiFS	292	0.81	27.1	−4.9	0.21	0.81
MODIST	502	0.79	26.5	−11.9	0.24	0.83
MODISA	388	0.86	21.1	−1.7	0.18	0.86
MERISRR	167	0.88	24.2	8.9	0.15	0.93
MERISNASARR	205	0.82	22.7	−10.6	0.18	0.91
VIIRS	38	0.80	27.1	−4.8	0.23	0.70
VIIRS 6 h	81	0.81	29.0	−12.7	0.21	0.78

Figure 3. Chlorophyll-a match-ups with individual sensors/products using standard algorithms. The maximum allowed time difference is 3 h, and at least five pixels out of nine are required to be valid. The red line is the one-to-one line, and the blue line is the reduced major axis linear regression.

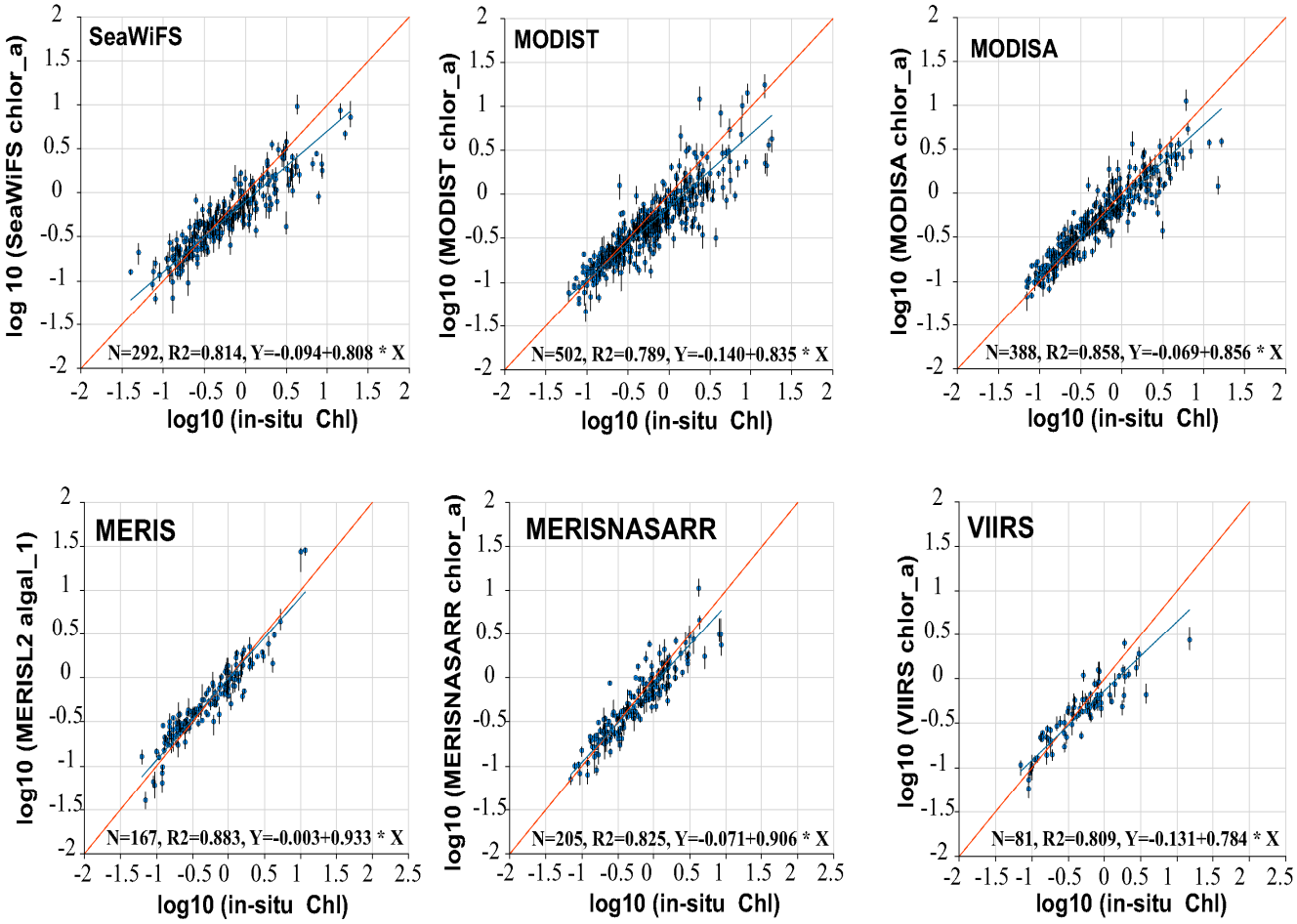
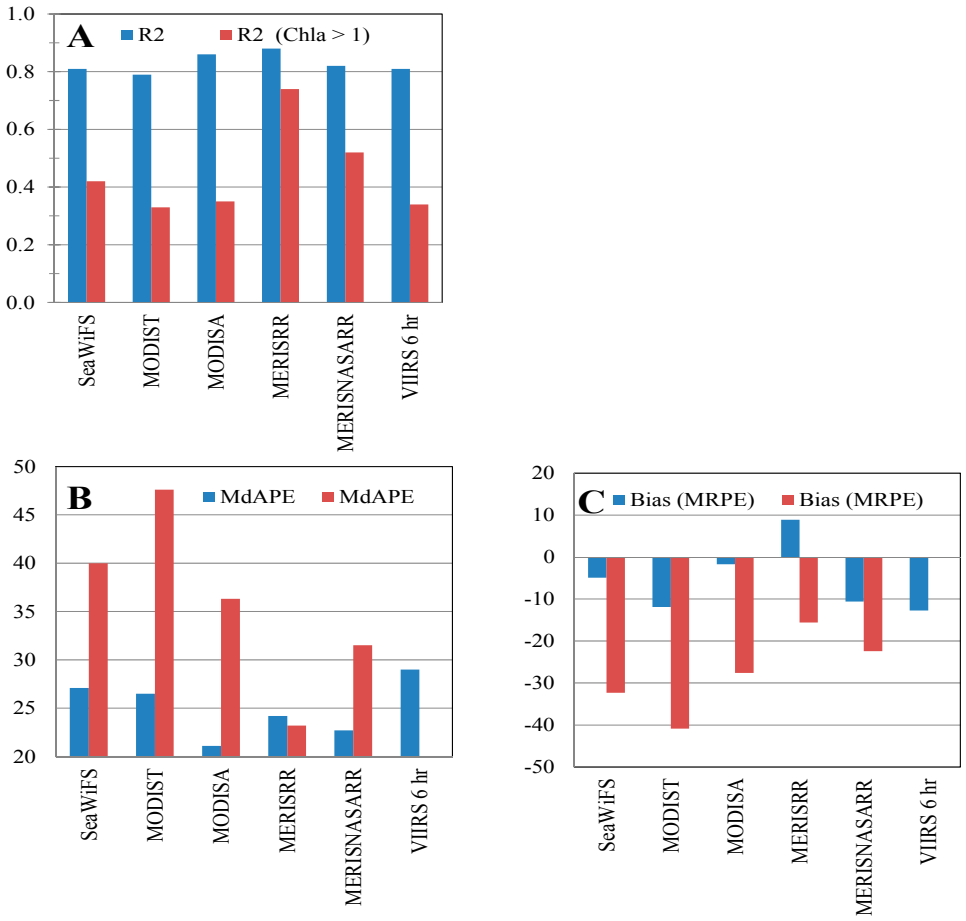


Table 3. Statistics for match-ups with *in situ* Chla ≥ 1 mg m^{-3} with up to a 3-h time difference (6 h for VIIRS) and at least five valid pixels. Column names are as in Table 2. The shaded rows are for nonstandard algorithms: RGCI, red-green chlorophyll index; NIR, using near-infrared bands. OC, ocean color.

Sensor	N	R^2	MdAPE	Bias (MRPE)	RMS	RmaSlope
SeaWiFS	73	0.42	40.0	−32.3	0.32	0.98
MODIST	128	0.33	47.6	−40.9	0.38	1.14
MODISA OC3	86	0.35	36.3	−27.6	0.29	0.96
MODISA RGCI	82	0.34	36.0	−14.3	0.24	0.66
MERISRR	39	0.74	23.2	−15.6	0.19	1.34
MERISRR NIR	33	0.08	118	118	0.62	2.2
MERISNASARR	58	0.52	31.5	−22.4	0.52	1.08
MERISNASARR NIR	47	0.11	271	271	0.72	2.4
VIIRS 6 h	14	0.34	46.2	−46.2	0.40	0.86

Figure 4. Statistics of chlorophyll-a retrievals by multiple satellite sensors for the full range of *in situ* Chla (blue) and for *in situ* Chla ≥ 1 mg m^{-3} (red): (**A**) coefficient of determination, R^2; (**B**) median absolute percent error (MdAPE, %); (**C**) bias or median relative percent error (MRPE, %).

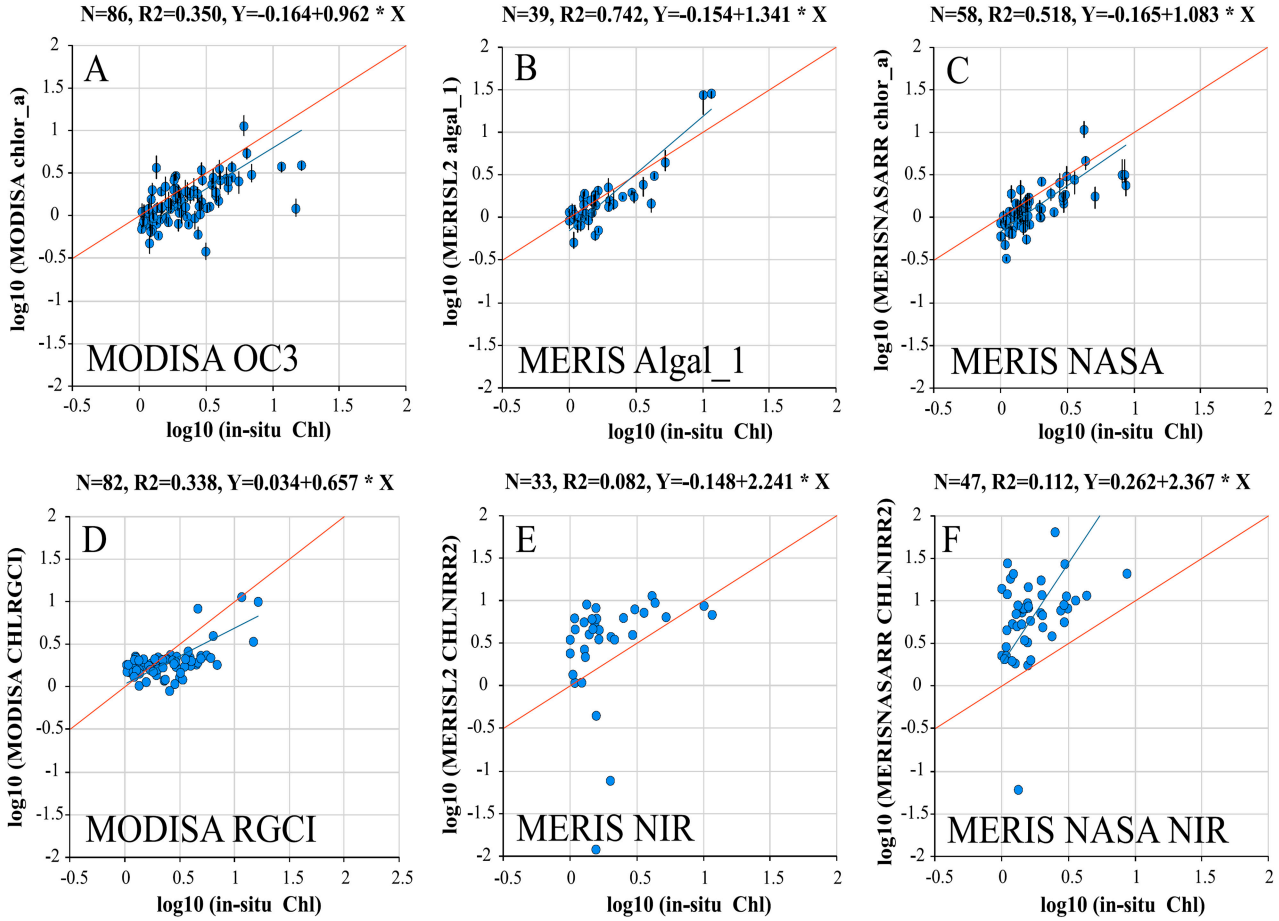


The median absolute percent error (MdAPE) over the full range of *in situ* Chla ranges from the lowest value of 21.1% for MODISA to the highest of 27.1% error for both SeaWiFS and VIIRS (Table 2 and Figure 4B). Bias (Figure 4C) shows that all sensors processed with NASA algorithms have a small negative bias or median underestimation (MRPE ranges from −1.7% for MODISA to −11.9% for MODIST), while the MERIS product processed with ESA algorithms has a small positive bias (MRPE = 8.9%).

The scatter and median errors increase drastically when evaluated for the range of medium and high Chla (Chla ≥ 1 mg m^{-3}, Table 3). In this Chla range, only MERISRR has relatively high R^2 (0.74) and relatively low absolute error (MdAPE = 23.2%) followed by MERISNASARR (R^2 = 0.52, MdAPE = 31.5%), while the other sensors have R^2 below 0.5 and MdAPE higher than 35%. VIIRS has the lowest R^2 and highest MdAPE for standard Chla products, but the number of match-ups is still small for definitive conclusions. To increase the number of VIIRS match-ups, we relaxed the timing criterion and used a 6-h time difference as the limit. At the medium and high Chla levels, all sensors underestimate pigment concentration compared to *in situ* Chla with a bias from −15.6% for MERISRR to about −40% for MODIST (Figure 4C). To determine if satellite retrievals of high Chla could be improved by implementing some alternative Chla algorithms that are designed to be suitable for high

Chla conditions, we also evaluated the Red-Green-Chlorophyll Index (RGCI, [18]) and the MERIS *Rrs*_709/*Rrs*_665 algorithms [19]. While the RGCI algorithm can be applied to most ocean color sensors, the near-infrared algorithm using *Rrs*_709 is applicable only to MERIS, as other major ocean color sensors do not have the 709-nm band. We compare the standard and high-Chla algorithms as applied to a subset of match-ups with *in situ* Chla > 1 mg m^{-3} (Figure 5). Regional tuning of these algorithms would potentially reduce the observed bias, but is not likely to reduce the scatter. No regional tuning was applied in Figure 5.

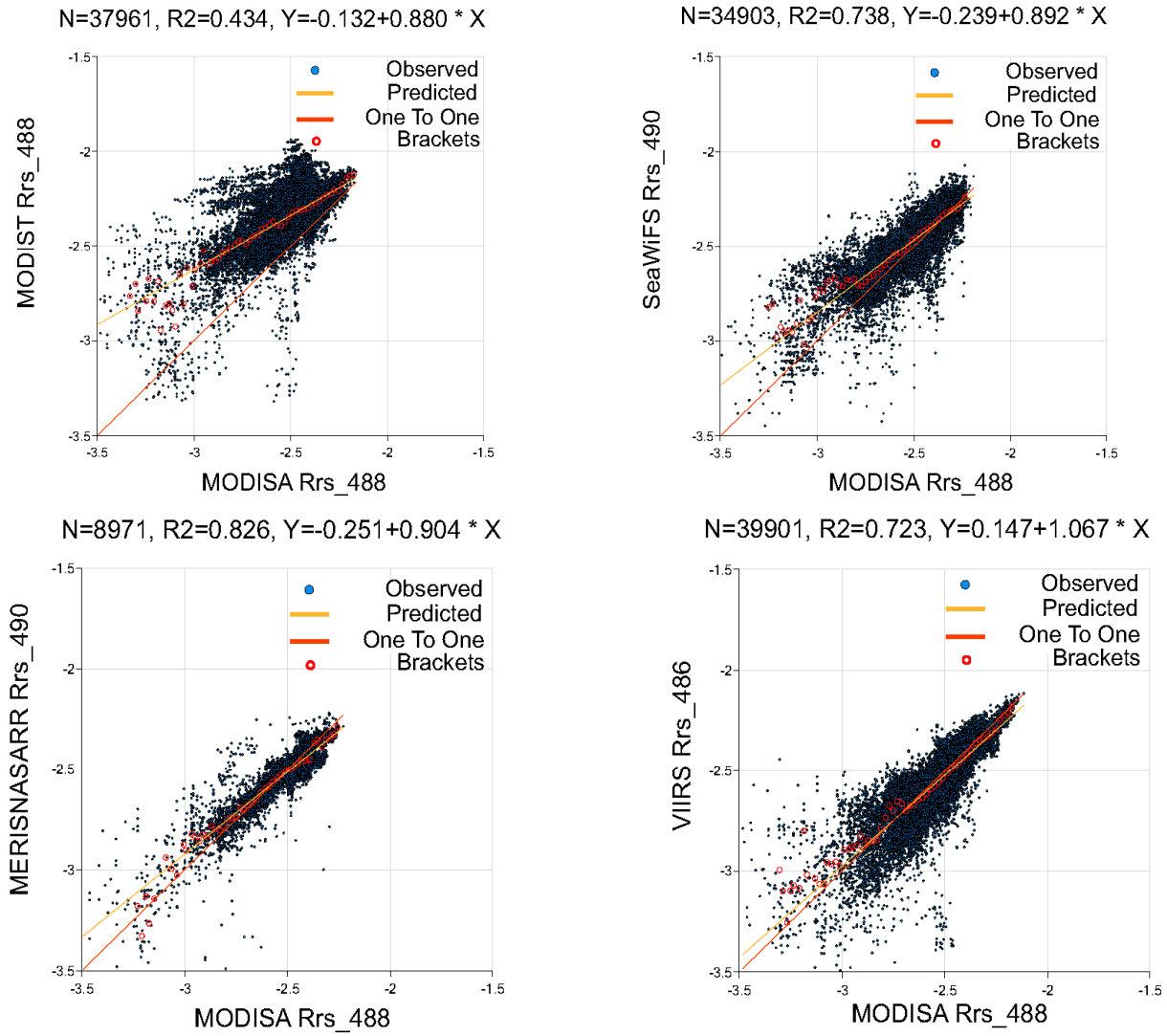
Figure 5. Chlorophyll-a match-ups with *in situ* Chla ≥ 1 mg m^{-3} using standard algorithms (**top**) and specialized (high-Chla) algorithms (**bottom**). (**A**) MODISA OC3; (**B**) ESA MERIS algal_1; (**C**) NASA MERIS; (**D**) MODISA Red-Green-Chlorophyll Index (RGCI); (**E**) MERIS Rrs_709/Rrs_665 applied to ESA MERIS Level-2 data; and (**F**) MERIS Rrs_709/Rrs_665 applied to NASA processed Level-2 data.



The empirical OC algorithms [6] use the maximum band ratio where the numerator band switches from a shorter wavelength (440 nm, the blue band) at low Chla to a progressively longer wavelength at higher Chla. For example, at high Chla, the numerator band is around 490 nm for MODISA, MODIST and VIIRS, whereas it is the 510-nm band for MERIS and SeaWiFS. We can therefore expect that the performance of a band ratio algorithm at medium and high Chla depends on the accuracy of *Rrs* retrievals of these bands. While we have no extensive and reliable ground truth to perform *Rrs* match-ups and to determine the accuracy of satellite-derived *Rrs* in the CC, we can compare the *Rrs*

measured by different satellite sensors with each other. We selected *Rrs* measured by MODISA as the common variable against which the other sensors are compared, as it overlaps temporally with all other sensors considered here and has a good calibration history. The band centers are slightly different for the different sensors, but for the purposes of this analysis, the differences are relatively minor. The scatter plots of satellite-derived *Rrs* at approximately 490 nm against *Rrs*_488 of MODISA (Figure 6) show that MERIS has the least scatter (MdAPE = 7.3%), followed by VIIRS (MdAPE = 9.5%) and SeaWiFS (MdAPE = 9.7%). A comparison with MODIST showed considerably more scatter (MdAPE = 40.7%).

Figure 6. Inter-sensor comparison of log10-transformed remote sensing reflectance (*Rrs*) at approximately 490 nm (488 nm for MODISA and MODIST, 490 nm for MERIS, 486 nm for VIIRS). This band is typically used for high Chla in the OC3 algorithm. Dots show same-day match-ups between the sensor along the same transect (Figure 1) during the first 99 days in 2012 (2004 for SeaWiFS). Bracket points are the median values of the points' corresponding medians of small brackets along the horizontal axis.



In order to explain the performance of the standard OC algorithms at medium and high Chla, we plotted subsets of these *Rrs* match-ups that correspond to medium and high Chla. According to the standard MODISA OC3 algorithm, pixels with Rrs_488/Rrs_547 < 0.8 correspond to Chla ≥ 3.3 mg m^{-3}.

Again, we see the best correspondence with MERIS (MdAPE = 22%), followed by VIIRS (MdAPE = 28%), SeaWiFS (MdAPE = 41) and MODIST (MdAPE = 140%) (Figure 7). Standard OCx algorithms use *Rrs* band ratios instead of the individual *Rrs* band values. This technique can reduce the error of the product if the *Rrs* errors are spectrally correlated. It appears that taking a ratio does not significantly reduce the errors for pixels with predicted at high Chla (Figure 8). The best correlation with MODISA band ratios is again observed for MERIS, followed by SeaWiFS. MODIST and VIIRS *Rrs* band ratios have practically no correlation with those of MODISA.

Figure 7. Inter-sensor comparison of $\log_{10}$-transformed remote sensing reflectance (*Rrs*) at approximately 490 nm as in Figure 6, but only for the subset of MODISA pixels with *Rrs*_488/*Rrs*_547 < 0.8. According to the standard MODISA OC3 algorithm, these pixels correspond to Chla ≥ 3.3 mg m^{-3}.

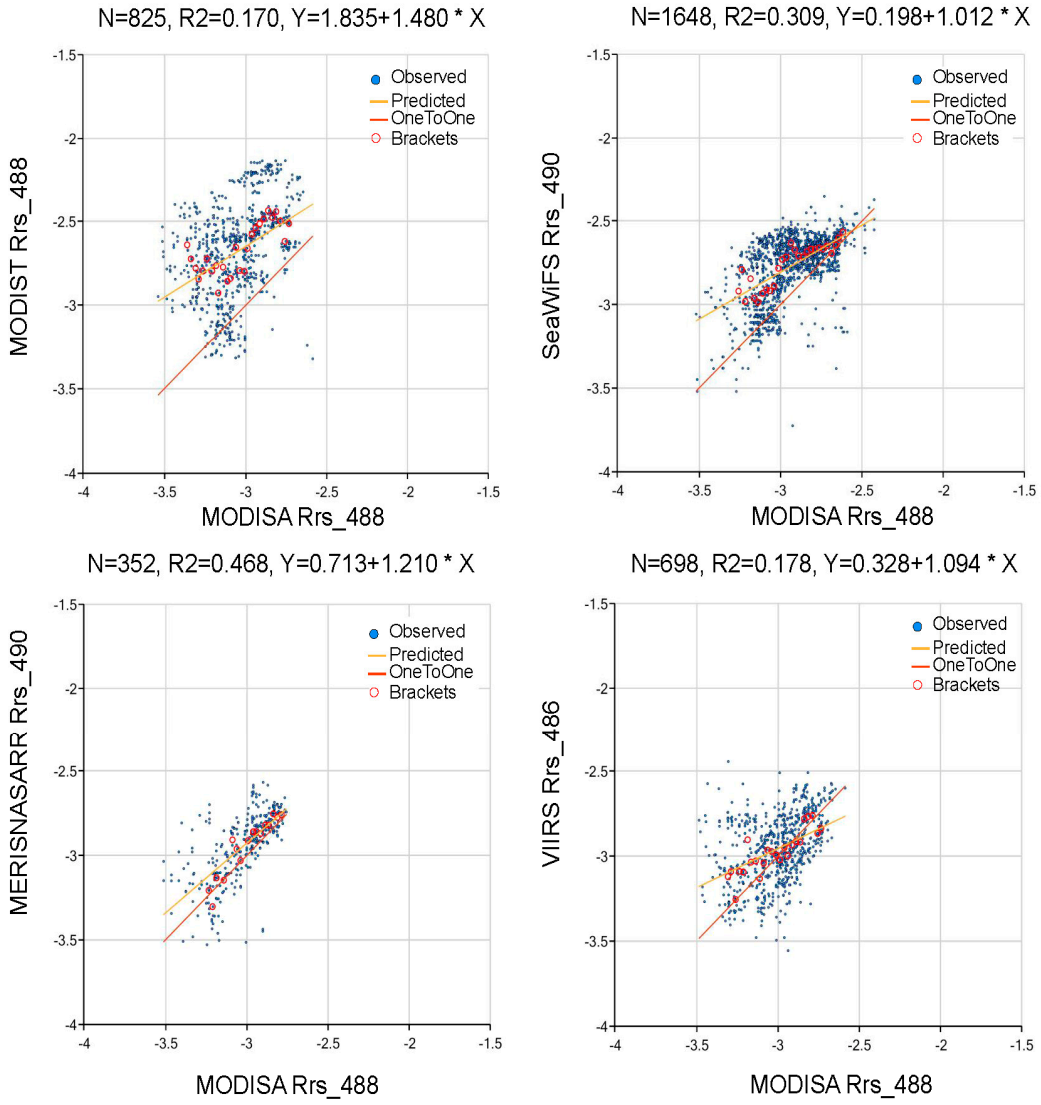
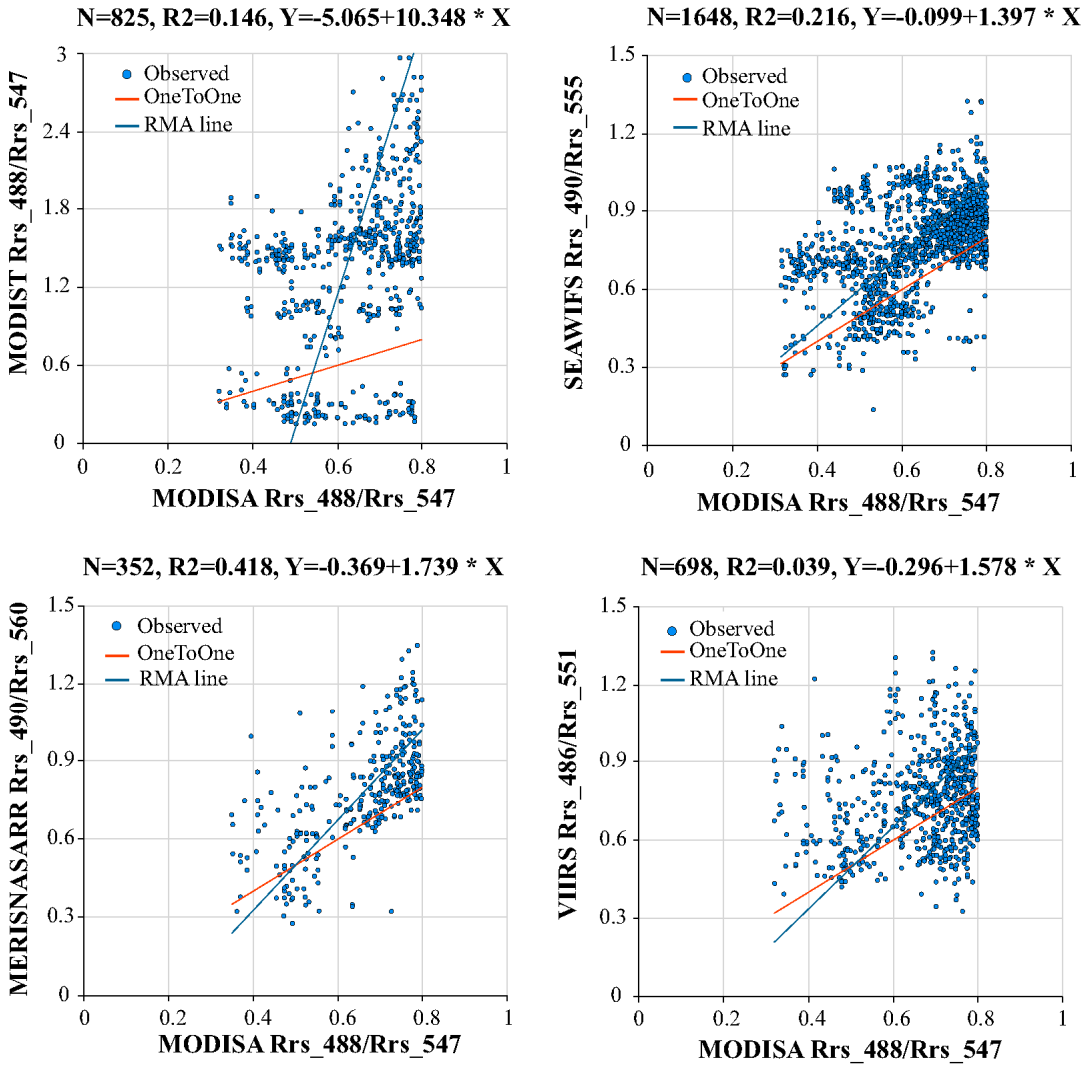


Figure 8. Inter-sensor comparison of *Rrs* ratios that are being used to detect high Chla by standard algorithms. The subset of MODISA pixels with *Rrs*_488/*Rrs*_547 < 0.8 (as in Figure 7) are used, corresponding to Chla ≥ 3.3 mg m^{-3}, according to the standard MODISA OC3 algorithm.



4. Discussion

We have shown that over the full range of *in situ* Chla, all satellite sensors estimate Chla reasonably well with median absolute percent errors below 30% and R^2 of about 0.8 or higher, which is consistent with other studies (e.g., [20]) and meets the original goal of 35% accuracy [21] set for SeaWiFS for retrieving ocean Chla in Case 1 waters. However, at medium and high Chla, the accuracy drops dramatically and in different ways for the different sensors. This is highly relevant to coastal management and the detection of phytoplankton blooms, particularly high biomass events, such as harmful algal blooms (HABs) [22,23]. Therefore, accurate satellite detection of Chla is required for those waters with Chla > 1 mg m^{-3}.

While MODISA has the lowest errors over the full range of Chla, there are very few valid Chla match-ups at higher concentrations. This is in contrast to the higher numbers of very high Chla values in the "raw" matchups (Figure 2). It is evident that MODISA retrievals of Chla at high levels become

highly variable from pixel to pixel, and those retrievals that are being filtered out by the match-up procedure are unreliable. This is in contrast to MERIS match-ups, which provide more reliable match-ups at medium and high Chla. For Chla > 1 mg m^{-3}, the ESA processed *algal_1* from MERISRR data has the highest accuracy (*i.e.*, lowest percent error) followed by the NASA processed *chlor_a* (also from MERISRR data). The disadvantage of MERIS data is that due to its orbit and swath width, there is less coverage compared to MODISA and SeaWiFS and, therefore, fewer match-ups. All of the other sensors have errors between 35% and 55%. Since VIIRS data have been available for a shorter period compared to other sensors (*i.e.*, since 2012), the number of high Chla match-ups with VIIRS is still small, and therefore, the error estimates are inconclusive. Preliminary analysis shows that VIIRS Chla estimates are similar to those made by SeaWiFS, but less accurate than the top performing sensors, MODISA and MERIS. VIIRS appears to be less reliable at medium and high Chla. All sensors have significant under-estimation at Chla > 1 mg m^{-3} (reported earlier in [5]), which is partly explained by the loss of high Chla match-ups due to their large pixel-to-pixel variability. These error patterns are consistent with the sensor-to-sensor match-ups of *Rrs*. It is interesting to note that while the original goals for SeaWiFS were to detect Chla with 35% accuracy and *Rrs* with 5% accuracy [21], the goal for Chla within the CC is met, but the current errors in *Rrs* that we estimated with an inter-sensor comparison are much higher. For the full range, the differences (MdAPE) with MODISA *Rrs_488* range from 7% for MERIS to 41% for MODIST, and for the *Rrs*, the range expected for high Chla, the MdAPE ranges from 22% for MERIS to 140% for MODIST. Chla estimates from MODIST are usually discarded due to calibration problems, which are shown by the high *Rrs* differences with other sensors. However, on average, MODIST Chla estimates are comparable to Chla estimates with SeaWiFS and VIIRS. These relatively good estimates of Chla by MODIST, in spite of the large errors in *Rrs* estimates (*cf*. Figures 3 and 6) are likely explained by the autocorrelation of the *Rrs* spectral bands, which reduces errors in band ratio algorithms.

It is difficult to perform a comprehensive validation of satellite *Rrs* using *Rrs* measured *in situ* due to the different ground resolution, inherently limited number of sites, variable Sun and weather conditions, *etc.* An indirect measure of the accuracy of satellite *Rrs* can be derived by comparing respective *Rrs* values derived from various satellites over the same pixel and same day. While some temporal delay (in hours) between satellite passes cannot be avoided, the daily *Rrs* measurements should have reasonable consistency, as they are assumed to be representative of the whole day. For the CC, the spatial and temporal decorrelation scales for phytoplankton biomass (FLH) derived from the MODIS-Aqua fluorescence line height (FLH) standard product also suggest that there is no inherent issue comparing data over a few hours' timespan [24]. We have shown that the differences in the distribution patterns of the scatter in Chla match-ups between different sensors can be explained by the distribution patterns of the respective sensor-to-sensor *Rrs* match-ups. When compared against MODISA, the *Rrs* measured by MERIS is clearly more consistent and less biased. At Chla > 1 mg m^{-3}, the likely band ratio combination in the OC3 algorithm for MODISA is *Rrs*_488/*Rrs*_547. When selecting a subset of MODISA pixels with *Rrs*_488/*Rrs*_547 < 0.8 that approximately corresponds to Chla ≥ 3.3 mg m^{-3} according to the standard NASA algorithm (OC3M version 6), the correlation between *Rrs* estimated by different sensors drops significantly. The respective R^2 is only about 0.2–0.3 between VIIRS and SeaWiFS and about 0.5 between MODISA and MERISNASARR. We can assume that the poor inter-sensor consistency in *Rrs* at high Chla is indicative of the poor accuracy of *Rrs* at

medium and high Chla. Using *Rrs* ratios instead of individual *Rrs* band values can reduce the errors if the *Rrs* errors are spectrally correlated; e.g., [25]. While ratioing improves Chla retrievals over the full range of *Rrs* (or Chla), it does not reduce the errors for pixels with low *Rrs*_490 (*i.e.*, those that are assumed to have high Chla, Figure 8). The best correlation with MODISA band ratios is observed for MERIS.

In order to improve the accuracy of Chla estimates at medium and high Chla, we need to either improve the accuracy of *Rrs* (*i.e.*, improve the atmospheric correction) or use a different algorithm. Of the sensors evaluated, MERIS clearly produces the best matchups for high Chla values. While some of the differences between sensors are clearly related to the instrumentation, the decrease in R^2, the increase in MdAPE and the switch from positive to negative bias (Figure 4) between MERISRR and MERISNASARR demonstrate that significant differences can be attributed to data processing and, presumably, the implementation of the atmospheric correction.

It appears that alternative algorithms specifically designed for high Chla waters, such as the Red-Green-Chlorophyll Index [18] and a band ratio algorithm using near-infrared bands [19], do not provide improvements in detecting high Chla in the CC. The algorithms using infrared bands, e.g., the MERIS 709 and 665 bands, are known to work well at high Chla with *in situ* data [26], but when applied to satellite *Rrs* with large errors, particularly at those infrared wavelengths, they have large errors and low accuracy compared to *in situ* data.

The biases observed in all of the ocean color platforms have potentially significant implications for our understanding of the dynamics and trends in the CC system. Multiple authors have reported significant increasing trends for *in situ* chlorophyll spanning multiple decades [9,27,28]. Satellite observations have the potential to provide synoptic spatial coverage and sustained temporal coverage; these data have also been used to infer trends within the CC [27,29,30], with decadal increases in Chla reported from these observations, consistent with the *in situ* data. However, using Monterey Bay as an example, an analysis of different data sets may lead to substantially different conclusions. A decadal (1988–2013) *in situ* time series exhibits a positive trend in Chla of 0.050/y (available from the Monterey Bay Aquarium Research Institute, http://www.mbari.org/bog/mb/Trends.htm). Using 4-km monthly data from a box bounding Monterey Bay (36.5–37.0°N, 121.75–122.25°W) for SeaWiFS (1998–2010) and MODISA (2002–2013), SeaWiFS Chla is increasing at 0.12/y (p = 0.001, Mann–Kendall test). In contrast, MODISA shows no significant trend (p = 0.672) and a decreasing Chla concentration (−0.04/y). When restricted to the overlapping time period (2002–2010), the MODISA results do not change, while the SeaWiFS results exhibit a slightly increased positive trend (0.15/y, p = 0.030). While SeaWiFS captures the observed (*in situ*) trend better than MODISA, SeaWiFS still exhibits more bias in matchups (Figure 4), and MODISA consistently exhibits higher monthly mean Chla than SeaWiFS for the overlapping period 2002–2010 ($p < 0.001$, paired t-test). By comparison, trends reported from state-space analysis of SeaWiFS data exhibit trends > 0.2 mg m^{-3} Chla per decade [31] for central California, exhibiting a consistent trend, but lower rate of change.

Given the interest in the development of environmental and climate data records [32] using satellite observations, these discrepancies in standard products using the most recent versions of the data highlight the potential difficulty of interpreting decadal trends. Moreover, advanced data assimilation techniques for real-time circulation and ecosystem models in the CC rely heavily on observational "truth" from satellite-derived chlorophyll. While we did not evaluate matchups for other coastal

regions, it is likely, given that the standard algorithms are derived using a large percentage of data from the CC, that similar biases exist (e.g., [33,34]).

5. Conclusions

We have shown that chlorophyll-a (Chla) estimates in the CC region by ocean color satellite sensors using standard algorithms are within the error limits of 35% over the full range of *in situ* Chla, but at *in situ* Chla > 1 mg m^{-3}, only products from MERIS (both the ESA produced *algal_1* and the NASA produced *chlor_a*) maintained reasonable accuracy. The loss in accuracy at medium and high Chla is caused by the poor retrieval of remote sensing reflectance. Accuracy is not improved through implementation of alternative algorithms, like the Red-Green-Chlorophyll Index [18] or band ratio algorithms using infrared bands. Uncertainties in satellite retrieval of medium to high chlorophyll values may affect the estimation of trends and may be biasing our interpretation of biomass and productivity in coastal waters of the CC, despite the large number of observations used in the development of the standard NASA ocean color algorithms.

Acknowledgments

Financial support was provided by the NASA Ocean Biology and Biogeochemistry Program Grants NNX09AT01G (M. Kahru and R.M. Kudela), NNX13AL28G (C.R. Anderson, R.M. Kudela and M. Kahru), NNX14AC42G (C.R. Anderson and R.M. Kudela), National Science Foundation (Grant OCE-1026607 to the CCE-LTER Program), the University of California Institute for Mexico and the United States (UC MEXUS) and Consejo Nacional de Ciencia y Tecnología, Mexico (CONACYT). Satellite data were provided by the NASA Ocean Color Processing Group and ESA MERIS team. We thank the CalCOFI and CCE-LTER programs, the NOAA SWFSC survey, the Monterey Bay Aquarium Research Institute and the Pacific Coastal Ocean Observing System for *in situ* data.

Author Contributions

M. Kahru assembled most of the datasets, performed the data analysis and wrote the paper. M.R. Kudela assembled part of the data, performed some of the data analysis and edited the manuscript. All authors assisted in the analysis and editing of the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Smith, R.C.; Wilson, W.H. Ship and satellite bio-optical research in the California bight. In *Oceanography from Space*; Gower. J.F.R. Ed.; Springer US: New York, NY, USA, 1981; Volume 13, pp. 281–294.
2. Kahru, M.; Mitchell, B.G. Empirical chlorophyll algorithm and preliminary SeaWiFS validation for the California Current. *Int. J. Remote Sens*. **1999**, *20*, 3423–3429.

3. Kahru, M.; Mitchell, B.G. Seasonal and non-seasonal variability of satellite-derived chlorophyll and CDOM concentration in the California Current. *J. Geophys. Res.* **2001**, *106*, 2517–2529.
4. Kostadinov, T.S.; Siegel, D.A.; Maritorena, S.; Guillocheau, N. Ocean color observations and modeling for an optically complex site: Santa Barbara Channel, California, USA. *J. Geophys. Res.* **2007**, *112*, doi:10.1029/2006JC003526.
5. Kahru, M.; Kudela, R.M.; Manzano-Sarabia, M.; Mitchell, B.G. Trends in the surface chlorophyll of the California Current: Merging data from multiple ocean color satellites. *Deep Sea Res. Part II*: *Top. Stud. Oceanogr.* **2012**, *77–80*, 89–98.
6. O'Reilly, J.E.; Maritorena, S.; Mitchell, B.G.; Siegel, D.A.; Carder, K.L.; Garver, S.A.; Kahru, M; McClain, C.R. Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* **1998**, *103*, 24937–24953.
7. O'Reilly, J.E.; Maritorena, S.; O'Brien, M.C.; Siegel, D.A.; Toole, D.; Menzies, D.; Smith, R.C.; Mueller, J.L.; Mitchell, B.G.; Kahru, M.; *et al. SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3*; NASA Technical Memorandum 2000–206892; NASA Goddard Space Flight Center: Greenbelt, MD, USA, May 2000; pp. 1–49.
8. Mitchell, G.; Kahru, M. Algorithms for SeaWiFS standard products developed with the CalCOFI bio-optical data set. Calif. Coop. Oceanic Fish. *Invest. Rep.* **1998**, *39*, 133–147.
9. Aksnes, D.L.; Ohman, M.D. Multi-decadal shoaling of the euphotic zone in the southern sector of the California Current System. *Limnol. Oceanog.* **2009**, *54*, 1272–1281.
10. McGaraghan, A.R.; Kudela, R.M. Estimating labile particulate iron concentrations in coastal waters from remote sensing data. *J. Geophys. Res.* **2012**, *117*, doi:10.1029/2011JC006977.
11. Kahru, M.; Lee, Z.; Kudela, R.M.; Manzano-Sarabia, M; Mitchell, B.G. Multi-satellite time series of inherent optical properties in the California Current. *Deep Sea Res. Part II*: *Top. Stud. Oceanogr.* **2013**, doi:10.1016/j.dsr2.2013.07.023.
12. Ohman, M.D.; Venrick, E.L. CalCOFI in a changing ocean. *Oceanography* **2003**, *16*, 76–85
13. Ryan, J.P.; Gower, J.F.R.; King, S.A.; Bissett, W.P.; Fischer, A.M.; Kudela, R.M.; Kolber, Z.; Mazzillo, F.M.; Rienecker, E.V.; Chavez, F.P. A coastal ocean extreme bloom incubator. *Geophys. Res. Lett*. **2008**, *35*, doi:10.1029/2008GL034081.
14. Ryan, J.P.;. Davis, C.O.; Tufillaro, N.B.; Kudela, R.M.; Gao, B.C. Application of the hyperspectral imager for the coastal ocean to phytoplankton ecology studies in Monterey Bay, CA, USA. *Remote Sens.* **2014**, *6*, 1007–1025.
15. Lorenzen, C.J. Chlorophyll b in the eastern North Pacific Ocean. *Deep-Sea Res.* **1981**, *28*, 1049–1056.
16. Werdell, P.J.; Bailey, S.W. An improved *in-situ* bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ*. **2005**, *98*, 122–140.
17. Bailey, S.; Werdell, P. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sens. Environ.* **2006**, *102*, 12–23.
18. Le, C.; Hu, C.; English, D.; Cannizzaro, J.; Kovach, C. Climate-driven chlorophyll-a changes in a turbid estuary: Observations from satellites and implications for management. *Remote Sens. Environ.* **2013**, *130*, 11–24.
19. Le, C.; Hu, C; Cannizzaro, J.; English, D.; Muller-Karger, F.; Lee, Z. Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sens. Environ.* **2013**, *129*, 75–89.

20. Hu, C.; Feng, L.; Lee, Z. Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water measurements. *Remote Sens. Environ*. **2013**, *133*, 168–182.
21. Hooker, S.B.; Esaias, W.E.; Feldman, G.C.; Gregg, W.W.; McClain, C.R. *An Overview of SeaWiFS and Ocean Color*; NASA Technical Memorandum 104566; NASA Greenbelt Space Flight Center: Greenbelt, MD, USA, July 1992; pp. 1–24.
22. Anderson, C.R.; Kudela, R.M.; Benitez-Nelson, C.; Sekula-Wood, E.; Burrell, C.T.; Chao, Y.; Langlois, G.; Goodman, J.; Siegel, D.A. Detecting toxic diatom blooms from ocean color and a regional ocean model. *Geophys. Res. Lett*. **2011**, *38*, doi:10.1029/2010GL045858.
23. Kudela, R.M.; Frolov, S.A.; Anderson, C.R.; Bellingham, J.G. Leveraging ocean observatories to monitor and forecast harmful algal blooms: A case study of the U.S. West Coast. In Proceedings of the Interagency Ocean Observation Committee IOOS Summit, Herndon, VA, USA, 13–16 November 2012.
24. Frolov, S.; Kudela, R.M.; Bellingham, J.G. Monitoring of harmful algal blooms in the era of diminishing resources: A case study of the U.S. West Coast. *Harmful Algae* **2013**, *21–22*, 1–12.
25. Jamet, C.; Loisel, H.; Kuchinke, C.P.; Ruddick, K.; Zibordi, G.; Feng, H. Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements. *Remote Sens. Environ*. **2011**, *115*, 1955–1965.
26. Gilerson, A.A.; Gitelson, A.A.; Zhou, J.; Gurlin, D.; Moses, W.; Ioannou, I.; Ahmed, S.A. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Opt. Express*. **2010**, *18*, 24109–24125.
27. Kahru, M.; Kudela, R.; Manzano-Sarabia, M.; Mitchell, B.G. Trends in primary production in the California Current detected with satellite data. *J. Geophys. Res.* **2009**, *114*, doi:10.1029/2008JC004979.
28. Rykaczewski, R.R.; Checkley, D.M. Influence of ocean winds on the pelagic ecosystem in upwelling regions. *PNAS* **2008**, *105*, 1065–1970.
29. Gregg, W.W.; Casey, N.W.; McClain, C.R. Recent trends in global ocean chlorophyll. *Geophys. Res. Lett.* **2005**, *32*, doi:10.1029/2004GL021808.
30. Kahru, M.; Mitchell, B.G. Ocean color reveals increased blooms in various parts of the World. *EOS Trans. AGU* **2008**, *89*, 170.
31. Thomas, A.C.; Mendelssohn, R.; Weatherbee, R. Background trends in California Current surface chlorophyll concentrations: A state-space view. *JGR Ocean*. **2013**, *118*, 5296–5311.
32. National Research Council. *Climate Data Records From Environmental Satellites*; The National Academies Press: Washington, DC, USA, 2004.
33. Alvarez, I.; Lorenzo, M.N. Analysis of chlorophyll a concentration along the Galician coast: Seasonal variability and trends. *ICES J. Mar. Sci.* **2012**, doi:10.1093/icesjms/fss045.
34. Beaulieu, C.; Henson, S.H.; Sarmiento, H.L.; Dunne, J.P.; Doney, S.C.; Rykaczewski, R.R.; Bopp, L. Factors challenging our ability to detect long-term trends in ocean chlorophyll. *Biogeosciences* **2013**, *10*, 2711–2724.