An Empirical Approach to Ocean Color Data: Reducing Bias and the Need for Post-Launch Radiometric Re-Calibration

Watson W. Gregg
Global Modeling and Assimilation Office
NASA/Goddard Space Flight Center
Greenbelt, MD 20771
watson.gregg@nasa.gov
(corresponding author)

Nancy W. Casey
Science Systems and Applications, Inc.
Lanham, MD 20706

John E. O'Reilly
NOAA/National Marine Fisheries Service Narragansett Laboratory
28 Tarzwell Drive
Narragansett, RI 02882

Wayne E. Esaias
Laboratory for Hydrospheric and Biospheric Processes
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

Accepted for Publication Remote Sensing of Environment 2009

Abstract

A new empirical approach is developed for ocean color remote sensing. Called the Empirical Satellite Radiance-In situ Data (ESRID) algorithm, the approach uses relationships between satellite water-leaving radiances and in situ data after full processing, i.e., at Level-3, to improve estimates of surface variables while relaxing requirements on post-launch radiometric re-calibration. The approach is evaluated using SeaWiFS chlorophyll, which is the longest time series of the most widely used ocean color geophysical product.

The results suggest that ESRID 1) drastically reduces the bias of ocean chlorophyll, most impressively in coastal regions, 2) modestly improves the uncertainty, and 3) reduces the sensitivity of global annual median chlorophyll to changes in radiometric re-calibration. Simulated calibration errors of 1% or less produce small changes in global median chlorophyll (<2.7%).

In contrast, the standard NASA algorithm set is highly sensitive to radiometric calibration: similar 1% calibration errors produce changes in global median chlorophyll up to nearly 25%. We show that 0.1% radiometric calibration error (about 1% in water-leaving radiance) is needed to prevent radiometric calibration errors from changing global annual median chlorophyll more than the maximum interannual variability

observed in the SeaWiFS 9-year record ($\pm$3%), using the standard method. This is much more stringent than the goal for SeaWiFS of 5% uncertainty for water leaving radiance.

The results suggest ocean color programs might consider less emphasis of expensive efforts to improve post-launch radiometric re-calibration in favor of increased efforts to characterize in situ observations of ocean surface geophysical products. Although the results here are focused on chlorophyll, in principle the approach described by ESRID can be applied to any surface variable potentially observable by visible remote sensing.

Introduction

Radiometric calibration is the foundation of remotely-sensed ocean color data. Calibration is strictly defined as a factor applied to at-satellite signals to bring them into agreement with some measure of truth. Specifically, it is the conversion of electronic responses from the sensor (digital counts) into radiometric units (typically mW cm^{-2} μm^{-1} sr^{-1}). Although activities such as removal of sensor artifacts, including trends, optical behavior, polarization, etc. are commonly lumped into the general usage "calibration" (Hooker et al., 2007), these are more precisely considered separate activities. Calibration involves identifying a single factor for each wavelength that represents the conversion from counts to radiances. Thus it is a spatially and temporally constant factor, varying only with wavelength, intended to remove bias. For modern NASA ocean color missions, the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), this is derived from pre-launch testing, and is refined on orbit using a single location near Hawaii (Bailey et al., 2008). This refinement is called vicarious calibration (Antoine et al., 2005), and actually incorporates some atmospheric correction error. We refer here to the latter process, i.e., refinement of calibration after launch, as post-launch radiometric re-calibration.

The quality of derived ocean color geophysical products (mainly chlorophyll, but others are considered) is ultimately dependent upon the radiometric calibration, and stringent criteria are set (McClain et al., 2004). Drawing from the results of the first ocean color mission, the Coastal Zone Color Scanner (CZCS), the goal of succeeding missions has been 5% uncertainty in water-leaving radiances (Hooker and Esaias, 1993). Successor missions, e.g., SeaWiFS and MODIS, were designed to meet these goals, and subjected to a comprehensive validation program consisting of global surface observations. Using state-of-the-art calibration (both pre-launch and post-launch), comprehensive sensor characterization, continuous trend analysis, complex atmospheric correction, and rigorous bio-optical algorithms, the data sets have gained broad usage by the Earth science community, a testament to the care and skill of the sensor design and data handling.

Yet the radiometric goals have not been achieved. Even using carefully selected in situ data sets for comparison, recent results show a bias in spectral normalized water-leaving radiances ranging from 0.2% at 412nm to 35% at 670 nm, and uncertainty ranging from 11% at 510nm to 65% at 670 nm (Franz et al., 2007). If we consider only the bands used for deriving chlorophyll (443, 490, 510 and 555nm), the bias range is -3.2% (555nm) to -5.8% (490nm), and the uncertainty 11% (510nm) to 16% (443nm). This is a problem deriving primarily from radiometric calibration and atmospheric correction bias. The inability to meet the mission goals, despite strong efforts in post-launch re-calibration technology and in situ radiometry, has major repercussions for the derived geophysical products. This is particularly true given that now, for the first time, we have an understanding of the variability of ocean chlorophyll from the decade-long observational record, and what uncertainty is needed. The 9-year record of SeaWiFS (1998-2006) showed that the minimum and maximum departures from the global chlorophyll median were -2.1% and +3.5%, a maximum range of 5.6%. We suggest first that global and large scale regional estimates of products is the most important objective of satellite observations, as it cannot be done with any other platform, at least at the present time. Satellites provide our only unaliased observation of global biology processes. We suggest second that the error of derived chlorophyll should be less than the observed natural departures if we are to support observations of climate change and variability, especially when we consider time-series that span >1 mission. This is a much more stringent requirement than originally conceived on the basis of the CZCS mission. We will show here that the present radiometric goals are not sufficient to meet this condition using the algorithm and processing sequence as currently configured by the NASA project.

We introduce the Empirical Satellite Radiance-In situ Data (ESRID) algorithm, a new approach for producing derived geophysical products. The approach is entirely empirical, and uses satellite water-leaving radiances, whatever their value and calibration, and derives relationships with co-located, coincident in situ data. Although ESRID is general in nature, we focus here on chlorophyll as it is the main ocean color geophysical product and vast repositories of in situ data are available. The results of the approach are compared to comprehensive global in situ data. Quantitative statistical analyses of the new approach and comparisons with NASA standard products are shown. The results are also weighted by the natural distributions of the geophysical products, to provide a more representative interpretation of the statistics, in contrast to previous analyses that reflect the distribution of the in situ validation data.

But most importantly, we show the effects of major and minor perturbations in radiometric calibration on global chlorophyll medians. In doing so we follow a path to lead us to a fundamental question of quantitative satellite ocean color: what is the global median chlorophyll?

Methods

In situ Data

The first step in development and analysis of a new empirical ocean color approach is to obtain in situ data. We aggregated 166589 observations of chlorophyll from the National Oceanographic Data Center (NODC; Conkright et al., 2002), NASA in situ (Werdell and Bailey, 2005), and Atlantic Meridional Transect (AMT; Aiken et al., 2000) archives. Of these, 53588 occurred during the SeaWiFS mission from late 1997 through 2005.

The quality of in situ chlorophyll data is evaluated using the Blended Analysis (Reynolds, 1988; Gregg and Conkright, 2001). The in situ and satellite data are degraded to 1-degree, seasonal resolution (using the arithmetic mean) to minimize local and transient variability. Obvious departures in the blended and difference (blend minus satellite data) fields provide clues to anomalous in situ data. In particular, we look for in situ data that create disruption within the satellite-derived major biological regimes, specifically, low chlorophyll-open ocean gyres, moderate chlorophyll tropical upwelling, and high chlorophyll-high latitude regions. If there are readily apparent anomalies, and the anomalies are associated with individual cruise data, then data from the entire cruise are removed. An example is a north-south Pacific cruise in winter 2005 (Figure 1), where the cruise track is apparent in the blended field, protruding into the South Pacific gyre. In this particular case, the results were reported to the data provider, who found an error in data submission, and corrected it.

For the SeaWiFS era, we removed data from 5 cruise data sets (1998, 2001, 2004, and 2 in 2005) and a single anomalous value in the North Central Atlantic gyre in 1999. This last value was the only individual value we eliminated while keeping the remainder of the cruise data. The total amount of in situ data removed was 12515, of which 12420 (99%) were along two Atlantic transects in 2005. The remaining data numbered 41073, which represented about 77% of the total original data (or 99.8% of all data excluding the Atlantic transects in 2005).

Initial analyses utilize what we refer to as in situ-weighted, or unweighted, statistics. This involves comparison of satellite chlorophyll data with in situ data, using all the coincident, co-located data. It is referred to as in situ-weighted, because the comparison follows the distribution of the in situ data.

Statistical analyses involve the percent error

$$\text{Percent Error (PE)} = \frac{S-I}{I} \times 100 \tag{1}$$

where S is the satellite (SeaWiFS) chlorophyll, and I is the in situ chlorophyll. The bias is estimated using the median of the percent errors:

$$\text{bias} = \text{median(PE)} \tag{2}$$

The median was chosen for error analysis because of the lognormal distribution of chlorophyll data (Campbell, 1995). Logarithm transforms are common in such circumstances but percent errors are difficult to interpret. The median is nearly independent of the distribution of the data and is thus a useful, simple, and easy-to-interpret representation of the bias regardless of the distribution, and naturally incorporates the percent error.

The uncertainty, or dispersion of the data, is represented by the Semi-Interquartile Range (SIQR)

$$\text{Semi-Interquartile Range (SIQR)} = \text{IQR(PE)} * 0.5 \tag{3}$$

The interquartile range (IQR) encompasses all data between the 25th percentile and the 75th percentile. One–half this value, the Semi-Interquartile Range (SIQR), is analogous to the standard deviation for normally-distributed data, in that the median $\pm$ SIQR contains 50% of the data. The difference is that the mean $\pm$ standard deviation encompasses 68% of the data.

We also report log statistics for a balanced presentation. Both in situ and SeaWiFS data are logarithmically-transformed (base 10) and log bias defined as

$$\text{log bias} = \frac{\sum [\log(S)-\log(I)]}{} \tag{4}$$

n

where n is the number of samples, and root-mean square (RMS) log error defined as

$$\mathrm{RMS} = \sqrt{\frac{\sum [\log(S)-\log(I)]^2}{n}} \qquad (5)$$

In situ-weighted (unweighted) statistics may not be representative of the validity of satellite data if they are distributed differently than satellite data. In the case of chlorophyll they are different (Table 1). The 9-year SeaWiFS data set showed a distribution across 6 chlorophyll brackets that contrasts with in situ data. In situ data are over-represented for the -0.5 to 0.0 and 0.0 to 0.5 log brackets (0.32 to 1.0 and 1.0 to 3.2 mg m^{-3}, respectively), leading to an over-consideration of high chlorophyll comparisons in the total. Conversely, in situ data are under-represented in low chlorophyll (-1.5 to -1.0 and -1.0 to -0.5, corresponding to 0.03 to 0.1 and 0.1 to 0.3 mg m^{-3}) where SeaWiFS data encompass 78% of the global total. If these differences in distribution are not accounted for, one may obtain a nominally good comparison that may disguise important errors in the satellite data. This can have major implications regionally, where satellite data may be dominated by particular portions of the satellite chlorophyll frequency distribution, resulting in erroneous estimates of global statistics.

To rectify this problem, global representations of bias and uncertainty are weighted according to the satellite frequency distribution for a more realistic representation of satellite data performance. These are called satellite-weighted statistics:

$$\text{Satellite-weighted bias} = \frac{\sum \mathrm{bias}(C_b)\, F_b}{\sum F_b} \qquad (6)$$

where C_b represents the log chlorophyll brackets shown in Table 1, bias (C_b) is the bias determined within each log chlorophyll bracket, and F_b is the satellite fraction (percentage/100) for each bracket. Uncertainty is determined similarly. The statistics for each chlorophyll bracket are determined independently, and only the global summary uses the satellite-weighting. For example, if the bias for each of the 6 log chlorophyll brackets is 10% except for the third bracket, which is 5%, then by Eq. 6, the satellite-weighted bias is

$$\frac{10.0*0.0087 + 10.0*0.2486 + 5.0*0.5436 + 10.0*0.1466 + 10.0*0.0381 + 10.0*0.0145}{0.0087 + 0.2486 + 0.5336 + 0.1466 + 0.0381 + 0.0145}$$

or in this case, 7.3%. The unweighted statistics, which follow the in situ distribution (Table 1), produce a value of 8.7%.

ESRID

ESRID uses the normalized water-leaving radiances from SeaWiFS at Level-3 (Standard Mapped Images), matches their locations with in situ data mapped to the same Earth grid, and derives statistical relationships. ESRID is applied a posteriori to the standard data processing procedure, shown in Table 2. It requires that all trends be removed a priori, just as in the vicarious calibration methodology (Evans and Gordon, 1994). For heritage considerations, we choose a fourth-order polynomial non-linear regression for chlorophyll, using the Maximum Band Ratio (MBR) concept (O'Reilly et al., 1998). In this algorithm, the reflectance ρ is computed as

$$\rho(\lambda) = \pi[L_w(\lambda)]_N / [F_o(\lambda)D] \tag{7}$$

where λ is the wavelength in nm, $[L_w(\lambda)]_N$ is the normalized water-leaving radiance, Fo is the extraterrestrial irradiance, and D is a correction for day of year. The reflectance ratio

$$R(\lambda) = \rho(\lambda)/\rho(555) \tag{8}$$

where λ= 443, 490, or 510nm, depending on which is highest, is chosen for each matchup and for the later computation of chlorophyll

$$\log \text{chl} = a_0 + a_1R + a_2R^2 + a_3R^3 + a_4R^4 \tag{9}$$

where $a_{0\text{-}4}$ represents empirical coefficients.

Although excellent relationships are derived from a simple non-linear regression application, biases are observed in the chlorophyll brackets and in the satellite-weighted global statistics. We refer to this direct approach as Level-3 regression.

To minimize these biases in chlorophyll brackets, ESRID utilizes the median of all matchups occurring within very small increments of in situ chlorophyll across the entire chlorophyll range. This can be seen in Figure 2. The increments are in logarithmic units. Since in situ data can be poorly sampled in some increments, most notably the high and low ends of the concentration spectrum, the increments are allowed to vary. We set a minimum of 5 observations for an increment to be a valid data point. Thus the first chlorophyll increment, where we have only a few in situ observations, spans the log range -1.979 to -1.891. An increment in the middle of the range, where there is an abundance of observations, uses our smallest allowable increment, 0.001 log chlorophyll. The median of all $[L_w(\lambda)]_N$ within the chlorophyll increment is associated with the mid-point of the increment to produce a data point for the regression (Figure 2). Tests with other minimum number observations produced similar results, but too large a value can produce poor results if there are few data points. For example, MODIS-Aqua for the period Sep 2002 through 2005 has about 1/5 as many matchups as SeaWiFS. In this case, a minimum of 50 values per increment leads to only 26 total data points for the regression. We consider 50 data matchup values to be necessary for a good regression fit, and 100 are preferable.

The effects of ESRID are most important on the derived best-fit polynomial (Figure 2). The fit of the un-transformed Level-3 regression is over-constrained by the massive number of observations occurring in the high range of chlorophyll concentrations, leading to a flatter regression slope (shown in the Results). ESRID partially equalizes the uneven distribution of the in situ observations by taking the median log reflectances over small chlorophyll increments, leading to a steeper slope.

A further refinement of ESRID is to apply separate non-linear regressions for coastal and open ocean regions. Here we define coastal regions as those where the bottom depth is <200m. There are no matchups in the coastal regions for low chlorophyll (< 0.1 mg m^{-3}), so we include the low chlorophyll data points from the open ocean to provide a regression fit. Similarly, we include matchup points to the 100m isobath for the open ocean regression fit, although the application of the open ocean empirical algorithm only applies for bottom depths ≥200m. These modifications provide enhanced error reduction and minimize discontinuities. Although this bifurcation of open ocean and coastal regions reduces the available number of data points for regression fits by about half for each domain, sufficient data remain for a meaningful regression for SeaWiFS. Data <5m depth are excluded to minimize the contribution of inland lakes.

ESRID empirical coefficients a_0 through a_4 are provided in Table 3, including open ocean and coastal values. We also provide coefficients for the un-bifurcated, single-coefficient set ESRID, where the coefficients are derived from global relationships.

Validation

Since ESRID utilizes in situ chlorophyll inherently, validation of results against the same in situ chlorophyll is problematic. However, there is sufficient disagreement between the ESRID results and the in situ chlorophyll that a comparison is not completely without merit. We consider this validation to set a lower bound on the estimate of error. To achieve a measure of independent validation, we also compare ESRID to in situ chlorophyll after withholding data. In this case, some of the data are used for development of the regression statistics in ESRID, and the rest for comparison. We find that a 50% split between method development and validation provides a reasonable estimate of the error. The data are divided systematically, every other value used for development and validation, respectively. This provides a measure of independence for the validation and establishes what we consider an upper bound to the error estimate. Other withholding scenarios were tested, 10:90% development:validation, 20:80%, and 40:60% (and the inverses), but in each case there were too few points in some chlorophyll brackets to establish either a reliable regression or meaningful error statistics.

In Situ Data Errors

An empirical methodology is dependent upon the validity of the in situ data. Unfortunately, there is little quantitative information on in situ chlorophyll data error in the major international archives. We attempt to provide evidence of the uncertainty of the in situ data for the combined NOAA/NASA/AMT data base by seeking 9km gridded values that occur 2 or more times within a 9km grid point. First the mean is computed. Then the uncertainty is estimated as the SIQR of each value within the grid from the grid point mean. The log RMS error is also computed. This only allows us to estimate uncertainty, not the bias, which remains unknown. It also conflates errors associated with spatial variability within the 9km grid point with measurement errors. However, it provides some estimate of the uncertainty at the satellite native Level-3 resolution.

Sensitivity to Radiometric Calibration

We estimate the sensitivity of ESRID to radiometric calibration by adjusting the values of top-of-the-atmosphere radiance, $L_t(\lambda)$ at each of the bands used to compute chlorophyll (Eq. 8) by fixed percentages: 10%, 5%, 1%, 0.5%, 0.1%. For example, a 10% calibration error is simulated by multiplying $L_t(\lambda)$ by 1.1. A 0.1% calibration error involves a multiplication by 1.001. We simulate negative calibration errors (-10%, -5%, etc.) similarly: a -10% error is represented by $L_t(\lambda)$ x 0.9, and a -0.1% error is $L_t(\lambda)$ x 0.999. The simulation requires that adjustments be applied at Level-1A (see Table 2) and a complete forward processing to Level-3 is executed. We use the algorithms and vicarious calibration from the NASA Version 3 SeaWiFS set (O'Reilly et al., 2000;

Robinson et al., 2000; Wang et al., 2000), although we use recent sensor trend data (Version 5.2, produced in 2007).

After adjustment of $L_t(\lambda)$, we use the NASA Standard and ESRID to compute daily mean chlorophyll fields at 9km, which are then used to obtain monthly means at the same resolution. These monthly means are the basis for annual mean chlorophyll fields. Then we use the median of these global annual mean chlorophyll fields to represent the central tendency. To summarize, the temporal aggregation involves use of the mean, while the central tendency uses the median over space. This conforms to the observation that chlorophyll is log-normally distributed in space (Campbell, 1995), but assumes that the distribution is more Gaussian in time.

We compute global annual chlorophyll medians for each of the simulated calibration errors described above, for each of the 443, 490, 510, and 555 nm bands. After each simulated calibration adjustment for each band, ESRID is re-derived using the new $[L_w(\lambda)]_N$ fields, with the in situ chlorophyll unchanged.

The effects of these simulated calibration errors using the NASA Standard method are shown for comparison. This represents the sensitivity of the NASA Standard and ESRID methods to changes in radiometric calibration.

In addition to the +10% to +0.1% sensitivity analyses, we investigate the sensitivity of global annual median chlorophyll to post-launch re-calibration factors (so-called vicarious gains) determined for actual NASA ocean color missions. Full Level-1A through Level-3 processing is employed, and comparison of the results with those without the re-calibration factors, provides an indication of the importance of re-calibration to the NASA Standard method and ESRID in practice. The re-calibration factors are obtained from the NASA Ocean Color Web (http://oceancolor.gsfc.nasa.gov) for SeaWiFS, MODIS-Aqua, and CZCS. We do not show MODIS-Terra in the Results but its re-calibration factors fall within the missions evaluated.

These sensitivity analyses are performed for 2001, and the derived annual median chlorophyll values for each calibration adjustment are compared to a reference. This reference uses the unadjusted $L_t(\lambda)$ for both ESRID and the NASA standard. We do not use data where any of the OC4 MBR radiances fall below 1 digital count, known as the noise equivalent radiance (Gordon, 1990). For the re-calibration factors from actual missions, the relationship between reference and test are reversed (factors applied represents the reference). This enables us to evaluate the effects of not applying the re-calibration factors.

A rule-of-thumb is that an adjustment of 1% in $L_t(\lambda)$, where the radiometric calibration applies, is equivalent to a 10% change in $[L_w(\lambda)]_N$ (Franz et al., 2007). However, in practice there is considerable variability in the relationship between $\Delta L_t(\lambda)$ and $\Delta[L_w(\lambda)]_N$. We investigate the relationship, as an attempt to translate our simulated calibration error results into water-leaving radiance errors, by evaluating a 1% change for all the SeaWiFS visible bands. The analysis is performed for 2 months, Jan and Jun, 2001.

Results

Interannual Variability in Global Median Chlorophyll

Interannual variability of global annual median chlorophyll from SeaWiFS shows minor deviation from the climatological value of 0.1845 mg m^{-3} over a 9-year record

(Figure 3). The maximum departure from the climatological median occurred in 1999 and 2000 at +3.5%. This corresponded to the peak years of the 1998-2001 La Niña event in the tropical Pacific. The minimum departure occurred in 2002 at -2.1%. This corresponded with an El Niño event that was not as strong as the 1997-1998 event, but the previous event switched in mid-1998, diminishing its effect on the annual median.

Error Statistics

Unweighted, or more accurately, in situ-weighted, statistics show that the Level-3 regression reduces bias to nearly 0, compared to nearly 17% for the NASA processing (Table 4). There is a modest reduction in uncertainty but a reduction in regression slope. These results are not necessarily compelling, because zero bias can be achieved if errors in different portions of the chlorophyll domain are cancelling out (i.e. overestimates at low chlorophyll cancel out underestimates at high chlorophyll).

When using satellite-weighting, statistics comparing satellite-derived chlorophyll and in situ chlorophyll from the standard NASA SeaWiFS processing and the new ESRID approach, show remarkably different results (Figure 4). To compensate for ESRID's lack of independence between development and validation, we include a test case where 50% of the data is withheld. Using satellite-weighted error statistics that represent the error corresponding to the distribution of global chlorophyll rather than the distribution of in situ sampling, we find that ESRID reduces the global bias to nearly zero. This is a major improvement over SeaWiFS, which indicates a bias of 28%. The improvement of ESRID in global uncertainty, represented by the SIQR, is not nearly as dramatic, indicating a reduction of only about 2-3%.

When these comparisons are broken into chlorophyll brackets, between 0.01 and 100 mg m^{-3} in log increments of 0.5, the bias reinforces the cumulative results (Figure 5) especially at the chlorophyll brackets that predominate in the oceans (see Table 1). A spike in the bias for the 50% withholding case in the -2 to -1.5 bracket (0.01 to 0.032 mg m^{-3}) is an artifact from the lack of data in this test: only 48 data points were available for algorithm development and only 28 for validation. The NASA Project SeaWiFS outperforms ESRID at the higher brackets, 0 to 2 (1 to 100 mg m^{-3}), but only slightly so at the highest bracket, 0.5 to 2. Combined, these brackets represent only 5.3% of the global data (Table 1).

Like the cumulative statistics, the partitioned uncertainty indicates only modest improvement for ESRID. There is degradation in uncertainty for ESRID at the highest bracket, 0.5 to 2.

Satellite-weighted error statistics divided into open ocean and coastal regions provide different views of these two very important ocean classifications (Figure 6). The open ocean statistics reflect the global statistics, with dramatic improvement by ESRID in bias (to nearly negligible values), and no improvement in uncertainty. We note that the uncertainty level generally meets the SeaWiFS target for chlorophyll of 35% (Hooker and Esaias, 1993).

The coastal results for NASA standard processing are disappointing, with 75% bias and 97% uncertainty. Ocean color sensors are widely considered to produce poorer results in coastal regions than open ocean, which is confirmed here.

ESRID, however, produces major improvements in coastal regions (Figure 6). Bias estimates fall from 75% to <17%, regardless of withholding. Large reductions in uncertainty are also apparent. We note that without the dual derivation and application of ESRID, that is, a single non-differentiated algorithm, we obtain results for the coasts similar to the NASA standard processing.

Effects of Post-Launch Re-Calibration Errors on Global Annual Median Chlorophyll

The sensitivity of global annual median chlorophyll estimates to radiometric calibration errors shows vastly different responses for the NASA standard algorithm and ESRID (Figures 7 and 8). Both the NASA standard method and ESRID produce changes in the global median many times the maximum interannual variability (about 3%) for absolute calibration errors greater than or equal to $\pm$5% at both 443 and 555nm. At $\pm$1% and $\pm$0.5% simulated calibration error, changes in the global median produced by ESRID are within our 3% requirement while the NASA standard is still many times our threshold. It is not until we reach a calibration error of $\pm$0.1% that the NASA standard shows acceptable insensitivity (Figures 7 and 8). These results translate to a maximum tolerance of 10% and 12% water-leaving radiance error at 443 and 555nm respectively for ESRID. For the NASA standard method, a maximum of tolerance of 1% is required in $[L_w(\lambda)]_N$. The sensitivity of the 490 and 510nm bands is less than 443 and 555nm (Table 5), and there is a difference whether there is a positive simulated calibration error or a negative error. At 490nm a positive error of 0.5% is within the interannual variability threshold using the NASA method, while a positive error of 1% is sufficient for 510nm. ESRID meets the threshold at $\pm$1% error, as with 443 and 555nm. For negative simulated errors at 490 and 510nm, all cases are within the 3% global median threshold for both methods.

The sensitivity of global annual median chlorophyll using the NASA Standard and ESRID methods in the context of the derived re-calibration factors for SeaWiFS is shown in Figure 9. If the re-calibration factor for the 443nm band is not applied using the NASA Standard method, serious errors in global median chlorophyll occur (23.9%). A slightly larger error occurs (25.6%) if the suite of re-calibration factors is not applied. For ESRID, none of the re-calibration factors changes the global median chlorophyll above the 3% interannual variability threshold, nor does the ensemble of all factors (Figure 9).

Similar high sensitivity of global annual median chlorophyll using the NASA Standard occurs if we apply the re-calibration factors from MODIS-Aqua and the CZCS to the nearest band for SeaWiFS (Figure 10). Note that in this analysis we are evaluating the simulated global median error in SeaWiFS, but using the re-calibration factors for MODIS and CZCS to evaluate the impact. Not using the re-calibration factors for 443 and 555nm from MODIS and CZCS produce changes in global median chlorophyll far in excess of our 3% target (up to a maximum of 64% for CZCS re-calibration factor associated with 555nm). In contrast, only one re-calibration factor causes ESRID to exceed the 3% threshold, namely the CZCS 555nm factor. Even then, the change in the global median is only -4.1%. Interestingly, the MODIS re-calibration factors at 443 and 555nm compensate when using the ensemble, producing a net change of 3.5%.

Ratio of Calibration Error to Water-Leaving Radiance Error

The mean ratio of radiometric calibration error to normalized water-leaving radiance error varies by wavelength (Table 6). There is also considerable variability in the ratio over monthly data, as exhibited by the standard deviation. For the SeaWiFS chlorophyll bands (443-555nm), the ratio ranges from 7.1% (490nm) to 12.3% (555nm). The mean over the chlorophyll bands is 9.4%, which is not very different from the rule-of-thumb of 10%.

In Situ Chlorophyll Error

The uncertainty of in situ chlorophyll data from the combined NOAA/NASA/AMT archives is estimated at 6.23% (log RMS = 0.10), with N = 65504 from 1979 to 2005 (Figure 11). There is no apparent change in time in annual estimates from 1979 to present ($r = 0.211$, $N = 27$, not significant at 95% probability level).

ESRID Global Chlorophyll and NASA Standard

Global annual mean chlorophyll derived from ESRID exhibits subtle differences from the NASA standard for 2005 (Figure 12). What is most apparent is the expansion of the mid-ocean gyres in ESRID and reduction in chlorophyll in coastal regions. The median difference of global annual median estimates for 1998-2005 is -14%.

Discussion

In the NASA standard ocean color processing, estimates of global annual median chlorophyll are quite sensitive to radiometric calibration. Simulated calibration errors of $\pm$1% produce changes in global medians of 20% to 25% for 555nm, and 13% to 15% for 443nm. These 2 bands represent the worst cases among the chlorophyll bands. These reported changes in global annual median chlorophyll compare unfavorably with the maximum observed interannual variability for the SeaWiFS 9-year chlorophyll record of about $\pm$3%. Even a minor change in radiometric calibration can produce large differences in global medians: a simulated $\pm$0.5% error at 555nm produces a 10 to 12% change in annual median. For 443nm this $\pm$0.5% error produces a 7% error. This represents 2 to 4 times the maximum observed interannual variability. Calibration errors of $\pm$0.1% are needed to meet the interannual variability target.

Considering that a radiometric calibration error of 0.1% corresponds to an error in normalized water-leaving radiance of approximately 1% (Table 6), our results suggest that a stricter standard is needed than the goal of 5% radiance uncertainty (Hooker and Esaias, 1993) using the NASA standard method. This assumes no contribution to uncertainty by atmospheric correction. Such radiometric uncertainty seems beyond the capability of modern methods, where calibration uncertainties of 0.57% to 0.85% are reported (see Figures 9 and 10) and radiance uncertainties of 11% to 16% have been observed (Franz et al., 2007). The calibration uncertainty produces an annual median chlorophyll error in the range of 4 to 7 times the maximum interannual variability. Considering the dependence of data products upon calibration and the current levels of uncertainty, the expenditure of effort on improved radiometric calibration is not misplaced. Our results suggest an alternative approach with ESRID, that takes advantage of empirical relationships using in situ observations.

ESRID is much less sensitive to errors in radiometric calibration. Simulated calibration errors of $\pm$1% in the 443 and 555nm wavelengths produce changes in global annual median chlorophyll within our target of <3%. This compares to a maximum 25% change using the NASA Standard method. Additionally, ESRID is capable of improving chlorophyll estimates from satellites. The bias of satellite-derived chlorophyll is reduced drastically compared to the NASA standard methodology, and there is a small reduction in uncertainty. The stability of ESRID for global median chlorophyll stands in contrast to the NASA standard methodology.

To place the sensitivity results in perspective, imagine ocean color without any post-launch radiometric re-calibration. Consider further, all the NASA ocean color missions flown to date. Without re-calibration, we would have to rely on pre-launch calibration. What would the impacts be on global annual median chlorophyll? Here we show that the change in global annual median chlorophyll, given the re-calibration adjustments for SeaWiFS determined by modern best efforts, exceeds 25% using the NASA Standard method (Figure 9). Post-launch radiometric re-calibration is essential to obtaining quality estimates of the global annual median. Using ESRID, however, a maximum change of only 2.1% is observed for SeaWiFS. In fact, only one case – the CZCS re-calibration adjustment at 555nm, produces a change in the global median that exceeds the interannual variability threshold we have defined here using ESRID (Figure 10). At 4.1%, it only exceeds the threshold by <2% in the global median. This compares to 64.5% for the NASA standard method. All other cases using ESRID: all the SeaWiFS bands, all the MODIS-Aqua bands (and Terra, although not shown), and all the CZCS bands except one, fall within our threshold without resorting to post-launch re-calibration.

How does ESRID achieve this reduction in bias and sensitivity to radiometric calibration errors? The main reason is the regression between satellite radiances and in situ chlorophyll. The coefficients a_0-a_4 in Eq. 9 in the best-fit polynomial absorb sensor bias, notably radiometric calibration. As simulated calibration errors change, the regression fit changes and the new coefficients absorb the new errors. Our use of median increments of log chlorophyll provides low bias over the entire range of chlorophyll, ensuring that our estimate of cumulative bias is not the result of error compensation of portions of the range. This enhances reliability of the results over the spectrum of chlorophyll values. The force-fit of satellite radiances to in situ chlorophyll drives ESRID to the same global median, except in the presence of very large sensor biases.

Level-3 Issues

Application of a bio-optical algorithm using Level-3 water leaving radiances is unconventional in ocean color data processing. Conventional methods apply the bio-optical algorithm as part of the Level-2 processing. Level-3 water-leaving radiances are averaged over 9km and over a day. Because of the non-linear nature of the bio-optical algorithm OC4, a fourth-order polynomial relationship between radiances (actually reflectances, from Eq. 7) and chlorophyll, the mean of radiances to derive chlorophyll is not equivalent to the mean of chlorophyll from Level-2 radiances. Despite this fact, analysis of our approach using Level-2 data (using the OC4 bio-optical algorithm) shows negligible differences in global and large scale oceanographic regional medians (Table 7).

A similar potential issue arises with the so-called near infrared (NIR)-correction. Siegel et al. (2000) suggested that at high chlorophyll concentration (> 2 mg m^{-3}), scattering by phytoplankton produced non-negligible upwelling radiances in the SeaWiFS NIR bands, at 765 and 865nm. If this radiance from the water was not accounted for, aerosol radiance estimates would be overestimated. They proposed an iterative methodology for removing this water-leaving radiance in the NIR bands (765 and 865nm), based on some assumptions of chlorophyll scattering properties. When chlorophyll concentrations exceed 2 mg m^{-3}, the iteration goes into effect, ending when two successive chlorophyll results are closely in agreement or after 10 tries, whichever comes first. The aerosol radiance is changed in the process, as is $[L_w(\lambda)]_N$. (Note that the Siegel et al. (2000) method is not actually used in the NASA standard processing, but the algorithm used, Stumpf et al., 2003, is similar in form and result).

Level-3 radiances have already accounted for NIR effects. Assuming the correction is valid, the effects should hold in application of the Level-3 radiances in ESRID. Moreover, the NIR algorithm does not depend strictly upon particular values of chlorophyll, but rather convergence to a repeat value (except of course the determination of when to implement the algorithm). It is worth pointing out that the algorithm is intended to apply only for chlorophyll concentrations >2 mg m^{-3}, which represents only about 2% of the chlorophyll in the global oceans. (In the NASA Standard processing, the correction is actually phased in around 0.7 mg m^{-3}). For these reasons, we consider the NIR issue minor in the application of ESRID.

Discontinuities between Open Ocean and Coastal Regions in ESRID

We performed a global survey of ESRID chlorophyll in coastal regions for Jun 2000 seeking evidence of discontinuities resulting from our bifurcated application. Very rarely do we visually observe clear discontinuities using our usual color scales, but they exist whether we can see them or not. When we focus in on small regions and adjust color scales to maximize the appearance of discontinuities, they are apparent. However, ESRID does not have to be applied in this two-domain coastal/open ocean distinction if discontinuities are a problem, especially if the transition zone between coastal and open ocean is a major focus. Coefficients for a global, single-domain application of ESRID are provided in Table 3. Statistics on the comparison with in situ data are provided in Table 8, along with the NASA Standard and ESRID two-domain. We note that this application provides similar results and statistics for the open ocean in our bifurcated approach, but coastal statistics approach the NASA standard processing results. For global analyses we prefer the reduced bias and uncertainty over the drawback of open ocean/coastal discontinuities.

How Much In Situ Data is Needed?

An estimate of the amount of in situ data needed for a reliable application of ESRID can be derived from our 50% withholding tests. Here, the amount of data used to derive the ESRID polynomial coefficients was reduced by half. This resulted in 2336 match-up data points used for the open ocean (compared to 4562) and 1782 for the coasts (compared to 3553). (Note that the total number of points for open ocean and coasts exceeds the total number of available match-up data, because of overlapping use of data between 100 and 200m bottom depth for development.)

We find that the global median changes a maximum of 1.4% in 2004 using the 50% withholding case compared to the no-withholding. The minimum difference was 0.7% in 2005. These values are well within the requirements for stability in global medians, as estimated from the SeaWiFS interannual variability, i.e., about 3%.

The question really is not how much in situ data is needed, but rather how it is distributed across the range of chlorophyll concentrations globally. The NOAA/NASA/AMT data have an overabundance at high concentrations (> 1 mg m^{-3}), but very little at low concentrations, especially at the lowest bracket (< 0.03 mg m^{-3}). This distribution contrasts with the global distributions observed by satellite (Table 1). What is needed is a more even distribution across the concentration range. Most immediately needed are more observations at low chlorophyll.

Implications of ESRID for Ocean Color

The ESRID results suggest that potentially profound changes in the implementation and management of ocean color are possible. Conventional ocean color data processing derives entirely from methods developed in the 1980's from the experience of the Coastal Zone Color Scanner. At that time, public repositories of in situ data did not exist and all satellite data procedures were required to be independent and self–contained. This is not to diminish the extraordinary strides in these procedures that have occurred since, including most relevantly radiometric calibration (Meister et al., 2003), but also bio-optical algorithms [both empirical (O'Reilly et al., 1998) and semi-empirical methods (Maritorena et al., 2002)], accounting for atmospheric effects in Rayleigh scattering and surface roughness (Wang, 2002), multiple scattering aerosols with interactions with Rayleigh scattering (Gordon and Wang, 1994), up front sensor characterization (Hooker and McClain 2000), and advanced methods to handle sensor idiosyncrasies, among many others. Rather our point is to acknowledge that these approaches had their genesis in the CZCS era. At the time that SeaWiFS and most other modern ocean color sensors were designed, knowledge of the nature and dynamics of global biological processes in the ocean was incomplete, and mission goals were chosen from the limited observations of the CZCS and the engineering capabilities of the late 1980's.

Given multi-year observations from SeaWiFS, we have our first opportunity to re-evaluate requirements. Our emphasis here is on global and large regional scale processes, which we consider the unique and primary purpose of satellite remote sensing. We show that at these scales, modern capabilities in radiometric calibration are inadequate. Specifically, for example, the 5% absolute water-leaving radiance target of SeaWiFS, which corresponds to about $\pm$0.5% calibration error, can lead to errors in global annual median chlorophyll 2 to 4 times the maximum interannual variability observed in the SeaWiFS 9-year record. Massive efforts and spending on radiometric calibration are not misplaced, considering the sensitivity of derived products on small changes.

These calibration efforts may not, however, be the most efficient and wisest choice to improve ocean color remote sensing. Our results with ESRID show that we can meet the requirements for global annual median chlorophyll stability with much larger radiometric uncertainties. Instead of relying upon more and better calibration, we achieve our objectives by utilizing the now extensive and high quality in situ data archives available to the general public. These archives, we emphasize, were not available until relatively recently.

The potentially profound management aspect that our ESRID results suggest is that it may be more beneficial, and certainly more straightforward, to intensify field campaigns collecting geophysical variables rather than new technological advances in radiometric calibration. At the end of the day, the field campaign intensification will increase our holdings of irreplaceable in situ observations, in addition to improving our ocean color observations.

ESRID essentially forces satellite data to agree with in situ data, within the limit of $\pm$1% calibration error. This is in contrast to the conventional processing, which only uses in situ data to evaluate satellite algorithm performance. The intimate, forced relationship of in situ and satellite data in ESRID produces maximum consistency between the archives, enhancing the value of both data holdings and promoting a unified description of ocean biology.

This is not to say that continued radiometric calibration efforts should be discontinued, or that conventional ocean color processing should be revamped. ESRID is intended to supplement ocean color processing, not supplant it. First, ESRID requires that the absolute calibration error be 1% or less. Second, ESRID requires that spatial and temporal variability of natural chlorophyll distributions be correctly represented in the first-time-through processing, which many of the advances listed above achieve. This means correcting for aerosols and Rayleigh scattering in space, and eliminating sensor-related trends in time. However, we note that with sufficient in situ data, ESRID could be applied on a temporal basis to remove trends. We note also that among all present and past NASA ocean color missions, post-launch radiometric re-calibration has only once changed the pre-launch calibration enough to matter for ESRID.

ESRID is capable of correcting or reducing any sensor/mission artifacts that result in a bias. This includes band-to-band misregistration, differences in band placement and width from sensor to sensor, most effects from out-of-band signals, and our focus here, radiometric calibration.

ESRID is not limited to chlorophyll. In principle, ESRID can be applied to any ocean surface variable that can be detected remotely by a visible signal. All that is needed is in situ observations of that variable, and good remotely-sensed representation of its radiometric spatial and temporal variability.

In the future, ESRID can play a role ameliorating the impacts of multi-purpose mission designs. For example, the next-generation ocean color sensor, the Visible and Infrared Imaging Radiometer Suite (VIIRS), is operated both as a real-time mission and a climate mission. These purposes conflict, and typically the real-time objectives dominate. Inadequate resources for re-processing data, shown to be critical for high quality observations needed for climate-relevant analyses, is a possibility. ESRID does not require re-processing in the conventional sense, where raw data must be re-computed from beginning to end. Instead, ESRID operates at Level-3, and can make the adjustments from there. ESRID can also play a role in relaxing the highly complex radiometric calibration requirements involved in proposed hyper-spectral sensors.

We acknowledge that by its nature ESRID will require a delay in data return while coincident in situ data are collected. But such a delay is inherent in the re-processing concept as currently defined as well. Vicarious calibration produced <17 radiometric observations y^{-1} for SeaWiFS (Franz et al., 2007), and <10 observations y^{-1} for MODIS. Considering how essential calibration is for the NASA Standard method, this suggests time on the order of years is required. ESRID may be able to provide quality products in less time, assuming the pre-launch calibration holds. Recall from the Methods that we suggest 100 observations at a median increment of 5. From the NOAA/NASA/AMT data record, the shortest amount of time to reach the required 500 observational matchups was Jun 10 to Sep 13 1999, a span of 95 days, or a little over 3 months.

Finally, ESRID holds the promise of providing a common foundation for constructing Climate Data Records (National Research Council, 2004) in ocean color. That foundation is in situ data, which in the case of chlorophyll, appears to retain its quality and consistency over decades in the public archives. The modern archives also extend well into the past to encompass historical missions, such as the CZCS. New missions can potentially take advantage of this approach, requiring an unprecedented cooperation between satellite and in situ observation organizations, enabling a common basis for understanding climatic changes in ocean geophysical products from the past to the future.

Acknowledgements

We thank the NASA SeaWiFS Project for in situ and satellite data. We also thank NODC for global in situ data and the British Oceanographic Data Centre for AMT in situ data. We also thank two anonymous reviewers for important remarks and suggestions. This work was supported by the NASA Carbon Cycle and REASoN programs.

References

Aiken, J., Rees, N., Hooker, S., Holligan, P., Bale, A., Robins, D., Moore, G., et al., 2000. The Atlantic Meridional Transect: Overview and synthesis of data. Progress in Oceanography 45: 257−312.

Antoine, D., A. Morel, H.R. Gordon, V.F. Banzon, and R.H. Evans, 2005. Bridging ocean color observations of the 1980a and 2000s in search of long-term trends. Journal of Geophysical Research 110: C06009, doi: 101029/2004JC002620.

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

Bailey, S.W., Hooker, S.B., Antoine, D., Franz, B.A., and Werdell, P.J., 2008. Sources and assumptions for the vicarious calibration of ocean color satellite observations. Applied Optics 47: 2035-2045.

Campbell, J.W., 1995. The lognormal distribution as a model for bio-optical variability in the sea. Journal of Geophysical Research, 100, 13237-13254.

Conkright, M.E., Antonov, J.I., Baranova, O., Boyer, T.P., Garcia, H.E., Gelfeld, R., Johnson, D., O'Brien, T.D., Smolyar, I., Stephens, C., 2002. World ocean database 2001, Vol. 1: Introduction. S. Levitus, Ed., NOAA Atlas NESDIS 42, US Govt. Printing Office, Washington, DC, 167 pp.

Evans, R.H. and H.R. Gordon, 1994. Coastal Zone Color Scanner "system calibration": A retrospective examination. Journal of Geophysical Research 99**:** 7293-7307.

Franz, B.A., S.W. Bailey, P.J. Werdell, and C.R. McClain, 2007. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. Applied Optics 46: 5068-5082.

Gordon, H.R., 1990. Radiometric considerations for ocean color remote sensors. Applied Optics 29: 3228-3236.

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

Gregg, W.W. and M.E. Conkright, 2001. Global seasonal climatologies of ocean chlorophyll: Blending in situ and satellite data for the Coastal Zone Color Scanner era. Journal of Geophysical Research, 106: 2499-2515.

Hooker, S.B. and W.E. Esaias, 1993. An overview of the SeaWiFS project. Eos, Transactions of the American Geophysical Union. 74: 241-246.

Hooker, S.B., and McClain, C.R., 2000. The calibration and validation of SeaWiFS data. Progress in Oceanography 45: 427-465.

Hooker, S.B., McClain, C.R., and Mannino, A., 2007. NASA strategic planning document: A comprehensive plan for the long-term calibration and validation of oceanic biogeochemical satellite data. NASA Technical Memorandum SP-2007-214152, 31 pp.

Maritorena, S., Siegel, D.A., and Peterson, A.R., 2002. Optimization of a semianalytical ocean color model for global-scale applications. Applied Optics 41: 2705-2714.

McClain, C.R., Feldman, G.C., and Hooker, S.B., 2004. An overview of the SeaWiFS project and strategies for producing a climate reearch quality global ocean bio-optical time series. Deep-Sea Research II 51: 6-42.

Meister, G., Abel, P., Barnes, R., Cooper, J., Davis, C., Fargion, G., Frouin, R., Godin, M., Korwan, D., Maffione, R., McClain, C., McLean, S., Menzies, D., Poteau, A., Robertson, J., and Sherman, J., 2003. Comparison of spectral radiance calibrations at oceanographic and atmospheric research laboratories. Metrologia 40**:** S93–S96.

NRC, 2004. Climate data records from environmental satellites. National Academies Press, Washington, DC, 136 pp.

O'Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder, S.A. Garver, M. Kahru, and C. McClain, 1998. Ocean color chlorophyll algorithms for SeaWiFS. Journal of Geophysical Research 103**,** 24937-24953.

O'Reilly, J.E., and 21 others, 2000. Ocean color chlorophyll a algorithms for SeaWiFS, OC2 and OC4: Version 4. NASA Technical Memorandum 2000-206892, Vol. 11, 49 pp.

Reynolds, R.W., 1988. A real-time global sea surface temperature analysis. Journal of Climate 1**:** 75-86.

Robinson, W.D., Schmidt, G.M., McClain, C.R., and Werdell, P.J., 2000. "Changes made in the operational SeaWiFS processing," In: SeaWiFS postlaunch calibration and validation analyses, Part 3, S.B. Hooker and E.R. Firestone, eds., NASA Tech. Memo. 2000-206892, Vol. 10, pp. 12-28.

Siegel, D.A., M. Wang, S. Maritorena, and W. Robinson, 2000. Atmospheric correction of satellite ocean color imagery: The black pixel assumption. Applied Optics 39: 3582-3591.

Stumpf, R. P., R. A. Arnone, R. W. Gould, P. M. Martinolich, and V. Ransibrahmanakul, 2003. A partially-coupled ocean-atmosphere model for retrieval of water-leaving radiance from SeaWiFS in coastal waters. In: NASA Technical Memorandum Vol. 22, Algorithm updates for the fourth SeaWiFS data reprocessing.

Wang, M., 2000. The SeaWiFS atmospheric correction algorithm updates, In: SeaWiFS postlaunch calibration and validation analyses, Part 1, S.B. Hooker and E.R. Firestone, eds., NASA Tech. Memo. 2000-206892, Vol. 9, pp. 57-63.

Wang, M., 2002. The Rayleigh lookup tables for the SeaWiFS data processing:

accounting for the effects of ocean surface roughness. International Journal of Remote Sensing 23: 2693-2702.
Werdell, P.J. and 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.

Appendix: Specific Implementation of ESRID

Here we describe the steps to implement ESRID, and provide the location of all code and data used for ESRID for SeaWiFS chlorophyll.

1) The first step is to acquire in situ chlorophyll data and SeaWiFS radiance data. The in situ data set has been described in the methods, and SeaWiFS Level-3 radiances are available at the NASA Ocean Color Web.
2) The second step is to associate the satellite radiances with the in situ chlorophyll, at daily, 9km resolution, creating matchups. We simultaneously convert to reflectances at the stage in preparation for the next step. This process is contained in the Fortran program matchdatafo.f.
3) Third, we compute the median of reflectances over small increments of in situ chlorophyll, performed in medrat.f, and fit a non-linear regression to the results. The non-linear regression we use is from the Interactive Data Language package, and specific routines are pltmedrat.pro and poly4.pro. This provides coefficients for a fourth-order polynomial that obeys the MBR method as described in O'Reilly et al (1998). Care is required to ensure a monotonic best-fit regression is obtained, if there is a double-back in the regression line, one will obtain multiple values for a given reflectance ratio. We avoid this by changing the minimum number of observations for a given chlorophyll increment. We find that this problem occurs at the high end of the chlorophyll range, and so only apply this modified minimum observation number for high values. One can check for this problem using the program polyratmatch4.f.
4) Once the coefficients for the polynomial are obtained, they can be applied to the NASA Level-3 radiances to produce new chlorophyll data. This occurs in the program adjchl.f. Of course, in our implementation steps 2-3 are executed separately for open ocean and coastal regions, but the code provided does not reflect this bifurcation for simplicity. All code and data are available at the NASA/GMAO web site (search for GMAO), under research, modeling, ocean biology modeling. Use of the code and data is unsupported.

Table 1. Frequency distributions of SeaWiFS and in situ data over 6 chlorophyll brackets. Frequencies are shown in percent and the brackets are defined as log chlorophyll. Un-transformed chlorophyll corresponding to the log brackets is shown in parenthesis.

Bracket	$mg\ m^{-3}$	SeaWiFS	In situ
[-2.0 to -1.5]	(0.01 to 0.03)	0.87%	1.70%
[-1.5 to -1.0]	(0.03 to 0.1)	24.86%	18.67%
[-1.0 to -0.5]	(0.1 to 0.3)	54.36%	26.22%
[-0.5 to 0.0]	(0.3 to 1.0)	14.66%	20.75%
[0.0 to 0.5]	(1.0 to 3.2)	3.81%	20.35%
[0.5 to 2.0]	(3.2 to 100.0)	1.45%	12.31%

Table 2. Levels of ocean color data processing.

Level-0	Raw uncorrected digital counts from the sensor
Level-1A	Level-0 sensor data with orbit and sensor information appended
Level-2	Radiometrically calibrated, atmospherically-corrected water-leaving radiances and geophysical products (e.g., chlorophyll) in the satellite coordinate system (orbit track and scan)
Level-3	Level-2 data placed on an Earth grid (longitude and latitude). Ocean color standard is 4320x2160, or about 9km resolution at the equator, equal angle for Standard Mapped Images (SMI)
ESRID	Apply regression statistics on Level-3 SMI using in situ data, and the median of reflectances over small chlorophyll increments

Table 3. Coefficients for ESRID

	a0	a1	a2	a3	a4
Coastal	0.3887	-4.0901	1.7775	4.9532	-5.2839
Open Ocean	0.4387	-3.8499	4.3706	-2.4844	-0.6622
Global	0.4393	-3.6461	1.6246	4.0033	-4.8224

Table 4. Statistics for NASA Project method and a Level-3 regression against in situ data where the data are unweighted (i.e., conforming to in situ weighting). Log error for bias and uncertainty (RMS) are shown in parenthesis. Slope and r^2 are for log statistics.

	Bias	Uncertainty	Slope	r^2	N
NASA Project	16.7% (0.073)	44.5% (0.305)	0.85	0.791	5994
Level-3 Regression	0.4% (0.000)	38.2% (0.291)	0.79	0.794	5956

Table 5. Response of the NASA Standard method and ESRID to radiometric calibration errors at 490 and 510nm, as indicated by change in global annual median chlorophyll from a reference (the unadjusted global median). Data for 443nm and 555nm are not shown here because they are shown in Figures 7 and 8.

Error	**10%**	**5%**	**1%**	**0.5%**	**0.1%**
490nm					
NASA Standard	-52.2%	-29.7%	-4.8%	-2.0%	-0.7%
ESRID	46.2%	12.5%	2.1%	1.4%	0.0%
510nm					
NASA Standard	-35.7%	-14.7%	-0.7%	0.0%	0.0%
ESRID	122.9%	31.8%	2.1%	0.7%	0.0%
Error	**-10%**	**-5%**	**-1%**	**-0.5%**	**-0.1%**
490nm					
NASA Standard	2.8%	2.8%	1.4%	0.7%	2.1%
ESRID	0.0%	0.0%	0.0%	-0.7%	0.0%
510nm					
NASA Standard	0.0%	0.0%	0.0%	0.0%	0.0%
ESRID	0.0%	0.7%	1.4%	0.7%	0.0%

Table 6. Relationship between a simulated calibration error of 1% in at-satellite radiance $L_t(\lambda)$ and normalized water-leaving radiances $[L_w(\lambda)]_N$. The evaluation was performed for Jan and Jun 2001. The mean ratio $\Delta[L_w(\lambda)]_N:\Delta L_t(\lambda)$ is shown in percent, along with standard deviation (sd). N = 8x10^7 for all bands except 670nm, where N = 6x10^7.

Band	Ratio $[\Delta[L_w(\lambda)]_N:\Delta L_t(\lambda)]\% \pm$ sd
410nm	14.6% (±26.3)
443nm	10.4% (±19.3)
490nm	7.1% (±11.2)
510nm	7.6% (±11.0)
555nm	12.3% (±16.7)
670nm	42.9% (±50.8)

Table 7. Global annual median chlorophyll and SIQR. Level-2 applies OC4 at Level-2 and then bins to Level-3. Level-3 applies OC4 in directly at Level-3. The largest difference by major oceanographic basin between the Level-2 and Level-3 processing is also shown.

	Global Median		Global SIQR		
	Level-2	Level-3	Level-2	Level-3	Maximum Difference, Region

1998	0.2018	0.2018	0.123	0.121	-1.30% South Pacific
1999	0.2118	0.2118	0.131	0.130	+1.39% North Atlantic
2000	0.2118	0.2118	0.133	0.130	+2.09% North Atlantic
2001	0.2089	0.2089	0.132	0.129	+2.10% North Pacific
2002	0.2004	0.2004	0.125	0.122	+2.09% North Pacific
2003	0.2061	0.2061	0.128	0.127	+2.09% North Indian
2004	0.2061	0.2046	0.128	0.127	+2.10% North Atlantic
2005	0.2061	0.2046	0.134	0.131	+2.10% North Pacific

Table 8. Statistics on the comparison of ESRID with in situ data, as it is applied separately for open ocean and coastal regions, contrasted with a single application globally. The NASA Standard statistics are shown for comparison. The statistics for ESRID-two-domain and NASA Standard have previously been shown in Figures 4 and 6. There is no withholding of data for these statistics.

	Coastal		Open Ocean		Global	
	Bias	Uncertainty	Bias	Uncertainty	Bias	Uncertainty
NASA Standard	75.0%	97.3%	18.6%	34.8%	28.0%	40.6%
ESRID	12.4%	59.5%	-1.8%	34.3%	0.7%	37.3%
ESRID global	49.4%	81.8%	-4.3%	31.5%	4.9%	37.2%

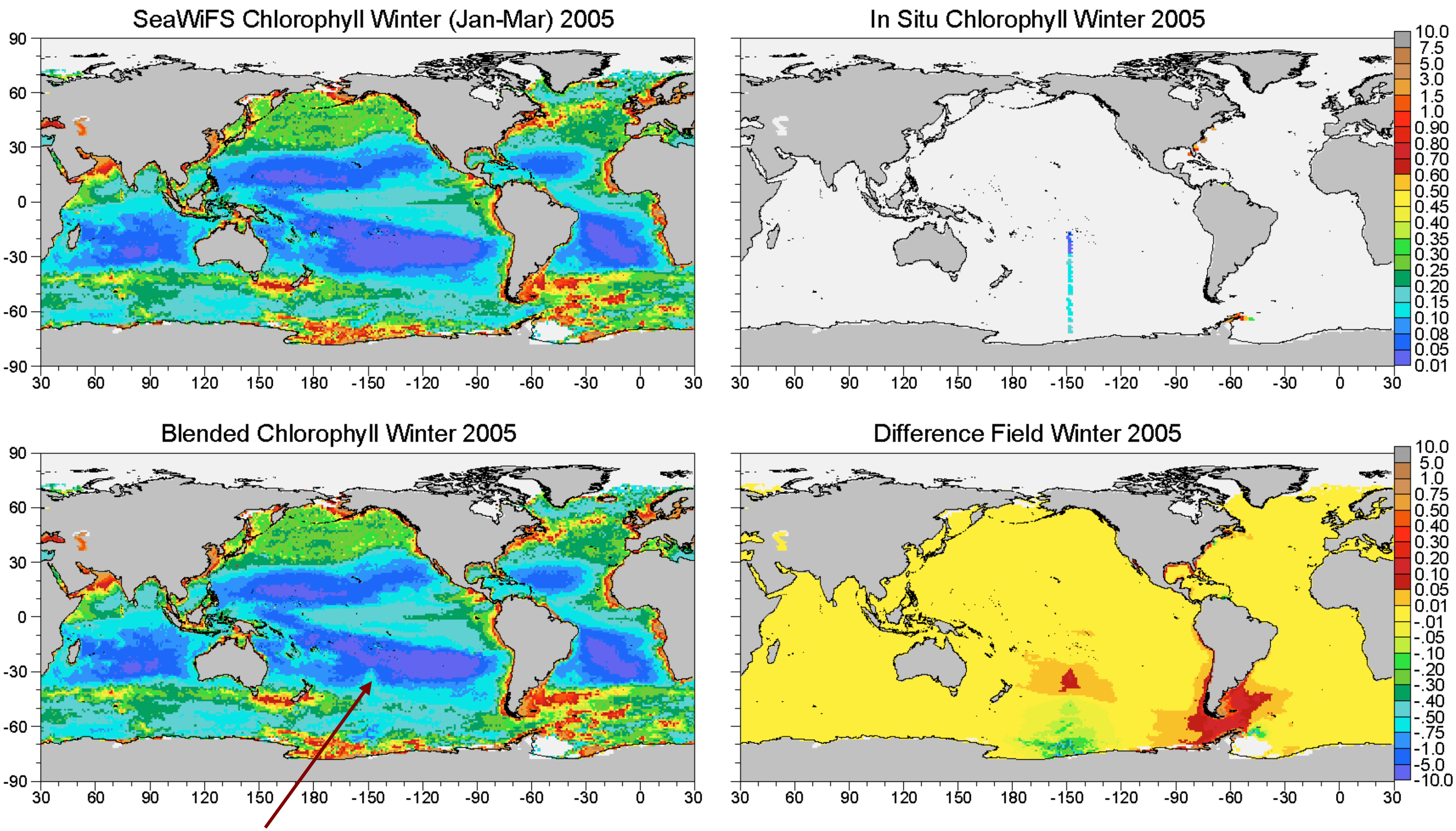


Figure 1. Example of the Blended Analysis for in situ data quality control. A north-south set of data from a single cruise in winter 2005 produced major distortion of the biological structure in the South Pacific. Adjustments from the Blended Analysis showed the cruise track in the blended chlorophyll, denoted by the arrow. The perturbations are apparent in the difference field, and align with the cruise track.

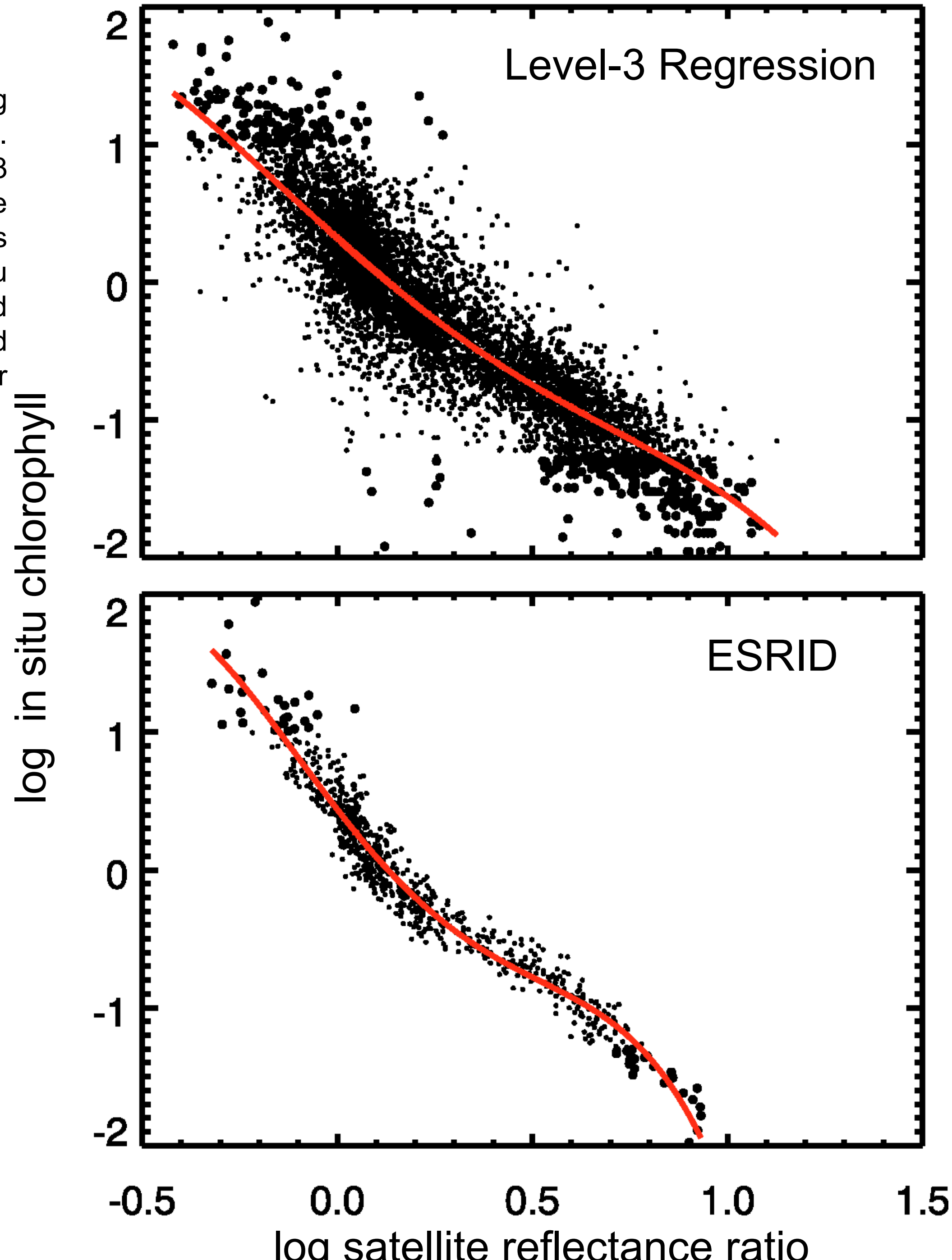


Figure 2. Scatterplot of satellite log reflectance ratio and log in situ chlorophyll. Top shows a simple non-linear Level-3 regression of in situ chlorophyll to satellite reflectance (solid red line). Bottom is ESRID with its median over small log in situ chlorophyll segments and the associated regression line. Matchup values at low and high chlorophyll are enhanced for increased visibility.

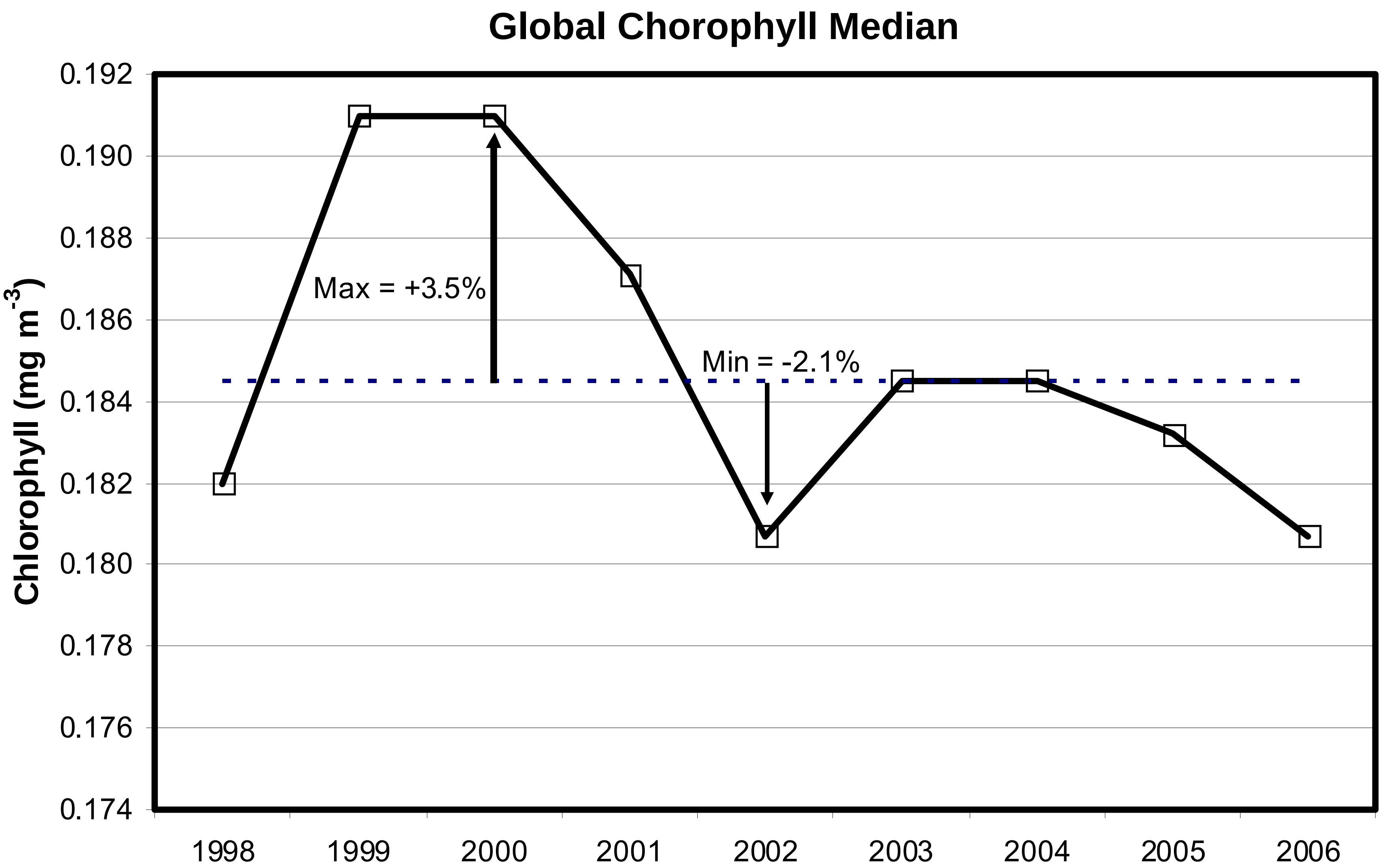
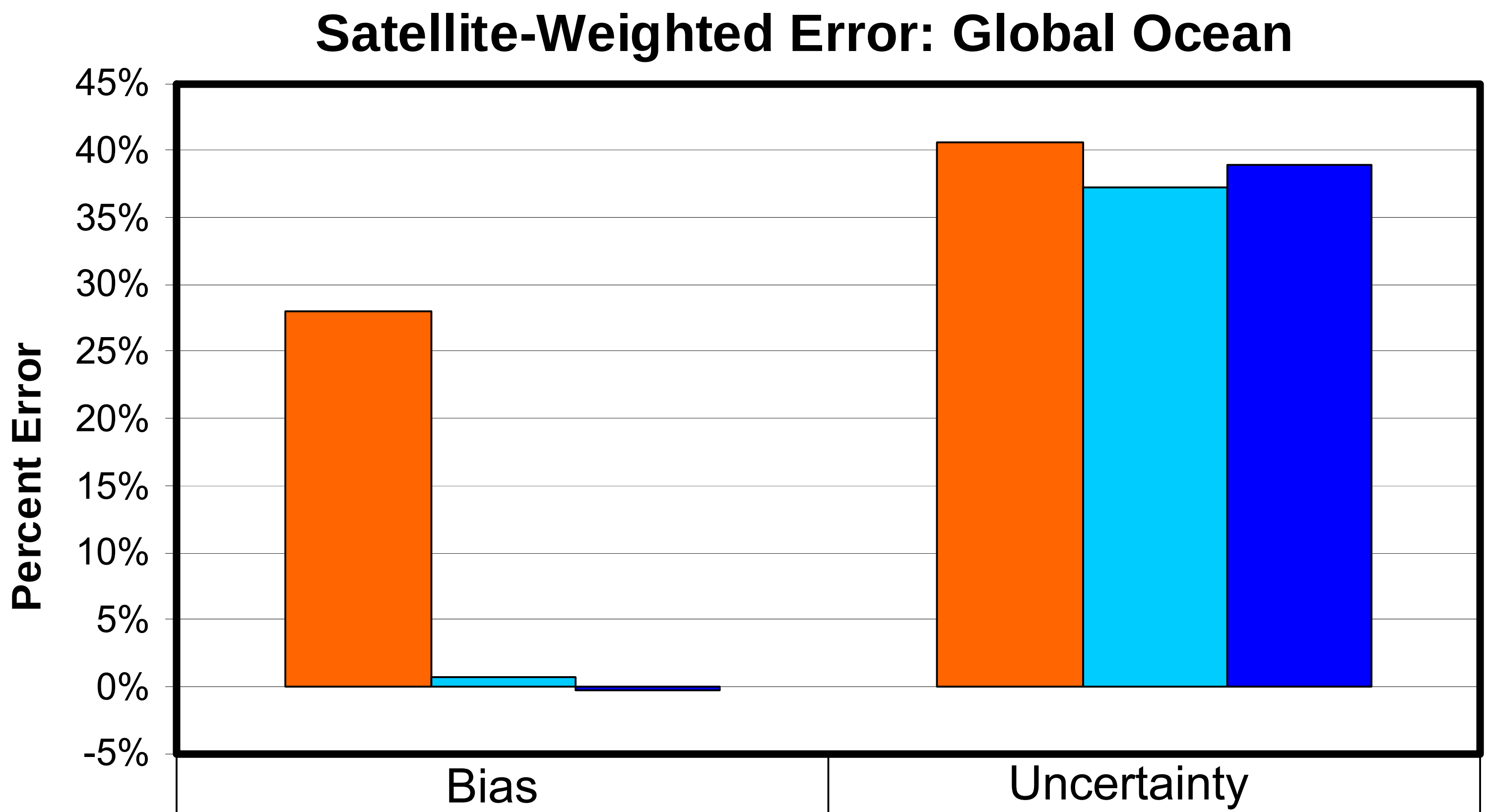


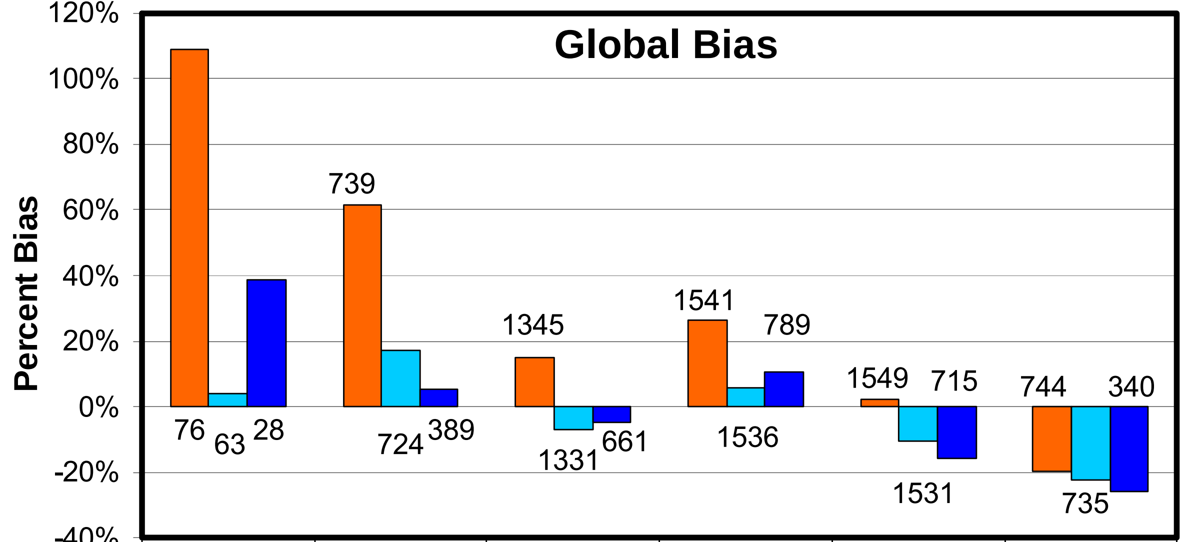
Figure 3. Global annual median global chlorophyll for the 9-year chlorophyll time series for SeaWiFS. The dashed line represents the climatological median. Minimum and maximum departures from the climatology are shown.



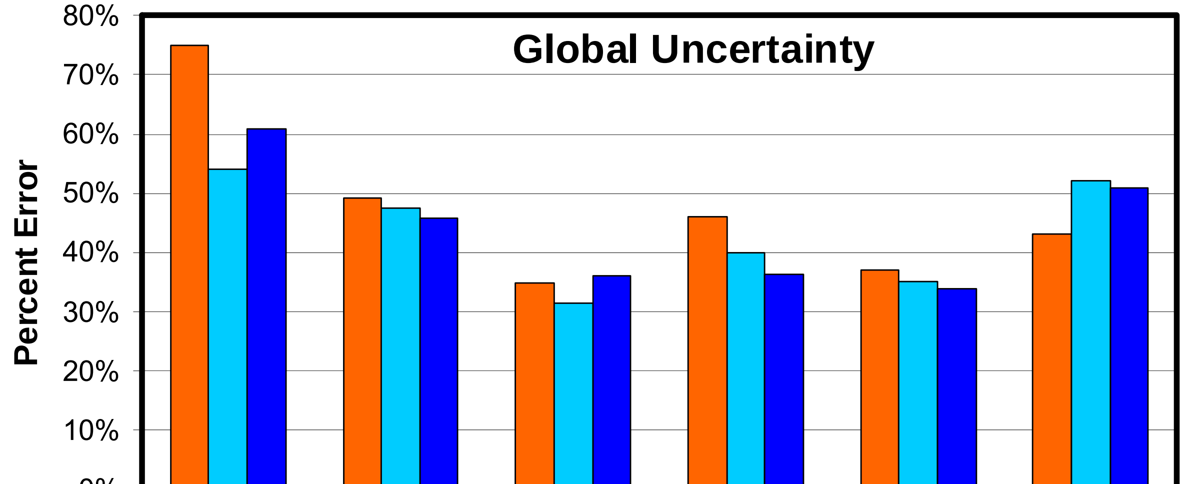
	Bias			Uncertainty	
SeaWiFS	N=5994	28.0%	(0.130)	40.6%	(0.298)
ESRID 0% w/h	N=5920	0.7%	(0.020)	37.3%	(0.291)
ESRID 50% w/h	N=2922	-0.3%	(0.028)	38.9%	(0.291)

Figure 4. Satellite- weighted error statistics (bias and uncertainty) for the global ocean. Statistics from SeaWiFS are compared with two estimates from ESRID: one with no withholding of data (0% w/h), representing a lower bound of the error estimate, and one with 50% withholding (50% w/h), representing an upper bound. The statistics are shown in the data table below the figures, with log error provided in parentheses.

Figure 5. Satellite- weighted error statistics for the global ocean, shown individually for 6 different chlorophyll brackets, indicated as log chlorophyll. Statistics from SeaWiFS are compared with two estimates from ESRID: one with no withholding of data (0% w/h),, and one with 50% withholding (50% w/h). The statistics are shown in the data table below the figures, with log error provided in parentheses. N is listed by the appropriate column in the bias plot.



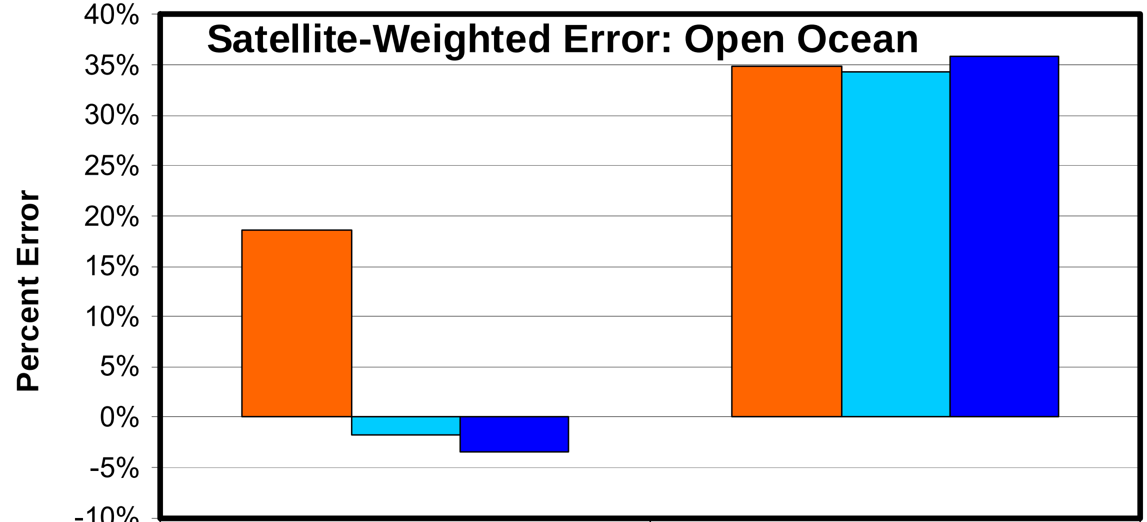
	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to 0.0	0.0 to 0.5	0.5 to 2.0
SeaWiFS	108.8%	61.5%	15.0%	26.2%	2.2%	-19.7%
ESRID 0% w/h	4.0%	17.1%	-7.0%	5.9%	-10.3%	-22.1%
ESRID 50% w/h	38.6%	5.4%	-4.7%	10.6%	-15.8%	-25.6%



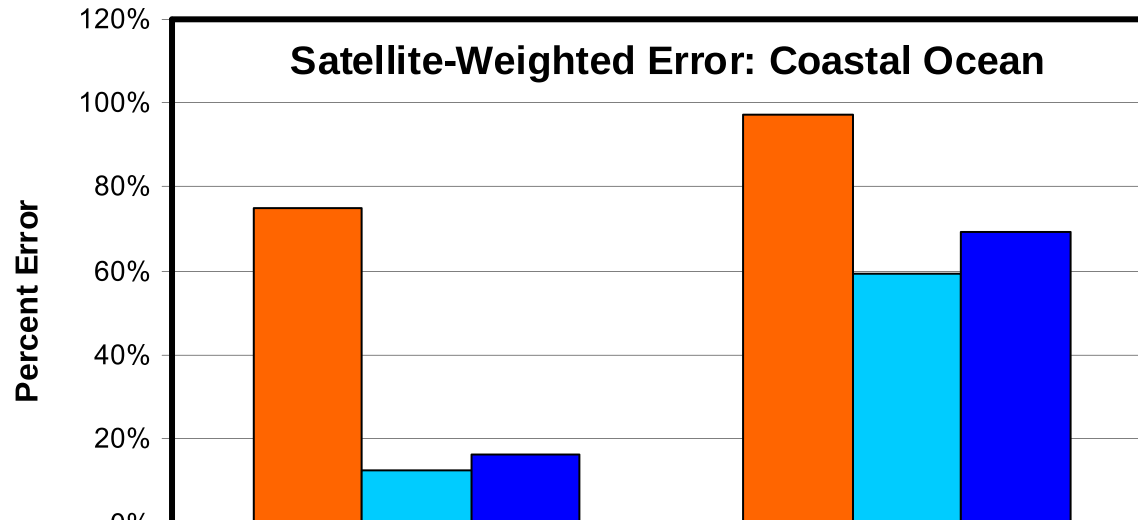
	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to 0.0	0.0 to 0.5	0.5 to 2.0
SeaWiFS	74.8%	49.2%	34.8%	46.1%	37.1%	43.1%
ESRID 0% w/h	54.0%	47.6%	31.4%	39.9%	35.0%	52.2%
ESRID 50% w/h	60.9%	45.9%	36.1%	36.4%	33.8%	50.9%

Chlorophyll Bracket (log units)

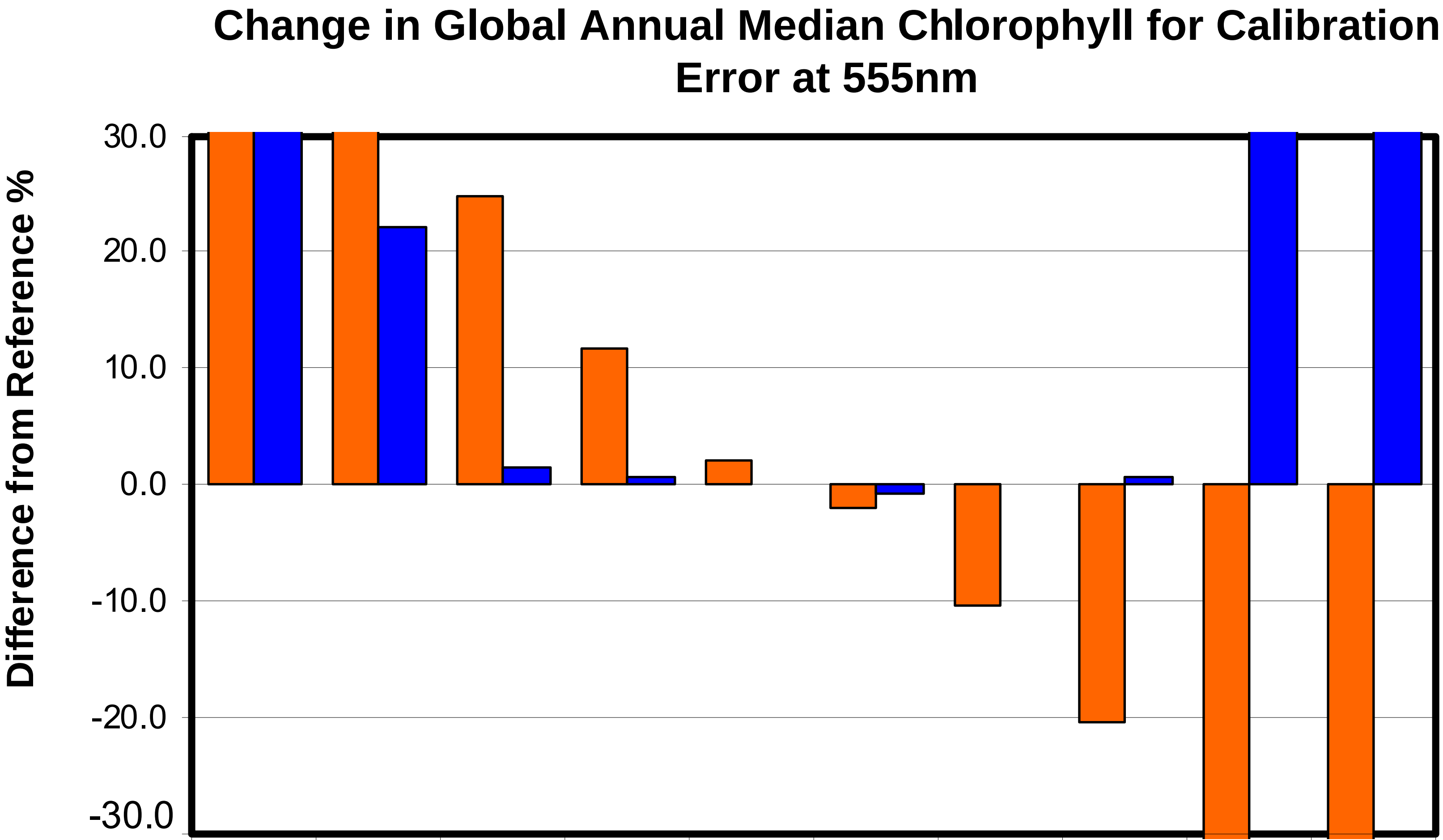
Figure 6. Satellite- weighted error statistics for the open ocean (bottom depth ≥ 200m) and the coastal ocean (depth < 200m). Statistics include no withholding of data (0% w/h), and 50% withholding (50% w/h). The statistics are shown in the data table below the figures, with log error provided in parentheses.



	Bias			Uncertainty	
SeaWiFS	N=3216	18.6%	(0.077)	34.8%	(0.249)
ESRID 0% w/h	N=3161	-1.8%	(-0.006)	34.3%	(0.261)
ESRID 50% w/h	N=1558	-3.4%	(0.001)	35.8%	(0.262)

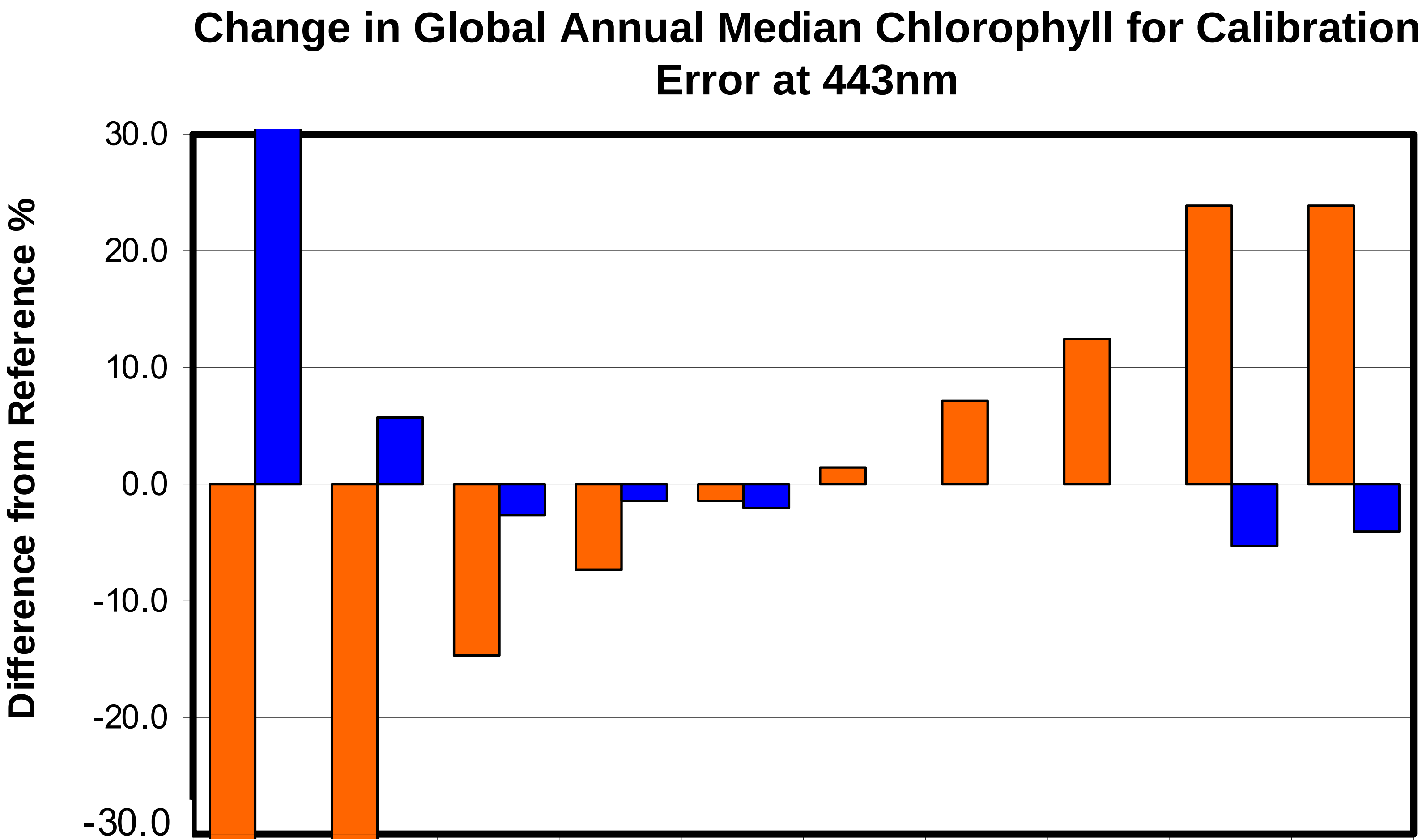


	Bias		Uncertainty	
SeaWiFS	75.0%	(0.314)	97.3%	(0.471)
ESRID 0% w/h	12.4%	(0.132)	59.5%	(0.414)
ESRID 50% w/h	16.4%	(0.134)	69.1%	(0.394)



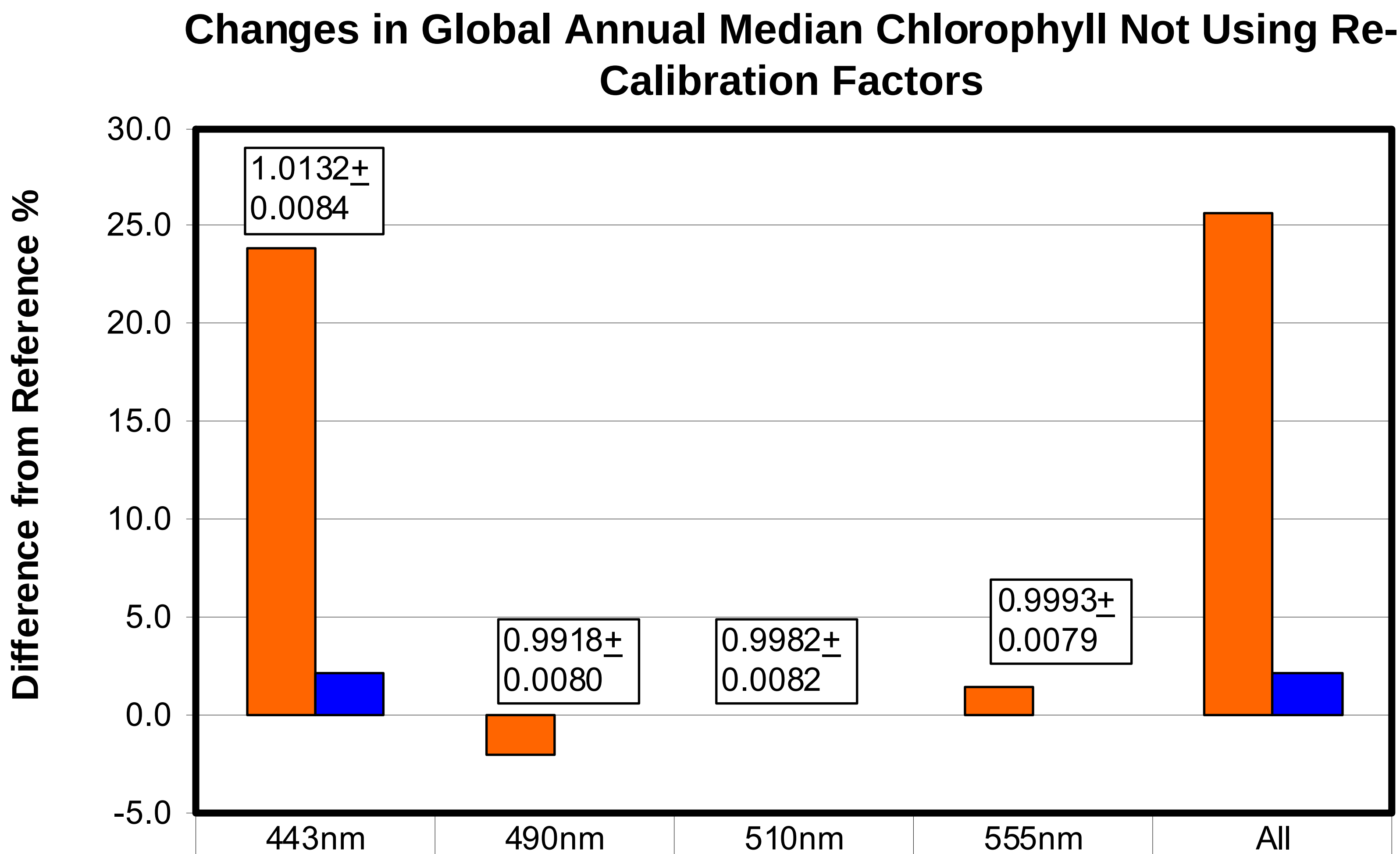
Calibration Error	10%	5%	1%	0.5%	0.1%	-0.1%	-0.5%	-1%	-5%	-10%
NASA Standard	722.1	223.6	24.7	11.7	2.1	-2.0	-10.5	-20.4	-76.7	-93.7
ESRID	38.4	22.2	1.4	0.7	0.0	-0.7	0.0	0.7	60.0	77.4

Figure 7. Sensitivity of global annual median chlorophyll to radiometric calibration for the NASA Standard method and ESRID. Simulated calibration errors are introduced at 555nm from 10% to -10%. The percent change in the global median chlorophyll is determined from a reference where the calibration is not adjusted. To meet the goal of a global median change <3% (maximum interannual variability observed in SeaWiFS), the NASA standard method requires an absolute calibration error $\leq$0.1%, while ESRID only requires an absolute error $\leq$1%.



Calibration Error	10%	5%	1%	0.5%	0.1%	-0.1%	-0.5%	-1%	-5%	-10%
NASA Standard	-70.8	-50.2	-14.7	-7.3	-1.4	1.4	7.1	12.5	23.9	23.9
ESRID	54.5	5.7	-2.7	-1.4	-2.0	0.0	0.0	0.0	-5.4	-4.1

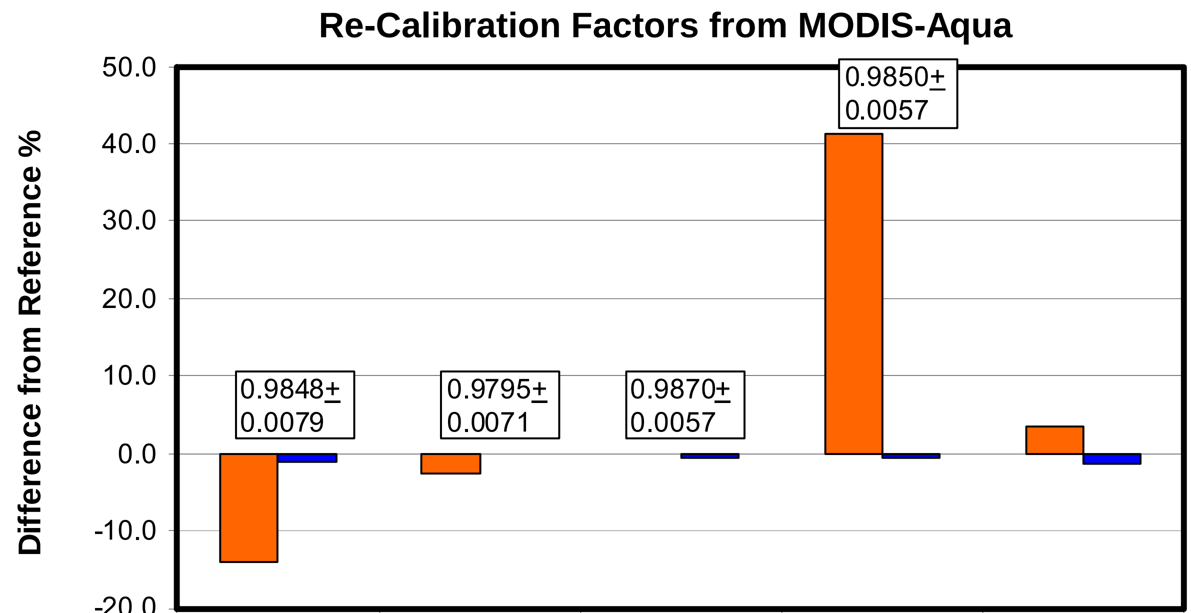
Figure 8. Sensitivity of global annual median chlorophyll to radiometric calibration for the NASA Standard method and ESRID. Simulated calibration errors are introduced at 443nm from 10% to -10%. The percent change in the global median chlorophyll is determined from a reference where the calibration is not adjusted. To meet the goal of a global median change <3%, the NASA standard method requires an absolute calibration error $\leq$0.1%, while ESRID only requires an absolute error $\leq$1%. This is the same as the requirements for calibration errors at 555nm.



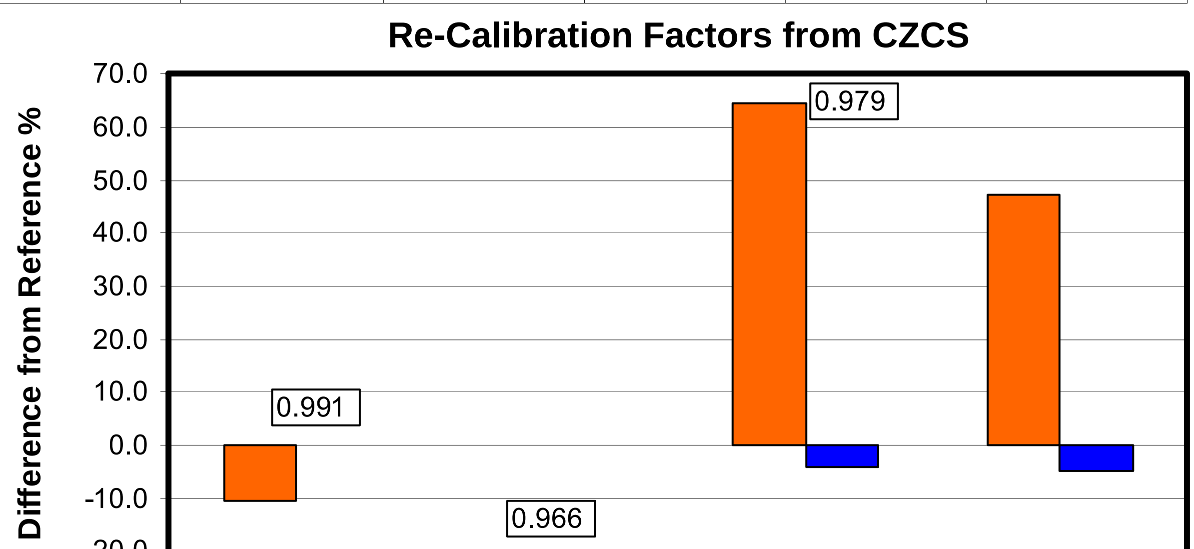
	443nm	490nm	510nm	555nm	All
NASA Standard	23.9	-2.0	0.0	1.4	25.6
ESRID	2.1	0.0	0.0	0.0	2.1

Figure 9. Changes in global annual median chlorophyll from not using the re-calibration factors (often referred to as vicarious gains) for SeaWiFS. The re-calibration factors ± standard deviation are shown above the data for each wavelength. N=147 (re-calibration data from NASA Ocean Color Web). The "All" column uses the ensemble of re-calibration factors for all bands.

Figure 10. Changes in global annual median chlorophyll from not using re-calibration factors derived for MODIS-Aqua and CZCS for nearest bands for SeaWiFS. This is not the sensitivity of MODIS and CZCS, but the sensitivity of SeaWiFS not using re-calibration factors derived from MODIS and CZCS. The re-calibration factors and standard deviations are shown. Standard deviations and N for CZCS was not available. N for MODIS-Aqua = 39.



	443nm	490nm	510nm	555nm	All
NASA Standard	-14.1	-2.7	0.0	41.3	3.5
ESRID	-1.0	0.0	-0.7	-0.7	-1.4



	443nm	510nm	555nm	All
NASA Standard	-10.5	0.0	64.5	47.2
ESRID	0.0	0.0	-4.1	-4.7

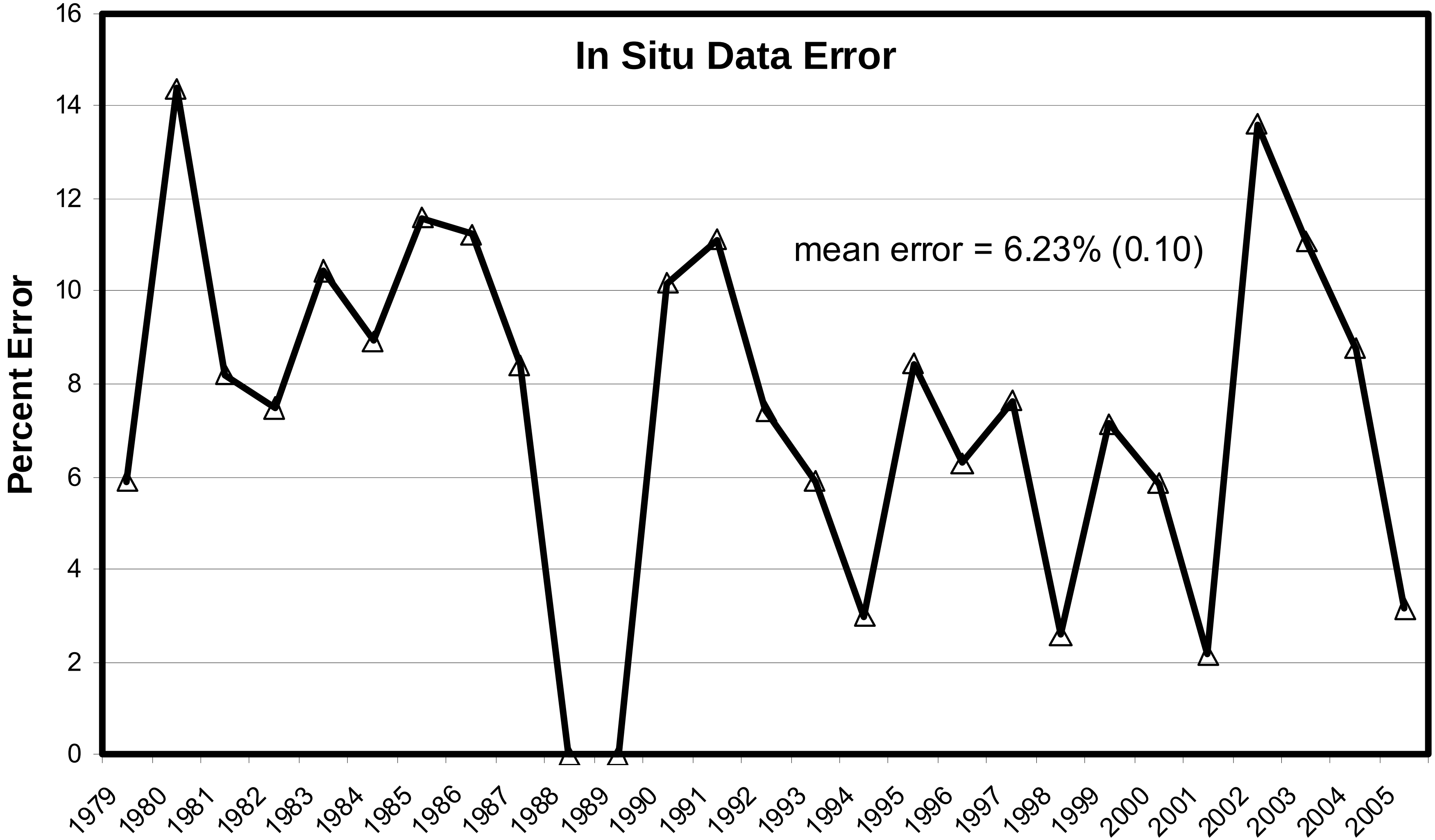


Figure 11. In situ data error estimated as the SIQR in percent at all 9km grid locations where 2 or more observations occurred. The Mean error over all years is shown with log RMS error in parentheses. N = 27.

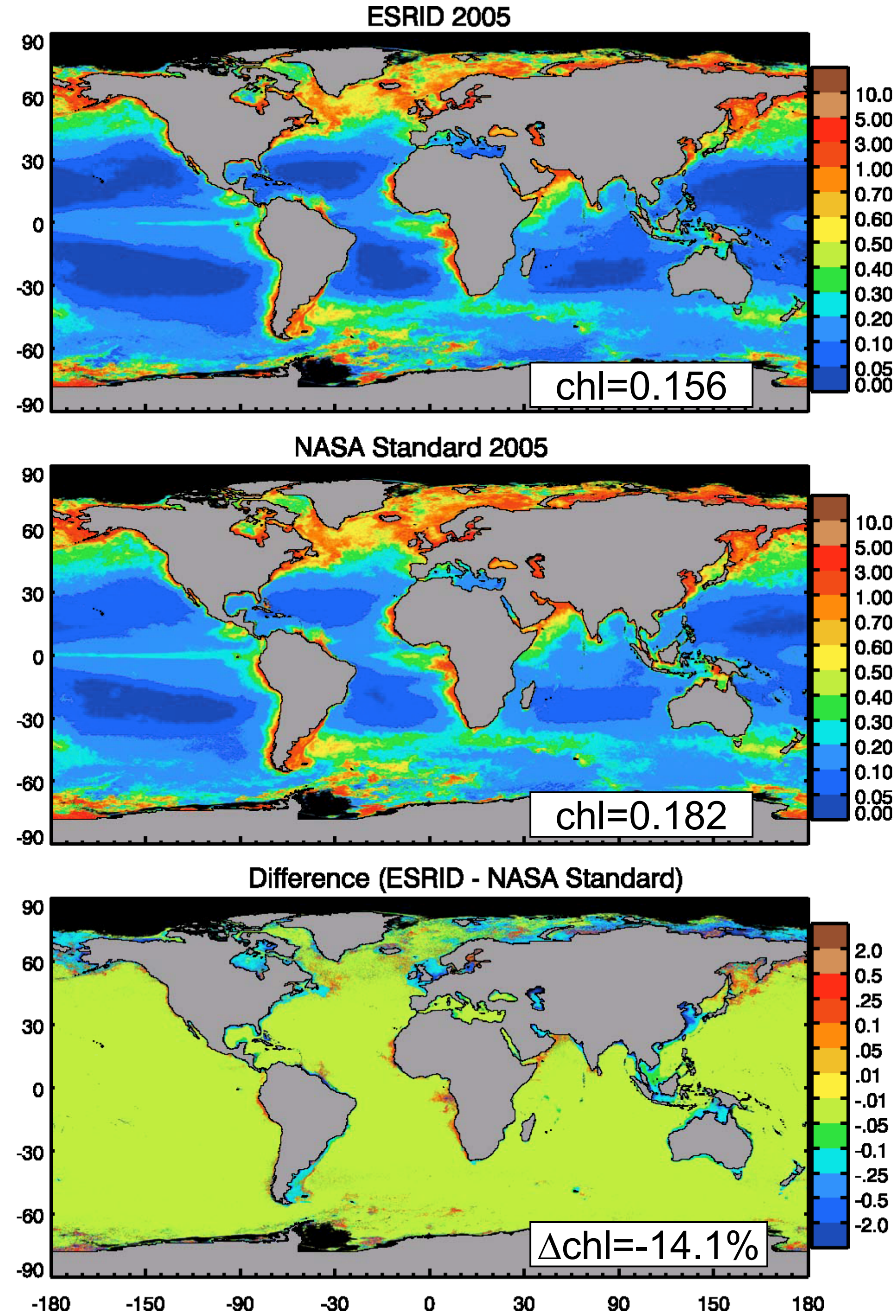


Figure 12. Global annual mean chlorophyll from SeaWiFS using ESRID, the NASA standard processing, and the difference for 2005. The global median chlorophyll values are shown in each plot, in units of mg m^{-3}, with the difference in percent in the difference plot.

An Empirical Approach to Ocean Color Data: Reducing Bias and the Need for Post-Launch Radiometric Re-Calibration

Authors:
Watson Gregg, Watson.Gregg@nasa.gov
Global Modeling and Assimilation Office

Nancy W. Casey, nancy_casey@ssaihq.com
Science Systems and Applications, Inc.

John E. O'Reilly, joreilly@mola.na.nmfs.gov
NOAA/National Marine Fisheries Service Narragansett Laboratory

Wayne E. Esaias, wayne.esaias@nasa.gov
Laboratory for Hydrospheric and Biospheric Processes

Submitted to: Remote Sensing of Environment 2009

Popular Summary:

Radiometric calibration is the foundation of remotely-sensed ocean color data. Calibration is strictly defined as a factor applied to at-satellite signals to bring them into agreement with some measure of truth. Specifically, it is the conversion of electronic responses from the sensor (digital counts) into radiometric units (typically mW cm^{-2} μm^{-1} sr^{-1}). Calibration involves identifying a single factor for each wavelength that represents the conversion from counts to radiances. Thus it is a spatially and temporally constant factor, varying only with wavelength, intended to remove bias. For modern NASA ocean color missions, the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer (MODIS), this is derived from pre-launch testing, and is refined on orbit using a single location near Hawaii.

The quality of derived ocean color geophysical products (mainly chlorophyll, but others are considered) is ultimately dependent upon the radiometric calibration. Yet the radiometric goals have not been achieved. Even using carefully selected in situ data sets for comparison, recent results show a bias in spectral normalized water-leaving radiances ranging from 0.2% at 412nm to 35% at 670 nm, and uncertainty ranging from 11% at 510nm to 65% at 670 nm. The inability to meet the mission goals, despite strong efforts in post-launch re-calibration technology and in situ radiometry, has major repercussions for the derived geophysical products. For the first time, we have an understanding of the variability of ocean chlorophyll from the decade-long observational record, and what uncertainty is needed. The 9-year record of SeaWiFS (1998-2006) showed that the minimum and maximum departures from the global chlorophyll median were -2.1% and +3.5%, a maximum range of 5.6%. We suggest first that global and large scale regional estimates of products is the most important objective of satellite observations, as it cannot be done with any other platform, at least at the present time. We suggest second that the error of derived chlorophyll should be less than the observed natural departures if we are to support observations of climate change and variability, especially when we consider time-series that span >1 mission. We will show here that the present radiometric goals are not sufficient to meet this condition using the algorithm and processing sequence as currently configured by the NASA project.

We introduce the Empirical Satellite Radiance-In situ Data (ESRID) algorithm, a new approach for producing derived geophysical products. In doing so we follow a path to lead us to a fundamental question of quantitative satellite ocean color: what is the global median chlorophyll?

The results suggest that ESRID 1) drastically reduces the bias of ocean chlorophyll, most impressively in coastal regions, 2) modestly improves the uncertainty, and 3) reduces the sensitivity of global annual median chlorophyll to changes in radiometric re-calibration. Simulated calibration errors of 1% or less produce small changes in global median chlorophyll (<2.7%). In contrast, the standard NASA algorithm set is highly sensitive to radiometric calibration: similar 1% calibration errors produce changes in global median chlorophyll up to nearly 25%.

The results suggest ocean color programs might consider less emphasis of expensive efforts to improve post-launch radiometric re-calibration in favor of increased efforts to characterize in situ observations of ocean surface geophysical products. Although the results here are focused on chlorophyll, in principle the approach described by ESRID can be applied to any surface variable potentially observable by visible remote sensing.