A band-ratio algorithm for retrieving open-lake chlorophyll values from satellite observations of the Great Lakes

Barry M. Lesht[a,*], Richard P. Barbiero[b], Glenn J. Warren[c]

[a] *CSC and Department of Earth and Environmental Sciences, University of Illinois at Chicago, 845 W. Taylor St., Chicago, IL 60607, USA*
[b] *CSC and Loyola University of Chicago, 1359 W. Elmdale Ave. Suite 2, Chicago, IL 60660, USA*
[c] *USEPA Great Lakes National Program Office, 77 W. Jackson Boulevard, Chicago, IL 60604, USA*

Abstract

The U.S. Environmental Protection Agency's Great Lakes National Program Office (GLNPO) has collected water quality data from the five Great Lakes annually since 1993. We used the GLNPO observations made since 2002 along with coincident measurements made by the Sea-viewing Wide Field-of-View Sensor (SeaWiFS) and the Moderate-resolution Imaging Spectroradiometer (MODIS) to develop a new band-ratio algorithm for estimating chlorophyll concentrations in the Great Lakes from satellite observations. The new algorithm is based on a third-order polynomial model using the same maximum band ratios employed in the standard NASA algorithms (OC4 for SeaWiFS and OC3M for MODIS). The sensor-specific coefficients for the new algorithm were obtained by fitting the relationship to several hundred matched field and satellite observations. Although there are some seasonal variations in some lakes, the relationship between the observed chlorophyll values and those modeled using the new coefficients is fairly stable from lake to lake and across

*Corresponding author

Email addresses: `blesht@gmail.com` (Barry M. Lesht), `gloeotri@sbcglobal.net` (Richard P. Barbiero), `warren.glenn@epa.gov` (Glenn J. Warren)

years. The accuracy of the satellite-derived chlorophyll estimates derived from the new algorithm was improved substantially relative both to the standard NASA retrievals and to previously published algorithms tuned to specific lakes. Monte-Carlo fits to randomly selected subsets of the observations allowed us to estimate the uncertainty associated with the retrievals purely as a function of the satellite data. Our results provide, for the first time, a single simple band ratio method for retrieving chlorophyll concentrations in the offshore “open” waters of the Great Lakes from satellite observations.
Keywords: Remote sensing, Ocean color, Chlorophyll *a*, Satellite observation

---

Introduction

The problem of estimating chlorophyll concentration in the surface waters of the Great Lakes from satellite observations is one that has challenged researchers for years. Although successful chlorophyll retrieval methods have been developed for large areas of the ocean (Yoder et al., 1993), efforts to develop new or adapt existing algorithms for use in the Great Lakes have met with, at best, mixed results (Lesht et al., 2012). The ocean algorithms are based for the most part on an empirical relationship between chlorophyll concentration and the ratio of the remote sensing reflectance (Rrs) measured by the satellite sensor at two wavelengths (bands). When applied to the Great Lakes, the variability in the performance of these algorithms has been attributed to the presence of confounding factors such as high concentrations of suspended material (Witter et al., 2009), high concentrations of dissolved organic material (Budd and Warrington, 2004), and phytoplankton populations dominated by particular organisms (Bergmann et al., 2004). The general consensus among workers in this area (Bukata et al., 1985; Mortimer, 1988; Li et al., 2004; Shuchman et al., 2006; Binding et al., 2008; Lohrenz et al., 2008; Binding et al., 2010) is that simple algorithms based on band ratios are not applicable to the Great Lakes because the Great Lakes, unlike the open ocean, are assumed to be optically complex “Case 2” waters (Morel and Prieur, 1977) and the factors that affect the color of the water are not dominated by phytoplankton pigments.

In theory, the influence of optically active non-algal substances, such as non-algal particulates (NAP, primarily suspended mineral particles), or colored dissolved organic material (CDOM), that would interfere with the chlorophyll retrievals based on band ratio methods can be calculated by using models that include the optical effects of these components explicitly. These calculations require knowledge of the spectrally resolved scattering and absorption properties of each optically active component (Preiur and Sathyendranath, 1981). Referred to here as the multi-component method, this approach was first applied to the Great Lakes by Bukata and colleagues (Bukata et al., 1978, 1979, 1981b,a, 1985, 1991a,b). In the multi-component method the spectral content of the incoming solar radiation reflected from the surface layer of water back to space is modeled as function of the spectral absorption and backscattering due to the combined effects of the color producing agents (CPAs, sometimes referred to as optically active constituents or OACs) present in the water. Attempts that have been made to apply multi-component methods to the Great Lakes (Bukata et al., 1985; Pozdnyakov et al., 2005; Shuchman et al., 2006) have not been entirely successful. Bukata et al. (1985) found that when applied to western Lake Ontario the multi-component method produced estimates that closely matched observed NAP concentrations and made acceptable estimates of CDOM concentration, but resulted in substantial underestimation of chlorophyll concentrations. Similarly, when Shuchman et al. (2006) compared multi-component estimates made from SeaWiFS observations with a limited set (two days) of field measurements of chlorophyll made in the vicinity of the Kalamazoo River outflow in Lake Michigan the model produced acceptable estimates of the NAP and CDOM observations, but underestimated the observed chlorophyll concentrations by an order of magnitude. More recently, however, Binding et al. (2012) developed a two-component (phytoplankton and mineral sediment) model for Lake Erie that is based on the red and near-infrared bands measured by MODIS. This model simultaneously estimates the concentrations of suspended mineral particles and chlorophyll and appears promising when applied to turbid and productive waters. Being based red and near-infrared wavelengths, this model should be fairly insensitive to the effect of CDOM absorption which is not included in the model.

No matter which components are included, the multi-component methods depend on the accuracy of the optical cross sections of the CPAs. Although work aimed at providing new estimates of these cross sections currently is underway (personal communication, G. Leshkevich, 2011), to the best of our knowledge, those determined by Bukata et al. (1981b) are the only optical cross sections measured in the Great Lakes that have been tabulated and published (Bukata et al., 1985). Other detailed optical characterization studies of the Great Lakes recently have been presented (Lohrenz et al., 2004; Binding et al., 2008; Effler et al., 2010; O'Donnell et al., 2010; Peng and Effler, 2010; Binding et al., 2012) but, with the exception of the Binding et al. (2012) study in Lake Erie, they do not present sufficient information to derive the spectral cross sections needed to apply a multi-component model. Until multi-component methods are proven and widely available, we believe that the empirical band ratio approach will provide the primary practical means of making quantitative estimates of chlorophyll concentrations in the lakes from satellite observations. Of course, because the complicating effects of non-algal substances can be significant, successful application the band ratio method will be limited to waters in which the optical properties are dominated by phytoplankton. As we will demonstrate below, however, waters where the band ratio method is most likely to be compromised by the presence of confounding substances constitute a small fraction of the Great Lakes (primarily embayments and shallow waters subject to frequent sediment resuspension).

The standard NASA retrieval algorithms are based on the work of O'Reilly et al. (1998) who conducted an extensive study comparing a large and diverse set of oceanic field measurements of chlorophyll concentrations with predictions made from a number of different retrieval algorithms. They found that, in general, the multi-component (or semi-analytical) methods did not perform as well as did band ratio methods. The band ratio methods are simple to apply and do not require detailed knowledge of the optical cross sections of the CPAs.

The fundamental assumption underlying the empirical band ratio retrieval methods is that the optical properties of the water are dominated by phytoplankton absorption of incoming solar radiation. Because chlorophyll-a absorbs most radiation at shorter (blue) wavelengths and very little in the middle (green) part of the spectrum (Bricaud and Stramski, 1990; Lohrenz et al., 2004), green light is preferentially reflected by algae. Thus, the ratio of the blue light reflected from the water (relatively sensitive to concentration of chlorophyll) to the reflected green light (relatively insensitive to chlorophyll concentration) should be inversely related to the concentration of phytoplankton in the water. By using a set of filters tuned to discrete narrow regions (bands) of the electromagnetic spectrum, satellite sensors like SeaWiFS and MODIS are designed to measure the spectral content of the light reflected from the surface in those bands most appropriate for calculating this blue/green ratio.

The choice of bands used to represent the blue and green portions of the spectrum varies between sensors and among the several empirical algorithms developed for each sensor. The current version of the standard NASA band ratio algorithm (see http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/) for SeaWiFS uses the maximum of the three bands $\{Rrs_{443}, Rrs_{489}, Rrs_{510}\}$ to represent the blue band and Rrs_{555} to represent the green band. For MODIS, the blue band is represented by the maximum of $\{Rrs_{443}, Rrs_{489}\}$ and the green band by Rrs_{547}. In both cases, the relationship between chlorophyll (Chl_a) and the band ratio is expressed as a fourth-order polynomial in $X = \log_{10}(Rrs_{blue}/Rrs_{green})$,

$$\log_{10}(Chl_a) = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + a_4 X^4. \tag{1}$$

The coefficients used in the standard NASA algorithms were determined by regression analysis of the large set of coincident *in situ* chlorophyll and Rrs measurements obtained from a wide variety of ocean waters described by O'Reilly et al. (1998). The data set (SeaBAM) used by NASA in this process is updated periodically and is publicly available (Werdell et al., 2003). No similar database exists for the Great Lakes, though as noted above efforts reportedly are underway to develop one (G. Leshkevich, personal communication, 2011).

It is important to understand that the standard NASA chlorophyll retrieval algorithms were intended to be global in scope. That is, for each sensor, one of the designers' goals was to develop a single relationship for estimating chlorophyll concentrations regardless of the time of year or area of the ocean being observed. This goal was achieved by tuning the candidate algorithms to the SeaBAM data, which were assembled by merging data from a number of different sources (O'Reilly et al., 1998) in which the measured chlorophyll concentrations ranged over four orders of magnitude (between 0.019 and 32.79 mg m^{-3}). Of course, because of the large bio-optical diversity in the ocean, it was explicitly recognized that no one single algorithm could be optimal in every situation or region. The expectation, rather, was that the estimates provided by the general algorithm would provide estimates that were within known and reasonable limits of accuracy.

Band ratio estimates of chlorophyll concentration based on the standard NASA algorithms have proven valuable for understanding biological processes in the Great Lakes (Lesht et al., 2002; Chen et al., 2004; Kerfoot et al., 2008, 2010; Barbiero et al., 2011). However, other studies have questioned the absolute accuracy of the standard NASA retrievals in the Great Lakes (Budd and Warrington, 2004; Li et al., 2004; Bergmann et al., 2004; Lohrenz et al., 2008; Witter et al., 2009; Watkins, 2009). We noted in our recent review of the applications of satellite ocean color algorithms to the Great Lakes (Lesht et al., 2012) that although the slopes, intercepts, and strength of the fits of the linear relationships between retrieved and observed chlorophyll varied from study to study, retrievals that were based on the standard NASA band ratio algorithms produced chlorophyll estimates that were linearly related to the concentrations measured in the field, contrary to expectations based on the assumption that the Great Lakes must be considered Case-2 waters. Lesht et al. (2012) showed that some variation in the results could be due to variations in the amounts of the confounding substances present, which undoubtedly differed among the published studies. Some variation might also to due to the limited extent of the data used in the underlying studies and/or from procedural differences among them.

A few studies have attempted to "tune" or optimize band ratio algorithms for particular regions in the Great Lakes. One common feature of these studies is that they have been limited to individual lakes. In some cases, this narrow focus stemmed from the researchers' proximity to or interest in the lake in question (Witter et al., 2009; Binding et al., 2012). In other cases the research was part of a larger program being done in a specific region (Li et al., 2004). Perhaps because of data limitations or because it has been assumed that retrieval algorithms must be lake-specific, no previous work has attempted to derive a single algorithm that would be applicable to all of the lakes. Such an algorithm, similar to the global ocean color algorithms that long have been used in the ocean, would greatly simplify and enhance efforts to employ satellite data for study of the Great Lakes. Although it is based on sampling done exclusively in offshore waters, the extensive GLNPO water quality monitoring data can be considered a Great Lakes analogy to the *in situ* portion of the SeaBAM database used by NASA to develop the global ocean algorithms. In this paper we describe our use of the GLNPO data to develop of a single chlorophyll retrieval algorithm for the Great Lakes that is appropriate for those regions of the lakes that are represented by the GLNPO monitoring program.

Methods

Satellite Data

All of the satellite data used in this study were processed with NASA's SeaDAS software (Baith et al., 2001), version 6.3. We began with daily SeaWiFS and MODIS L1A imagery obtained from NASA's Ocean Color Data archive (http://oceancolor.gsfc.nasa.gov). These image files, which were extracted geographically to limit the imagery to the individual lakes, included every daytime overpass from shortly after launch (September 1997 for

SeaWiFS and July 2002 for MODIS) to the end of the SeaWiFS mission (December 2010) and through December 2011 for MODIS (which still operates). We used the appropriate versions of the SeaDAS *l2gen* module to convert the raw L1A raw radiance values to L2 geophysical variable values, adopting the default SeaDAS atmospheric correction scheme that involves a 2-band model selection with an iterative near infrared (NIR) correction (Bailey et al., 2010). For days on which two L1A images of a lake were collected we kept only the image with the more favorable viewing geometry. To avoid computational artifacts, we did not further resample or grid the L2 files but rather used the original L2 data values in all of our subsequent analyses.

Field Observations

Figure 1 shows the locations of the eighty stations sampled regularly by GLNPO between 1998 and 2011. Surveys were made twice a year, generally in April to monitor spring conditions and again in August to collect data when the lakes are stratified. The actual ranges of dates sampled over the years are 11 March through 15 May in the spring and 30 July through 30 September in the summer. Although we included a two shallow water stations that were primarily intended to sample benthos in our analysis, the regular GLNPO monitoring stations were located offshore in the open waters of each lake by design, and except for those in the shallow western and central basins of Lake Erie, all are in water that is greater than 30 m deep. All the regular monitoring stations in the central basin of Lake Erie are in water that is greater than 20 m deep.

[Figure 1 here.]

At each station, samples for chlorophyll were taken at discrete depths throughout the entire water column with Niskin bottles mounted on a SeaBird Carousel Water Sampler.

For the present study, averages of samples collected from the isothermal upper (10 m) water column for each station/survey were used. Chlorophyll-a, uncorrected for pheophytin, was determined on a Turner Designs 10-AU fluorometer following the method of Welschmeyer (1994).

Data Screening and Matching

We matched the field data values to the corresponding pixels in the satellite imagery as follows: For each field sample we identified the satellite images that were recorded on or within one day of the date of field collection. Beginning with the image closest in time to the field sampling, we determined the extent to which the image was contaminated by cloud cover. The amount of cloud contamination was calculated by dividing the number of cloudy water pixels by the total number of water pixels. We discarded images that were more than 80% cloud covered and then checked the next closest image. If none of the images recorded within a day of the field sample collection passed through this initial screen, then no match was made for that field observation. For each accepted image we identified the pixel corresponding to the field location by using geometrical correlation between the station location and the image pixel locations (see Appendix A for details). We accepted the observation for analysis only if the all the pixels within a 5x5 pixel box surrounding the sampling location were cloud free and valid, as indicated by the following SeaDAS data quality flags: ATMFAIL, LAND, HIGLINT, HILT, STRAYLIGHT, CLDICE, CHLFAIL, NAVFAIL (http://oceancolor.gsfc.nasa.gov/VALIDATION/flags.html).

Of the 2126 individual GLNPO samples collected between 1998 and 2011 our matching process resulted in a total of 1035 station/pixel pairs for SeaWiFS (1998-2010) and 974 (2002-2011) for MODIS. We eliminated a number of the matched observations because one or more of the reflectance values were negative, possibly indicating that the atmospheric correction model overestimated the contributions of scattering aerosols (Bailey et al., 2010). Because Barbiero et al. (2011) found evidence that the GLNPO chlorophyll measurements made prior to 2002 were biased toward low values, we further decided to limit the SeaWiFS data to the period 2002-2010 (the end of the SeaWiFS mission). Our final data set consisted of 454 matches for SeaWiFS and 782 matches for MODIS. The distribution of matched samples/images by lake and year is shown in Table 1. Only three samples (1 for SeaWiFS and 2 for MODIS) came from the shallow benthos stations.

[Table 1 here.]

Model Selection, Fitting, and Evaluation Statistics

After conducting an extensive analysis of different combinations of band ratios and functional forms, O'Reilly et al. (1998) found that a fourth-order polynomial relating $\log_{10}(Chl_a)$ to $X = log_{10}(Rrs_{blue}/Rrs_{green})$ best represented the SeaBAM data (as noted above, Rrs_{blue} is $\max[Rrs_{443}, Rrs_{489}, Rrs_{510}]$ for SeaWiFS and is $\max[Rrs_{443}, Rrs_{489}]$ for MODIS; Rrs_{green} is Rrs_{555} for SeaWiFS and Rrs_{547} for MODIS) . Rather than experiment with different band ratios and functional forms, we chose to use these same ratios and polynomial model in our study, but we found that the Great Lakes data could be represented adequately by a third-order relationship

$$\log_{10}(Chl_a) = a_0 + a_1 X + a_2 X^2 + a_3 X^3. \tag{2}$$

We used the same tuning method described by O'Reilly et al. (1998) to determine the coefficients in Eq. 2. This procedure uses an iterative process in which the model coefficients (a_0, a_1, a_2, a_3 in Eq. 2) are adjusted until the intercept and slope of the linear relationship between $\log_{10}(Chl_a^{model})$ and $\log_{10}(Chl_a^{insitu})$ were zero and one respectively. In contrast to standard linear regression in which the objective is to determine a set of model coefficients that minimizes the sum of the squared differences between the modeled and observed values, the iterative method is aimed at determining the set of model coefficients that produces a 1:1 relationship between the modeled and observed values. Although the error sum of squares may be larger (relative to the standard regression result) when the model coefficients are determined by using the iterative method, the method facilitates comparison between different models by constraining the slope and intercept (O'Reilly et al., 1998; Campbell and O'Reilly, 2006).

We based our assessment of model performance on statistics calculated from the log-transformed data. The log transformation is appropriate (Campbell and O'Reilly, 2006) both because the data values vary over several orders of magnitude and because the log transformed chlorophyll is more normally distributed than the untransformed data (Fig 2). Our evaluation statistics include the slope (b) and intercept (a) of the best fit regression line between the (log transformed) model estimates (P_i) and observed values (O_i), the bias, or the difference between the means of the estimates ($\overline{P}$) and the observations ($\overline{O}$) (negative bias indicates that the predicted values underestimate the observed values), the ratio of the standard deviations of the estimates (σ_p) and observations (σ_o), the refined index of agreement (d_r) (Willmott et al., 2011), the root mean squared error ($RMSE$), the percent unsystematic error ($\%USE$) (Willmott, 1982), and the mean absolute error (MAE) (Willmott et al., 2009). For comparison with other studies, we also included the Pearson's correlation coefficient (r), though this statistic has been shown to be overly sensitive to high extreme values (Willmott, 1982; Legates and McCabe, 1999; Moriasi et al., 2007) and is less useful than measures based on the absolute difference between the estimates and observations (Campbell and O'Reilly, 2006). Because several of these statistics may be unfamiliar, we provide their formal definitions in Appendix B.

[Figure 2 here.]

The slope and intercept of the regression line indicate how well the estimates match the observations. We used a type II (or reduced major axis) model to compute the regression. This type of model is appropriate when both variables are subject to uncertainty (Press and Teukolsky, 1992; Press et al., 1992). Bias measures the average tendency of the estimates to be larger or smaller than the observations; ideally the bias would equal zero. Similarly, comparing the standard deviations of the estimates and observations shows how well the model reproduces the overall variation in the data. Both $RMSE$ and MAE are error indices that are useful because they characterize the error in the units of the variable of interest. Because it is based on the squared error, $RMSE$ tends to exaggerate large errors and MAE is the preferred statistic (Willmott et al., 2009). In either case, the lower the ratio of $RMSE$ or MAE to the standard deviation of the observations, the better the model (Moriasi et al., 2007). Ideally, models would be free of systematic error (Willmott, 1982). In better models, the $\%USE$, which is the unsystematic proportion of the $RMSE$ approaches one. In addition to evaluating the overall performance of the model, we calculated evaluation statistics for each lake, for each lake and season, and for each year.

The refined index of agreement (Willmott et al., 2011) is a dimensionless statistic, bounded by ± 1, that provides a summary measure of how well the model estimates reproduce the data. Based on the absolute values of the difference between the estimates and observations, this statistic is not overly sensitive to high extreme values. The d_r statistic is a measure of how well the model (algorithm) predicts the observations relative to how well the observations could be predicted by the observed mean. A perfect model, one for which $P_i = O_i$, would result in a d_r value of 1. If the sum of the absolute differences between the predicted and observed values ($\Sigma|O_i - P_i|$) is very large relative to the sum of the absolute deviations of the observations around their mean ($\Sigma|O_i - \overline{O}|$) or if there is very little observed variability, then d_r will approach -1. The value of d_r will be zero when sum of the absolute value of the differences between the predictions and observations is twice the sum absolute differences of the observations about the observed mean. Models with $d_r = 0.5$ result in predictions that are equivalent to using the observed mean as the predictor and "good" models should have d_r values > 0.5 indicating that the sum of the absolute predicted deviations is less than the sum of the absolute observed deviations. However, because characterization of model performance using values of d_r is somewhat arbitrary (Legates and McCabe, 2012), we use the statistic as a relative indicator. When evaluating the success of the models applied to our data, we primarily considered the slope, the intercept, the MAE, and the $\%USE$.

Estimation of Retrieval and Parameter Uncertainties

We have only one set of matched data for each sensor so we are unable to validate our results with a completely independent set of observations. Furthermore, standard methods for assessing the uncertainty associated with the model fits are inapplicable to our data because the model coefficients were determined by using the iterative method described above rather than by using simple least-squares regression. To address these issues we adapted a dual Monte Carlo resampling approach (Wei et al., 2008) to estimate the uncertainties of our models. In this two step process we first selected (with replacement) a random subset of half the observations. Assuming that both the selected chlorophyll and maximum band ratio values were samples from independent, normally-distributed, random variables we then perturbed the observed values by a random error term scaled to an assumed accuracy for the measured variables (5% for maximum band ratio (Bailey and Werdell, 2006) and 10% for chlorophyll) and used these perturbed values as the basis for a new fit. We repeated these random selection processes 1000 times to generate an ensemble of model coefficients that could be used to estimate confidence intervals for the predictions as a function of the maximum band ratio. Volpe et al. (2011) used a similar method to determine confidence intervals for estimates of remote sensing reflectance as a function of water turbidity.

We used a subsampling approach (Hartigan, 1969) to estimate the uncertainty in the model parameters. Because the number of our matched samples was fairly large, we did not apply the "leave out one" jackknife analysis adopted by the few other remote sensing studies that attempted a similar analysis (Volpe et al., 2011; Novoa et al., 2012). Rather, we partitioned the full data set into halves by years and determined the model coefficients for each of the five-year partitions. Because the total data set consisted of ten years, there are 252 unique five-year partitions. We also determined how well the model tuned to each partition predicted the observations in the complementary partition and used the model evaluation statistics described above to assess the results.

Results

Satellite images matched with in situ *chlorophyll observations*

Chlorophyll concentrations measured by GLNPO between 2002 and 2011 ranged between 0.19 mg m^{-3} and 33.55 mg m^{-3}, with a geometric mean of 1.37 mg m^{-3}. The distribution of the measurements is approximately log-normal, with a slight skew toward larger values. Histograms of the subsets of chlorophyll values that were matched with the SeaWiFS and MODIS observations (Fig. 2) are very similar to overall distribution indicating that the matching process resulted in samples representative of the overall population of observations. The minimum, maximum, and geometric mean of the SeaWiFS-matched chlorophyll values were 0.24 mg m^{-3}, 24.03 mg m^{-3}, and 1.29 mg m^{-3}. For MODIS, these values were 0.22 mg m^{-3}, 32.69 mg m^{-3}, and 1.30 mg m^{-3}.

Model fit

Model coefficients derived from GLNPO data (Table 2) resulted in improved fits for both MODIS and SeaWiFS sensors, compared to standard NASA models (Fig. 3). Recalling that the coefficients determined for the Great Lakes Fit (GLF) models (Table 2) were constrained to result in a slope of one and intercept of zero, the values for d_r, %USE, and MAE were 0.780, 0.976, 0.142 for MODIS and 0.758, 0.956, and 0.158 for SeaWiFS, respectively. For comparison, the slope, intercept, and statistics for the standard NASA models were 0.892, -0.074, 0.761, 0.640, 0.154 for MODIS-OC3M and 0.844, -0.048, 0.739, 0.631, and 0.170 for SeaWiFS-OC4. When based on the entire dataset, the standard NASA relationships tend to underestimate the chlorophyll concentration for both sensors. A larger set of evaluation statistics comparing the GLF model to the standard NASA algorithms is given in Table 3.

[Table 2 here.]

[Figure 3 here.]

[Table 3 here.]

Plots (after O'Reilly et al. (1998)) showing the relationships between the observed chlorophyll values and those predicted using both the standard NASA algorithms and the GLF model for MODIS and SeaWiFS are shown in Fig. 4. Because the GLF model coefficients were determined under the constraint of producing a slope and intercept equal to one and zero respectively, the relative improvement in the performance of the fitted models is best illustrated by the changes in the quantile-quantile and frequency distribution plots. The plots indicate greater deviation between predicted and observed values at higher chlorophyll concentrations in the standard NASA models (OC3M and OC4) compared to the GLF models. This underestimation of chlorophyll at high concentrations by the O3M and OC4 models can be seen by the divergence in relative frequencies between model and *in situ* values at high chlorophyll concentrations Fig. 4. At the most extreme values, however, all models show notable deviation between observed and predicted values.

[Figure 4 here.]

Fits by lake and by year

The highest chlorophyll concentrations observed in our study come from Lake Erie where there also seems to be a distinct seasonal bifurcation in the relationship between $\log_{10}(Chl_a^{insitu})$ and $\log_{10}$ of the maximum band ratio (MBR), especially for MODIS (Fig. 4 bottom row). This bifurcation is evident when the GLF model is applied separately by season (Fig. 5 and Table 3). The seasonal difference is most pronounced in Lake Erie (for both MODIS and SeaWiFS), where spring values were underestimated by the model and summer values were overestimated. The slope of the relationship between the values predicted by the GLF models and the observations also shows seasonal dependence in both Lakes Ontario (MODIS and SeaWiFS) and Superior (SeaWiFS).

When applied to the individual lakes the MODIS GLF model performs very well (slope very close to 1, intercept near 0) in Lakes Huron and Michigan (Table 4). The model tends to under-predict the observations made in Lakes Erie and Superior, though the slopes are still greater than 0.92. In Lake Ontario, the slope is higher than would be expected (1.45), but this determination is based on considerably fewer samples than in the other lakes. The SeaWiFS GLF results are similar for Lakes Erie, Huron, Michigan, and Ontario with slopes between 0.91 and 1.16. The Lake Superior slope is somewhat higher (1.25). In both cases with high slopes (MODIS Ontario and SeaWiFS Superior) the standard deviation of the model predictions is higher than the standard deviation of the observations, suggesting the presence of some outlier observations. We did not attempt to identify or remove possible outliers in this analysis.

[Figure 5 here.]

[Table 4 here.]

When the GLF model is applied to all lakes by year the quality of the predictions are remarkably stable (Table 5). For MODIS the slope of the relationship between the values predicted by the GLF and the observations varies between 0.911 (2007) and 1.150 (2005) and MAE between 0.108 (2010) and 0.179 (2011). For SeaWiFS the slope varies between 0.940 (2007) and 1.417 (2010) and MAE between 0.129 (2003) and 0.186 (2006). We note, however, that the SeaWiFS sensor experienced problems throughout 2010 before failing completely in December of that year so the results for 2010 may be suspect.

[Table 5 here.]

Comparison with published regionalized models

Only a few studies have been published in which researchers attempted to improve the local accuracy of chlorophyll retrievals by fitting new band ratio models to data collected in the Great Lakes. Li et al. (2004) (L-2004) used *in situ* optical measurements and least squares methods to optimize the fit of the OC4 algorithm (Eq. 1) to chlorophyll data they collected in Lake Superior. The set of optimized coefficients are given as {0.3815, -1.6837, 2.5054, -0.5899, -0.6505} (L-2004, page 452). When applied to our data, the results (Fig. 6, bottom panel) show that a 4th order model with these coefficients blows up at higher values of the maximum band ratio and predicts unrealistically low chlorophyll concentrations (relative to the GLNPO chlorophyll observations). This problem is avoided in the 3rd order GLF model (Fig. 6, top panel) which produces reasonable predictions (intercept = 0.032, slope = 1.247, MAE = 0.121) over the narrow range of chlorophyll values (0.5 mg m^{-3} to 1.9 mg m^{-3}) that were observed in Lake Superior.

[Figure 6 here.]

Witter et al. (2009) (W-2009) used data collected in Lake Erie to develop a set of "regional algorithms" that were tuned both for whole lake and for the three individual lake basins. Rather than use an algorithm in the same form as Eq. 1, W-2009 found that the expression $Chl_a = 10^{a+bR+cR^2}$, where $R = \log(Rrs_{490}/Rrs_{555})$ and a, b, and c are a set of coefficients specific to the whole lake, and western, central, and eastern basins, resulted in a statistically improved relationship (relative to estimates from the standard NASA algorithms) between the calculated and observed chlorophyll values, though the tuned estimates still tended to underestimate the observed values. When we applied the W-2009 models to our data (Fig. 7, first column), we found that although the slope of the fit for the whole-lake was close to one, the model was biased low (observed values were higher than the modeled values). Seasonal differences in the relationship between the modeled and observed chlorophyll are seen in both the GLF and W-2009 models with the apparent slope for the spring data being lower than that for the summer data. Although it is not basin-specific, the GLF model produced estimates in the eastern and central basins with MAE values of 0.180 and 0.186 (log units) respectively. The overall (not seasonally separated) GLF predictions in the more turbid and productive western basin had an MAE of 0.322. The MAE values for the W-2009 model were 0.305, 0.422, and 0.343 in the eastern, central, and western basins respectively. The GLF tended to over predict the lower range of chlorophyll values observed in the spring in all basins. The over prediction was largest in the western basin.

[Figure 7 here.]

The Binding et al. (2012) model (B-2012) uses the multi-component approach to simultaneously estimate the concentrations of suspended mineral particles and chlorophyll in Lake Erie from MODIS data. Overall, we found that this multicomponent model did not perform as well as did the GLF when compared to the GLNPO observations (Fig. 8). The intercept, slope, and MAE values for the GLF model were 0.087, 0.917, and 0.209 compared to -0.318, 1.226, and 0.382 for B-2012. The B-2012 model, however, was developed primarily for application to turbid and productive waters and can result in artificially low concentration values (sometimes negative estimates) in clearer waters. As is the case for the GLF model, there seems to be some seasonal dependence in the B-2012 predictions, with the slope of the relationship between the modeled and observed chlorophyll values lower in the spring than in the summer at higher chlorophyll values (> 4 mg m^{-3}) where the B-2012 model should be most accurate.

[Figure 8 here.]

Parameter and prediction uncertainty

The distributions of the model parameters determined from the 252 unique five-year partitions show that for all the parameters the mean values of the distributions are very close to the values obtained by fitting to the entire dataset (Fig. 9). Performance of the models fit to the complementary five-year partitions is comparable to the performance of overall model. The mean value of d_r for the partitioned subsets is 0.772 for MODIS and 0.743 for SeaWiFS with ranges of [0.736, 0.798] and [0.674, 0.791] respectively, suggesting that the model calibration is robust. A complete listing of the fit statistics for the independent data sets is given in Table 6.

[Figure 9 here.]

[Table 6 here.]

The estimated chlorophyll prediction error is shown as a function of the observed maximum band ratio in Fig. 10. By enumerating the Monte Carlo generated values in a number of bins along the MBR axis, we were able to estimate empirical confidence intervals for the model predictions (the 80% interval is listed in Table 7 along with the $\pm 1\sigma$ interval). For both MODIS and SeaWiFS, the GLF predictions become very uncertain when $\log_{10}$(MBR) values are very low, (< -0.3 for MODIS, < -0.2 for SeaWiFS). In our data, however, observations in this range are fairly rare (Fig. 11). Throughout most of the range of observed MBR values, the estimated accuracy of the retrieved chlorophyll concentrations is better than 30%.

[Figure 10 here.]

[Table 7 here.]

Figure 11 here.]

Discussion

The GLF models presented here represent the first chlorophyll retrieval algorithms tuned to data from all five Laurentian Great Lakes. Our study is unique in both its spatial and temporal extent, covering all five lakes and including ten years of data and our results clearly demonstrate a consistent relationship between satellite-measured blue-green reflectance ratios and surface chlorophyll concentrations in the offshore waters of the Great Lakes represented by the GLNPO monitoring program. Based on several appropriate statistical measures, including the slope and intercept of the linear relationship between the modeled and observed log-transformed chlorophyll concentrations, the mean absolute error, and the revised index of agreement, the GLF model outperformed both the standard ocean-derived algorithms (OC4 for SeaWiFS and OC3M for MODIS, O'Reilly et al. (1998)) as well as regionally-tuned, lake-specific, algorithms developed for Lake Erie (Witter et al., 2009; Binding et al., 2012) and Lake Superior (Li et al., 2004).

Based on our Monte-Carlo simulations, we estimate that the accuracy of the GLF predictions throughout most of the expected concentration range in the offshore waters of the Great Lakes is better than the ±35% criterion established for the standard NASA algorithms for Case 1 ocean waters (McClain et al., 1992). We note that the errors in reflectance and chlorophyll assumed above are intended to represent random measurement variability and not the systematic variability that would result from the contributions of non-algal substances to the radiance values. This latter (and likely larger) source of variability is represented by the random selection of samples. For example, if the set of selected observations includes samples in which non-algal substances dominate the reflectance spectrum (and band ratio), then the resulting variability should be reflected in the predicted chlorophyll concentrations. Of course, the degree to which this latter source of variability is included in our data depends on the extent to which interfering substances influenced the input observations. Because the GLNPO data were primarily collected in offshore waters where the effects of interferences would tend to be minimal, our estimates will likely underestimate the uncertainty that might be associated with observations made near the shore, in very turbid waters, or waters with high concentrations of CDOM.

The GLF is based on a third-order polynomial (Eq. 2) rather than on the fourth-order polynomial used in the NASA OC3M and OC4 algorithms (Eq. 1). O'Reilly et al. (1998) note that adding the higher order (fourth) term in their relationship served to improve the fit at the lowest chlorophyll values. Because the lowest chlorophyll concentration observed in the Great Lakes is an order of magnitude larger than the lowest value in the ocean data set (0.19 mg m^{-3} in the Great Lakes versus 0.019 mg m^{-3} in SeaBAM) and the highest chlorophyll concentrations are comparable (33.55 mg m^{-3} in the lakes data versus 32.79 mg m^{-3} in the ocean) the extra term is unnecessary for modeling the Great Lakes. The retrievals based on the NASA algorithms are biased low throughout the entire range of observed chlorophyll concentrations but the greatest differences between the NASA retrievals and those obtained using the GLF occur at higher concentrations. This difference may represent a compositional distinction between the coastal ocean samples that contribute to the higher chlorophyll concentrations in the SeaBAM data and those (primarily Lake Erie) samples that contribute the high chlorophyll values in the GLNPO data.

Because the Lake Erie samples appear to drive the major differences between the NASA and GLF results, we repeated the GLF analysis on the data base eliminating the Lake Erie samples. As would be expected the reduced-set (without Lake Erie) GLF model coefficients differ from those derived using the entire data set. The new a_0, a_1, a_2, and a_3 values for MODIS are {0.3269, -2.7992, 1.2031, and 1.9369}. For SeaWiFS the coefficient set is {0.3889, -2.6479, 0.4819, -1.1660}. As measured by the change in slope, intercept, d_r, MAE, and $\%MAE$, the overall GLF fit to the remaining lakes was improved only marginally. On the other hand, the ability of the NASA algorithms to model the GLNPO observations was much improved, with slopes much closer to one (1.013 for MODIS, 0.964 for SeaWiFS). The NASA algorithms, however, still had much higher bias than the GLF fits (0.072 for MODIS and 0.028 for SeaWiFS).

Although some samples were obtained in coastal waters the SeaBAM dataset was drawn primarily from Case 1 non-polar waters in which optical properties are dominated by phytoplankton and their associated products (O'Reilly et al., 1998). The coefficients for the standard OC3M and OC4 equations were derived from SeaBAM and thus, these coefficients would not be expected to perform well in more optically complex Case-2 waters in which non-algal derived substances, such as mineral suspended solids and colored dissolved organic matter, significantly influence optical properties. Our results show that a band ratio model can be used successfully in the offshore areas of the Great Lakes. This fact implies that, at least on a statistical basis, these waters are similar to the ocean Case 1 waters in which the optical properties are dominated by phytoplankton.

Many studies, however, have shown that the standard band ratio equation form can still be appropriate for optically complex waters when fitted to a regional dataset. For example, McKee et al. (2007) derived sets of OC4 coefficients for two optical water types a study assessing the applicability of blue/green reflectance ratios to estimate chlorophyll in the Irish and Celtic Seas. The resulting models performed well, indicating that the standard multiple band ratio equation form can be appropriate for shelf seas. Werdell et al. (2007) successfully developed a tuned OC3 type algorithm for estimating chlorophyll concentrations in Chesapeake Bay. A regionally-tuned version of OC4 also improved the accuracy of SeaWiFS retrievals in the Yellow and East China Seas (Siswanto et al., 2011), although an alternative equation form (Tassan, 1994) proved superior at high TSM concentrations. In the Baltic Sea, however, developed tuned standard algorithms for both MODIS and SeaWiFS substantially reduced bias in chlorophyll retrievals, but were still deemed unsatisfactory in view of large associated RSME (Darecki and Stramski, 2004).

None of the few previous instances where band ratio model tuning has been attempted in the Great Lakes has resulted in accurate retrievals. In W-2009 the tuned models provided better estimates than did the standard NASA models but the tuned estimates still were lower than the observations. Li et al. (2004) (L-2004) however, were unable to find a tuned model that was significantly better than the NASA algorithm. When applied to our observations, the W-2009 basin-specific models (Witter et al., 2009) generally under-predicted the *in situ* data. As mentioned above, Barbiero et al. (2011) found evidence that the GLNPO chlorophyll measurements made prior to 2002 were likely too low. Because most of the data used by W-2009 to calibrate their model were obtained from GLNPO surveys made between 1998 and 2002 this low bias might account for under-prediction of W-2009. The negative bias in the tuned W-2009 model is also seen in the results for the individual basins. Li et al. (2004) suggest that the low ratio of chlorophyll-*a* to CDOM precludes the use of empirical approaches in Lake Superior. However, this conclusion is based on the poor performance of L-2004 model in the coastal waters sampled by L-2004 where CDOM concentrations may be high due to riverine inputs. As might be expected, the L-2004 also fails when applied to our observations, which come entirely from offshore regions where the GLF model provides reasonably good predictions in Lake Superior (Fig. 6).

The tendency of both the MODIS and SeaWiFS GLFs to underestimate the observed chlorophyll in Lakes Erie and Ontario (Fig. 5) in the spring and to overestimate chlorophyll in the summer may be due to seasonal differences in the phytoplankton population. Specific absorption coefficients of phytoplankton can vary due to differences in pigment composition, cell size and amount and distribution of pigment within the cell, (e.g., Sathyendranath et al. (1987)). As a result, phytoplankton community composition will impact estimation of chlorophyll from ocean-color data (Carder et al., 1999; Sathyendranath et al., 2001). For instance, Stuart et al. (2000) have shown that diatom populations exhibit lower specific absorption coefficients, relative to prymnesiophytes, in the Labrador Sea due to increased pigment packaging and increased intra-cellular chlorophyll a, while Bergmann et al. (2004) have hypothesized that the accuracy of blue/green reflectance ratios in the Great Lakes can be compromised by phycobilin-containing algae. Sathyendranath et al. (2004) used a model to estimate the effects of changes in the dominance of diatoms on the reflectance ratio and found that, for equal values of chlorophyll concentration, diatom-dominated populations would tend to have higher reflectance ratios than populations of mixed phytoplankton. Therefore, dominance of phytoplankton communities by larger diatoms might lead to underestimation of retrieved chlorophyll. While the differences in the calculated ratios by Sathyendranath et al. (2004) were relatively small (~5%), the effect on concentration estimates would be amplified at low band ratio (high concentration) values because of the polynomial form of the retrieval models. Spring phytoplankton community composition in Lake Erie, particularly in the central and eastern basins, is notable for being dominated by the large-celled diatom *Aulacoseira islandica* (GLNPO, unpublished data). Because the GLF coefficients were tuned to the complete dataset, the GLF estimates will generally fall between the extremes defined by pure diatom and mixed plankton populations in more productive regions of the lakes. The relatively wide confidence intervals associated with low band ratio values (Fig. 10) reflect the effect of this seasonal bifurcation. Seasonal differences in the fit are much less pronounced in the other lakes because chlorophyll values are generally low (higher band ratio values).

In addition to differences due to changes in the phytoplankton populations, chlorophyll overestimates in Lake Ontario during the summer may also result from whiting events (Peng and Effler, 2010) that cause high reflectance values in the green portion of the spectrum (Wortman, 2005). High green reflectance values would tend to reduce the observed band ratio and result in higher estimated chlorophyll values. GLNPO summer sampling during the years 2005, 2006, and 2007 coincided with peaks in satellite observed values of Rrs_{555} (J. Watkins, personal communication, 2012) and these three years accounted for almost half of the total number of matched samples from Lake Ontario in our data (Table 1).

No previous study of chlorophyll retrievals done in the Great Lakes has used such an extensive set of data, nor has one attempted to provide any characterization of the uncertainty associated with the estimates. Although it is likely that random variations in the quantities of interfering substances contribute most to the uncertainty associated with the GLF retrievals, other factors that are difficult to ignore also may have some affect. Among these are errors in the basic measurements of chlorophyll and reflectance (including inaccuracy in the atmospheric correction algorithm embedded in the radiometric calibrations) and temporal and spatial differences in the matching of the *in situ* and satellite observations. Our Monte-Carlo approach was intended to simulate the combined effects of these error sources. The question of uncertainty becomes most important when satellite data are used to estimate absolute values of chlorophyll and to assess the significance of apparent changes in concentration in space or over time. For example, using SeaWiFS data Barbiero et al. (2011) found that spring chlorophyll levels in Lake Huron declined by approximately 50% between 1998-2002 and 2003-2006. The SeaWiFS estimated average Lake Huron southern basin spring (April-May) chlorophyll concentration in 2003-2006 was ~1.0 mg m^{-3}, a decline of approximately 0.8 mg m^{-3} from the values estimated for the 1998-2002 period. Based on the empirical confidence limits shown in Table 7, a change of this magnitude is unlikely (<10%) to be an artifact of the retrieval uncertainty.

Conclusion

Algorithms based on the blue-green band ratio are among the most simple of the methods designed to retrieve surface water chlorophyll concentration from satellite observations. The practical utility of the band ratio method results from this simplicity. Estimates of blue and green reflectance from satellite sensors are readily available and the computation required to convert reflectance values into estimates of chlorophyll concentration is straightforward and easily implemented. By using a single set of sensor-dependent coefficients, the GLF model makes it possible to make consistent estimates of chlorophyll concentration across the lakes without the necessity of adjusting coefficients on the basis of location, season, or year.

Discovering the limits of band ratio methods applied to the Great Lakes is an ongoing process. We expect our results to be most applicable to the offshore waters represented by the GLNPO monitoring program. The regular GLNPO monitoring program does not include sampling in the major Great Lakes embayments such as Green Bay, Saginaw Bay, the North Channel, Georgian Bay, and the Bay of Quinte. Because these areas are outside the our sampling universe, we would not necessarily expect that our results would be applicable to satellite observations of these waters, nor would we necessarily expect our results to be applicable to shallow or nearshore waters strongly influenced by sediment resuspension or the presence of high concentrations of CDOM. Some work has been reported in which satellite observations have been used to determine if waters are Class-1 or Class-2 (Lee and Hu, 2006; Matsushita et al., 2012) and to classify inland waters before choosing a retrieval algorithm that has been tuned to water type (Le et al., 2011; Li et al., 2012). Such methods may be applicable to the Great Lakes and we are exploring that possibility using the GLNPO data.

Further work is needed to determine the causes of the apparent seasonal bifurcation in the relationship between observed chlorophyll and the maximum band ratio at higher chlorophyll concentrations and to understand how errors in the retrievals might be related to other properties of the surface water that are observable by satellite. Validation of the

GLF model with independent data also is very desirable. Although the GLF performs well, other algorithmic approaches also should be explored. Given the appropriate optical cross-sections, the two-component model developed for Lake Erie by Binding et al. (2012) based on two bands in the red and near-infrared is fairly simple to apply and shows promise for providing simultaneous estimates of both chlorophyll and suspended mineral concentrations in productive regions of the lakes. Similarly, the five-band algorithm developed by Gohin et al. (2002) that incorporates both the blue-green ratio as well as radiances at two other wavelengths was successful when applied to the coastal Bay of Biscay (Gohin et al., 2005), the English Channel (Gohin et al., 2002), and the Bay of Bengal and Arabian Sea (Tilstone et al., 2011) and should be investigated using data from the Great Lakes.

A simple band ratio method using a single set of sensor-specific coefficients can provide consistent estimates of chlorophyll concentrations in the offshore surface waters of Great Lakes with accuracy comparable to that required for oceanic estimates. The uncertainty associated with the chlorophyll retrievals also can be estimated from the satellite data making it possible to assign confidence limits to the estimates. Because the model is independent of lake and time, application of the GLF to satellite images of the Great Lakes provides the means for quantitative analysis of differences within and between lakes and over time. Applying the GLF to both historical and contemporary satellite observations should greatly facilitate use of this imagery in studies of phytoplankton processes in the Great Lakes.

Acknowledgements

This work was supported by the USEPA Great Lakes National Program Office as part of EPA Contract No. EP-C-06-085, Scientific and Technical Support with CSC under the direction of Louis Blume, Project Manager. We gratefully acknowledge the constructive comments of two anonymous reviewers.

Appendices

A. Matching station locations with image pixels

We let *plat*, *plon* represent the latitude and longitude of the image pixels and *flat*, *flon* the latitude and longitude of the field station. To determined the image pixel corresponding to the field location we let

$$\begin{aligned} uvp[0] &= \cos(plat) * \cos(plon) \\ uvp[1] &= \cos(plat) * \sin(plon) \\ uvp[2] &= \sin(plat) \end{aligned}$$

and

$$\begin{aligned} uvf[0] &= \cos(flat) * \cos(flon) \\ uvf[1] &= \cos(flat) * \sin(flon) \end{aligned}$$

$$uvf[2] = \sin(flat)$$

and then calculate the dot product between the uvp and uvf vectors,

$$dot = uvp[0] * uvf[0] + uvp[1] * uvf[1] + uvp[2] * uvf[2].$$

Finding the maximum value of the dot product yields the location of the desired image pixel. Once this pixel was located, we checked to ensure that it and all pixels within a 5-km radius of it were cloud free and had valid data (as determined by the status of the L2 data quality flags). Only if these criteria were met did we accept the station and image pair for further analysis

B. Model comparison statistics

Following Willmott (1982) and Willmott et al. (2011), given a set of N paired observations, O_i and model predictions, P_i we define the following statistics that appear in the text. Following Campbell and O'Reilly (2006) these statistics are based on the log-transformed variables.

$\overline{O} = N^{-1} \sum_{i=1}^{N} O_i$, the mean of the observations.

$\overline{P} = N^{-1} \sum_{i=1}^{N} P_i$, the mean of the predicted values.

$\hat{P}_i = a + bO_i$, the linear fit prediction of P_i, where a and b are the intercept and slope of the least-squares regression of P on O.

$$MSE_s = N^{-1} \sum_{i=1}^{N} (\hat{P}_i - O_i)^2,$$ the systematic error of the model.

$$MSE_u = N^{-1} \sum_{i=1}^{N} (P_i - \hat{P}_i)^2,$$ the unsystematic error of the model.

$MSE = MSE_s + MSE_u$, the mean square error.

$\%USE = MSE_u / MSE$, the percent unsystematic error.

$$RMSE = [N^{-1} \sum_{i=1}^{N} (P_i - O_i)^2]^{0.5},$$ the root mean square error.

$$MAE = N^{-1} \sum_{i=1}^{N} |P_i - O_i|,$$ the mean absolute error.

The revised index of agreement (d_r),

$$d_r = \begin{cases} 1 - \dfrac{\sum_{i=1}^{n} |P_i - O_i|}{2\sum_{i=1}^{n} |O_i - \overline{O}|}, & \text{when} \\ \sum_{i=1}^{n} |P_i - O_i| \leq 2\sum_{i=1}^{n} |O_i - \overline{O}| & \\ & \\ \dfrac{2\sum_{i=1}^{n} |O_i - \overline{O}|}{\sum_{i=1}^{n} |P_i - O_i|} - 1, & \text{when} \\ \sum_{i=1}^{n} |P_i - O_i| > 2\sum_{i=1}^{n} |O_i - \overline{O}| & \end{cases}$$

References

Bailey, S., Werdell, P., 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., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. Optics Express 18, 7521–7527.

Baith, K., Lindsay, R., Fu, G., McClain, C., 2001. SeaDAS, a data analysis system for ocean-color satellite sensors. EOS Transactions of the American Geophysical Union 82, 202.

Barbiero, R.P., Lesht, B.M., Warren, G.J., 2011. Evidence for bottom-up control of recent shifts in the pelagic food web of Lake Huron. J. Great Lakes Research 37, 78–85.

Bergmann, T., Fahnenstiel, G., Lohrenz, S., Millie, D., Schofield, O., 2004. Impacts of a recurrent resuspension event and variable phytoplankton community composition on remote sensing reflectance. J. Geophys. Res. 109.

Binding, C.E., Greenberg, T.A., Bukata, R.P., 2012. An analysis of MODIS-derived algal and mineral turbidity in lake erie. J. Great Lakes Research 38, 107–116.

Binding, C.E., Jerome, J.H., Bukata, R.P., Booty, W.G., 2008. Spectral absorption properties of dissolved and particulate matter in Lake Erie. Remote Sensing of Environment 112, 1702–1711.

Binding, C.E., Jerome, J.H., Bukata, R.P., Booty, W.G., 2010. Suspended particulate matter in Lake Erie derived from MODIS aquatic colour imagery. International Journal of Remote Sensing 31, 5239–5255.

Bricaud, A., Stramski, D., 1990. Spectral absorption coeffficients of living phytoplankton and nonalgal bigenous matter: A comparison between the Peru upwelling area and the Sargasso Sea. Limnology and Oceanography 35, 562–582.

Budd, J., Warrington, D., 2004. Satellite-based sediment and chlorophyll a estimates for Lake Superior. Journal of Great Lakes Research– Special Issue on Lake Superior 30, 459–466.

Bukata, R.P., Bruton, J.E., Jerome, J.H., 1985. Application of direct measurements of optical parameters to the estimation of lake water quality indicators. Scientific Series 140. Inland Waters Directorate, National Water Research Institute. Canada Centre for Inland Waters, Burlington, Ontario, Canada.

Bukata, R.P., Bruton, J.E., Jerome, J.H., Jain, S., Zwick, H.H., 1981a. Optical water quality model of Lake Ontario. 2: Determination of chlorophyll-*a* and suspended mineral concentrations of natural waters from submersible and low altitude optical sensors. Applied Optics 20, 1704–1714.

Bukata, R.P., Jerome, J.H., Bruton, J.E., Bennett, E., 1978. Relationship among optical transmission, volume reflectance, suspended sediment concentration, and chlorophyll-*a* concentration in Lake Superior. J. Great Lakes Research 4, 456–461.

Bukata, R.P., Jerome, J.H., Bruton, J.E., Jain, S., 1979. Determination of inherent optical properties of Lake Ontario coastal waters. Applied Optics 18, 3926–3932.

Bukata, R.P., Jerome, J.H., Bruton, J.E., Jain, S., Zwick, H.H., 1981b. Optical water quality model of Lake Ontario. 1: Determination of the optical cross sections of organic and inorganic particulates in Lake Ontario. Applied Optics 20, 1696–1703.

Bukata, R.P., Jerome, J.H., Kondratyev, K.Y., Pozdnyakov, D.V., 1991a. Estimation of organic and inorganic matter in inland lakes: optical cross sections of Lakes Ontario and ladoa. J. Great Lakes Research 17, 461–469.

Bukata, R.P., Jerome, J.H., Kondratyev, K.Y., Pozdnyakov, D.V., 1991b. Satellite monitoring of optically-active components of inland waters: an essential input to regional climate change impact studies. J. Great Lakes Research 17, 470–478.

Campbell, J.W., O'Reilly, J.E., 2006. Metrics for quantifying the uncertainty in a chlorophyll algorithm: explicit equations and examples using the OC4.v4 algorithm and NOMAD data. Papers from the Ocean Color Bio-optical Algorithm Mini-Workshop (27-29 September 2005).

Carder, K.L., Chen, F., Lee, Z., Hawes, S., 1999. Semianalytical moderate-resolution imaging spectrometer algorithms for chlorophyll-*a* and adsorption with bio-optical domains based on nitrate-depletion temperatures. J. Geophys. Res. 104, 5403–5421.

Chen, C., Wang, L., Ji, R., Budd, J.W., Schwab, D.J., Beletsky, D., Fahnenstiel, G.L., Vanderploeg, H., Eadie, B.J., Cotner, J., 2004. Impacts of suspended sediment on the ecosystem of Lake Michigan: A comparison between the 1998 and 1999 plume events. J. Geophys. Res. 109.

Darecki, M., Stramski, D., 2004. An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea. Remote Sensing of Environment 89, 326–350.

Effler, S.W., Perkins, M., Peng, F., Strait, C., Weiderman, A.D., Auer, M.T., 2010. Light-absorbing components in lake superior. J. Great Lakes Research 356, 656–665.

Gohin, F., Druon, N., Lampert, L., 2002. A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. International Journal of Remote Sensing 23, 1639–1661.

Gohin, F., Loyer, S., Lunven, M., Labry, C., Froidefond, J.M., Delmas, D., Huret, M., Herbland, A., 2005. Satellite-derived parameters for biological modelling in coastal waters: Illustration over the eastern continental shelf of Bay fo Biscay. Remote Sensing of Environment 95, 29–46.

Hartigan, J., 1969. Using subsample values as typical values. Journal of the American Statistical Association 64, 1303–1317.

Kerfoot, W.C., Budd, J.W., Green, S.A., Cotner, J.B., Biddanda, B.A., Schwab, D.J., Vanderploeg, H.A., 2008. Doughnut in the desert: Late-winter production pulse in southern Lake Michigan. Limnology and Oceanography 53, 589–604.

Kerfoot, W.C., Yousef, F., Green, S.A., Budd, J.W., Schwab, D.J., Vanderploeg, H.A., 2010. Approaching storm: Disappearing winter bloom in Lake Michigan. J. Great Lakes Research 36, 30–41.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., Zhang, H., 2011. Remote estimation of chlorophyll a in optically complex waters based on optical classificaiton. Remote Sensing of Environment 115, 725–737.

Lee, Z., Hu, C., 2006. Global distribuiton of Case-1 waters: an analysis from SeaWiFS measurements. Remote Sensing of Environment 101, 270–276.

Legates, D.R., McCabe, G.J., 1999. Evaluating the use of “goodness of fit” measures in hydrologic and hydroclimatic model validation. Water Resources Research 35, 233–241.

Legates, D.R., McCabe, G.J., 2012. A refined index of model performance: a rejoinder. International Journal of Climatology .

Lesht, B.M., Barbiero, R.P., Warren, G.J., 2012. Satellite ocean color algorithms: a review of applications to the Great Lakes. J. Great Lakes Research 38, 49–60.

Lesht, B.M., Stroud, J.R., McCormick, M.J., Fahnenstiel, G., Stein, M.L., Welty, L.J., Leshkevich, G.A., 2002. An event-driven phytoplankton bloom in southern Lake Michigan observed by satellite. Geophysical Research Letters 29, 5 pp.

Li, H., Budd, J.W., Green, S.A., 2004. Evaluation and regionalization optimization of bio-optical algorithms for central lake superior. Journal of Great Lakes Research–Special Issue on Lake Superior 30, 443–458.

Li, Y., Wang, Q., Wu, C., Zhao, S., Xu, X., Wang, Y., Huang, C., 2012. Estimation of chlorophyll a concentration using NIR/Red bands of MERIS and classification procedure in inland turbid water. IEEE Transactions on Geoscience and Remote Sensing 50, 988–997.

Lohrenz, S., Fahnenstiel, G., Schofield, O., Millie, D., 2008. Coastal sediment dynamics and river discharge as key factors influencing coastal ecosystem productivity in southeastern Lake Michigan. Oceanography 21, 61–69.

Lohrenz, S.E., Fahnenstiel, G.L., Millie, D.F., Schofield, O.M.E., Johengen, T., Bergmann, T., 2004. Spring phytoplankton photosynthesis, growth, and primary production and relationships to a recurrent coastal sediment plume and river inputs in southeastern Lake Michigan. J. Geophys. Res. 109.

Matsushita, B., Yang, W., Chang, P., Yang, F., Fukushima, T., 2012. A simple method for distinguishing global Case-1 and Case-2 waters using SeaWiFS measurements. ISPRS Journal of Photogrammetry and Remote Sensing 69, 74–87.

McClain, C.R., Esaias, W.E., Barnes, W., Guenther, B., Endres, D., Hooker, S.B., Mitchell, B.G., Barnes, R.A., 1992. SeaWiFS Calibration and Validation Quality Control Plan. NASA Tech. Memo. 104566, Vol. 3. NASA Goddard Space Flight Center. Greenbelt, Maryland 20771.

McKee, D., Cunningham, A., Dudek, A., 2007. Optical water type discrimination and tuning remote sensing band-ratio algorithms: Application to retrieval of chlorophyll and $k_d(490)$ in the Irish and Celtic seas. Estuarine and Coastal and Shelf Science 73, 827–834.

Morel, A., Prieur, L., 1977. Analysis of variation in ocean color. Limnology and Oceanography 22, 709–722.

Moriasi, D., Arnold, J., Liew, M.W.V., Bingner, R., Harmel, R., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 50, 885–900.

Mortimer, C., 1988. Discoveries and testable hypotheses arising from coastal zone color scanner imagery of southern Lake Michigan. Limnology and Oceanography 33, 203–226.

Novoa, S., Chust, G., Valencia, V., Froidefond, J.M., Morichon, D., 2012. Estimation of chlorophyll-*a* concentration in waters over the continental shelf of the Bay of Biscay: a comparison of remote sensing algorithms. International Journal of Remote Sensing 32, 8349–8371.

O'Donnell, D.M., Effler, S.W., Strait, C.M., Leshkevich, G.A., 2010. Optical characteristics and pursuit of optical closure for the western basin of lake erie through *in situ* measurements. J. Great Lakes Research 36, 736–746.

O'Reilly, J.E., Maritorena, S., Mitchell, B., Siegel, D.A., , Carder, K., Garver, S., Kahru, M., McClain, C., 1998. Ocean Color Chlorophyll a Algorithms for SeaWiFS. J. Geophys. Res. 103, 24,937–24,953.

Peng, F., Effler, S.W., 2010. Characterizations of individual suspended mineral particles in western lake erie: implications for light scattering and water clarity. J. Great Lakes Research 36, 686–698.

Pozdnyakov, D., Shuchman, R., Korosov, A., Hatt, C., 2005. Operational algorithm for the retrieval of water quality in the Great Lakes. Remote Sensing of Environment 9, 352–370.

Preiur, L., Sathyendranath, S., 1981. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials. Limnology and Oceanography 26, 671–689.

Press, W., Teukolsky, S.A., 1992. Fitting straight line data with errors in both coordinates. Comput. Phys. 6, 274–276.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, 40 West 20th Street, New York, New York 10011-4211. 2nd edition.

Sathyendranath, S., Cota, G., Stuart, V., Maass, H., Platt, T., 2001. Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. International Journal of Remote Sensing 22, 249–273.

Sathyendranath, S., Lazzara, L., Preiur, L., 1987. Variations in the spectral values of specific absorption of phytoplankton. Limnology and Oceanography 32, 403–415.

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., Maass, H., 2004. Discrimination of diatoms from other phytoplankton using ocean-colour data. Marine Ecology - Progress Series 272, 59–68.

Shuchman, R., Korosov, A., Hatt, C., Pozdnyakov, D., Means, J., Meadows, G., 2006. Verification and application of a bio-optical algorithm for Lake Michigan using SeaWiFS: a 7-year inter-annual analysis. J. Great Lakes Research 32, 258–279.

Siswanto, E., Tang, J., Yamaguchi, H., Ahn, Y.H., Ishizaka, J., Yoo, S., Kim, S.W., Kiyomoto, Y., Yamada, K., Chiang, C., Kawamura, H., 2011. Empirical ocean-color algorithms to retrieve chlorophyll-*a*, total suspended matter, and colored dissolved organic matter absorption coefficients in the Yellow and East China Seas. J. Oceanogr .

Stuart, V., Sathyendranath, S., Head, E., Platt, T., Irwin, B., Maass, H., 2000. Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea. Marine Ecology - Progress Series 201, 91–106.

Tassan, S., 1994. Local algorithms using SeaWiFS data for the retrieval of phytoplankton, pigments, suspended sediment, and yellow substances in coastal waters. Applied Optics 33, 2369–2378.

Tilstone, G.H., Angel-Benavides, I.M., Pradhan, Y., Shutler, J.D., Groom, S., Sathyendranath, S., 2011. An assessment of chlorophyll-a algorithms available for SeaWiFS in coastal and open areas of the Bay of Bengal and Arabian Sea. Remote Sensing of Environment 115, 2277–2291.

Volpe, V., Silvestri, S., Marani, M., 2011. Remote sensing retrieval of suspended sediment concentration in shallow waters. Remote Sensing of Environment 115, 44–54.

Watkins, J.M., 2009. Comparison of shipboard and satellite measurements of surface water temperature and chlorophyll *a* in Lake Ontario. Aquatic Ecosystem Health & Management 12, 271–280.

Wei, H., Nearing, M., Stone, J., Breshears, D., 2008. A dual Monte Carlo approach to estimate model uncertainty and its application to the rangeland hydrology and erosion model. Transactions of the ASABE 51, 515–520.

Welschmeyer, N., 1994. Fluorometric analysis of chlorophyll-*a* in the presence of chlorophyll-*b* and pheopigments. Limnology and Oceanography 39, 1985–1992.

Werdell, P.J., Bailey, S., Fargio, G., Pietras, C., Knobelspiesse, K., Feldman, G., McClain, C., 2003. Unique data repository facilitates ocean color sensor validation. EOS Transactions of the American Geophysical Union 84, 377.

Werdell, P.J., Franz, B.A., Bailey, S.W., Jr., L.W.H., Feldman, G.C., 2007. Approach for the long-term spatial and temporal evaluation of ocean color satellite data products in a coastal environment, in: Frouin, R.J. (Ed.), Coastal Ocean Remote Sensing, pp. G1–G12.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bulletin Amer. Meteorol. Soc. 63, 1309–1313.

Willmott, C.J., Matsuura, K., Robeson, S.M., 2009. Ambiguities inherent in sums-of-squares-based error statistics. Atmospheric Environment 43, 749–752.

Willmott, C.J., Robeson, S.M., Matsuura, K., 2011. A refined index of model performance. International Journal of Climatology .

Witter, D.L., Ortiz, J.D., Palm, S., Heath, R.T., Budd, J.W., 2009. Assessing the the application of SeaWiFS ocean color algorithm to Lake Erie. Journal of Great Lakes Research 35, 361–370.

Wortman, S.R., 2005. Satellite Observations of Lake Michigan Whiting Events, 1997-2006. Master's thesis. University of Illinois at Chicago. 845 W. Taylor, St., Chicago, IL 60607.

Yoder, J., McClain, C., Feldman, G., Esaias, W., 1993. Annual cycles of phytoplankton chlorophyll concentrations in the global ocean: A satellite view. Global Biochem. Cycles 7, 181–193.

Tables

MODIS	Lake					
Year	Erie	Huron	Michigan	Ontario	Superior	Total
2002	12	7	0	4	8	31
2003	32	11	16	1	21	81
2004	28	12	6	6	20	72
2005	23	18	21	15	21	98
2006	28	6	16	15	22	87
2007	18	12	9	12	22	73
2008	29	25	12	5	28	99
2009	22	21	11	13	16	83
2010	11	11	10	9	26	60
2011	30	20	3	11	34	98
Total	233	143	104	84	218	
SeaWiFS						
Year	Erie	Huron	Michigan	Ontario	Superior	Total
2002	29	3	12	7	8	59
2003	24	10	13	1	4	52
2004	23	9	13	4	5	54
2005	23	14	14	12	9	72
2006	35	13	12	2	8	70
2007	17	7	13	12	8	57
2008	5	9	1	0	3	18
2009	18	11	6	5	0	40
2010	1	6	7	2	16	32
Total	175	82	91	45	61	

Table 1: Number of Matched Field/Satellite Samples

Sensor	a_0	a_1	a_2	a_3
MODIS	0.3429	-3.3925	3.3412	0.7857
SeaWiFS	0.4006	-4.0975	10.6576	-16.4647

Table 2: Coefficients of the GLF model ($log_{10}(Chl_a^{mod}) = a_0 + a_1{*}X + a_2{*}X^2 + a_3{*}X^3$) for MODIS and SeaWiFS. X = $\log_{10}$ of the maximum band ratio (MBR).

Model	Intercept	Slope	r	Bias	$\sigma_{pred}/\sigma_{obs}$	d_r	RMSE	%USE	MAE	N
All data										
OC3M	-0.074	0.892	0.891	0.090	0.893	0.761	0.203	0.640	0.154	782
MOD3-GLF	-0.000	1.040	0.891	-0.006	1.041	0.780	0.191	0.976	0.142	782
OC4	-0.048	0.844	0.841	0.078	0.845	0.739	0.229	0.631	0.170	454
SWF4-GLF	0.001	1.052	0.834	-0.012	1.053	0.758	0.237	0.956	0.158	454
Spring data										
OC3M	-0.060	0.742	0.893	0.106	0.742	0.756	0.231	0.391	0.172	408
MOD3-GLF	0.006	0.875	0.899	0.016	0.876	0.798	0.191	0.759	0.143	408
OC4	-0.013	0.789	0.797	0.065	0.789	0.719	0.266	0.586	0.200	252
SWF4-GLF	0.038	1.023	0.791	-0.044	1.022	0.736	0.283	0.892	0.188	252
Summer data										
OC3M	-0.085	1.094	0.924	0.073	1.094	0.771	0.167	0.808	0.135	374
MOD3-GLF	-0.002	1.263	0.917	-0.029	1.264	0.760	0.192	0.889	0.141	374
OC4	-0.083	0.913	0.911	0.095	0.913	0.776	0.173	0.582	0.133	202
SWF4-GLF	-0.038	1.067	0.906	0.029	1.067	0.797	0.161	0.963	0.121	202

Table 3: Statistics of NASA and GLF models fit to data [$log_{10}(Chl_a^{mod})$ vs. $log_{10}(Chl_a^{insitu})$] with seasonal subsets (see Methods for explanation of abbreviations in column headings).

Lake	Intercept	Slope	r	Bias	$\sigma_{pred}/\sigma_{obs}$	d_r	RMSE	%USE	MAE	N
MODIS										
Erie	0.087	0.917	0.675	-0.037	0.917	0.624	0.268	0.745	0.209	233
Huron	-0.005	1.010	0.723	0.007	1.011	0.640	0.117	0.868	0.091	143
Michigan	-0.063	0.999	0.767	0.063	0.999	0.608	0.128	0.670	0.105	104
Ontario	-0.118	1.448	0.601	-0.023	1.447	0.415	0.214	0.976	0.181	84
Superior	-0.007	0.938	0.390	0.002	0.938	0.472	0.143	0.649	0.108	218
SeaWiFS										
Erie	0.110	0.906	0.639	-0.059	0.906	0.619	0.302	0.704	0.215	175
Huron	0.037	1.165	0.729	-0.010	1.166	0.613	0.139	0.961	0.102	82
Michigan	-0.068	0.937	0.724	0.066	0.937	0.603	0.134	0.609	0.109	91
Ontario	-0.087	1.030	0.662	0.077	1.031	0.576	0.190	0.714	0.138	45
Superior	0.032	1.247	0.466	-0.016	1.246	0.430	0.154	0.864	0.121	59

Table 4: Statistics of GLF model fit to data [$log_{10}(Chl_a^{mod})$ vs. $log_{10}(Chl_a^{insitu})$] for all years by lake (see Methods for explanation of abbreviations in column headings).

Year	Intercept	Slope	r	Bias	$\sigma_{pred}/\sigma_{obs}$	d_r	RMSE	%USE	MAE	N
					MODIS					
2003	-0.008	0.948	0.921	0.019	0.948	0.817	0.157	0.881	0.124	81
2004	0.021	1.116	0.854	-0.036	1.117	0.740	0.169	0.947	0.116	72
2005	-0.028	1.150	0.887	-0.000	1.150	0.750	0.220	0.999	0.149	98
2006	0.045	1.109	0.916	-0.059	1.109	0.779	0.176	0.886	0.139	87
2007	-0.056	0.911	0.884	0.070	0.911	0.764	0.216	0.739	0.166	73
2008	0.022	0.914	0.895	-0.015	0.914	0.805	0.204	0.829	0.151	99
2009	-0.029	0.964	0.885	0.035	0.964	0.778	0.179	0.870	0.136	83
2010	-0.049	1.131	0.914	0.042	1.132	0.734	0.135	0.898	0.108	60
2011	0.017	1.107	0.907	-0.037	1.107	0.785	0.230	0.974	0.179	98
					SeaWiFS					
2002	0.093	1.039	0.761	-0.099	1.039	0.685	0.252	0.771	0.176	59
2003	-0.046	0.970	0.916	0.056	0.970	0.803	0.164	0.819	0.129	52
2004	0.004	1.012	0.812	-0.006	1.011	0.744	0.201	0.915	0.136	54
2005	-0.079	1.113	0.881	0.052	1.114	0.738	0.226	0.947	0.161	72
2006	0.082	1.169	0.842	-0.114	1.170	0.756	0.298	0.853	0.186	70
2007	-0.041	0.940	0.808	0.055	0.939	0.769	0.288	0.809	0.180	57
2008	0.064	1.324	0.922	-0.086	1.325	0.790	0.220	0.712	0.130	18
2009	-0.033	0.954	0.900	0.045	0.954	0.822	0.200	0.851	0.140	40
2010	0.050	1.417	0.586	-0.034	1.418	0.307	0.179	0.943	0.144	32

Table 5: Statistics of GLF model fit to data [$log_{10}(Chl_a^{mod})$ vs. $log_{10}(Chl_a^{insitu})$] from all lakes by year. MODIS began producing data after the GLNPO sampling was completed in 2002 (see Methods for explanation of abbreviations in column headings).

Statistic	Mean	σ	Minimum	Maximum
MODIS				
Intercept	-0.001	0.023	-0.061	0.063
Slope	1.057	0.089	0.878	1.348
r	0.888	0.009	0.831	0.906
Bias	0.008	0.029	-0.066	0.081
d_r	0.772	0.013	0.736	0.798
RMSE	0.200	0.023	0.163	0.324
%USE	0.937	0.053	0.724	1.000
MAE	0.146	0.010	0.124	0.182
SeaWiFS				
Intercept	0.002	0.046	-0.096	0.112
Slope	1.062	0.102	0.864	1.342
r	0.695	0.033	0.605	0.786
Bias	0.013	0.057	-0.126	0.142
d_r	0.743	0.020	0.674	0.791
RMSE	0.247	0.028	0.186	0.319
%USE	0.889	0.082	0.572	1.000
MAE	0.167	0.012	0.137	0.199

Table 6: Statistics of $log_{10}(Chl_a^{mod})$ vs. $log_{10}(Chl_a^{insitu})$ for five-year partitioned model fits. Coefficients calibrated with five-year partitions were used to model the observed chlorophyll in the complementary partitions (see Methods for explanation of abbreviations in row labels).

$\overline{log_{10}(MBR)}$	$\overline{log_{10}(Chl_a^{mod})}$	$Chl_a^{mod} - \sigma$ (mg/m^3)	Chl_a^{mod} (mg/m^3)	$Chl_a^{mod} + \sigma$ (mg/m^3)	10% quantile (mg/m^3)	90% quantile (mg/m^3)
			MODIS			
0.349	-0.383	0.381	0.414	0.449	0.371	0.456
0.249	-0.280	0.472	0.525	0.584	0.457	0.609
0.149	-0.087	0.698	0.819	0.960	0.663	1.022
0.050	0.184	1.239	1.527	1.882	1.156	2.039
-0.050	0.522	2.599	3.330	4.267	2.385	4.677
-0.150	0.913	6.050	8.179	11.058	5.557	12.447
-0.250	1.354	13.763	22.595	37.093	12.759	44.610
			SeaWiFS			
0.349	-0.436	0.306	0.366	0.438	0.289	0.454
0.249	-0.237	0.502	0.579	0.669	0.479	0.701
0.149	-0.027	0.800	0.941	1.106	0.762	1.174
0.049	0.248	1.409	1.770	2.223	1.310	2.429
-0.050	0.635	3.142	4.320	5.939	2.854	6.712
-0.149	1.187	8.834	15.386	26.798	8.050	34.042

Table 7: Estimated uncertainty ($\pm 1\sigma$ and empirical 80% confidence intervals) for GLF chlorophyll retrievals. Overbars indicate the averages of the Monte Carlo run values falling in 0.1 bins of $log_{10}(MBR)$. Chl_a^{mod} is $10.^{(log_{10}(Chl_a^{mod}))}$ and $Chl_a^{mod} \pm \sigma$ is $10.^{(log_{10}(Chl_a^{mod}) \pm \sigma(log_{10}(Chl_a^{mod}))}$.

List of Figures

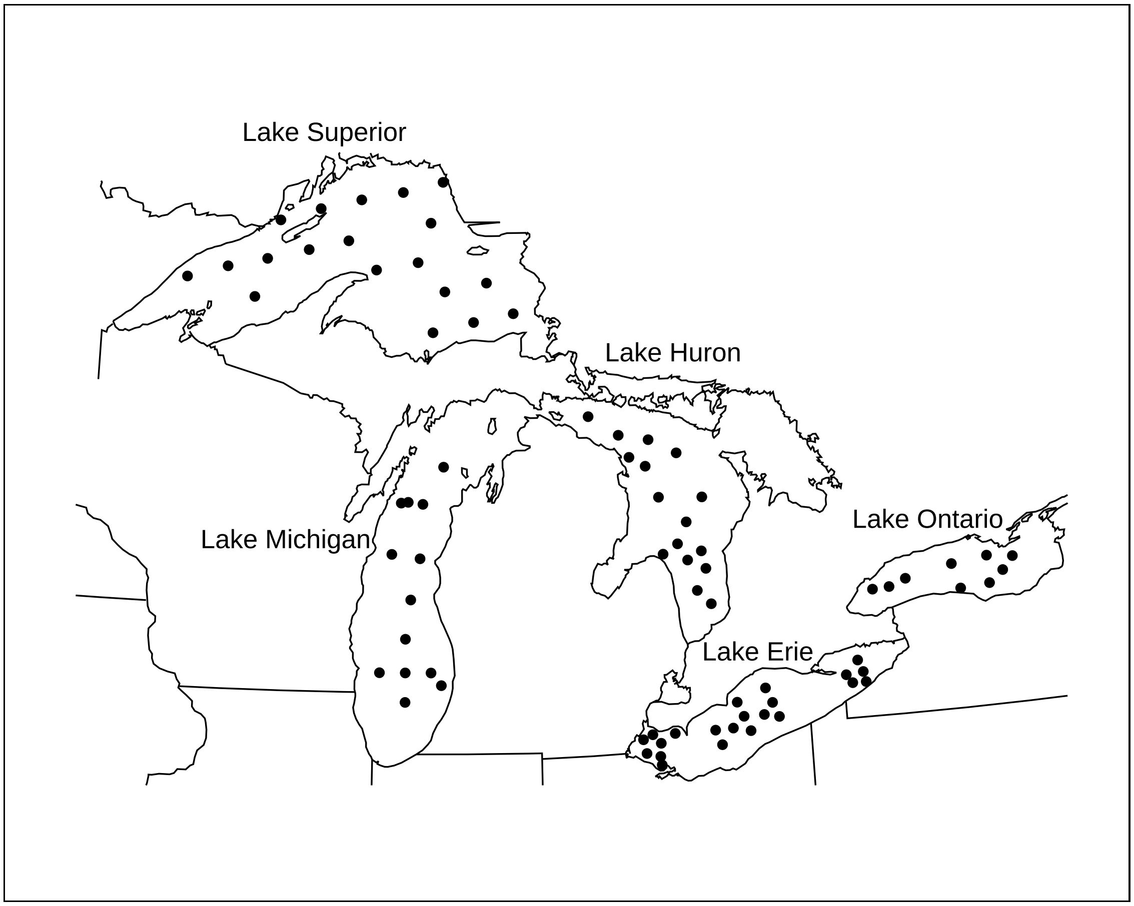

Figure 1: Locations of GLNPO WQS stations.

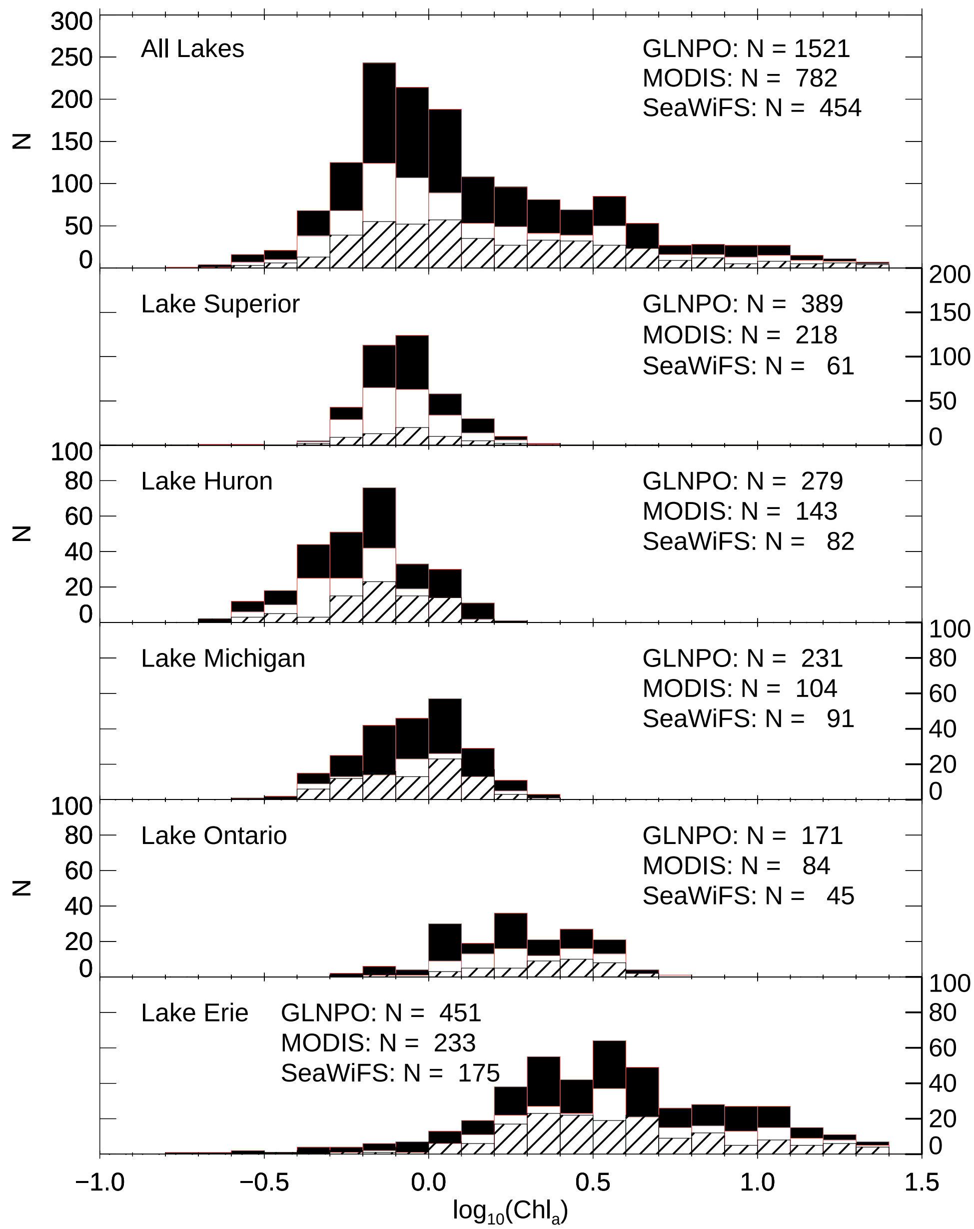

Figure 2: Histograms of GLNPO chlorophyll-a measurements, 2002-2011. Black bars represent the entire set of field data, the white bars represent the field data that were matched with MODIS observations (2002-2011), and the hatched bars represent the field data that were matched with SeaWiFS observations (2002-2010).

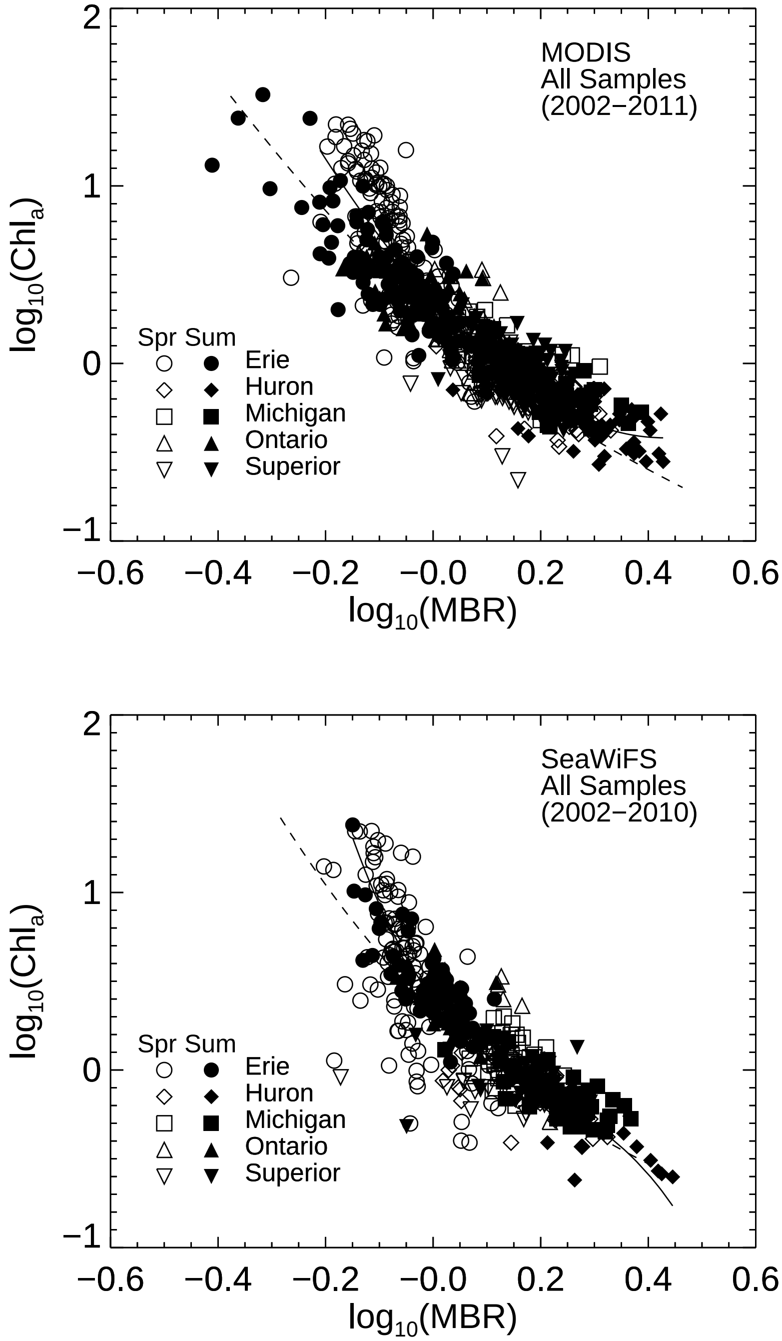

Figure 3: $\log_{10}$(observed chlorophyll) vs. $\log_{10}$(maximum band ratio) for MODIS (top panel) and SeaWiFS (bottom panel). Dashed lines show the standard NASA algorithms. The GLF model for each set is shown by the solid lines.

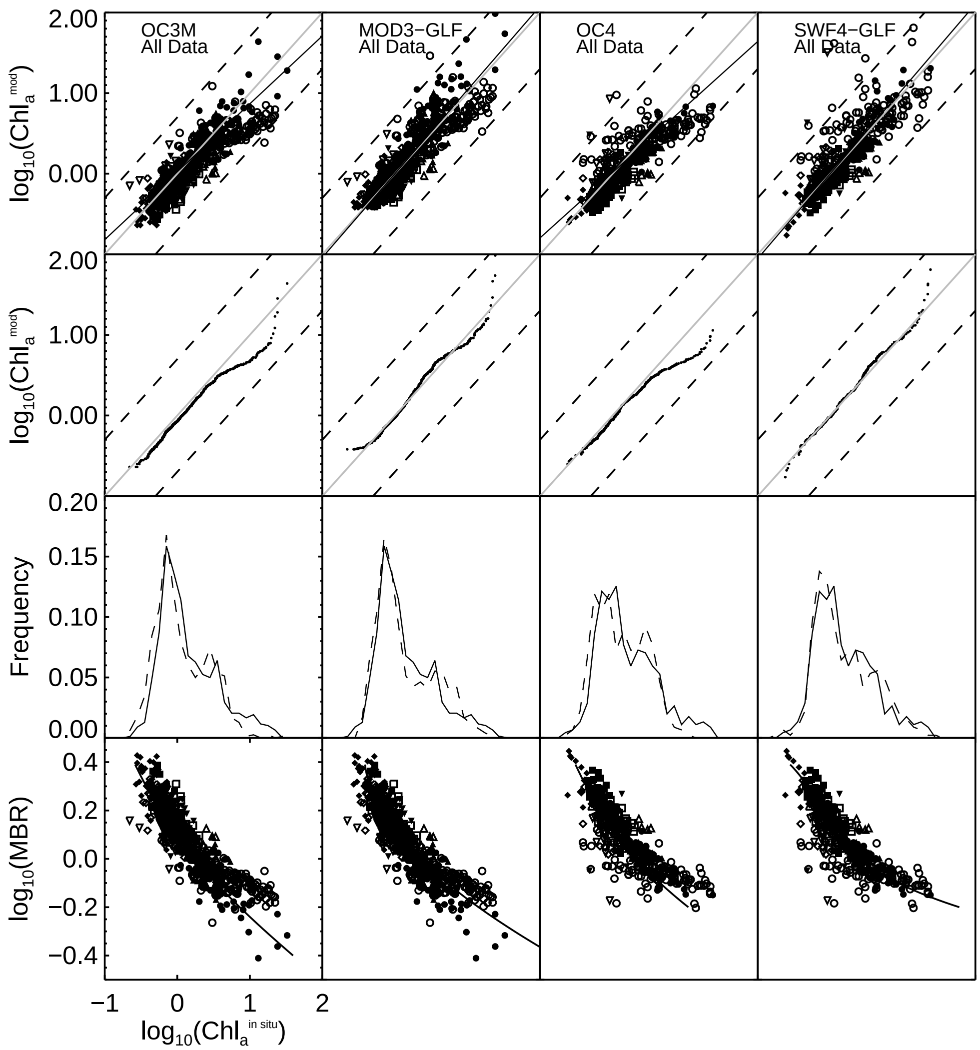

Figure 4: Comparisons between modeled (Chl_a^{mod}) and observed chlorophyll (Chl_a^{insitu}) data: From top to bottom: Scatterplots (1:1 indicated by gray line, regression indicated by black line); quantile-quantile plots; relative frequency of *in situ* (solid line) and modeled (dashed line) values; maximum band ratio versus *in situ* Chl_a (symbols) and maximum band ratio (MBR) versus model (curve). Note that the x-axes for each row of figures are shown in column 1. Also shown in the second panel from the top are lines indicating model:*in situ* ratios of 1:5 and 5:1. From left to right columns are NASA OC3Mv6, MODIS GLF, NASA OC4v6, and SeaWiFS GLF.

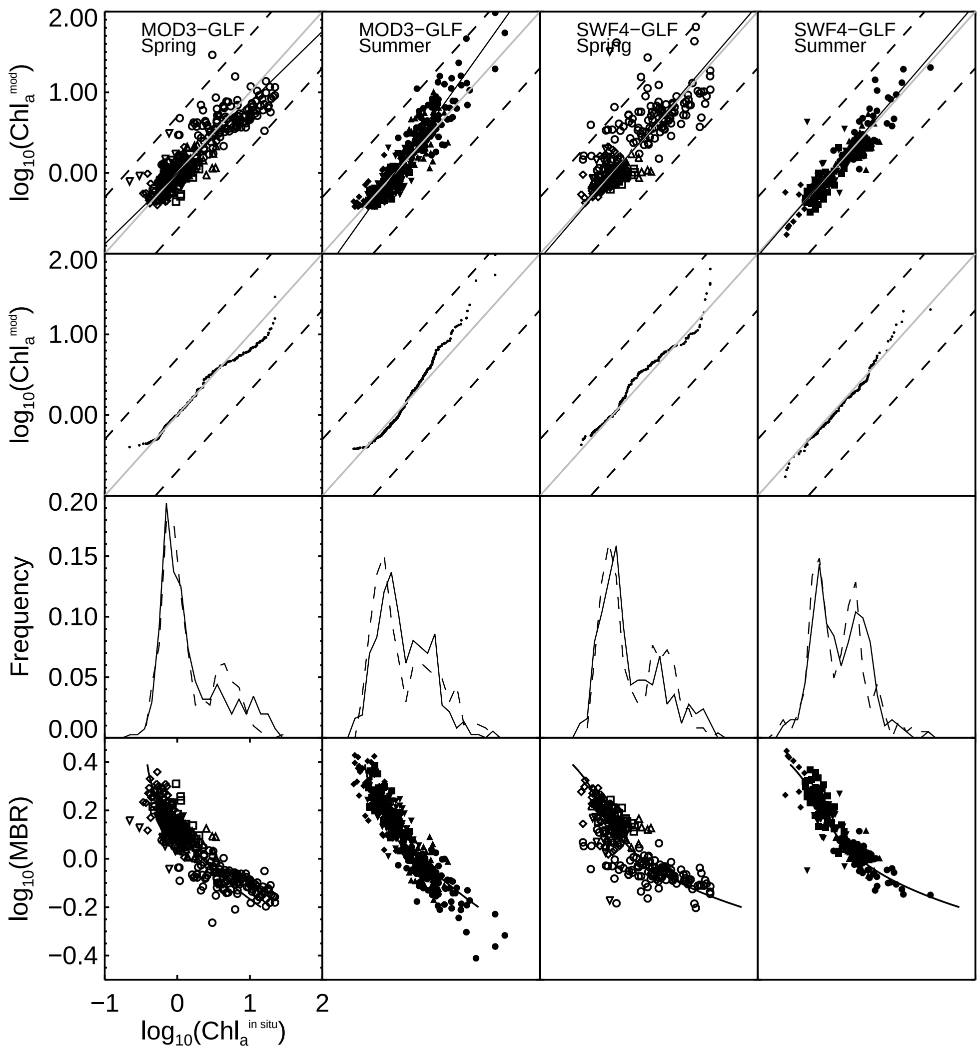

Figure 5: Comparison of GLF modeled (Chl_a^{mod} and observed chlorophyll (Chl_a^{mod}) data by sensor and season. Panels from top to bottom are as described in Fig. 4. From left to right the columns are MODIS GLF spring, MODIS GLF summer, SeaWiFS GLF spring, SeaWiFS GLF summer. Symbols are the same as those used in Fig. 3 and Fig. 4

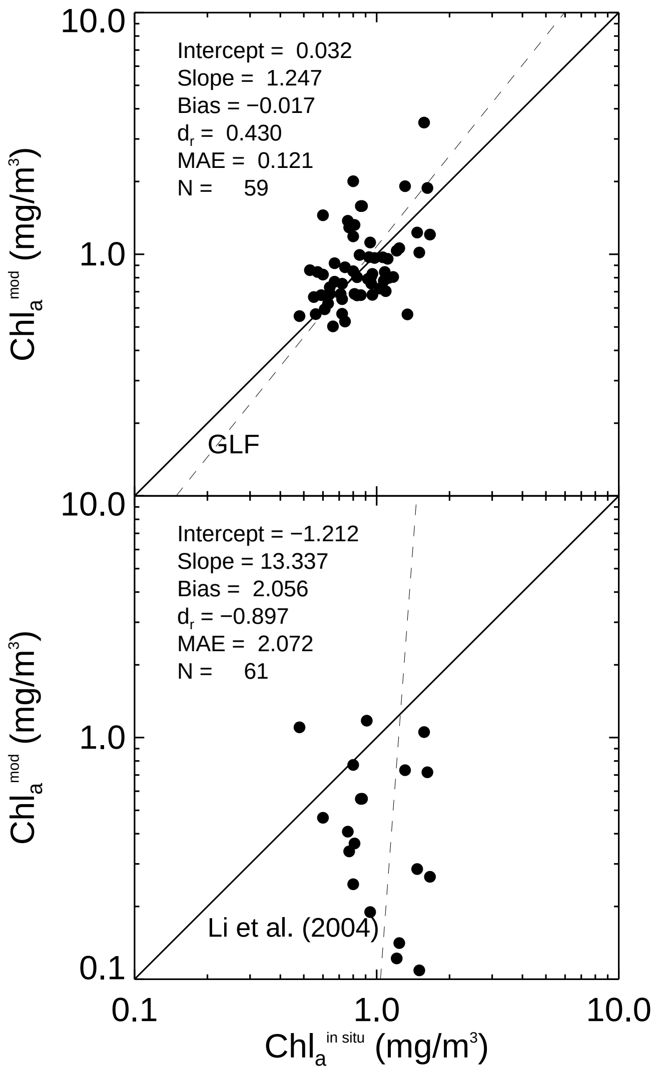

Figure 6: Chlorophyll (Chl_a^{mod}) in Lake Superior predicted by the GLF model (top panel) and the Li et al. (2004) model (bottom panel) versus observed chlorophyll (Chl_a^{insitu}). Results for the Li et al. (2004) model that produced values of chlorophyll < 0.1 mg/m^3 are not plotted nor included in the statistics (see Methods for description of abbreviations in legend). Two outliers with GLF modeled chlorophyll values > 10 mg/m^3 are not plotted nor included in the statistics shown. Solid line is 1:1 and dashed line is the model regression.

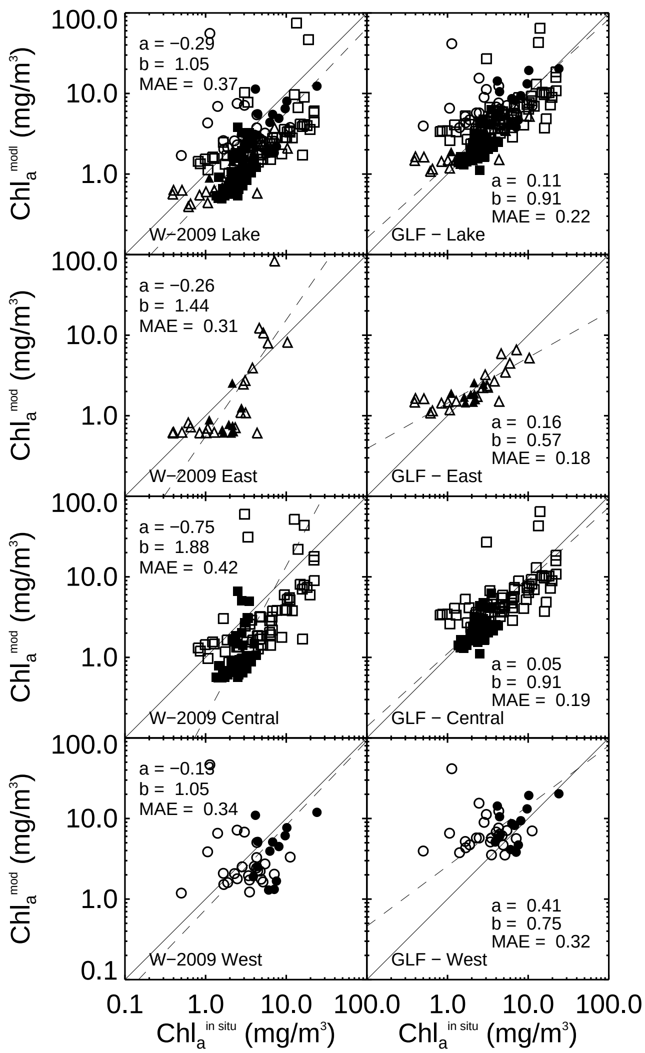

Figure 7: Chlorophyll (Chl_a^{mod}) in Lake Erie predicted by the GLF model (right column) and the Witter et al. (2009) whole lake and basin-tuned models (left column) versus observed chlorophyll. Rows are (from top) all Lake Erie stations, eastern basin stations (triangles), central basin stations (squares), and western basin stations (circles). Open symbols represent samples collected in the spring and filled symbols those collected in the summer. Solid line is 1:1 and dashed line is model regression. See methods for explanation of statistics abbreviations.

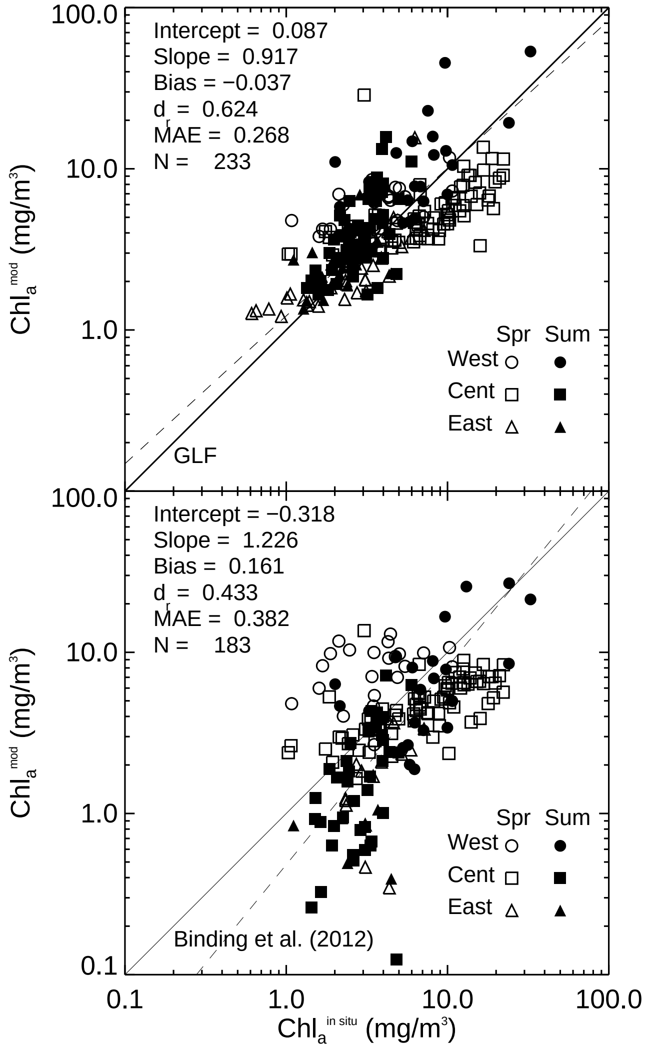

Figure 8: Chlorophyll (Chl_a^{mod}) in Lake Erie predicted by the GLF model (top) and the Binding et al. (2012) model (bottom) versus observed chlorophyll (Chl_a^{insitu}). Results for the Binding et al. (2012) model were limited to those that produced non-negative values of chlorophyll. Lake Erie basins are designated by symbols (open for spring samples and solid for summer samples). Solid line is 1:1 and dashed line is model regression. See Methods for explanation of abbreviations in legend.

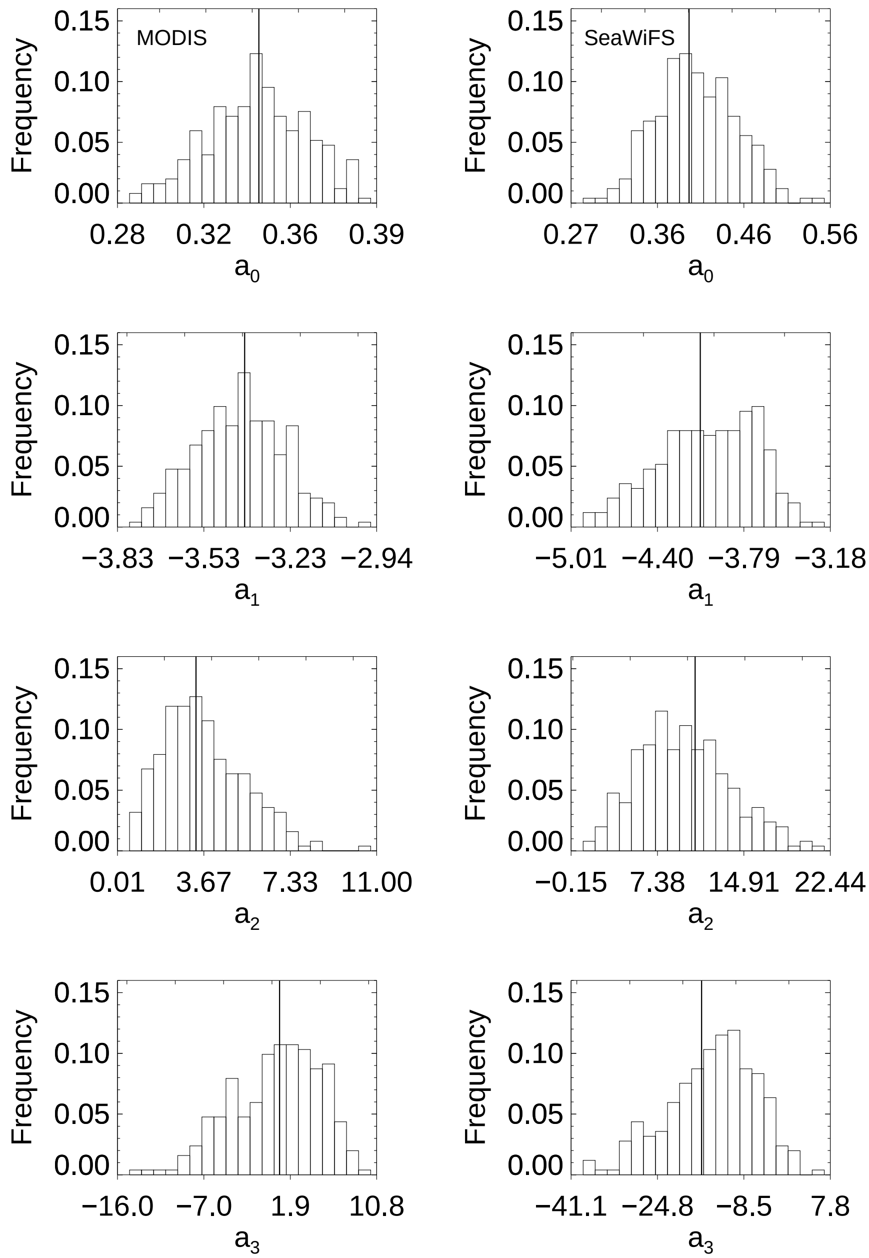

Figure 9: Frequency distributions of GLF model parameters (a_0, a_1, a_2, a_3) obtained from five-year subsets of the complete data record. Left column shows results for MODIS, right column for SeaWiFS. Vertical lines indicate parameter values determined from fit to entire dataset.

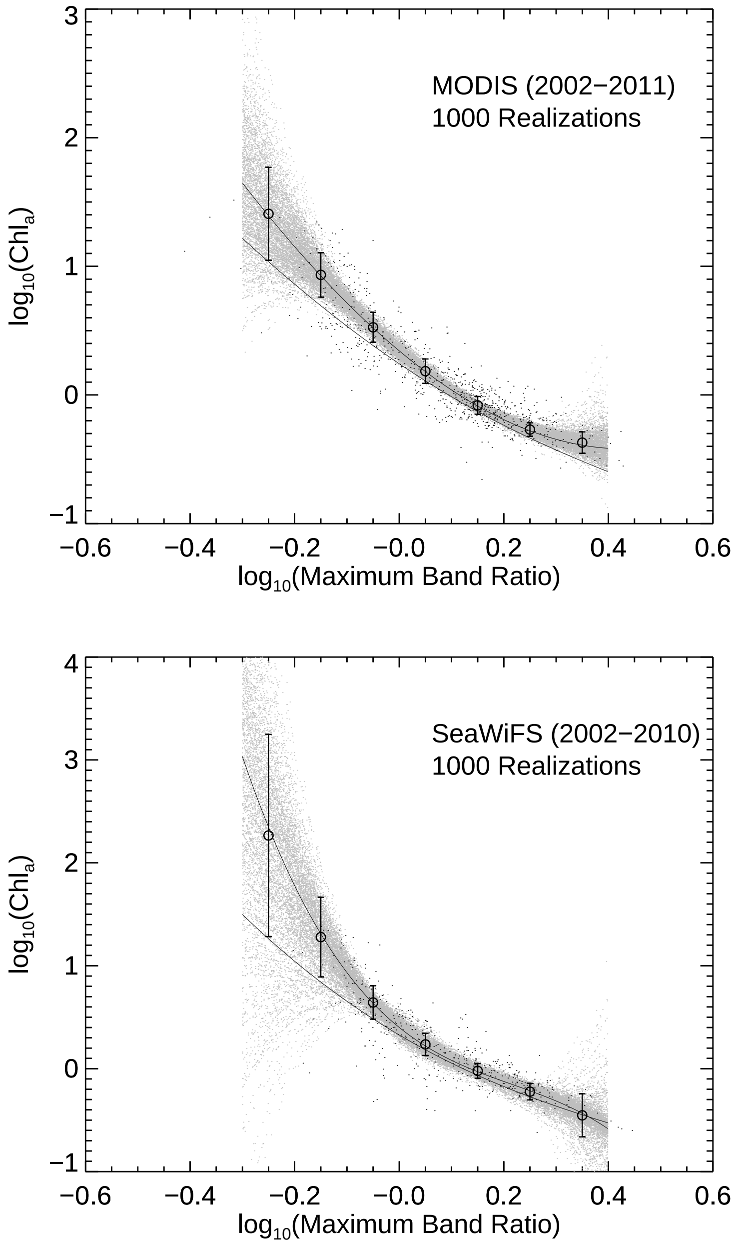

Figure 10: Monte-Carlo (M-C) fits of the 3rd order model (Eq. 2) relating $\log_{10}$(chlorophyll) to $\log_{10}$(maximum band ratio) for MODIS (top panel) and SeaWiFS (bottom panel). Samples from M-C runs are plotted in gray. The points and error bars show the average ± one standard deviation of the samples within 0.1 intervals of $\log_{10}$(maximum band ratio). The upper curve is the GLF model and the lower curve is the standard NASA algorithm.

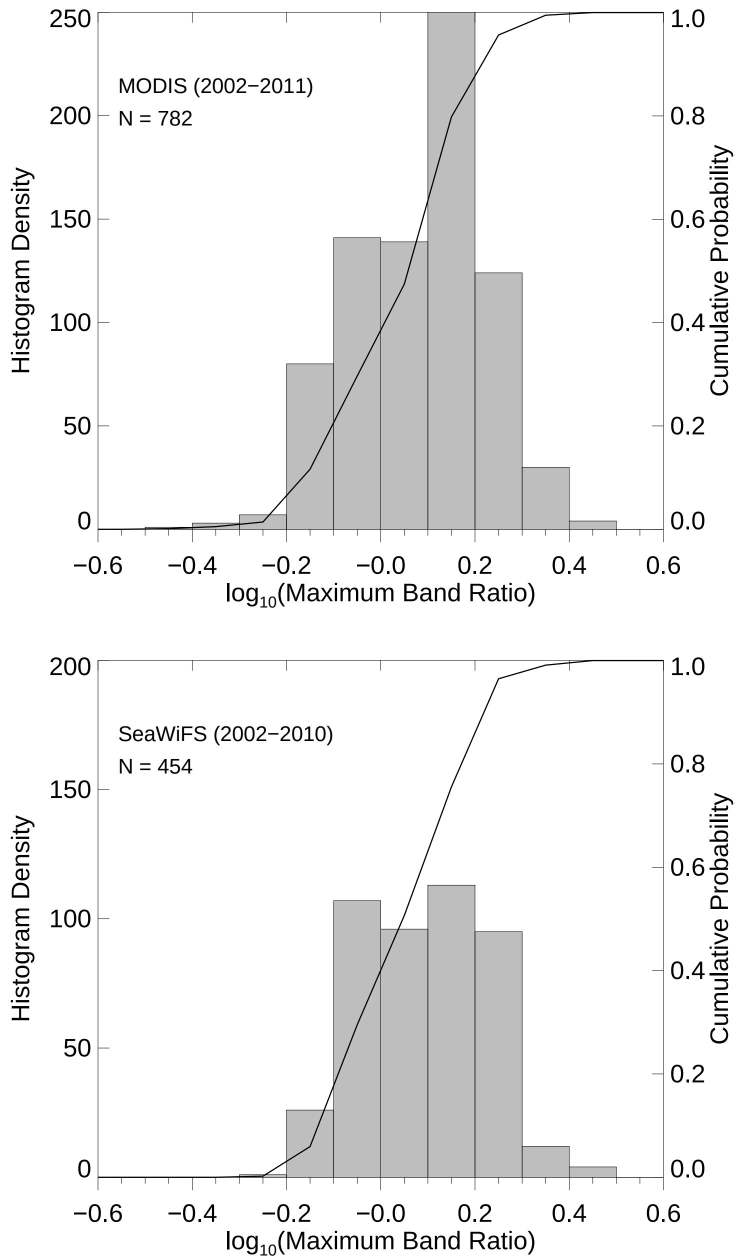

Figure 11: Histograms of $\log_{10}$(maximum band ratio) from the MODIS (top panel) and SeaWiFS (bottom panel) observations that were matched with GLNPO field samples.