# Satellite ocean color algorithms: a review of applications to the Great Lakes

Barry M. Lesht<sup>a,\*</sup>, Richard P. Barbiero<sup>b</sup>, Glenn J. Warren<sup>c</sup>

<sup>a</sup>CSC and Department of Earth and Environmental Sciences, University of Illinois at Chicago, 845 W. Taylor St., Chicago, IL 60607, USA

<sup>b</sup>CSC and Loyola University of Chicago, 1359 W. Elmdale Ave. Suite 2, Chicago, IL 60660, USA <sup>c</sup>USEPA Great Lakes National Program Office, 77 W. Jackson Boulevard, Chicago, IL 60604, USA

#### Abstract

We review the literature relating to the retrieval of chlorophyll concentrations in the Great Lakes from satellite observations. Most studies show that the satellite estimates of lake chlorophyll concentration are linearly related to the observed concentrations, though they tend to overestimate concentrations at lower values and underestimate them at higher values. Deviations from a consistent, accurate, linear relationship can be attributed to temporal and spatial variations in the inherent optical properties of the color producing agents in the water as well as to varying concentrations of interfering substances such as suspended non-algal particles and colored dissolved organic matter. We confirmed these results by using a simple optical model to examine the sensitivity of the retrieved chlorophyll values to the concentrations of interfering substances and to differences in model parameters. Because the spatial and temporal optical properties of the Great Lakes are unpredictable, no retrieval method is likely to produce accurate results all the time. The papers we reviewed show that simple band ratio algorithms can provide chlorophyll estimates that are proportional to *in situ* concentrations. The bulk of the literature suggests that the band ratio methods will be of most value in regions where the concentrations of interfering substances such as dissolved organic material or suspended sediments are minimal. Because of these limitations we recommend that future papers presenting chlorophyll analysis based on satellite data provide confirming field observations that include measurements of suspended particles and dissolved organic carbon. We also recommend that Great Lakes scientists explore novel methods for retrieving chlorophyll concentrations from satellite observations that have proven useful in other optically complex waters.

Preprint submitted to Journal of Great Lakes Research

<sup>\*</sup>Corresponding author

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

Keywords: Remote sensing, Ocean color, Chlorophyll a, Satellite observation

#### Introduction

The number of studies that have used satellite observations to explore biological features of the Great Lakes has increased dramatically in recent years. Of the published research that has focused on estimation of chlorophyll *a* concentration, there are two general classes of papers; those that report using satellite data to explore the processes that affect the phytoplankton community size and distribution in the lakes (Lesht et al., 2002; Chen et al., 2004; Kerfoot et al., 2008; Lohrenz et al., 2008; Kerfoot et al., 2010; Barbiero et al., 2011) and those that present evaluations the methods used to convert satellite observations to estimates of chlorophyll *a* concentration (Bergmann et al., 2004; Budd and Warrington, 2004; Li et al., 2004; Shuchman et al., 2006; Witter et al., 2009). Although many of these studies use the same equations, often termed retrieval algorithms, to estimate chlorophyll concentration, their conclusions with respect to the applicability and success of the estimations can be quite different. In fact, because the literature can seem contradictory and confusing, we believe that a review of these papers is useful not only to assess the research that has been published, but also to understand the applicability and limitations of the methods and to suggest avenues for future research.

Our review includes studies of each of the lakes and includes data collected by three ocean color sensors; the Coastal Zone Color Scanner (CZCS), the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and the Moderate Resolution Imaging Spectroradiometer (MODIS). In analyzing these studies, we attempted to determine and describe the important aspects of the methods used by the original authors. However, because the original works varied in their purposes and approaches they also varied in the amounts of method-ological detail provided and for some papers we were unable to extract technical specifics that might have been useful. In a few cases we used the published data to reanalyze the original results in an attempt to clarify or generalize them. These cases are indicated in the text. Our review, however, is not intended to be a detailed study-by-study critique nor to be a presentation of new field or laboratory observations.

In the following section we briefly describe the history of United States efforts in satellite remote sensing of ocean color. Other space agencies, notably those in Japan and Europe, also have launched ocean color monitoring satellites, but these have not been used to any extent in studies of the Great Lakes. We then present the basic theory behind the retrieval algorithms used to convert the satellite observations to geophysical values, focusing first on multi-component algorithms and then on the more commonly used empirical band ratio algorithms. Because applications of empirical band ratio methods are much more common in the literature we devote the subsequent section to the studies that are based on these algorithms. Finally, we discuss and summarize our findings and conclusions.

#### Satellite remote sensing of ocean color

Satellite-borne ocean color sensors measure the sunlight scattered back toward the satellite from substances present in surface layer of the water. The depth to which the sunlight penetrates the water depends on the composition of the water and can vary from less than a meter in very turbid water to tens of meters in clear water where absorption and scattering by the water itself is the major light attenuating process. The first satellite ocean color sensor, the CZCS, was designed to determine if observations from space could be used to identify and quantify the suspended and dissolved substances that affect the color of ocean waters. CZCS measured the scattered sunlight at four discrete bands in the visible portion of the spectrum (each 20 nm wide, centered at 443 nm, 520 nm, 550 nm, and 670 nm). Intended as a proof-of-concept instrument, the CZCS operated from October 1978 through June 1986 when the sensor failed. Although many studies used the data collected in the 91-month life of the system to examine the distribution of phytoplankton in the oceans (Gordon et al., 1983; Singh et al., 1983; Pan et al., 1988) no similar applications were done in the Great Lakes, probably because of problems with the atmospheric correction algorithm (Tanis, 1984). A major result of the CZCS mission was the conclusion that although it was a successful proof-of-concept experiment, uncertainties in the quality of the radiometric data indicated that future ocean color sensors needed to be more stable and have more reliable radiometric calibrations (Evans and Gordon, 1994).

These requirements for radiometric calibration and stability were incorporated into the follow-on sensors: SeaWiFS and MODIS. SeaWiFS, was launched in September 1997 and produced data through early December of 2010. MODIS-A, which was launched in June 2002, still is producing data. The SeaWiFS sensor has detectors at six bands in the visible (412 nm, 443 nm, 490 nm, 510 nm, 555 nm, 670 nm) which were selected to facilitate global studies of ocean color. Unlike SeaWiFS which was designed specifically for ocean color monitoring, MODIS was designed for a wide range of Earth system studies. MODIS has 36 bands with different spatial and radiometric resolutions. Seven bands in the visible (412 nm, 442 nm, 488 nm, 531 nm, 551 nm, 667 nm, 678 nm) are intended primarily for ocean color studies. Both the SeaWiFS and MODIS instruments have higher radiometric resolution and wider dynamic range than did CZCS so the sensors are more sensitive and less likely to be saturated in regions with high concentrations of scattering material.

#### **Retrieval algorithms**

The numerical processes used to convert the radiometric measurements made by the satellite-borne sensors into interpretable geophysical values are referred to as retrieval algorithms. Typically, the geophysical variable (*e.g.* chlorophyll concentration) is expressed as a function of either the normalized radiances (radiant flux per unit area per unit solid angle per unit wavelength interval) or the normalized remote sensing reflectances – the ratio of the upwelling normalized radiance leaving the water to the incident irradiance (power per unit area per unit wavelength interval) on the water surface at some number of bands.

The basic assumption underlying all such algorithms is the existence of a unique relationship between the spectral content of the scattered sunlight measured by the satelliteborne sensor and the concentrations of color producing agents (CPAs, sometimes referred to as optically active components or OACs) in the water with which the sunlight interacts. Sunlight passing downward through the air-water interface is absorbed and scattered by the water and by the other materials dissolved and suspended in the water. This absorption and scattering is wavelength dependent and each CPA has a different absorption and/or scattering spectrum. In addition to the water itself, three CPAs usually are considered. These are suspended minerals or, more generally, non-algal particulates (NAP), colored dissolved organic material (CDOM), and phytoplankton.

As an example, Fig. 1 shows the average phytoplankton absorption spectrum determined by Lohrenz et al. (2004) for Lake Michigan from sampling done as part of the Episodic Events Great Lakes Experiment (EEGLE) program (http://www.glerl.noaa. gov/eegle/). This figure, which is typical of phytoplankton absorption spectra reported by others (Bukata et al., 1991b; Bergmann et al., 2004; Witter et al., 2009) shows that most communities of phytoplankton absorb most radiation at shorter (blue) wavelengths and very little in the middle part of the spectrum (green). As a result, green light is preferentially scattered from algae-rich waters. The radiance received at the satellite is proportional to the amount of light scattered, which, in the absence of other CPAs, depends on the concentration of chlorophyll. If other CPAs are present, the radiance also depends on the concentrations of those substances and on their absorption and scattering properties.

Retrieval algorithms can be classified by the assumptions made about the optical complexity of the water. Semi-analytical or multi-component models attempt to model the effects of multiple interacting substances. Empirical algorithms are based on the assumption that a single optical component (*e.g.* chlorophyll) is dominant. Both have been applied to studies of the Great Lakes, though because empirical algorithms are simpler to apply and are used in standard processing software, studies based on the empirical algorithms are much more common.

#### Semi-analytical and multi-component model algorithms

Because phytoplankton may not be the only CPA in the water, a complete retrieval algorithm would include the effects all of the constituents that interact with the incoming solar radiation. By considering each of these constituents in one retreival, multi-component algorithms are intended to determine simultaneously the concentrations of each from the sampled radiance spectrum. Many such algorithms have been developed and applied to oceanic and coastal waters (see, for example, recent papers by Garcia et al. (2006) and Chang et al. (2007)); here we concentrate on those that have been used in the Great Lakes.

The first application of these methods to the Great Lakes was the work done by Bukata and colleagues (Bukata et al., 1978, 1979, 1981b,a, 1985, 1991a,b). Basing their methods on the findings of Gordon et al. (1975), Bukata et al. (1979) used field measurements of upwelling and downwelling spectral irradiance along with measurements of spectral attenuation to determine the diffuse reflectance and attenuation coefficient profiles in five bands at ten stations along a nearshore (maximum depth 37 m) transect in western Lake Ontario. They used these calculations to compute the inherent optical properties (absorption coefficient (*a*), scattering coefficient (*b*), total attenuation coefficient (*c*), backscattering probability (*B*), forwardscattering probability (F = 1 - B), volume scattering function ( $\beta(\Theta)$ ), and scattering albedo ( $\omega_0$ ) of the water at the sampled locations.

By combining these calculations of the inherent optical properties with simultaneous measurements of chlorophyll- *a*, NAP, and dissolved organic carbon (DOC) concentrations (as a proxy for CDOM), Bukata et al. (1981b), were able to determine the optical cross sections of the CPAs in nearshore western Lake Ontario. The optical cross sections are essentially the absorption and scattering spectra characteristic of each constituent normalized by the concentration of the constituent. That is, the total absorption and scattering properties of the water (measured by the inherent optical properties) may be expressed as

$$a(\lambda) = a_w(\lambda) + x\hat{a}_{Chl}(\lambda) + y\hat{a}_{NAP}(\lambda) + z\hat{a}_{CDOM}(\lambda),$$
(1)  

$$b(\lambda) = b_w(\lambda) + x\hat{b}_{Chl}(\lambda) + y\hat{b}_{NAP}(\lambda),$$
(1)  

$$(Bb)(\lambda) = Bb_w(\lambda) + x\hat{B}b_{Chl}(\lambda) + y\hat{B}b_{NAP}(\lambda)$$

in which  $a(\lambda), b(\lambda)$ , and  $Bb(\lambda)$  are the total absorption, scattering, and backscatter coefficients at wavelength  $(\lambda)$ , x, y, and z are the concentrations of chlorophyll-a, NAP, and CDOM, and  $\hat{a}_i, \hat{b}_i$ , and  $\hat{B}b_i$  are the absorption, scattering, and backscattering cross sections of the  $i^{th}$  constituent (Chl, NAP, CDOM). The absorption and scattering coefficients of water are designated  $a_w, b_w$ , and  $Bb_w$ . Thus, given simultaneous measurements of the constituent concentrations and the inherent optical properties, the optical cross sections may be estimated by multivariate regression.

Once the optical cross sections were determined, Bukata et al. (1981a) showed that by

following Gordon et al. (1975) the irradiance reflectance  $R(0, \lambda)$  just below the free surface could be expressed as

$$R(0,\lambda) = \sum_{n=0}^{N} r_n(0) X^n$$
(2)

where  $r_n(0)$  are polynomial expansion coefficients determined by Gordon et al. (1975), and

$$X = \frac{(Bb)(\lambda)}{a(\lambda) + (Bb)(\lambda)}$$
(3)

as determined from equation set (1).

By calculating the propagation of the radiance through the water-air interface and subsequently through the atmosphere, it is possible to determine the theoretical spectra of radiances that would be received at a satellite sensor as a function of the constituent concentrations. The spectra shown in Fig. 2 are the result of this type of calculation for the simple case when chlorophyll and water are the only CPAs.

The process of using these equations to retrieve the concentrations of the constituents in the water from the radiance measured at the satellite is based on inverting the calculations. Measured radiance spectra and known (or hypothesized) optical cross sections are used along with non-linear optimization methods to retrieve the in-water concentrations of the several constituents. To test the methodology, Bukata et al. (1985) used an independent set of radiance data from Lake Ontario along with their directly determined cross sections to calculate concentrations of NAP (primarily mineral sediments), CDOM (parameterized by dissolved organic carbon), and chlorophyll in the lake (their Figures 33-35). Their model produced "excellent" predictions for NAP, "possibly acceptable" predictions for CDOM, but "nonexistent" predictions for chlorophyll. Specifically, the predicted chlorophyll values were much lower than the observed values and often close to zero.

Bukata et al. (1985) tested several possible explanations for the failure of the model to predict the observed chlorophyll concentrations in Lake Ontario. They concluded that the problem resulted from a combination of factors, including (a) dominant influence of the non-phytoplankton (NAP and CDOM) components on the reflectance spectrum and (b) the optical cross sections they used (which were determined from their original data set) were inappropriate for the second, independent, data set. When they used the second data set to derive new optical cross sections, they found that the absorption spectrum for chlorophyll differed considerably from the one derived from the first data set. Predictions made using the new set of cross sections were much improved.

A recent series of overlapping papers (Pozdnyakov et al., 2005a,b; Shuchman et al., 2006), applied the same bio-optical model used by Bukata et al. (1985) to archival observations made in Lake Michigan. The focus of the first two papers in this series is on

the optimization method used to invert the bio-optical model to obtain in-water concentrations. The authors use a combination of the Levenberg-Marquardt algorithm employed by Bukata et al. (1985) and a neural network model to speed the calculations to the point where the model could be implemented in an operational scheme. The third paper, described below, is focused on using the combined optimization method to examine chlorophyll patterns in Lake Michigan between the years 1998 and 2004.

To validate the performance of the algorithm, Shuchman et al. (2006) compared field measurements of chlorophyll made in the vicinity of the Kalamazoo River outflow on single days in July and September 2003 with satellite retrievals averaged over a nine-pixel (~10 km<sup>2</sup>) area surrounding the sampling locations. The satellite data were primarily from SeaWiFS and though the processing details are not described, Shuchman et al. (2006) report that the satellite data involved application of an early form of atmospheric correction that commonly resulted in non-physical negative radiance values in the blue portion of the spectrum or other distortions. We note that in presenting the exact same data set and comparisons Pozdnyakov et al. (2005b) state that the data were processed with the "MUMM-MSL12 software" presumably referring to the more sophisticated MUMM (Ruddick et al., 2000) atmospheric correction method developed for use in turbid and inland waters that was imbedded into NASA's SeaDAS software (Baith et al., 2001).

Retrievals of chlorophyll-*a*, NAP, and CDOM in Shuchman et al. (2006) were obtained using the optical cross sections determined for Lake Ontario by Bukata et al. (1991a). Although no statistics were presented, the comparison between the retrieved and sampled chlorophyll concentrations was described as "good." The term "good consistency" also was used when the authors compared retrieved values of chlorophyll, NAP, and CDOM to historical data collected as part of the EEGLE program during the springs of 1998-2000. Although the authors point to several uncertainties associated with their estimates, including use of an imperfect optical model, possible problems with atmospheric correction, disparity between the point measurements made in the field and the pixel-averaged estimates obtained from the satellite, and temporal mismatch between the satellite data and field samples, they conclude that the "results should be considered satisfactory for the analysis of patterns and seasonal distributions and interactions although actual concentrations may be in error." Indeed, the algorithm underestimated the EEGLE chlorophyll concentrations by at least an order of magnitude (see their Figure 10). Like Bukata et al. (1985), the retrievals of NAP and CDOM were more in line with commonly accepted values.

A different multi-component approach is illustrated by the work of Li et al. (2004) in their study of Lake Superior. These researchers used a model by Carder et al. (1999) that is based on the approximate relationship between remote sensing reflectance ( $R_{rs}$ ) and adsorption and backscattering,  $R_{rs}(\lambda) \cong b_b(\lambda)/a(\lambda)$ . In this relationship  $\lambda$  is wavelength,  $b_b$  is the total backscattering resulting from the combination of backscatter from the water  $(b_{bw})$  and from suspended particles  $(b_{bp})$ . Total absorption is calculated as  $a(\lambda) = a_w(\lambda) + a_\varphi(\lambda) + a_d(\lambda) + a_g(\lambda)$ , in which the subscripts  $w, \varphi, d$ , and g refer to water, phytoplankton, detritus (NAP), and gelbstoff (CDOM). In this model the absorption and backscattering functions are determined empirically and a rather complicated calculation yields chlorophyll concentration.

Applied to Lake Superior using a default set of parameters, this semi- analytical approach resulted in predictions of chlorophyll concentration that Li et al. (2004) characterize as having "better agreement" with the field results than the empirical algorithms tested (see section below), though examination of their Figure 10 does not offer strong support for this statement. The authors used their data to optimize the model parameters for Lake Superior data in an attempt to improve the relationship between the predicted and observed measurements. Although the optimization reduced the variability of the predicted values, the new relationship was no more successful at reproducing the field data than was the unoptimized model. The authors conclude that more work needs to be done to develop an algorithm to retrieve chlorophyll concentrations more accurately in Lake Superior.

Estimating the optical cross- sections (the spectrally dependent absorption and scattering functions) of the various CPAs in the water can be quite involved. Until very recently the only published study was that of Bukata et al. (1985). Work on development of a multiyear, multi-lake database of optical properties reportedly is currently underway (personal communication from an anonymous reviewer) and in the past few years a number of new optical characterization studies (Lohrenz et al., 2004; Binding et al., 2008; Effler et al., 2010; O'Donnell et al., 2010; Peng and Effler, 2010) have been presented. Although none of the published studies was explicitly extended to development of new chlorophyll retrieval algorithms, they should provide a more solid basis for further exploration of the method. However, these recent papers also point to complications that may make it even more difficult to apply multi-component algorithms. For example, Binding et al. (2008) showed that the NAP absorption spectrum in Lake Erie depends on the ambient level of CDOM absorption, probably because some of the CDOM is in the form of colloids or adhered to the surface of mineral particles. In addition, the Binding et al. (2008), O'Donnell et al. (2010), and Peng and Effler (2010) work in Lake Erie and the Effler et al. (2010) study in Lake Superior all demonstrate substantial spatial variations in the measured optical cross sections.

#### Empirical algorithms for retrieving values of chlorophyll- a

Because chlorophyll is the dominant color-producing agent in the open ocean (Barale, 1991) most open-ocean retrieval algorithms either ignore the effects of the other CPAs that might be present or assume that they co-vary with chlorophyll. These algorithms have been developed by assuming a relationship between the chlorophyll concentration

and some function of the radiance (or reflectance) values measured by the satellite. In the most general mathematical terms, we can express this as

$$C_{chl} = f(g(R(\lambda))),$$

in which  $C_{chl}$  is the calculated concentration of chlorophyll, f and g are functions, and  $R(\lambda)$  represents the satellite measured radiances in different bands (wavelengths),  $\lambda$ . The function g represents some expression involving radiance values determined at one or more wavelengths. The function f represents some formulaic relationship between the dependent variable chlorophyll concentration and the independent variable radiances passed through the function g.

Algorithms of this sort are termed empirical because the forms of the functions f and g and their coefficients are determined by statistical regression of radiance and chlorophyll measurements obtained from field data rather than based purely on theory. One common feature of the empirical algorithms is that the function  $g(R(\lambda))$  is based on the logarithm of a ratio of two of the sensor bands. Because chlorophyll absorbs strongly in the blue and weakly in the green (see Fig. 1), all other things being equal, the ratio of blue to green radiance received at the satellite should be inversely proportional to the concentration of chlorophyll (Fig. 2).

In preparation for operation of the current generation of satellite ocean color sensors, NASA assembled a large database of *in situ* observations of bio-optical conditions from around the world. This database, which is continually updated, has been used to develop and evaluate a number of empirical retrieval algorithms (O'Reilly et al., 2000). An example subset of these algorithms, used by the first three ocean-color monitoring satellites, is listed in Table 1.

Although algorithms of this sort have been successful when applied to satellite observations of the open ocean, it was not at all clear that they would be equally successful when applied to more optically complex (so-called "Case 2", using the classification of Morel and Prieur (1977)) waters where substances other than chlorophyll might have significant optical effects. Reflectance spectra from these waters would be more complicated (Bukata et al., 1985) and, in theory, the simple band ratio algorithms would be "invalid" (Binding et al., 2008) . Nearshore areas of the Great Lakes that are influenced by riverine inputs of dissolved organics and resuspension of bottom sediments as well as open water areas affected by whitings or with significant concentrations of other suspended particles would likely fall into this category of optical complexity.

Because the Great Lakes rarely are completely free of NAP or CDOM, use of band ratio algorithms in the Great Lakes has been approached with some skepticism. In his classic study using CZCS observations to study southern Lake Michigan, Mortimer (1988) used the spectra presented by Bukata et al. (1985) as the basis for his remark that

The chief challenge facing remote-sensing in coastal waters is the disentanglement of the optical contributions of chlorophyll (Chl), suspended minerals (SM), and dissolved organic carbon (DOC).

More recent papers (Li et al., 2004; Binding et al., 2008; Lohrenz et al., 2008; Binding et al., 2010) also refer to this problem when discussing the difficulty of obtaining chlorophyll concentration values from satellite measurement of the Great Lakes using ratio algorithms.

Bukata et al. (1991b) is the earliest work that attempted to evaluate the ocean data retrieval methods for application to the Great Lakes. Working in Lakes Ladoga (Russia) and Ontario these researchers examined the ability of empirical ocean retrieval algorithms to estimate chlorophyll-*a* concentrations in inland waters. They tested several contemporary (*ca* 1985) algorithms, including those developed by Gordon et al. (1983) for use with the Coastal Zone Color Scanner (CZCS). Bukata et al. (1991b) used the empirical algorithms and *in situ* measurements of upwelling radiance in Lake Ladoga to calculate chlorophyll concentrations and compared these calculations with field measurements of chlorophyll. Although no statistics were presented in the paper, it was apparent from their Figure 1 that the ocean algorithms were poor predictors of the measured values. The failure was attributed to the optical complexity of the largely nearshore waters that were sampled; the authors suggested that similar optical complexity also might be a problem throughout the Great Lakes.

Re-examination of Bukata et al. (1991b) and of the results of more recent work suggests that the failure of the algorithms tested may not have resulted entirely from the optical complexity of the waters, but rather because the functional forms of the algorithms they tested may have been inherently inadequate. Figure 3 shows that the current CZCS algorithm, which is based on a polynomial power function, is much more sensitive over the range of reflectance ratios tested by Bukata et al. (1991b) than are the simple power function algorithms they examined, especially at higher concentrations.

Without more information about the methods used to collect the biological and optical data presented in Bukata et al. (1991b), it is impossible to completely reassess these results. We note, however, that an approximate reconstruction of the relationship between the measurements and the predictions made by using the best algorithm reported in Bukata et al. (1991b) shows a weak, though significant at the 10% level (p = 0.06), correlation ( $C_{pred} = 0.24 * C_{obs} + 2.66, r = 0.36$ ). The correlation ( $C_{pred} = 0.39 * C_{sat} + 3.72, r = 0.37$ ) is slightly better (p = 0.05), though the intercept is larger, if we use the current CZCS algorithm (O'Reilly et al., 2000) with the approximate reflection ratios obtained

by inverting the Bukata et al. (1991b) results (Fig. 4). The intercept using the current CZCS algorithm may be biased high because the older algorithm examined by Bukata et al. (1991b) was based on the  $R_{520}/R_{550}$  ratio. The current algorithm uses either this ratio or the  $R_{443}/R_{550}$  ratio depending on the relative magnitudes of  $R_{443}$  and  $R_{520}$ . Because reflection in the blue wavelengths is higher at lower chlorophyll concentrations, the ratio  $R_{443}/R_{550}$  would be higher than the  $R_{520}/R_{550}$  ratio and the predicted concentrations lower. Of course, it also is possible that enhanced absorption in the blue region of the spectrum due to the presence of interfering substances (*e.g.* CDOM) would affect the ratio as well.

#### Use of satellite ocean color observations to study the Great Lakes

The standard ocean color products distributed by NASA include estimates of chlorophylla concentrations are obtained from band ratio algorithms. Since SeaWiFS began producing usable data late in 1997 several studies have used the SeaWiFS standard chlorophyll products derived from one of the empirical algorithms shown in Table 1 to study processes in the Great Lakes. Other studies, which are discussed below, have focused on evaluation and, in some cases, local or regional modification of these standard SeaWiFS algorithms.

#### Studies that used standard products to examine processes

Lesht et al. (2002) was the first work to use SeaWiFS chlorophyll observations to examine biological processes in the Great Lakes. Chlorophyll concentrations were estimated by using the OC2v2 algorithm (O'Reilly et al., 2000), which was the NASA standard at the time, and the paper documented an early summer phytoplankton bloom in southern Lake Michigan that occurred in response to a transient wind event. The paper included a qualitative comparison of the retrieved chlorophyll values with a limited set of field measurements, but did not attempt any statistical analysis or detailed verification of the satellite observations with *in situ* values. The authors noted that the magnitude of the satellite estimates both before and during the phytoplankton bloom were consistent with historical measurements of chlorophyll concentrations in Lake Michigan and their calculations of primary production, which were based on the satellite derived chlorophyll values, compared well with published *in situ* values.

In a modeling study comparing conditions in Lake Michigan in 1998 and 1999 Chen et al. (2004) used a few SeaWiFS chlorophyll images along with observations made during the EEGLE program. Chen et al. (2004) concluded that their modeled spatial patterns of chlorophyll distribution agreed well, in a qualitative way, with the SeaWiFS images which were based on the OC2v2 algorithm. They did not, however, attempt to make a quantitative comparison based on the modeled and observed concentration values.

More recently, Kerfoot et al. (2008) also used EEGLE field observations and a few selected SeaWiFS images to examine the springtime spatial pattern of chlorophyll concentration in southern Lake Michigan. The authors used a coastal atmospheric correction method (Stumpf et al., 2003) intended for use in Case 2 waters, and generated chlorophyll images by using the SeaWiFS OC2v2, OC2v4 and OC4v4 algorithms. The satellite estimates were compared with EEGLE program field measurements of chlorophyll "taken along transects at 5-10 m depth from 22 Mar 98 - 11 Sep 99." The association between the field measurements ( $C_a$ ) and the satellite estimates ( $C_{sat}$ ) was excellent  $(C_a = -0.087[\pm 0.171] + 1.079[\pm 0.102]C_{sat}, r^2 = 0.874, N = 18)$ . The paper, however, was not clear as to whether the  $C_{sat}$  values used in the regression came from OC2v4 or OC4v4 algorithms nor was was there comment about the potential interference that might result from the presence of high concentrations of either CDOM or NAP though at least some of the samples were collected during the March 1998 EEGLE "plume" event (Lohrenz et al., 2008). A subsequent paper (Kerfoot et al., 2010) used similar satellite data to examine long term changes in the occurrence of a late winter phytoplankton bloom in Lake Michigan. A note in this paper suggests that the regression described above (Kerfoot et al., 2008) was based on OC2v4 retrievals. Kerfoot et al. (2010) collected additional samples in the southern basin of Lake Michigan from 21-23 April 2008 and again found excellent correspondence between the satellite derived chlorophyll and the field observations, though the slope differed significantly from 1 ( $C_a = -0.022[\pm 0.082] + 1.209[\pm 0.058]C_{sat}$ ,  $r^2 =$ 0.900, N = 51).

Lohrenz et al. (2008) used SeaWiFS OC2 retrievals from Lake Michigan to produce chlorophyll values as input to a primary production model. No other information about processing of the satellite data was provided in this paper. They found that the satellite derived chlorophyll values were "generally comparable in magnitude to determinations of chlorophyll in discrete samples." However, the authors report that some of the satellite estimates from inshore stations were "anomalously high" relative to the field values and attribute this overestimation to the presence of high concentrations of CDOM originating from the St. Joseph River. In addition to overestimation of chlorophyll values near the river outflow, the authors found that in 1999 and 2000 the OC2 algorithm produced chlorophyll values that were lower than field values collected in offshore waters, an effect they attributed to relatively high abundances of cryptophytes observed in the phytoplankton population during 1999. Because cryptophytes absorb light preferentially in the green region of the spectrum they would tend to distort the chlorophyll estimates based on the standard band ratio algorithm. Examination of Lohrenz et al. (2008) Figure 5, however, suggests that with the exception of the nearshore (within  $\sim 5 \text{ km}$  from shore) values from 1999, the retrieved chlorophyll values are quite reasonable. Using the data plotted in this figure to approximate the root-mean-squared error between retrieved and

sampled chlorophyll values yields RMSE ~0.5 mg/m<sup>3</sup> (about 18% of the mean observed chlorophyll value) overall for each of the two different analytical methods (high-pressure liquid chromatography and fluorometric assay) they used to determine the field chlorophyll concentrations. The average RMSE between paired chlorophyll values determined by the two analytical methods was ~0.37 mg/m<sup>3</sup>. Lohrenz et al. (2008) conclude (their p. 67) by stating that their findings are "consistent with those of other studies that have found satellite estimates of chlorophyll to correctly represent spatial and temporal trends in coastal waters despite bias in absolute accuracy."

Finally, Barbiero et al. (2011) used SeaWiFS OC4v4 chlorophyll observations made from 1998 though 2006 to demonstrate that the magnitude of the spring bloom in Lake Huron has declined dramatically since 2003. The decline in chlorophyll concentration suggested by the satellite data corresponded to field observations of declines in phytoplankton biovolume and to increases in spring dissolved silica and secchi depth that also began in 2003 (Barbiero et al., 2009). The satellite data were processed from Level 1A to Level 2 with NASA's SeaDAS software using an advanced iterative atmospheric correction algorithm ((Bailey et al., 2010). Pixels were mapped to a standard grid and estimated chlorophyll values extracted from locations corresponding to the U.S. EPA's Great Lakes National Program Office (GLNPO) open water monitoring stations. The field values were obtained by using standard analytical methods. The relationship between 2002-2006 satellite and laboratory chlorophyll values was linear ( $C_{sat} = 0.0145 + 1.000C_{lab}$ ) and statistically indistinguishable from a 1:1 relationship (F = 70.76; p < 0.001;  $r^2 = 0.50$ .)

#### Studies that use in situ biological observations and satellite data to evaluate algorithms

Rather than use satellite data to examine limnological processes, several studies have focused on evaluation of the retrieval algorithms. We first consider those that use chlorophyll values derived from the radiances measured at the satellite along with the field observations of chlorophyll. These studies used satellite and field data that were approximately matched in space and time. Because some of the satellite imagery may be obscured by cloud cover and because, unlike the satellite measurements which are essentially averages over  $\sim 1 \text{ km}^2$  areas, the field samples are point measurements, some variability due to temporal and spatial mismatching is to be expected. Furthermore, because the satellite retrievals also depend on the method used to remove the effects of the atmosphere from the radiances measured at the satellite this type of study is subject to an additional degree of uncertainty. However, despite these limitations, these studies are most like those used to explore processes and shed considerable light on the applicability of the algorithms for limnological studies. In particular, they demonstrate how satellite chlorophyll retrievals might be expected to correspond to field observations (so-called "ground truth"). We describe those studies that are based solely on *in situ* observations of chlorophyll and radiance, thus eliminating the temporal, spatial, and atmospheric uncertainties, in the following section.

Working in Lake Superior during the springs of 1998-1999, Budd and Warrington (2004) compared field values of chlorophyll obtained from several coastal transects off the north shore of the Keweenaw peninsula with satellite estimates made using the OC2v2 algorithm. The satellite data were pre-processed to remove negative radiance values. Negative radiance values typically occur because the atmospheric correction method overestimates the contribution of atmospheric aerosol scattering to the satellite radiances, especially at short wavelengths (Stumpf et al., 2003). The imagery was screened to accept only those images that matched the date of field sampling and that had fewer than 10% pixels covered by cloud. The satellite values used in the comparisons were averages of the retrieved values within 3 by 3 pixel boxes surrounding the sampling locations. Field chlorophyll values ranged from 0.1 mg/m<sup>3</sup> to 2.5 mg/m<sup>3</sup> with all but two of the field observations below 0.5 mg/m<sup>3</sup>; these two large values were considered outliers and eliminated from the analysis. Although the slope of the relationship between the satellite and field values deviated considerably from unity, the association between the remaining retrieved ( $C_{OC2}$ ) and measured (*Chl*) values was excellent ( $C_{OC2} = 3.1(Chl) + 0.4, r^2 = 0.87, N = 18$ ). The authors concluded that the tendency of OC2v2 to overestimate the field measurements resulted from high concentrations of CDOM in the water. Because the relationship between the retrieved and observed values was linear, they suggest multiplying the OC2v2 results by a constant factor of 0.3 to obtain more accurate estimates.

Witter et al. (2009) compared field measurements collected in Lake Erie to SeaWiFS retrievals using twelve empirical algorithms. The field data used in this work were taken from the spring and summer GLNPO monitoring surveys done between 1998 and 2002 and represented all three Lake Erie basins. The field data were screened to eliminate high (> 20 mg/m<sup>3</sup>) values of chlorophyll concentration and samples from within 2 km of land. The satellite data were processed using the same atmospheric correction method used by Budd and Warrington (2004) and screened with very similar criteria. Witter et al. (2009), however, accepted images if fewer than 20% of their pixels were covered by cloud rather then the 10% cloud limit required by Budd and Warrington (2004). Comparisons were made with the field data only if all pixels within 5 pixels of the sampling locations were cloud-free.

In general, the correlations between the retrieved and measured values was good (0.61 < r < 0.73). However, for nine of the twelve algorithms the slope of the correlation was less than 0.7 and the intercept was greater than 2.0 mg/m<sup>3</sup>. In other words, most of the algorithms seemed to be biased high at low chlorophyll concentrations and biased low at higher concentrations. When stratified by basin, the agreement between the retrieved and measured values improved from western basin to eastern basin, suggesting that interfer-

ence from suspended sediments in the western and central basins negatively affected the retrievals.

Witter et al. (2009) went on to use their data to develop a "regional algorithm" for Lake Erie. In essence, they used regression to optimize a different form of empirical algorithm specific to their observations. Specifically, Witter et al. (2009) found that the expression  $C = 10^{a+bR+cR^2}$ , where  $R = \log(Rrs_{490}/Rrs_{555})$  and *a*, *b*, and *c* are a set of coefficients specific to the western, central, and eastern basins, provided a statistically improved (relative to the standard NASA algorithms they tested) relationship between the calculated and observed chlorophyll values in Lake Erie.

Because the regional algorithm described above still is based on a reflectance ratio albeit with different coefficients, it still should be subject to the same types of interferences that are expected to affect the standard algorithms though it is possible that the coefficient changes also changed the sensitivity of the algorithm to the presence of NAP and CDOM. The authors state that improvement in fit between the measured values and those obtained by using the optimized retrieval was small in the western basin and better in the central and eastern basins, though the retrieved values still tended to underestimate the field values. An important conclusion of this work is that "an algorithm with a power-law form may be effective" in a region of the lake where interference from suspended sediment is minimal.

Although he used MODIS data rather than SeaWiFS data in his work, Watkins (2009) evaluated the OC4v4 algorithm using matched satellite and field data collected during April, August, and September 2003 in Lake Ontario. The satellite data were processed by using the standard NASA procedures that at the time included an iterative near-infrared atmospheric correction method. Like Witter et al. (2009), the Lake Ontario results showed good association between the satellite retrievals and field data ( $r^2 = 0.62$ ) at open water (> 30 m) stations, though the satellite estimates underestimated the field results ( $C_{sat} = 0.42(Chl_a) + 1.01$ ). The algorithm tended to overestimate the chlorophyll measured at nearshore stations, especially in April when, the author notes, "suspended sediment and dissolved organic matter are prevalent."

#### Studies that use in situ optical and biological measurements to evaluate algorithms

Only two studies (Bergmann et al., 2004; Li et al., 2004) report using *in situ* radiometric observations along with biological sampling to evaluate the standard empirical chlorophyll algorithms in the Great Lakes. In both cases, surface radiometers, profiling spectral radiometers, and profiling absorption and attenuation meters were used to measure the optical properties of the water column at each sampling station. These optical observations provided direct measurements of the remote sensing reflectance values that were input to the retrieval algorithms, thus avoiding the uncertainty associated with atmospheric correction models and the spatial and temporal mismatch between satellite and biological

sampling.

Bergmann et al. (2004) did their work in southern Lake Michigan as part of the EE-GLE program. Although the main aim of their research was to examine how resuspended sediments alter the optical conditions in the water, they also tested the predictive ability of seven empirical chlorophyll retrieval algorithms and reported the results in a single paragraph at the end of their paper. These authors state that "there was a strong correlation between measured and calculated chlorophyll concentration at most stations." However, they also found that the chlorophyll retrievals underestimated the field data by about 45% at stations where the phytoplankton community was dominated by cryptophytes. They attributed this underprediction to the high absorption of green wavelengths by the cryptophytes. Interestingly, it was the offshore water that was dominated by cryptophytes, so in contrast to the other studies discussed below, the algorithms performed better in the nearshore than offshore. At those stations at which fewer than 40% of the phytoplankton were cryptophytes, the slope and  $r^2$  values for regressions relating the field data to chlorophyll values calculated by using SeaWiFS OC2v2, OC4v4, and MODIS/OC3M were 1.04/0.95, 0.73/0.83, and 0.70/0.89 respectively. These results did not seem to be affected by the presence of high concentrations of suspended sediment.

Finally, using similar field methods, Li et al. (2004) evaluated nine empirical algorithms and one "semi-analytical" algorithm using data collected in Lake Superior. We only discuss the empirical algorithms here; the semi-analytical algorithm was discussed above in the section on multicomponent models.

Chlorophyll concentrations in Lake Superior generally are very low. The range of values reported by Li et al. (2004) was ~0.5 to ~1.2 mg/m<sup>3</sup>. In contrast to the results reported by Budd and Warrington (2004) for the same waters, none of the calculated chlorophyll values using several algorithms was correlated with the field measurements. Li et al. (2004) attempted to find an optimized regional algorithm by manipulating the coefficients of the OC4v4 algorithm, but the optimized algorithm did not improve the results. In all cases the empirical algorithms overestimated the field measurements. The authors attribute this to the low ratio of chlorophyll- *a* to dissolved organic material in the coastal Lake Superior waters they sampled. Budd and Warrington (2004) also speculated that CDOM would negatively affect the retrievals though they did not present any CDOM measurements.

#### Discussion

Like many inland and coastal waters, the Great Lakes usually are thought of as optically complex (Case 2) and many researchers (as noted above) have assumed that the band ratio algorithms will not be at all useful for study of these waters. Unfortunately, the few published studies that evaluate multi-component models developed for Case 2 waters showed that they were not successful predictors of chlorophyll concentration. Because multi-component models would include the effect of all the significant CPAs, they could, in theory, accommodate spatial and temporal variations in the concentrations of the substances that interfere with the band ratio algorithms. To accomplish this, however, the optical cross sections of the CPAs would have to be determined and furthermore, some assessment would have to be made as to the spatial and temporal variability of those cross sections. As we noted above, work intended to develop a database of optical cross sections for the Great Lakes is currently underway. None of this work that might be directed toward development of retrieval algorithms has been published as of this writing. We hope that when it is, it will be fully accessible to the broad community and be sufficiently general so that it can be widely implemented. Until that time, researchers interested in using satellite imagery to study phytoplankton processes the Great Lakes will have to rely on the commonly available and easily implemented band ratio methods.

Table 2 shows that with one exception (Li et al., 2004), all the studies that used the band ratio algorithms and collected matched field data showed a good correspondence between retrieved and observed values, although only a few demonstrated a 1:1 relationship. Clearly, because these published studies represent the work of different investigators who used different field and satellite data processing methods and who studied different areas of different lakes at different times in different years it is not surprising that the results differ in detail. What is surprising, in our view, is the consistent significant linearity of the relationships. This linearity indicates to us that though they are not ideal, the band ratio algorithms can provide estimates that are reflective of the true chlorophyll concentrations in the Great Lakes. The degree to which these estimates are useful will depend, of course, on the purpose for which they are intended.

Given that the waters of the Great Lakes are never completely free of interfering CPAs and that the phytoplankton assemblages will vary between and within lakes as functions of space and time (Fahnenstiel and Scavia, 1987), what then can account for the consistent linearity between predicted and observed chlorophyll values found when the band ratio algorithms have been used in Great Lakes research? Because the standard empirical algorithms are based on analysis of oceanic measurements, it certainly is reasonable to assume that they would not be appropriate for the Great Lakes where the phytoplankton populations differ from oceanic populations and also vary considerably within and between lakes and from season to season. The question is how much does this variation affect the retrievals? We might further ask if such variations can account for differences in the results demonstrated in the published work? We attempted to address these questions by using the Bukata et al. (1985) multicomponent optical model (Eqs 1-3 above) in a diagnostic mode.

As noted above, the empirical algorithms are based on the assumption that chlorophyll is the dominant CPA in the water and, if other CPAs are present, they co-vary with chlorophyll concentration. Considering the situation when the absorption and scattering processes result only from the presence of chlorophyll and water, then the empirical algorithms must parameterize these processes in their functional forms and parameter values. Assuming the absorption and scattering properties of pure water are universal, the specific forms of the empirical algorithms must represent some global average of the scattering and absorption properties of phytoplankton and co-variates.

We tested the question of how the retrieved chlorophyll values would depend on the specified absorption spectrum by using the Bukata et al. (1985) multicomponent optical model along with the two field measured phytoplankton absorption spectra determined by Bukata et al. (1985) and the one determined by Lohrenz et al. (2004). We used the model to produce theoretical reflectance spectra that might be expected in Case 1 (phytoplankton dominated) waters for several different values of chlorophyll concentration. We then sampled these spectra at the wavelengths used by the current NASA standard empirical algorithm for SeaWiFS (OC4v6) and compared the retrieved chlorophyll values to those that were used to generate the spectra (Fig. 5).

The results shown in Fig. 5 show that, assuming the Bukata et al. (1985) model and phytoplankton backscattering cross-section are appropriate, that the retrievals indeed depend on the form of the phytoplankton absorption spectrum. Despite the differences, the results are linear, though the slopes vary between 0.5 and 1.5. The highest slope corresponds to the case where the assumed phytoplankton absorption spectrum (Bukata et al. (1985) spectrum C) does not show a distinct absorption peak in the blue. That spectrum resembles the phytoplankton absorption spectra presented by Binding et al. (2008) in which CDOM (highly absorbent at short wavelengths) may have been adsorbed on the plankton particulates or otherwise retained on the filter being analyzed. The other two spectra are more typical with distinct absorption peaks in both the blue and red regions.

To see if the possible presence of CDOM and NAP could account for the diverse results reported by the studies listed in Table 2, we again used the model to simulate OC4v6 chlorophyll retrievals for cases in which both CDOM and NAP were allowed to vary over a range of values. For this exercise we used the Lohrenz et al. (2004) phytoplankton absorption spectrum along with the spectral model of NAP absorption from Binding et al. (2008) and the backscattering functions provided by Bukata et al. (1985). Because the model components have been drawn from studies that were done at different times and in different lakes, we do not expect the results to be exact or definitive, but rather to help illustrate how, in theory, the presence of these interfering substances might confound the band ratio retrieval algorithms. The results are plotted in Fig. 6. We note that Bukata et al. (1985) parameterized CDOM absorption in terms of the more commonly measured DOC

concentration which we used in the model. Much of DOC is not colored, and Binding et al. (2008) found that absorption due to CDOM could range from 0.08 m<sup>-1</sup> to 0.75 m<sup>-1</sup> in Lake Erie. When normalized by the observed concentration of DOC, Binding et al. (2008) found the average specific absorption coefficient at 440 nm ( $\hat{a}_{CDOM}(440)$ ) in Eq. 2 above) was 0.102 m<sup>2</sup> g<sup>-1</sup> which agreed well with the value 0.115 m<sup>2</sup> g<sup>-1</sup> determined in Lake Ontario by Bukata et al. (1985).

In the absence of CDOM (Fig. 6a), adding any suspended particles (indicated by the TSM concentrations shown in the figure legends) reduces the slope and increases the intercept of the relationship between retrieved and modeled chlorophyll values thus producing predictions that are lower than model input values at higher concentrations. We should repeat here that we view these model results as qualitative rather than quantitative. Observed TSM concentrations in more turbid regions of the Great Lakes (e.g. the western Lake Erie or nearshore areas subject to sediment resuspension) can reach 20 g/m<sup>3</sup> shown as the highest TSM concentration in Fig. 6. At this TSM level, using the backscatter and absorption spectra noted above, we found that the retrieved chlorophyll values became nearly independent of the model input chlorophyll values. The modeled effect of adding particles primarily depends on the scattering properties of the particles. The particle scattering cross sections presented by Bukata et al. (1985) range from 0.052 to 0.034  $m^2g^{-1}$ with the higher values occurring at shorter wavelengths. Unfortunately, very few detailed studies of particle scattering have been done in the Great Lakes and we know of no other published backscattering cross section for Great Lakes particulates. In the absence of particulates, increasing the concentration of CDOM (Figs. 6b, 6c, and 6d) changes the relationship between retrieved and model input values from linear to logarithmic, with very rapid increases in the predicted values at low concentrations. Higher concentrations of CDOM also tend to increase the sensitivity of the retrieval to changes in particulate concentrations and result in reduced slopes and higher intercepts.

The results of the studies listed in Table 2 generally are consistent with these qualitative model results. The tendency of the retrievals from Lakes Erie and Ontario (Witter et al., 2009; Watkins, 2009) to underestimate the field values suggests the presence of NAP especially at higher chlorophyll concentrations. The Lake Ontario results (Watkins, 2009) particularly were poor in the nearshore where NAP and CDOM were prevalent. When chlorophyll and suspended particulate concentrations are low but CDOM is present, such as in the coastal Lake Superior case studied by Budd and Warrington (2004), the slope of the modeled regression will be greater than 1 near the origin, but will flatten out as chlorophyll concentration increases, more or less as was observed (*e.g.* the "outlier" values identified by Budd and Warrington (2004)). The Lake Huron results in Barbiero et al. (2011) were based on samples collected over a number of years in the open lake where both CDOM and NAP concentrations are expected to be low and conditions most like "Case 1" waters. In the coastal areas of southern Lake Michigan sampled by Kerfoot et al. (2008) and Kerfoot et al. (2010) the authors comment on minimal effects of CDOM (Kerfoot et al., 2010) and on the insensitivity of the relationship between retrieved and observed chlorophyll concentration on the presence of suspended particulates (Bergmann et al., 2004). In Lake Michigan, however, the composition of the phytoplankton community was found to affect the relationship because of differences in the absorption spectrum (Bergmann et al., 2004).

The reasons satellite data are so attractive for limnological studies is that, relative to traditional ship-based sampling, they are (a) synoptic and (b) available year-around with near daily frequency. No other sampling method affords the possibility for simultaneous measurement of properties over such large areas or for conducting studies based on spatially resolved time series. Given that the optical properties of the lakes (*i.e.* the concentrations of CPAs and composition of the phytoplankton communities) will no doubt vary in both space and time, it certainly is reasonable to question the value of satellite retrievals based on any single empirical algorithm. Because the optical cross sections on which the multi-component algorithms depend also will vary in space and time, making simultaneous retrievals of all the CPAs similarly questionable unless some scheme for dynamic (in both space and time) adjustment of the optical cross sections is developed.

Several authors (Lohrenz et al., 2004; Shuchman et al., 2006; Lohrenz et al., 2008) comment on the utility of satellite imagery for estimating spatial and temporal trends regardless of the absolute accuracy of the algorithms even though this would strictly require that the algorithms be broadly applicable over space and time. To determine how unpredictable variations in NAP and CDOM might affect chlorophyll retrievals made by using the empirical algorithms we conducted a Monte-Carlo experiment in which we randomly and independently selected the concentrations of the interfering CPAs input to the Bukata et al. (1985) model from log-normal distributions assumed typical of the distributions in the Great Lakes. For each realization of the model we also randomly selected the phytoplankton absorption spectrum from those determined by Bukata et al. (1985) and Lohrenz et al. (2004), the NAP absorption spectrum from those presented by Bukata et al. (1985) and that derived by Binding et al. (2008), and the CDOM absorption from data published by Bukata et al. (1985), Binding et al. (2008), and Effler et al. (2010). The model was run 10,000 times and a randomly selected subset of 250 of the 10,000 realizations was used to generate the results shown in Fig. 7.

The relationship between the retrieved and modeled chlorophyll shown in Fig. 7 is representative of the Monte-Carlo results. The best fit relationship is linear with a slope less than one and a positive intercept, similar to those presented in most of the papers we have reviewed above (Bergmann et al., 2004; Budd and Warrington, 2004; Kerfoot et al., 2008; Witter et al., 2009; Watkins, 2009; Kerfoot et al., 2010; Barbiero et al., 2011). The

increased scatter at higher values of chlorophyll shown in Fig. 7 results from the three different chlorophyll absorption cross sections input to the model (see Fig. 5). This variability emphasizes the importance of using accurate optical cross sections when applying a multi-component optical model such as the one on which this simulation was based.

Each of the field studies cited above involved collection of data from different locations at different times and in some cases over areas as large as an entire lake and over time spans of several years. If we assume that the samples collected as part of these studies randomly varied in concentrations of interfering substances and optical properties, then our Monte-Carlo study simulates the field data. Unfortunately, none of the field studies report simultaneous measurement CDOM or NAP with the chlorophyll values. Though the Monte-Carlo model is only qualitative, the linear relationship between the retrieved and modeled chlorophyll values found in Fig. 7 and the linear relationships between observed chlorophyll in the papers reviewed suggests that in a statistical sense, the empirical algorithms can reflect the chlorophyll concentrations in the Great Lakes despite local variations in the concentrations and composition of the CPAs.

One additional aspect of the studies we reviewed should be mentioned in the context of assessing the uncertainty associated with application of the band ratio methods. Many of the field measurements reported in the work we reviewed were collected in the spring. The EEGLE sampling in Lake Michigan (Bergmann et al., 2004; Lohrenz et al., 2004, 2008) took place during March and April. Kerfoot et al. (2010) used Lake Michigan data collected in April and Barbiero et al. (2011) used GLNPO data from spring surveys (March and early April) in their study of Lake Huron. Although several of the other studies we cite used data collected throughout the year, only the Budd and Warrington (2004) study in Lake Superior formally tested the dependence of the retrieval on season; they found no seasonal effect. Because the number of samples that is collected in any single survey is limited, the tendency has been to pool all the available data. Without more specific information about timing of the field sampling, we cannot evaluate this source of variability.

By design, we have restricted our review to the published literature that deals explicitly with retrieval of chlorophyll concentrations from satellite observations of the North American Great Lakes. These papers are dominated by studies using data from SeaWiFS or MODIS and band ratio algorithms. We would be remiss, however, if we did not also note that novel chlorophyll retrieval methods based on different spectral bands and on other sensors have been applied to other optically complex waters (coastal, estuarine, and inland) where band ratio algorithms have been found lacking. Although not within the scope of our review, because these methods offer possible alternatives that could be exploited in studies of the Great Lakes, we describe them briefly here.

In March, 2002 the European Space Agency (ESA) launched a satellite (ENVISAT)

that carried a variety of instruments intended for environmental monitoring. Among these instruments was the Medium Resolution Imaging Spectrometer (MERIS) which was designed primarily for ocean color applications. In contrast to the fixed spectral bands provided by SeaWiFS and MODIS, MERIS samples 15 spectral bands between 390 nm and 1040 nm, the exact location and the width of which are programmable and can be modified in orbit by ground control. For ocean color applications, the fifteen bands are distributed between 412.5 nm and 900 nm. Because Sun-stimulated chlorophyll fluoresces near 685 nm, three MERIS bands (665 nm, 681.25 nm, and 708.75 nm) are concentrated in this region of the spectrum. Gower et al. (1999) describe how these bands can be used in an algorithm based on the Fluorescence Line Height (FLH) to estimate chlorophyll concentrations in regions where retrievals based on band ratio algorithms are complicated by the presence of NAP and CDOM. Other researchers have explored using FLH based on these MERIS band as well as the MODIS band at 678 nm but found erratic results when a substantial amount of NAP was present.

At least three other algorithms with potential for improving chlorophyll retrievals in optically complex waters have been the subject of recent study. The first, known as the maximum chlorophyll index (MCI), is based on the amplitude of the MERIS 708.75 nm band. Gower et al. (2008) discuss the algorithm as well as its application for detection of blooms and other transient events. This algorithm seems especially well suited to blooms of cyanobacteria (Alikas et al., 2010; Binding et al., 2011). Dall'Olmo et al. (2005) explored the use of SeaWiFS and MODIS red and near-infrared (NIR) bands for retrieving chlorophyll concentration from observations of turbid water. These algorithms were used, along with a similar three-band algorithm based on MERIS observations by Gitelson et al. (2007) in a study of chlorophyll in Chesapeake Bay. In this highly productive system, Gitelson et al. (2007) found that the algorithms could account for approximately 80% of the variability in chlorophyll observations with a relative error of about 20%. Finally, a completely different approach that uses five bands has been proposed by Gohin et al. (2002). The basis of this algorithm is a look-up table developed from field observations in which chlorophyll concentration is related to a data triplet consisting of the standard OC4 band ratio and the radiances at 412 nm and 555 nm. In a comparison of algorithm performance conducted in the Bay of Bengal and the Arabian Sea, Tilstone et al. (2011) found this five-band algorithm was the most accurate for both coastal and open ocean areas.

Only one paper (Gons et al., 2008) has assessed the use MERIS data or these algorithms for research in the Great Lakes. Gons et al. (2008) chose two diverse areas, indicative of the extremes of trophic conditions in the lakes, to examine the performance the FLH and red-to-NIR algorithms in oligotrophic (Keweenaw Bay, Lake Superior) and eutrophic (Green Bay, Lake Michigan) waters. They found that the red-to-NIR algorithm was applicable to Green Bay, but not to Keweenaw Bay where the method resulted in negative values. Although they found a strong linear relationship between observed chlorophyllla concentrations and those predicted by the FLH algorithm in Keweenaw Bay, when the algorithm was applied to an entire scene, results from closely located pixels exhibited considerable noise, suggesting that conditions in Lake Superior may be very close to or below the MERIS detection limit.

#### Conclusion

Modeling the performance of the empirical algorithms for different combinations of interfering substances and properties shows that the empirical band ratio retrieval algorithms can produce estimates of chlorophyll concentrations that are proportional to model input values. This finding is consistent with numerous studies reporting significant linear relationships between chlorophyll determined from satellite observations using band ratio algorithms and that measured *in situ*. Furthermore, the published evidence suggests that even if the absolute accuracy of the empirical retrievals may be questioned, the spatial and temporal patterns in chlorophyll concentrations derived from satellite data correspond to what is known about the lakes from other types of studies (Lesht et al., 2002; Chen et al., 2004; Shuchman et al., 2006; Kerfoot et al., 2008; Lohrenz et al., 2008; Watkins, 2009; Barbiero et al., 2011). Because they are subject to interference from other CPAs, the empirical algorithms should provide better estimates of chlorophyll concentration in regions where the influence of CDOM and NAP are likely to be minimal (e.g. the open waters of the lakes). In nearshore areas, in regions where the lakes are affected by significant riverine inputs, and in waters subject to frequent resuspension of bottom sediments (e.g. western Lake Erie and other shallow areas), the retrievals may be seriously distorted by the presence of these interfering substances and should be used with great care.

Ideally studies based on satellite observations would include some field measurements for confirmation which would include simultaneous determinations of the concentrations of the other CPAs present in the waters. Although the empirical band ratio methods currently are the most accessible, recent application of other algorithms and data from new sensors have shown promise in optically complex waters. More studies assessing algorithms such as FLH, MCI, and red-NIR modeling are needed to determine their applicability to the Great Lakes. Additional studies of the optical properties of the Great Lakes also are needed to better understand the regional and temporal variations of these properties. Because of the sensitivity of multi-component models to variations in the input optical cross sections, accurate determination of these cross sections are particularly important. We are hopeful, of course, that the latest work (Binding et al., 2010; Effler et al., 2010; O'Donnell et al., 2010; Peng and Effler, 2010) and other ongoing, not yet published, studies will provide new insight and extend the application of satellite remote sensing to 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 thank two anonymous reviewers for their helpful comments.

- Alikas, K., Kangro, K., Reinart, A., 2010. Detecting cyanobacterial blooms in a large North European lakes using the Maximum Chlorophyll Index. Oceanologia 52, 237– 257.
- 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.
- Barale, V., 1991. Sea-surface color in the field of biological oceanography. International Journal of Remote Sensing 12, 781–793.
- Barbiero, R.P., Balcer, M., Warren, G., Lesht, B.M., 2009. Recent changes in the lower food web of Lake Huron. 52nd Conference on Great Lakes Research, International Association for Great Lakes Research.
- 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., Jerome, J.H., Bukata, R.P., Booty, W.G., Letourneau, G., 2011. An assessment of MERIS algal products during an intense bloom in Lake of the Woods. J. Plankton Research 33, 793–806.
- 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.
- 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.
- 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.
- Chang, C.H., Liu, C.C., Wen, C.G., 2007. Integrating semianalytical and genetic algorithms to retrieve the constituents of water bodies from remote sensing of ocean color. Optics Express 15, 252–265.
- 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.

- Dall'Olmo, G., Gitelson, A.A., Rundquist, D.C., Leavitt, B., Barrow, T., Holz, J.C., 2005. Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. Remote Sensing of Environment 96, 176–187.
- Effler, S.W., Perkins, M., Peng, F., Strait, C., Weiderman, A.D., Auer, M.T., 2010. Lightabsorbing components in lake superior. J. Great Lakes Research 356, 656–665.
- Evans, R.H., Gordon, H.R., 1994. Coastal zone color scanner "system calibration": A retrospective examination. J. Geophys. Res. 99, 7293–7307.
- Fahnenstiel, G.L., Scavia, D., 1987. Dynamics of lake michigan phytoplankton: Recent changes in surface and deep communities. Canadian Journal of Fisheries and Aquatic Sciences 44, 509–514.
- Garcia, V., Signorini, S., Garcia, C., McClain, C., 2006. Empirical and semi-analytical chlorophyll algorithms in the southwestern atlantic coastal region (25-40°s and 60-45°w). International Journal of Remote Sensing 27, 1539–1562.
- Gitelson, A.A., Schalles, J.F., Hladik, C.M., 2007. Remote chlorophyll-*a* retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sensing of Environment 109, 464–472.
- 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.
- Gons, H.J., Auer, M.T., Effler, S.W., 2008. MERIS satellite chlorophyll mapping of oligortophic and eutrophic waters in the Laurentian Great Lakes. Remote Sensing of Environment 112, 4098–4106.
- Gordon, H.R., Brown, O.B., Jacobs, M.M., 1975. Computed relationship between the inherent and apparent optical properties of a float homogeneous ocean. Applied Optics 14, 417–427.
- Gordon, H.R., Clark, D.K., Brown, J.W., Brown, O.B., Evans, R.H., Broenkow, W.W., 1983. Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates. Applied Optics 22, 20–36.
- Gower, J., Doerffer, R., Borstad, G.A., 1999. Interpretation of the 685nm peak in waterleaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. International Journal of Remote Sensing 20, 1771–1786.

- Gower, J., King, S., Goncalves, P., 2008. Global monitoring of plankton blooms using MERIS MCI. International Journal of Remote Sensing 29, 6209–6216.
- 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. Limology 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.
- 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.
- 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.
- 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.
- Morel, A., Prieur, L., 1977. Analysis of variation in ocean color. Limnology and Oceanography 22, 709–722.
- Mortimer, C., 1988. Discoveries and testable hypotheses arising from coastal zone color scanner imagery of southern Lake Michigan. Limnology and Oceanography 33, 203–226.
- 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., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., Kahru, M., Chavez, F.P., Strutton, P., Cota, G.F., Hooker, S.B., McClain, C.R.,

Carder, K.L., Müeller-Karger, F., Harding, L., Magnuson, A., Phinney, D., Moore, G.F., Aiken, J., Arrigo, K.R., Letelier, R., Culver, M., 2000. Ocean Color Chlorophyll a Algorithms for SeaWiFS, OC2 and OC4: Version 4. SeaWiFS postlaunch technical report series. Vol 11, Part 3, Ch. 2. NASA-TM-2000-206892. NASA Goddard Space Flight Center. Greenbelt, Maryland.

- Pan, D., Gower, J., Borstad, G., 1988. Seasonal variation of the surface chlorophyll distribution along the british columbia coast as shown by czcs satellite imagery. Limnology and Oceanography 33, 227–244.
- 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., Korosov, A., Grassl, H., Pettersson, L., 2005a. An advanced operational algorithm for operational retrieval of water quality from satellite data in the visible. International Journal of Remote Sensing 26, 2669–2687.
- Pozdnyakov, D., Shuchman, R., Korosov, A., Hatt, C., 2005b. Operational algorithm for the retrieval of water quality in the Great Lakes. Remote Sensing of Environment 9, 352–370.
- Ruddick, K.G., Ovido, F., Rijkeboer, M., 2000. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. Applied Optics 39, 897–912.
- 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 Sea-WiFS: a 7-year inter-annual analysis. J. Great Lakes Research 32, 258–279.
- Singh, S., Cracknell, A., Charlton, J., 1983. Comparison between czcs data from 10 july 1997 and simultaneous *in situ* measurements for southeastern scottish waters. International Journal of Remote Sensing 4, 755–784.
- Stumpf, R.P., Arnone, R., Jr., R.G., Martinolich, P., Ransibrahmanakul, V., 2003. A partially-coupled ocean-atmosphere model for retrieval of water leaving radiance from SeaWiFS in coastal waters, in: Patt, F. (Ed.), Algorithm Updates for the Fourth SeaWiFS Data Reprocessing. NASA Goddard Space Flight Center, Greenbelt, Maryland. volume 22.
- Tanis, F., 1984. Phase 2 development of Great Lakes algorithms for Nimbus-7 Coastal Zone Color Scanner. Final Report 157900-23-F. Mich. Environ. Res. Inst.. Ann Arbor, Michigan.

- 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.
- 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.
- 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.

Table 1: Sampling of algorithms for retrieving chlorophyll-*a* concentration ( $C_{chl}$ ) from satellite observations of remote sensing reflectance. A more complete list of algorithms may be found in O'Reilly et al. (2000). *R* in these functions is  $\log_{10}$  of the indicated band ratio. CZCS refers to the Coastal Zone Color Scanner, SeaWiFS to the Sea-viewing Wide Field Spectrometer, MODIS to the Moderate-resolution Imaging Spectrometer.

Algorithm	Functional Representation	Bands Employed (R)
CZCS/OC3C	$C_{chl} = 10.0^{(0.362 - 4.066R + 5.125R^2 - 2.6458R^3 - 0.597R^4)}$	443 > 520/550
SeaWiFS/OC4v4	$C_{chl} = 10.0^{(0.366 - 3.067R + 1.930R^2 + 0.649R^3 - 1.532R^4)}$	443 > 490 > 510/555
MODIS/OC3M	$C_{chl} = 10.0^{(0.2830 - 2.753R + 1.4578R^2 - 0.6598R^3 - 1.4038R^4)}$	443 > 490/550

Reference	Lake	Algorithm: Regression	Notes
Bergmann et al. (2004)	Michigan	OC2v2: $C_{sat} = 0.81 * C, r^2 = 0.66$	No intercept given. All stations.
		OC2v2: $C_{sat} = 1.04 * C, r^2 = 0.95$	No intercept given. Non-cryptophyte
			stations.
		OC2v4: $C_{sat} = 0.68 * C, r^2 = 0.69$	No intercept given. All stations.
		OC2v4: $C_{sat} = 0.78 * C, r^2 = 0.83$	No intercept given. Non-cryptophyte
			stations.
		OC4v4: $C_{sat} = 0.65 * C, r^2 = 0.68$	No intercept given. All stations.
		OC4v4: $C_{sat} = 0.73 * C, r^2 = 0.83$	No intercept given. Non-cryptophyte
			stations.
Budd and Warrington (2004)	Superior	OC2v2: $C_{sat} = 3.1 * C + 0.4, r^2 = 0.87$	Without two outliers
Li et al. (2004)	Superior	OC2v2: $C_{sat} = 5.12 * C - 1.05, r^2 = 0.05$	Based on <i>in situ</i> radiances
		OC4v4: $C_{sat} = 3.82 * C - 0.50, r^2 = 0.05$	
Kerfoot et al. (2008)	Michigan	OC2v4: $C = 1.08 * C_{sat} - 0.087, r^2 = 0.84$	Regressed $C_{field}$ on $C_{sat}$ ; 1998-1999
			data; March-September
Watkins (2009)	Ontario	OC4v4: $C_{sat} = 0.42 * C + 1.01, r^2 = 0.62$	Offshore (> 30 m) sites; March-
			December
Witter et al. (2009)	Erie	OC2v4: $C_{sat} = 0.46 * C + 2.06, r^2 = 0.72$	All basins, spring and summer open-
			lake data
		OC4v4: $C_{sat} = 0.75 * C + 2.21, r^2 = 0.70$	
Kerfoot et al. (2010)	Michigan	OC2v4: $C = 1.209 * C_{sat} - 0.022$	21-23 April 2008 data
Barbiero et al. (2011)	Huron	OC4v4: $C_{sat} = 1.00 * C + 0.01, r^2 = 0.50$	2002-2006 spring open lake data

Table 2: Regression results from those studies with quantitative comparisons between chlorophyll values retrieved with standard OC2v2, OC2v4, and OC4v4 empirical algorithms and field measurements.

32

## List of Figures

1	Average Lake Michigan phytoplankton absorption spectrum, normalized to 440 nm, from Lohrenz et al. (2004). Data were collected during the springs of 1998-2000 from stations along three nearshore transects in the southeast portion of the lake. Vertical lines show the positions of the six	
	SeaWiFS bands in the visible portion of the spectrum	35
2	Theoretical subsurface reflection (non-dimensional) spectra for varying concentrations of chlorophyll- <i>a</i> calculated by using the method described by Bukata et al. (1991a). Optical cross sections for chlorophyll- <i>a</i> and for pure water are those determined by Bukata et al. (1985) for Lake Ontario and later used by Shuchman et al. (2006). Results shown are for the case when there are no other optically active component ( <i>e.g.</i> , suspended minerals and DOC) in the water. Vartical lines show the positions of the six	
	SeaWiFS bands in the visible portion of the spectrum.	36
3	Comparison of chlorophyll- <i>a</i> concentrations predicted by the current CZCS algorithm (O'Reilly et al., 2000) and four CZCS algorithms tested by	00
	Bukata et al. (1991b) for a range of reflectance ratios.	37
4	Chlorophyll- <i>a</i> concentrations predicted by the current CZCS algorithm (O'Reilly et al., 2000) and the best algorithm from Bukata et al. (1991b) versus measured chlorophyll values from Lake Ladoga. Values of the field measurements and of the reflectance ratios used to calculate the points on this figure were estimated from Fig. 1 in Bukata et al. (1991b). The solid line is 1:1, the dashed line is the best fit for Bukata algorithm, and the	
5	dash-dotted line is the best fit for the the current CZCS algorithm Chlorophyll- <i>a</i> concentrations ( $C_r$ ) retrieved by using the current SeaWiFS algorithm (OC4v6) versus modeled chlorophyll concentrations ( $C_m$ ) in the absence of any CDOM or NAP calculated with the Bukata et al. (1985) model for three chlorophyll absorption spectra. Dotted line indicates 1:1 relationship. Plotted points were selected randomly and used to calculate	38
	the linear regression coefficients.	39
6	Chlorophyll- <i>a</i> concentrations retrieved by using the current SeaWiFS algorithm (OC4v6) versus modeled chlorophyll concentrations assuming varying concentrations of NAP for CDOM concentrations of 0.0 g/ $m^3$ (panel a) 1.0 g/ $m^3$ (panel b) 2.0 g/ $m^3$ (panel c) 3.0 g/ $m^3$ (panel d) Solid	
	line indicates a 1:1 relationship. Plotted points were selected randomly.	40



Figure 1: Average Lake Michigan phytoplankton absorption spectrum, normalized to 440 nm, from Lohrenz et al. (2004). Data were collected during the springs of 1998-2000 from stations along three nearshore transects in the southeast portion of the lake. Vertical lines show the positions of the six SeaWiFS bands in the visible portion of the spectrum.



Figure 2: Theoretical subsurface reflection (non-dimensional) spectra for varying concentrations of chlorophyll-a calculated by using the method described by Bukata et al. (1991a). Optical cross sections for chlorophyll-a and for pure water are those determined by Bukata et al. (1985) for Lake Ontario and later used by Shuchman et al. (2006). Results shown are for the case when there are no other optically active component (*e.g.*, suspended minerals and DOC) in the water. Vertical lines show the positions of the six SeaWiFS bands in the visible portion of the spectrum.



Figure 3: Comparison of chlorophyll-*a* concentrations predicted by the current CZCS algorithm (O'Reilly et al., 2000) and four CZCS algorithms tested by Bukata et al. (1991b) for a range of reflectance ratios.



Figure 4: Chlorophyll-*a* concentrations predicted by the current CZCS algorithm (O'Reilly et al., 2000) and the best algorithm from Bukata et al. (1991b) versus measured chlorophyll values from Lake Ladoga. Values of the field measurements and of the reflectance ratios used to calculate the points on this figure were estimated from Fig. 1 in Bukata et al. (1991b). The solid line is 1:1, the dashed line is the best fit for Bukata algorithm, and the dash-dotted line is the best fit for the the current CZCS algorithm.



Figure 5: Chlorophyll-*a* concentrations ( $C_r$ ) retrieved by using the current SeaWiFS algorithm (OC4v6) versus modeled chlorophyll concentrations ( $C_m$ ) in the absence of any CDOM or NAP calculated with the Bukata et al. (1985) model for three chlorophyll absorption spectra. Dotted line indicates 1:1 relationship. Plotted points were selected randomly and used to calculate the linear regression coefficients.



Figure 6: Chlorophyll-*a* concentrations retrieved by using the current SeaWiFS algorithm (OC4v6) versus modeled chlorophyll concentrations assuming varying concentrations of NAP for CDOM concentrations of 0.0 g/m<sup>3</sup> (panel a), 1.0 g/m<sup>3</sup> (panel b), 2.0 g/m<sup>3</sup> (panel c), 3.0 g/m<sup>3</sup> (panel d). Solid line indicates a 1:1 relationship. Plotted points were selected randomly.



Figure 7: Chlorophyll-*a* concentrations ( $C_{OC4}$ ) retrieved by using the current SeaWiFS algorithm (OC4v6) versus modeled chlorophyll concentrations ( $C_m$ ) for randomly selected combinations of chlorophyll, CDOM and NAP concentrations and randomly selected absorption spectra for chlorophyll, NAP, and CDOM. Dashed line indicates a 1:1 relationship. and the solid line shows the best fit to the plotted points. The plotted points were selected randomly.