An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters

Article *in* Remote Sensing of Environment · March 2014

DOI: 10.1016/j.rse.2013.11.021 · Source: PubMed

4 authors:

Timothy S Moore
University of New Hampshire
23 PUBLICATIONS **516** CITATIONS
SEE PROFILE

Mark Dowell
European Commission
29 PUBLICATIONS **1,212** CITATIONS
SEE PROFILE

Shane Bradt
University of New Hampshire
7 PUBLICATIONS **41** CITATIONS
SEE PROFILE

Antonio Ruiz-Verdu
University of Valencia
15 PUBLICATIONS **335** CITATIONS
SEE PROFILE

Remote Sensing of Environment 143 (2014) 97–111

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters

Timothy S. Moore [a,*], Mark D. Dowell [b], Shane Bradt [c], Antonio Ruiz Verdu [d]

[a] Ocean Process Analysis Laboratory, University of New Hampshire, Durham, NH 03824, USA
[b] European Commission—Joint Research Centre, Institute for Environment and Sustainability, TP 272, via E. Fermi 2749, I-21027 Ispra, VA, Italy
[c] Department of Biological Sciences, University of New Hampshire, Durham, NH 03824, USA
[d] Image Processing Laboratory (IPL). Universitat de València Catedrático José Beltrán, 2E-46980 Paterna (València), Spain

ARTICLE INFO

Article history:
Received 15 July 2013
Received in revised form 16 November 2013
Accepted 22 November 2013
Available online xxxx

Keywords:
Remote sensing
Instruments sensors, and techniques
Computational methods and data processing
Biosignatures and proxies
Bio-optics

ABSTRACT

Bio-optical models are based on relationships between the spectral remote sensing reflectance and optical properties of in-water constituents. The wavelength range where this information can be exploited changes depending on the water characteristics. In low chlorophyll-*a* waters, the blue/green region of the spectrum is more sensitive to changes in chlorophyll-*a* concentration, whereas the red/NIR region becomes more important in turbid and/or eutrophic waters. In this work we present an approach to manage the shift from blue/green ratios to red/NIR-based chlorophyll-*a* algorithms for optically complex waters. Based on a combined *in situ* data set of coastal and inland waters, measures of overall algorithm uncertainty were roughly equal for two chlorophyll-*a* algorithms—the standard NASA OC4 algorithm based on blue/green bands and a MERIS 3-band algorithm based on red/NIR bands—with RMS error of 0.416 and 0.437 for each in log chlorophyll-*a* units, respectively. However, it is clear that each algorithm performs better at different chlorophyll-*a* ranges. When a blending approach is used based on an optical water type classification, the overall RMS error was reduced to 0.320. Bias and relative error were also reduced when evaluating the blended chlorophyll-*a* product compared to either of the single algorithm products. As a demonstration for ocean color applications, the algorithm blending approach was applied to MERIS imagery over Lake Erie. We also examined the use of this approach in several coastal marine environments, and examined the long-term frequency of the OWTs to MODIS-Aqua imagery over Lake Erie.

1. Introduction

Water quality properties are used as primary indicators for assessing lake and coastal water environmental viability by agencies to guide resource management and public safety decisions. These water quality properties include chlorophyll-*a* concentration, total suspended matter, Secchi depth, and nutrient concentrations, as well as the plant and animal species that inhabit these environments. Of these, chlorophyll-*a* concentration is arguably the most comprehensive environmental descriptor as it a measure of algal biomass and indicator of water clarity. *In situ* sampling remains the most accurate way of determining chlorophyll-*a* concentration, yet the use of satellite remote sensing for routine and synoptic chlorophyll-*a* monitoring has been increasing in the last decade in these types of environments (e.g., Binding, Jerome, Bukata, & Booty, 2010; Hunter, Tyler, Carvalho, Codd, & Maberly, 2010; Kloiber, Brezonik, Olmanson, & Bauer, 2002; Kutser, 2004; Olmanson, Brezonik, & Bauer, 2013; Yacobi et al., 2011).

Historically, the main applications of ocean color satellites and bio-optical algorithms have been directed towards open-ocean conditions. The optical properties of these environments are largely dictated by the concentration of phytoplankton and covarying material in the water, and have been referred to as 'case 1' waters (Morel & Prieur, 1977). Optical models designed to retrieve geophysical properties (e.g., chlorophyll-*a* concentration) in case 1 water have been modeled using the spectral light field in the blue-green part of the spectrum (e.g., Maritorena, Siegel, & Peterson, 2002; O'Reilly et al., 1998). These models begin to break down in environments where the optical properties are governed by materials other than phytoplankton—the so-called 'case 2' waters. Coastal regions and inland waters are highly susceptible to case 2 conditions from land effects (e.g., runoff of sediments, nutrients and organic matter) and re-suspension of sediments from shallow bottoms. In addition, the concentrations of particles including phytoplankton can be much higher compared to open ocean environments. As a consequence, bio-optical algorithms developed for the open ocean are less effective in more optically-complex waters found in coastal and inland waters (Melin et al., 2011; Moore, Campbell, & Dowell, 2009).

The development of bio-optical algorithms for eutrophic conditions more common to lakes and coastal regions has focused on wavelengths in the red and near-infrared (NIR) region of the light spectrum (Gitelson, Gurlin, Moses, & Yacobi, 2011; Gower, King, Borstad, & Brown, 2005; Hu et al., 2010; Matthews, Bernard, & Robertson, 2012;

* Corresponding author at: Ocean Process Analysis Laboratory, University of New Hampshire, USA. Tel.: +1 603 862 0690.
E-mail address: timothy.moore@unh.edu (T.S. Moore).

0034-4257/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.11.021

Yacobi et al., 2011). These algorithms achieve higher performance in highly eutrophic conditions compared to the open ocean case 1 algorithms (Gilerson et al., 2010), but often times it is not known which algorithm is best suited for a particular place or time in ocean color image scenes that contain both types of optical cases. The iconic case 1/case 2 system view that has predominated the view of aquatic optical classification for the last several decades is actually not an objective classification system, but a way to think about where and when algorithms are appropriate. If, as the evidence suggests, bio-optical algorithms perform better under certain situations and worse at times under different conditions, then a classification scheme is needed that can differentiate the environment and choose the more appropriate algorithm for the given environmental conditions.

Previous studies focused on optical classification of coastal and inland waters for bio-optical algorithm development/selection have been tested in a variety of environments. Melin et al. (2011) utilized a classification scheme that could select and blend type-specific bio-optical algorithms between two water types in the Adriatic Sea. Based on cluster analysis of remote sensing reflectance data, Vantrepotte, Loisel, Dessailly, and Meriaux (2012) showed that optical classes developed from coastal *in situ* data could identify the types in ocean color imagery. In addition, these optical classes could sufficiently represent other coastal areas not included in their data set. Le et al. (2011) developed a bio-optical classification scheme based on data from several lakes in China which could all be considered case 2 waters. Their results showed that optical classes were successful in selecting the best-performing algorithm for given optical case 2 conditions. Lubac and Loisel (2007) and Feng, Campbell, Dowell, and Moore (2005) developed optical classes for case 2 waters in the English Channel and Tokyo Bay, respectively. These studies support the emerging view that optical classes and algorithms vary within coastal and lake environments for waters that could collectively be termed as case 2, and not just between case 1 and case 2 waters.

Observed reflectance spectra from different case 2 waters share common features, as their optics are governed by similar factors including eutrophic/hypereutrophic trophic conditions, high loads of suspended sediments and colored dissolved organic matter (CDOM). Thus, coastal and lake environments may benefit from a common classification scheme using aggregate data that avoids specific regions or separation of fresh and marine waters. We have previously proposed a classification and blending scheme based on *optical water types* (OWTs) for oceanic regional and global scales (Moore, Campbell, & Feng, 2001; Moore et al., 2009). This global OWT system was based on open ocean and coastal waters with low to moderate levels of chlorophyll-*a*, and is not designed for lakes with moderate to extreme values of chlorophyll-*a*. Therefore, we are interested in adapting the application of the OWT method to inland lakes formed from a new data set representative of these types of waters. However, our interest is also to generalize complex optical water types across both coastal and inland water bodies to provide continuity from freshwaters to marine environments.

Our main objective is to 1) describe optical water types for coastal and lake environments which share similar levels of optical complexities, and to provide transition along the continuum of optical conditions between optical environments. Furthermore, we aim 2) to assess the feasibility of an optical classification system for blending the retrievals from multiple bio-optical algorithms for these optically-complex waters. The goal was not to advocate or promote any single algorithm, but to use two existing algorithms as case studies for the proposed classification framework.

2. Methods

To achieve the main goal, we sought to implement a classification system through defining *optical water types* from *in situ* remote sensing reflectance (R_{rs}) measurements covering a wide range of optical conditions in coastal and inland lake waters. This involves identifying the water types, sorting the data into respective subsets, and developing membership functions for the water types. The membership functions are the heart of the classification method, and the class memberships produced are used as the basis of weighting factors for blending algorithm retrievals into a single product when applied to satellite imagery. This process will be detailed in the following Sections.

2.1. Data sets

We assembled an aggregate data set from multiple sources with two requirements: 1) the reflectance measurements must have hyperspectral resolution collected from an above-water or near-surface radiometer, and 2) have co-measured chlorophyll-*a* measurements. These requirements were needed for two reasons: first, we are focused on capturing spectral features throughout the visible spectrum and into the NIR, and therefore need hyperspectral resolution. This also provides flexibility in adapting the derived OWT spectral reflectance characteristics to any satellite-specific band configuration and accommodates the use of existing and planned algorithms. Secondly, the need for reflectances in the red/NIR limits the use of profiling sensors that may not be sensitive enough to resolve the light field when descending through the water column. Additionally, our main interest was in evaluating chlorophyll-*a* products, although the conceptual approach applies to other bio-optical products from other types of algorithms (e.g., semi-analytical).

Several data sets from numerous freshwater lakes and coastal marine sites were combined. They comprise three main sources: a data set collected by the University of New Hampshire (UNH) in various northeastern US lakes as well as the Great Salt Lake in Utah (Bradt, 2012); a data set from Spain covering many lakes and trophic conditions (Ruiz-Verdu, Simis, de Hoyos, Gons, & Pena-Martinez, 2008); and a data set obtained from NASA's SeaBASS archive primarily from U.S. coastal marine sites (Werdell et al., 2003). All reflectance data were collected with hyperspectral instruments, which were binned at 3 nm intervals from 400 to 800 nm. The reflectance data were visually examined individually, and obvious erroneous spectra were not included in the final data set. The total number of reflectance data that passed our inspection with co-measured chlorophyll-*a* data was 488 points (Table 1).

In all cases, we are basing our analysis on the remote sensing reflectance denoted as the vector R_{rs}, which is defined as the ratio of the upwelling spectral light field to the downwelling spectral irradiance. All of our *in situ* data are in the above-water form $R_{rs}(0+)$; that is, it is the remote sensing reflectance just above the air–water interface. In our analyses with optical water types, we have converted the above-water form to below-water form, $R_{rs}(0-)$; that is, the remote sensing reflectance just below the air–sea interface. These two quantities are directly related, and we use the standard NASA conversion from above-water to below-water:

$$R_{rs}(0-) = \frac{R_{rs}(0+)}{0.52 + 1.7 * R_{rs}(0+)} \quad (1)$$

For clarity, the $R_{rs}(0+)$ form is used for all chlorophyll algorithm input.

Table 1
In situ data set summary.

Data set	N	Location	Source
UNH	140	NH lakes, Great Salt Lake	UNH
Spain	179	Assorted Spanish lakes	CEDEX
NASA	169	Coastal marine, U.S.	SeaBASS, NASA

2.1.1. Spanish lake data

The data set from Spain comprises 179 stations from over 60 lakes in a variety of trophic conditions, ranging from clear waters to those with monospecific blooms of cyanobacteria. The data were collected between 2001 and 2005 in several field campaigns conducted by the Spanish Centre for Hydrographic Studies (CEDEX). Measurements included above-water hyperspectral radiometry with an ASD-FR instrument, and chlorophyll concentration measured by HPLC techniques. The details of the sites and other measuring techniques/specifications are contained in Simis et al. (2007) and Ruiz-Verdu et al. (2008).

2.1.2. UNH lake data

The data set from UNH comprises 140 stations from over 60 lakes in the six New England states as well as the Great Salt Lake representing a variety of trophic conditions. These lakes were generally less eutrophic than those found in the Spanish data set. These data were collected between 2005 and 2009. The spectral measurements were based on methods adapted from Dall'Olmo and Gitelson (2005) and Doxoran et al. (2006); a profile of near-surface under-water light conditions was collected in each lake using with a dual radiometer system, and the profile data used to calculate remote sensing reflectance. Note, these data were not collected above water but near the surface; measurements were taken at 3–5 depths between the surface and 1 m for each profile, including one near-surface measurement at either 0.05 m or 0.10 m. The profile were used to estimate Rrs(0−), which was converted to Rrs(0+). Chlorophyll-*a* concentration was measured by spectrophotometric techniques. Further details of the sites and methods are contained in Bradt (2012).

2.1.3. SeaBASS coastal marine data

The NASA archive of bio-optical data—the SeaWiFS Bio-optical Archive and Storage System (SeaBASS)—was mined to acquire coastal marine data (Werdell et al., 2003). This archive was searched with restriction to 1) above-water hyperspectral R_{rs} and 2) co-measured chlorophyll. This search resulted in a data set that was assembled from multiple sources, but all were primarily from coastal waters around the continental United States. Specifically, data sets were combined from these SeaBASS sources: NOAA_CCMA, NORTHSEA, NRL, UCSB, and USF. Data were downloaded and re-assembled to coincide with the spectral range and wavelength intervals previously mentioned. Chlorophyll-*a* data were separately downloaded, and then matched to their respective radiometric measurements. A total of 169 matched points were obtained. These data were combined with the freshwater lake data.

2.2. Bio-optical algorithms under evaluation

Several recent published reports have described and compared the accuracy and use of chlorophyll-*a* algorithms for application in lake environments (Gitelson et al., 2011; Matthews et al., 2012; Odermatt, Gitelson, Brando, & Schaepman, 2012; Yacobi et al., 2011). These reports cover a wide variety of algorithms, including the traditional open-ocean empirical blue/green band-ratio algorithms and multiple-band red/NIR approaches. In this study, we examined two of the leading algorithms that showed good performance from the aforementioned studies—the standard NASA case 1 blue/green band-ratio algorithm OC4 (O'Reilly et al., 1998) and a 3-band MERIS algorithm (Gitelson et al., 2011), termed Mer-3B hereafter.

The NASA OC4 algorithm (version 6) uses a band-ratio in a 4th order polynomial exponential form defined as:

$$Chl_{OC4} = 10^{(0.327-2.994X+2.721X^2-1.225X^3-0.568X^4)} \quad (2)$$

where X is the log10 of the maximum ratio of the triplet Rrs443, Rrs490, Rrs510 to Rrs555.

The Mer-3B algorithm as taken from Gitelson et al. (2011) is defined as follows:

$$Chl_{MER-3B} = 243.86 \times \left({}^{1}\!/_{Rrs665} - {}^{1}\!/_{Rrs708}\right) \times Rrs753 + 23.17 \quad (3)$$

We note that we used these two algorithms on the *in situ* data and MERIS image retrievals. The NASA OC3 algorithm is the OC4 algorithm adapted to MODIS wavelengths without the 510 nm channel. Currently, MODIS-Aqua data lack the necessary channels to implement the Mer-3B algorithm. Thus, we did not apply chlorophyll-*a* algorithms to MODIS-Aqua imagery, although we classified MODIS-Aqua images.

2.3. Error definitions

One goal of the study was to assess the performance of select chlorophyll-*a* algorithms across the different optical water types. The metrics we used as performance indicators were the root-mean-square error (RMSE), bias (the average difference) and the median absolute relative error (MARE). The RMSE is the more comprehensive metric because it combines the mean and variance of the error distribution into a single term (Szeto, Campbell, Moore, & Werdell, 2011). The bias should be considered as well, as it speaks to systematic offsets that can reveal other sources of error. We also use the MARE in evaluating of chlorophyll-*a* performance.

We define RMSE as:

$$\mathrm{RMSE} = \sqrt{{}^{1}\!/_{N}\sum_{i=1}^{N}(\log10Chl_{mod} - \log10Chl_{meas})^2} \quad (4)$$

the bias as:

$$\mathrm{bias} = {}^{1}\!/_{N}\sum_{i=1}^{N}(\log10Chl_{mod} - \log10Chl_{meas}) \quad (5)$$

and the MARE expressed as a percentage:

$$\mathrm{MARE} = 100 * median\left[\frac{|Chl_{mod} - Chl_{meas}|}{Chl_{meas}}\right] \quad (6)$$

where Chl_{meas} is the *in situ* chlorophyll-*a* measurement and Chl_{mod} is the algorithm-derived chlorophyll-*a* concentration. We note that the calculations of bias and RMSE were performed on log-transformed chlorophyll-*a* data. It has long been known that in the ocean, the chlorophyll-*a* distribution is lognormal (Campbell, 1995). Our data set spanned 3–4 orders of magnitude in terms of chlorophyll-*a* concentration, and exhibits a lognormal distribution for chlorophyll-*a* (Fig. 1). Thus, for balanced error assessment across this dynamic range, log-transformation of chlorophyll-*a* is preferred for RMSE and bias.

2.4. Data clustering

Our choice for cluster analysis was the fuzzy *c*-means (FCM) algorithm (Bezdek, 1981) which has been used in many classification studies from medical image processing to remote sensing. A suite of cluster validity functions for deriving metrics was used to guide the choice of the optimal number of clusters. We note that the *k*-means clustering algorithm gives similar results and also could have been used in our application.

The FCM algorithm was applied to the *in situ* R_{rs} data. The FCM algorithm produces a fuzzy clustering of the data into a specified number of clusters (herein denoted as *c*). The basic function of this algorithm is to choose clusters that minimize the distance between the data points and the prototype cluster centers (or cluster means). Cluster centers are iteratively adjusted until optimization criteria are met (e.g., maximum number of iterations or minimum change residual). The clustering routine then returns the mean reflectance vectors for the *c* classes, and a matrix containing the memberships of each point to each class.

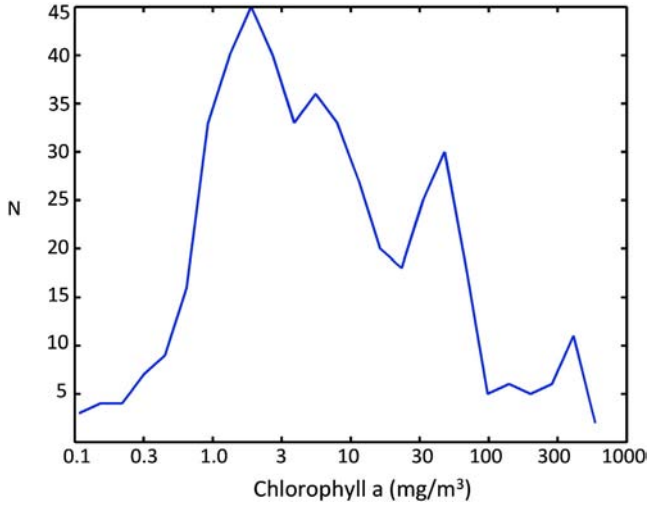

Fig. 1. Histogram of the combined in situ chlorophyll-*a* data set (N = 488).

The number of clusters, *c*, was an input to the FCM clustering routine, but the optimum number was not known in advance. We experimented with values of *c* ranging from 2 to 20, and evaluated the results with the aid of the cluster validity measures (Bezdek, Li, Attikiouzel, & Windham, 1997). Ten primary validity measures were used to objectively assess the optimal number of clusters, and the details are contained in Moore et al. (2009). Ideally, the validity measures should indicate the same optimum choice. This occurred in the present case, and we arrived at $c = 7$ as a best choice for our data set. The issue of the number of clusters is discussed later.

Once the clusters were identified, the reflectance spectra were sorted according to the highest membership value, and the mean R_{rs} and covariance matrix were calculated for each cluster. These statistics now define our optical water types, and are subsequently used in the membership function (defined below).

2.5. Algorithm blending using a fuzzy membership function

Our classification and blending scheme is based on the concept of fuzzy logic (Zadeh, 1965), and assigns class memberships for an observation to specified optical classes. The advantages of using a fuzzy logic approach are that 1) ambiguous boundaries between different classes can be captured with graded memberships and 2) the memberships can be used as weighted coefficients for class-specific bio-optical algorithms. This allows the blending of retrieved chlorophyll-*a* values from multiple algorithms without step-wise discontinuities that might arise under a 'this class or that class' approach.

The heart of the classification approach is the *membership function*. We defined our membership function using the Mahalanobis distance between the observation and the class mean vector, which is then input into a chi-square probability function. [The details of this membership function are contained in Moore et al. (2009), and will not be repeated here to the same detail.] The output is a number between 0 and 1, and represents the degree of 'closeness' of the observation to the mean. A value of 1 indicates the mean and the observation are identical. Values of zero indicate observations very far away from the mean. Two key statistics needed for the membership function are the reflectance mean vector and the covariance matrix, and are particular to each cluster. For any given cluster, the mean vector describes the basic magnitude and shape of its members. The spread of values at each wavelength shape the covariance matrix. Clusters with greater point spread will extend memberships to reflectance vectors further away from the mean compared to clusters with tighter distributions.

In this study, we used two chlorophyll-*a* algorithms. Thus, two weighting factors which sum to one are needed to blend the algorithm retrievals into one product. The choice of which algorithm is assigned to each water type is governed by the performance of each algorithm based on our *in situ* data set. Ultimately, we will end up with several water types assigned to each algorithm. The class memberships assigned to each of the algorithms are normalized by the membership sum to arrive at a weighting factor. The weights are given by:

$$w_i = \frac{\sum_1^n f_n}{\sum_1^c f_c} \tag{7}$$

where w_i is the normalized weighting factor for algorithm *i*, f_n is the fuzzy membership to the *n*th water type assigned to algorithm *i*, and f_c is the fuzzy membership to water type *c*. The blended chlorophyll-*a* product is then determined by:

$$Chl_{blend} = \sum_{i=1}^{2} w_i Chl_i \tag{8}$$

where Chl_i is the chlorophyll-*a* product from algorithm *i*. The underlying assumption is that the observed reflectance vector belongs to one of the defined water types. However, this is not always the case, which we will explore in the Discussion section.

2.6. Satellite imagery and test sites

Our main image location is Lake Erie, one of the Great Lakes in North America. It is an ideal test site—it is a large freshwater body amenable to remote sensing that is optically complex, and is considered a mesotrophic/eutrophic lake (Binding, Greenberg, & Bukata, 2012). It is subject to sediment re-suspension and whitings (Binding, Jerome, Bukata, & Booty, 2007; Binding et al., 2010), floating harmful algal blooms (Bridgeman, Chaffin, & Filbrun, 2013; Wynne, Stumpf, Tomlinson, & Dyble, 2010), and influences from colored dissolved organic matter (CDOM) from river input (Binding et al., 2010). Thus, it potentially contains a wide variety of optical water types.

Satellite images were obtained over Lake Erie from NASA's Ocean Biology Processing Group at Goddard Space Flight Center. Individual MODIS-Aqua and MERIS image scenes were processed at UNH from level-1 to level-2 using the standard default parameters in SeaDAS version 6.4. MODIS-Aqua images were subsequently remapped to a standard projection at 500 m resolution using band-interpolation provided by SeaDAS. The processed MODIS-Aqua daily image series ran from 2002 through 2012. The level-2 MERIS images (full resolution with nominal spatial resolution of 300 m) were also remapped to the same 500 m resolution and projection as the MODIS-Aqua imagery. The MERIS image series ran from 2009 through the end of 2011. Select MERIS scenes from several coastal marine areas were also processed using SeaDAS from level 1 to 2 for the Chesapeake Bay and the Yellow Sea.

3. Results

3.1. Cluster results

The cluster analysis separated and differentiated subsets based on both the shape and magnitude of R_{rs} and resulted in seven optimal clusters (Fig. 2). This number was deemed best based on a suite of cluster validity functions. When seven clusters are specified, the relation of data points to each other and cluster centers (mean vectors) in terms of compactness and separation aspects are collectively in a better configuration compared to other cluster choices. However, this number could change with a different input data set—either by adding to the existing data set or using a different one altogether.

The differences between clusters can be more readily seen when their reflectance means are plotted together (Fig. 3, Table 2). Collectively, these R_{rs} means form the coastal/lake *optical water types* (OWTs). They

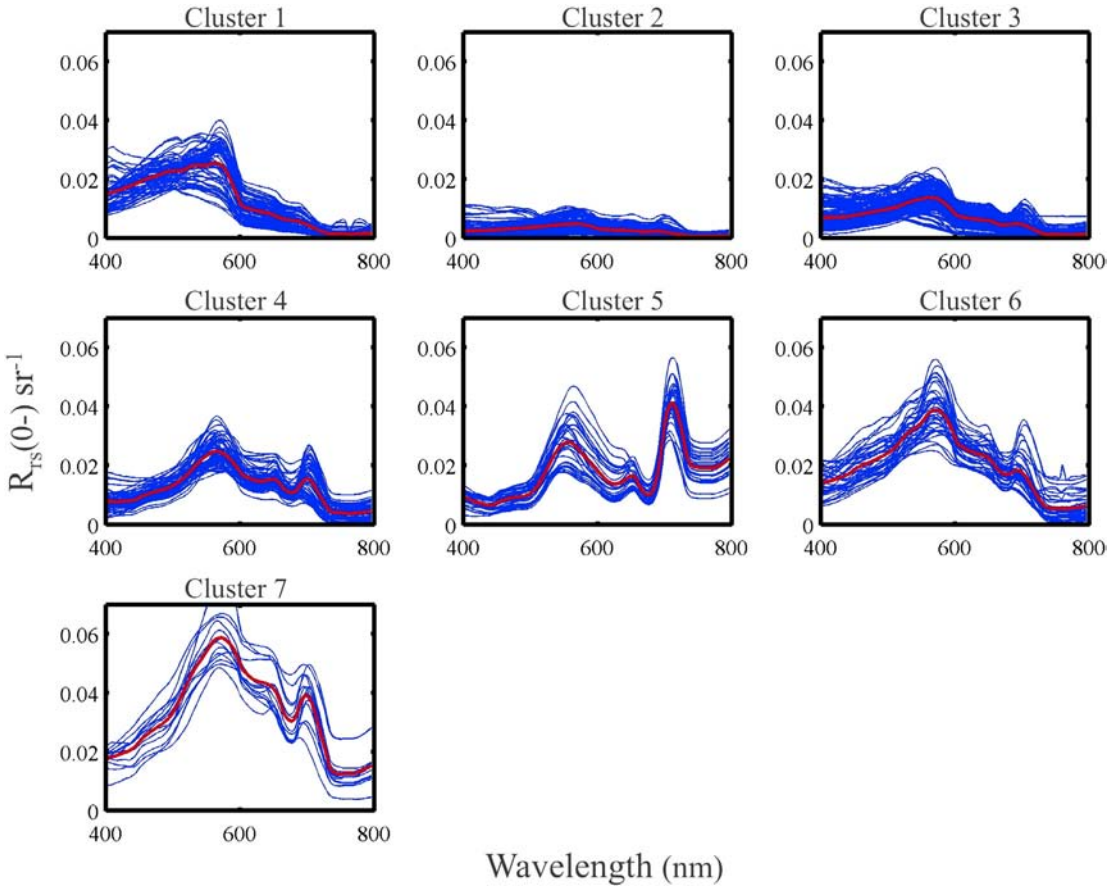

Fig. 2. Reflectance data sorted into the seven clusters from the fuzzy *c*-means cluster analysis (N = 488); blue lines: individual station reflectance data; red lines: mean reflectance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are representations of averaged conditions governed by the optical properties of the water column and ultimately depend on the absorption and scattering properties of the in-water constituents (e.g., phytoplankton and non-phytoplankton particles). These conditions are not unique to any particular lake, region, or a result of differences between freshwater and marine waters. Table 3 shows the distribution of the data from each data set across the OWTs. While not completely evenly spread, the distributions of each data set cover multiple OWTs with marine and freshwater stations found in the same clusters.

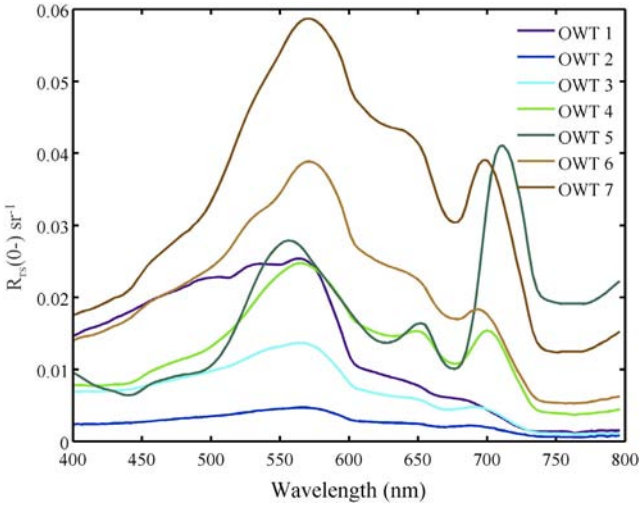

Fig. 3. The reflectance means of the seven optical water types (OWTs). The OWT means and covariance matrices serve as the basis for the membership function.

The OWTs were organized based on spectral features and ascending chlorophyll-*a* concentration system (Table 4). The OWTs show a pattern of increasing absorption in the blue/green for low red/NIR features (OWTs 1 through 3), followed by increasing peak magnitude at 555 nm (types 4 through 7). OWTs 1 through 5 show increasing chlorophyll-*a* concentration, while OWTs 6 and 7, which have lower mean chlorophyll-*a* values than OWT 5. In practice, the order of the OWTs is not important.

In general, the OWT R_{rs} spectra contain unique and separate characteristics. OWTs one through three have low overall spectral magnitude, and show relatively flat features from 600 nm onward compared to the other OWTs. Low particle concentration combined with water absorption could result in these types of reflectance spectra. Conversely, OWTs 4 through 7 show higher overall magnitudes and more features especially in the red/NIR region. These OWTs all show peaks around 700 nm, but are different from each other in magnitude. OWT 5 shows a strong peak at 700 nm compared to its overall magnitude in the green part of the spectrum. This peak is characteristic of strong particle backscattering and has been associated with high algal particle concentration (Gower et al., 2005; Zimba & Gitelson, 2006; Gilerson et al., 2007).

All of the OWTs show a reflectance peak to some degree at or near 555 nm, and is most pronounced in OWTs 4, 5, 6 and 7. The peak at 555 nm can be attributed to enhanced particle scattering from living (e.g., phytoplankton) and non-living (e.g., sediments) sources (Ahn, Bricaud, & Morel, 1992; Kutser, 2004). Other secondary peaks are seen at or near 650 nm in these OWTs. There are also spectral regions of similarity or overlap between several types. OWT 1 and 6 overlap in the blue region from 400 to 500 nm, but clearly separate afterwards. OWTs 3 and 4 share similar magnitude and shape from 400 to 700 nm, but are separated by the peak height at 709 nm. While it is

Table 2
Mean Rrs(0−) spectra of the optical water types at the MERIS wavelengths.

	Wavelength (nm)									
OWT	412	443	490	510	560	620	665	680	709	753
1	0.0157	0.0181	0.0223	0.0227	0.0252	0.0096	0.0063	0.0058	0.0034	0.0014
2	0.0023	0.0025	0.0033	0.0036	0.0047	0.0027	0.0020	0.0021	0.0016	0.0006
3	0.0070	0.0072	0.0093	0.0103	0.0135	0.0067	0.0046	0.0045	0.0036	0.0011
4	0.0077	0.0081	0.0121	0.0148	0.0245	0.0152	0.0125	0.0110	0.0138	0.0039
5	0.0081	0.0064	0.0095	0.0132	0.0278	0.0141	0.0125	0.0109	0.0410	0.0191
6	0.0146	0.0171	0.0229	0.0261	0.0372	0.0252	0.0185	0.0171	0.0149	0.0053
7	0.0183	0.0212	0.0305	0.0379	0.0571	0.0446	0.0339	0.0311	0.0352	0.0123

not possible to definitely associate these features to unique constituents without more complete optical information, phytoplankton are playing a significant role in the shape of the reflectance spectrum.

3.2. Chlorophyll-a properties and algorithm performance

Values for the mean and median of the chlorophyll-*a* concentration increased from OWT 1 to OWT 5 (Table 4). OWT 6 and 7, while still having high mean chlorophyll-*a* levels, are lower than OWT 5 but comparable to OWTs 3 and 4 in terms of concentration values. It should be noted that with the exception of OWTs 5 and 7, all the OWTs contain station data with minimum chlorophyll-*a* levels at or below 1 mg/m^3. In addition, all types contain station data with maximums that exceed 10 mg/m^3. Chlorophyll-*a* values between 1 and 10 mg/m^3 are an important range, as it is the region where the OC4 and Mer-3B algorithms transition in terms of performance (Fig. 4). In the case of OC4, when chlorophyll-*a* exceeds 10 mg/m^3 there is an increase in algorithm uncertainty. The opposite is true of Mer-3B, which exhibits better performance at higher chlorophyll-*a* values. This is reflected in the error statistics.

It is important to note that there were a significant number of negative chlorophyll-*a* retrievals when using this algorithm (N = 186 out of 488). This arises when Rrs665 is smaller than Rrs709 (Eq. (3)), and typically occurs when the chlorophyll-*a* concentration is less than 10 mg/m^3. These points were excluded for the bias and RMSE calculations because of the log transformation. These were present in OWTs 1, 2, 3, 4 and 6. However, they were included in the calculations for the MARE. However, the RMSE and MARE reveal the same overall performance patterns.

The results of the RMSE, MARE and bias for each chlorophyll-*a* algorithm are shown in Tables 5, 6 and 7, and graphically shown in Fig. 4. The RMSE for both algorithms based on the entire data set are similar (RMSE of 0.416 and 0.437 for OC4 and Mer-3B, respectively), while the MARE is higher overall for Mer-3B (66.1% and 146.1% for OC4 and Mer-3B, respectively). Considering the whole data set, the overall RMSE does not reflect the performance of the algorithms for individual water types. For both OC4 and Mer-3B, lower RMSE's for some OWTs are compensated by higher RMSE's in environments where the algorithm performs poorly. Thus, overall RMSE is not particularly informative regarding each algorithm's effectiveness in certain environmental conditions.

However, algorithm errors derived for individual OWTs reveal the performance for a narrower range of optical and environmental conditions. Regarding OC4, RMSE generally increases progressing from OWT 1 to OWT 7. In contrast, RMSE for Mer-3B generally decreases over the same progression of OWTs. The OC4 algorithm has lower RMSE to Mer-3B in OWTs 1 through 3, ranging from 0.295 to 0.330. From OWTs 4 to 7, Mer-3B has lower RMSE with the exception of OWT 6, ranging from 0.194 to 0.357. OWT 6 shows better performance with OC4, with an RMSE of 0.345 compared to 0.438 for Mer-3B. These patterns also hold for the MARE over the same OWTs. OWTs 1, 2, 3 and 6 have lower MARE for OC4 compared to Mer-3B, and OWTs 4, 5 and 7 have lower MARE for Mer-3B compared to OC4.

The OC4 algorithm showed a systematic underestimation or negative bias (Table 7) in OWTs 3 through 7 (except OWT 6) and an overestimation (positive bias) for the lower range (OWTs 1 and 2). In contrast, the Mer-3B algorithm showed a large positive bias in the OWTs 1 and 2 (chlorophyll-*a* values under 10 mg/m^3), and a smaller bias in the higher OWTs and chlorophyll range.

3.3. Algorithm blending

Memberships to the OWTs are the key for selecting and blending algorithm retrievals. By assigning the best performing algorithm (i.e., lowest RMSE) to a particular water type, the memberships can be used to weight chlorophyll-*a* retrievals into a blended product. To arrive at a blended chlorophyll-*a* product, the memberships need to be converted to normalized weights per Eq. (6). For the present algorithm assessment, the weighting factor for the OC4 product is the sum of the normalized memberships for OWTs 1, 2, 3 and 6, and the weighting factor for the Mer-3B products is the sum of the normalized memberships for OWTs 4, 5 and 7.

The blending method (Section 2.5) was applied to the *in situ* data set using R_{rs} as input to the classification algorithm and the chlorophyll-*a* algorithms to derive a blended chlorophyll-*a* product. The overall RMSE improved for the blended chlorophyll-*a* product, reducing to 0.320 from 0.416 to 0.437 for the OC4 and Mer-3B algorithms, respectively (Table 5). The overall MARE also improved to 47.9% compared to 66.1% and 146.1% for OC4 and Mer-3B, respectively (Table 6). In addition to lower RMSE and MARE, the blended chlorophyll-*a* product shows improvement in the bias, and while not eliminated completely it is reduced with the blended product (Table 7).

Table 3
In situ data distribution across OWT.

OWT	Spanish lakes	UNH lake data	SeaBASS coastal	Total
1	26	18	21	65
2	24	71	23	118
3	42	15	91	148
4	42	6	22	70
5	28	4	0	32
6	12	18	12	42
7	5	8	0	13
Total	179	140	169	488

Table 4
In situ chlorophyll-*a* characteristics of the OWTs.

OWT	Chl min	Chl Median	Chl Mean	Chl Max	Std. Dev
1	0.41	1.07	1.97	26.61	3.83
2	1.08	2.21	3.11	14.20	2.67
3	0.53	5.25	6.62	41.12	6.76
4	0.93	11.28	20.85	129.33	23.10
5	10.35	137.02	206.37	705.0	177.50
6	0.45	4.68	10.80	69.44	15.01
7	2.21	33.82	36.00	108.51	26.57

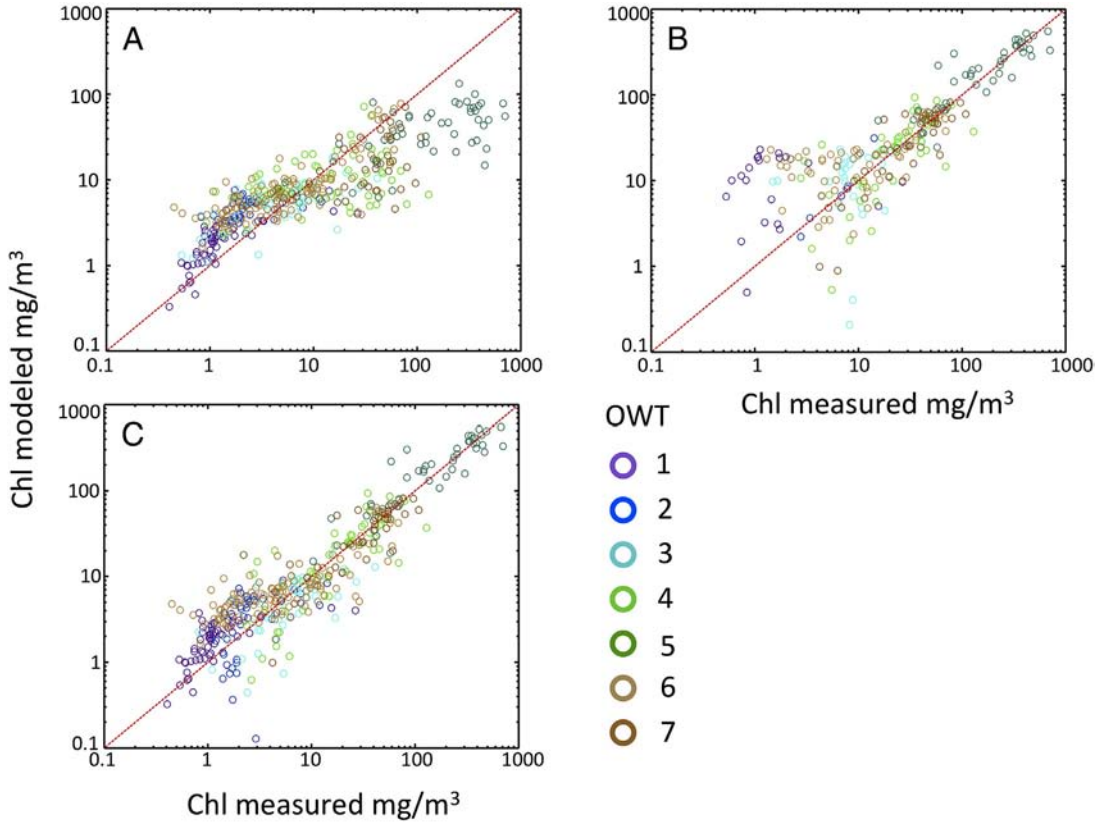

Fig. 4. Modeled versus measured chlorophyll-*a* plots. A: OC4 versus *in situ* chlorophyll-*a*; B: Mer-3B versus in situ chlorophyll-*a*; C: the blended versus *in situ* chlorophyll-*a*.

3.4. Satellite image application

The OWTs can be mapped from ocean color imagery using the membership functions with image spectral reflectance as input. Since the OWT base statistics were formed from hyperspectral data, the spectral channels from any given satellite can be matched and used to select the necessary wavelengths from the OWT base vectors to form satellite-specific membership functions. We will show two examples using a MERIS and MODIS-Aqua image over Lake Erie for our test application.

Fig. 5 shows the raw unnormalized membership maps for a MERIS image from September 3, 2011 over Lake Erie. For this case, the R_{rs} bands used in the classification were 443 nm, 490 nm, 510 nm, 560 nm, 620 nm, 667 nm, 680 nm, 709 nm, and 748 nm. The 412 nm channel was available but not used because of questionable accuracy in this image. This particular image shows a complex optical environment across the lake, with a large plume containing a known cyanobacteria bloom entering and developing in the lake from the southwestern corner, while the central and eastern basins were optically and environmentally different. The membership maps shown in Fig. 5 portray the degree to which pixels are identified to each of the OWTs (OWTs 5 and 7 are not shown as there were very few pixels with any membership to either of these OWTs). The membership map for OWT 1 shows strong membership in a spatially coherent region which dominates the central part of the lake, intertwined with patches that are shades of blue (intermediate membership) to black (no or low membership). The fuzzy membership map for OWT 3 shows a spatially coherent feature stretching from the southwest corner of the lake across the central portion to the northern shore, and weaker but significant membership in the far eastern basin. Similarly, the dark regions are areas where there is no membership to this OWT (i.e., the image spectra are not similar to the characteristic mean vector of OWT 3). The 'hard' image is the expression of the dominant OWT (i.e., the OWT with the maximum membership) for the image and shows these two OWT dominating much of the lake.

Fig. 6 shows the OWT membership maps for the same-day MODIS-Aqua image as in Fig. 5. For this case, the R_{rs} bands used in the

Table 5
RMSE for the OC4, Mer-3B and the blended chlorophyll-*a* products (in log10 units) by OWT.

OWT	OC4[a]	MERIS 3-Band[a]	Blend[a]
1	*0.295*[A]	0.910	0.286
2	*0.330*[A]	0.425	0.370
3	*0.307*[A]	0.517	0.303
4	0.479	*0.357*[A]	0.310
5	0.670	*0.194*[A]	0.197
6	*0.345*[A]	0.438	0.384
7	0.460	*0.289*[A]	0.254
All	0.416	0.437	0.320

A—denotes optimal algorithm for that water type.
[a] Excludes chlorophyll retrievals that were negative from MERIS 3-band algorithm.

Table 6
Median Absolute Relative Error (MARE) for the OC4, Mer-3B and the blended chlorophyll-*a* products (in percent) by OWT.

OWT	OC4 (%)	MERIS 3band (%)	Blend (%)
1	72.3	1278	64.6
2	109.8	1174	62.7
3	41.2	216.8	48.8
4	71.0	71.2	51.4
5	73.1	22.6	23.7
6	56.1	196.8	55.8
7	62.8	24.3	20.0
All	66.1	146.1	47.9

Table 7
Bias for the OC4, Mer-3B and the blended chlorophyll-*a* products (in log10 units) by OWT.

OWT	OC4[a]	MERIS 3-Band[a]	Blend[a]
1	0.167	0.687	0.264
2	0.110	0.210	0.103
3	−0.105	0.105	−0.095
4	−0.234	−0.041	−0.193
5	−0.554	0.054	0.054
6	0.021	0.185	0.167
7	−0.314	0.043	−0.043
All	−0.194	0.115	0.023

[a] Excludes chlorophyll-*a* retrievals that were negative from MERIS 3-band algorithm.

classification were 443 nm, 489 nm, 531 nm, 547 nm, 665 nm, 678 nm, 681 nm, and 748 nm (the 412 nm channel was left out). While the band sets used to derive the memberships were different between MODIS-Aqua and MERIS, the maps are very similar. A key difference is the lack of the 620 nm and 709 nm channels in MODIS-Aqua. This leads to some differences between the mapped OWT distributions. The overall degree of difference depends on satellite band differences and the quality of image reflectance data. For this particular image, omitting these bands (620 nm and 709 nm) had minor influence on the OWT maps. This occurs because the information contained in the OWTs and their membership functions are still distinct with the reduced bands (i.e., the MODIS-Aqua bands). Since the membership functions are based on collective shape and magnitude of reflectance, the role of any single band—while important—is tempered.

We applied the blending scheme to the same MERIS image over Lake Erie (Fig. 7). In this image, most of the central and eastern part of Lake Erie is weighted towards the OC4 algorithm, while the western part is weighted towards the Mer-3B algorithm. The Mer-3B retrievals for the central and eastern region were either zero or below a noise threshold for the 753 nm channel. We used a value of 0.00025 sr^{-1} as a noise cutoff value based on the global standard deviation for eutrophic water from a recent NASA analysis (http://oceancolor.gsfc.nasa.gov/ANALYSIS/global/ar2013.0m_ar2013.0m/tables/).

4. Discussion

4.1. The nature of optical water types

The notion of optical water types has existed in the literature since Jerlov (1951) introduced the term to describe marine environments that had different optical characteristics based on light attenuation. Light attenuation is directly linked to the absorption and scattering properties of the water column, and influences the spectral reflectance at the surface of the water. Our use of reflectance as the basis for optical water types is analogous to land cover classification schemes. An important difference is that land cover classification bases land classes on true end-members, whereas in our application the aquatic water types are representations of average optical conditions, and do not necessarily describe true end-member classes.

In our scheme, we have adopted a fuzzy classification approach towards assigning pixels or observations to the optical water type classes. The fuzzy memberships do not represent the proportion of each water type within the pixel, as is the case with land remote sensing. This implies that end-member classes within a pixel can be resolved with increased spatial resolution. Since mixing occurs at the microscopic scale in aquatic environments, such differentiation is not possible. A pure end member class that could be described is that of pure water, and that condition does not exist in freshwater or marine environments. The optical water types themselves are representations of optical conditions that are snapshots of a continuum. However, we have identified water types with clearly different reflectance characteristics and by association optical properties, and the functional use of the water types is equivalent to an end-member in the membership function.

4.2. How representative are the OWTs?

Currently, we do not know whether all 'possible' optical classes in coastal and lake systems are represented in this scheme. We are limited by the data used in this study. While we sought to generalize the variety of optical types from coastal to lake systems, it is possible and even likely that the data collected do not represent all conditions. It is likely that

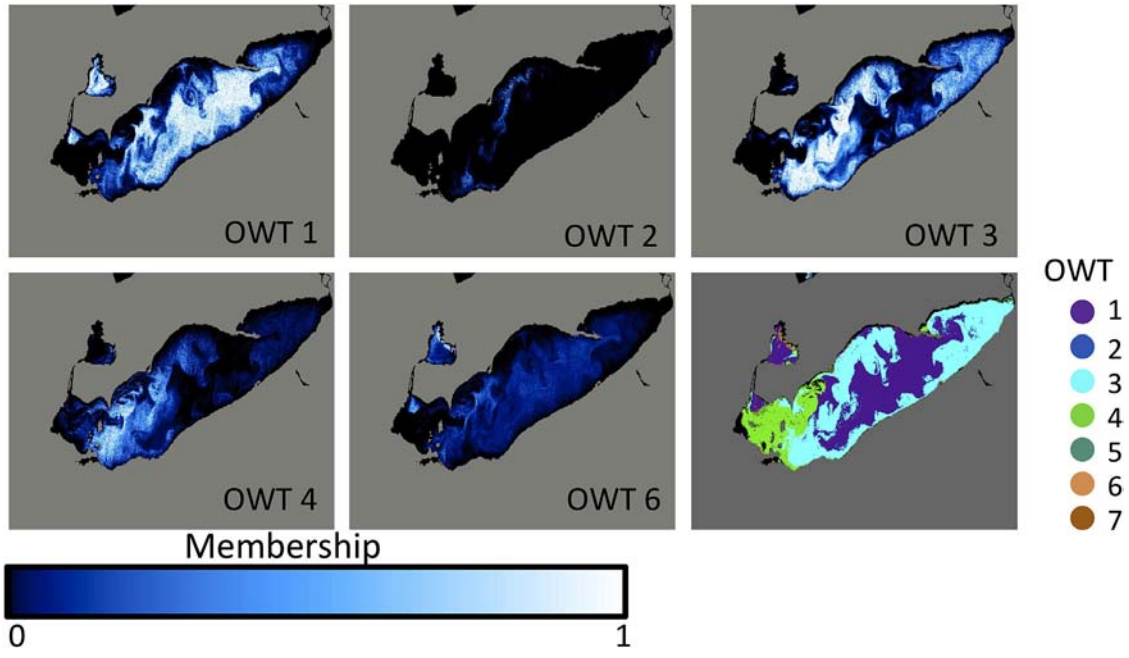

Fig. 5. Membership maps for a MERIS image over Lake Erie—Sept. 3, 2011. White areas denote high membership; black areas denote zero or low membership; shades of blue are intermediate values. The dominant OWT—expressed as the water type with the highest membership—is shown in the lower right panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

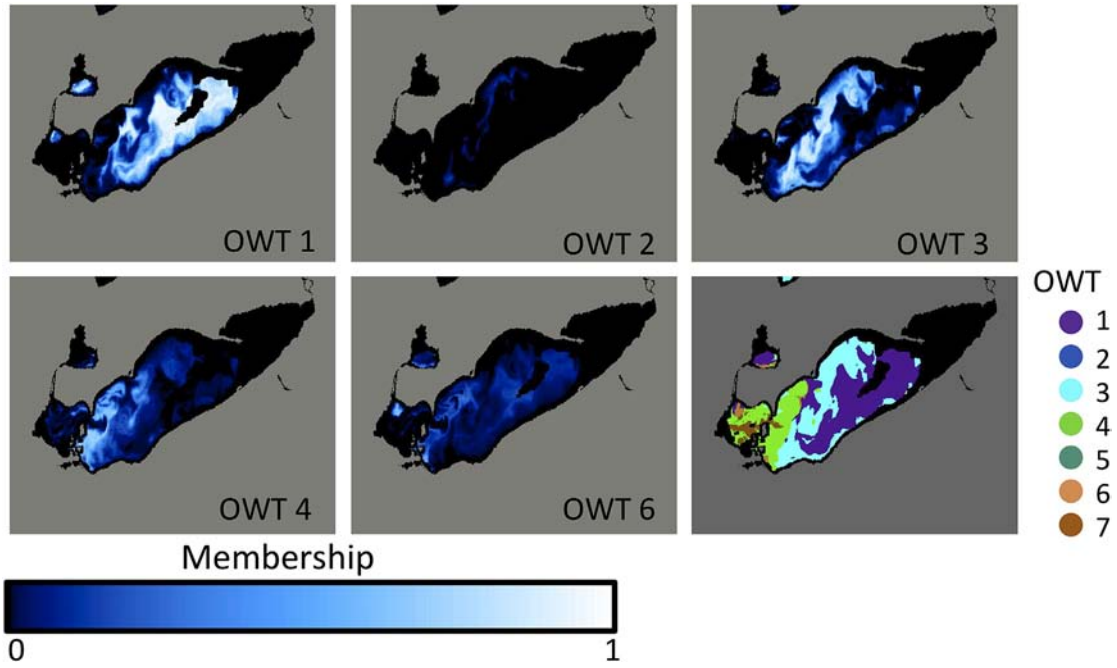

Fig. 6. Membership maps for a MODIS-Aqua image over Lake Erie—Sept. 3, 2011. As in Fig. 5, white areas denote high membership; black areas denote zero or low membership; shades of blue are intermediate values. The dominant OWT—expressed as the water type with the highest membership—is shown in the lower right panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

different systems may only express a subset of the OWTs for their range of optical variability, and would not indicate either way how representative the OWTs relate to all conditions. Conversely, it is also likely that a system or region may express low affinity to any of the OWTs, and could signify that certain optical conditions are not represented in our scheme.

To gain insight into this question, we examined long-term patterns of the dominance of each OWT (or the lack of) in Lake Erie from MODIS-Aqua data. The *dominant* OWT for each pixel was based on the OWT with the maximum membership. The frequency of dominant OWT expression was generated by counting the dominant OWT for each pixel for individual daily scenes over the entire mission time period (2002–2012), and normalizing by the total count of valid water pixels (i.e., passed atmospheric correction). Dominant OWT pixels (numerator) were counted if the sum of the un-weighted membership exceeded a threshold. In our study, we used a threshold value of 0.10.

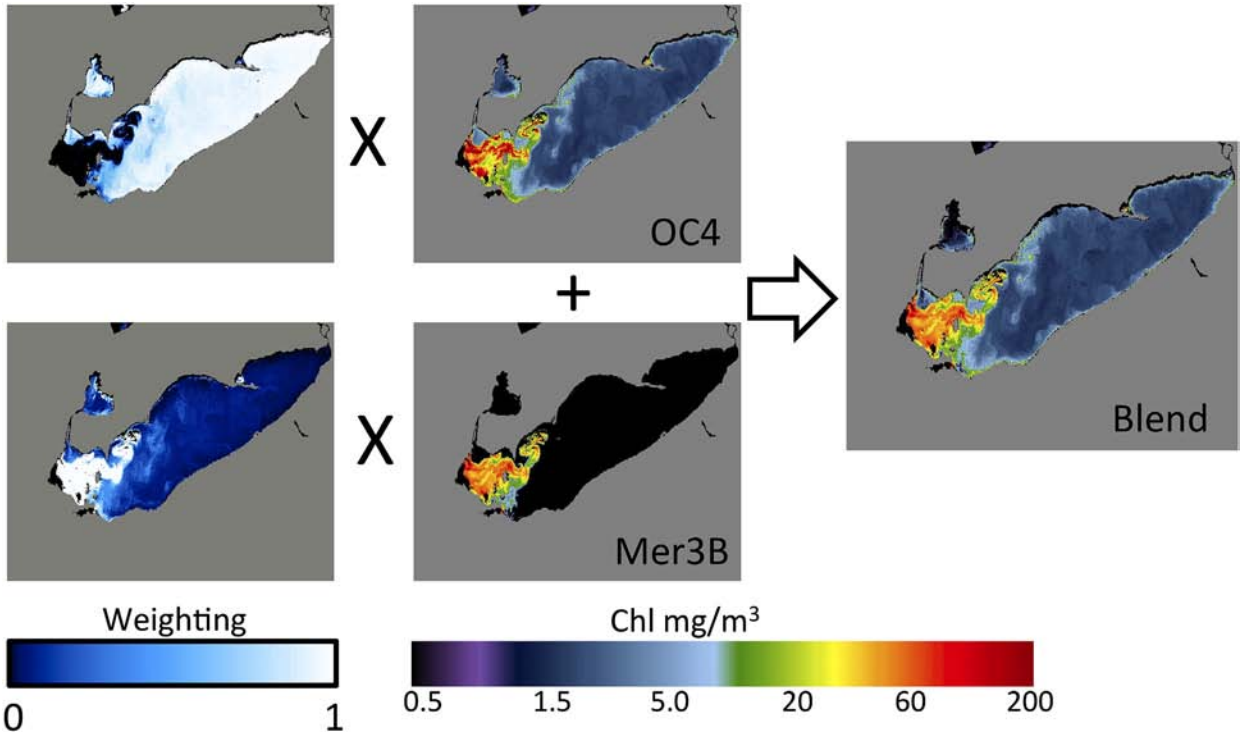

Fig. 7. Two chlorophyll-*a* products (based on OC4 and Mer-3B) are blended into a single image using the normalized OWT memberships to weight the two chlorophyll-*a* products from the MERIS image on Sep. 3, 2011 (same as Fig. 5).

This value is adopted from Vantrepotte et al. (2012) and Melin et al. (2011)—two studies that also used the Mahalanobis distance for classifying ocean color pixels. These threshold levels are subjective and arbitrary, as the Mahalanobis distance is subject to the choice and accuracies of the particular wavelengths used in the calculation. Pixels that did not meet this criterion were counted towards 'missing' cases (cloud/ice and pixels without radiance were excluded). In the case of the MERIS Sep. 3, 2011 image, 78.6% of the pixels had membership sums greater than a sum threshold of 0.10. In comparison, 87.7% of the pixels were valid when the threshold was reduced to 0.01.

The frequency maps show the *persistence* of dominant OWTs over time (Fig. 8). Red colors indicate regions that are dominated by a water type for a majority of the time. Whereas long-term chlorophyll-*a* distributions reveal patterns related to the ratio of blue to green reflectance, OWT persistence maps in effect show the long-term distribution of the overall spectral character (shape and magnitude) of remote sensing reflectance. This information reveals important characteristics of the overall spatial distribution of radiance patterns, and the use of potential chlorophyll-*a* algorithms for this region.

Much of central and eastern basins of Lake Erie are dominated by OWT 3. In our analysis, OC4 had better performance in OWT 3 than Mer-3B. The chlorophyll-*a* range for this OWT (Table 4) extends from a low of 0.53 to 41.1 mg/m^3, with a mean of 6.62 mg/m^3. This range agrees with chlorophyll-*a* values presented in Binding et al. (2012) for Lake Erie. In these areas, waters are strongly associated with OWT 3 more than 50% of the time, and higher in portions of the eastern basin. In contrast, the western basin is associated with OWTs 1, 4 and 5. OWTs 4 and 5 are types that are characterized by high chlorophyll-*a* and show better performance with the Mer-3B algorithm. From this perspective and in regards to Lake Erie, different chlorophyll-*a* algorithms are more appropriate for different parts of the lake and at different times.

The persistence maps can also be used to locate areas of 'low membership'. These areas might reflect potential missing water types or problems with atmospheric correction. Fig. 9 shows the mapped frequency distribution of low memberships (membership sums less than the 0.10 threshold) for Lake Erie, and the average unnormalized membership sum for the entire mission period. Overall, areas with low memberships (and low membership sums) occur along the central and eastern shorelines up to 50% of the time. In western Lake Erie low membership frequency occurs up to 10% of the time, although the average membership sums are at or greater than one. Membership sums greater than one are permitted in our scheme, as ultimately these are normalized to a sum of one when used as weights. These areas reflect high membership to multiple water types, which may indicate that some OWTs are close in proximity to one another and can compensate for times when there is low membership in the average.

It is unclear if areas of low membership are related to radiance errors from atmospheric correction (e.g., aerosol characterization, land adjacency effects, shallow water bottom reflection), a lack of representation in the current OWT characterization, or sensitivities of the classification algorithm. All of these sources could cause memberships to be low to the OWTs. Vantrepotte et al. (2012) attribute quality issues with satellite R_{rs} as a significant factor for the presence of unclassified pixels in their scheme, but also include missing spectral shapes from the training set as a possibility. Further insight can be gained by directly comparing the satellite radiances with the mean OWT vectors. These are shown in Fig. 10 for the same MERIS image over Lake Erie from Sep. 3, 2011. While these spectra are from a MERIS image, the analysis is representative for MODIS-Aqua as well.

For each OWT, extracted image spectra generally follow the shape and magnitude of the mean vector of the OWT. However, the membership values to the OWTs could be high or low. For OWT 5, the spectra shown have memberships to this OWT of less than 0.10, but it is still the dominant OWT. For this case, it can be seen that at 709 nm the reflectance values from the image are much less than the OWT mean value, but overall these spectra are closest to the mean vector for that OWT. The accuracy of image reflectance values at this wavelength is unknown without additional ground truth.

Fig. 10 also shows R_{rs} spectra with negative values at 412 nm for OWTs 2, 3, 4 and 6. The memberships to OWTs 2 and 6 (Fig. 5) are lower than 0.25. The accuracy of radiances at this wavelength has also been a long-standing issue with ocean color data especially in coastal regions (Bailey, Franz, & Werdell, 2010; Zibordi et al., 2009). This is a result of the atmospheric correction scheme and the problems associated with aerosol characterizations near land. In addition, non-zero

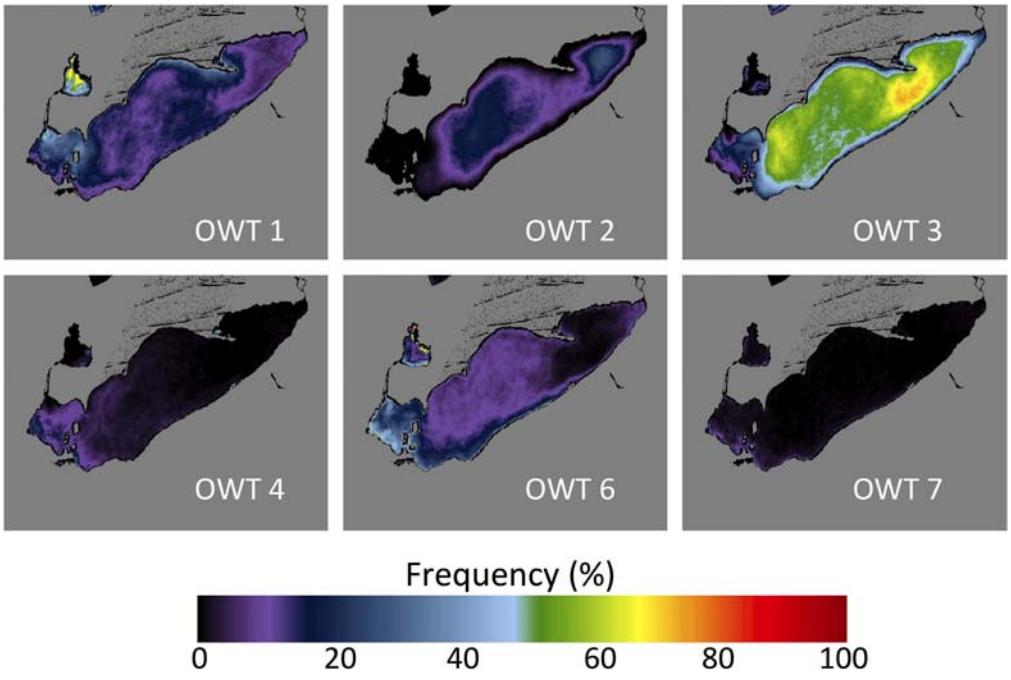

Fig. 8. OWT persistence maps for MODIS-Aqua from 2002 through 2012 over Lake Erie (OWT 5 not shown due to lack of persistence). Persistence is defined as the frequency of the dominant OWT of valid pixels (membership sum > 0.10) normalized to the total number of water pixel counts (i.e., pixel that passed atmospheric correction).

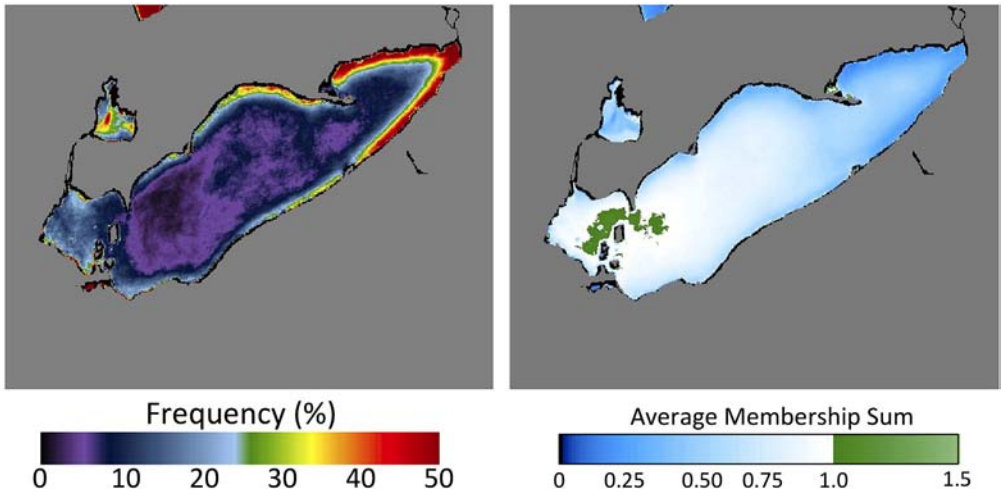

Fig. 9. Left: OWT persistence map for of low membership values (<0.10) for MODIS-Aqua from 2002 through 2012 over Lake Erie Note: Color scale range is from 0 to 50%. Right: Average membership sum of non-normalized memberships for same date range. Green area is where memberships sum are greater than one. Shore areas along the central and eastern edges of the lake are most susceptible to low membership conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reflectance in the NIR region impacts all bands in atmospheric correction as the NIR region is used for assessing aerosol models that are interpolated from the NIR to all other wavelengths. For these reasons, we omitted this band from the membership function (see Section 3.3). It is worth noting that Melin et al. (2011) and Vantrepotte et al. (2012) also discarded this wavelength from their classification schemes. Ultimately, *in situ* radiance validation data are needed to determine satellite radiance accuracies.

One of our aims was to assess how applicable the lake/coastal OWTs are to coastal regions. To investigate this aim, we applied the classification to a MERIS image over Chesapeake Bay and the Yangtze River/Yellow Sea—two coastal regions with turbid waters. Across much of the scenes, high membership sums are present along with areas of membership sums greater than one, indicating multi-type overlap (Fig. 11). The hard OWTs also show reasonable and expected distribution patterns.

In both images, turbid waters were present, and in some areas failed using the default atmospheric correction. Turbid areas that did pass were assigned to OWTs 6 and 7. Yet, the membership sums in these areas are extremely low. After examination of the satellite radiance fields from these areas, the spectra are nearest to OWT 6 (or 7) relative to the other OWTs, but significant differences exist in at least a few bands when compared to the mean spectra for OWT 6 (or 7). As in the case of Lake Erie, it is unclear if the low memberships are a result of atmospheric correction, a lack of OWT representation or both.

4.3. Concerning satellite red/NIR accuracy

In general, the accuracy of satellite radiances in the red/NIR remains problematic in ocean color data, and is a concern to bio-optical algorithms that rely on wavelengths in this region and also to the classification scheme. The issues of accuracy stem from atmospheric correction assumptions in the NIR region, which begin to break down in turbid waters (Goyens, Jamet, & Schroeder, 2013). In addition, aerosol properties can be difficult to model and remove in such regions as well. New techniques are emerging for improving accuracy of radiances in coastal, turbid areas (e.g., Doron, Belanger, Doxoran, & Babin, 2011; Ruddick, De Cauwer, & Park, 2006; Stumpf, Arnone, Gould, Martinolich, & Ransibrahmanakul, 2003; Wang, Son, & Shi, 2009), which should eventually lead to more accurate radiances in these bands. In this study, images were processed with default settings in SeaDAS. Mouw et al. (2013) found that the standard atmospheric correction worked best over Lake Superior. However, this evaluation was based on the blue-green region. Binding et al. (2012) showed good agreement between modeled and satellite-derived radiances at 667 nm and 748 nm for Modis-Aqua imagery in Lake Erie. Those images were processed using the default iterative NIR scheme within SeaDAS, and offer some insight on the use of the default scheme for this region.

We did not set out to test and evaluate different atmospheric correction schemes. Such an evaluation would need to be site-specific, and require match-up data between *in situ* and satellite-derived radiances that include the red/NIR region. Thus, assessing how effective OWT membership maps are at characterizing the underlying water environment is impaired by errors in atmospheric correction and the dependencies of the membership functions on the information in the red/NIR region. Further analysis is needed to evaluate accuracy in Lake Erie and other regions.

4.4. Is the chlorophyll product improved?

Our assessment of a blended chlorophyll-*a* product approach showed lower RMSE and MARE than for either of the single algorithms over the entire range of the *in situ* data set. From our analysis, each algorithm generally performed best at certain ranges of chlorophyll-*a*. The OC4 algorithm, representing blue/green band ratio algorithms in general, performed better when chlorophyll-*a* was less than 10 mg/m3. The Mer-3B algorithm, representing the red/NIR-based algorithms, performed best above this chlorophyll-*a* value. These findings support the view that bio-optical algorithms tuned and developed for specific ranges of conditions perform better than an algorithm tuned to a larger variety of conditions, and a blended algorithm approach is superior to single algorithms when considering the entire dynamic range of environmental conditions.

In this study, we examined only two algorithms, yet a more comprehensive analysis using a wider range of algorithms is needed. However, the framework for the image classification, algorithm selecting and retrieval blending capabilities is flexible and can readily integrate different algorithms that have yet to be tested in the system. In addition, a regional versus global algorithm study should be conducted, as it has been shown that Mer-3B has global application (Gitelson et al., 2011). The capabilities to dynamically identify water types and blend retrievals from multiple algorithms while reducing error are the advantages of the OWT scheme. We have shown how the OWTs can be used in a single image to blend spatially-varying optical environments.

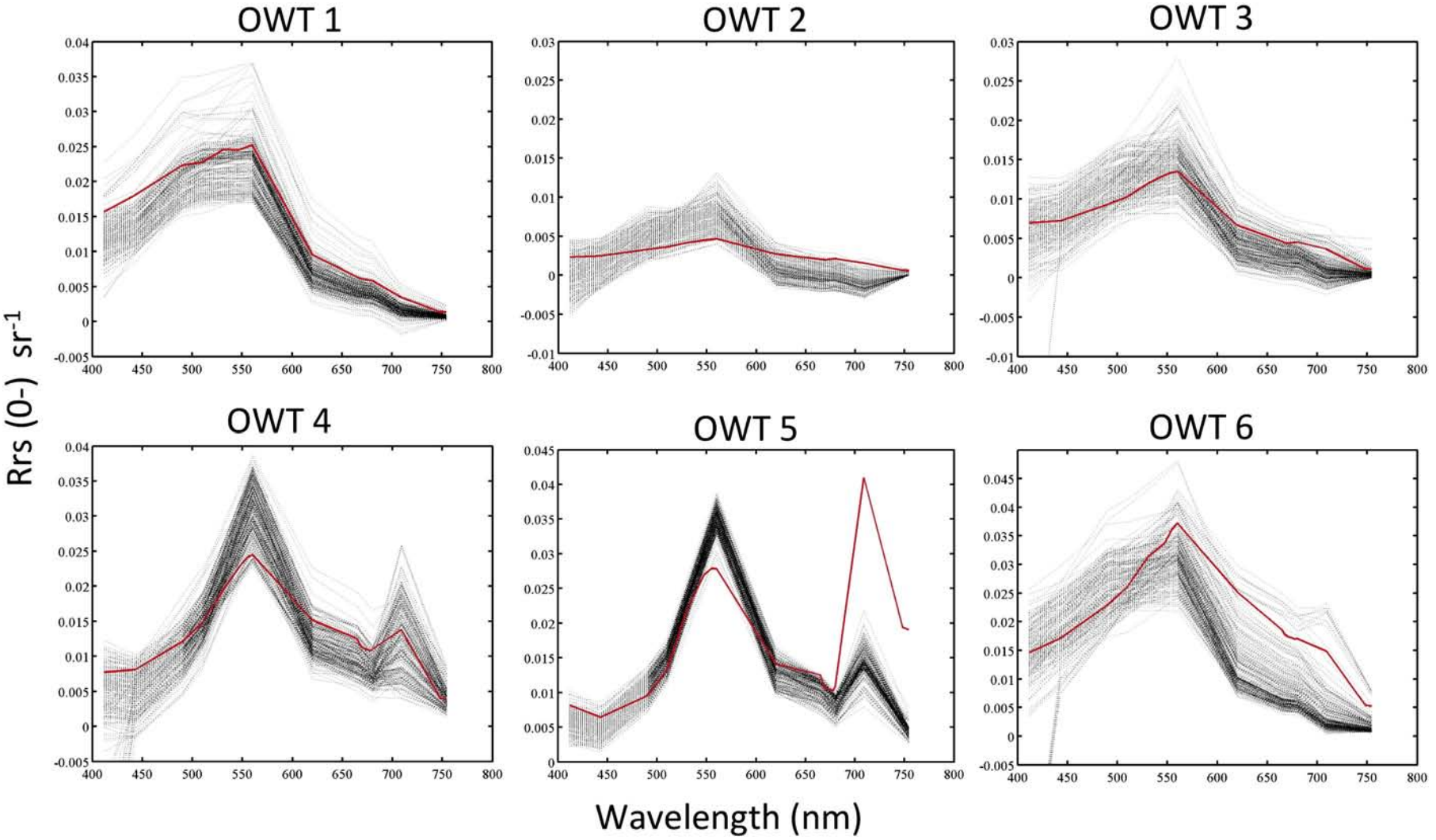

Fig. 10. Extracted satellite pixels (subset of N = 100 for clarity) associated with each OWT (from the maximum membership) for the MERIS image from Sep. 3, 2011 over Lake Erie. Black lines indicate image spectra, and red lines indicate the OWT mean reflectance vector. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

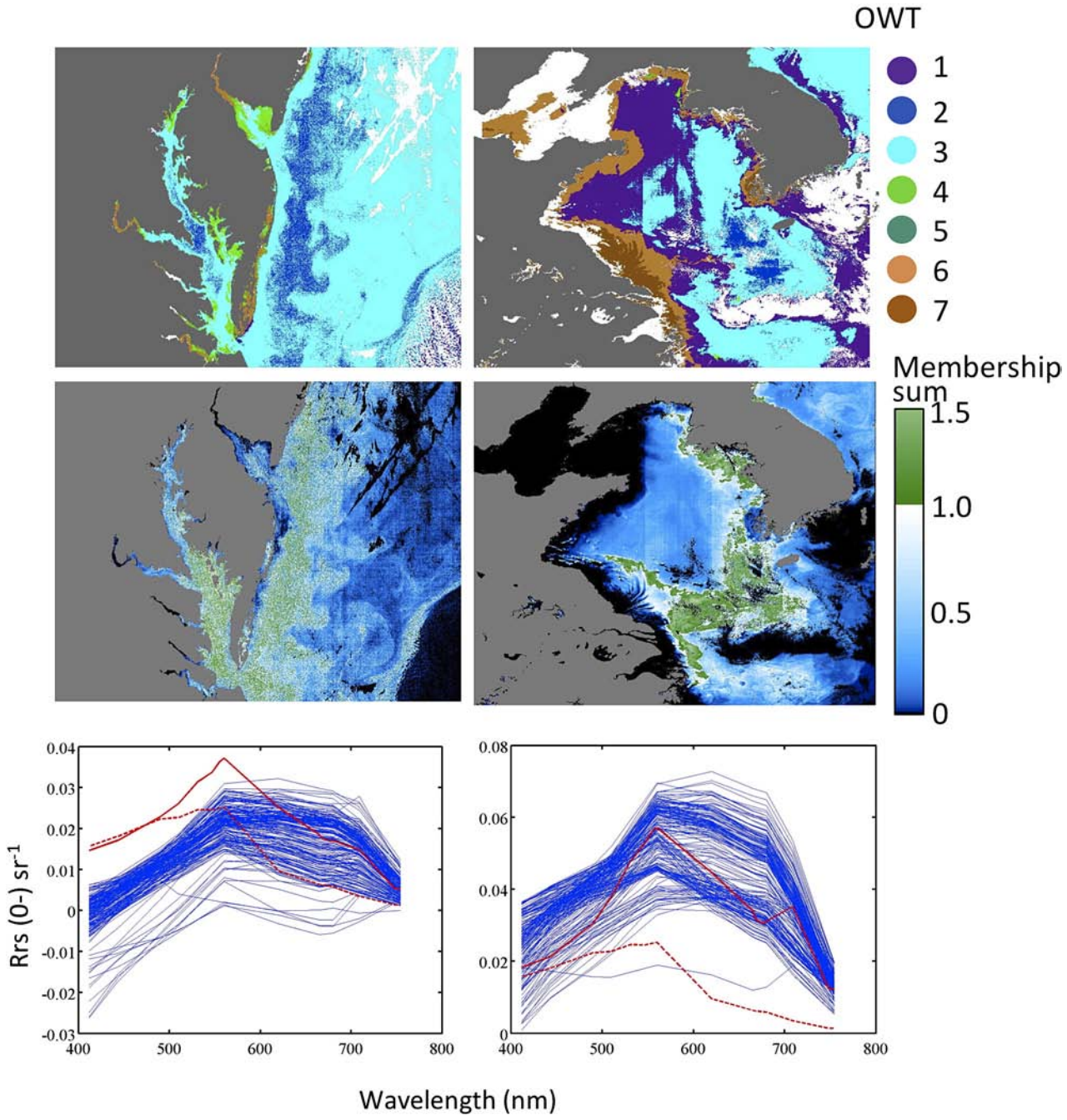

Fig. 11. Left panels: MERIS image from Oct. 14, 2008 over Chesapeake Bay. Top: Hard classification; middle: Membership sum—green areas indicate membership sums greater than 1, dark areas no membership or cloud; bottom: image spectra (blue) from low membership areas associated with OWT 6; solid red line: OWT 6 mean vector, dashed red line: OWT 1 mean vector. Right panels: MERIS image from June 5, 2011 over the Yellow Sea. Top: Hard classification; middle: Membership sum—green areas indicate membership sums greater than 1, dark areas no membership or cloud; bottom: image spectra (blue) from low membership areas associated with OWT 7; solid red line: OWT 7 mean vector, dashed red line: OWT 1 mean vector. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.5. Future impacts

It is possible that a water type could be associated with a unique optical phenomenon such as a coccolithophore bloom (e.g., Moore, Dowell, & Franz, 2012). We recognize that floating algal mats—such as those resulting from cyanobacteria blooms—could also have a unique optical signature. Kutser (2004) used a classification scheme for detecting cyanobacteria blooms based on reflectance spectral characteristics by pattern matching for a region off the coast of Finland. There are other instances of floating or near-surface blooms in the ocean having unique optical properties (Gower, King, & Goncalves, 2008; Subramanium, Brown, Hood, Carpenter, & Capone, 2002). Matthews et al. (2012) have shown that floating algae and algal scum have distinct optical properties. The data set we used in the characterization of the OWTs did contain samples from lakes with high concentrations of floating cyanobacteria at the surface. While we need to assess this further, it is conceivable that under the current method unique optical types could be detected as such provided that the optical signature is represented in one of the OWTs—either existing or as a future addition.

A final consideration is how the lake/coastal OWTs developed here relate to the oceanic OWTs developed from NOMAD (Moore et al., 2009). While the two systems were developed independently from

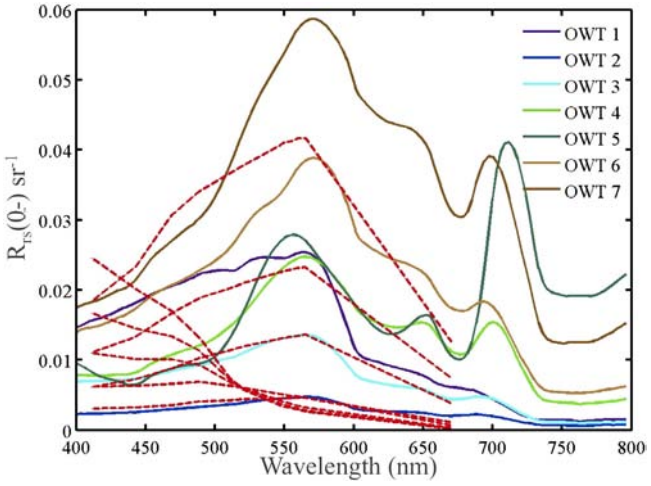

Fig. 12. The reflectance means of the seven optical water types combined with the NOMAD-based optical water types (red dashed lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

each other, there is the question of whether they should be merged. Fig. 12 shows the NOMAD OWTs with the current lake/coastal OWTs. There is some degree of overlap with the existing OWTs. Specifically, lake OWTs 1, 2 and 3 have very similar shape and magnitude to three of the NOMAD-based OWTs and are somewhat redundant, and to a lesser degree OWT 6 is also similar to an existing NOMAD OWT. It also highlights the new water type forms that have emerged from analysis of the present data set. From this view, the feasibility in merging the two systems seems possible. Merging the systems would potentially provide water type coverage from oligotrophic areas of the open ocean to highly eutrophic areas in coastal waters and lakes, and cover greater optical range in environments than either water type system alone.

5. Summary

We have presented an architectural framework aimed at the improvement of remote sensing ocean color products through the integration of bio-optical algorithms tuned over specific ranges of environmental conditions. We have shown the feasibility of such an approach using two chlorophyll-*a* algorithms—one based on blue/green relationships and the other based on relationships in the red/NIR region of the light spectrum. This method is based on the notion of optical water types, and is an extension of the global ocean optical water type (OWT) method of Moore et al. (2009) adapted to coastal/inland waters. Using a weighted blending scheme, the method effectively combines retrievals from multiple algorithms into a seamless product across different optical conditions present in water bodies both spatially and temporally. This method is constructed so as to be applicable to any ocean color satellite, and is a common reference framework that can be tailored to different satellites and algorithms. We expect that other algorithms not considered here could be better in terms of retrieval accuracy for different OWTs and/or different regions. These can be added to the existing framework without changing the base OWT characteristics and configuration. We also foresee the possibility of either adding new OWTs or modifying their current characteristics with the addition of new data. As such, our characterization of the inland lake/coastal OWTs is provisional.

Acknowledgments

This research has been funded by the joint NSF/NIH grant R01 ES021929-02. The authors would like to thank the contributors to SeaBASS who make their data available for open scientific use, to Carsten Brockmann for thoughtful discussion and suggestions on the research, and to the three reviewers who spent time evaluating our manuscript and provided valuable feedback.

References

Ahn, Y. H., Bricaud, A., & Morel, A. (1992). Light backscattering efficiency and related properties of some phytoplankters. *Deep-Sea Research, 39*, 1835–1855.
Bailey, S. W., Franz, B.A., & Werdell, J. P. (2010). Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Optics Express, 18*(7), 7521–7527.
Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms.* New York Plenum.
Bezdek, J. C., Li, W. Q., Attikiouzel, Y., & Windham, M. P. (1997). A geometric approach to cluster validity for normal mixtures. *Soft Computing, 1*, 166–179.
Binding, C. E., Greenberg, T. A., & Bukata, R. P. (2012). An analysis of MODIS-derived algal and mineral turbidity in Lake Erie. *Journal of Great Lakes Research, 38*, 107–116.
Binding, C. E., Jerome, J. H., Bukata, R. P., & Booty, W. G. (2007). Trends in water clarity of the lower Great Lakes from remotely sensed aquatic color. *Journal of Great Lakes Research, 33*, 828–841.
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.
Bradt, S. R. (2012). *Development of bio-optical algorithms to estimate chlorophyll in the Great Salt Lake and New England lakes using in situ hyperspectral measurements.* (PhD Thesis). The University of New Hampshire.
Bridgeman, T. B., Chaffin, J.D., & Filbrun, J. E. (2013). A novel method for tracking western Lake Erie *Microcystis* blooms, 2002–2011. *Journal of Great Lakes Research, 39*, 83–89.
Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research, 100*(C7), 13,133–13,754.
Dall'Olmo, G., & Gitelson, A. A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll-*a* concentration in turbid productive waters: Experimental results. *Applied Optics, 44*(3), 412–422.
Doron, M., Belanger, S., Doxoran, D., & Babin, M. (2011). Spectral variations in the near-infrared ocean reflectance. *Remote Sensing of Environment*, http://dx.doi.org/10.1016/j.rse.2011.01.0105.
Doxoran, D., Cherukuru, R., & Lavender, S. (2006). Use of reflectance band ratios to estimate suspended and dissolved matter concentrations in estuarine waters. *International Journal of Remote Sensing, 26*, 1763–1769.
Feng, H., Campbell, J. W., Dowell, M.D., & Moore, T. S. (2005). Modeling spectral reflectance of optically complex waters using bio-optical measurements from Tokyo Bay. *Remote Sensing of Environment, 99*, 232–243.
Gilerson, A., Zhou, J., Hlaing, S., Ioannou, I., Schalles, J., Gross, B., Moshary, F., & Ahmed, S. (2007). Fluorescence component in the reflectance spectra from coastal waters. *Dependence on water composition, Optics Express, 15*(24), 15702–15722.
Gilerson, A., Gitelson, A., Zhou, J., Gulrin, D., Moses, W., Ioannou, I., & Ahmed, S. (2010). Algorithms for remote sensing of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express, 18*(23), 24109–24125.
Gitelson, A. A., Gurlin, D., Moses, W. J., & Yacobi, Y. Z. (2011). Remote estimation of chlorophyll-*a* concentration in inland, estuarine and coastal waters. In Q. Weng (Ed.), *Advances in Environmental Remote Sensing.* CRC Press.
Gower, J., King, S., Borstad, G., & Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing, 26*, 2005–2012.
Gower, J., King, S., & Goncalves, P. (2008). Global monitoring of plankton blooms using MERIS MCI. *International Journal of Remote Sensing, 29*, 6209–6216.
Goyens, C., Jamet, C., & Schroeder, T. (2013). Estimation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sensing of Environment, 131*, 63–75.
Hu, C., Lee, Z., Ma, R., Yu, K., Li, D., & Shang, S. (2010). Moderate Resolution Imaging Spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China. *Journal of Geophysical Research, 115*.
Hunter, P. D., Tyler, A. N., Carvalho, L., Codd, G. A., & Maberly, S.C. (2010). Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sensing of Environment, 114*, 2705–2718.
Jerlov, N. G. (1951). Optical studies of ocean water. *Reports of the Swedish Deep-Sea Expedition, 3*, 1–59.
Kloiber, S. M., Brezonik, P. L., Olmanson, L. G., & Bauer, M. E. (2002). A procedure for regional lake water clarity assessment using Landsat multispectral imagery. *Remote Sensing of Environment, 82*, 38–47.
Kutser, T. (2004). Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnology and Oceanography, 49*, 2179–2189.
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 classification. *Remote Sensing of Environment, 115*, 725–737.
Lubac, B., & Loisel, H. (2007). Variability and classification of remote sensing reflectance spectra in the eastern English Channel and southern North Sea. *Remote Sensing of Environment, 110*, 45–58.
Maritorena, S., Siegel, D. A., & Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics, 41*, 2705–2714.
Matthews, M. W., Bernard, S., & Robertson, L. (2012). An algorithm for detecting trophic status (chlorophyll-*a*), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment, 124*, 637–652.
Melin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J. -F., et al. (2011). Multi-sensor satellite time series of optical properties and chlorophyll-*a* concentration in the Adriatic Sea. *Progress in Oceanography, 91*, 229–244.

Moore, T. S., Campbell, J. W., & Dowell, M.D. (2009). A class-based approach for characterizing the uncertainty of the MODIS chlorophyll product. *Remote Sensing of Environment*, *113*, 2424–2430.

Moore, T. S., Campbell, J. W., & Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1764–1776.

Moore, T. S., Dowell, M.D., & Franz, B.A. (2012). Detection of coccolithophore blooms in ocean color satellite imagery: A generalized approach for use with multiple sensors. *Remote Sensing of Environment*, *117*, 249–263.

Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, *22*, 709–722.

Mouw, C. B., Chen, H., McKinley, G. A., Effler, S., O'Donnell, D., Perkins, M. G., et al. (2013). Evaluation and optimization of bio-optical inversion algorithms for remote sensing of Lake Superior's optical properties. *Journal of Geophysical Research*, *118*, 1696–1714, http://dx.doi.org/10.1002/jgrc.20139.

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

Odermatt, D., Gitelson, A., Brando, V. E., & Schaepman, M. (2012). Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of the Environment*, *118*, 116–126.

Olmanson, L. G., Brezonik, P. L., & Bauer, M. E. (2013). Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: the Mississippi River and its tributaries in Minnesota. *Remote Sensing of Environment*, *130*, 254–265.

Ruddick, K. G., De Cauwer, V., & Park, Y. -J. (2006). Seaborne measurements of near infrared water-leaving reflectance: the similarity spectrum for turbid waters. *Limnology and Oceanography*, *51*, 1167–1179.

Ruiz-Verdu, A., Simis, S. G. H., de Hoyos, C., Gons, H. J., & Pena-Martinez, R. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sensing of Environment*, *112*, 3996–4008.

Simis, S. G. H., Ruiz-Verdu, A., Dominguez-Gomez, J., Pena-Martinez, R., Peters, S. W. M., & Gons, H. J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment*, *106*, 414–427.

Stumpf, R. P., Arnone, R. A., Gould, R. W., Martinolich, P.M., & Ransibrahmanakul, V. (2003). A partially coupled ocean–atmosphere model retrieval of water-leaving radiance from SeaWiFS in coastal waters. *Algorithm Updates for the Fourth SeaWiFS Data Reprocessing. NASA Technical Memorandum. SeaWiFS Postlaunch Technical Report Series, vol. 22*. (pp. 2003–206892).

Subramanium, A., Brown, C. W., Hood, R. R., Carpenter, E. J., & Capone, D.G. (2002). Detecting Trichodesmium blooms in SeaWiFS imagery. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *49*, 107–121.

Szeto, M., Campbell, J. W., Moore, T. S., & Werdell, P. J. (2011). Are the world's oceans optically different? *Journal of Geophysical Research*, http://dx.doi.org/10.1029/2011JC007230.

Vantrepotte, V., Loisel, H., Dessailly, D., & Meriaux, X. (2012). Optical classification of contrasted coastal waters. *Remote Sensing of Environment*, *123*, 306–323.

Wang, M., Son, S., & Shi, W. (2009). Evaluation of MODIS SWIR and NIR-SWIR atmospheric correction algorithms using SeaBASS data. *Remote Sensing of Environment*, *113*, 635–644.

Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., et al. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Trans AGU*, *84*, 38,377.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., & Dyble, J. (2010). Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meterological data. *Limnology and Oceanography*, *55*, 2025–2036.

Yacobi, Y. Z., Moses, W. J., Kaganovsky, S., Sulimani, B., Leavitt, B. C., & Gitelson, A. A. (2011). NIR-red reflectance-based algorithms for chlorophyll-*a* estimation in mesotrophic inland and coastal waters: Lake Kinneret case study. *Water Research*, *45*(7), 2428–2436.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zibordi, G., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Melin, F., et al. (2009). AERONET-OC: A network for the validation of ocean color primary radiometric products. *Journal of Atmospheric and Oceanic Technology*, *26*, 1634–1651.

Zimba, P. V., & Gitelson, A. (2006). Remote estimation of chlorophyll concentration in hyper-eutrophic aquatic systems: Model tuning and accuracy optimization. *Aquaculture*, *256*, 272–286.