Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters

Patrick L. Brezonik a,⁎, Leif G. Olmanson b, Jacques C. Finlay c, Marvin E. Bauer b

a Department of Civil, Environmental, and Geo-Engineering, University of Minnesota, Minneapolis, MN 55455, USA
b Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA
c Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108, USA

A R T I C L E   I N F O

Article history:
Received 7 January 2014
Received in revised form 25 March 2014
Accepted 30 April 2014
Available online 5 July 2014

Keywords:
CDOM
DOC
Chlorophyll
Hyperspectral imagery
Mississippi River
Lake Superior
Remote sensing
Landsat
Sentinel
Lakes
Rivers
Turbidity
Suspended sediment
Water quality

A B S T R A C T

A combination of new measurements and analysis of historical data from several geographic regions was used to address four issues that affect the reliability and interpretation of colored dissolved organic matter (CDOM) measured by remote sensing of inland waters. First, high variability of CDOM levels in lakes and rivers was found at seasonal and multi-year time scales and at shorter intervals in some rivers and lakes. Coefficients of variation (CVs) of 30%–50% for absorptivity at 440 nm (a440) were common in historical and new data sets we examined. CDOM values used to calibrate imagery thus should be measured close to the image acquisition date, preferably within 1–2 months in lakes and a few days in large rivers, unless it can be shown that CDOM levels are temporarily stable over longer (or shorter) time periods in a given aquatic system. Second, spectral slopes (S) for CDOM in the visible range vary little over time (even over multi-year periods) within sites. Substantial variation was found between sites, however, and most spectra showed a change in slope near 460 nm. Values of S400–460 for waters with moderate to high CDOM levels generally were within a narrow range (−0.014–0.018) and similar to reported S values in the near UV. Values of S400–460 for waters with low CDOM generally were smaller and more variable, as were values for S600–650 for all waters. Overall, the variability of spectral slopes in the visible range should not have a large effect on the reliability of a440 estimates made from remote sensing, which in many models involve reflectance measurements at wavelengths >500 nm. Third, although a strong correlation (r² = 0.925) was found between CDOM levels and DOC concentrations in 34 surface waters sampled in 2013, the standard error of estimate suggests an uncertainty of ~ ±20% in predicting DOC at a440 =5 m−1 (a moderate CDOM level). Moreover, CDOM–DOC relationships for unpublished data sets we analyzed and those reported in the literature indicate that both the fraction of DOC that is colored and slopes of regressions between CDOM and DOC are highly variable in space and time. Prediction of DOC concentrations in water bodies from CDOM levels (whether measured in the laboratory or by remote sensing) thus is associated with considerable uncertainty. For the present, this implies that field sampling is required to verify DOC concentrations predicted from remotely sensed CDOM measurements until we have a better understanding of variations in DOC–CDOM relationships. Fourth, shapes of reflectance spectra for CDOM-rich waters varied greatly depending on the concentrations of other constituents (suspended solids and chlorophyll) that affect the optical properties of water. Nonetheless, it is not obvious from our results for several predictive models that different remote sensing algorithms are needed to calculate CDOM levels accurately for waters where CDOM is the only variable affecting reflectance versus waters where other constituents also affect the spectra. The best band or ratio models for simulated Landsat 8, Sentinel-2 and Sentinel-3 bands from field measured reflectance spectra yielded high r² values (0.84–0.86) for a440. The broader Landsat 8 bands worked nearly as well for a440 as the narrower Sentinel band sets and hyperspectral bands, probably because CDOM is characterized by a broad exponential increase in absorbance with decreasing wavelength rather than specific peaks or troughs in absorbance or reflectance.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Colored dissolved organic matter (CDOM) is considered to be the major component of the dissolved natural organic matter (DNOM) in natural waters. CDOM also has many important effects on aquatic ecology and chemistry, as well as on the suitability of surface waters for human use. Understanding the distribution and dynamics of CDOM...
thus is broadly important to monitoring, assessment and management of surface waters, and remote sensing is especially attractive for monitoring purposes because of its ability to make measurements at regional and even larger scales.

Many algorithms have been proposed to retrieve CDOM values from remotely sensed imagery; empirical, semi-analytical, matrix inversion and optimization methods all have been proposed. Zhu et al. (2014) recently compared 15 of these methods using data from Saginaw Bay in Lake Huron, and they found that most methods do not provide reliable estimates of CDOM levels. Even those that provided the best estimates tended to yield underestimates at high CDOM levels and overestimates at low CDOM. In spite of the recent spate of publications on CDOM measurement by remote sensing, many questions remain unanswered, and procedures to measure CDOM accurately across a wide range of water quality conditions remain an elusive goal.

Development of reliable methods to retrieve CDOM information from spectral reflectance data is difficult. Indeed, among the major water quality variables measurable by remote sensing (e.g., suspended solids, chlorophyll, Secchi depth), for several reasons CDOM may be the most difficult to measure accurately in inland waters. First, CDOM absorbs but does not scatter or reflect light. Hence the upwelling (water-leafing) light signal is very small in CDOM-rich waters. Second, CDOM has no absorbance troughs or peaks, such as are found for plant pigments; instead light absorption by CDOM follows a simple quasi-exponential decrease with increasing wavelength. There are no wavelength bands in the visible spectrum uniquely associated with CDOM that can be used for measurement purposes. In addition, the exponential decline in absorbance with increasing wavelength means that its effects on reflectance spectra are higher in the blue region, where atmospheric correction is more difficult (Kutser, Pierson, Kallio, Reinart, & Sobek, 2005) than in the green and red regions generally preferred to measure water quality variables in freshwaters (Gilerson et al., 2010; Gitelson, 1992; Moses et al., 2011). Third, measurement of low to moderate levels of CDOM in optically complex waters (those containing mineral suspended solids, algae and associated organic particles) is especially difficult because light scattering by these particles dominates their reflectance spectra.

Aside from issues related to the direct effects of CDOM and other water quality constituents on reflectance spectra, several other CDOM-related issues must be resolved to achieve reliable remote sensing protocols for CDOM across a wide range of inland water conditions. This paper focuses on four important questions:

1. How variable are CDOM levels temporally (intra-seasonally to inter-annually) in a water body?
2. How do the spectral properties of CDOM vary in the visible range (~400–700 nm) among surface waters and over time in a given water body?
3. How constant is the relationship between CDOM measured as absorbance at 440 nm ($a_{440}$) and DOC measured as mg per liter?
4. How do reflectance spectra of inland waters vary as a function of CDOM levels and concentrations of other light-absorbing and scattering constituents, such as chlorophyll and suspended particles, and does the wavelength combination used to retrieve CDOM from reflectance spectra depend on whether CDOM is the predominant property affecting a water body’s inherent optical properties (IOPs) or constituents like suspended solids and chlorophyll also are important?

We address these questions by (i) analysis of historical and unpublished data of others and collection of new data, including a sampling program on waters with diverse optical properties, (ii) measurement of ground-level detailed reflectance spectra on these waters, (iii) use of these spectra to simulate Sentinel 2 and 3 and Landsat 8 sensors, and (iv) evaluation of new and published models to retrieve CDOM values from reflectance data.

2. Background

2.1. Why are these questions important?

Variation in CDOM levels within and across ecosystems (question 1) is important relative to whether contemporaneous ground-based samples are needed to calibrate predictive algorithms. CDOM generally is thought to be a relatively stable property of lakes (probably less so in flowing waters), but whether this stability allows use of ground-based samples collected within a month, six months, a year, or even longer is not well known. Variation in the visible spectral properties of CDOM (question 2) is important because remote sensing techniques to measure CDOM differ from laboratory and in situ methods. Most empirical remote sensing methods used for inland waters are based on reflectance at wavelengths > 500 nm (e.g., Kutser et al., 2005; Menken, Brezonik, & Bauer, 2006), but laboratory methods are based on absorbance at ~410–440 nm, and in situ probes for CDOM (e.g., Chen, 1999; Gardner, Chen, & Berry, 2005) are based on fluorescence at ~450 nm (excitation wavelength of ~350 nm). The extent to which CDOM absorbance spectra vary among waters could affect the reliability of estimating CDOM in terms of $a_{440}$ from reflectance values measured at longer wavelengths. The constancy of the relationship between CDOM and DOC (question 3) is important relative to the use of CDOM as a surrogate for DOC, which is a goal of biogeochemists interested in quantifying the role of freshwater DOC in the global carbon cycle. Remote sensing is attractive because it can measure CDOM at regional scales, but to be useful in global carbon-cycle models, the data need to be convertible to mass or molar units of carbon. The geographic extent to which DOC–CDOM relationships reported in the literature apply is not well understood. Finally with regard to question 4, many bands or wavelength combinations have been used to estimate CDOM by remote sensing. Relationships using wavelengths ($\lambda > 500$ nm) are common (see Section 2.3) despite the fact that CDOM absorbance is highest at low $\lambda$ and absorptivity is low at $\lambda > 500$ nm. Studies are needed to constrain the conditions under which published algorithms can be used reliably, to determine why the algorithms work, and to evaluate whether other algorithms might be better.

2.2. Nature and importance of CDOM

By definition, CDOM is the portion of the dissolved natural organic matter in natural waters that absorbs visible light. It has two sources: (1) allochthonous—derived from decomposition of woody plants in terrestrial environments, and (2) autochthonous—derived from decomposition of algae and aquatic vegetation within a water body. In general, the former is more highly colored than the latter. Both CDOM types commonly are called “aquatic humic matter,” and they consist mostly of fulvic acids (Aiken, 1985). Chemically they are mixtures (many thousands of structures in any given water body) of moderately large molecules with an average molecular weight around 400–600 and a broad distribution of molecular weights (200–800, or more). CDOM molecules have aromatic, phenolic, and acidic functional groups (for a review, see Brezonik & Arnold, 2011).

CDOM affects pH and alkalinity, forms chemical complexes with metals (with implications for their bioavailability and toxicity), and acts as a transport agent for metals and nonpolar organic contaminants. It also plays an important role in aquatic photochemistry, serving as a source of reactive radicals. Photochemical processes also transform CDOM, e.g., affecting its availability as a substrate for microorganisms. CDOM effects on aquatic ecology include: (1) the chemical effects described above, (2) decreasing light penetration (reducing the depth of water with enough light for net primary production), (3) changes in the thermal properties of water bodies (Fee, Hecky, Kasian, & Cruikshank, 1996; Houwer, 2006), and (4) changes in the spectral distribution of light, which may affect the composition of aquatic plant and phytoplankton communities. Finally, CDOM is important in human use of water. It exerts a chlorine demand, making it more costly to
disinfect water and is thought to be the primary component of DNOM serving as the precursor for toxic disinfection byproducts like chloroform (Bellar, Lichtenberg, & Kroner, 1974; Minear & Amy, 1995; Rook, 1974).

Although CDOM is considered to be a relatively stable property of natural waters, few studies have focused on this topic. Pace and Cole (2002) showed that summer levels of CDOM in 20 northern Michigan lakes varied substantially but synchronously over a six year period, and they related the variations to climatic conditions affecting inputs of allochthonous organic matter from watersheds. Erm et al. (2002) found only small variations in CDOM on a time scale of months during summer in some Estonian lakes. Kutser et al. (2005) used the above information to limit the time between image acquisition and field data collection to 4 weeks in developing a predictive algorithm for CDOM. Several studies (Cotner, Biddanda, Makino, & Stets, 2004; Pace, Cole, 2002; Shank, Zepp, Väheatalo, Lee, & Barrels, 2010; Zafiriou, 2002; Zhang et al., 2013) have shown seasonal decreases in CDOM levels and linked the losses to photobleaching (i.e., the loss of chromophoric functional groups in DNOM caused by photochemical reactions). Climate can alter CDOM levels by changing the amount of photobleaching and altering the sources of DOC to lakes (Pace & Cole, 2002). Large temporal variations in CDOM levels can occur in flowing waters in response to changes in flow paths through watersheds. For example, Brinkman and Hozalski (2011) found large variations during 2007 in DOC and water absorbance at 254 nm ($A_{254}$), an indirect measure of DOC and CDOM, in the Mississippi River near Minneapolis, Minnesota. Highest DOC values were associated with spring runoff and autumn rains, but $A_{254}$ peaked only in fall.

The spectral properties of CDOM have been studied extensively in the UV and blue region of visible light (up to ~450 nm). Plots of the natural logarithm of absorbance versus wavelength ($\lambda$) are quasi-linear, but the slopes of such plots, called the “spectral slope,” $S$, depend on the range of wavelengths used. Values of $S$ for various wavelength segments have been shown to be useful in estimating some chemical properties of DNOM components, such as the ratio of fulvic acid to humic acid and the mean molecular weight of fulvic acid (Carter, Chen, Lee, Hawes, & Kamyszowski, 1999) and CDOM (Helms et al., 2008). In contrast, little has been reported on the spectral properties of freshwater CDOM at longer wavelengths. Kirk (1994) reported absorptivity ($a_\lambda$) spectra of filtered water over the range 350–700 nm for ten Australian water bodies, and most spectra showed small but measurable values of $a$ at $\lambda$ ~ 600 nm [highest value for $a_{600} = -0.8$ m$^{-1}$]. All the spectra appeared to show an exponential decrease in $a_\lambda$ with increasing $\lambda$, but quantitative results (e.g., plots of ln($a$) versus $\lambda$ or values of $S$) were not given.

Many studies (e.g., Blough, Zafiriou, & Bonilla, 1993; Gorham et al., 1998; Spencer, Butler, & Aiken, 2012; Tranvik, 1990) have reported good correlations between DOC (measured in mg or mmol per liter) and CDOM (measured either in chlorophlate units or as $a_\lambda$, where $\lambda$ is in the range ~400–440 nm). The geographic extent to which such relationships apply outside the areas in which data were collected for a given relationship is uncertain, however. No universal relationship can exist between CDOM and DOC because algal-derived DOC has lower color intensity than DOC from the decomposition of woody vegetation (humic-derived DOC), and DOC from human sources (e.g., wastewater effluent) is nearly uncolored. Many water bodies with high CDOM levels have low levels of phytoplankton because CDOM absorbs sunlight and decreases the euphotic zone. Such lakes may exhibit good correlations between CDOM and DOC.

2.3. Measurement of CDOM by remote sensing

Several efforts to retrieve CDOM information from satellite imagery on marine systems have used semi-analytical and matrix inversion techniques (e.g., Brando, Dekker, Park, & Schroeder, 2012; Carder et al., 1999; Hoge & Lyon, 1996; Lee, Carder, & Arnone, 2002). Other workers have used empirical (regression) approaches to retrieve CDOM data on marine and coastal waters from satellite imagery (e.g.; Del Castillo & Miller, 2008; D’Sa & Miller, 2003; Mannino, Russ, & Hooker, 2008). CDOM levels are much lower in marine waters than those of interest in freshwaters; $a_{400}$ generally is < 1 m$^{-1}$ in coastal waters and < 0.1 m$^{-1}$ in open ocean waters. (Marine scientists report CDOM as absorptivity at 412 nm rather than 420 or 440 nm, which is commonly used by freshwater scientists.) Such values usually would be regarded by limnologists as indicative of negligible CDOM in freshwaters, in which $a_{400}$ may range up to ~40 m$^{-1}$ or more. Concentrations of other constituents (e.g., suspended solids and chlorophyll) affecting the intrinsic optical properties (IOPs) of water are also much higher in many freshwaters than in the oceans, thus rendering freshwater systems very optically complex. Because of these differences, methods derived for marine systems are not likely to be directly applicable to freshwaters, and we focus here on the latter systems.

Kutser, Herlevi, Kallio, and Arst (2001) described a semi-analytical model to retrieve values of CDOM, chlorophyll, and suspended matter from hyperspectral data on lakes collected with an airborne spectrometer, and Zhu, Yu, Tian, Chen, and Gardner (2011) used a semi-analytical inversion technique to retrieve separate wavelength-dependent values of the absorption coefficients, $\epsilon_{CDOM}$ and $\epsilon_{NAP}$ (NAP = nonalgal particles), from hyperspectral data gathered by ship on the Mississippi and Atchafalaya Rivers. Previous semi-analytical models typically combined $\epsilon_{CDOM}$ and $\epsilon_{NAP}$ into one coefficient ($\epsilon_{CDOM+NAP}$) because CDOM and NAP have similar spectral shapes and slopes.

Most previous studies to estimate CDOM in freshwaters, however, used empirical (regression) methods to analyze reflectance data from ground-, aircraft- or satellite-based spectrometers, and most published relationships involve nonlinear (power) relationships and wavelengths > 500 nm (Ficek, Zapadka, & Dera, 2011; Griffin, Frey, Rogan, & Holmes, 2011; Kutser, 2012; Kutser, Tranvik, & Pierson, 2009; Kutser et al., 2005; Menken et al., 2006), generally with moderate $r^2$ values (~0.6–0.75). Why $R_{670}/R_{571}$ worked best for Menken et al. (2006), for example, and why the ratio of Advanced Land Imager (ALI) band 2 (525–605 nm) to band 3 (630–690 nm), $R_2/R_3$, worked best for Kutser et al. (2005, 2009) are not well understood. The nonlinearity of the relationships also increases prediction uncertainties at high CDOM levels. Several have questioned why longer wavelengths work best (Kallio et al., 2001; Menken et al., 2006). According to Witte et al. (1982), nonlinearity in CDOM-reflectance relationships can be explained by the governing bio-optical equations. The presence of wave-lengths near 670 nm (including Landsat Thematic Mapper TM 3 and ALI band 3) in many relationships probably is related to correction for chlorophyll effects on reflectance values.

Menken et al. (2006) examined relationships at $\lambda$ < 450 nm, where CDOM absorbance is high. Interference from chlorophyll led to poor correlations when they included all 15 of their lakes, including some highly eutrophic ones, but stronger relationships were found for lakes with chlorophyll $a$ < 10 µg/L and when a term was added to correct for chl $a$. Kutser et al. (2005) noted that atmospheric correction is difficult in the blue region (<500 nm), especially for lakes with low reflectance, and concluded that use of green bands (e.g., Landsat TM and ALI bands 2) was preferable. Nonetheless, Brezonik, Menken, and Bauer (2005) found a good relationship ($R^2 = 0.77$) using Landsat TM1 and TM1:TM4 to fit CDOM data on 15 Minnesota lakes with wide ranges of CDOM and chlorophyll, and Griffin et al. (2011) reported an $R^2$ of 0.78 for a regression of data from a Siberian river involving ln($a_{400}$) and Landsat TM or ETM + bands 1–3.

3. Methods

3.1. Data sources

The temporal variability of CDOM in surface waters was evaluated by examining unpublished data for U.S. water bodies in a cold-temperate region, the Upper Midwest (Wisconsin and Minnesota) and
a warm-temperate/subtropical region (Florida). Data sources included (i) two multi-year surveys of lakes in Florida in the 1960s and 1970s led by the senior author, (ii) a long-term record (1989 to present) for seven lakes in north-central Wisconsin from the University of Wisconsin’s Long-Term Ecological Research (LTER) program, (iii) a one-year record of CDOM and DOC in the Mississippi River near Minneapolis, Minnesota (Brinkman & Hozalski, 2011), (iv) a lake water quality database of the Minnesota Pollution Control Agency (MPCA), and (v) an ongoing study of the St. Louis River Estuary (SLRE) near Duluth, Minnesota and Superior, Wisconsin led by co-author JCF.

The first Florida survey was conducted in 1968–70 to evaluate limnological characteristics of Florida lakes (Shannon & Brezonik, 1972a) and factors affecting trophic state conditions (Shannon & Brezonik, 1972b). Fifty-five lakes were sampled 5–7 times from late 1968 to mid-1970; a few were sampled more frequently or over longer time periods as part of other studies (e.g., Lake Mize: 21 times from 1966 to 1970; Keirn & Brezonik, 1971). The second survey was conducted in 1977–78 to evaluate potential effects of acidic deposition on the chemistry and biology of Florida lakes (Brezonik, Crisman, & Schulze, 1984; Hendry & Brezonik, 1984). Twenty acid-sensitive (low alkalinity) lakes in north- and south-central Florida were sampled quarterly. Most of the lakes had low CDOM levels ($A_{440} < 2.0 \text{ m}^{-1}$), but three had moderate levels. Color data from the two surveys were published only in highly summarized form, and temporal variations in CDOM were not examined until the present study. Color was measured in both surveys as absorbance at 420 nm in a 1-cm cell, compared to a standard curve prepared from chloroplatinate standard solutions, and recorded as color in standard chloroplatinate units (CPU; e.g., Eaton, Clesceri, Rice, & Greenberg, 2005). For this study the data were transformed to absorbance at 420 nm ($A_{420}$) using the measured absorbance of a chloroplatinate standard and then converted absorptivity at 420 ($a_{420}$) by the relationship:

$$a_{420} = 2.303 A_{420}/\ell \text{ (m}^{-1})$$

where $\ell$ is the cell path length (in m). Values of $a_{420}$ were converted to the common measure of CDOM, $A_{440}$, using a spectral slope of 0.014.

The Wisconsin LTER data were obtained from http://lter.limnology.wisc.edu. With occasional exceptions, samples have been collected four times per year since fall 1989 corresponding to spring (after ice-out—April or May), summer (August), autumn (October or November) and winter (under-ice conditions—January or February) from two high-colored bog lakes, one colored drainage lake, and four low-CDOM seepage lakes in Vilas County, north-central Wisconsin; Winslow (2012) recently analyzed the data, with emphasis on the UV part of the spectra. Absorbance of filtered water is measured using 5 or 10-cm quartz cells over the range 200–800 nm. We downloaded absorbance at 440 nm for the lakes over the period of record and converted the data to $A_{440}$ using Eq. (1). We also downloaded absorbance spectra in the visible range for selected lakes and time periods to analyze the variability in spectral slopes, $S$.

Brinkman and Hozalski (2011) compiled a record of CDOM and DOC measurements on the Mississippi River near Minneapolis, MN in terms of daily-averaged absorbance values at 254 nm over most of calendar year 2009. Absorbance values over the range 250–500 nm for spring summer and fall samples from the same site (obtained from B. Brinkman, pers. comm., 2013) were converted to In($A$), and regressions of In($A$) versus $\lambda$ yielded identical slopes, $S = 0.018$. We used this value to convert $A_{254}$ to $A_{440}$. Although some uncertainty is likely in such conversions, our primary purpose was to examine variations in $A_{440}$ over time rather than compare absolute values of the results to directly measured values of $A_{440}$. For this purpose we think the errors in extrapolation are small.

The MPCA and MN Department of Natural Resources monitor lakes across the state as part of a long-term program called SLICE. We obtained a data summary for the program containing 511 lake records (Monson, B., MPCA, pers. comm., 2013) and used it to analyze CDOM–DOC relationships. CDOM is measured in SLICE by visual comparison of filtered water to chloroplatinate standards and reported in CPU; color measured by the visual method corresponds to that measured spectrophotometrically at 440 nm using chloroplatinate for calibration (Cuthbert & del Giorgio, 1992). We converted the data to $A_{440}$ using Eq. (1) and the absorbance at 440 nm of a chloroplatinate standard.

Data for the St. Louis River Estuary were collected as part of an ongoing project examining nutrient transport and transformations in the highly modified freshwater estuary between Duluth, MN and Superior, WI. Water in the estuary comes from a combination of riverine inputs, seiches from Lake Superior, and adjacent urban and industrial runoff. Water samples for CDOM and other analyses presented in this paper were collected seven times from February 28 to September 19, 2013, at 6–30 stations along the lower St. Louis River and Superior and Allouez Bays.

3.2. Study sites

In summer 2013 we sampled 35 surface water sites (lakes and rivers) as part of a campaign to measure their optically-related water quality characteristics simultaneously with the measurement of reflectance spectra by a ground-based hyperspectral radiometer system and nearly contemporaneously with imagery from several satellites (Landsat 7 and 8, HICO on the International Space Station). Imagery from these sensors will be evaluated in a subsequent manuscript. Sites were selected to obtain a wide range of CDOM levels in CDOM-dominated systems and in systems where chlorophyll (algae) and inorganic suspended matter also affect reflectance. Sites in Minnesota were located in: (i) the Minneapolis–St. Paul metropolitan area, (ii) the Sturgeon Lakes region near Side Lake (St. Louis County), (iii) Big Sandy lake chain (Aitkin County), and (iv) the St. Louis River Estuary near Duluth, MN and Superior, WI. A few sites also were sampled in north-central Wisconsin. Table 1 summarizes locations and other limnological information on the sites.

3.3. Field and laboratory water quality measurements

At each site, we collected near-surface samples to measure CDOM, chlorophyll $a$, suspended solids, and dissolved organic carbon (DOC). Secchi depth (SD) also was measured. In addition, several background limnological variables (dissolved oxygen and temperature profiles, pH, conductivity) were obtained at most sites. Sampling methods and laboratory analytical procedures followed standard limnological and water quality practice; see Tsui and Finlay (2011) for details. Absorbance spectra of filtered water samples (0.45 μm GF) were measured at 1 nm intervals over the range 250–700 nm using 1 (250–400 nm) and 5 cm (375–700 nm) quartz cells in a dual beam Shimadzu UV-1601PC spectrophotometer. CDOM was determined from the absorptivity calculated at 440 nm, and spectral slopes were calculated over various ranges on In-transformed absorbance values. SUVA$_{254}$ was calculated according to U.S. EPA Method 415.3 (Potter & Wimsatt, 2003) as $100 \times A_{254}/[DOC]$, where $A_{254}$ is the sample absorbance in a 1 cm cell at 254 nm, [DOC] is the concentration of dissolved organic carbon in mg/L, and 100 is a unit conversion factor to express the results in L/mg·m (or equivalently in m$^2$/g).

Descriptive statistics (means, standard deviation (std. dev.), coefficient of variation (CV)) were calculated in Excel 2010; regression analyses were done in Excel 2010, Systat 12.0, or JMP Pro 10. Box plots and cluster analyses were computed in Systat 12.0.

3.4. Hyperspectral reflectance measurements

The hardware and software system used to collect hyperspectral reflectance data was developed by the Center for Advanced Land
Management Information Technologies (CALMIT), University of Nebraska-Lincoln. The dual system consists of two Ocean Optics USB-2000+ spectroradiometers with attached fiber-optic cables, a Spectralon calibration panel, and laptop computer with software to record and process data. The end of one fiber-optic cable, which is pointed toward the sky on top of a 4 m pole, has a diffuser tip to measure downwelling sunlight. The other cable, attached to a thin metal tube pointed toward the sky on top of a 4 m pole, has a diffuser tip to measure reflected light) above or just below the water surface away from the boat with minimal shading. The software, CDAP-2, calculates spectral reflectance by dividing upwelling light by downwelling light in each narrow (~0.4 nm) band over the range ~370–900 nm. The configuration allows simultaneous measurement of incoming and reflected radiation, thus making calculated reflectance values considerably less dependent on illumination and weather conditions than when upwelling and downwelling irradiance are measured sequentially (Chipman, Olmanson, & Gitelson, 2009).

Reflectance measurements were obtained at 32 of the 35 sites (see Table 1) by first collecting 3–4 calibration scans of the Spectralon panel and then collecting 3–4 scans just below the water surface. Each scan consists of 25 measurements that are averaged by the software, and median values from the calibration scans are used to compute reflectance values for each scan. In processing the data, the individual scans were filtered for outliers and averaged for more robust reflectance spectra. To smooth out noise we calculated running averages using 5–9 individual reflectance measurements representing a range of ~1–3 nm. The final spectra were created by averaging the consistent below-water spectra; we did not include a few below-water spectra in the average that were notably different from the rest of the spectra at a few sites.

Processing of the spectral data to correct for the immersion factor and produce the remote sensing reflectance at nadir, Rrs(λ) = Rrs−1(λ) (sr−1), was done using protocols described by Gitelson et al. (2008); Gitelson, Gurlin, Moses, and Barrow (2009), Chipman et al. (2009), and Yacobi et al. (2011). The reflectance data also were used to simulate spectral band sets of current and planned regional satellite systems, including the Operational Land Imager (OLI) sensor on Landsat 8, the Multi-Sensor Instrument (MSI) sensor on the Sentinel-2 (anticipated launch in 2014), and the Ocean and Land Colour Instrument (OLCI) on Sentinel-3 (likely to be launched 1–2 years after Sentinel-2). Simulations were accomplished by calculating the average values from the reflectance spectra for the range of each band.

### 4. Results and discussion

#### 4.1. Temporal variability of CDOM

Although CDOM levels were stable in several Florida lakes for much of the 1968–70 sampling period (e.g., Lake Suggs, Lake Ten; Fig. 1a,b),
large changes occurred between some relatively short sampling intervals, and some lakes (e.g., Lake Hawthorne and Mize; Fig. 1c,d) showed high temporal variability throughout the study period. Overall, the average coefficient of variation (CV) of $a_{440}$ for 22 lakes with mean $a_{440} > 5$ m$^{-1}$ (range = 5.1–27.1 m$^{-1}$) was 50.5% (range = 18%–96%). A similar CV (45%) was found for ten additional lakes with lower CDOM levels (range of mean $a_{440}$ = 2–5 m$^{-1}$). There was no trend in CV related to mean CDOM concentration, and seasonal patterns were not apparent in the CDOM levels. The large variations in CDOM likely were caused by variations in rainfall and runoff, but the lakes are hydrologically complicated and rainfall is spatially patchy in northern Florida. Consequently, the data are not sufficient to discern direct relationships between rainfall amounts and CDOM levels.

Only three lakes in the 1978–79 Florida lake survey had moderate CDOM (range of mean $a_{440}$ = 3.5–9.4 m$^{-1}$); CDOM in the other lakes was low (range of $a_{440}$ = 0.25–1.54 m$^{-1}$; grand average = 0.81 m$^{-1}$). Box plots of data for the two categories of lakes for each sampling period (Fig. 2) were used to evaluate whether CDOM levels varied by season. All 20 lakes are in landscapes with porous, sandy soils; they lack surface stream drainage and their water is derived directly from atmospheric precipitation and subsurface seepage. As a result, water residence times for the lakes are relatively long (typically several years), and CDOM levels,
especially in the low-CDOM lakes, likely were driven more by “in-lake” processes than by inputs from the surrounding landscape. The median value of $a_{440}$ for the low-color lakes (Fig. 2a) varied only slightly by season, and seasonal means are not significantly different (one-way ANOVA, $p = 0.64$). Nonetheless, temporal variations within individual lakes were moderately high; coefficients of variation ranged from 9% to 74% (mean of 31%). For the three lakes with moderate to high color, CVs for the seasonal data ranged from 8% to 34%. Although the median $a_{440}$ and interquartile values for spring shown in Fig. 2b are lower than those for the other seasons, high variability of the within-season data resulted in no significant difference in seasonal mean $a_{440}$ values (one-way ANOVA, $p = 0.86$). The strength of this conclusion, however, is limited by the small number of lakes in this category.

CDOM variability in the three colored Wisconsin LTER lakes was generally high over the 23-year record (Fig. 3). CVs for the highly colored bog lakes were 26% (Crystal) and 34% (Trout); for moderately colored Allequash Lake, CV was 51%. A long-term trend of increasing CDOM was found in Crystal Bog, especially since the mid-1990s (Fig. 3c); mean $a_{440}$ increased from $-4 \text{ m}^{-1}$ to $-8 \text{ m}^{-1}$ over the period of record. Similar trends were not found in the other six lakes, but episodic spikes in $a_{440}$ (3–4 times higher than the long-term average) were apparent in Allequash Lake during the period 2002–2011 (Fig. 3a) and in low-color Big Muskellunge Lake (Fig. 3b) during 2003–2009. Occasional (but less frequent) spikes in $a_{440}$ were observed in long-term records for three other low-CDOM LTER lakes (Crystal, Sparkling, Trout) during the period 1999–2008 (data not shown). Although some spikes occurred concurrently in the low-color lakes, this was not always the case. Highly localized hydrologic events evidently play a role in producing episodes of high CDOM levels in these lakes.

---

**Fig. 2.** Seasonal box plots for $a_{440} \text{ (m}^{-1})$ in Florida lake survey on effects of acid deposition (Hendry & Brezonik, 1984): (a) 17 low-color lakes; (b) three moderate color lakes.

**Fig. 3.** Time trends for $a_{440} \text{ (m}^{-1})$ in four Wisconsin LTER lakes: (a) Allequash, (b) Big Muskellunge, (c) Crystal Bog, (d) Trout Bog. Data source: Color — Trout Lake Area 1989–current, North Temperate Lakes Long Term Ecological Research program (http://lter.limnology.wisc.edu), NSF, E. H. Stanley, Center for Limnology, University of Wisconsin-Madison.
Seasonal box plots of the three colored LTER lakes (Fig. 4) show only small differences in spring, summer and fall medians for Crystal Bog. The mean values for Trout Bog were significantly different by season ($p = 0.001$), and those for Lake Allequash were significant at $p = 0.1$ but not at $p = 0.05$. The long-term trend of increasing CDOM in Crystal Bog and high inter-annual variability of CDOM in Lake Allequash may have obscured seasonal trends in these lakes. Overall, the results are consistent with the idea that photobleaching can affect seasonal variations in CDOM levels in colored lakes (e.g., Pace & Cole, 2002), but other causes also could be responsible for the observed trends. Although box plots for the four low-CDOM lakes showed some differences in median $a_{440}$ (data not shown), no consistent patterns were found for either the maximum or minimum in median $a_{440}$, and differences among seasons were not statistically significant for any of these lakes.

CDOM levels in the Mississippi River near Minneapolis, Minnesota during 2007 (Fig. 5a) varied between 0.71 and 4.3 m$^{-1}$ (mean = 1.8 m$^{-1}$; std. dev. = 0.8 m$^{-1}$; CV = 43%). Peak values were associated with a high-flow period in autumn, but the lowest value occurred during a moderately high flow event (~300 m$^3$/s) in early June; other low values occurred consistently during an extended low flow period in summer. The low CDOM values during late spring and summer probably reflect the increased importance of low-CDOM groundwater as a source for river flow during a period of low rainfall. High CDOM values during the fall high flow period suggest increased importance

---

Fig. 4. Seasonal box plots for $a_{440}$ (m$^{-1}$) in three Wisconsin LTER lakes with moderate to high CDOM: (a) Allequash Lake, (b) Crystal Bog Lake, (c) Trout Bog Lake. See Fig. 3 for source of data.

Fig. 5. (a) DOC, mg/L, (red squares) and $a_{440}$, m$^{-1}$, (blue diamonds) versus date in 2007 in Mississippi River at Minneapolis water intake; (b) $a_{440}$ versus flow, m$^3$/s, at same site in 2007; (c) DOC versus flow at same site in 2007; (d) DOC versus $a_{440}$ for the data in (a). Values of $a_{440}$ calculated from data in Brinkman and Hozalski (2011); other data from B. Brinkman (Sweet Briar College, pers. comm., 2013).
of contributions from forested and wetland areas in northern Minnesota during large-scale frontal events, and the low CDOM levels in the June event suggest that runoff was from more localized storms near the Minneapolis–St. Paul region. Overall, the relationship between $a_{440}$ and river flow (Fig. 5b) exhibited much scatter ($r^2 = 0.35$), reflecting the complexity of factors affecting transport of DNOM during runoff events—e.g., nature of antecedent conditions (wet versus dry) and intensity and duration of rainfall. The relationship between DOC and flow (Fig. 5c) was even more scattered ($r^2 = 0.10$), reflecting the complexity of water sources with different CDOM levels that contribute to river flow during the year.

CDOM levels in the SLRE showed large seasonal variations in 2013 (Fig. 6). Lowest levels occurred in late winter and highest values in summer, although not necessarily consistently at all stations. Ratios of maximum to minimum $a_{440}$ values ranged from ~2.5 to 5.5 across the sites.

In summary, CDOM data from many sites, both lakes and rivers, in the U.S. Upper Midwest and Florida show large temporal variability at seasonal, annual, and multi-year time scales. The major cause of this variability likely is hydrologic variations in the contributing watershed (e.g., variable input of runoff containing colored humic material from forests and wetlands). Internal processes, including autochthonous CDOM production, photobleaching and other removal mechanisms, likely are also involved in lakes. The time scales of these internal processes are much longer (months) than the processes that affect chlorophyll levels in water bodies, which can produce important variations.
concentration changes at time scales of days. In the absence of major hydrologic influences, CDOM in standing waters is unlikely to exhibit large intra-seasonal changes, but we still lack closely spaced measurements to verify this supposition. In contrast, major changes in river and stream loadings can occur at time scales of hours to days in response to rainfall-runoff events, as Fig. 5a illustrates for the Mississippi River, where high rainfall caused 440 to nearly double in one week in October 2007.

The above findings have important implications regarding the collection of CDOM calibration data for remote sensing imagery. Unless it can be shown that CDOM levels are temporally stable in a given aquatic system, such data should be collected close to the image acquisition date, preferably within 1–2 months in lakes and within a few days in large rivers, and perhaps even closer in time in small rivers and streams or if major changes in precipitation conditions are expected. In contrast, Cardille, Leguet, and del Giorgio (2013) recently concluded that there was no bias toward an incorrect model when using field calibration data as much as 1–7 years older than satellite images, even though the use of older CDOM data did cause somewhat greater scatter than found in models with near concurrent use of older CDOM data. Slope values for the relatively low r² values for the 2001–2002 period were limited. Data to analyze seasonal and long-term trends in CDOM concentrations were missing in one or both seasons for some years. Consequently, only for the idea that the source and nature of DOC are different in low-CDOM environments than in high-CDOM systems is provided in Fig. 7b, which shows that SUVA254 increases with the logarithm of CDOM (SUVA254 = 1.145 ln(440) + 1.76; r² = 0.90) across all the sites. SUVA254, an indicator of the aromatic content of DNOM, generally is higher in allochthonous humic matter than in autochthonous DNOM.

Table 2
Summary of spectral slopes in visible range for Wisconsin LTER sites.

<table>
<thead>
<tr>
<th></th>
<th>Allequash Lake (n = 6)</th>
<th>Crystal Bog (n = 15)</th>
<th>Trout Bog (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spring</td>
<td>Summer</td>
<td>Spring</td>
</tr>
<tr>
<td>S400–460</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0124</td>
<td>0.0147</td>
<td>0.0122</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0026</td>
<td>0.0015</td>
<td>0.0007</td>
</tr>
<tr>
<td>C.V.</td>
<td>20.7</td>
<td>10.5</td>
<td>5.5</td>
</tr>
<tr>
<td>S460–550</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0102</td>
<td>0.0117</td>
<td>0.0089</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0022</td>
<td>0.0031</td>
<td>0.0016</td>
</tr>
<tr>
<td>C.V.</td>
<td>21.1</td>
<td>26.3</td>
<td>17.5</td>
</tr>
</tbody>
</table>

4.2. Variability of CDOM spectral slopes

Plots of ln(A) versus λ in the visible range (400–700 nm) were quasi-linear for all historical data and measurements on samples collected in 2013 (results not shown). Most samples, however, showed a small decrease in slope around 450–470 nm. In addition, many samples with low CDOM values had very low absorbance values above ~650 nm, often approaching detection limits of the spectrophotometer even with 5–10 cm cells, which resulted in noisy data for this region. For analysis of spectral slopes (S) we truncated all spectra at 650 nm and computed separate slopes for the ranges 400–460 and 460–650 nm.

The variability of spectral slopes across the 35 sites sampled in summer 2013 was large; ranges for S400–460 and S460–550 were 0.007–0.017 and 0.004–0.013, and mean values were 0.013 and 0.009, respectively. Closer inspection of the results, however, showed that slope values were related to CDOM levels (a440). For most of the sites (23) S400–460 had a narrow range (0.014–0.018; Fig. 7a); these values are similar to values of S typically reported for DOC in the near UV. However, a group of low-CDOM sites (a440 < 5 m⁻¹) had values ≤ 0.010 (mean = 0.0088, n = 10). These patterns may indicate that the DNOM in most of the water bodies, especially sites with moderate–high CDOM (a440 > 5), was dominated by allochthonous (humic-rich) material, but DNOM in most sites with low CDOM was dominated by autochthonous (or anthropogenic) material. Four exceptions (low-CDOM, high S460–550) are lakes (Holland, Child, Superior, and Sturgeon SW Bay) with low–moderate levels of DOC (5–8 mg/L). Although their CDOM levels were low, it seems likely that some of the DNOM was allochthonous. Further support for the idea that the source and nature of DOC are different in low-CDOM environments than in high-CDOM systems is provided in Fig. 7b, which shows that SUVA254 increases with the logarithm of CDOM (SUVA254 = 1.145 ln(440) + 1.76; r² = 0.90) across all the sites. SUVA254, an indicator of the aromatic content of DNOM, generally is higher in allochthonous humic matter than in autochthonous DNOM.

S460–550 values were smaller than corresponding S400–460 values in almost every case, and the difference in means was highly significantly (p < 0.000, paired t-test). The general pattern observed for trends in S400–460 versus a440 was found for S460–550, but scatter was somewhat larger and only one low-CDOM site (a440 < 5), Sturgeon Lake SW Bay, had a slope > 0.009.

Seasonal variations in spectral slopes were very small for the SLRE stations, and variations between stations also were small. The mean and standard deviation of S400–460 for 81 spectra collected on 6–20 sites over seven sampling dates from late February to mid-September was 0.0156 ± 0.0009 (CV = 5.6%); corresponding values for S460–550 were 0.0109 ± 0.0012 (CV = 10.9%). No seasonal changes were found for either slope parameter. Similarly, values of S400–460 for the two LTER bog lakes with high CDOM (Table 2) showed no differences between spring and summer. S460–550 was higher in summer than in spring in Trout Bog but not in Crystal Bog. The variability of S460–460 over the period of analysis (1990–2012) was small for the bog lakes (CV = 4%–6%), but S460–550 was more variable (CV range: 4%–20%). Some variability in the latter data may reflect difficulties in dealing with increasingly low absorbance values at longer wavelengths (e.g., Green & Blough, 1994). Data to analyze seasonal and long-term trends for Allequash Lake were limited. Most August spectra from 2001 to 2011 contained negative values at wavelengths > 650 nm, and samples were missing in one or both seasons for some years. Consequently, only...
not be an important factor contributing to uncertainty in that variations in spectral slope values over time or across sites should not affect CDOM

mer 2013

r2 = 0.925 for the relationship between DOC and CDOM, and no major variations in surveys likely would lead to poor DOC–CDOM relationships. These are likely explanations for the poor relationship (r2 = 0.30) between DOC and CDOM in the MPCA lake database (Fig. 9). Similarly, poor DOC–CDOM relationships may exist in rivers like the Upper Mississippi (Fig. 5d) that have multiple source waters with differing DOC and CDOM levels that vary in relative importance over time, as described in Section 4.1.

In contrast, when data sets are restricted to exclude water bodies with heavy impacts from human activities that yield DOC with low color, tight DOC–CDOM relationships can be observed. Such is the case when we examine a group of moderately to highly colored lakes in our 2013 studies: the Big Sandy and Sturgeon chains of lakes in north-central Minnesota (Fig. 10a). A close linear relationship (r2 = 0.968) over a large CDOM range (a440 = 1.8–25.1 m−1) was found for these 12 lakes.

As Stanley, Powers, Lottig, Buffam, and Crawford (2012) recently described, human activities such as intensive agriculture or input of wastewater effluent can have large effects on the nature of DOC in rivers and streams. For example, hypereutrophic lakes often have high concentrations (>10 mg/L) of autochthonous DOC, which has a low color intensity per unit of DOC, and wastewater effluents typically have 10–20 mg/L of DOC (called EFOM, effluent organic matter; Shon, Vigneswaran, & Snyder, 2006) that is low in color per unit of DOC. In addition, Xiao, Sara-Aho, Hartikainen, and Vähätalo (2013) and Köhler et al. (2013) recently showed that the association of dissolved iron with DOC affects its color intensity, and spatial variations in iron concentrations thus can affect CDOM–DOC relationships. Inclusion of water bodies affected by human activities and geographic variations in dissolved iron concentrations in surveys likely would lead to poor DOC–CDOM relationships.
A weaker relationship ($r^2 = 0.72$) was found for DOC and CDOM measurements from seven 2013 dates at various locations in the St. Louis River Estuary near Duluth, MN (Fig. 10b). However, closer inspection showed that the outliers in the figure all were from one date in early spring (May 7). River flow had peaked shortly before this date due to snowmelt but still was high. Rather than being randomly distributed, the May 7 data have a separate DOC–CDOM relationship ($r^2 = 0.7$, excluding one outlier) with a higher slope (2.0) and higher DOC per unit of CDOM; i.e., the DOC had a lower CDOM content than that for the six other 2013 sampling dates ($r^2 = 0.925$; slope = 1.28), which ranged from February 28 to September 18. Although the nature and source of the less colored DOC in early May are unknown, the finding supports our concern that prediction of DOC concentrations from CDOM levels must be done cautiously because DOC–CDOM relationships in colored surface waters are not constant.

Several studies have reported close correlations between CDOM, expressed in chlorophyllinatetrin units or as $\lambda$, (where $\lambda$ is in the range 400–450 nm), and DOC concentration, but others have focused on the variability in the relationship (e.g., Müller et al., 2013; Wallage & Holden, 2010; Worrall & Burt, 2010). Gorham et al. (1998) reported that the ratio of color expressed in chloroplatinate units to DOC (mg/L) when all the data were pooled. The average color/DOC ratio ($\alpha_440$) in peatland porewaters in northern England, which for non-peat groundwater was 0.87–0.93 for five of six sampling dates over a one-year period for CDOM–DOC relationships on 42 streams in Nova Scotia (one date had an $r^2$ of 0.61). The best-fit relationships were not linear, and involved log or nth root (where n = 2–5) transformations. Moreover, slopes of the relationships were different for each sampling period, suggesting that the fraction of DOC that was colored varied with season. The CDOM/DOC ratio rose with increasing DOC concentration, and the authors inferred from this that the quality of DNOM changed with increasing wetland influence. Similarly, Wallage and Holden (2010) found an $r^2$ of 0.77 for 1047 pairs of measurements of DOC and chlorophyllinatetrin absorbance at 400 nm for peatland porewaters in northern England, and they showed that the relationship suffered from heteroscedasticity (i.e., variance was not constant across the range of the data), which calls into question the predictive reliability of the regressions. Moreover, different DOC–color relationships were found as a function of depth within peat profiles and when the samples were grouped according to land management practices in the peatlands.

Worrall and Burt (2010) examined a large, long-term (1974–2005) database of color and DOC on 70 rivers in the United Kingdom and found a trend in the ratio of color expressed in chlorophyllinatetrin units to DOC (mg/L) when all the data were pooled. The average color/DOC ratio decreased from 6.5 in 1974 to a minimum of 1.5 in 1989 and then slowly rose to ~4 by 2005. Variations in DOC concentration and flow were not sufficient to explain the variations, and the authors suggested they may reflect changes in soil source water chemistry. Müller et al. (2013) recently analyzed data from 911 lakes in Sweden and reported that the ratio $\alpha_440$/TOC (TOC = total organic carbon, nearly all of which is DOC in these lakes) decreased in nonlinear fashion with increasing water residence time, $\tau_w$, in a lake ($\tau_w = V/Q_{out}$, where $V$ = lake volume and $Q_{out}$ = volumetric discharge rate from a lake) or with total water residence time, $\tau_{w,TOT}$, which for non-headwaters lakes accounts for water residence times in upstream lakes in a drainage basin. Although there was considerable scatter in the relationships, the predicted value of $\alpha_440$/TOC for $\tau_w < 1$ y was $\sim 0.7$, but this dropped to ~0.2 for $\tau_w > 7$ y. Similarly, Köhler et al. (2013) explained differences in long-term CDOM trends (Kutser, 2012) in multi-basin Lake Mälaren, Sweden as resulting from in-lake processes. Increased terrestrial inputs of organic matter to the western basins of this lake over the past 40+ years resulted in a more than doubling of CDOM levels in those basins, but levels in the easternmost basin increased by less than 25%. The authors concluded that the difference in trends resulted from rapid loss of dissolved iron from the western basins, which was strongly correlated with a loss in color. The nature of the DOC also shifted from more colored allochthonous material to less colored autochthonous material as water moved from west to east in the lake. Overall, the above analyses suggest that the quality of DOC in lakes changes over time, with fresh DOC draining into lakes from forested and wetland areas being more highly colored and older DOC less colored.

In contrast to the cautionary evidence described above, Spencer et al. (2012) recently concluded from a study of 30 rivers, that CDOM could be used as a surrogate for DOC in most major rivers of North America. Important exceptions included the Colorado, Columbia, Rio Grande and St. Lawrence (which drains the Great Lakes). Characteristics of DNOM measured on the latter four rivers, for which $r^2 > 0.2$ for the DOC-$\alpha_440$ regression, suggested their DNOM was autochthonous, antherogenic, and/or photodegraded allochthonous material. It should be noted that the $r^2$ values for relationships between DOC and $\alpha_440$ were $\leq 0.5$ for more than a third (11) of the 30 rivers, and the mean $r^2$ value for the 26 rivers with $r^2 > 0.2$ was 0.686 (range = 0.306–0.961). Although $r^2$ values for regressions between DOC concentration and $\alpha_254$ were much stronger (mean = 0.935; range = 0.822–0.999 for the same 26 rivers), this part of the spectrum is not accessible by the sensors used in satellite imagery.

4.4. Effects of CDOM and other water quality characteristics on reflectance spectra

Reflectance spectra were measured over the period August 8 to September 18, 2013 at 32 of our 35 sites, and samples were collected at the same time to measure water quality variables that affect the optical properties of water; see Table 1 for a list of sites and water quality values. To aid in assembling the reflectance spectra into groups of lakes with similar optical characteristics, we performed a cluster analysis on all sites using three water quality variables that affect the optical properties of water bodies and thus affect reflectance spectra: $\alpha_440$, chlorophyllinatetrin $\alpha$ (chl $\alpha$), and SD. We were unable to use TSS, an important factor affecting the optical properties of water, as a clustering variable because of missing data at several sites, but SD, an integrative measure of water clarity, includes the influence of TSS. We evaluated several clustering options and found that log-transformed values of $\alpha_440$ and chl $\alpha$ and reciprocal values for SD (SD $^{-1}$), along with average linkage as the joining algorithm and a Euclidean distance measure, yielded easily interpretable clusters.

Overall, the analysis grouped the sites into two large categories (Fig. 11) based primarily on CDOM (high and low CDOM levels) at a distance measure (similarity index) of 1.2. At a higher level of similarity, (distance measure of 0.8), the procedure yielded 5 clusters that could be interpreted in terms of major optical properties: A. moderate color, highly eutrophic (based on chl $\alpha$ and SD); B. high color, mesotrophic to eutrophic; C. high color, oligotrophic (dystrophic); D. low color, oligotrophic; and E. low color, moderately eutrophic.

The shapes of the measured reflectance spectra differed considerably across the clusters, as representative spectra for each cluster (Fig. 12) show. The two spectra presented for cluster B show that spectra for sites in this cluster exhibited some variations. In general, however, spectra in a given cluster showed high similarity (Fig. 13). If we assume that each cluster is supposed to represent waters with similar optical properties (given that the clustering variables represent important optical properties of water) and should have similar reflectance spectra, it is apparent that some misclassifications occurred. For example, the spectra for Lakes Okomis and Harriet, which were clustered in class D (spectra 4 and 5, Fig. 13a), appear to belong to the low color, moderately eutrophic class (cluster E) represented by Lake Calhoun (E3, Fig. 12); all three spectra show the characteristic chlorophyll trough near 670 nm and the peak near 705 nm characteristic of scattering by phytoplankton. In contrast, the other spectra in class D (Fig. 13a) lack both features.

Similarly, three class B lakes – Johnson, Section 11, and West Sturgeon (site IDs 13, 16, 21) – have spectra (Fig. 13c) that fit those of the dystrophic lakes (class C) better than those of other sites in class B (high color, mesotrophic–eutrophic). Overall, class B shows the highest
heterogeneity in reflectance spectra Fig. 13c). Several SLRE sites (Allouez Bay, Pokegama Creek, Superior Bay, and St. Louis River Sta. 8; sites 25, 27, 30 and 31 in Fig. 13c) show evidence of suspended solids: high reflectance across a broad band from ~580 to 710 nm) and small evidence of the 670-nm chlorophyll trough and 705-nm phytoplankton scattering peak. Other sites, such as Big Sandy Lake East Basin and Big

Fig. 11. Cluster diagram for the 35 sampling sites in Table 1 based on three optical properties: SD, CDOM, and chlorophyll a.

Fig. 12. Representative reflectance spectra for clusters A–E identified in Fig. 11. Numbers next to spectra correspond to site IDs in Table 1.
different shapes depending on whether CDOM is the only important optical property or other properties (TSS and chlorophyll) also play important roles (see Fig. 13bc). In the former cases, the spectra are simple—nearly flat and very low reflectance values across the visible range. In the latter cases spectra also show the reflectance characteristics of the other optical properties, and the influence of CDOM appears to be mainly in causing lower reflectance, especially at wavelengths < ~600 nm. CDOM-rich waters typically are associated with forested regions that tend to be nutrient poor and have little potential for erosion-causing elevated TSS levels. On a global basis, the spectra of these lakes are thus probably most common. Nonetheless, CDOM-rich waters whose spectra are heavily influenced by chlorophyll and/or TSS are common in Minnesota, as illustrated by the SLRE, Big Sandy chain of lakes and Mississippi River (Table 1, Fig. 13bc). They also are likely to be common in other regions where the natural conditions that produce high CDOM levels intersect with human activities like agriculture and urban development.

4.5. CDOM predictive model development

The spectra discussed above lead to an obvious question related to estimating CDOM from reflectance data: given the variety of spectral shapes for waters containing CDOM, can we produce accurate CDOM estimates across the range of spectra using one empirical model, or are different models needed depending on whether CDOM is the only IOP affecting reflectance or other properties like chl a or TSS also are important? We addressed this question by using the in situ hyperspectral data to simulate the measurements that would be obtained in various bands (see Fig. 12 for band locations) of Landsat 8 (launched in 2013) and the forthcoming Sentinel-2 and Sentinel-3 satellites of the European Space Agency, as well as the narrow (hyperspectral) bands in the Hyperion and HICO satellite sensors. With improved spectral, spatial and temporal coverage compared with existing satellites, the Sentinel satellites are expected to provide an important advancement in capabilities for remote sensing of inland water resources. Landsat 8 also has improved characteristics compared with Landsat 7, including an additional blue band that may be beneficial for estimating CDOM. By simulating data from these sensors using our in situ spectra, we can explore how well they can measure CDOM (and other optical characteristics of water) in complex inland waters without atmospheric disturbances.

We used stepwise regression with the sensor band simulations for Sentinel-2, Sentinel-3, and Landsat 8 to explore the best fit band or band ratio model using natural log (ln) transformed $a_{440}$ as the dependent variable and all the non-transformed and ln-transformed bands and band ratios for each sensor as independent variables. The best models for $a_{440}$ based on the simulated Landsat 8, Sentinel-2 and Sentinel-3 bands and the field-measured reflectance spectra yielded $r^2$ values of 0.84 to 0.86 (Table 3). The narrow, well-positioned bands of Sentinel-3 worked best, but they may not be needed for CDOM because the wide bands of Landsat worked almost as well. Models using the new Landsat 8 blue band, OLI 1, worked slightly better than models using the traditional Landsat blue band but not as well as the ratio of the Landsat green and red bands. This finding is reassuring for historical measurements of CDOM since it suggests that the Landsat archive can be used, although Kutser (2012) found that low radiometric sensitivity of the earlier Landsat sensors led to poor results for waters high in CDOM. Models that worked best for the well-positioned Sentinel bands were ratios of ~500 and ~750 nm.

We also evaluated the published model forms of Kutser et al. (2005), Menken et al. (2006), Ficek et al. (2011), and Griffin et al. (2011); see Table 3 for band combinations and model forms. The best Landsat 8 model has nearly the same spectral bands and mathematical form as the model of Kutser et al. (2005) and actually had a higher $r^2$ (0.84) than the original relationship (0.73) reported by Kutser et al. The model of Menken et al. (2006), which was based on hyperspectral data, yielded a nearly identical fit ($r^2 = 0.83$). Although derived

---

Fig. 13. Reflectance spectra for clusters A–D in Fig. 11: (a) spectra for cluster D; (b) spectra for clusters A (1,6,7,9) and C (18,35); (c) spectra for cluster B. Numbers next to spectra correspond to site IDs in Table 1.
independently, this model and that of Kutser et al. have the same form and use essentially the same band ratios (although Menken et al. used much narrower bands than the ALI bands used by Kutser et al., 2005). The model of Ficek et al. (2011), which has the same form and a ratio of narrow bands similar to that of Menken et al. (2006), yielded a slightly higher $r^2$ value (0.84). Overall, however, it appears that narrow bands are not needed to obtain close fitting relationships for CDOM and reflectance data. An alternative model of Menken et al. (2006), which uses reflectance at 412 nm and the ratio of reflectance at 412 nm to 670 nm to correct for effects of chlorophyll, yielded poor fit ($r^2 = 0.38$) to our data set, lending some support to the view that use of short wavelengths (e.g., <450 nm), where CDOM absorbs light most strongly, is not the best way to measure CDOM by remote sensing. Finally, the Landsat-based model of Brezonik et al. (2005) yielded a mediocre fit ($r^2 = 0.53$), and the model of Griffin et al. (2011), which worked well for data from a Siberian river, performed poorly with our data set ($r^2 = 0.21$). Zhu et al. (2014) also found poor performance of this model using data from an estuarine area of Saginaw Bay (Lake Huron).

The models we evaluated for $a_{440}$ also produced strong relationships between reflectance ratios and DOC ($r^2 = 0.81$), but it must be noted that DOC itself does not have optical properties, and the relationships work because of the strong correlation between CDOM and DOC in our data set (Fig. 8). As discussed earlier, additional field studies would be needed to verify the relationship (or establish a new one) between CDOM and DOC for other areas.

As a step toward determining whether different models are needed for accurate estimates of CDOM in the three classes of high CDOM spectra (A, B, C in Fig. 11), we calculated the % difference between measured and predicted $a_{440}$ for each class (Table 4). For this analysis we moved three water bodies with simple reflectance spectra dominated by CDOM that were misclassified as cluster B into cluster C (the high CDOM, dystrophic class). In terms of absolute values of the % differences between measured and predicted $a_{440}$, the Sentinel-2 algorithm produced the best results across the three CDOM clusters; on average, the other three models we evaluated exhibited about the same amount of scatter. Relative to the three clusters, predictions for cluster B were clearly the most scattered, and cluster A predictions had slightly less scatter than those for cluster C. Insofar as cluster B had the largest number of sites and its reflectance spectra exhibited greater diversity than the spectra in clusters A and C, the higher differences between measured and predicted $a_{440}$ values for cluster B suggests that different predictive models for each of the three types of high-CDOM waters may produce more accurate results. A more definitive answer to the question might be obtained by developing and comparing separate predictive models for each CDOM class, but because of the small number of sites in classes A and C (see Fig. 11), we did not attempt to do this. Additional data collection may allow this to occur in the future.

Finally, we can ask whether the spectra in Figs. 12 and 13 help us to understand why the algorithms of Kutser et al. (2005), Menken et al. (2006) and Ficek et al. (2011) work, given that they all feature nonlinear relationships of reflectance ratios at wavelengths >500 nm, where the influence of CDOM should be relatively small compared to that at lower wavelengths. One key difference between the spectra for sites with low CDOM levels and those with high CDOM levels is a difference in slopes of the spectra in the range ~570–650 nm. For low-CDOM waters, the spectra all decline with increasing wavelength in this region, but for high-CDOM waters the opposite trend occurs (increasing reflectance with increasing wavelength). Further studies should be undertaken to evaluate these trends in a more quantitative manner.

### 5. Summary and conclusions

1. CDOM levels in lakes and rivers of the U.S. Upper Midwest and lakes in Florida are highly variable at seasonal and multi-year time scales and at shorter intervals in some rivers and lakes, with coefficients of variation of 30%–50% for $a_{440}$ common in the historical and new data sets we examined. CDOM values used to calibrate imagery thus should be measured close to the image acquisition date, preferably within 1–2 months in lakes and a few days in large rivers, unless

### Table 3

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation form*</th>
<th>Coefficients</th>
<th>Adj. $r^2$</th>
<th>$r^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kutser et al. (2005)</td>
<td>$\ln(a_{440}) = a_0 + a_1\ln(\text{OLI2/OLI4})$</td>
<td>1.582, -1.507</td>
<td>0.833</td>
<td>0.838</td>
<td>0.473</td>
</tr>
<tr>
<td>Brezonik et al. (2005)</td>
<td>$a_{440} = a_0 + a_1(\text{OLI2} + \text{OLI2/OLI5})$</td>
<td>2.304, -255.88, -0.2542</td>
<td>0.496</td>
<td>0.525</td>
<td>1.232</td>
</tr>
<tr>
<td>Menken et al. (2006)</td>
<td>$a_{440} = a_1(R_{500}/R_{700})^{a_2}$</td>
<td>1.734, -1.321</td>
<td>0.821</td>
<td>0.826</td>
<td>0.490</td>
</tr>
<tr>
<td>Ficek et al. (2011)</td>
<td>$a_{440} = a_1(R_{500}/R_{700})^{a_2}$</td>
<td>-1.665, -2007, 9.723</td>
<td>0.337</td>
<td>0.380</td>
<td>5.912</td>
</tr>
</tbody>
</table>

* Equations for published models were reformatted to a common measure of CDOM ($a_{440}$) and common set of bands—Landsat 8 OLI 1–5 (except the hyperspectral models, which are unchanged). Original equations used absorbance or absorptivity at different wavelengths for CDOM; Kutser et al. (2005) reported their equation in the equivalent format $a_{440} = a_1(\text{ALI2}/\text{ALI3})^{a_2}$; the ALI bands have nearly same ranges as OLI bands 3 and 4.

### Table 4

<table>
<thead>
<tr>
<th>Model</th>
<th>A (moderate CDOM, highly eutrophic)</th>
<th>B (high CDOM, mixed eutrophy and high TSS)</th>
<th>C (high CDOM, oligo-trophic/ dystrophic)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ave. $^b$</td>
<td>Range $^b$</td>
<td>Ave. $^b$</td>
</tr>
<tr>
<td>Landsat 8 (Kutser et al., 2005)</td>
<td>30.0</td>
<td>25.3–43.4</td>
<td>39.9</td>
</tr>
<tr>
<td>Menken et al. (2006)</td>
<td>37.8</td>
<td>28.3–51.0</td>
<td>39.5</td>
</tr>
<tr>
<td>Sentinel-3</td>
<td>15.7</td>
<td>13.2–16.9</td>
<td>42.9</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>21.2</td>
<td>0.3–38.1</td>
<td>27.7</td>
</tr>
</tbody>
</table>

* $^a$ Clusters A, B, and C are defined in Fig. 11; misclassified sites 13, 19 and 21 in cluster B were moved to cluster C for this analysis.

$^b$ Average or range of absolute values of % differences between measured and predicted $a_{440}$.

$^c$ Average of % differences between measured and predicted $a_{440}$ taking sign of difference into account.
data exist to show that levels are temporally stable over longer (or shorter) time periods in a given aquatic system.

2. Spectral slopes ($S$) for CDOM in the visible range typically show a change in slope near 460 nm. Slopes vary little over time, even over multi-year periods, within sites, but variations occur between sites. Values of $S_{460-440}$ for waters with $a_{440} > 5$ m$^{-1}$ generally occur in a narrow range, 0.014–0.018, similar to reported $S$ values in the near UV, but values of $S_{400-460}$ for waters with low CDOM generally are smaller and more variable, as are values for $S_{460-650}$ across the range of CDOM levels. Overall, we conclude that the variability of $S$ in the visible range should not have an important effect on the reliability of $a_{440}$ estimates made from remotely sensed reflectance measurements at wavelengths > 550 nm, at least for waters with moderate to high CDOM levels.

3. A strong correlation ($r^2 = 0.925$) was found between CDOM levels and DOC concentrations in 34 surface waters sampled in 2013, but the SEE still suggests an uncertainty of ±20% in predicting DOC at moderate CDOM levels ($a_{440} = 5$ m$^{-1}$). From CDOM–DOC relationships for other data sets we analyzed and those reported in the literature we conclude that the fraction of DOC that is colored varies widely, and slopes of regressions between CDOM and DOC are highly variable. Prediction of DOC concentrations in water bodies from CDOM levels alone, whether measured in the laboratory or by remote sensing, thus is subject to substantial uncertainty, and until we have a better understanding of variations in DOC–CDOM relationships, additional field sampling is essential to verify DOC concentrations predicted from remotely sensed CDOM measurements.

4. Shapes of in situ reflectance spectra for CDOM-rich waters vary greatly depending on the concentrations of suspended solids and chlorophyll that also affect the optical properties of water. Nonetheless, from our analysis of results for several predictive models we are not able to verify that different algorithms would improve the accuracy of $a_{440}$ calculated from remote sensing data for waters where CDOM alone affects reflectance versus waters where other constituents affect the spectra.

5. The best band or band ratio models for simulated Landsat 8, Sentinel-2 and Sentinel-3 bands from field-measured reflectance spectra yielded high $r^2$ values (0.83–0.86) for $a_{440}$. The Landsat 8 bands work nearly as well for $a_{440}$ as the narrower Sentinel bands and hyperspectral bands, probably because CDOM is characterized by a broad exponential increase in absorbance with decreasing wavelength rather than having specific peaks or troughs in absorbance or reflectance.

Acknowledgments

We thank the following individuals for supplying data for our analysis: Emily Stanley, Univ. of Wisconsin, Madison, North Temperate Lakes LTER Program; Bruce Monson, Minnesota Pollution Control Agency, St. Paul; and Bethany Brinkman and Ray Hazolaki, Univ. of Minnesota. PLB thanks his former sampling and lab crews and graduate students at the Univ. of Florida, especially R. Vorton, E. Shannon, and C. Hendry, for their dedication in producing the historical databases on Florida lakes. We thank Sandra Brovold, Nolan Kleijan, and Adam Worm, Dept. of Ecology, Evolution, and Behavior, Univ. of Minnesota, Luke Loken, Univ. of Wisconsin-Madison, and Chip Small, Univ. of St. Thomas, for field and lab assistance with the 2013 SLRE data, and the Lake Supe-

rior North Estuarine Research Reserve and then reserve manager Shon Schooler for sampling assistance on the St. Louis River Estuary. We also greatly appreciate the assistance of Bryan Leavitt and Rick Perk, CALMlT, University of Nebraska, for the loan of their spectroradiometer system and their helpful advice in operating the system and interpreting the data. PLB thanks the Univ. of Minnesota Office of the Vice President for Research (UMOVR) and UM Retirees Association for an award from the Professional Development Grants Program for Retirees that provided critical financial support. JCF acknowledges support from the Minnesota Sea Grant Program supported by the NOAA office of Sea Grant, United States Department of Commerce, under grant No. R/RegHCE-09-12. LGO and MB also acknowledge support from the UMOVR for funding under the U-Spatial: Spatial Science and Systems Infrastructure grant.

References


