

2016 Project Abstract

For the Period Ending June 30, 2019

PROJECT TITLE: Assessment of Surface Water Quality with Satellite Sensors

PROJECT MANAGER: Jacques C Finlay

AFFILIATION: University of Minnesota

MAILING ADDRESS: 140 Gortner Lab, 1479 Gortner Avenue

CITY/STATE/ZIP: Saint Paul, MN 55108

PHONE: 612 624 4672

E-MAIL: jfinlay@umn.edu

WEBSITE: <http://cbs.umn.edu/finlay-lab/home>

FUNDING SOURCE: Environment and Natural Resources Trust Fund

LEGAL CITATION: M.L. 2016, Chp. 186, Sec. 2, Subd. 04i

APPROPRIATION AMOUNT: \$345,000

AMOUNT SPENT: \$345,000

AMOUNT REMAINING: \$ 0

Sound bite of Project Outcomes and Results

This project developed methods to measure key water quality parameters across the state's waterbodies using satellite imagery. Current ground-based sampling has limited ability to measure status and trends in water quality across our lakes and rivers. Our innovative methods can routinely measure major water quality indicators in Minnesota's 10,000+ lakes.

Overall Project Outcome and Results

Water quality monitoring is essential for managing Minnesota's surface waters, maintaining the services they provide, and detecting changes caused by environmental stressors. Direct measurements of water quality are possible, however, in only a small fraction of the thousands of lakes and river miles in the state. Methods developed in this project allow use of increasingly frequent satellite observations to measure water clarity and the three key water quality indicators that control it: algae, colored dissolved organic matter (CDOM), and suspended solids. Because these parameters have distinct impacts on water quality, the ability to measure them directly across the state's waters enables comprehensive assessment of water quality status and trends and increases understanding of the causes and consequences of water quality degradation. We developed methods to relate direct measurements of water quality to satellite imagery, assessed atmospheric correction techniques and validated methods using independent datasets. We applied these methods to measure water quality parameters on lakes >10ha, and provide the information at <https://lakes.rs.umn.edu>. Examples are included in appendix 1. Our methods extract information at seasonal to annual scales for algae, CDOM and suspended solids in lakes at state, regional, county, and watershed scales. Water quality parameters were linked to disinfection byproduct formation potential in drinking water treatment and degradation of contaminants driven by sunlight. CDOM levels were closely related to formation rates of two classes of disinfection byproducts, trihalomethanes and haloacetic acids, and to the production photo-induced reactive intermediates that degrade pesticides. This information can be used with remote sensing to assess pesticide persistence and suitability of surface waters for drinking water sources. An ongoing LCCMR project uses methods developed here with automated imagery acquisition and analysis to gather information on lake conditions at potentially a weekly basis. Project outcomes are summarized at <https://water.rs.umn.edu/> for use by researchers, managers, lake associations and the public.

Project Results Use and Dissemination

Information from this project has been disseminated through five ways during the three year project:

- 1) Data produced in the project is now freely available at <https://lakes.rs.umn.edu/>
- 2) Numerous presentations at meetings, agencies, and academic institutes have been given.
- 3) Five research articles documenting technical methodology and water quality relationships have been published. These publications have been included in our final report. Several other publications are being developed, and several others are planned.
- 4) We have engaged with state and agency partners in data gathering and interpretation, resulting in one publication, and plans for future collaborations with MPCA on refining water quality standards in the state based on project findings.
- 5) A widely used website for remote sensing of water quality (<https://water.rs.umn.edu/>) has been thoroughly revised with updated and expanded content.



Environment and Natural Resources Trust Fund (ENRTF) M.L. 2016 Work Plan Final Report

Date of Report: August 13, 2019

Final Report

Date of Work Plan Approval: June 7, 2016

Project Completion Date: June 30, 2019

PROJECT TITLE: Assessment of Surface Water Quality with Satellite Sensors

Project Manager: Jacques Finlay

Organization: University of Minnesota

Mailing Address: 140 Gortner Lab, 1479 Gortner Avenue

City/State/Zip Code: St. Paul, MN 55108

Telephone Number: (612) 624-6777

Email Address: jfinlay@umn.edu

Web Address: <http://cbs.umn.edu/finlay-lab/home>

Location: Statewide

Total ENRTF Project Budget:

ENRTF Appropriation: \$345,000

Amount Spent: \$345,000

Balance: \$0

Legal Citation: M.L. 2016, Chp. 186, Sec. 2, Subd. 04i

Appropriation Language:

\$345,000 the second year is from the trust fund to the Board of Regents of the University of Minnesota for a statewide assessment of water quality using new satellite sensors for high frequency measurement of major water quality indicators in lakes and rivers. This appropriation is available until June 30, 2019, by which time the project must be completed and final products delivered.

I. PROJECT TITLE: Assessment of Surface Water Quality with Satellite Sensors

II. PROJECT STATEMENT:

Minnesota's abundant surface waters face multiple threats related to land use change, eutrophication and invasive species. Water clarity is a key water quality indicator that can be monitored on many thousands of lakes in Minnesota with a simple Secchi disk or remotely using satellites. While water clarity data are highly valuable due to the ability to measure most water bodies on a frequent basis, current methods cannot distinguish between the three factors that affect water clarity in our lakes and rivers: algae, colored dissolved organic matter, and suspended solids (e.g., clay minerals). Because they have distinctly different impacts on water quality, the ability to measure them directly increases our understanding of the causes and consequences of water quality degradation. Technology now exists to make this advance. **This proposal takes an innovative next step in statewide assessment of water quality using new satellite sensors to remotely measure major water quality indicators in Minnesota's 10,000 lakes at high frequency and low cost.**

Background

The three major water quality drivers that are the focus of our project have distinct sources and impacts on surface waters. Algae are present in all surface waters and are a primary resource for lake and river food webs. *Algal biomass* is strongly affected by levels of phosphorus and nitrogen. When these nutrients are present at high levels, they result in eutrophication. Eutrophication reduces water quality by degrading habitat for native species, enhancing invasive species, reducing water quality for human use, and causing algal dominance by cyanobacteria, including species that produce harmful liver and neurotoxins. As a consequence, ability to monitor changes in algal biomass is a key part of understanding functioning of Minnesota's surface waters. *Colored dissolved organic matter* (CDOM) also occurs in all surface waters, and is the most abundant organic matter pool in many settings, especially in forested watersheds with wetlands. CDOM strongly affects water quality because it mobilizes metals and hydrophobic chemicals, serves as a source of reactive photochemical intermediates, controls many aquatic ecosystem processes, and has negative effects on production of safe drinking water. The paucity of data for CDOM distribution limits our understanding of the quality and characteristics of surface waters in Minnesota and the US in general. *Suspended sediment* enters surface waters during large storm events that cause erosion. Sediments can also be reintroduced to surface water when sediments from lake and river bottoms are resuspended due to wind, or invasive species such as the common carp which destroy aquatic plants and stir up sediment deposits.

Conventional methods for water quality assessment rely heavily on manual collection and analyses of surface water samples to characterize these constituents and other water quality variables. While Minnesota has an excellent surface water quality monitoring program, water quality conditions are highly variable in space, and can change very rapidly due to climate, land management and invasive species. Effective lake management requires long-term water quality information on a large number of lakes and streams so that managers can take into account differences among lakes, as well as changes through time for individual lakes at the watershed or regional scale. Unfortunately, only a small percentage of inland waters are currently monitored regularly by conventional methods, and long term (i.e. at least 20 years) water quality data are lacking for most inland water bodies. Comprehensive assessment of water quality is not practical with conventional point sampling methods due to limited resources, and thus water quality data are sparse in many areas. Furthermore, traditional ground-based monitoring programs target larger recreational lakes and thus are not randomly selected. Extrapolation from these lakes to the larger population would likely bias conclusions. Remote sensing methods can provide a solution to many of these issues and generally complement ground based monitoring by providing frequent and extensive measurements in virtually all lakes larger than 10 acre surface area in the state. This project will develop methods to allow assessment of an unprecedented number of lakes, and lays the groundwork for future remote sensing based monitoring programs that can provide comprehensive data and near real-time monitoring of important water quality variables.

The quality of our lakes and rivers directly affects the availability of clean drinking water and habitat for fish and other wildlife. In much of Minnesota, lake and river water quality is influenced mainly by excessive algae (from nutrients), colored organic matter from decaying woody plants in forests and wetlands, and suspended solids from stormwater runoff and invasive carp. These factors thus have distinct causes, and each also requires different management responses in watersheds. Each also has distinct effects on water resources: too much algae decreases water quality and habitat; high levels of colored organic matter decrease fish growth and interfere with natural contaminant degradation processes and drinking water treatment processes; suspended solids destroy fish habitats and clog waterways. The ability to detect these problems in lakes on a regular basis would provide an early warning system for changes and allow management of watersheds and surface waters in ways that are specific to the particular stressors causing the degradation.

Major Project Objectives

Our goals are to: 1) develop remote sensing methods to permit routine measurement of colored organic matter, algae, and suspended solids levels in Minnesota's waters; 2) apply these methods to our 10,000 lakes and large rivers, creating a database and corresponding maps; and 3) explore how variations in these water quality indicators influence the fate of contaminants (e.g., pesticides, mercury) and the suitability of water bodies to serve as drinking water supplies. Dissemination of information gained in this project on lake conditions and methods for satellite based monitoring of conditions will reach diverse stakeholder in the state and the nation via publications, integration into Lake Brower and meeting presentations.

General Project Activities and Methods

The project develops methods to use recently enhanced satellite based remote sensing capabilities to measure key water quality parameters: algae, dissolved color, and suspended sediments. The project will acquire optical remote sensing imagery from Landsat and Sentinel satellites concurrently with sampling of surface waters for these three parameters. Using these datasets, analyses will be conducted to determine the most robust relationships between remotely sensed reflectance measurements and field conditions. These models will be validated with independent datasets. Data describing distributions of these parameters will be integrated into Lake Browser, and disseminated to state and federal agencies via presentations and peer reviewed publications.

Project Significance

By developing methods to measure algae, dissolved color, and suspended sediments by remote sensing, this project will provide the means to greatly increase understanding of our water resources in Minnesota. The capabilities developed in this project will enable measurement of indicators of key ecosystem parameters (e.g. algal biomass, light regimes, thermal properties) and functions (e.g. fish productivity, contaminant reactions). Future monitoring programs using methods developed here could gain high frequency records (up to one measurement per week with normal weather conditions) on all lakes and larger river reaches in the state. This information complements detailed ground based sampling programs at a smaller number of sites that provides information that can be used to interpret patterns detected by regional analyses conducted via satellite based monitoring.

III. OVERALL PROJECT STATUS UPDATES:

Project Status as of January 25 2017:

Our major activities focused on initiation of sampling of lakes in Minnesota to collect data for use in calibrating and validating remote sensing methods to measure surface water quality (Activity 1), as well as initiation of contaminant behavior sampling (Activity 2) and outreach activities (activity 3). A total of 152 lakes were sampled during 2016, including repeat sampling of 14 sites. Sites were selected to include the major lake rich ecoregions of Minnesota, and spanned wide range of lake size. Methods for accessing and sampling were adjusted and refined to improve efficiency of sample collection and to minimize sample holding time. Preliminary assessment of satellite imagery from Landsat (8) and Sentinel (2A) sensor indicates that water quality data to relate to

satellite imagery is available for 2-3 dates for summer 2016, with additional clear-sky imagery available in autumn. The last half of the year was unusually wet and cloudy, which limited availability of clear imagery that corresponded with our field sampling. With greater satellite capacity available in 2017 (due to the expected availability of a new satellite, Sentinel 2B), more targeted sampling in response to clear imagery, and the likelihood of better conditions for image acquisition, we expect to continue construction of a robust dataset for development of remote sensing characterization of Minnesota's surface waters. Complementing our Activity 1 progress, we initiated research to relate surface water chemical composition to pollutant behavior and drinking water quality. Twenty-five large volume samples were collected for detailed chemical characterization to help inform subsequent sampling and laboratory experimentation. The post-doctoral fellow who will be central to Activity 2 was identified and hired. This scientist joined the project in January 2017. Activity 3 efforts included initiation of website enhancements and presentations.

Project Status as of July 2017:

During this period, our Activity 1 efforts focused on completion of laboratory analyses for samples collected in the previous year, methods development for satellite imagery, and initial analyses of remote sensing imagery from 2016 to measure concentrations of CDOM in MN lakes. For satellite imagery method development, we evaluated image normalization atmospheric correction methods and evaluated their performance for regional mapping of water quality. CDOM maps of Minnesota were created for 2015 Landsat 8 imagery and used to compare among alternative methods. Contaminant research (Activity 2) focused on completion of laboratory experiments in selected lake water samples to evaluate the reactivity of CDOM with chlorine to form disinfection byproducts (DBPs). The Uniform Formation Conditions (UFC) test was conducted in duplicate (or greater replication) to determine the DBP formation potential (DBFP) for 24 lake samples. The formation potentials upon chlorination were obtained for two regulated DBP classes, trihalomethanes (THMs) and haloacetic acids (HAAs). A trial-and-error approach was used to evaluate the chlorine demand for each lake sample. Analytical methods for THMs and HAAs were developed to improve the accuracy of product quantification. Bromide levels of raw lake samples were measured to evaluate the formation of brominated DBP species. Correlations of chlorine demand, THM formation potential or HAA formation potential versus various indicators of DOM concentration and composition were built. The results indicate that THMs and HAAs follow different formation patterns upon chlorination, but both have positive correlation with CDOM, DOC, UV_{254} and $SUVA_{254}$. These results will contribute to generation of statewide DBFP maps and assessments of water quality in Minnesota's lakes and rivers. Dissemination efforts (Activity 3) focused on redesign and update of the water.rs.umn.edu website for dissemination of research results and educational materials derived from the project. Project related presentations were shared at the ASLO Aquatic Sciences Meeting in Honolulu, Hawaii and the American Chemical Society (Environmental Chemistry Division) meeting in San Francisco, California.

Project Status as of January 2018:

Our Activity 1 research focused on methods development for satellite imagery and atmospheric correction methods, model development for CDOM and sampling of lakes across Minnesota to collect data for calibration and validation of remote sensing imagery needed to measure surface water quality (Activity 1). A total of 191 lakes and rivers were sampled in summer 2017. Including repeat sampling of 41 sites, a total of 273 samples collected. Many of the repeat sampling sites were re-sampled once, while 12 were sampled on a monthly basis. Sites were selected to include all major ecoregions of the state, and spanned wide range of lake size and land use. Methods were further refined to improve efficiency of sample collection and to minimize sample holding time. Preliminary assessment of satellite imagery from Landsat and Sentinel satellites indicates availability of high quality imagery corresponding to water quality sampling for the entire state during 2017. The data collected in 2017 will be used to develop and improve quantitative relationships between satellite observations and major water quality parameters (e.g., chlorophyll, total suspended solids, and turbidity) that are the focus of this project. Contaminant work (Activity 2) focused on the investigation of the role of CDOM in photodegradation of pesticides. The rate of formation ($R_{f,T}$) of triplet excited states of dissolved natural organic matter ($^3DOM^*$) was examined for 24 lake samples. One herbicide, 2,4,6-Trimethylphenol (TMP), was used as a probe compound to evaluate the formation of $^3DOM^*$. Photochemical experiments of TMP under simulated

sunlight were performed in triplicate to measure $R_{f,T}$. We initiated inhibition experiments to estimate an inhibition factor (IF) to correct the photodegradation rate of TMP at high CDOM levels. A positive correlation between CDOM and degradation rate will be built, which can be coupled with satellite CDOM assessments and contribute to the statewide assessments of pesticide photodegradation upon the impact of CDOM in Minnesota's lakes and rivers. Dissemination efforts (Activity 3) included continued updating and expansion of the water.rs.umn.edu website and several presentations, including two at the annual Water Resources Conference in October 2017.

Project Status as of July 2018:

Advances were made in all three project activities during the first half of 2018. Our Activity 1 research focused on methods development for satellite imagery and atmospheric correction, model development for water clarity, chlorophyll and CDOM. Final CDOM maps were created based on estimates of CDOM for the state's lakes, and statistics calculated at the ecoregion level for 2015 and 2016 using Landsat 8 imagery. We assessed the influence of CDOM on water clarity, and developed methods to identify water bodies highly affected by CDOM where water quality assessments based on Secchi depth (SD) measurements are biased. We presented and discussed our initial findings to MPCA. We continue working with data collected in 2017 by our group and others to develop and improve quantitative relationships between satellite observations and the water quality parameters (e.g., chlorophyll, total suspended solids, and turbidity) that are a major focus of the project. We investigated methods for Sentinel 2 imagery that work for optically complex lakes. Sentinel 2 imagery from two late summer 2017 dates was used to develop atmospheric correction methods, starting with water clarity data. We assembled available Secchi depth data from agency sources and tested atmospheric correction methods and SD models for use for both Sentinel 2 and Landsat 8 imagery. Dissemination efforts (Activity 3) included continued updates to and expansion of the water.rs.umn.edu website and the Lake Browser, and six presentations to agencies and academic institutions.

Project Status as of January 2019:

Advances were made in all three project activities during the first half of 2018. Our Activity 1 research focused on methods development for satellite imagery and atmospheric correction, model development for water clarity, chlorophyll and CDOM. Final CDOM maps were created based on estimates of CDOM for the state's lakes, and statistics calculated at the ecoregion level for 2015 and 2016 using Landsat 8 imagery. We assessed the influence of CDOM on water clarity, and developed methods to identify water bodies highly affected by CDOM where water quality assessments based on Secchi depth (SD) measurements are biased. We presented and discussed our initial findings to MPCA. We continue working with data collected in 2017 by our group and others to develop and improve quantitative relationships between satellite observations and the water quality parameters (e.g., chlorophyll, total suspended solids, and turbidity) that are a major focus of the project. We investigated methods for Sentinel 2 imagery that work for optically complex lakes. Sentinel 2 imagery from two late summer 2017 dates was used to develop atmospheric correction methods, starting with water clarity data. We assembled available Secchi depth data from agency sources and tested atmospheric correction methods and SD models for use for both Sentinel 2 and Landsat 8 imagery. Dissemination efforts (Activity 3) included continued updates to and expansion of the water.rs.umn.edu website and the Lake Browser, and six presentations to agencies and academic institutions.

Overall Project Outcomes and Results:

Water quality monitoring is essential for managing Minnesota's surface waters, maintaining the services they provide, and detecting changes caused by environmental stressors. Direct measurements of water quality are possible, however, in only a small fraction of the thousands of lakes and river miles in the state. Methods developed in this project allow use of increasingly frequent satellite observations to measure water clarity and the three key water quality indicators that control it: algae, colored dissolved organic matter (CDOM), and suspended solids. Because these parameters have distinct impacts on water quality, the ability to measure them directly across the state's waters enables comprehensive assessment of water quality status and trends and increases understanding of the causes and consequences of water quality degradation. We developed methods

to relate direct measurements of water quality to satellite imagery, assessed atmospheric correction techniques and validated methods using independent datasets. We applied these methods to measure water quality parameters on lakes >10ha, and provide the information at <https://lakes.rs.umn.edu>. Examples are included in appendix 1. Our methods extract information at seasonal to annual scales for algae, CDOM and suspended solids in lakes at state, regional, county, and watershed scales. Water quality parameters were linked to disinfection byproduct formation potential in drinking water treatment and degradation of contaminants driven by sunlight. CDOM levels were closely related to formation rates of two classes of disinfection byproducts, trihalomethanes and haloacetic acids, and to the production photo-induced reactive intermediates that degrade pesticides. This information can be used with remote sensing to assess pesticide persistence and suitability of surface waters for drinking water sources. An ongoing LCCMR project uses methods developed here with automated imagery acquisition and analysis to gather information on lake conditions at potentially a weekly basis. Project outcomes are summarized at <https://water.rs.umn.edu/> for use by researchers, managers, lake associations and the public.

IV. PROJECT ACTIVITIES AND OUTCOMES:

ACTIVITY 1: *Build advanced methods for measuring water quality in surface waters of Minnesota*

Description: Physically-based predictive relationships will be developed to determine dissolved organic matter, chlorophyll (algae), and suspended solid levels from available satellite data. The predictive relationships will be developed or “calibrated” using approximately 150 lake and river sites, plus additional ~50 sites used for validation. We will evaluate the frequency with which state or region-wide assessments of these water quality indicators are possible within a given year by assessing availability of clear imagery and testing assumptions related to efficacy of automated image processing.

Field sampling methods. To obtain sufficient calibration data to map chlorophyll, suspended solids (SS), and CDOM at the statewide scale, we will sample ~200 lakes in summer 2017 to measure their optically-related water quality characteristics nearly contemporaneously with Landsat 8 and/or Sentinel-2 image acquisitions. We will select sites to obtain a wide range of CDOM, chlorophyll and SS values, including systems where one of these variables dominates the optical (reflectance) properties of the water body and systems where various combinations of the three variables also affect reflectance. Samples will be collected across the state in proportion to the density of lakes in a given region.

We will leverage sampling to be done in 2016 on a related study to refine sampling procedures for the 2017 field season and to collect preliminary calibration data. To the extent possible, we also will use Secchi depth (SD) data from the citizen monitoring program and water quality data from ongoing monitoring programs of the MPCA to expand the database for calibration/validation of images. The usability of these ancillary data will depend on close in time the sampling was done in relation to acquisition of the satellite images we will use for the statewide mapping.

Field data collected at each site will include clarity (as Secchi depth), temperature, dissolved O₂, pH and conductivity. Samples will be collected for measurement in the laboratory of the following water quality variables: dissolved organic carbon (DOC), total suspended solids (TSS), volatile suspended solids (VSS, a measure of suspended organic matter), turbidity, color (CDOM by absorbance measurements on filtered samples), and chlorophyll *a*. SD will be measured with a standard 20-cm disk from the sunny side of the boat using the average of the depths where the disk just disappears and re-appears when raised from below. Dissolved oxygen, temperature, pH and conductivity will be measured using calibrated field meters. Water samples for dissolved organic carbon (DOC) and CDOM will be filtered through ashed glass filters (Whatman GF/F) in the field and stored in the dark in washed polycarbonate or ashed glass bottles at 4 C until analysis. Field-filtered chlorophyll *a* samples will be stored frozen in foil until analysis following acetone extraction. Water samples for TSS, VSS and turbidity will be stored on ice at collection and processed on return to the lab.

UV/Vis absorbance scans (250-700 nm) will be obtained in 1- and/or 5-cm quartz cells on filtered (Whatman GF/F) samples using a Shimadzu UV-1601PC UV-Visible spectrophotometer. Results will be used to compute α_{440} , SUVA₂₅₄, and spectral slope, *S*, over the ranges: 275-295, 350-400, and 400-460 nm. Filtered water will be analyzed for DOC using a Sievers 900 Portable Analyzer or a Shimadzu Vcpn Analyzer calibrated with dilutions of potassium hydrogen phthalate. Blanks prepared with ultrahigh purity water will be used to confirm that contamination is not occurring in sample handling. Procedures from *Standard Methods* will be used to measure TSS, VSS, turbidity, and chlorophyll.

Two state-of-the-art characterization techniques, fluorescence spectroscopy and Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR-MS), will be used to obtain information on the chemical composition of fractionated dissolved organic matter (DOM) collected from selected lakes. We will relate the compositional information with data from satellite images and results from DOM reactivity testing. Fluorescence excitation-emission matrices (EEMs) will be generated on filtered samples (Millipore 0.2 μ m nitrocellulose membrane filters) using a Horiba Aqualog fluorometer over an excitation range of 200-400 nm (5 nm intervals) and emission range of 290-550 nm (2 nm intervals). Scans will be corrected for variations in Raman scattering, lamp intensity, and cuvette imperfections. Contour plots representing the resulting matrix will be created and parallel factor analysis will be performed using reference components identified in similar work of others. FT-ICR-MS analyses will be done at the Old Dominion University Major Instrumentation Cluster, Norfolk, VA (<http://www.sci.odu.edu/sci/cosmic/index.shtml>) by electrospray ionization using a Bruker Daltonics 12 Tesla Apex Qe FT-ICR-MS following standard procedures. We have worked previously with this group to characterize DOM from wetlands; they analyze samples on a fee for service basis.

Satellite imagery processing. Cloud-free satellite imagery for Minnesota will be downloaded from the Earth Resources Observation and Science (EROS) Center operated by USGS in Sioux Falls, South Dakota. We will rely primarily on Sentinel-2, which has some advantages of image size and frequency of repeat coverage, but will also use Landsat 8 imagery. Image processing will be based on methods we developed previously.

We will evaluate several methods for atmospheric correction: the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), based on the MODerate resolution atmospheric TRANsmission (MODTRAN) v5 algorithm, and QUick Atmospheric Correction (QUAC), Atmospheric Correction for OLI 'lite' (ACOLITE), along with Landsat Surface Reflectance Data Products provided by EROS. Areas obscured by clouds and haze will be masked out during preprocessing, and images will be processed in ERDAS IMAGINE. After preprocessing, spectral data for each lake will be extracted using the signature editor and used for modeling and data analysis.

Final image processing will use procedures that produce superior results, as determined by comparison with in situ data. To use satellite imagery to develop colored organic matter, algae, and suspended solids measurements, we will need in situ lake data that represent the wide range of optical conditions found in the imagery. Measurements of CDOM, chlorophyll and TSS will be used to develop robust algorithms and determine whether they work well under all optical conditions or if different algorithms are needed for optically complex waters and also to evaluate atmospheric correction methods. If we find that different algorithms are needed, we will develop screening techniques to group water bodies into appropriate optical classes prior to applying retrieval algorithms. The maps produced for Minnesota will include over 10,000 lakes and many large rivers.

Summary Budget Information for Activity 1:

ENRTF Budget: \$ 205,000
Amount Spent: \$ 205,000
Balance: \$ 0

Outcome	Completion Date
---------	-----------------

<i>1. Measurements of algae, colored organic matter, and suspended solids in 125 selected lake and river sites to obtain a data set for developing predictive relationships</i>	<i>December 2017</i>
<i>2. Analysis of field and satellite data to develop predictive relationships to permit routine monitoring of algae, organic matter, and suspended solids in the state's waters</i>	<i>February 2018</i>
<i>3. A method for comprehensive water quality monitoring for Minnesota's 10,000 lakes</i>	<i>June 2018</i>

Activity Status as of January 25 2017:

Our research in Activity 1 since project initiation focused on sampling lakes for major water quality constituents and acquisition of complementary satellite imagery. Samples for algae, colored organic matter, and suspended solids were collected from a set of 152 sites during the summer. Fourteen sites were sampled repeatedly (3-4 dates) to assess and characterize temporal variation. Sites ranged from very small (e.g., 14 acres) to some of the largest inland lakes in the state (e.g., Mille Lacs). Analyses of chemical parameters, algal biomass, and suspended sediment are ongoing. A total of 191 samples of colored organic matter have been analyzed, including repeat sampling of selected sites. Analyses of dissolved organic carbon and total dissolved nitrogen samples are 85% complete, and all available dissolved inorganic carbon samples have been analyzed. Algal biomass and suspended sediments were not collected at a few sites owing to logistical restraints. All available suspended sediment samples and algal biomass samples will be analyzed by February 2017.

Thus far, the lakes sampled in Minnesota show a wide range of optical and chemical measurements. For instance, colored organic matter absorption at 440 nm ranges from 0.3 m⁻¹ (Six Lake in western MN) to 32.5 m⁻¹ in Itasca County's Johnson Lake. Lakes within close proximity to each other can vary quite dramatically in colored organic matter, likely based upon differences in hydrology and land use. For instance, Burntside Lake absorption measured 0.8 m⁻¹, while less than 10 miles away Blueberry Lake measured 21.6 m⁻¹. Despite this heterogeneity, lakes in the Twin Cities metropolitan region tended to be low in color while most highly colored lakes were found in the Northern Lakes and Forests ecoregion, in the northeast portion of the state.

Similarly, we found a wide range in algal biomass and total suspended solids. Chlorophyll *a*, a proxy for algal biomass, ranged from 1.4 µg/L to 99 µg/L. TSS varied from 0.2 mg/L to 24.0 mg/L. Total suspended solids tended to be correlated with algal biomass, particularly at low concentrations (R² = 0.63). This, combined with available suspended sediment data, indicates that algal biomass is likely the primary control on the amount of total suspended solids and that mineral suspended sediments are a minor contributor in most lakes included in the study, to date. Exceptions to this likely are related to re-suspension of bottom sediments in lakes sampled during stormy or windy conditions. More comprehensive analysis will be performed upon completion of all laboratory analyses.

This dataset of lake chemical and optical properties provides a rich database upon which to develop remote sensing models for monitoring Minnesota's 10,000 lakes. The year 2016 was an unusual climatic year – indeed, it was the wettest year on record for the Twin Cities, and August was the wettest month of the year. The timing of rain events and overall cloudiness of the summer and early autumn overlapped with much of our sampling activities. Late summer and early fall were targeted for our most intensive sampling efforts because of the seasonal timing of algal blooms and because the historical satellite record indicates that this is an optimal period for acquisition of imagery. Despite the high cloud cover in 2016, there are several Landsat 8 and Sentinel 2A images that can be used for model development. We anticipate many more opportunities for imagery in 2017 with Sentinel 2A at full operation (images every 10 days vs. every 20 in 2016), launch of Sentinel 2B (March 2017), and the likelihood of more favorable climatic conditions.

Activity Status as of July 2017:

Our activity 1 research focused on methods development for satellite imagery, model development for CDOM and completing analysis of 2016 water quality analyses and preparation for the 2017 field season. During the late spring, we identified 100+ lakes to target for sampling during the field season that would include a large variety of watershed types and lake characteristics across the state. Undergraduate technicians were hired for

summer research activities. Initial sampling in May and June targeted 12 lakes that would be sampled on a monthly basis throughout the summer, as well as opportunistic sampling of other sites. Data collected in 2015 were used for initial development of relationships between field and satellite based measurement of CDOM. We evaluated one normalization (Radiometric Rectification (RR)) and three atmospheric correction methods (EROS Surface Reflectance v. 2.2 (SR), ACOLITE, FLAASH using five paths of Landsat 8 imagery (Figure 1). The RR and SR

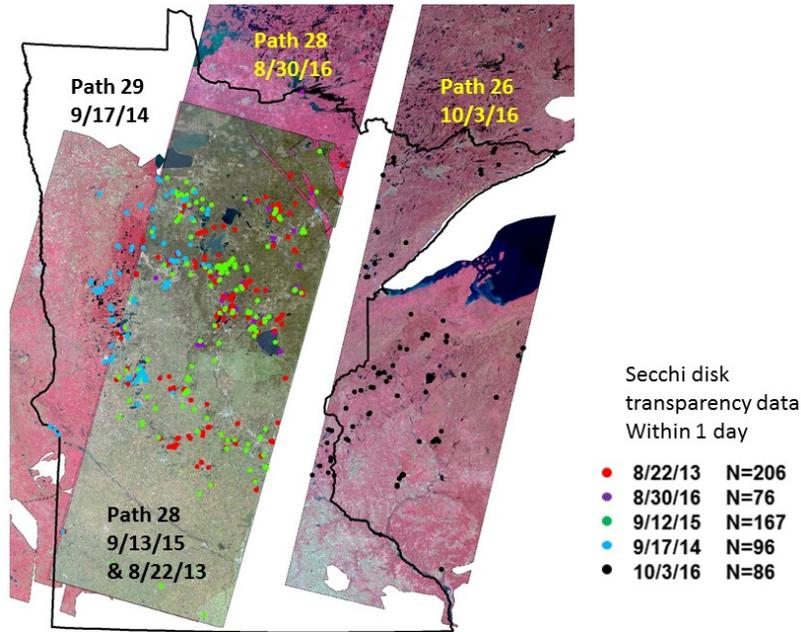


Figure 1. Landsat 8 imagery and available Secchi disk transparency data used to evaluate atmospheric correction and normalization techniques for regional water quality mapping.

methods both showed promise as consistent normalization techniques for regional mapping of water quality. CDOM maps of lakes in Minnesota were created for 2015 using Landsat 8 imagery using the RR and SR corrected imagery and compared for consistency. Both produced reasonable results, with the SR method less labor intensive and amendable to automated image processing methods. Over the next period we will develop a more robust model using CDOM data from 2015 and 2016 and apply the same universal

model to map CDOM across the state.

Activity Status as of January 2018:

Our Activity 1 research focused on methods development for satellite imagery, model development for CDOM and extensive field sampling and laboratory analyses of samples collected, and access and integration of data collected by partners. During the summer and early fall we sampled 191 lakes for major water quality parameters needed to develop algorithms to relate satellite imagery to concentrations of algae, colored organic matter, and suspended sediments. Sampling was done with a UMN group of students and staff, and via collaboration and coordination with other groups. These included the Itasca Biological Station, the Red Lake Nation, DNR, the Minneapolis Parks Board, Minnesota Science Museum St. Croix Research Lab, and citizen volunteers from the Brainerd and Lake Vermillion areas to expand our ability to collection data from a large number of diverse lakes. We coordinated with MPCA in order to leverage data collections for our projects. The large majority of samples have been analyzed, and we are in the final stages of review for use. The best satellite imagery from 2017 has been acquired and stored for use. Once field data are finalized (in January 2018) we will begin to perform the necessary analyses needed to develop and improve methods to measure algae, colored organic matter, and suspended sediments across all the states' lakes.

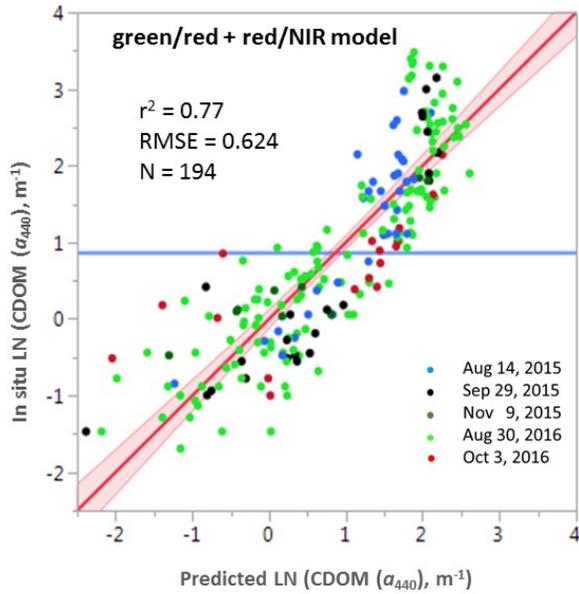


Figure 2. Landsat 8 EROS Surface Reflectance data from multiple dates used to create a universal model that was applied to 2015 and 2016 Landsat 8 imagery to create CDOM maps.

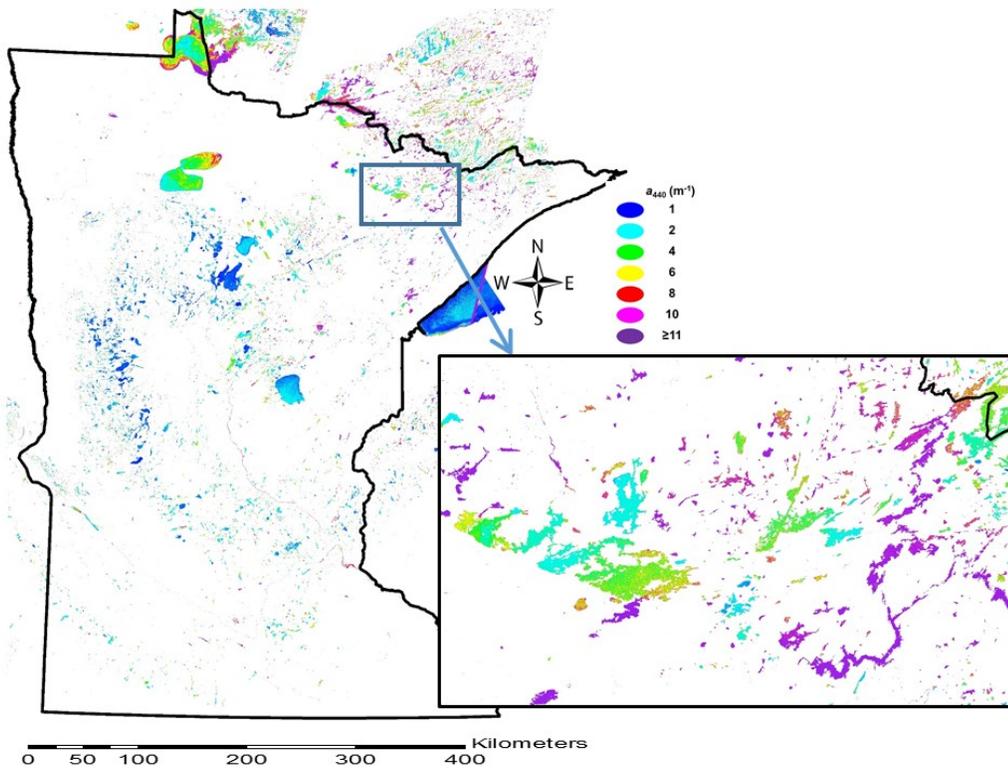


Figure 3. 2015 CDOM map for Minnesota created using the universal two-variable model (green/red + red/NIR), $a_{440} = -5.478 \cdot OLI3/OLI4 - 0.633 \cdot OLI4/OLI5 + 8.135$ for SR imagery.

We evaluated four additional atmospheric correction methods (EROS Analysis Ready Data SR (ARD), SeaDAS l2gen, Modified atmospheric correction for inland waters (MAIN) and ESA Case-2 Regional/Coast Color (C2RCC) using five paths of Landsat 8 imagery (Figure 1). The SR method that showed promise in the previous period was used to create a more robust CDOM model from 2015 and 2016 in situ data to create a universal model (Figure 2). The universal model was applied 2015 and 2016 Landsat 8 SR imagery to create 2015 and 2016 CDOM maps (Figures 3 and 4 respectively).

Activity 2 requires minor revision of the target dates for completion of Outcomes 2 and 3, to June 2018 and August 2018, respectively. These changes are the result of three factors: acquisition of more samples and data than planned, shifts in personnel effort to accommodate other projects, and the renovation of Finlay's laboratory, which slowed some analyses this past fall. Despite these adjustments, Activity 2 and the project as a whole will remain on track.

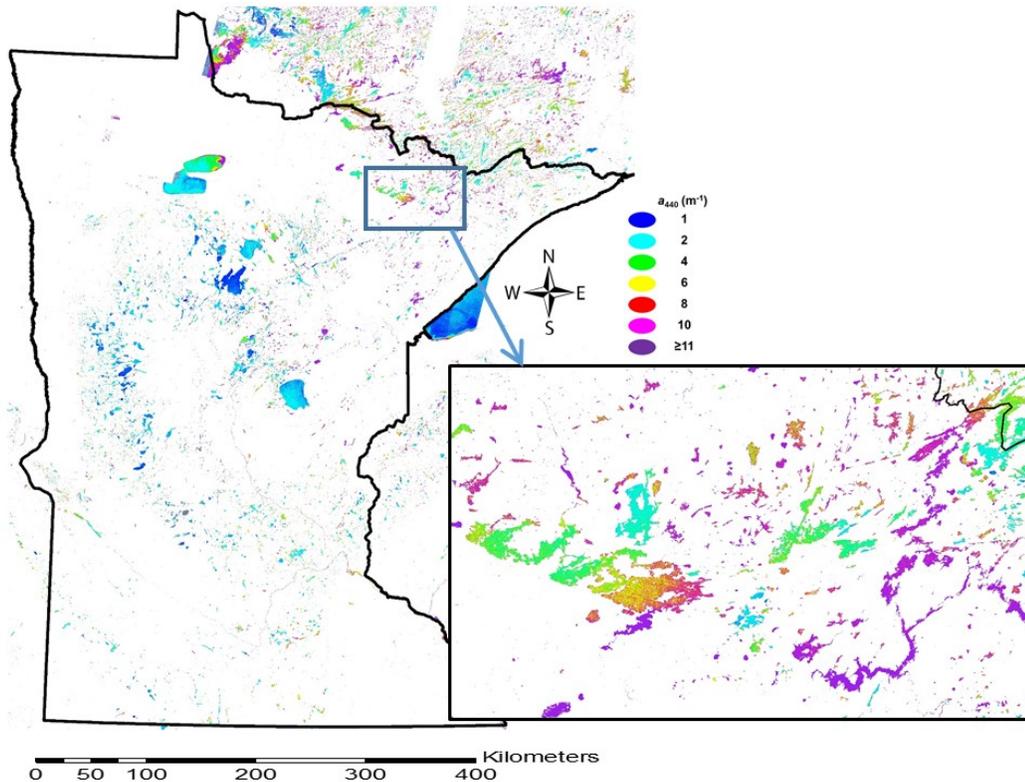


Figure 4. 2016 CDOM map for Minnesota created using the universal two-variable model (green/red + red/NIR), $a_{440} = -5.478 * OLI3 / OLI4 - 0.633 * OLI4 / OLI5 + 8.135$ for SR imagery.

Activity Status as of July 2018:

The EROS SR method that showed earlier promise was used to create an improved CDOM model based on 2015 and 2016 in situ data that could be applied across atmospherically corrected images (hereafter referred to as a "universal model"). This model was applied to 2015 and 2016 Landsat 8 SR imagery to create final 2015 and 2016 maps that delineate CDOM levels in lakes across the state (Figures 5 and 6, respectively).

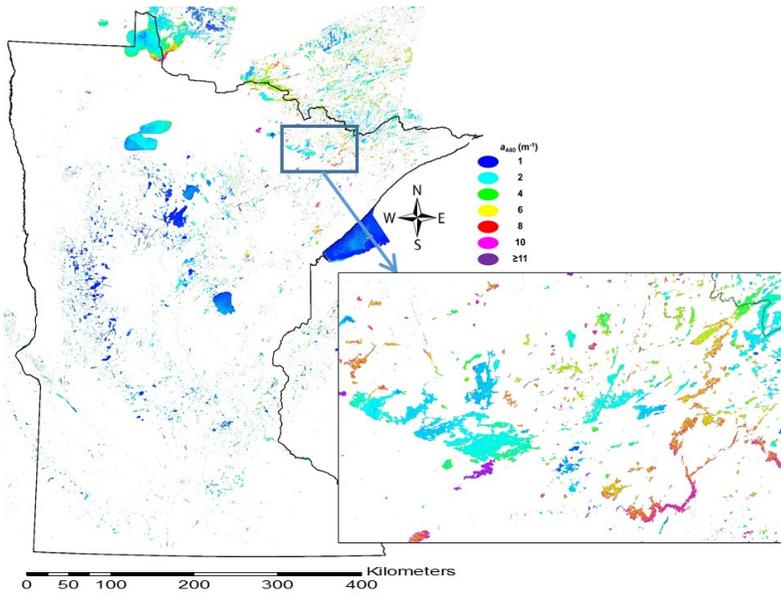


Figure 5. Final 2015 CDOM map for Minnesota created using the two-variable universal model (green/red + red/NIR), $a_{440} = -5.478 \cdot OLI3/OLI4 - 0.633 \cdot OLI4/OLI5 + 8.135$ for Landsat 8 SR imagery.

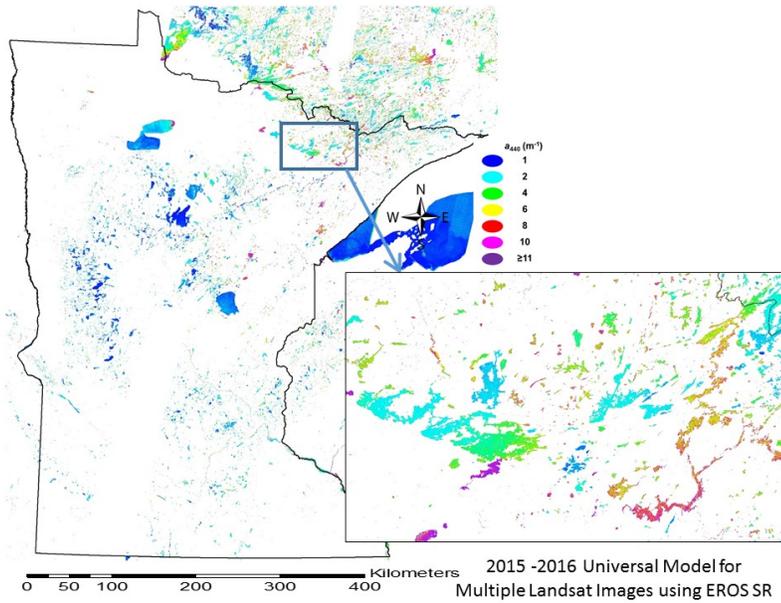


Figure 6. Final 2016 CDOM map for Minnesota created using the two-variable universal model (green/red + red/NIR), $a_{440} = -5.478 \cdot OLI3/OLI4 - 0.633 \cdot OLI4/OLI5 + 8.135$ for Landsat 8 SR imagery.

We are using data for CDOM distributions across the state to understand implications for water quality and ecosystems. A recent manuscript, accepted pending minor revisions, establishes relationships between CDOM and dissolved organic carbon (DOC), an important controlling variable in freshwater aquatic ecosystems. These analyses show that in many northern Minnesota lakes, where dissolved organic matter comes primarily from terrestrial sources, CDOM is a reliable proxy for DOC, and satellite remote sensing can thus be used to estimate both parameters across broad spatial scales (Griffin et al., *accepted, July 2018*). DOC and CDOM are not strongly correlated in lakes where phytoplankton or macrophytes produce significant amounts of organic matter,

however, as is often the case in central and southern Minnesota. Satellite remote sensing is useful for monitoring CDOM in these lakes, but resulting DOC estimates are not as reliable.

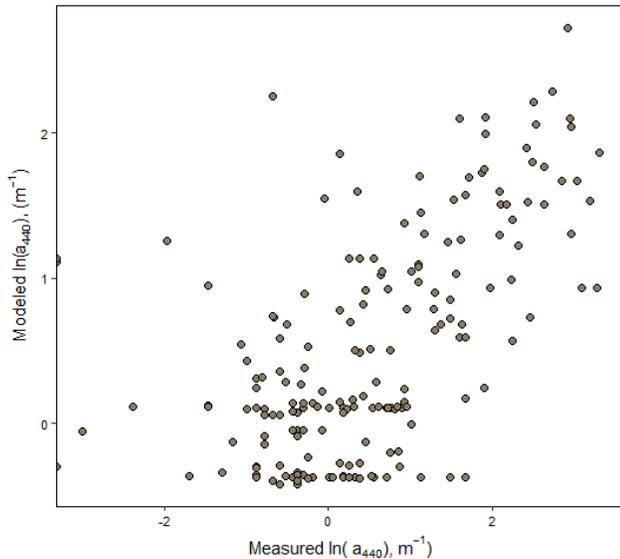


Figure 7: Variability in CDOM measured in lakes across the state was modeled, using watershed forested wetlands and lake area.

We are currently assessing landscape controls on CDOM using the data collected through Activity 1 and ancillary GIS data. Additional watershed data come from the MN DNR Hydrography and Watershed data suites, and land use data were extracted from the University of Minnesota Remote Sensing and Geospatial Analysis Laboratory's 2013 Minnesota Land Cover dataset. Initial analyses have focused on modeling CDOM from field-sampled lakes and headwater systems. The percentage of forested wetlands within the watershed is the largest driver of CDOM in these initial explorations, explaining 25-35% of variability in CDOM. For instance, Figure 7 shows field-measured $\ln(a_{440})$ versus CDOM across all of MN modeled using forested wetland cover and lake surface area, which together explain 40% of CDOM

variance. Other potential controls were tested but not significant in regression modeling, including lake area to basin ratios, percent watershed agriculture or impervious surface, and Strahler stream order. We are continuing to assess watershed controls on CDOM, using satellite-derived estimates from all lakes larger than 15 acres. Major regions in the state may differ markedly in terms of the landscape characteristics that most affect CDOM, and so these processes are being assessed on the scale of major watersheds (e.g., St. Louis River or Minnesota River watersheds), as well as on a statewide basis.

We developed atmospheric correction methods for use with both Sentinel-2 and Landsat 8 imagery. A promising method, the Modified Atmospheric correction for INland waters (MAIN), was evaluated for water clarity (Secchi depth, SD) during this period. We used SD data and Sentinel 2 imagery from August 23, 2017 (Figure 8a) that had some cloud cover but was clear for a large portion of Minnesota and an overlapping September 12, 2017 image (Figure 8b) that was clear throughout the state. In-situ SD data (hereafter referred to as SD_f , where subscript f stands for field) collected from the Citizen Lake Monitoring Program (CLMP) for calibration included 504 and 249 measurements within ± 3 days of the clear portions coinciding with the Sept. 12 and Aug. 23 images, respectively (total of 753 measurements across the Western Cornbelt Plains (WCBP), North Central Hardwood Forests (NCHF), and Northern Lakes and Forests (NLF) ecoregions). The SD_f point locations were converted to a polygon shapefile with a 50 m circular buffer around the centroid. Corresponding mean remote sensing reflectance values (R_{rs}) of pixels within the polygons were extracted and tabulated.

An important factor influencing the strength and reliability of ground-satellite relationships for water clarity is the size of the "time window" between ground and satellite observations (Kloiber, et al., 2002). To address this issue, regression statistics were calculated using the relationship between $\ln(SD_f)$ and the ratio of the Sentinel-2 R_{rs} values for the red and blue bands centered around 664 and 497 nm, designated bands B4 and B2, respectively. These bands have been used in all our previous studies for retrieval of SD data from satellite

imagery. We successively increased the time window between satellite overpass and ground observations and found that data collected within one day of the image provided the best R^2 and RMSE while still maintaining a large calibration dataset. This resulted in a final dataset for further analysis of 365 SD_f values with satellite imagery data (Figure 8(c-d)).

We used multi-stepwise linear regression for water clarity model development with log-transformed SD, $\ln(SD_f)$, as the dependent variable and Sentinel-2 R_{rs} values for bands B1-B8A (443 nm to 865 nm). All Sentinel-2 band ratio permutations were considered as independent variables using JMPPro 14. Models also were developed for simulated Landsat 8 bands based on the Sentinel-2 bands corresponding to those present on the Landsat 8-OLI sensor) by excluding the three red-edge (B5-B7) and one NIR band (B8) not available on the OLI sensor to establish a robust water clarity model when both sensors are used for time series analysis.

We examined many two-term regression models for the Sentinel-2 imagery, and the best model generated from the combined SD dataset has the form:

$$\ln(SD_{S-2}) = a(R_{rs}(B2)/R_{rs}(B4)) + b(R_{rs}(B5)) + c$$

where coefficients a , b , and c were fit to the calibration data by regression analysis, $\ln(SD_{S2})$ is the natural logarithm of the Sentinel-2-derived SD, and B2, B4, and B5 are the Sentinel-2 blue, red and NIR bands. From the combined dataset, the $\ln(SD_{S2})$ prediction model generated a high R^2 of 0.81, an RMSE of 0.360 m and $p < 0.0001$. The band ratio term $(R_{rs}(B2)/R_{rs}(B4))$ was found to contribute the most to estimating $\ln(SD_f)$ from the 365 inland water bodies.

Three additional models that also had reasonable fits ($R^2 > 0.79$) included the same band ratio as in the above equation as the most significant predictor. Interestingly, one regression generated a two-term model containing the NIR band centered at 842 nm ($R^2 = 0.79$, RMSE = 0.377) within the band ratio $R_{rs}(B3)/R_{rs}(B6)$ ratio, further supporting the usefulness of NIR reflectance to assess water quality characteristics in optically complex waters.

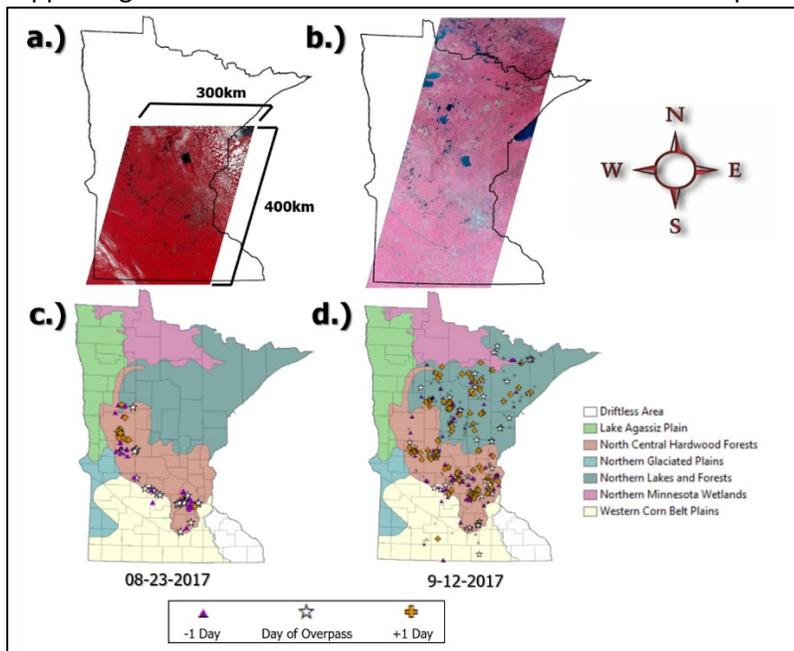


Figure 8. False-color Sentinel-2 image composites (R/G/B: B8A/B5/B4) of Minnesota on (a) August 23, 2017 and (b) September 12, 2017. (c-d) Corresponding time-window qualified SD_f sample locations (± 1 day within satellite overpass, $n = 365$) collected by CLMP across the WCBP, NCHF, and NLF ecoregions.

An additional $\ln(\text{SD})$ model was developed to enable subsequent cross-sensor comparisons between Landsat and Sentinel. This model was based on Sentinel-2 imagery using only Landsat-8 simulated bands as independent variables ($\ln(\text{SD}_{sL8})$). Using a similar two-term linear regression analysis, the following equation was generated:

$$\ln\text{SD}_{sL8} = a(R_{rs}(B2)/R_{rs}(B4)) + b(R_{rs}(B2)/R_{rs}(B3))$$

As expected, the B2/B4 (blue/red) band ratio also was the highest contributing factor in predicting $\ln\text{SD}_f$. The $\ln\text{SD}_{sL8}$ model generated an R^2 of 0.79, RMSE of 0.38 m, and $p < 0.0001$. A visual comparison of SD maps from the best performing RMSE two-term Sentinel-2 and simulated Landsat 8 models is shown in Figure 9.

We are currently exploring models to estimate algal biomass (i.e. chlorophyll) levels in Minnesota's lakes. Initial results show some promising models for chlorophyll that take advantage of the Sentinel-2 red edge (705 nm) band. These analyses will be completed in 2018; assembly of water quality data for suspended solids and turbidity analyses is ongoing.

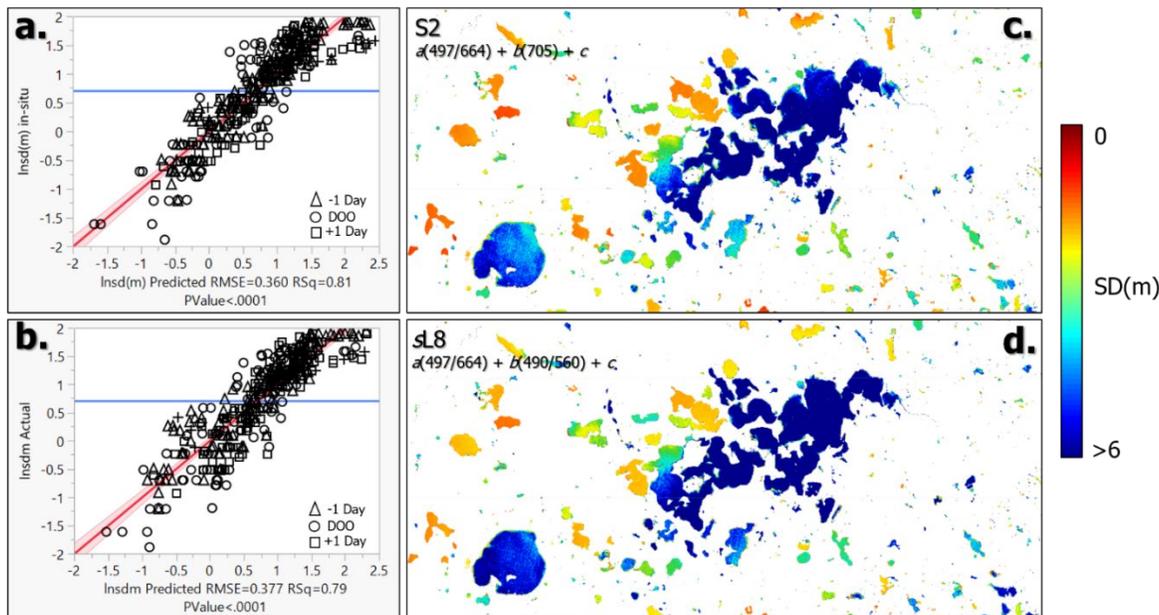


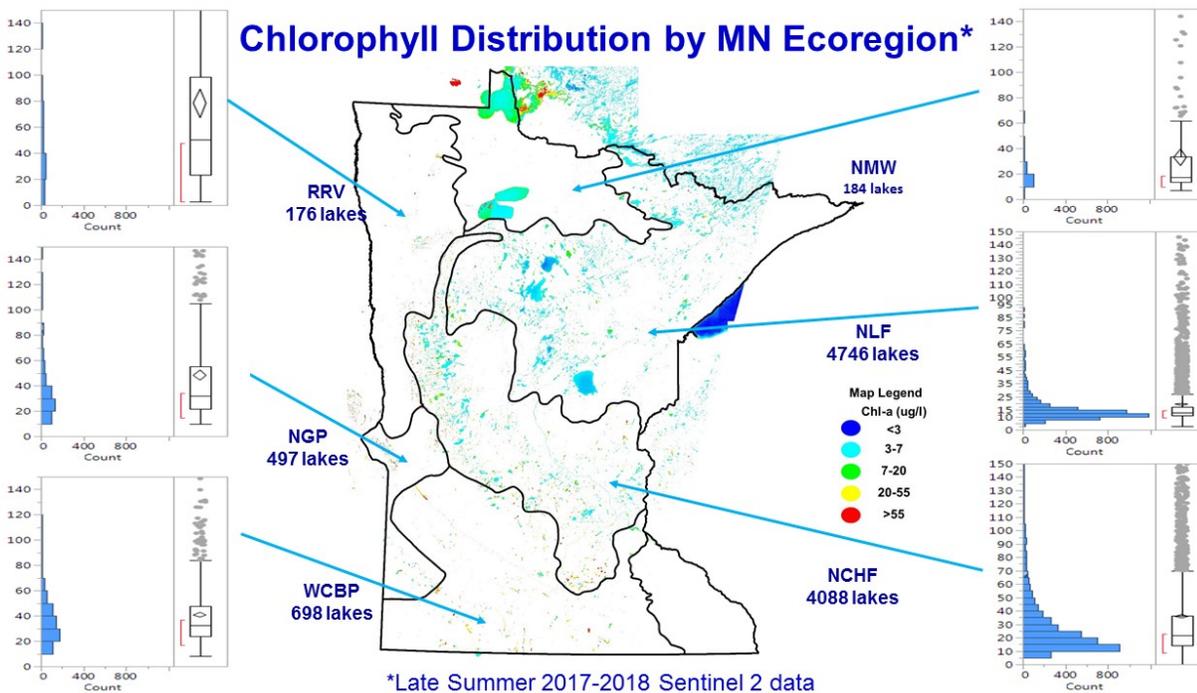
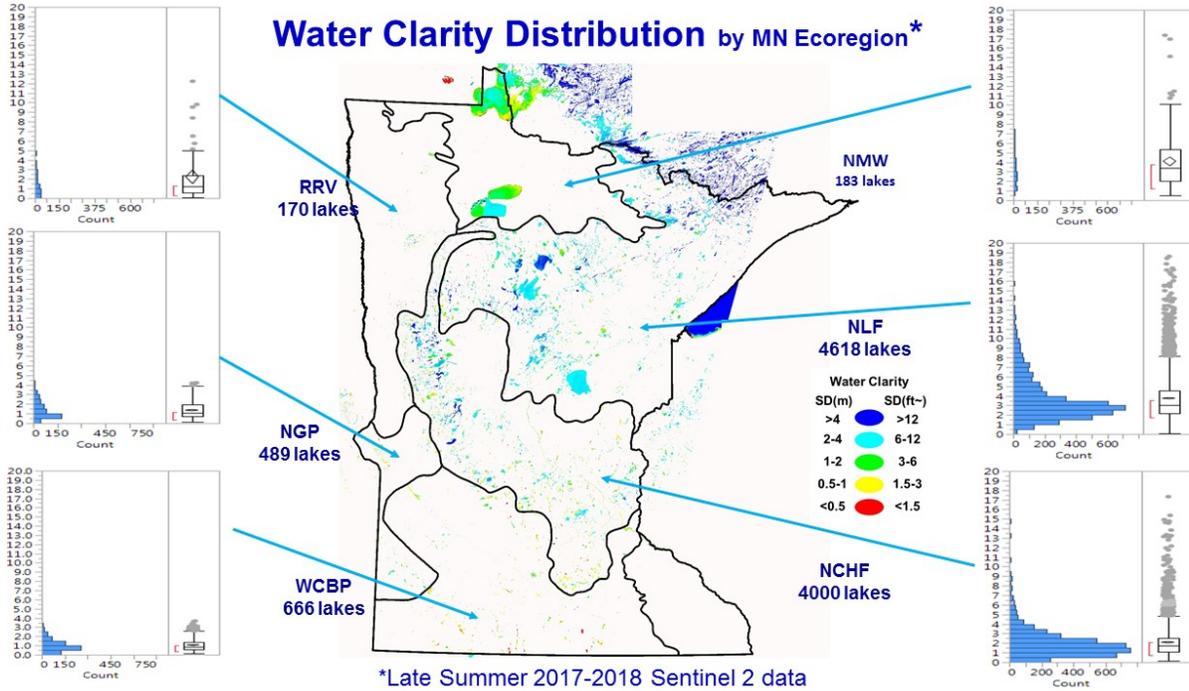
Figure 9. Average regression plots for (a) $\ln\text{SD}_{\text{Sentinel 2 (S2)}}$ and (b) $\ln\text{SD}_{\text{simulated Landsat 8 (sL8)}}$ from the calibration dataset collected on Aug. 23 and Sept. 12. DOO = Day of overpass (c-d) Corresponding pseudo-color SD_{S2} and SD_{sL8} maps of Lake Minnetonka, MN and surrounding water bodies.

Activity Status as of January 2019:

During the summer of 2018, we sampled 30 lakes to expand our dataset for our remote sensing of TSS, chl- a , and CDOM from Sentinel 2 imagery. Sampling included lakes in the Twin Cities Metro Area and nearby agriculturally dominated regions. Sampling was coordinated to correspond with clear-sky imagery from August 30th and September 23rd. Lakes ranged from clear, with few algal blooms (chl- a of 2.46 $\mu\text{g/L}$ in Square Lake) to hypereutrophic (172 $\mu\text{g/L}$ in Elysian Lake).

Using remote sensing advances made in this and another LCCMR project, we estimated CDOM concentration and water volume for 12,000 lakes. Data for lake volume from an LCCMR project led by Professor John Nieber was combined with CDOM to estimate the amount of terrestrial DOC (tDOC) in Minnesota lakes. CDOM and tDOC pool sizes are important because they influence nutrient and contaminant cycling and greenhouse gas dynamics. Our estimates indicate that 1.1 Tg C of DOC is stored in the water column of Minnesota lakes. To put this in perspective, the Mississippi River annually discharges about 1.9 Tg C to the Gulf of Mexico. We plan to submit a related manuscript for publication in early 2019

We developed a preliminary chlorophyll map using 2017 and 2018 Sentinel 2 imagery to cover the entire state for the standard late summer index period (July 15 to Sept 15). Sentinel imagery from September 12, 2017 and associated field data were used to develop quantitative models for chlorophyll (R^2 of 0.69, RMSE of 0.646, and $p < 0.0001$). This model and the water clarity model previously developed (Figure 9) were applied to Sentinel-2 imagery for Sept 12, 2017 and August 21, 2018 to create chlorophyll and water clarity maps for the entire state. Using these and the CDOM maps created earlier, lake-average distributions of water clarity, chlorophyll and CDOM were calculated for each ecoregion (Figure 11, respectively). Validation of models and final statewide maps will be completed for all water quality variables during the final stages of the project.



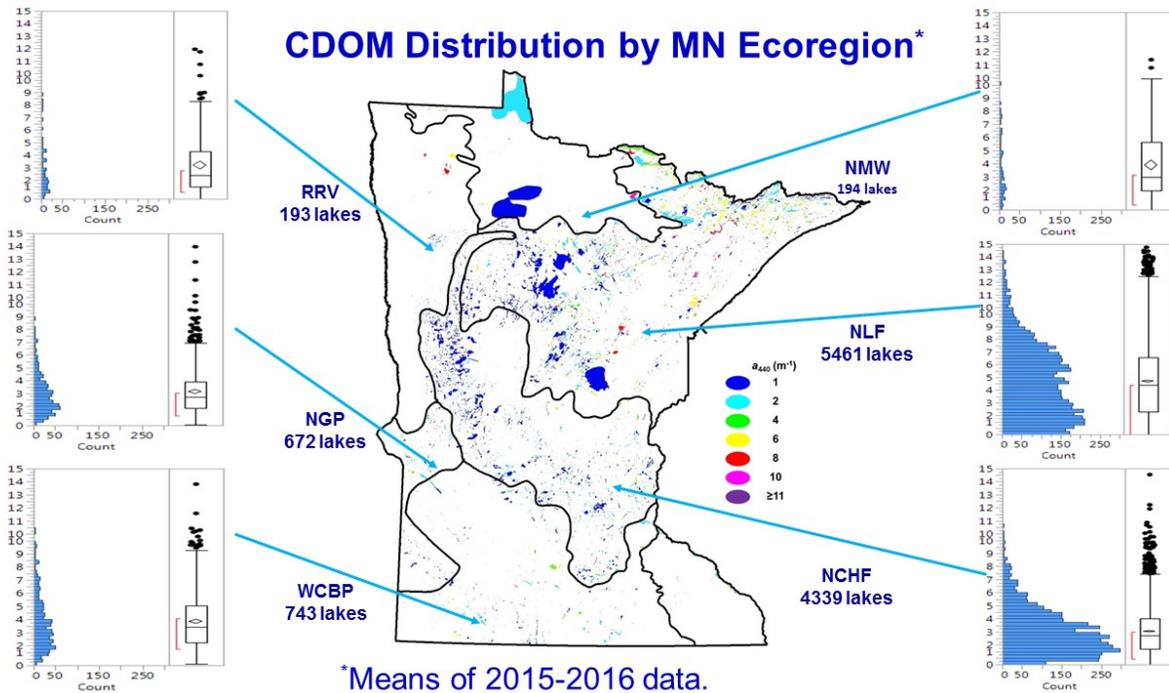


Figure 11 Statewide maps of water clarity, chlorophyll a , and CDOM for all lakes including summaries by ecoregion using remote sensing methods developed in this project. Landsat 8 was used for the CDOM map.

In collaboration with colleagues from the MPCA, we assembled a large database from our 2014-2017 ground-based sampling program and the MPCA's 2015-2017 lake monitoring program and used it to analyze interrelationships among the principal water quality variables that affect Secchi depth (SD) measurements. These variables are chlorophyll a , the main measure of algal abundance in lakes, total suspended matter (TSM), which reflects the abundance of both algae and mineral particles (clays, soil particles), and CDOM, the colored component of natural dissolved organic matter. The analysis showed clearly the conditions under which water clarity, as measured by SD, is controlled by one or several of the three driving variables. SD is one of the most common measurements on lakes and is widely used as a metric of lake trophic conditions. The MPCA uses SD to assess water quality impairment problems caused by eutrophication, and the agency established ecoregion-based SD criteria to aid in their decision-making on needs for watershed and lake interventions to manage such issues. For example, in the NLF ecoregion, the MPCA uses an SD criterion of 2 meters (m) for warm- and cool-water lakes; SD values below this criterion are assumed to indicate eutrophication-caused impairment relative to designated uses, such as body-contact recreation. Our analysis showed that many NLF lakes have SD values less than 2 m because of high CDOM levels, even though chlorophyll concentrations are low and the lakes do not suffer impairment due to eutrophication. We established guidelines for CDOM levels that render use of the 2-m SD criterion inapplicable for eutrophication assessments. In collaboration with MPCA scientists, we wrote a scientific paper documenting our findings. The paper was favorably reviewed for publication in the journal *Ecological Applications* and was recently accepted for publication in 2019.

Final Report Summary:

Development of methods for statewide assessments of water quality using new satellite capacity required three major components. First, we collected field data concurrently with high quality satellite imagery to relate optical image signals to water quality parameters. Second, we analyzed image processing methods needed to normalize imagery. Third, we developed and tested robust methods to use satellite imagery to measure water quality indicators. Field data collections (#1) were largely described previously and are briefly summarized here. Our data collection efforts were supplemented by collaboration and coordination with MPCA, citizen volunteers, and other groups. For CDOM, 250 measurements were coincident within 30 days of clear Landsat 8 and 134 for clear

Sentinel 2 imagery. Because of the relative stability of CDOM over short periods (Brezonik et al. 2015), these data were suitable for model calibration and validation. The other water quality variables (chlorophyll, TSS and turbidity) are commonly measured throughout the state but are more temporally variable. Because of this variability, only data collected within one day of the image date was used for calibration and validation. For chlorophyll, 213 measurements within one day of clear Sentinel-2 imagery were available from August 23, 2017 and September 12, 2017. For TSS and turbidity, 46 and 75 measurements, respectively, were available.

Our work on image processing (#2) resulted in major advancements in development of reliable methods for atmospheric correction (AC), which is needed to normalize imagery, which is applicable to both the Landsat L8/OLI and Sentinel S2/MSI sensor systems, and also is adaptable to future imagery from multiple sensor types. In our investigation of atmospheric correction methods, we discovered important differences among correction methods and between the specific versions of methods used. Many of these methods are under active development, and although modifications tend to improve performance for water quality sensing, they make models developed from previous version obsolete. This is a large problem relative to our goals for long term monitoring and automated image processing (discussed below). To address this issue, we evaluated the most widely used atmospheric correction methods, including the most recent versions of EROS SR, EROS Analysis Ready Data SR (ARD), SeaDAS I2gen, Modified atmospheric correction for inland waters (MAIN), ATmospheric REMoval (ATREM), ACOLITE (AC) and ESA Case-2 Regional/Coast Color (C2RCC) using three Minnesota paths of Landsat 8 imagery. Although all the evaluated methods could be used for water quality assessments, the EROS SR and MAIN discussed in the previous period performed best. MAIN has the added capacity for harmonization of remote sensing data from multiple sensors (Page et al. 2019). Following careful evaluation of these methods, we concluded that MAIN is best suited for present and future needs. The EROS SR AC method that showed promise in earlier periods was modified by EROS and is different from the version that we used to create a CDOM model and the 2015 and 2016 CDOM maps described previously. Because of the likelihood of further changes to EROS methods, we selected MAIN for all atmospheric correction tasks associated with image processing. For consistency, we reanalyzed earlier CDOM models with MAIN atmospheric correction.

The third major component of our work was development of models to quantitatively relate corrected satellite imagery to key surface water quality variables. We used a semi-empirical approach to develop robust algorithms for CDOM, chlorophyll, TSS and turbidity using field data described earlier. To explore the potential of Landsat 8 and Sentinel 2 to measure water quality variables we implemented a bootstrap forest technique (JMP®, Version 14. SAS Institute Inc., Cary, NC) that determined the most robust bands and band-ratio combinations to model the water quality variables. The calibration water quality dataset that corresponded with clear Landsat 8 or Sentinel 2 imagery was used as the dependent variable and MAIN-derived mean R_{rs} values from bands B1-B5 for Landsat 8 and B1-B8A for Sentinel 2 (Table 1) and all band-ratio permutations as independent variables (26 and 81 total terms, respectively).

Table 1. Sensor characteristics for L8/OLI and S2/MSI, including band center, bandwidth, spatial resolution and signal-to-noise ratio (SNR). SNR values have been scaled for radiances observed over clear coastal waters (adopted from Pahlevan et al., 2017).

<u>Landsat-8/OLI</u>											
Band ID	B1	B2	B3	B4	-	-	-	-	B5	B6	B7
Band center (nm)	443	482	561	655	-	-	-	-	865	1609	2201
Bandwidth (nm)	20	65	60	40	-	-	-	-	30	85	190
Resolution (m)	30	30	30	30	-	-	-	-	30	30	30
Signal-to-Noise Ratio (SNR)	284	321	223	113	-	-	-	-	45	10.1	7.4
<u>Sentinel-2/MSI</u>											
Band ID	B1	B2	B3	B4	B5	B6	B7	B8	B8A	B11	B12
Band center (nm)	444	497	560	664	705	740	783	842	865	1610	2190
Bandwidth (nm)	20	55	35	30	15	15	15	15	20	9	175

Resolution (m)	60	10	10	10	20	20	20	10	20	20	20
Signal-to-Noise Ratio (SNR)	439	102	79	45	45	34	26	20	16	2.8	2.2

The bootstrap forest technique uses decision trees to associate imagery data (R_{rs} data) with in situ (lake) data, chosen randomly to determine the terms that predict a lake variable based on the highest total sum of squares (SSTO) (Hastie et al., 2009). Prediction consistency of the bootstrap decision for each term was evaluated by splitting the samples into training (70%) and validation (30%) datasets and running 10,000 iterations. Only the two highest contributing terms that produced the highest coefficient of determination (R^2) with in situ data were used to develop models using the Landsat 8 and Sentinel 2 (using only the bands coincident with Landsat 8) (5 bands, 443–865 nm). A second model was developed for Sentinel 2 using all available bands (9 bands, 443–865 nm), including the red-edge bands (near 705 nm) that improved measurement of chlorophyll in other studies. Overall model accuracy was assessed based on how well the model forecast in situ water quality parameters.

After a review of the previous literature on satellite-based models for the water quality parameters and further exploration of various two-term regression models using Landsat 8 and/or Sentinel 2- R_{rs} imagery for CDOM, chlorophyll, TSS and turbidity, we identified the best models using the datasets described above. The two-term models took the form:

$$\text{Ln}(\text{water quality variable}) = a(\text{term 1}) + b(\text{term 2}) + c$$

where term 1 and term 2 represent R_{rs} data for various bands or band ratios; coefficients a , b , and c were fit to the calibration data by regression analysis; and $\text{Ln}(\text{water quality variable})$ is the natural logarithm of the satellite-derived water quality variable for a given pixel.

We also developed one term models for turbidity that had the form:

$$\text{Ln}(\text{water quality variable}) = a(\text{term 1}) + c$$

Model development. Model terms and statistics for the five focal water quality variables are summarized in Table 2. For water clarity (measured as Secchi depth, SD), CDOM (measured as a_{440}) and TSS, the Landsat 8 bands worked almost as well as the Sentinel 2 bands (even better in the case of a_{440}). Although the difference in R^2 between the Landsat 8 and Sentinel-2 models for CDOM (a_{440}) (0.85 and 0.73, respectively) could be partly due to differences in the datasets used for the two models, it likely is largely due to the lower signal-to-noise ratio of the Sentinel-2 MSI sensor for the bands used in the a_{440} model (Page et al. 2019 & Table 1). The differences in fit between the models for Sentinel-2 and models developed with only the Sentinel 2/Landsat 8 coincident bands (Table 1) for chl a and to a lesser degree turbidity and CDOM indicates that the additional spectral bands in the MSI sensor improve the relationships for chl a, turbidity and CDOM. When the difference in signal-to-noise ratio is taken into account, Landsat 8 is well suited for all of these water quality variables except possibly chl a. The red edge bands (at 705 and 740 nm) of Sentinel-2 are well suited for all these water quality variables due to the improved spectral bands although the lower signal to noise ratio of this sensor may decrease the relationship.

Table 2. Water quality model terms and statistics

Water quality variable	Sensor	Term 1 ¹	Term 2 ¹	R ²	R ² range	RMSE	RMSE range	N
Water Clarity	Landsat 8	655/490	655	0.78	0.78-0.79	0.352	0.346-0.360	507
Water Clarity	Sentinel 2	655/490	705	0.79	0.79-0.80	0.337	0.330-0.342	348
Water Clarity	S2_LS	655/490	560	0.78	0.78-0.79	0.348	0.339-0.355	348
a440	Landsat 8	655/560	865/560	0.85	0.85-0.86	0.487	0.479-0.493	250
a440	Sentinel 2	655/560	740/560	0.76	0.74-0.79	0.628	0.589-0.654	134
a440	S2_LS	655/560	865/560	0.73	0.71-0.76	0.664	0.626-0.683	134
Chl a	Sentinel 2	490/705	665/560	0.78	0.76-0.80	0.569	0.551-0.593	213
Chl a	S2_LS	490/665	665/560	0.72	0.69-0.75	0.641	0.614-0.655	213
TSS	Sentinel 2	490/655	740/842	0.84		0.486		46
TSS	S2_LS	490/665	665/842	0.83		0.503		46
Turbidity	Sentinel 2	705/560		0.84		0.480		75
Turbidity	S2_LS	490/665		0.80		0.532		75

¹ Band center wavelength in nm and listed as Landsat 8 band centers for coincident sentinel 2 bands

Together, these analyses show strong relationships that can be used to measure and map water quality parameters for all lake conditions in the state. Initial results of this approach are displayed on the LakeBrowser (<https://lakes.rs.umn.edu/>) for water clarity (SD), CDOM (a_{440}) and chlorophyll a. In addition, more extensive data for the initial findings for TSS and turbidity are displayed on our web site (<http://water.rs.umn.edu>).

Outcomes for Activity 1, in particular the MAIN-based models for water clarity, CDOM, chlorophyll, TSS and turbidity, and other advances made by our project will be implemented for routine water quality monitoring across Minnesota by the ongoing LCCMR project *Providing Critical Water-Quality Information for Lake Management* (J. Peterson, PI). Our completed project developed the basis for an automated monitoring system that uses combined information from both Landsat and Sentinel satellite imagery. The increased availability of Sentinel and Landsat imagery, along with cloud-based and supercomputing capabilities, enables the development of automated routines using the improved spectral, spatial, radiometric and temporal resolution of these satellites for water quality monitoring and fisheries management. This new system is currently in testing phase; once complete, it will be implemented for the whole state. All water quality data produced has been made available in the Lake Browser as described in Activity 3.

Literature cited

- Brezonik, P. L., L. G. Olmanson, J. C. Finlay, and M. E. Bauer. 2015. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sens. Environ.* 157: 199-215. DOI: 10.1016/j.rse.2014.04.033.
- Hastie, T. J., Tibshirani, R. J., and Friedman, J. H. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer-Verlag.
- Olmanson, L., Brezonik, P., and Bauer, M., 2015. Remote sensing for regional lake water quality assessment: capabilities and limitations of current and upcoming satellite systems. In *Advances in Watershed Science and Assessment*, 111-140. doi.org/10.1007/978-3-319-14212-8_5
- Pahlevan, N., Sarkar, S., Franz, B.A., Balasubramanian, S.V., and He, J., 2017. Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. *Remote Sens. Environ.* 201, 47-56. doi.org/10.1016/j.rse.2017.08.033.
- Page, B.P. and Olmanson, L.G., and Mishra, D. R., 2019. A harmonized image processing workflow using Landsat-8 and Sentinel-2 for mapping water clarity in optically complex lake systems. *Remote Sens. Environ.*

ACTIVITY 2: Relate surface water composition to pollutant fate and drinking water quality

Description: Lakes and rivers in Minnesota are influenced by algae, colored dissolved organic matter, and suspended solids, but we do not fully understand how these constituents affect drinking water production or

the fate of pollutants. We will conduct studies to understand how these features of surface waters, measured for all lakes in the state, influence formation of disinfection byproducts during drinking water treatment and degradation of pesticides and pharmaceuticals in surface waters. Experiments will be performed in laboratory systems that simulate drinking water treatment and natural processes (the latter leveraged through an active LCCMR project led by Arnold).

Laboratory experiments will be performed with water samples collected during field sampling (in Activity 1) to assess the reactivity of the natural organic matter in the sampled waters. Water samples will be collected from a variety of lakes that represent a range of water quality conditions including CDOM levels and algae levels. The water samples will be filtered and stored in the dark in the refrigerator until further use.

The first set of experiments will be designed to assess the reactivity of the organic matter with chlorine to form so-called disinfection byproducts (DBPs). DBP formation is a significant concern in the water treatment industry because chlorine is the most common chemical used for disinfection and because many DBPs are known or suspected human carcinogens. In our DBP experiments, we will focus on a specific class of DBP compounds referred to as trihalomethanes (THMs), one of which is the organochlorine compound chloroform. We will perform THM formation potential tests on each water sample. In the test, excess chlorine is added to a water sample, and the sample is incubated at 25°C for 7 days. At the end of the incubation period, the chlorine residual is measured and then quenched with sodium sulfite. Then, the THMs are analyzed via gas chromatography with an electron capture detector.

The second set of experiments is designed to assess the role of CDOM in sunlight-mediated or ‘photochemical’ reactions. Photochemistry is important in the fate of pollutants in surface water bodies such as lakes because sunlight can react both directly with pollutants and indirectly via formation of reactive oxygen species (ROS) when sunlight reacts with other constituents in the water. CDOM can play two roles in photochemistry. One is to absorb sunlight and diminish the role of direct photolysis and photochemically-mediated contaminant degradation. Conversely, CDOM can act as a photosensitizer and convert sunlight into ROS (such as hydroxyl radical). These ROS can then react with the pollutants to transform or degrade them. In our experiments, the target pollutants will be pesticides. A few pesticides will be selected to represent the range of chemicals commonly used in Minnesota including atrazine and imidacloprid (i.e., Roundup). We will select water samples that have different spectral characteristics based on the satellite and laboratory measurements. These experiments will be performed by adding the pesticide to the water sample and incubating the sample in a solar simulator that mimics the light spectrum of natural sunlight. At the end of the incubation period, the target contaminant will be analyzed by appropriate techniques (e.g., liquid chromatography). Additional experiments will be performed with the same water samples but without added pollutants to quantify the formation of specific ROS during exposure to simulated sunlight using chemical probes. Leveraging current LCCMR funding for Arnold, we will evaluate whether the satellite and laboratory measured spectral characteristics provide predictive power in regards to ROS production.

After the experiments are completed, we will investigate relationships between indicators of DOM composition and content generated in Activity 1, including DOC, CDOM, chlorophyll *a*, spectral slope, UV absorption, EEMs indicators, and molecular structure information from FT-ICR-MS, and both DBP formation potential and the potential for photochemical intermediate formation. We will investigate single parameter correlations as well as multivariate relationships. These analyses will help develop the use of satellite data to perform large-scale assessments of water treatability and photochemical contaminant degradation rates.

Summary Budget Information for Activity 2:

ENRTF Budget: \$ 95,000
Amount Spent: \$ 95,000
Balance: \$ 0

Outcome	Completion Date
---------	-----------------

1. Evaluate the influence of algae and organic matter on formation of disinfection byproducts upon chlorination	December 2018
2. Identify waters least and most likely to degrade pesticides by photolysis	December 2018

Activity Status as of January 2017:

We collected large volume (20-liter) carboys of water from 25 surface waters (mostly large lakes) during the 2016 sampling season for subsequent laboratory studies on the effects of CDOM on formation of disinfection by-products during the chlorination process in drinking water treatment and on *in situ* photochemical intermediate production. All samples were filtered (0.45 μm) immediately after collection, transported in a large freezer, and stored in a cold room in the dark prior to the onset of laboratory studies. Water bodies from which the samples were collected were selected to provide a wide range of CDOM levels, and the emphasis was on large lakes that could serve as potential drinking water supplies in the future. The list includes Rainy Lake, Lake of the Woods, Lake Mille Lacs, Lake Bemidji, and several sites on the Mississippi River. These samples were collected under the Activity 1 budget, since sites were used for both activities.

We hired a postdoctoral scientist to lead the laboratory studies on photochemistry and treatment process effects. She received her Ph.D. in environmental engineering from Temple University in December and arrived on campus in early January. She currently is completing lab safety training requirements and will be starting the lab studies by the end of January 2017.

Activity Status as of July 2017:

Our Activity 2 research focused on evaluating the reactivity of CDOM in 24 lake water samples with chlorine to form disinfection byproducts (DBPs). A trial-and-error exercise was performed to evaluate the chlorine demand for each lake sample. The Uniform Formation Conditions (UFC) test was conducted in at least duplicate to determine the DBP formation potential (DBPFP) upon chlorination. Analytical methods for two regulated DBP classes, trihalomethanes (THMs) and haloacetic acids (HAAs), were developed to improve the accuracy of product quantification. Bromide levels of raw lake samples were measured to estimate the formation concentration of brominated DBP species.

Two THM species including chloroform and bromodichloromethane (TCM and BDCM), and five HAA species including monochloro-, monobromo-, dichloro-, bromodichloro-, and trichloroacetic acid (MCAA, MBAA, DCAA, BDCAA, and TCAA) have been observed as chlorination products. Specifically, TCM, DCAA and TCAA are predominant over other species. The chlorine demand of 24 lake samples varied from 3.7 to 61.5 mgCl_2/L . The THM formation potential (THMFP) we obtained ranges from 104.0 to 2019.1 $\mu\text{g}/\text{L}$, while HAAFP ranges from 592.1 to 2819.7 $\mu\text{g}/\text{L}$.

We built a series of correlations between chlorine demand, THMFP or HAAFP and various indicators of DOM concentration and composition. No correlation was observed between chlorine demand and other indicators such as total suspended solids (TSS), chlorophyll α and so on. SUVA_{254} has a linear correlation with specific chlorine demand, while DOC, CDOM, and UV_{254} correlate well with chlorine demand.

Similar to chlorine demand, positive and strong correlations were observed between THMFP and CDOM, DOC and UV_{254} , indicating that the CDOM, DOC and UV_{254} level are key drivers of the THM yields, especially in lower ranges. Meanwhile, SUVA_{254} and specific THMFP are well correlated. Meanwhile, a lower level of CDOM, DOC or UV_{254} has a linear correlation with HAAFP, respectively, which is similar with THMFP. However, as the level of CDOM, DOC or UV_{254} increases, HAAFP starts to level off, leading to a logarithmic correlation. On the other hand, SUVA_{254} and specific HAAFP were poorly correlated, suggesting that the effectiveness of SUVA_{254} in predicting HAA yields is water-specific. Specifically, different trends were observed among mono HAAFP, di HAAFP and tri HAAFP, indicating that they have different precursors. These results indicate that THMs and HAAs follow different formation patterns upon chlorination, but both have positive correlation at low levels of CDOM, DOC,

UV254 and SUVA254, which can contribute to generate statewide DBPFP maps and perform assessments of water quality in Minnesota's lakes and rivers.

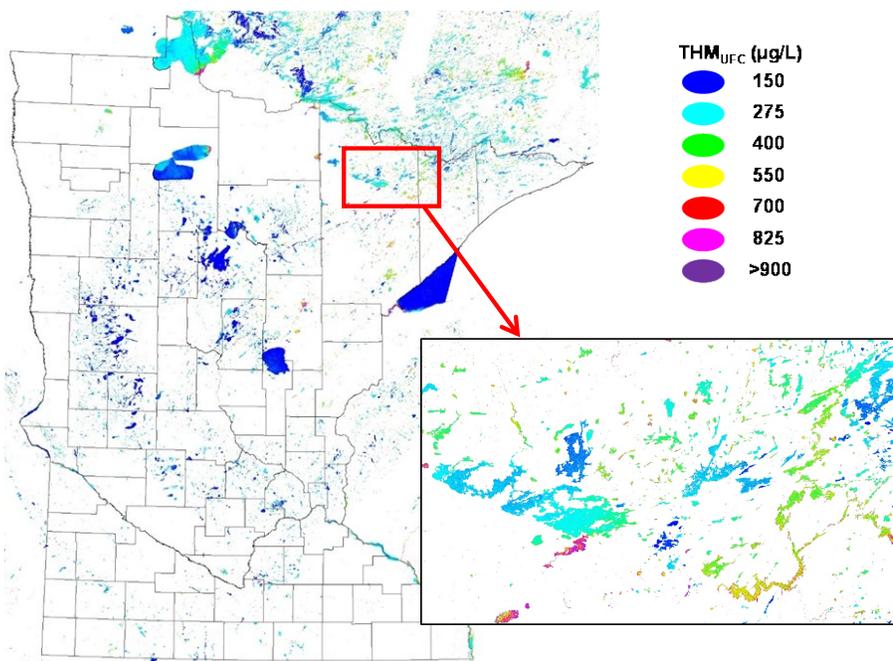
Activity Status as of January 2018:

Our Activity 2 research focused on investigating the role of CDOM in photodegradation of pesticides. The rate of formation ($R_{f,T}$) of triplet excited states of dissolved natural organic matter (${}^3\text{DOM}^*$) was examined for 24 lake samples. One representative herbicide, 2,4,6-Trimethylphenol (TMP), was used as a probe compound to evaluate the formation of ${}^3\text{DOM}^*$. Photochemical experiments of TMP under simulated sunlight were performed in triplicate to measure $R_{f,T}$.

We found that TMP degradation follows pseudo-first order kinetics. A linear correlation was observed between the pseudo-first order rate constant, $k_{\text{obs,TMP}}$, and CDOM at low levels. However, a positive nonlinear trend that approached plateau was observed between $k_{\text{obs,TMP}}$ and CDOM increasing to higher concentration (logarithmic fit). We initiated inhibition experiments to estimate an inhibition factor (IF), which can be used to correct the $k_{\text{obs,TMP}}$ upon high CDOM levels. A linear correlation between $k_{\text{obs,TMP}}$ and CDOM level will be built, which can be coupled with satellite CDOM assessments and contribute to the statewide assessments of pesticide photodegradation upon the impact of CDOM in Minnesota's lakes and rivers.

Activity Status as of July 2018:

We established correlations between CDOM and the concentration of various DBP species in 24 lake samples, and proved that CDOM can be used as a predictor for chlorine demand and DBP formation. We generated pixel level maps to evaluate distribution of chlorine demand and DBP_{UFC} levels for > 10,000 lakes throughout



Minnesota, based on models of CDOM-chlorine demand and CDOM- DBP_{UFC} , as well as Landsat 8 images obtained close in time to sampling. A much higher level of THM_{UFC} (> 150 $\mu\text{g/L}$) in northeast Minnesota, especially in the NLF ecoregion, and relatively lower THM_{UFC} were obtained in the NCHF), the Northern Minnesota Wetlands (NMW), and the WCBP ecoregions. Similarly, the NLF generally has higher levels of HAA_{UFC} (> 800 $\mu\text{g/L}$). The distribution of DBP_{UFC} levels in the four ecoregions clearly reflected their current situations of urban and agricultural land uses. Consistent with DBP_{UFC} , a higher chlorine demand (> 5 mgCl_2/L) was observed in the NLF.

Figure 10. Distribution of lake THM_{UFC} levels in Minnesota using the two-variable universal model (green/red + red/NIR), $a_{440} = -5.478 \cdot \text{OLI3}/\text{OLI4} - 0.633 \cdot \text{OLI4}/\text{OLI5} + 8.135$ for Landsat 8 SR imagery. Inset map shows detail in the Ely, MN area.

Our efforts to elucidate the role of CDOM in photochemical experiments focused on determining the steady-state concentrations of different photochemically produced reactive intermediates (PPRIs), including ${}^3\text{DOM}^*$, ${}^1\text{O}_2$, and $\bullet\text{OH}$ under simulated sunlight. Because the photodegradation of pesticides in aquatic environments

may be facilitated via photosensitized processes in which PPRIs are involved. Knowing the concentrations of PPRIs plays an important role in predicting the fate of pesticides in natural waters. For $^3\text{DOM}^*$, we found that the degradation rate constant of its probe TMP ($k_{\text{obs,TMP}}$) followed a positive logarithmic trend across the entire CDOM/DOC range. We investigated the influence of DOC inhibition on TMP photodegradation using 4-carboxybenzophenone (CBP) as a $^3\text{DOM}^*$ model, and corrected the nonlinear correlation between $k_{\text{obs,TMP}}$ and CDOM. We also examined the formation of $^1\text{O}_2$ using furfuryl alcohol (FFA) as a probe. We found that FFA degradation follows pseudo-first order kinetics, and linearly correlated with CDOM level, suggesting that $^3\text{DOM}^*$ is the primary source for $^1\text{O}_2$ in sunlit natural waters. By monitoring the photodegradation of p-nitroanisole (PNA) as an actinometer, we measured the rate of light absorption by the samples, and calculated the rate of formation ($R_{f,T}$) of $^3\text{DOM}^*$ and the efficiency of $^3\text{DOM}^*$ formation (the apparent quantum yield, AQY_T). In addition, we quantified the formation of $\bullet\text{OH}$ by the formation of 2-hydroxyterephthalic acid (hTPA) using terephthalic acid (TPA) as a probe. We will apply the established correlations between CDOM and the formation of PPRIs to generate state-wide maps, which can be used to further assess the treatment necessary to use Minnesota's lakes and rivers as drinking water sources.

Activity Status as of January 2019:

We made advances in elucidating the relationship between CDOM (a_{440}) and the generation of PPRIs upon photolysis. The DOM compositional parameters we investigated include the total rate of light absorption by the water samples (R_a), the formation rates (R_f), the steady-state concentrations and apparent quantum yields (Φ) of three PPRIs. We found that R_a was strongly and positively related to a_{440} levels, as found in other recent studies. For $^3\text{DOM}^*$ and $^1\text{O}_2$, the R_f and the steady state concentrations increased linearly with increasing a_{440} , while Φ were independent of a_{440} , but linearly correlated with E2/E3 (the ratio of UV absorbance at 250 to 365 nm), which is a direct proxy for the degree of photobleaching and an inverse proxy for DOM molecular weight. Consistent with previous studies, no correlation was observed between E2/E3 and $\Phi_{\bullet\text{OH}}$. Our ongoing work will apply the relationships between R_f and R_a values with a_{440} to provide estimates of the rate of contaminant loss via indirect photolysis within lakes. These relationships can then be scaled up in maps of reactive species production rates across all surface waters in Minnesota that will allow prediction of contaminant transformation rates. An example is shown in Figure 12.

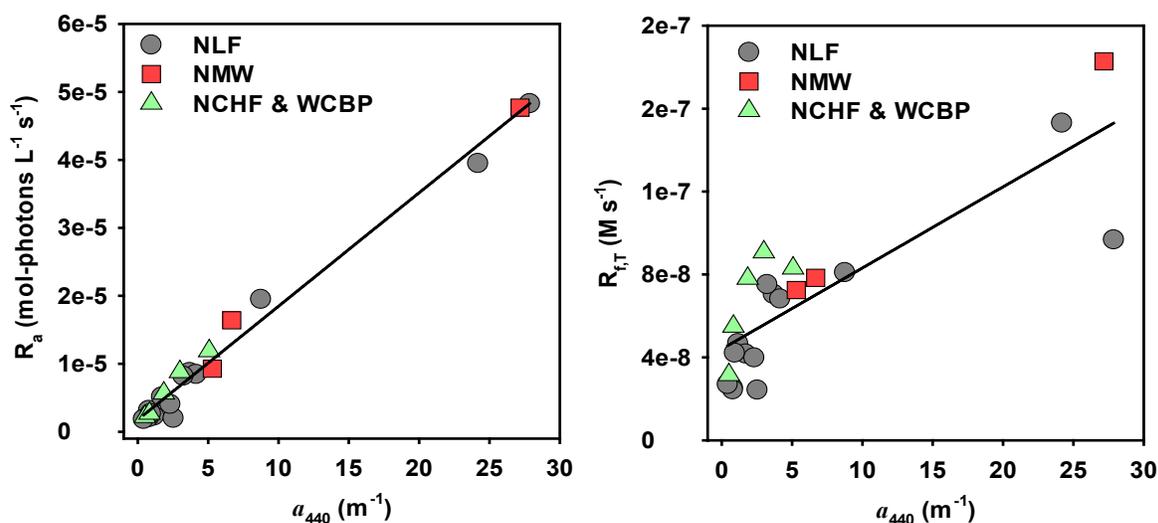


Figure 12. Correlations of CDOM (a_{440}) with R_a and $R_{f,T}$ of $^3\text{DOM}^*$. The equations for the regressions shown are: (a) R_a (mol-photons $\text{L}^{-1} \text{s}^{-1}$) = $1.6695\text{E-}006 \cdot a_{440}$ (m^{-1}) + $1.7708\text{E-}006$; (b) $R_{f,T}$ (M s^{-1}) = $3.9125\text{E-}009 \cdot a_{440}$ (m^{-1}) + $4.4028\text{E-}008$.

Final Report Summary:

Laboratory experiments were completed to quantify the effects of CDOM on chlorine demand and the formation of two classes of halogenated disinfection byproducts (DBPs): trihalomethanes (THMs) and haloacetic

acids (HAAs). Using 24 lake samples collected across Minnesota, we found that CDOM (a_{440}) is a useful predictor for both chlorine demand and trihalomethane formation potential (THMFP); strong linear correlations were observed between CDOM and these two parameters. For low- to moderately-colored waters ($a_{440} \leq 11 \text{ m}^{-1}$), a_{440} relationships with mono-, di- and tri-HAA formation potentials (HAAFPs) were best fit by a linear relationship, but over the entire CDOM range, di- and tri-HAA were best fit by a logarithmic relationship. Besides CDOM, other measured parameters related to DOM, including DOC, UV_{254} , and $SUVA_{254}$, also were observed to significantly affect DBPs yields, but a variety of other water quality indicators, including pH, TSS, chlorophyll a, and bromide, did not correlate with CDOM and did not affect the formation of DBPs.

Laboratory-measured CDOM levels from 194 lakes were used to develop a predictive relationship between measured CDOM and reflectance data obtained from Landsat 8 imagery. Correlations relating chlorine demand and THMFP and HAAFP values with CDOM, coupled with satellite CDOM assessments, were able to estimate chlorine demand and both THMFP and HAAFP values for low and moderate color surface waters larger than 12 acres in the state of Minnesota. The maps show relatively high levels of chlorine demand and disinfection byproduct formation potentials in northeastern Minnesota lakes, especially in the NLF ecoregion. The maps also suggest that only 22% of Minnesota lakes would meet both the THM and HAA maximum contaminant levels even if pre-disinfection treatment could remove 75% of DBP precursors. However, the models showed an increase in uncertainty for highly colored surface waters ($a_{440} > 11 \text{ m}^{-1}$), and CDOM and disinfection byproduct formation potential values were severely under-predicted due to various factors, including strong absorption of light resulting in a weak signal obtained by the satellite sensors. Overall, this study demonstrated that satellite imagery can be used to evaluate potential drinking water sources and water treatability metrics.

We measured the formation of photochemically-produced reactive intermediates (PPRIs) for 24 lake samples under simulated sunlight. These PPRIs include triplet excited states of dissolved organic matter ($^3\text{DOM}^*$), singlet oxygen ($^1\text{O}_2$), and hydroxyl radical ($\bullet\text{OH}$). We measured their rates of formation (R_f), steady-state concentrations ($[\text{PPRI}]_{ss}$) and apparent quantum yields (Φ). The results showed that the total rate of light absorption by the water samples (R_a) and the rates of formation of $^3\text{DOM}^*$ and $^1\text{O}_2$ ($R_f, ^3\text{DOM}^*$, and $R_f, ^1\text{O}_2$, respectively), as well as steady state concentrations of triplet excited state organic matter ($^3\text{DOM}^*$) and singlet oxygen ($^1\text{O}_2$) linearly increased with increasing CDOM levels. The production rate of hydroxyl radical ($R_f, \bullet\text{OH}$) correlated linearly with CDOM, while the steady state concentration ($[\bullet\text{OH}]_{ss}$) versus CDOM revealed a logarithmic relationship. The efficiencies by which $^3\text{DOM}^*$ and $^1\text{O}_2$ were produced (which are known as quantum yields, and given the symbols $\Phi_{app, ^3\text{DOM}^*}$, and $\Phi_{^1\text{O}_2}$) increased linearly with increasing ratios of the absorption coefficients at 250 nm and 365 nm (termed E2/E3), and were negatively correlated with $SUVA_{254}$, suggesting that photobleached CDOM, which tends to have a lower average molecular weight and lower aromaticity, leads to more efficient $^3\text{DOM}^*$ and $^1\text{O}_2$ production. On the other hand, the quantum yield for hydroxyl radical production, $\Phi_{app, \bullet\text{OH}}$, exhibited no significant correlation with either E2/E3 or $SUVA_{254}$. The correlations relating CDOM with rates of formation of PPRIs, steady state concentrations of PPRIs, or the rate of light absorbance were coupled with satellite CDOM assessments to develop maps of the production rates and concentrations PPRIs across all surface waters in Minnesota. These maps can be used with known rate constants for contaminants to predict the maximum rate of transformation of contaminants in sunlit surface waters. Further work is needed to adjust the findings for different light intensity, which varies seasonally. This work demonstrates that satellite remote sensing can be an effective tool to provide regional estimates for loss rates of contaminants via indirect photolysis in surface waters.

ACTIVITY 3: Dissemination and application for surface water monitoring and management

Description: We have maintained a web site (<http://water.umn.edu>) for fifteen years to provide public access to results of our remote sensing studies on lake water clarity and quality. The Minnesota Lake Browser provides easy access of information for individual lakes or regions for seven time periods at five-year intervals from 1975 to 2008. Lake Browser is a Google Earth format mapping application that allows users to query for individual

lakes or zoom into an area of interest to get not only the information for individual lakes but for all of the lakes in the area.

We will update the Lake Browser site to include the new data gained in Activity 1 during 2016 and 2017. This information will be used to construct a web-accessible statewide database of colored dissolved organic matter, algae, and suspended solid levels and descriptions of their relevance to citizens. Lake Browser will use information based on both pixel level maps, that allow identification of within lake differences, and lake level maps, that enable statistical analysis of geospatial and temporal trends. These additions will greatly enhance the site by providing up to date and more detailed information for >10,500 lakes. The site is currently extensively used with on average 8,000 unique visitors per month, and we expect that data and products provided by this project will increase use of Lake Browser substantially.

We expect at least 3-4 publications in high-impact, refereed journals will result from this research. Results will be presented at local and national conferences on environmental engineering, ecosystem ecology, humic substances, and remote sensing. The PIs have close relationships with scientists in regional water management agencies, and we will organize meeting for staff of those agencies (e.g., MPCA) and at state conferences to demonstrate capacity and effectiveness of remote sensing methods and to enhance their integration into water quality monitoring programs and assessments.

Summary Budget Information for Activity 3:

ENRTF Budget: \$ 45,000
Amount Spent: \$ 45,000
Balance: \$ 0

Outcome	Completion Date
<i>1. Enhancements and expansion of Lake Browser</i>	<i>June 2019</i>
<i>2. Peer reviewed papers describing methods for submission to peer reviewed journals</i>	<i>June 2019</i>
<i>3. Presentations to MPCA describing remote sensing capabilities and methods</i>	<i>June 2019</i>

Activity Status as of January 25 2017:

Our efforts toward dissemination of project products and applications for education, monitoring and management for Activity 3 included two presentations (described in more detail under our Section V update), sharing of information with MPCA on project activities, and an initial update of the water.umn.edu website to include CDOM. A statewide CDOM map was added to Lake Browser this fall (<http://lakes.gis.umn.edu/cdom/>). This map was created with previously collected data, and will be useful to target lakes with a wide range of water quality conditions for sampling in summer 2017. Field monitoring results from 2017 will be used to refine and improve mapping of CDOM and other water quality variables in this project using remote sensing. These results will continue to be integrated into the water.rs.umn.edu and Lake Browser websites for easy access by agencies, managers and citizens.

Activity Status as of July 2017:

Efforts begun in the previous period continued, largely focused on work on website improvements and expansions that describe this project and provide comprehensive explanation of the water quality parameters now available via remote sensing, and the methods being developed in this project. The water.umn.edu site has been readdressed as water.rs.umn.edu as part of a broader revision of the Water Resource Center’s website and Digital Water initiative. Project related presentations were shared at the ASLO Aquatic Sciences Meeting in Honolulu, Hawaii and the American Chemical Society meeting in San Francisco, California.

Activity Status as of January 2018:

Revision of the water.rs.umn.edu website is nearing completion, with final editing underway. A key component of the site, Lake Browser, will be expanded and updated when substantial new data generated by this project

are available late in 2018 and 2019. Two project presentations were made at the 2017 Water Resources Conference. Discussions with MPCA about application of project results to trophic classifications of the states lakes are ongoing and will continue once analysis of 2017 water quality data has been completed. CDOM data for the Northern Lakes and Forest Ecoregion were tabulated from our CDOM maps and disseminated to the MPCA for use to identify lakes with high CDOM where water clarity data should not be used for nutrient impairment assessment.

Activity Status as of July 2018:

The revised and updated version of the water.rs.umn.edu website was implemented (launched) in February. It has a new homepage and organization. In addition to revising the material on monitoring lake clarity, we made major additions including new sections (pages) on chlorophyll *a*, phycocyanin, turbidity and suspended matter (TSM), and colored dissolved organic matter (CDOM). The sections for mapping river water quality and aquatic plants were also revised to be consistent with other topics and pages. Publications were updated and a new page on our research team (people) was added. New results from this and related projects will be added as they are available.

A key component of the water website has been the Lake Browser (lakes.rs.umn.edu) that has approximately 70,000 annual visits. It currently has lake clarity for seven dates from 1975 to 2008, and CDOM for 2015. We are working with staff of the U-Spatial (uspatial.umn.edu) to convert it to ArcGIS Server, a newer and easier to work with platform. The first phase, to be completed in August, is transferring the capability/functionality of the current MapServer version to ArcGIS Server. Lakes will be searchable by lake number and county name, as well as lake name, and selected lakes will have a dotted outline and symbol at their centroid. The next phase will add new data for chlorophyll, turbidity and CDOM. The third phase will add enhancements, including maps and statistical analyses on distributions by ecoregion, watershed and county, plus temporal trends for clarity. New data from current projects, including lake clarity for 2010 and 2015 from a new project sponsored by the MPCA, plus data from the current LCCMR project chlorophyll, suspended sediments and CDOM, will be added as they become available.

Other Activity 3 advancements included presentations at the St. Louis River Summit, the University of Minnesota HAB Workshop, Minnesota Supercomputing Institute Exhibition, and several other academic settings, as well as advancement on four publications related to our project.

Activity Status as of January 2019:

Lake Browser (lakes.rs.umn.edu) was successfully transferred to the ArcGIS Server. This transition represents a major improvement in the capacity and usefulness of the resources. Lakes now easily are searchable by lake number and county name, as well as lake name, and selected lakes have a dotted outline and symbol at their centroid. We will continue to add new data from this and other projects during 2019 and beyond. These data include lake clarity for 2010 and 2015 from Landsat imagery and for water clarity, chlorophyll, turbidity and CDOM using Sentinel 2 imagery from recent years. Further enhancements are in development, including maps and statistical analyses on parameter distributions by ecoregion, watershed and county, as well as greater detail on temporal trends in water clarity.

Other Activity 3 advancements included presentations at the Water Resources Conference, Minnesota Supercomputing Institute and several other academic settings, as well as advancement on four publications for peer reviewed journals arising from our work.

Final Report Summary:

The three outcomes for Activity 3 are now complete. First, we have extensively revised and updated our websites — *Remote Sensing of Water Quality* (<https://water.rs.umn.edu/>) and the LakeBrowser (<http://lakes.rs.umn.edu/>) — that have been providing information on remote sensing of water resources to agencies, researchers and citizens since 2002. The new version (released in August 2019) of the Remote Sensing

of Water Quality site describes how remote sensing provides key information about water quality, an overview of how remote sensing works, methods of extracting water quality properties from remote sensing data, and results of classifications of lake water quality, including clarity, chlorophyll, colored dissolved organic matter (CDOM), turbidity and suspended solids. Additional pages summarize past studies on aquatic plant mapping and mapping water quality of large rivers, along with research publications and presentations, current faculty, staff and students contributing to projects, and news on project activities and accomplishments.

An expanded and enhanced version of the LakeBrowser was also released in August 2019. **We highly encourage readers of this report to check out the new LakeBrowser at <http://lakes.rs.umn.edu/>.** LakeBrowser now includes nine statewide classifications of lake clarity from 1975 to 2015, along with recent classifications of chlorophyll *a* and CDOM. The site has been converted to an ArcGIS Server with improved capability for search and data download, additional base maps (including high resolution imagery and land cover), and display of temporal trends and geographic patterns. Users may now search for individual lakes by name or ID, or select a lake from a map. Reports for individual lakes includes lake clarity by year, and chlorophyll, and CDOM levels, along with a percentile ranking of clarity and the land cover in a 1000-foot (shoreland) buffer around the lake. Statistical summaries of lake clarity by year are available for lake clarity at state, county, ecoregion and watershed levels. Several examples of the various capabilities for display of water quality properties are shown in Appendix 1. The LakeBrowser is viewed by several thousand visitors each month and its data are used by state and county agencies, researchers, and citizens. We expect that its use will increase with the addition of new data and other capabilities added by this project.

Second, five manuscripts describing methods and project results have been published in peer reviewed journals. These papers describe (1) methods for correction of satellite imagery for atmospheric interference (Page et al. 2019), (2&3) relationships between CDOM, dissolved organic carbon and iron, providing information needed to interpret and use CDOM data (Brezonik et al. 2019a, Griffin et al. 2019), (4) relationships between remoted sensed lake water quality parameters and disinfection byproduct formation potential in drinking water treatment (Chen et al., In press), and (5) the influences of CDOM on water clarity based water quality standards in Minnesota (Brezonik et al. 2019b). The results of the last paper are described in more detail below because it directly addresses our third outcome for engaging with MPCA in interpretation of results. All five of these manuscripts have been included at attachments and the full citations are listed below. Several additional papers are under development.

Third, in collaboration with researchers at the Minnesota Pollution Control Agency (MPCA), we analyzed a large data set to evaluate the role of CDOM in affecting Secchi depth (SD) values in lakes. SD is a primary metric to assess trophic state because it is controlled in many lakes by algal population densities, measured in terms of chlorophyll-*a* (chl-*a*) concentration. Two other water quality variables also affect SD: non-algal suspended solids (SS_{NA}) and colored dissolved organic matter, CDOM, expressed as a_{440} . Using a database of 1460 samples from ~ 625 lakes comprised of our ground-based calibration data and MPCA lake monitoring data from 2015-2017, we analyzed relationships among these variables. Special focus was placed on CDOM levels that influence SD and implications for the Minnesota SD standards used to assess eutrophication impairment. Log-transformed chl-*a*, total suspended solids (TSS), and SD were strongly correlated with each other; log a_{440} had major effects on log SD but was more weakly correlated with log chl-*a* and log TSS. Regression models for log SD and 1/SD based on

the variables chl-*a*, SS_{NA}, and CDOM explained ~ 80% of the variance in SD, but substantial differences in the form of the best-fit relationships were found among major ecoregions. High chl-*a* concentrations (> 50 µg/L) and TSS (> 20 mg/L) rarely occurred in lakes with high CDOM ($a_{440} > \sim 4 \text{ m}^{-1}$). Moreover, all lakes with $a_{440} > 8 \text{ m}^{-1}$ had SD ≤ 2.0 m despite low chl-*a* (< 10 µg/L) in most lakes. Further statistical analyses revealed that CDOM starts to have significant effects on SD at $a_{440} > \sim 4 \text{ m}^{-1}$. Thus, SD is not an accurate trophic state metric in moderately to highly colored lakes. We concluded that Minnesota's 2-m SD criterion should not be the sole metric to assess eutrophication impairment in warm/cool-water lakes of the Northern Lakes and Forest ecoregion. This work was published in the journal *Ecological Applications* in early 2019.

Manuscripts published arising from this project:

Brezonik, P. L., J. C. Finlay, C. G. Griffin, W. A. Arnold, E. H.

Boardman, N. Germolus, R. M. Hozalski, and L. G.

Olmanson. 2019a. Iron influence on dissolved color in lakes of the Upper Great Lakes States. *Plos One*

14:e0211979.

Brezonik, P. L., R. W. Bouchard Jr, J. C. Finlay, C. G. Griffin, L. G.

Olmanson, J. P. Anderson, W. A. Arnold, and R. Hozalski.

2019b. Color, chlorophyll a, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. *Ecological Applications* **29**:e01871.

Chen, Y. et al. 2019. Assessment of the Chlorine Demand and Disinfection Byproduct Formation Potential of 1 Surface Waters via Satellite Remote Sensing. *Water Research*. **165**, 115001

Griffin, C. G., J. C. Finlay, P. L. Brezonik, L. Olmanson, and R. M. Hozalski. 2018. Limitations on using CDOM as a proxy for DOC in temperate lakes. *Water Research* **144**:719-727.

Page, B.P. and L. Olmanson, and D.R. Mishra. 2019. A harmonized image processing workflow using Landsat-8 and Sentinel-2 for mapping water clarity in optically complex lake systems. *Remote Sens. Environ.* **231** 111284.

V. DISSEMINATION:

Description: Project results will be communicated using a range of outlets. A primary mode of dissemination is the expansion of Lake Browser. This website provides content for diverse users including citizen scientists, homeowners, classes, natural resource managers, researchers at agencies and academic institutions. Results will also be disseminated in the peer reviewed literature, and in presentations made at conferences and at state agencies.

Activity Status as of January 25 2017:

Communication of project progress included a presentation to scientists and managers at NASA Goddard Space Flight Center in Greenbelt, Maryland (Washington D.C. area) titled "Using new enhanced satellite remote sensing systems for regional water quality measurements in optically complex inland waters". (Travel funding supplied by a related project). In addition, a presentation was given at the 2016 Minnesota GIS/LIS Consortium Conference in Duluth titled "Regional lake water quality measurements beyond water clarity using new enhanced satellite remote sensing systems". The NASA visit resulted in an article titled "Minnesota: Land of the Many-Colored Lakes" in NASA's *Earth Observatory Images of the Day* highlighting the preliminary CDOM map we developed from 2015 Landsat 8 imagery (<http://earthobservatory.nasa.gov/IOTD/view.php?id=88971>).

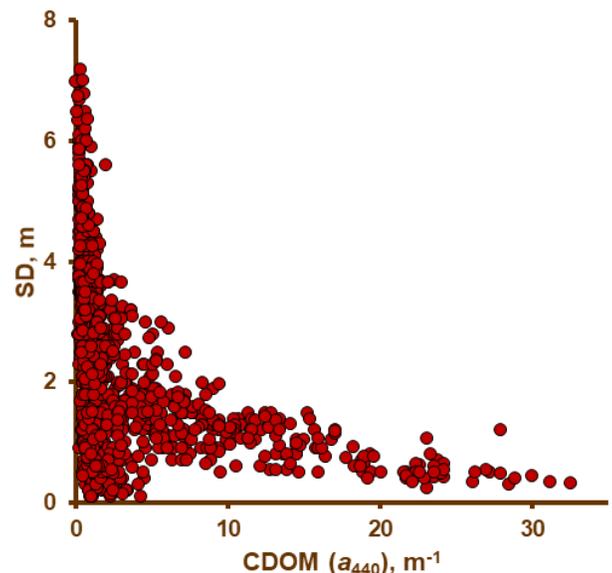


Figure 13 SD encompasses a wide range in low-color lakes, reflecting how much chlorophyll (algae) is present, but high colored lakes have low SD because of the light-absorbing properties of CDOM. Data points represent SD and CDOM values for lakes sampled in 2015-2018.

Activity Status as of July 2017:

Communication of project progress included a presentation at the ASLO Aquatic Sciences Meeting in Honolulu, Hawaii titled “Using Landsat 8 and Sentinel 2 data for regional water quality measurements in optically complex inland waters”. (Travel funding was supplied by a related project). Another project-related presentation was made in a symposium at the American Chemical Society (Environmental Chemistry Division) National meeting in San Francisco, California titled “Cross-scale Advances in CDOM Biogeochemistry: From Molecular to Eco-Regional Perspectives.” The presenter (a project co-PI) covered travel and conference expenses from personal funds.

Activity Status as of January 2018:

Communication of project progress included a presentation to scientists and managers at the Pecora Conference in Sioux Falls, South Dakota titled “Water Quality Measurements of Optically Complex Inland Waters Using New Enhanced Landsat 8 and Sentinel 2 Imagery” (travel funding was supplied by America View). In addition, two presentations were given at the Water Resources Conference in St. Paul, Minnesota.

Activity Status as of July 2018:

We have made substantial enhancements to a widely used website, and presented project results in seminars at six academic and research institutions. Dissemination of project findings to MPCA has been initiated with further collaborations planned for fall.

Activity Status as of January 2019:

We have completed upgrade and enhancements to a widely used website for lake water quality data and presented project results in seminars at six academic and research institutions. Dissemination of project findings to MPCA and DNR is ongoing, and has resulted in development of a collaborations on a journal article submitted for publication in fall 2018.

Final Report Summary:

Project findings have been disseminated through diverse means. We have extensively revised and updated our websites — *Remote Sensing of Water Quality* (<https://water.rs.umn.edu/>) and the LakeBrowser (<http://lakes.rs.umn.edu/>) — that have been used for dissemination of information on remote sensing and water resources since 2002. The new version (released in August 2019) of the Remote Sensing of Water Quality site describes how remote sensing provides key information about water quality, an overview of how remote sensing works, methods of extracting water quality properties from remote sensing data, and results of classifications of lake water quality, including clarity, chlorophyll, colored dissolved organic matter (CDOM), turbidity and suspended solids. Additional pages summarize past studies on aquatic plant mapping and mapping water quality of large rivers, along with research publications and presentations, current faculty, staff and students contributing to projects, and news on project activities and accomplishments.

An expanded and enhanced version of the LakeBrowser was also released in August 2019 and we highly encourage readers of this report to explore its capabilities (<http://lakes.rs.umn.edu/>.) LakeBrowser now includes nine statewide classifications of lake clarity from 1975 to 2015, along with recent classifications of chlorophyll *a* and CDOM. The site has been converted to an ArcGIS Server with improved capability for search and data download, additional base maps (including high resolution imagery and land cover), and display of temporal trends and geographic patterns. Users may now search for individual lakes by name or ID, or select a lake from a map. Reports for individual lakes includes lake clarity by year, and chlorophyll, and CDOM levels, along with a percentile ranking of clarity and the land cover in a 1000-foot (shoreland) buffer around the lake. Statistical summaries of lake clarity by year are available for lake clarity at state, county, ecoregion and watershed levels. Several examples of the various capabilities for display of water quality properties are shown in Appendix 1. The LakeBrowser is viewed by several thousand visitors each month and its data are freely available for use by state and county agencies, researchers, and citizens. We expect that use will increase substantially with the addition of new data and other capabilities added by this project.

Collaborations with MPCA and DNR developed in this project resulted in a joint authored journal publication in 2019. These results will be used to help development of more accurate water quality standards for water clarity by distinguishing between effects of algae (which are increased by excessive nutrient loading) from those of CDOM (which come from natural sources), so they will have a broad impact on water monitoring and assessment. Results from this project have been presented at more than 10 meetings, including six at the Minnesota Water Resources Conference. Finally, five manuscripts presenting the results of this work have been published in academic journals (citations given above denoted with *), and several manuscripts are in preparation.

VI. PROJECT BUDGET SUMMARY:

A. ENRTF Budget Overview:

***This section represents an overview of the preliminary budget at the start of the project. It will be reconciled with actual expenditures at the time of the final report.**

Budget Category	\$ Amount	Overview Explanation
Personnel:	\$ 303,000	Salary support for project partners, postdoctoral researchers, a junior scientist, and undergraduate researchers
Professional/Technical/Service Contracts:	\$ n/a	
Equipment/Tools/Supplies:	\$ 11,000	Laboratory supplies and analytical time/costs
Capital Expenditures over \$5,000:	\$ 20,000	Multi parameter water quality datalogger
Fee Title Acquisition:	\$ n/a	
Easement Acquisition:	\$ n/a	
Professional Services for Acquisition:	\$ n/a	
Printing:	\$ 1,000	
Travel Expenses in MN:	\$ 6,000	In state travel for sample collection and presentation of results
Other:	\$ 4,000	Laboratory services
TOTAL ENRTF BUDGET:	\$345,000	

Explanation of Use of Classified Staff: N/A

Explanation of Capital Expenditures Greater Than \$5,000: YSI Sensor package - A Yellow Springs Instruments (YSI) sonde capable of measuring and data logging multiple parameters including algal pigments and DOM fluorescence is requested. This instrument will be used to verify the correspondence between remotely sensed parameters and field observed values. The equipment will be used throughout the project for collection of data to needed to ensure that sampling sites are representative, to examine assumptions related to effects of shoreline and water depth effects on satellite data, and to monitor pH, oxygen and organic matter levels in laboratory experiments. We will continue to use the equipment for purposes related to the proposed research throughout the life of the instrument as new satellite sensors added to the ones currently available.

Number of Full-time Equivalent (FTE) Directly Funded with this ENRTF Appropriation: 4.73 total

Number of Full-time Equivalent (FTE) Estimated to Be Funded through Contracts with this ENRTF Appropriation: N/A

B. Other Funds:

Source of Funds	\$ Amount Proposed	\$ Amount Spent	Use of Other Funds
Non-state			
n/a			
State			
n/a	\$	\$	
TOTAL OTHER FUNDS:	\$	\$	

VII. PROJECT STRATEGY:

A. Project Partners:

Project Partners Receiving Funds:

- Dr. Jacques Finlay (Project Manager, University of Minnesota)
- Patrick Brezonik (Co PI, University of Minnesota)
- Leif Olmanson (Co PI, University of Minnesota)
- William Arnold (Co PI, University of Minnesota)
- Raymond Hozalski (Co PI, University of Minnesota)

B. Project Impact and Long-term Strategy: This project directly addresses LCCMR funding priorities in *Water Resources* and *Foundational Natural Resource Data and Information*. Our project brings together expertise in remote sensing, aquatic ecology, contaminant cycling, water quality analysis, and drinking water treatment to advance our abilities to detect and understand spatial and temporal patterns in water quality. Our past development of remote sensing methods for water clarity, funded in part by LCCMR, has allowed routine monitoring of >10,000 Minnesota lakes. Expansion of these capabilities through the use of new satellite capabilities to include organic matter, algal abundance, and suspended sediments will be a major step in the development of more cost-effective and spatially comprehensive methods to monitor, understand and manage Minnesota’s freshwater resources. Because water quality affects fisheries, drinking water, ecosystem integrity, and human enjoyment of water bodies, results from our project will be of immediate use to the Minnesota Pollution Control Agency and the Department of Natural Resources in decision making and prioritization of resources. At the end of this project, we will be able to provide these and other relevant agencies with the basic tools needed to initiate their own use of remote sensing techniques as operational tools for frequent, statewide assessments of surface water quality throughout the state. Subsequent integration by natural resources agencies of satellite-based monitoring based on the methods we will develop would be enabled by increasing the GIS expertise of the agencies, thus allowing them greater ability to process high-quality satellite imagery, which is becoming increasingly available at no cost. It is feasible that such an advanced monitoring program could be in place following the end of this project in summer 2019 if resources were devoted to this task. We also expect that techniques for detection of blue-green algae (the primary cause of harmful algal blooms) will become feasible within the next few years, and we further envision subsequent development of ways to further automate image processing to the point that near real-time assessments of important water quality conditions will be possible.

C. Funding History:

Funding Source and Use of Funds	Funding Timeframe	\$ Amount
"Solar Driven Destruction of Pesticides, Pharmaceuticals, Contaminants in Water" Arnold is currently investigating pesticide/pharmaceutical fate in wetlands and the role of DOM plays in photolysis. The techniques developed will be used on the lake and river samples in this project.	July 1, 2014-June 30,2017	\$291,000

VIII. FEE TITLE ACQUISITION/CONSERVATION EASEMENT/RESTORATION REQUIREMENTS:

A. Parcel List: N/A

B. Acquisition/Restoration Information: N/A

IX. VISUAL COMPONENT or MAP(S): See attached figure.

X. RESEARCH ADDENDUM: Research addendum is an unfunded but well reviewed proposal to USGS from 2016

XI. REPORTING REQUIREMENTS:

Periodic work plan status update reports will be submitted no later than January 2017, July 2017, January 2018, July 2018, and January 2019. A final report and associated products will be submitted between June 30 and August 15, 2019.

Field supplies including bottles, gloves, and filters required for collection, transport, and storage of samples, and for preparation for lab manipulations												
YSI Sensor package - A YSI sonde capable of measuring and data logging multiple parameters including algal pigments and DOM fluorescence is requested. This instrument will be used to verify the correspondence between remotely sensed parameters and field observed values. The equipment will be used throughout the project for collection of data to needed to ensure that sampling sites are representative, to examine assumptions related to effects of shoreline and water depth effects on satellite data, and to monitor pH, oxygen and organic matter levels in laboratory experiments. We will continue to use the equipment for purposes related to the proposed research throughout the life of the instrument as new satellite sensors added to the ones currently available. All other equipment needed for the project is currently available for the use of the research team in on-campus laboratories or in shared (core) facilities.	\$20,000	\$20,392										
Printing							\$1,000	\$1,000	\$0	\$1,000	\$0	
Travel expenses in Minnesota	\$4,000	\$4,000	\$0		\$0		\$2,000	\$2,000	\$0	\$6,000	\$0	
Travel funds are requested for travel to field sites in year 1 and 2. This includes vehicle rental, mileage, hotel, and meals, estimated according to UMN guidelines, for PIs, post docs, and undergraduates during year 1 and for the field sampling campaign in the second year. In addition, funds are requested for travel to meetings with collaborators and state agencies, and for registration at the Minnesota Resources Conference in years 2 and 3.												
Other	\$4,000	\$4,000	\$0		\$0		\$0	\$0		\$4,000	\$0	
Lab Services - Lab services includes costs of sample analyses at external labs for metals and major ions via ICP and IC in the Dept. of Earth Sciences at the University of Minnesota (\$31.50 per sample), and charges for FT-ICR-MS analyses (\$100 per sample) to characterize DOC/DOM properties.												
COLUMN TOTAL	\$205,000	\$205,000	\$0		\$95,000	\$95,000	\$0	\$45,000	\$45,000	\$0	\$345,000	\$0

Sources and nature

Dissolved organic matter (DOM) in lakes and rivers is derived from two natural processes: production by aquatic organisms and leaching of decaying vegetation from terrestrial sources. Both sources yield a wide array of organic molecules of varying sizes and chemical composition that form the aggregate water quality constituent we call DOM.

Anthropogenic sources, such as wastewater effluent and urban and agricultural runoff, also contribute to the DOM pool in many waters. DOM differs chemically depending on its source, and understanding the relationships between sources and DOM composition is still a work in progress.

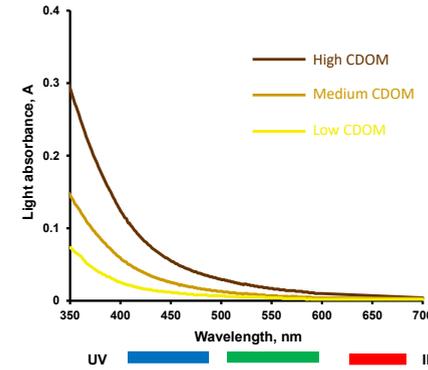
Colored dissolved organic matter (CDOM), which in high concentrations is associated with runoff from forested wetlands, is a complicated mixture of tannins and breakdown products of woody vegetation. CDOM molecules absorb light increasingly with decreasing wavelength in the blue and UV regions, giving rise to a yellow-brown tint to CDOM-rich waters. CDOM is the most abundant DOM fraction in natural waters of forested watersheds with wetlands, which are common in northern areas of the Great Lakes States.



CDOM can be measured from 1) smartphone photos of a Secchi disk placed at a fixed depth in lakes, or 2) from photos of white buckets filled with lake water.

Measurement of CDOM

Traditionally, CDOM is measured by filtering a water sample to remove suspended particles and measuring the light absorbance in the blue region (~ 440 nm) with a spectrophotometer. An older method compares the color of filtered water to calibrated colored glass disks. We have developed methods to measure CDOM in lakes using satellite imagery, which provides the ability to examine landscape patterns in CDOM levels at local to regional scales. We also have found that smartphone cameras can measure CDOM by photographing the white part of Secchi disks at a fixed depth in the water. An app that uses RGB information from digital photos is being developed for use by citizen monitoring programs.



CDOM light absorption increases exponentially as wavelength decreases in the blue and UV regions.

For more information:

Jacques Finlay, Professor
DEPARTMENT of ECOLOGY, EVOLUTION and BEHAVIOR
jfinlay@umn.edu

Patrick Brezonik, Professor Emeritus
DEPARTMENT of CIVIL, ENVIRONMENTAL, and GEO- ENGINEERING
brezonik@umn

Support to develop this brochure was provided by National Science Foundation grant CBET 1510332, citizen science supplement, from the Environmental Engineering Program.

<https://water.rs.umn.edu>



CFANS
COLLEGE OF FOOD, AGRICULTURAL
AND NATURAL RESOURCE SCIENCES

**WATER RESOURCES
CENTER**



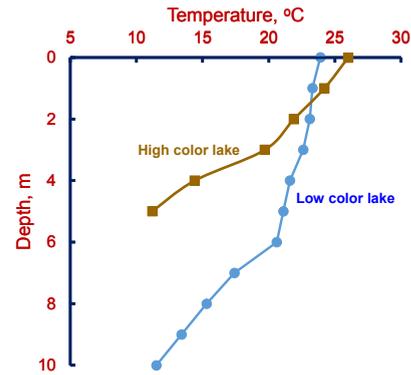
CFANS
COLLEGE OF FOOD, AGRICULTURAL
AND NATURAL RESOURCE SCIENCES

**WATER RESOURCES
CENTER**



Ecological importance

CDOM plays a central role in the behavior of freshwater ecosystems. It attenuates sunlight, and this affects physical, chemical, and biological processes in surface waters. CDOM regulates heat transfer to water, controlling lake temperatures, mixing and stratification, and UV impacts on biota are ameliorated by high CDOM levels. CDOM tends to lower the pH of water; it forms chemical complexes with heavy metals, leading to increased concentrations in colored waters, and it affects transfer of contaminants such as mercury into food webs. CDOM is an important source of reactive intermediates in the indirect photolysis of aquatic organic contaminants, and photobleaching of CDOM (time scale of weeks to months) is important in transforming bio-refractory CDOM molecules into forms more susceptible to biodegradation.



By reducing visible light, CDOM suppresses primary productivity and also affects productivity of higher trophic levels. CDOM shifts the metabolic balance in lakes toward heterotrophy, stimulating carbon burial. The key role of CDOM in regulating biogeochemical cycles and aquatic food webs is a paradigm that has emerged from research of the past two decades. In addition, CDOM affects nitrogen, phosphorus, and sulfur cycling through effects on microbial metabolism.

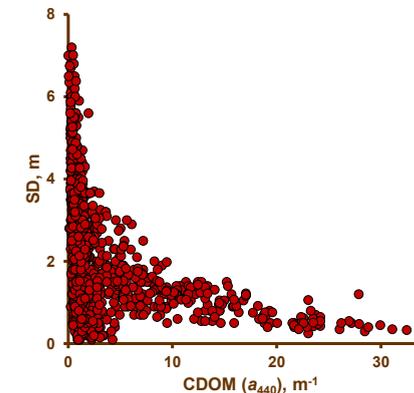
High iron concentrations often are found in highly colored waters, and it is common to hear people attribute the color to the presence of iron. This actually is not correct. Our recent studies have shown that almost all the brown color in lakes and wetlands of northern Minnesota comes from organic matter and not the varying levels of iron associated with it.

“I was surprised to learn that lake color was not due to iron.”

Unnamed homeowner,
South Lake Sturgeon.

CDOM and Secchi depth

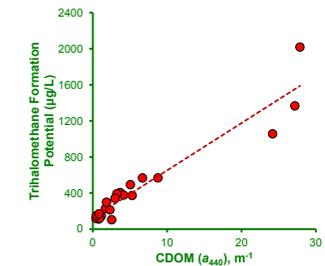
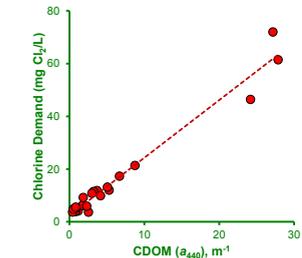
Secchi depth (SD), a measure of water clarity, is widely used to assess the eutrophication status of lakes and their suitability for recreation. In many lakes SD is controlled by chlorophyll (a measure of algal abundance), but in northern Minnesota, high color levels, which absorb light, often are more important SD controls than chlorophyll. Caution thus must be used in interpreting SD values strictly as a measure of trophic state.



SD encompasses a wide range in low-color lakes, reflecting how much chlorophyll (algae) is present, but high colored lakes have low SD because of the light-absorbing properties of CDOM. Data points represent SD and CDOM values for lakes sampled in 2015-2018.

CDOM and drinking water treatment

Although DOM is not directly harmful to human health, it has negative effects on the production of safe drinking water. DOM, especially CDOM, increases consumption of treatment chemicals such as chlorine, and CDOM levels are linearly related to the formation of potentially harmful disinfection byproducts such as trihalomethanes and haloacetic acids when water is chlorinated.





Color, chlorophyll *a*, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment

PATRICK L. BREZONIK,^{1,5} R. WILLIAM BOUCHARD JR.,² JACQUES C. FINLAY,³ CLAIRE G. GRIFFIN,³ LEIF G. OLMANSON,⁴
JESSE P. ANDERSON,² WILLIAM A. ARNOLD,¹ AND RAYMOND HOZALSKI¹

¹Department of Civil, Environmental, and Geo Engineering, University of Minnesota, Minneapolis, Minnesota 55455 USA

²Environmental Analysis and Outcomes Division, Minnesota Pollution Control Agency, 520 Lafayette Road, Saint Paul, Minnesota 55155-4194 USA

³Department of Ecology, Evolution and Behavior, University of Minnesota, Saint Paul, Minnesota 55455 USA

⁴Department of Forest Resources, University of Minnesota, Saint Paul, Minnesota 55455 USA

Citation: Brezonik, P. L., R. W. Bouchard Jr., J. C. Finlay, C. G. Griffin, L. G. Olmanson, J. P. Anderson, W. A. Arnold, and R. Hozalski. 2019. Color, chlorophyll *a*, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. *Ecological Applications* 29(3):e01871. 10.1002/eap.1871

Abstract. Secchi depth (SD), a primary metric to assess trophic state, is controlled in many lakes by algal densities, measured as chlorophyll-*a* (chl-*a*) concentration. Two other optically related water quality variables also directly affect SD: non-algal suspended solids (SS_{NA}) and colored dissolved organic matter (CDOM, expressed as the absorption coefficient at 440 nm, *a*₄₄₀). Using a database of ~1,460 samples from ~625 inland lake basins in Minnesota and two other Upper Midwest states, Wisconsin and Michigan, we analyzed relationships among these variables, with special focus on CDOM levels that influence SD values and the Minnesota SD standards used to assess eutrophication impairment of lakes. Log-transformed chl-*a*, total suspended solids (TSS), and SD were strongly correlated with each other; log(*a*₄₄₀) had major effects on log(SD) but was only weakly correlated with log(chl-*a*) and log(TSS). Multiple regression models for log(SD) and 1/SD based on the three driving variables (chl-*a*, SS_{NA}, and CDOM) explained ~80% of the variance in SD in the whole data set, but substantial differences in the form of the best-fit relationships were found between major ecoregions. High chl-*a* concentrations (> 50 µg/L) and TSS (> 20 mg/L) rarely occurred in lakes with high CDOM (*a*₄₄₀ > ~4 m⁻¹), and all lakes with *a*₄₄₀ > 8 m⁻¹ had SD ≤ 2.0 m despite low chl-*a* values (<10 µg/L) in most lakes. Further statistical analyses revealed that CDOM has significant effects on SD at *a*₄₄₀ values > ~4 m⁻¹. Thus, SD is not an accurate trophic state metric in moderately to highly colored lakes, and Minnesota's 2-m SD criterion should not be the sole metric to assess eutrophication impairment in warm/cool-water lakes of the Northern Lakes and Forest ecoregion. More generally, trophic state assessments using SD in regions with large landscape sources of CDOM need to account for effects of CDOM on SD.

Key words: chlorophyll *a*; colored dissolved organic matter; dissolved colored organic matter; ecoregion; lakes; Secchi depth; total suspended solids; trophic state; Upper Midwest.

INTRODUCTION

Secchi depth (SD), the most common indicator of lake water clarity and quality, has long played an important role in defining lake trophic state (Nürnberg 1996, Heiskary and Wilson 2008, Lottig et al. 2014). Along with chlorophyll and total phosphorus, SD is one of the three metrics in Carlson's (1977) trophic state indices. The State of Minnesota uses SD in numeric standards to determine whether or not a water body meets the water quality conditions that support its designated beneficial uses, e.g.,

aquatic life and recreational uses such as swimming (Heiskary and Wilson 2008). SD is used extensively by citizen monitoring programs to track trends in lake trophic status, owing to its simplicity and low measurement costs (e.g., Lottig et al. 2014, Heiskary and Egge 2016), and it also is useful to estimate other optical properties, such as the diffuse attenuation coefficient for photosynthetically active radiation, *K*_{PAR} (Lee et al. 2018).

As an integrative measure of water clarity, SD is determined primarily by three variables. Algal biomass, usually measured as chlorophyll *a* (chl-*a*), has a major influence of SD, and SD is often used as a proxy for algal levels. A second factor is non-algal suspended solids (SS_{NA}), including clays and other suspended minerals, that often are affected by storm events, particularly in rivers and reservoirs. The third is colored

Manuscript received 13 September 2018; revised 30 November 2018; accepted 14 January 2019. Corresponding Editor: Ryan S. King.

⁵E-mail: brezonik@umn.edu

dissolved organic matter (CDOM), which is composed largely of humic substances. In many lakes affected by eutrophication, algal biomass (*chl-a*) is the primary determinant of SD. Several studies have reported hyperbolic relationships between SD and *chl-a* and linear relationships between log SD and log *chl-a* (e.g., Carlson 1977). SS_{NA} can be a key determinant of water clarity in surface waters where watershed conditions promote soil erosion but is not commonly a major factor in natural recreational lakes of the Upper Midwest (UMW) states. Colored dissolved organic matter similarly is not important in many lakes, but high levels are common in regions dominated by forests and wetlands, such as the Northern Lakes and Forests (NLF) ecoregion (Fig. 1) in northeastern Minnesota, northern Wisconsin, Upper Michigan and the northern half of Lower Michigan (Griffin et al. 2018) and in similar ecoregions in the northeastern and southeastern U.S. (Omernik 1987, Omernik and Griffith 2014).

Colored dissolved organic matter levels are increasing in some temperate and boreal regions, e.g., Scandinavia (Haaland et al. 2010), for reasons still not fully understood. Lakes where CDOM affects SD levels thus may become more common in the future as a result of increasing precipitation (Leech et al. 2018), declining acidification (Monteith et al. 2007), or both. Evidence for the widespread occurrence of this “browning” phenomenon in the Upper Midwest is inconclusive (Brezonik et al. 2015), however, and several studies have

shown that CDOM temporal trends in individual lakes are not monotonic but driven by variations in climate and hydrologic conditions (Pace and Cole 2002, Jane et al. 2017, Carpenter and Pace 2018, Cormann et al. 2018, Leech et al. 2018).

Limnologists have long recognized that SD can be controlled by CDOM. Based on data from 470 northeastern Wisconsin lakes, Juday and Birge (1933) concluded that color had more important effects on lake SD than plankton did, and they found an inverse hyperbolic relationship between SD and lake color. Brezonik (1978) found a linear relationship between inverse SD ($1/SD$) and color using in situ mesocosms to which a concentrated source of humic color was added. Data from a Florida lake survey also showed a strong regression relationship between $1/SD$ and color and turbidity (Brezonik 1978). These studies showed that CDOM strongly influences SD at moderate to high levels of CDOM and low levels of algal biomass.

Less well known are the relationships between CDOM, SD, *chl-a*, and total suspended solids (TSS) in lakes with higher and more variable algal biomass and mineral turbidity. When both CDOM and algae co-occur at levels that affect SD, failure to consider the influence of CDOM will bias interpretations of lake impairment and trophic state based solely on SD. Consequently, the U.S. EPA (2000) recommended that development of water quality standards to assess lake eutrophication should consider CDOM as a potential confounding factor of trophic status measurements.

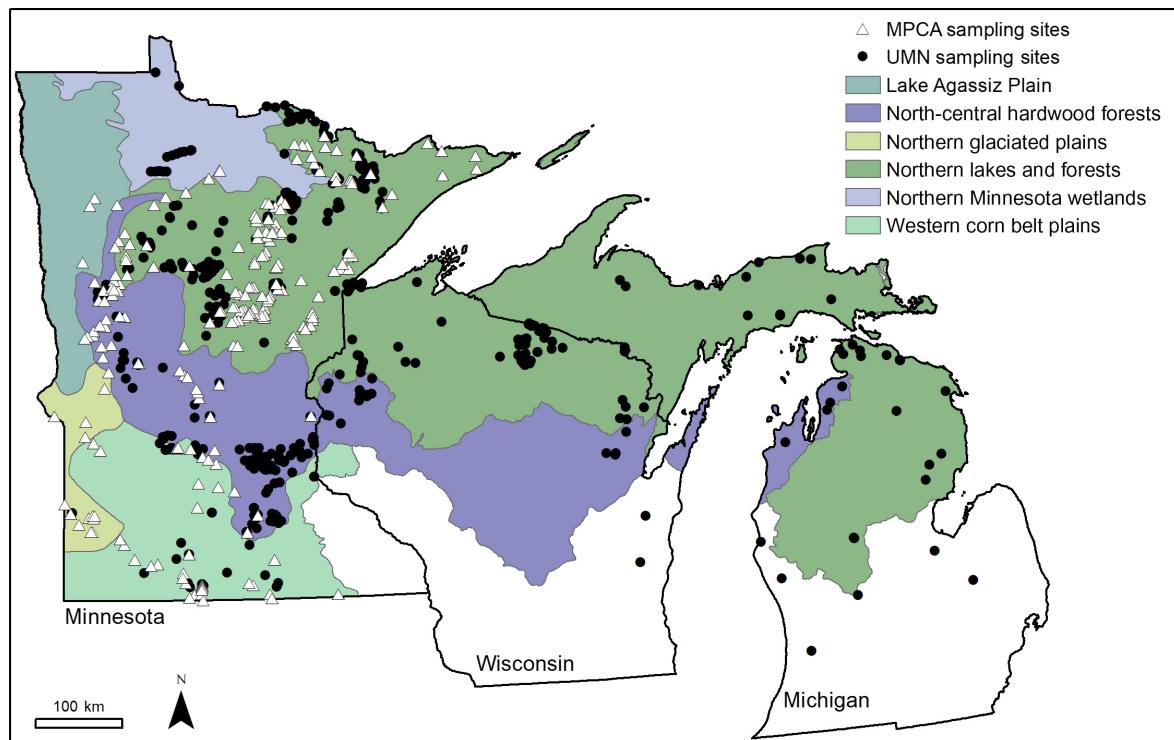


FIG. 1. Map of Upper Midwest states showing Minnesota ecoregions delineated by Omernik and Griffith (2014). Database includes lakes from all ecoregions except for DLA, where only a few small lakes occur.

Despite the now broadly recognized importance of CDOM in lake ecosystems, e.g., the nutrient-color paradigm (Williamson et al. 1999, 2014, Webster et al. 2008), its quantitative influence on SD is poorly known, partly because CDOM is measured much less often than SD and other trophic indicators (chl-*a* and total phosphorus). A recent study on coastal waters suggested that CDOM should be routinely monitored to help interpret water clarity monitoring (Harvey et al. 2015), but CDOM impacts have not yet been formally integrated with lake assessments. Here we use a large database on lakes of the UMW to examine relationships among four common optical water quality parameters, chl-*a*, TSS, CDOM, and SD, and evaluate the CDOM levels that affect interpretation of SD as a trophic state metric.

METHODS

Study area and data sources

This study focused on the large, lake-rich NLF ecoregion that extends across the UMW states of Minnesota, Wisconsin, and Michigan (Fig. 1) and comparisons with two ecoregions to its south: the North Central Hardwood Forest (NCHF) and Western Corn Belt Plains (WCBP; Omernik and Griffith 2014). The NLF is heavily forested (approximately 50%, mixed conifers and hardwoods). About one-third of its area is wetlands and lakes (Homer et al. 2015), and the high proportion of forest and wetlands in the NLF leads to an abundance of lakes with high CDOM. The ecoregion has little urban and agricultural land (4% and 7%, respectively). In contrast, nearly one-half (48%) of the NCHF is used for agriculture and 9% is urban; forests account for one-quarter of the ecoregion, and wetlands constitute 10% of land cover. Lakes are abundant throughout the NCHF, but high CDOM levels are uncommon (Griffin et al. 2018). The WCBP, southernmost ecoregion in Minnesota, is dominated by agricultural land (83%) and has only ~4% forest. It tends to have higher chlorophyll and smaller SD values than the NLF and NCHF.

We have been studying CDOM and mapping its abundance in UMW lakes since 2014 using a combination of field campaigns and satellite imagery (e.g., Brezonik et al. 2015, Olmanson et al. 2016, Griffin et al. 2018). Ground-based sampling in 2014–2015 was focused in the NLF and NCHF in northern Minnesota (Fig. 1A). In 2016 sampling was expanded to include NLF and NCHF portions that extend across Wisconsin and Michigan, as well as the Northern Minnesota Wetlands (NMW), an ecoregion in north-central Minnesota that has only ~100 lakes, very few of which are road accessible (Fig. 1B). Sample collection in 2017 was extended to the WCBP and other ecoregions in central, western, and southern Minnesota.

The Minnesota Pollution Control Agency (MPCA) routinely monitors ~150 lakes across the state each year for water quality assessments. Since 2015 they have included CDOM in their measurements. We combined

the 2014–2017 UMN data from ground-based sampling (708 site-date measurement sets) with 754 sets of 2015–2017 measurements on ground-based samples by the MPCA to produce a data set of 1,462 site-date measurements with little overlap between the two data sources. Many lakes were sampled more than once in both studies, and all observations (site-date combinations) were treated separately; i.e., multiple samples from a lake were not averaged. The final data set includes data from 251 MPCA lake basins and 382 UMN basins.

Sampling and analysis methods

Sampling procedures and field and laboratory analyses followed standard limnological practices. Detailed methods are described elsewhere (Egge et al. 2018, Griffin et al. 2018). In brief, UMN water samples were collected from ~0.25 m below the lake surface, and the MPCA collected a 0–2 m integrated sample of the epilimnion. Samples were stored in acid-washed and triple-rinsed polycarbonate or high-density polyethylene bottles and filtered for chl-*a* and dissolved constituents within 24 h of collection. Chl-*a* was filtered from water with 0.22 μm cellulose nitrate filters (0.45 μm glass fiber filters for MPCA) and stored frozen until analysis by fluorometry after 90% acetone extraction. Total suspended solids (TSS) was measured as the additional dry weight after filtration and drying at 105°C, normalized by volume. Water for CDOM analysis was filtered through 0.45 μm Geotech High Capacity filters and stored in the dark at 4°C in pre-ashed 40-mL amber glass bottles until analysis within 1 month of collection. Absorbance at 440 nm, measured using a Shimadzu (Columbia, Maryland, USA) 1601UV-PC dual beam spectrophotometer through 1- or 5-cm quartz cuvettes against a nanopure water blank, was converted to Napierian absorption coefficients (Kirk 1994) using:

$$a_{440} = \frac{2.303 A_{440}}{l} \quad (1)$$

where a_{440} is the absorption coefficient at 440 nm, A_{440} is absorbance at 440 nm, and l is cell path length (m). Absorbance scans were blank-corrected before conversion. CDOM values are reported as a_{440} .

Statistical analyses

The data were assembled into an Excel 2016 spreadsheet. Distributional statistics and principal components analysis based on inter-parameter correlation coefficients were analyzed using JMP Pro 13.1 (SAS Institute, Inc., Cary, North Carolina, USA). Generalized simple and multiple regression analyses were conducted using the Akaike information criterion (AIC) to select the best models, and least-squares regressions then were run on these models. Concentrations of non-algal suspended solids (SS_{NA}) and their log transforms were estimated as the chl-*a* detrended values of TSS, calculated as the

residuals from a regression of $\log(\text{TSS})$ vs. $\log(\text{chl-}a)$ using the whole database.

Physically based predictive equation for SD

A physically based relationship between SD and its controlling constituents (chl-*a*, CDOM, and TSS) is derived as follows for use in evaluating the effects of the three controlling variables on SD. Lorenzen (1980) and Megard et al. (1980) demonstrated that SD can be expressed in terms of an equation of the form

$$\text{SD} = \frac{\ln I_o/I_z}{k'_{\text{NA}} + k'_A \text{chl-}a} \quad (2)$$

where SD is the depth (in m) at which the Secchi disk disappears, I_z is the light intensity at the depth of disappearance, I_o is the light intensity at the water surface, k'_{NA} is the light attenuation coefficient for all non-algal constituents causing either light absorption or scattering, and k'_A is the light attenuation coefficient caused by algal biomass, expressed as chl-*a* concentration. The usual assumption is that $I_z = 0.1 I_o$ at the depth of Secchi disk disappearance (Brezonik 1978, Megard et al. 1980). The term $\ln(I_o/I_z)$ thus becomes a constant, 2.303, that is subsumed into the attenuation coefficients, and Eq. 2 becomes

$$1/\text{SD} = k_{\text{NA}} + k_A \text{chl-}a \quad (3)$$

where $k_{\text{NA}} = k'_{\text{NA}}/2.303$ and $k_A = k'_A/2.303$.

The coefficient for light attenuation by non-algal constituents, k_{NA} , is a composite coefficient that can be expanded into terms for the specific constituents causing light attenuation; e.g., for most natural waters

$$k_{\text{NA}} = k_w + k_{\text{NAP}} \times \text{NAP} + k_{\text{CDOM}} \times a_{440} \quad (4)$$

where k_w is the light attenuation coefficient for water itself; k_{NAP} is the light attenuation coefficient for non-algal particles (clays, other minerals, and organic particles not derived from or associated with algal activity); NAP is the concentration (mg/L) of these particles; k_{CDOM} is the attenuation coefficient for light absorption by CDOM; and a_{440} is the measure of CDOM. The attenuation coefficient for water itself is small compared to the other factors and not important in most lakes of interest here. Exceptions in Minnesota might be Lake Superior and the deep, ultra-oligotrophic lakes found in abandoned iron mine pits of the Mesabi Range, where SD values of 15 m or more can occur. For simplicity, we assumed k_w is negligible in the following analysis. NAP was calculated as the chl-*a* detrended TSS, i.e., SS_{NA} . Eq. 3 thus becomes

$$1/\text{SD} = k_A \times \text{chl-}a + k_{\text{NAP}} \times \text{SS}_{\text{NA}} + k_{\text{CDOM}} \times a_{440}. \quad (5)$$

Eq. 5 was used to evaluate the effects of chl-*a*, SS_{NA} , and a_{440} on SD in UMW lakes.

RESULTS

Distributions and relationships among optically important variables

The database comprised wide ranges of all four water quality variables (Fig. 2, Table 1) and encompassed a wide range of trophic conditions (ultra-oligotrophic to hypereutrophic) and lake types. All four variables had skewed distributions with a preponderance of data at the low end of the range and long tails; log transformations yielded more Gaussian-like distributions (Fig. 2). Consequently, most of our statistical analyses were performed on log-transformed data. Ecoregion-specific analyses were performed for the NLF, NCHF, and WCPB, which contain 90% of Minnesota's lakes and encompass a wide range of land uses and ecological conditions; each had 90+ sites in our database. The database had only 14–34 unique site/date measurement sets for the three other Minnesota ecoregions, and they were only included in statewide analyses.

Large differences were found between the ecoregions (also see Heiskary et al. 1987, Olmanson et al. 2014). Nearly all the lakes with high color occurred in the NLF (Griffin et al. 2018), which also had the highest mean, median, maximum, and interquartile values for a_{440} . Only 2.5% of the a_{440} values were $> 4.2 \text{ m}^{-1}$ in the NCHF, and the maximum a_{440} in the WCBP was only 4.4 m^{-1} . Mean and median concentrations of chl-*a* were in the oligotrophic range in the NLF, higher in the NCHF, and much higher in the WCBP, but extreme values $>100 \text{ } \mu\text{g/L}$ in the NLF and NCHF and $>700 \text{ } \mu\text{g/L}$ in the WCBP indicated that hypereutrophic conditions occur in all three ecoregions. Mean and median SD values showed the opposite trend, with highest values in the NLF and lowest values (indicative of hypereutrophy) in the WCBP, but the NLF also had many low SD values ($<1 \text{ m}$) associated with high CDOM levels. The highest SD in the database, 19.5 m, was measured in a deep, ultra-oligotrophic abandoned iron mine pit in the NLF. TSS concentrations generally were low to moderate (75 percentile values $<10 \text{ mg/L}$ except in the WCBP), but a few exceptions indicative of high SS_{NA} were found in all three ecoregions. Mean and median TSS values followed the same ecoregional pattern as chl-*a*, suggesting that algal biomass may be an important contributor to TSS in UMW lakes.

A correlation matrix of the four variables for the whole data set (Table 2; Appendix S1; Fig. S1) showed that log-transforms of chl-*a*, TSS, and SD were strongly correlated with each other ($r = 0.76\text{--}0.78$), but $\log a_{440}$ was only moderately correlated with $\log(\text{SD})$ ($r = 0.53$) and very weakly correlated with $\log(\text{chl-}a)$ and $\log(\text{TSS})$. The high correlation between chl-*a* and TSS supports the above suggestion that algal biomass is an important contributor to TSS in UMW lakes and that the two variables are not independent in their effects on SD.

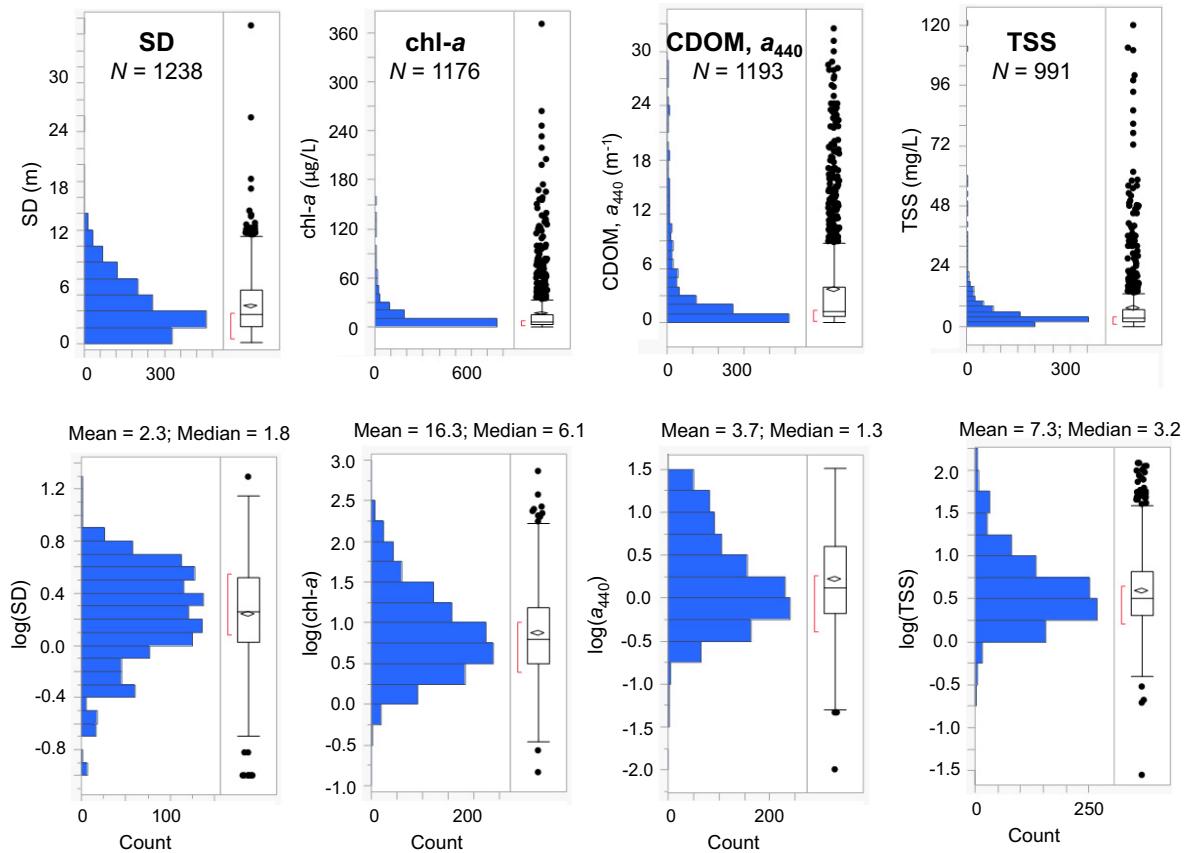


FIG. 2. Histograms and box-and-whisker plots of data distribution for the four optically related water quality variables and their log-transformations. SD, Secchi depth; CDOM a_{440} , colored dissolved organic matter, expressed as the absorption coefficient at 440 nm; TSS, total suspended solids. Mean and median values are for untransformed data. Box mid lines, median value; box edges, 25th and 75th quartile values; whiskers, 5 and 95% quantiles; points, outliers.

A principal components analysis on log-transformed values of SD, chl-*a*, TSS, and a_{440} showed that TSS and chl-*a* had similar and high loadings on the first principal component (PC1; Fig. 3). SD also had a high loading on PC1 but in a negative direction. PC1 thus can be viewed essentially as a composite trophic state variable. In contrast, PC2 had high loading from a_{440} and much lower loadings from the other three variables; PC2 thus appears to represent a (humic) color component that is largely orthogonal to the trophic state component. The essentially orthogonal relationship between TSS and a_{440} in the principal components analysis (Fig. 3) is similar to that found by Olson et al. (2018) between turbidity (an optical property related to TSS) and a_{440} in glacially fed alpine lakes of the Canadian Rocky Mountains.

Given the above findings, the negative linear relationship between log(SD) and log(chl-*a*) for all the data (Fig. 4A) was expected, but the fit ($R^2 = 0.61$) was lower than others have reported for similar relationships (e.g., $R^2 = 0.86$, Carlson 1977; $R^2 = 0.82$ and 0.69 ; Nürnberg 1996). Many outliers in Fig. 4A are below the line of

best fit; in these cases, SD was less than expected based on chl-*a* concentration, implying that some other factor (s) also affected SD. Removal of samples with $a_{440} > 3.0 m^{-1}$, which was found to be a limiting value for CDOM domination by allochthonous sources (Griffin et al. 2018), improved the fit to $R^2 = 0.76$. A few sites, mostly with high TSS, however, still fell far from the regression line (Fig. 4B).

The three highly correlated variables, chl-*a*, TSS, and SD, all had complicated relationships with CDOM (a_{440}). High values of the first two (Fig. 5A, B) were essentially orthogonal to the a_{440} distribution, with chl-*a* concentrations $> \sim 50 \mu g/L$ and TSS $> 20 mg/L$ occupying narrow ranges of CDOM (a_{440} generally $< 3.5 m^{-1}$ and $4.5 m^{-1}$, respectively), supporting the nutrient-color paradigm (Williamson et al. 1999). Of the 203 samples with both measured a_{440} and chl-*a* values, only two samples with $a_{440} > 5 m^{-1}$ had chl-*a* $> 50 \mu g/L$. One value ($57 \mu g/L$; Fox Lake, Minnesota, USA) was from a shallow, bog-stained NLF lake with both wetlands and agricultural activity in its riparian zone. The other value ($64 \mu g/L$, shallow, wetland-dominated Turner Lake, near

TABLE 1. Statistical summaries of optical water quality variables for all data and for data separated by ecoregion.

Parameter	a_{440} (m^{-1})	Chl- <i>a</i> ($\mu g/L$)	SD (m)	TSS (mg/L)
All data				
Mean	3.68	16.74	2.31	7.26
Median	1.31	6.33	1.80	3.20
Standard deviation	5.45	36.28	1.72	12.94
SEE	0.16	1.06	0.05	0.41
Maximum	32.5	721	19.5	120
75th percentile	4.01	15.36	3.3	6.58
25th percentile	0.67	3.14	1.0	2.00
<i>N</i>	1,193	1,177	1,238	991
NLF ecoregion				
Mean	4.79	8.6	2.60	3.77
Median	1.93	5.1	2.13	2.80
Standard deviation	6.22	10.6	1.76	4.67
SEE	0.22	0.4	0.06	0.18
Maximum	32.5	124.8	19.51	93.4
75th percentile	6.7	10.0	3.6	4.4
25th percentile	0.7	2.8	1.3	2.0
<i>N</i>	793	749	823	655
NCHF ecoregion				
Mean	1.23	26.11	2.12	8.77
Median	0.97	8.19	1.57	4.80
Standard deviation	1.00	29.25	1.65	14.09
SEE	0.06	1.76	0.1	0.98
Maximum	8.27	264	7.4	120
75th percentile	1.51	25.4	2.8	9.2
25th percentile	0.57	3.7	0.9	2.4
<i>N</i>	268	275	248	205
WCBP ecoregion				
Mean	1.48	73.9	0.77	29.9
Median	1.20	42.3	0.52	21.5
Standard deviation	0.88	97.5	0.64	24.6
SEE	0.10	105.2	0.07	2.8
Maximum	4.42	721	3.4	111
75th percentile	2.13	111	1.1	39.2
25th percentile	0.84	17.3	0.3	11.2
<i>N</i>	73	86	92	76

Note: Ecoregions are NLF, northern lakes and forests; NCHF, north central hardwood forest; and WCBP, western corn belt plains. SEE, standard error of estimate.

TABLE 2. Correlation matrix of Pearson *r* values for the four water quality variables.

	$\log(a_{440})$	$\log(\text{chl-}a)$	$\log(\text{SD})$	$\log(\text{TSS})$
$\log(a_{440})$	1.000	0.253	-0.528	0.147
$\log(\text{chl-}a)$		1.000	-0.761	0.778
$\log(\text{SD})$			1.000	-0.759
$\log(\text{TSS})$				1.000

Brainerd, Minnesota, USA) is an anomaly insofar as the next highest chl-*a* value in two seasons of monthly sampling was only 27 $\mu g/L$. Similarly, of the 192 samples with both measured a_{440} and TSS, only two with $a_{440} > 5 m^{-1}$ had TSS > 15 mg/L. Both were from sites

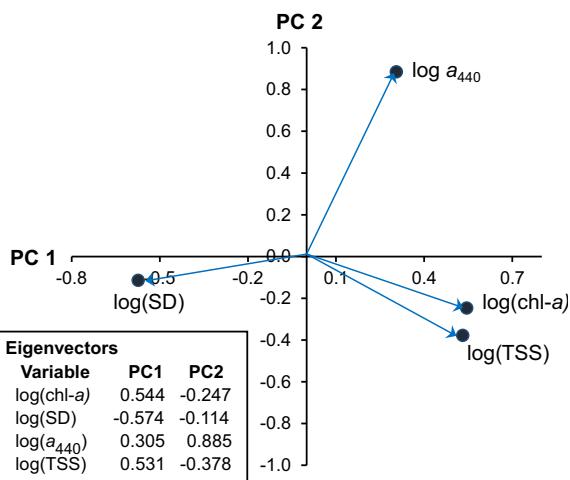


FIG. 3. PC1 vs. PC2 for principal component analysis of the four optical water quality variables. Inset box shows eigenvectors for the first two principal components.

in the St. Louis River Estuary of Lake Superior that are influenced by runoff from the Pokegama River, which drains a region with highly erodible clay soils (Roesler et al. 2018).

The distribution of SD vs. a_{440} was somewhat broader than that of chl-*a* or TSS (Fig. 5C), but almost no SD values > 3.0 m occurred for $a_{440} > 3 m^{-1}$, and no SD values > 2 m were found for $a_{440} > 8 m^{-1}$ (Fig. 5D). High CDOM levels apparently are antithetical to production of high levels of algal biomass and related organic suspended solids in UMW lakes, as Thrane et al. (2014) found for boreal lakes. Despite the low chl-*a* concentrations associated with high-CDOM waters, their SD values are small because of light absorption by CDOM. Similarly, catchment factors that promote high export of CDOM into UMW lakes, such as wetlands and other poorly drained landscape features, apparently are not favorable for high export of mineral or non-algal organic suspended solids into lakes.

SD predictive relationships

We evaluated relationships between SD and its controlling variables singly and in combination using log-transformed values for the whole data set and separately for the NLF, NCHF, and WCBP ecoregions using generalized regression analysis and the AIC to select the best models (Table 3). For the whole database, the best-fit relationship ($R^2 = 0.80$) included all three predictor variables, but a two-variable relationship with chl-*a* and a_{440} was found to be best ($R^2 = 0.76$) for the NLF. Although the three-variable model had slightly higher R^2 , addition of TSS added little explanatory power and increased the AIC, probably because TSS concentrations in this ecoregion are low and generally associated with chl-*a*. In contrast, a two-variable model using chl-*a* and TSS was the

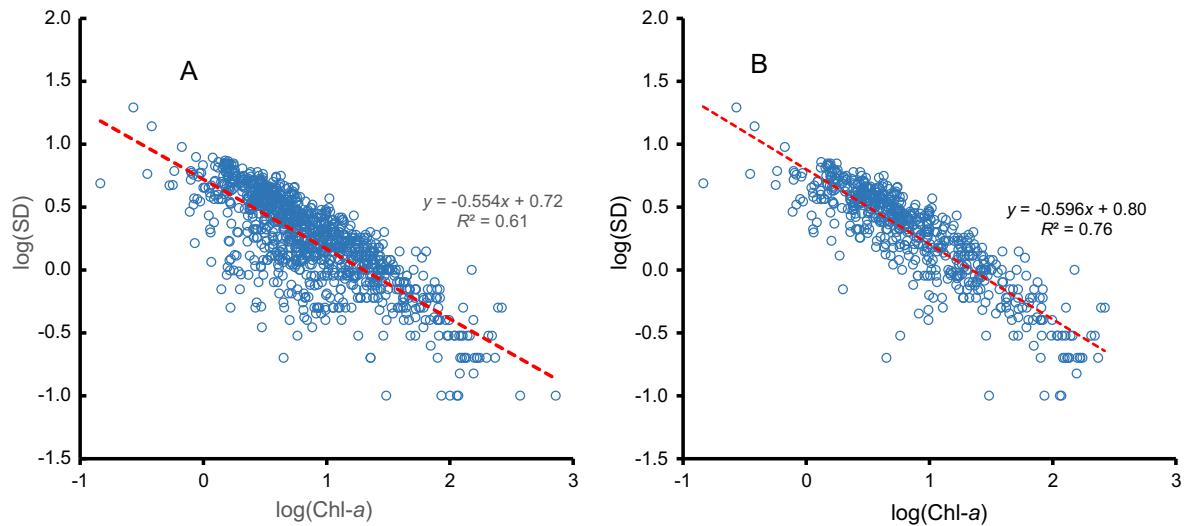


FIG. 4. (A) $\log(\text{SD})$ vs. $\log(\text{chl-}a)$ for all samples in the UMN-MPCA 2014–2017 database; (B) same plot with samples having $a_{440} > 3 \text{ m}^{-1}$ removed.

best predictive model for the NCHF, and CDOM was not an important predictor of SD in the NCHF owing to the relative homogeneity of lake color in the ecoregion. More ambiguous results were found for the WCBP, with two two-variable models yielding similar R^2 and AIC values (Table 3). R^2 values for both models were lower than those for the best models for the whole database and NLF and NCHF ecoregions, reflecting more compressed variable ranges in this nutrient-rich ecoregion (Table 1).

In terms of simple bivariate relationships, $\log(\text{TSS})$ and $\log(\text{chl-}a)$ had nearly the same predictive values of $\log(\text{SD})$ for the whole database (Table 3), which likely reflects the important role of algae as a TSS source in lakes of this study. In contrast, $\log(\text{SS}_{\text{NA}})$ was a poor predictor ($R^2 = 0.08$), supporting the idea that SS_{NA} is not an important control on SD in most UMW lakes. $\log(a_{440})$ by itself also was a poor predictor of $\log(\text{SD})$ for the whole database, no doubt because of the L-shaped distribution of the $\text{SD}-a_{440}$ relationship (Fig. 5C).

Because of the strong correlation between $\text{chl-}a$ and TSS, we were concerned that including both variables in the regression could “double-count” the influence of $\text{chl-}a$. A three-term SD regression in which TSS was replaced by SS_{NA} gave the same fit (R^2 and RMSE), but increased the equation coefficient for $\log(\text{chl-}a)$ (Table 3). The RMSE for the three-term $\log(\text{SD})$ equation (0.163) translates to an SD of 1.45 m.

The above log-log relationships can be criticized as empirical, and the physically based relationship between SD and its controlling constituents (Eq. 5) derived in the Methods section was applied to the $\text{chl-}a$, a_{440} , and SS_{NA} data for all sites in the database using multiple regression analysis. Although the R^2 (0.79) for this equation is comparable to that for the three-term log SD

equation (0.80, Table 3), the RMSE (0.54) translates to larger uncertainty in SD (1.85 m). Moreover, the large intercept value implies that the equation applies only to relatively low SD values; an $\text{SD} > 2.8 \text{ m}$ would require negative values for one or more of the predictor variables, which clearly is not possible.

Reasons for the limitations of the $1/\text{SD}$ equation are found in the predictor variable distributions (Fig. 5A–C). High $\text{chl-}a$ and TSS concentrations occurred nearly exclusively at low a_{440} , and these high values caused low SD (Fig. 4, Table 2); low SD is equivalent to high $1/\text{SD}$. High $\text{chl-}a$ and TSS at some sites with low a_{440} thus promoted high $1/\text{SD}$ values, which distorted the $1/\text{SD}$ vs. a_{440} relationship, effectively decreasing the slope of the best-fit line between $1/\text{SD}$ and a_{440} . This distortion can be seen in a plot of $1/\text{SD}$ vs. a_{440} (Fig. 6A), which includes only sites with $\text{chl-}a < 10 \mu\text{g/L}$. High $1/\text{SD}$ values ($> \sim 1 \text{ m}^{-1}$) for sites with $a_{440} < 3 \text{ m}^{-1}$ were associated with high TSS (typically 15–20 mg/L). Given the low $\text{chl-}a$ concentrations at these sites, the TSS likely was mostly SS_{NA} (e.g., clay minerals or non-algal organic matter from allochthonous sources or macrophytes). These high values pulled the regression line upward at low a_{440} and decreased the slope and fit (R^2) of the relationship, indicating that the relationship should not be used to quantify CDOM effects on SD. A plot of $1/\text{SD}$ vs. a_{440} for sites with $\text{chl-}a < 10 \mu\text{g/L}$ and $a_{440} > 3 \text{ m}^{-1}$ effectively eliminated the sites affected by high SS_{NA} (Fig. 6B).

DISCUSSION

Relationships of SD with other optical variables

SD often is used as a surrogate for estimating algal biomass and trophic state and, in many lakes, SD alone

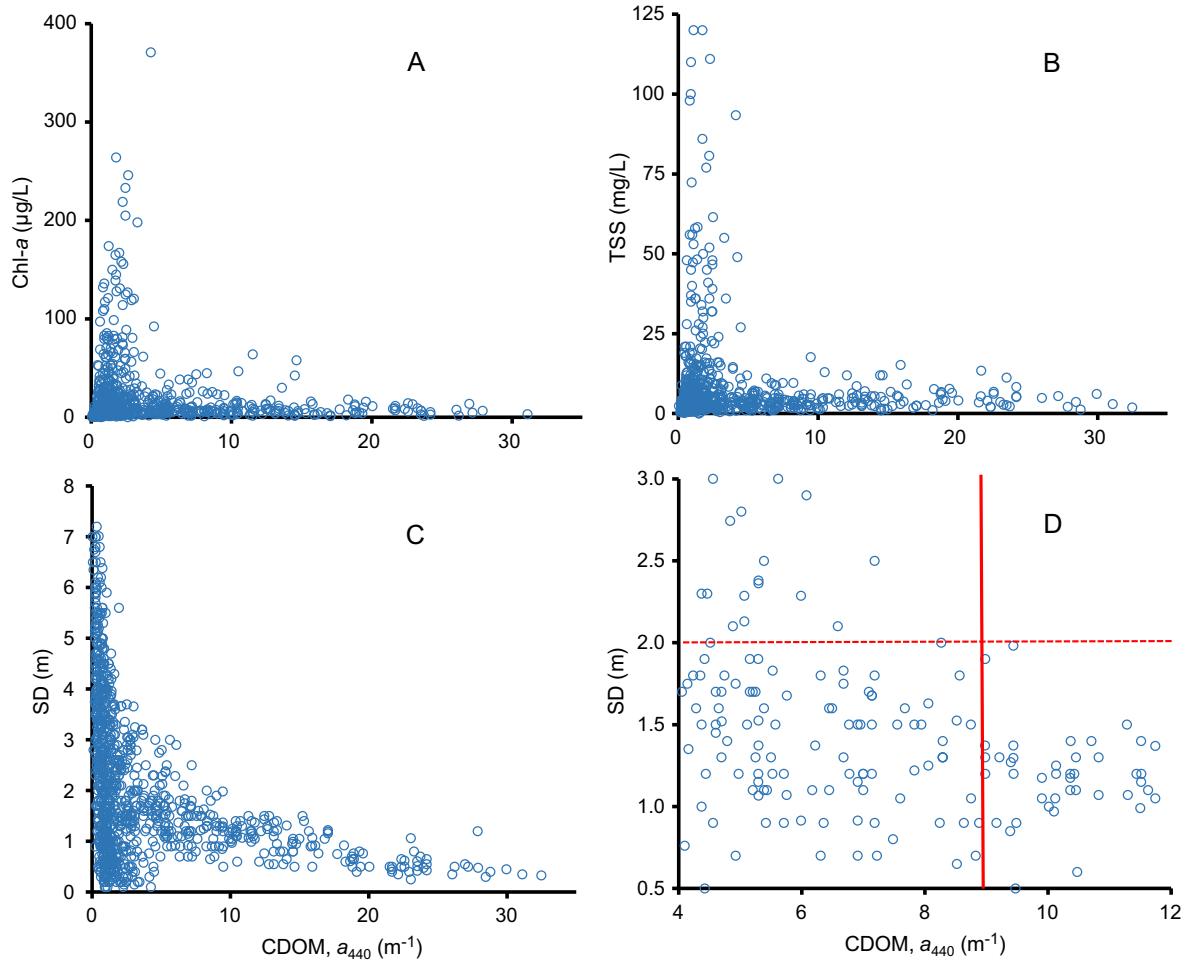


FIG. 5. (A) Chl-*a*, (B) TSS, and (C) SD vs. CDOM (a_{440}) for all UMN-MPCA 2014–2017 data. (D) Enlargement of SD vs. CDOM plot focusing on range where CDOM limits SD to <2.0 m. Horizontal dashed red line denotes 2-m SD water quality standard for warm/cool-water northern lakes and forests lakes in Minnesota; vertical red line indicates CDOM level above which no lakes in the database had SD \geq 2.0 m.

is a good estimator of trophic state (Carlson 1977, Nürnberg 1996, Heiskary and Wilson 2008). SD is also a good measure for evaluating recreational suitability, as users often respond to water clarity when deciding whether a water body is suitable for recreation (Smeltzer and Heiskary 1990, Heiskary and Wilson 2005). For these reasons and the widespread availability of data, the U.S. EPA recommended the use of SD to develop lake and reservoir nutrient standards (U.S. EPA 2000).

SD thus is now widely accepted for trophic state assessment, but our results indicate that it cannot be applied uniformly. Moreover, there is a long history of studies demonstrating that algal production (chl-*a* concentration) does not always control SD. Juday and Birge (1933) found an inverse curvilinear relationship between SD and lake color similar to the trend in our SD- a_{440} data (Fig. 5C). They obtained the relationship by dividing data on 470 lakes from northeastern Wisconsin into 11 groups with SD ranges of 0–0.9 m to 7.5–9.4 m and

plotting mean SD for the groups vs. mean color determined visually and reported in platinum-cobalt units (PCU) for the lakes in each group. Differences in measurement methods and reporting units between our results and those of Juday and Birge make quantitative comparisons difficult, but the inverse curvilinear trends in the two data sets are similar.

Brezonik (1978) found close-fitting, straight-line relationships between 1/SD and color in experiments where a concentrated humic color source was added to mesocosm enclosures in two Florida lakes. Color was determined colorimetrically and reported in PCU. The slope of the 1/SD-color (PCU) relationship was 0.0040, which translated to a slope of ~ 0.074 when the absorption coefficient of chloroplatinate at 440 nm was used to convert color in PCU to a_{440} in m^{-1} . This is similar to the slope (0.069) for the 1/SD vs. a_{440} relationship of colored lakes ($a_{440} > 3 m^{-1}$) in Fig. 6B but lower than that for the three-variable model based on all UMW data (first

TABLE 3. Regression equations to predict log(SD) and 1/SD†.

Data set	Variables‡	AIC	Adjusted R^2	RMSE	N	Equation
Best equation based on the Akaike information criterion (AIC)						
All	chl- a , a_{440} , TSS	-560	0.80	0.16	718	$\log(\text{SD}) = 0.722 - 0.202 \times \log(\text{chl-}a) - 0.240 \times \log(a_{440}) - 0.446 \times \log(\text{TSS})$
All	chl- a , a_{440} , TSS	1162	0.79	0.54	719	$1/\text{SD} = 0.790 + 0.0095 \times \text{Chl-}a + 0.052 \times a_{440} + 0.054 \times \text{TSS}$
NLF	chl- a , a_{440}	-531	0.75	0.15	581	$\log(\text{SD}) = 0.619 - 0.283 \times \log(\text{chl-}a) - 0.334 \times \log(a_{440})$
NCHF	chl- a , TSS	-148	0.87	0.14	150	$\log(\text{SD}) = 0.88 - 0.26 \times \log(\text{chl-}a) - 0.334 \times \log(\text{TSS})$
WCBP	chl- a , TSS	-47	0.73	0.17	72	$\log(\text{SD}) = 0.687 - 0.153 \times \log(\text{chl-}a) - 0.594 \times \log(\text{TSS})$ §
WCBP	TSS, a_{440}	-49	0.74	0.15	58	$\log(\text{SD}) = 0.535 - 0.655 \times \log(\text{TSS}) - 0.184 \times \log(a_{440})$ §
Bivariate relationships						
All	chl- a	-132	0.60	0.229	1081	$\log(\text{SD}) = 0.711 - 0.545 \times \log(\text{chl-}a)$
All	TSS	-56	0.60	0.234	925	$\log(\text{SD}) = 0.628 - 0.677 \times \log(\text{TSS})$
All	a_{440}	457	0.27	0.305	977	$\log(\text{SD}) = 0.292 - 0.341 \times \log(a_{440})$
All	SS_{NA}	690	0.08	0.357	879	$\log(\text{SD}) = 0.231 - 0.403 \times \log(\text{SS}_{\text{NA}})$

Note: SS_{NA}

† All regressions and coefficients significant at $P < 0.0001$ except as noted in footnote §.

‡ All independent variables were log-transformed except in row 2 (regression vs. 1/SD).

§ $P = 0.009$ for chl- a and 0.04 for a_{440} in the WCBP regression equations.

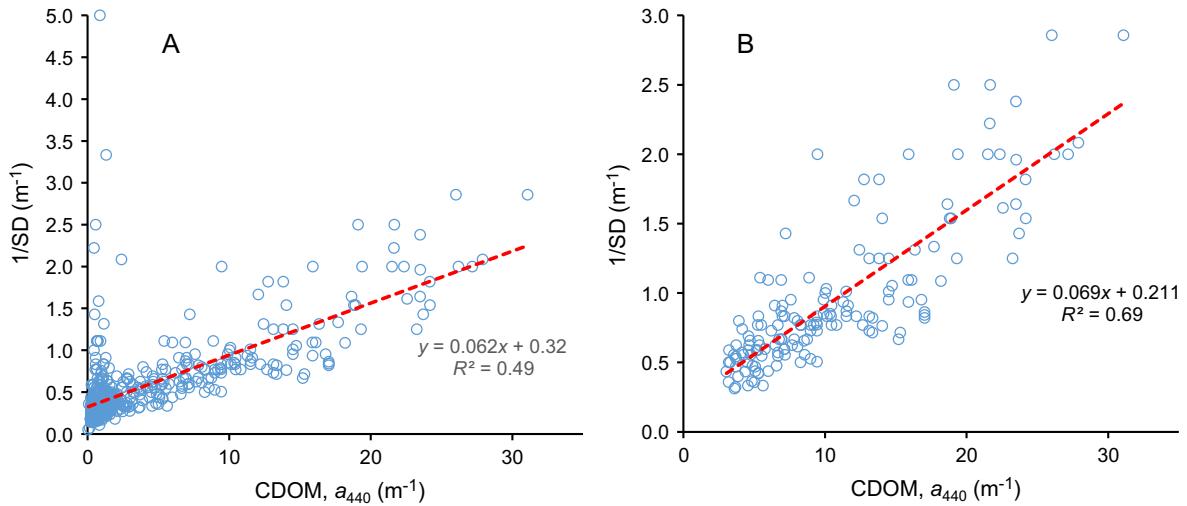


FIG. 6. (A) The 1/SD vs. a_{440} and regression line for all sites in the database with chl- $a \leq 10 \mu\text{g/L}$; (B) same plot and regression but for sites with chl- $a \leq 10 \mu\text{g/L}$ and $a_{440} > 3.0 \text{ m}^{-1}$.

equation in Table 3). As described above, the latter was influenced by high TSS and chl- a values at low a_{440} , which raised the regression line at low color and decreased the slope of the relationship.

Several previous studies also have reported regression relationships between log SD or 1/SD and optical water quality variables. Using data from a survey on 55 lakes in north and central Florida (Shannon and Brezonik 1972), Brezonik (1978) found a strong relationship ($R^2 = 0.89$) between 1/SD and color and turbidity: $1/\text{SD} = 0.106 + 0.128(T) + 0.0025(C)$, where T was laboratory-measured turbidity in standard formazin (nephelometric) turbidity units and C was color measured in PCU by colorimetry. Turbidity accounts for light attenuation by all suspended particles, both

algal and non-algal, and color accounts for light attenuation for light-absorbing CDOM. A regression of 1/SD vs. color and chl- a was not as strong ($R^2 = 0.63$). Canfield and Hodgson (1983) reported an R^2 of 0.79 for a regression of $\ln(\text{SD})$ vs. $\ln(\text{chl } a)$ and $\ln(\text{color})$ based on data from 165 Florida lakes. Nürnberg (1996) found an R^2 of 0.88 for a regression of $\log(\text{SD})$ vs. $\log(\text{chl-}a)$ and $\log(\text{color})$ for 33 lakes from northeastern North America and an R^2 of 0.79 for another data set of 91 lakes worldwide.

Although the three-term regression models (both for $\log[\text{SD}]$ and 1/SD) explained a high proportion (~80%) of the variance in SD for the whole database, they still did not account for ~20% of the SD variance, and the models had high RMSE values (1.45–1.85 m) that limit

their usefulness for SD predictions. The unexplained variance in the models likely reflects measurement uncertainties in the variables, especially SD, values for which depend on such factors as sunlight intensity/angle, wave activity, and observer biases. Uncertainties from model assumptions also cannot be ruled out. For example, chl-*a* is an imperfect measure of the effect of algae on SD because it does not take into account differences in “packaging”; i.e., a given chl-*a* concentration associated with algal cells clumped into visibly large particles will have less effect on SD than the same concentration associated with discrete free-floating cells. Moreover, chl-*a* to cell-volume relationships differ between algal taxa, and chl-*a* concentrations within algal cells vary with environmental conditions (Reynolds 1984, Felip and Catalan 2000). Similarly, TSS is a mass-related variable, but the amount of light scattered by suspended particles depends on particle numbers, shapes, and surface properties more than on mass itself. Thus, the multiple regression results should be viewed as tools to understand how limnological characteristics affect SD, rather than as equations that exactly predict SD given certain physical/optical measurements.

The PCA results (Fig. 3) support the ideas of the nutrient-color paradigm (Williamson et al. 1999, Webster et al. 2008, Fergus et al. 2016), which views trophic (food web) processes in lakes as defined by two orthogonal drivers: CDOM and TP. The former promotes heterotrophy and the latter autotrophy with attendant effects on algal (chl-*a*) production. Although the influence of CDOM on algal biomass and production is less clear at large (continental) scales (Havens and Nurnberg 2004, Yuan and Pollard 2014), more consistent negative impacts have been observed at local to regional scales (e.g., Karlsson et al. 2009, Fergus et al. 2016), mediated by the strong reduction of light availability. Recent research shows that despite modest positive impacts on nutrient availability, CDOM suppresses primary production and food web production in colored lakes (Karlsson et al. 2009, Thrane et al. 2014, Creed et al. 2018).

Implications for SD trophic state standards for CDOM-rich lakes

We used our results to evaluate the influence of CDOM on SD and specifically the levels of CDOM that interfere with interpreting SD as a measure of algal abundance and lake trophic state. The State of Minnesota has adopted eutrophication standards for warm- and cold- (i.e., trout) water NLF lakes. For example, the standards for warm-water NLF lakes include water quality criteria of 30 µg/L for total phosphorus (TP), 9 µg/L for chl-*a*, and 2.0 m for SD as June–September averages (Heiskary and Wilson 2005). Nonattainment of this standard occurs when the TP criterion is exceeded and either or both chl-*a* and SD are in non-attainment. In evaluating whether warm-water NLF lakes satisfy

water quality conditions for their designated beneficial uses, use of the SD criterion is based on the assumption that SD is controlled by algal abundance. Inspection of the SD vs. a_{440} plots in Fig. 5C, D shows that no site with $a_{440} > 8 \text{ m}^{-1}$ had an SD > 2.0 m. Of the 128 sites having $a_{440} > 8.0 \text{ m}^{-1}$ and also having chl *a* data, chl *a* was < 10 µg/L in 100 cases; 20 had chl-*a* of 10–20 µg/L, and only 8 had chl-*a* > 20 µg/L. It thus is apparent that SD is limited to values < 2.0 m primarily by CDOM for sites with $a_{440} > 8 \text{ m}^{-1}$ and that an SD limit of 2.0 m is unlikely to be a realistic trophic state criterion for such waters.

The above result can be considered an “upper limit,” however, insofar as it is highly likely that a_{440} values < 8 m^{-1} also affect SD. To evaluate the a_{440} level at which this begins to occur (relative to the 2-m SD criterion), we used the regression results described above. Insertion of the MPCA’s chl-*a* criterion of 9 µg/L for NLF lakes and SD = 2 m into the 1/SD predictive equation (Table 3) led to a negative a_{440} , even at $SS_{\text{NA}} = 0$, however, and clearly that is not possible. As described above, high chlorophyll and/or TSS concentrations and associated low SD that occurred at some low-CDOM sites distorted the 1/SD vs. a_{440} relationship (Fig. 6A). The plot of 1/SD vs. a_{440} for sites with chl-*a* < 10 µg/L and $a_{440} > 3 \text{ m}^{-1}$ (Fig. 6B), however, eliminated sites affected by high SS_{NA} . The best-fit line for this data set yielded $a_{440} = 4.2 \text{ m}^{-1}$ for an SD of 2 m. We suggest that this is a reasonable value for limiting use of the 2-m SD water quality standard as an indicator of eutrophication impairment in NLF lakes. As a minimum, it should serve as a threshold where additional response data (e.g., chl-*a*) are needed to assess lake eutrophication. Because SD standards vary by state and ecoregion, the specific CDOM threshold value may be different in other settings.

The above analysis is not to suggest that lakes with high CDOM ($a_{440} > \sim 4 \text{ m}^{-1}$) cannot have eutrophication problems, although high levels of CDOM tend to promote heterotrophy rather than autotrophy and may suppress algal growth, thus counteracting some effects of eutrophication (e.g., Williamson et al. 1999, Webster et al. 2008). Instead, these results indicate that SD is not an effective predictor of such problems in lakes with color higher than $a_{440} \sim 4 \text{ m}^{-1}$.

Restricting the use of SD as a water quality standard based on a threshold level of CDOM indicates that greater reliance on indicators such as TP and chl-*a* is needed to assess impacts of eutrophication in the NLF. TP, TN, and chl-*a* are commonly used trophic state parameters that are linked to key ecosystem services (Keeler et al. 2012) and water quality standards (Heiskary and Wilson 2008). Colored dissolved organic matter has been shown to influence all of these parameters in oligotrophic lakes (e.g., Karlsson et al. 2009). Increasing CDOM levels in many lakes across northern Europe and parts of North America (Monteith et al. 2007, Williamson et al. 2015, Cormann et al. 2018) also may be

increasing the effects of CDOM on trophic state metrics in these regions. Although these influences may be somewhat subtle compared to human influences on nutrients, they may require further modification of trophic state standards in moderate-to-high CDOM lakes, where availability and cycling of nutrients contrast strongly with those of low-CDOM waters.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grant (CBET 1510332), the Minnesota Environmental and Natural Resources Trust Fund, as recommended by the Legislative-Citizen Commission on Minnesota Resources, the Minnesota Pollution Control Agency's lake water quality assessment program, and the University of Minnesota's Office of the VP for Research and Retirees Association, U-Spatial Program, Sea Grant Program, and Agricultural Experiment Station. We thank numerous collaborators, research staff, and students for assistance in sample collection and analysis and colleague Marvin Bauer for advice and encouragement throughout the larger project, of which this paper is a component.

LITERATURE CITED

- Brezonik, P. L. 1978. Effect of organic color and turbidity on Secchi disk transparency. *Journal of the Fisheries Research Board of Canada* 35:1410–1416.
- Brezonik, P. L., L. G. Olmanson, J. C. Finlay, and M. E. Bauer. 2015. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sensing of Environment* 157:199–215.
- Canfield Jr., D. E., and L. M. Hodgson. 1983. Prediction of Secchi disc depths in Florida lakes: impact of algal biomass and organic color. *Hydrobiologia* 99:51–60.
- Carlson, R. E. 1977. A trophic state index for lakes. *Limnology & Oceanography* 22:361–369.
- Carpenter, S. R., and M. L. Pace. 2018. Synthesis of a 33-yr series of whole-lake experiments: Effects of nutrients, grazers, and precipitation-driven water color on chlorophyll. *Limnology & Oceanography Letters* 3:419–427.
- Corman, J., B. Bertolet, N. Casson, S. Sebestyen, R. Kolka, and E. Stanley. 2018. Nitrogen and phosphorus loads to temperate seepage lakes associated with allochthonous dissolved organic carbon loads. *Geophysical Research Letters* 45:5481–5490.
- Creed, I. F., A.-K. Bergström, C. G. Trick, N. B. Grimm, D. O. Hessen, J. Karlsson, K. A. Kid, et al. 2018. Global change-driven effects on dissolved organic matter composition: Implications for food webs of northern lakes. *Global Change Biology* 24:3692–3714.
- EGGE, L., L. ENGEL, P. ANDERSON, and K. O'HARA. 2018. Standard operating procedures. Intensive watershed monitoring. Lake water quality sampling. Minnesota Pollution Control Agency, St. Paul, Minnesota, USA. document wq-s1-16. <https://www.pca.state.mn.us/sites/default/files/wq-s1-16.pdf>
- Felip, M., and J. Catalan. 2000. The relationship between phytoplankton biovolume and chlorophyll in a deep oligotrophic lake: decoupling in the spatial and temporal maxima. *Journal of Plankton Research* 22:91–150.
- Fergus, C. E., A. O. Finley, P. A. Soranno, and T. Wagner. 2016. Spatial variation in nutrient and water color effects on lake chlorophyll at macroscales. *PLoS ONE* 11:e0164592.
- Griffin, C. G., J. C. Finlay, P. L. Brezonik, L. G. Olmanson, and R. M. Hozalski. 2018. Limitations on using CDOM as a proxy for DOC in temperate lakes. *Water Research* 144:719–727.
- Haaland, S., D. Hongve, H. Laudon, G. Riise, and R. D. Vogt. 2010. Quantifying the drivers of the increasing colored organic matter in Boreal surface waters. *Environmental Science & Technology* 44:2975–2980.
- Harvey, E. T., S. Kratzer, and A. Andersson. 2015. Relationships between colored dissolved organic matter and dissolved organic carbon in different coastal gradients of the Baltic Sea. *Ambio* 44(Suppl. 3):S392–S401.
- Havens, K. E., and G. K. Nurnberg. 2004. The phosphorus-chlorophyll relationship in lakes: potential influences of color and mixing regime. *Lake & Reservoir Management* 20:188–196.
- Heiskary, S. A., and L. Egge. 2016. Page 24. A review of Secchi transparency trends in Minnesota lakes. Minnesota Pollution Control Agency, St. Paul, Minnesota, USA.
- Heiskary, S. A., and C. B. Wilson. 2005. Page 176. Minnesota lake water quality assessment report: Developing nutrient criteria. Minnesota Pollution Control Agency, St. Paul, Minnesota, USA.
- Heiskary, S. A., and C. B. Wilson. 2008. Minnesota's approach to nutrient criteria development. *Lake & Reservoir Management* 24:282–297.
- Heiskary, S. A., C. B. Wilson, and D. P. Larsen. 1987. Analysis of regional patterns in lake water quality: using ecoregions for lake management in Minnesota. *Lake & Reservoir Management* 3:337–344.
- Homer, C. G., J. A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. D. Herold, J. D. Wickham, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* 81:345–354.
- Jane, S. F., L. A. Winslow, C. K. Remucal, and K. C. Rose. 2017. Long-term trends and synchrony in dissolved organic matter characteristics in Wisconsin, USA, lakes: Quality, not quantity, is highly sensitive to climate. *Journal of Geophysical Research Biogeosciences* 122:546–561.
- Juday, C., and E. A. Birge. 1933. The transparency, the color, and the specific conductance of the lakes in northeastern Wisconsin. *Transactions of the Wisconsin Academy of Science, Arts, and Letters* 28:205–259.
- Karlsson, J., P. Byström, J. Ask, L. Persson, and M. Jansson. 2009. Light limitation of nutrient-poor lake ecosystems. *Nature* 460:506–508.
- Keeler, B. L., S. Polasky, K. A. Brauman, K. A. Johnson, J. C. Finlay, A. O'Neill, K. Kovacs, and B. Dalzell. 2012. Linking water quality and well-being for improved assessment and valuation of ecosystem services. *Proceedings of the National Academy of Sciences* 109:18619–18624.
- Kirk, J. T. O. 1994. *Light and photosynthesis in aquatic ecosystems*, 2nd edition. Cambridge University Press, Cambridge, UK.
- Lee, Z., S. Shang, K. Du, and J. Wei. 2018. Resolving the long-standing puzzles about the observed Secchi disk relationships. *Limnology & Oceanography* 63. <https://doi.org/10.1002/lno.10940>
- Leech, D. M., A. I. Pollard, S. G. Labou, and S. E. Hampton. 2018. Fewer blue lakes and more murky lakes across the continental U.S.: Implications for planktonic food webs. *Limnology & Oceanography* 63:2661–2680.
- Lorenzen, M. W. 1980. The use of chlorophyll-Secchi disk relationships. *Limnology & Oceanography* 25:371–372.
- Lottig, N. R., T. Wagner, E. N. Henry, K. S. Cheruvilil, K. E. Webster, J. A. Downing, and C. A. Stow. 2014. Long-term

- citizen-collected data reveal geographical patterns and temporal trends in lake water clarity. *PLoS ONE* 9:e95769.
- Megard, R. O., J. C. Settles, H. A. Boyer, and W. S. Combs Jr. 1980. Light, Secchi disks, and trophic states. *Limnology & Oceanography* 25:373–377.
- Monteith, D. T., J. L. Stoddard, C. D. Evans, H. A. De Wit, M. Forsius, T. Høgåsen, A. Wilander, B. L. Skjelkvåle, D. S. Jeffries, and J. Vuorenmaa. 2007. Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. *Nature* 450:537.
- Nürnberg, G. K. 1996. Trophic state of clear and colored, soft- and hardwater lakes with special consideration of nutrients, anoxia, phytoplankton and fish. *Lake & Reservoir Management* 12:432–447.
- Olmanson, L. G., P. L. Brezonik, and M. E. Bauer. 2014. Geospatial and temporal analysis of a 20-year record of Landsat-based water clarity in Minnesota's 10,000 lakes. *Journal of the American Water Resources Association* 50:748–761.
- Olmanson, L. G., P. L. Brezonik, J. C. Finlay, and M. E. Bauer. 2016. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sensing of Environment* 185:119–128.
- Olson, M. H., J. M. Fischer, C. E. Williamson, E. P. Overholt, and N. Theodore. 2018. Landscape-scale regulators of water transparency in mountain lakes: implications of projected glacial loss. *Canadian Journal of Fisheries and Aquatic Science* 75:1169–1176.
- Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118–125.
- Omernik, J. M., and G. E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environmental Management* 54:1249–1266.
- Pace, M. L., and J. J. Cole. 2002. Synchronous variation of dissolved organic carbon and color in lakes. *Limnology & Oceanography* 47:333–342.
- Reynolds, C. S. 1984. *The ecology of freshwater phytoplankton*. Cambridge University Press, Cambridge, UK.
- Roesler, C., M. Roberts, and C. Vang. 2018. Page 96. Saint Louis River Estuary clay-influenced bay assessment 2017. Technical Report. Wisconsin Department of Natural Resources, Madison, Wisconsin, USA.
- Shannon, E. E., and P. L. Brezonik. 1972. Limnological characteristics of north and central Florida lakes. *Limnology & Oceanography* 17:97–110.
- Smeltzer, E., and S. A. Heiskary. 1990. Analysis and applications of lake user survey data. *Lake & Reservoir Management* 6:109–118.
- Thrane, J.-E., D. O. Hessen, and T. Andersen. 2014. The absorption of light in lakes: negative impact of dissolved organic carbon on primary productivity. *Ecosystems* 17:1040–1052.
- U.S. EPA. 2000. Nutrient criteria technical guidance manual: Lakes and reservoirs. EPA-822-B00-001. United States Environmental Protection Agency, Washington, D.C., USA.
- Webster, K. E., P. A. Soranno, K. S. Cheruvilil, M. T. Bremigan, J. A. Downing, P. D. Vaux, T. R. Asplund, L. C. Bacon, and J. Connor. 2008. An empirical evaluation of the nutrient-color paradigm for lakes. *Limnology & Oceanography* 53:1137–1148.
- Williamson, C. E., D. P. Morris, M. L. Pace, and O. G. Olson. 1999. Dissolved organic carbon and nutrients as regulators of lake ecosystems: Resurrection of a more integrated paradigm. *Limnology & Oceanography* 44:795–803.
- Williamson, C. E., J. A. Brentrup, J. Zhang, W. H. Renwick, B. R. Hargreaves, L. B. Knoll, E. P. Overholt, and K. C. Rose. 2014. Lakes as sensors in the landscape: Optical metrics as scalable sentinel responses to climate change. *Limnology & Oceanography* 59:840–850.
- Williamson, C. E., E. P. Overholt, R. M. Pilla, T. H. Leach, J. A. Brentrup, L. B. Knoll, E. M. Mette, and R. E. Moeller. 2015. Ecological consequences of long-term browning in lakes. *Scientific Reports* 5:18666.
- Yuan, L. L., and A. I. Pollard. 2014. Classifying lakes to improve precision of nutrient-chlorophyll relationships. *Freshwater Science* 33:1184–1194.

SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.1871/full>

DATA AVAILABILITY

The data upon which the analyses reported are based are available from the Data Repository for U of MN (DRUM) at: <https://doi.org/10.13020/01wt-jg66>.

RESEARCH ARTICLE

Iron influence on dissolved color in lakes of the Upper Great Lakes States

Patrick L. Brezonik^{1*}, Jacques C. Finlay², Claire G. Griffin², William A. Arnold¹, Evelyn H. Boardman^{2#a}, Noah Germolus^{1#b}, Raymond M. Hozalski¹, Leif G. Olmanson³

1 Department of Civil, Environmental, and Geo- Engineering, University of Minnesota, Minneapolis, MN, United States of America, **2** Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN, United States of America, **3** Remote Sensing Laboratory, Department of Forest Resources, University of Minnesota, St. Paul, MN, United States of America

^{#a} Current address: Fitzgerald Environmental Associates, LLC, Colchester, VT, United States of America.

^{#b} Current address: Parsons Laboratory, MIT, Cambridge, MA, United States of America.

* brezonik@umn.edu



OPEN ACCESS

Citation: Brezonik PL, Finlay JC, Griffin CG, Arnold WA, Boardman EH, Germolus N, et al. (2019) Iron influence on dissolved color in lakes of the Upper Great Lakes States. PLoS ONE 14(2): e0211979. <https://doi.org/10.1371/journal.pone.0211979>

Editor: Lee W. Cooper, University of Maryland Center for Environmental Science, UNITED STATES

Received: November 5, 2018

Accepted: January 24, 2019

Published: February 13, 2019

Copyright: © 2019 Brezonik et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data uploaded to the DRUM public repository are all the data that others would need to replicate all our results. The DOI is: <https://doi.org/10.13020/kk4m-zx88>.

Funding: PLB received a grant from the University of Minnesota Office of the Vice President for Research and the University of Minnesota's Faculty Retirees Association, no grant number provided (<https://umra.umn.edu>). RMH, PLB, and JCF received a grant from the National Science Foundation, CBET 1510332 (www.nsf.gov). JCF, PLB, LGO, and RMH received a grant from the

Abstract

Colored dissolved organic matter (CDOM), a major component of the dissolved organic carbon (DOC) pool in many lakes, is an important controlling factor in lake ecosystem functioning. Absorption coefficients at 440 nm (a_{440} , m^{-1}), a common measure of CDOM, exhibited strong associations with dissolved iron (Fe_{diss}) and DOC in 280 lakes of the Upper Great Lakes States (UGLS: Minnesota, Wisconsin, and Michigan), as has been found in Scandinavia and elsewhere. Linear regressions between the three variables on UGLS lake data typically yielded R^2 values of 0.6–0.9, suggesting that some underlying common processes influence organic matter and Fe_{diss} . Statistical and experimental evidence, however, supports only a minor role for iron contributions to a_{440} in UGLS lakes. Although both DOC and Fe_{diss} were significant variables in linear and log-log regressions on a_{440} , DOC was the stronger predictor; adding Fe_{diss} to the linear a_{440} -DOC model improved the R^2 only from 0.90 to 0.93. Furthermore, experimental additions of Fe^{III} to colored lake waters had only small effects on a_{440} (average increase of $0.242 m^{-1}$ per 100 $\mu g/L$ of added Fe^{III}). For 136 visibly stained waters (with $a_{440} > 3.0 m^{-1}$), where allochthonous DOM predominates, DOM accounted for $92.3 \pm 5.0\%$ of the measured a_{440} values, and Fe_{diss} accounted for the remainder. In 75% of the lakes, Fe_{diss} accounted for $< 10\%$ of a_{440} , but contributions of 15–30% were observed for 7 river-influenced lakes. Contributions of Fe_{diss} in UGLS lakes to specific UV absorbance at 254 nm ($SUVA_{254}$) generally were also low. Although Fe_{diss} accounted for 5–10% of measured $SUVA_{254}$ in a few samples, on average, 98.1% of the $SUVA_{254}$ signal was attributable to DOM and only 1.9% to Fe_{diss} . DOC predictions from measured a_{440} were nearly identical to those from a_{440} corrected to remove Fe_{diss} contributions. Overall, variations in Fe_{diss} in most UGLS lakes have very small effects on CDOM optical properties, such as a_{440} and $SUVA_{254}$, and negligible effects on the accuracy of DOC estimated from a_{440} , data for which can be obtained at broad regional scales by remote sensing methods.

Minnesota Environmental and Natural Resources Trust fund, as recommended by the Legislative-Citizen Commission on Minnesota Resources, no grant number provided (<https://www.lccmr.leg.mn/>). JCF received a grant from the Minnesota Sea Grant, no grant number provided (www.seagrant.umn.edu). LGO also received salary support from the University of Minnesota's U-Spatial Program and Agricultural Experiment Station. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Research associating iron (Fe) concentrations and organic color (now called colored dissolved organic matter, or CDOM) in surface waters extends back to studies in Finland [1] and Sweden [2], but its nature and significance were poorly understood for many decades. CDOM plays a major role in the ecological functioning of lakes by affecting light penetration, temperature structure, metal bioavailability, and photochemical processes. Several recent studies, e.g., [3,4], have implicated Fe as a factor in the long-term increases observed in CDOM across Scandinavia [5, 6] and some other temperate regions—the so-called “browning” phenomenon [7]. Increasing total Fe (Fe_T : dissolved plus particulate Fe) in 27 of 30 Swedish rivers was estimated to account for an average of 25% of the variations in CDOM and up to 74% in northern Sweden [2]. Ekström et al. [8] proposed that long-term CDOM trends in Swedish rivers could be related to increasing Fe mobilization driven by increasing temperature and river discharge that increase the probability of anoxic conditions conducive to Fe solubilization.

Whether the Fe-CDOM relationship is actually causative or merely correlative may affect the use of CDOM, which can be retrieved on regional scales from satellite imagery, e.g., [9,10], to estimate concentrations of DOC, a major component in the aquatic carbon cycle. If Fe affects absorption coefficients (a_λ) at the wavelength (λ) used to quantify CDOM, variations in dissolved Fe or in the fraction of a_λ caused by Fe could affect the accuracy of DOC estimated from a_λ . Here we address this issue for lakes in the U.S. Upper Great Lakes States (UGLS).

Most recent studies on the influence of Fe on CDOM have focused on Swedish lakes. Based on observations from multi-basin Lake Mälaren, Köhler et al. [11] found decreasing dissolved Fe (Fe_{diss}) as water flowed through the basins, with concurrent declines in CDOM and a shift from colored allochthonous material to less colored autochthonous DOM. Weyhenmeyer et al. [4] found a linear relationship between dissolved organic carbon (DOC) and CDOM (measured as absorption coefficients at 420 nm, a_{420}) in a large dataset from Sweden and Canada, but the carbon-specific a_{420} (a_{420}/DOC) increased nonlinearly, approaching an asymptotic value, with increasing Fe_T , which the authors considered to be all Fe_{diss} . Based on these findings, the authors inferred that Fe_{diss} affected apparent CDOM levels (i.e., absorption coefficients at 420 nm, a_{420}) and concluded that Fe_{diss} , pH, water residence time, and colored DOC all may be important factors for regional changes in lake browning. Alternative explanations for the browning phenomenon, including climate change [12] and recovery from acidification by atmospheric acid deposition, e.g., [5,13], are not necessarily inconsistent with a role for Fe.

Effects of Fe on UV absorbance are well studied, but effects in the visible range are less well known. Weishaar et al. [14] found that absorbance at 254 nm (A_{254}) increased with Fe^{III} at the same rate in solutions with or without DOM. Poulin et al. [15] found A_{254} increased linearly with Fe^{III} in DOM-containing solutions but found no effect for added Fe^{II} . They concluded that Fe^{III} should be accounted for in measurements of specific UV absorbance at 254 nm ($SUVA_{254}$; i.e., A_{254} normalized by DOC) and provided an equation to make such corrections. Maloney et al. [16] found a nonlinear increase in carbon-specific absorptivity, a_{320}/DOC in the Fe_{diss} range of 1–4 mg/L in a humic-rich lake and reported that the spectral slope in the range 280–400 nm decreased as Fe_{diss} increased from 0.0 to 0.5 mg/L. They hypothesized that Fe_{diss} likely would affect light conditions in the visible range but made no measurements in this region. Kritzberg and Ekström [3] and Xiao et al. [17] reported that adding Fe^{III} to CDOM-containing waters linearly increased absorptivity at 410–420 nm. Adding Fe^{III} (1600–3600 $\mu\text{g/L}$) to humic and fulvic acid reference materials also decreased spectral slopes in the UV range [17].

We have been studying characteristics of CDOM in UGLS lakes and mapping its distribution by field studies and satellite imagery [10,18,19]. The occurrence of major iron ore deposits

in Minnesota led us to question whether Fe contributes to observed CDOM levels (measured as a_{440}) and/or affects its other optical properties, and whether that could affect DOC values inferred from a_{440} . This study had three primary objectives: (1) quantify the association between Fe and a_{440} (our measure of CDOM) in surface waters of three UGLS ecoregions; (2) experimentally determine whether the association is causal, and if so evaluate the extent of Fe contributions to a_{440} and other optical properties that characterize CDOM; and (3) quantify the influence of Fe variability on DOC estimated from a_{440} in natural waters.

Methods

Sampling sites

We collected 450 samples from 280 water bodies (mostly lakes) in northern and central Minnesota, Wisconsin, and Michigan over the period 2014–2018 (Fig 1). Sampling occurred during summer (June–September). Nearly all 2014 and 2015 samples were from two lake-rich ecoregions of northeastern and east-central Minnesota: Northern Lakes and Forest (NLF) and North-Central Hardwood Forest (NCHF), which together contain ~ 9800 of the state’s ~12,000 lakes. The NLF is ~ 50% forested, and nearly a third of its area is wetlands or lakes. Agriculture and urban land cover constitute only small portions of this ecoregion (7 and 4%, respectively). In contrast, the NCHF is ~ 48% agricultural land and 9% urban. Forest cover constitutes 25% and wetlands ~ 10% of the NCHF. In 2016, sampling was extended to NLF

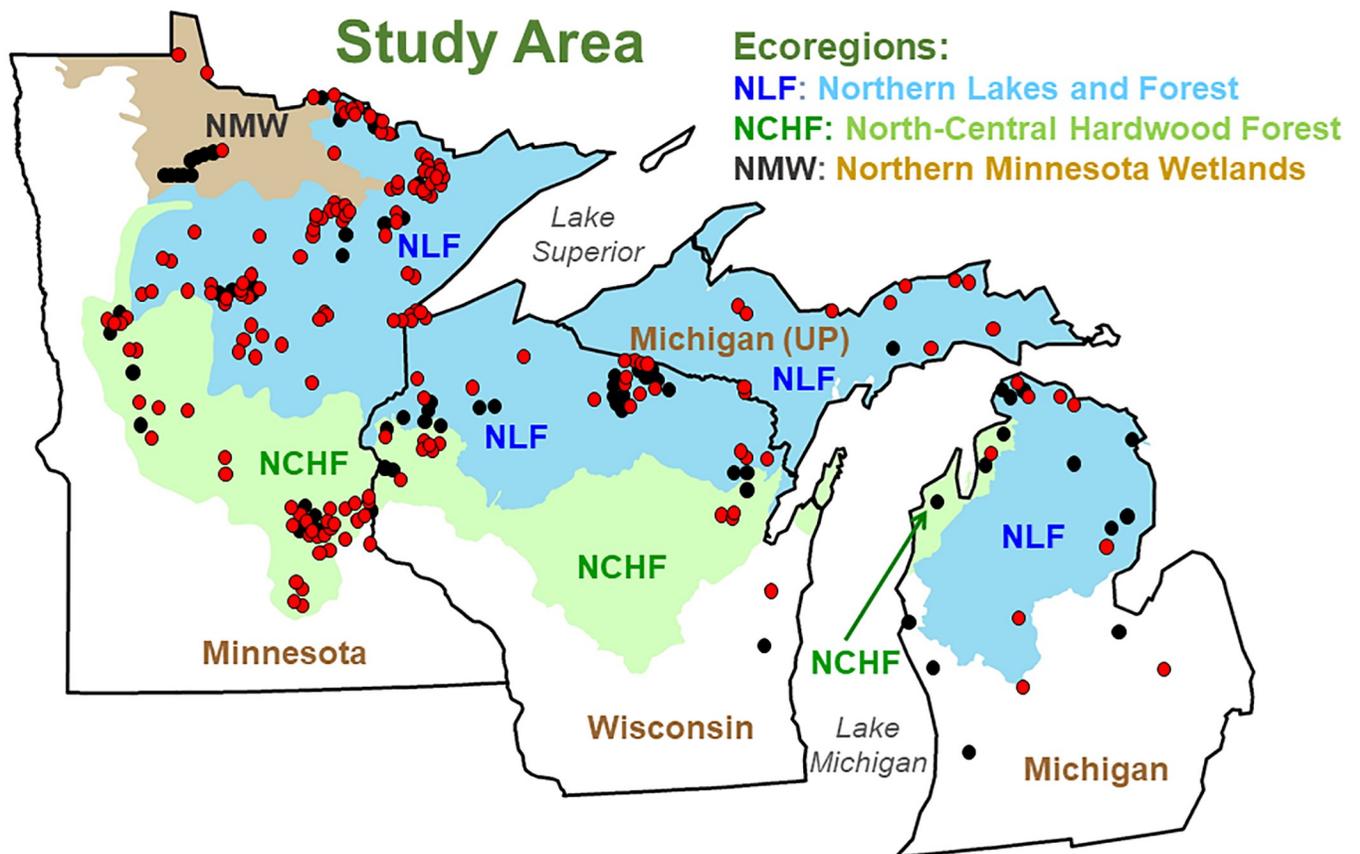


Fig 1. Map of study area showing ecoregions and sampling sites. Red circles: sites with Fe_{diss} , a_{440} and DOC data; black circles: additional sites with only a_{440} and DOC data.

<https://doi.org/10.1371/journal.pone.0211979.g001>

and NCHF areas in Wisconsin and Michigan and the Northern Minnesota Wetlands (NMW) ecoregion, which has < 200 lakes, only a few of which are road-accessible.

Water usually was collected by small boat or kayak in the open water area, but collections were made from the ends of docks on small lakes where boat access was not feasible. Sites were selected to include a diversity of lake types, CDOM levels, and catchment land cover. The vast majority of sites were lakes, but six large rivers and six impoundments on large rivers were included. NLF lakes were in forested catchments (mixed conifers and hardwoods with substantial wetlands) with little to no human development; supplemental sampling in 2018 focused on the NLF ecoregion, where the vast majority of CDOM-rich lakes occur in the study area. Some NCHF lakes were in minimally developed catchments, but most were in urban to exurban areas and a few had catchments with row-crop agriculture.

Sampling/Field procedures

Water samples were collected from ~ 0.25 m depth using acid-washed, triple-rinsed polycarbonate or high-density polyethylene bottles and stored on ice until processed, usually the same evening. Secchi depth (SD) was measured by standard limnological procedures. Samples were collected at various depths on a few NLF lakes in 2018 to determine effects of stratification on CDOM and Fe_{diss}. Raw water was filtered through 0.45 μm Geotech trace-metal-certified capsule filters or pre-combusted (4 h at 450 °C) 0.7 μm Whatman glass fiber filters. Filtered water for DOC and Fe_{diss} analyses was acidified using 0.1 mL of 2 M HCl per 50 mL of sample and refrigerated (DOC) or frozen (Fe_{diss}) in pre-cleaned glass or plastic bottles, respectively. Unfiltered water for Fe_T analyses was acidified with 1 mL of concentrated HNO₃ per 50 mL sample and stored in the same manner as the Fe_{diss} samples. Un-acidified filtered water for CDOM analysis was refrigerated in 40 mL glass vials with no headspace. Filter blanks (DI water) showed no measurable DOC or CDOM. Chlorophyll-*a* (chl-*a*) was collected by vacuum filtration of water samples onto 0.22 μm cellulose nitrate filters that were then stored frozen until analysis.

Analytical methods

Absorbance was measured within a month of sample collection by scanning from 250 to 700 nm using a Shimadzu 1601UV-PC dual beam spectrophotometer with 1 cm or 5 cm quartz, depending on CDOM levels, and nanopure water in the reference cell. We tested whether length of storage affected *a*₄₄₀ measurements on filtered, refrigerated samples from three colored lakes with *a*₄₄₀ values of 5–30 m⁻¹ and found no detectable decreases in *a*₄₄₀ after one month of storage, which agrees with other studies [20]. Samples were allowed to warm to room temperature on the benchtop prior to measurements. Absorbance was converted to Napierian absorption coefficients using:

$$a_{\lambda} = 2.303A_{\lambda}/\ell \quad (1)$$

where: *a*_λ is the absorption coefficient (m⁻¹) and *A*_λ is absorbance, both at wavelength λ, and ℓ is cell path length (m). Absorbance was blank-corrected before conversion. CDOM is reported as absorption coefficient (m⁻¹) at 440 nm, *a*₄₄₀. SUVA₂₅₄ (L mg⁻¹ m⁻¹) was calculated by dividing absorbance at 254 nm by DOC concentration (mg/L), after correcting for cell path length. Contributions of Fe_{diss} to SUVA₂₅₄ were calculated using the equation of Poulin et al. [15]; subtraction of the Fe_{diss} contribution yielded DOM-based values, SUVA_{254,DOM}. Spectral slopes (*S*_{λ2-λ1}) were calculated from absorbance data for three wavelength regions (275–295, 350–400, and 400–460 nm) by taking the natural logarithm (ln) of *A* and computing slopes in

Excel or by nonlinear fit of absorptivity data to Eq (2):

$$a_{\lambda} = a_{\lambda_{ref}} \exp\{-S(\lambda - \lambda_{ref})\} \quad (2)$$

where λ_{ref} is a reference wavelength and S is the slope.

DOC was measured on a Shimadzu TOC L_CSN analyzer. Chl-*a* was measured by fluorometry after 90% acetone extraction of the chl-*a* filters. Fe_{diss} and Fe_T were analyzed in triplicate with 200 µg/L of yttrium added as an internal standard on a Thermo Scientific iCAP 6500 DUO ICP-OES or iCAP 7600 DUO ICP-OES instrument. Fe_{diss} was not analyzed on some low-CDOM waters sampled in 2016 that, based on 2014–2015 results, were expected to have low Fe_{diss} nor on some high-CDOM samples from lakes sampled multiple times in 2016. Based on analysis of the 2014–2016 data, we collected additional samples in 2018 from some rivers and lakes fed by rivers to measure Fe_T and Fe_{diss} and calculated particulate Fe (Fe_{part}) by difference.

Fe addition experiment. The effect of adding Fe^{III} on a_{440} was measured for surface water samples from six northern Minnesota lakes with a range of ambient a_{440} and Fe_{diss}. The lake waters were circumneutral (pH 6.0–8.0). We used Fe^{III} because Poulin et al. [15] found no effect of Fe^{II} on UV absorbance of CDOM-containing solutions. Fe^{III} is the thermodynamically stable form in oxic water at circumneutral pH, which suggests that Fe^{III}-humic complexes predominate in surface waters. A 500 mL aliquot of filtered lake water (0.7 µm glass fiber filters) was placed in a 1.0 L beaker on a magnetic stirrer, and five 0.6 mL increments of a solution containing 77.1 mg/L of Fe^{III} were added sequentially. The additions were designed to yield measurable increases in Fe^{III} (total of 460 µg/L over the five increments) but not over-saturate the DOM. The ratio Fe^{III}/DOC was < 1 µmol/mg for the highest additions, lower than reported iron-binding capacities for humic materials, e.g., [21,22]. We also added similar amounts of Fe^{III} to deionized water to determine whether a_{440} increased from uncomplexed Fe^{III} and to 0.01 M EDTA to determine whether a_{440} increased when Fe^{III} was added to a colorless chelating agent. The Fe^{III} solution was prepared in 0.1 M HNO₃ from reagent-grade Fe₂(SO₄)₃·nH₂O, and the resulting Fe^{III} concentration was determined by triplicate ICP-OES analysis. After each Fe^{III} increment, sample pH was adjusted to within 0.1 of its ambient value by dropwise addition of 1 M NaOH. A preliminary experiment showed that the acidic Fe^{III} solution decreased the pH enough to affect the measured a_{440} . The effect of pH on CDOM absorbance is well known [23]. After pH stabilization, 5 mL aliquots were stored in the dark at 4 °C with no further filtration until absorbance was measured, ~ 24 h later.

Data analysis. All observations (site-date combinations) were treated as separate data points; i.e., multiple samples from a lake across or within years were not averaged. Statistical analyses were done in JMP Pro 13.1 except for some simple regressions done in Excel 2016. Initial data inspection showed that distributions for a_{440} , DOC, and Fe_{diss} were skewed to low values but otherwise well distributed over the range of observed values (S1 Fig). Natural log transforms yielded more Gaussian-looking distributions but still did not satisfy the Shapiro-Wilks test for normality. Unless stated otherwise, statistical results are reported for untransformed data. In addition to simple and multiple regression analyses on subsets of the untransformed and log-transformed data, we analyzed relationships between a_{440} and “de-trended” values of DOC and Fe_{diss}. The de-trended DOC analysis regressed the residuals from a regression of DOC vs. Fe_{diss} (i.e., the variance in DOC not explained by Fe_{diss}) against a_{440} . The de-trended Fe_{diss} analysis similarly used the residuals from a regression of Fe_{diss} vs. DOC (i.e., the variance in Fe_{diss} not explained by DOC) in a regression vs. a_{440} .

Results and discussion

Overview of water quality variables in study lakes

Broad ranges of a_{440} , DOC, and Fe_{diss} and two basic limnological variables, SD and chl-*a*, were measured in the study, and large differences were found between the two major ecoregions (NLF and NCHF; Table 1). Median, mean and maximum values of a_{440} , Fe_{diss} , and DOC were substantially higher for NLF lakes (dominated by forests) than NCHF lakes (dominated by agriculture). The median chl-*a* in NLF lakes was 4.1 $\mu\text{g/L}$ (range 0–25 $\mu\text{g/L}$), and the median in NCHF lakes was 7.6 $\mu\text{g/L}$ (range 1–98 $\mu\text{g/L}$). Lakes with obvious color (defined here as $a_{440} > 3.0 \text{ m}^{-1}$) also had low chl-*a*, nearly all $< 20 \mu\text{g/L}$ [19]. The SD range was more limited (0.4–5.5 m) in NCHF lakes than NLF lakes (0.3–19.5 m), where high SD values were associated with deep, ultra-oligotrophic mine pit lakes. NLF lakes with $a_{440} > 3 \text{ m}^{-1}$ generally had SD $< 3 \text{ m}$, and CDOM levels were the controlling factor for SD in highly colored lakes [24], some of which had SD values as low as 0.3 m. Higher mean than median values for the five variables (especially for a_{440} and Fe_{diss}) in both ecoregions are indicative of non-normal (skewed) distributions.

A principal components analysis to examine relationships among the above five variables showed that 90% of the variance was explained by the first two principal components (PCs) (Fig 2). DOC, a_{440} , and Fe_{diss} were clustered together with high positive loadings on PC1, which accounted for 67.2% of the variance. SD also had a high PC1 loading but in a negative direction. PC2, which accounted for 22.5% of the variance, was driven by a high negative loading of chl-*a* and smaller positive loadings of SD and Fe_{diss} . Overall, the results support the idea that a_{440} , Fe_{diss} , and DOC behave similarly as variables but behave differently from chl-*a* and SD.

Our 2014–2016 measurements were on near-surface samples because most CDOM effects of interest are near-surface phenomena. Other recent studies on Fe-CDOM interactions, e.g.,

Table 1. Summary statistics for a_{440} , DOC, and Fe_{diss} and two basic limnological variables in the NLF and NCHF ecoregions.

	$a_{440} \text{ m}^{-1}$	DOC mg/L	$Fe_{diss} \mu\text{g/L}$	Chl- <i>a</i> $\mu\text{g/L}$	SD m
NLF					
Mean	6.03	12.2	247	4.74	2.3
Standard deviation	7.20	7.8	342	3.54	2.1
Median	2.76	9.2	118	4.06	1.7
Standard error of mean	0.40	0.4	23	0.25	0.1
Skewness	1.6	1.1	2.1	1.9	3.8
Minimum	0	2.5	1	0.02	0.3
Maximum	32.47	36.1	1858	24.9	19.5
Interquartile range	8.33	9.6	346	4.2	1.9
Number of samples	317	313	212	195	234
NCHF					
Mean	1.41	8.0	37	16.64	1.7
Standard deviation	1.03	2.7	64	21.49	1.2
Median	1.15	7.7	14	7.56	1.4
Standard error of mean	0.10	0.3	8	2.46	0.2
Skewness	1.40	0.8	3.5	1.90	1.5
Minimum	0.10	3.1	1	1.22	0.4
Maximum	5.30	17.8	391	98.70	5.5
Interquartile range	1.38	3.4	36	15.0	1.5
Number of samples	104	105	64	76	53

<https://doi.org/10.1371/journal.pone.0211979.t001>

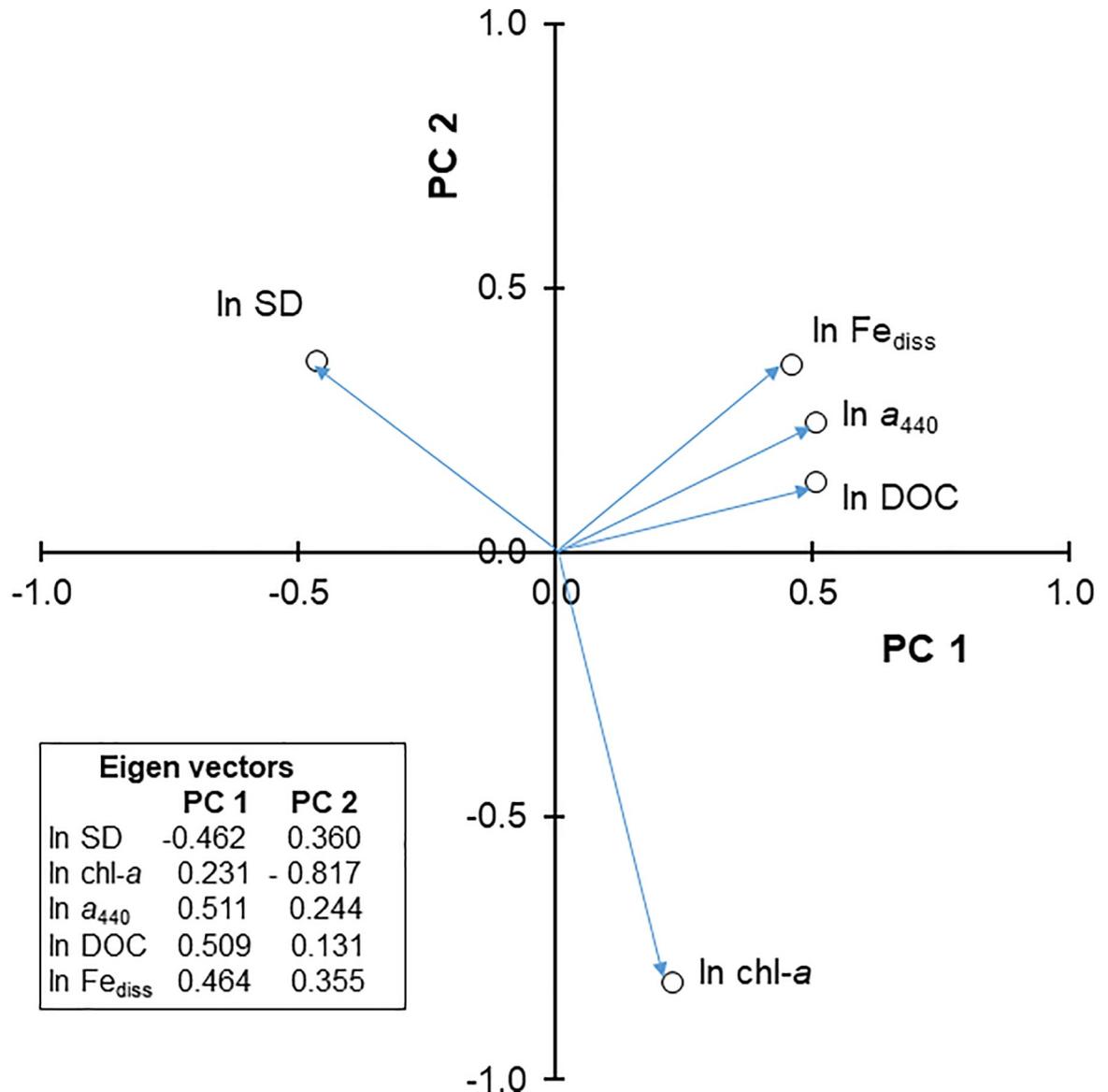


Fig 2. Plot of eigenvalues for the first two principal components of a principal components analysis of five variables (a_{440} , DOC, Fe_{diss} , chl-a, and SD) for the whole data set.

<https://doi.org/10.1371/journal.pone.0211979.g002>

[3,4], also focused on surface water samples. Data from summer 2018 for lakes with a range of near-surface a_{440} indicate that a_{440} and Fe_{diss} may vary with depth, with a_{440} decreases and Fe_{diss} increases in near-bottom waters of highly colored lakes (SI Table). These trends could be caused by seasonal variations in CDOM and in-lake cycling processes for Fe and CDOM, a topic beyond the scope of this paper.

All 2014–2016 samples analyzed for Fe were filtered (0.7 μm filters) prior to analysis, and the results are defined operationally as Fe_{diss} , which comprises Fe in true solution, including that complexed by DOM, and colloidal Fe associated with macromolecular DOM and hydrous Fe oxide particles too small to be retained on filters. We excluded particulate Fe (Fe_{part}) associated with filterable particles from analysis because we considered it inappropriate to include Fe in plankton or mineral particles. Three lines of evidence support this decision.

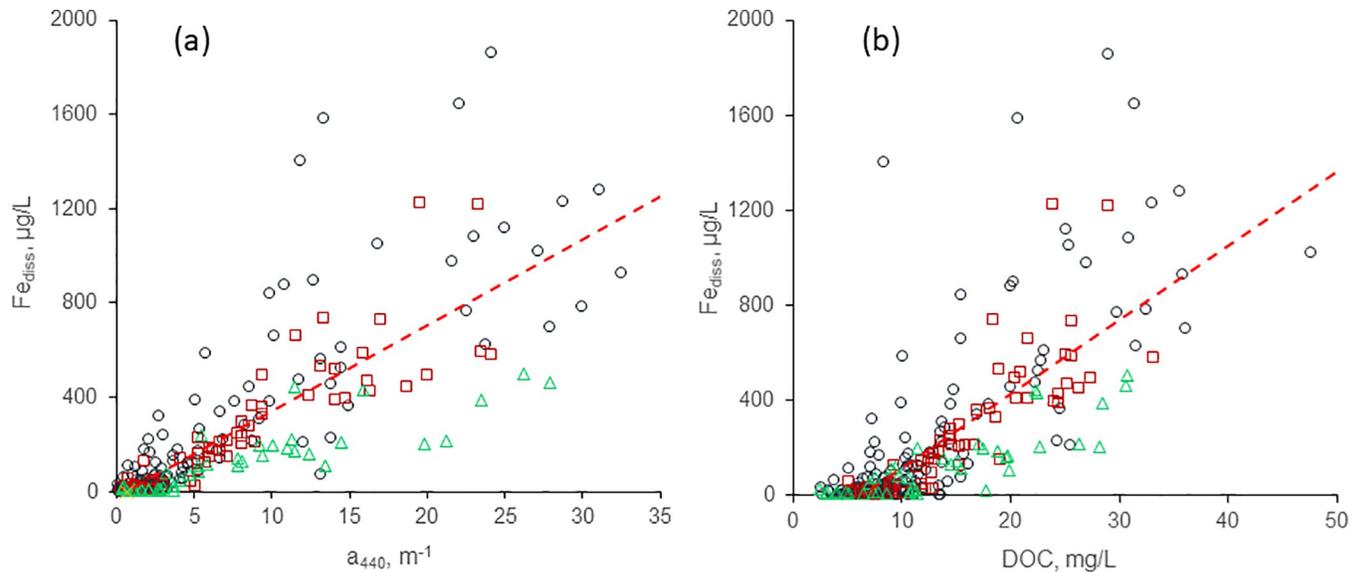


Fig 3. (a) a_{440} and (b) DOC vs. Fe_{diss} for sampling years 2104–2016; triangles = 2014, squares = 2015, circles = 2016; dashed lines are regression fits for all years (see Table 1 for regression statistics).

<https://doi.org/10.1371/journal.pone.0211979.g003>

First, analyses of Fe_T and Fe_{diss} on samples collected in 2018 from moderate- to high-CDOM lakes showed that Fe_{part} was only a small fraction of Fe_T (average of 10.3%, range 0–23%; S2 Table), and most of Fe_T was Fe_{diss} . These samples were from lakes with moderate-to-high Fe_{diss} concentrations, and most of the lakes had river inflows. Samples from the associated rivers had a higher fraction of Fe_{part} —average of ~27%, or ~20% when one sample from a high runoff event was excluded (S2 Table). These results agree with the findings of Weyhenmeyer et al. [4], who reported that Fe_{part} was not an important component of Fe_T in Swedish lakes. Kritzberg and Ekström [3] found that Fe_{part} was an important fraction of Fe_T in Swedish rivers. Our more limited sampling found that Fe_{part} was more important in rivers than lakes, but Fe_{diss} was still dominant in rivers.

Second, concentrations of total suspended matter (TSM) were generally < 10 mg/L in lakes with $a_{440} > 3.0$ m^{-1} ; the average TSM for 53 lakes sampled in 2016 with $a_{440} > 3.0$ m^{-1} was only 3.8 mg/L. Third, CDOM-rich UGLS lakes occur in highly vegetated, forested catchments, where soil erosion is low, similar in terrain and ecological conditions to the Swedish and Canadian lakes where Fe_{part} was found not to be important [4].

Fe_{diss} is linearly correlated a_{440} and DOC

Strong correlations were found between Fe_{diss} and a_{440} , our measure of CDOM, as well as for Fe_{diss} and DOC, for each year and for the complete data set (Fig 3, Table 2). Similar correlations were obtained for log-transformed data (S3 Table). Values of Fe_{diss} and a_{440} generally were higher and more scattered in 2016 than in the two previous years, probably for two reasons. First, unusually high precipitation across Minnesota in 2016 broke many daily and monthly records at individual locations, likely resulted in higher export of Fe and DOM from catchments to lakes, and thus led to higher concentrations. Second, we sampled three times as many sites in 2016 than in 2014 or 2015. These sites covered a larger geographic range and had a greater proportion of catchments in agricultural, urban, or mixed-use landscapes, resulting in a greater diversity of geochemical conditions among sites than for previous years, thus accounting for the greater scatter. Because of the inter-annual differences, R^2 for the total data

Table 2. Dissolved Fe-CDOM and Fe-DOC relationships.

Year	N	Regression equation ^a	R ²	RMSE ^b	Slope SE ^c
Fe_{diss} (µg/L) vs. a₄₄₀ (m⁻¹)					
2014	46	Fe _{diss} = 17.0×a ₄₄₀ - 5.8	0.77	66	1.4
2015	61	Fe _{diss} = 37.1×a ₄₄₀ - 34.3	0.77	131	2.6
2016	175	Fe _{diss} = 43.1×a ₄₄₀ - 21.9	0.73	184	2.0
All	282	Fe _{diss} = 36.4×a ₄₄₀ - 23.3	0.67	182	1.5
Fe_{diss} (µg/L) vs. DOC (mg/L)					
2014	42	Fe _{diss} = 14.7×DOC - 76	0.71	76	1.5
2015	61	Fe _{diss} = 32.8×DOC - 241	0.68	154	2.9
2016	177	Fe _{diss} = 36.2×DOC - 218	0.62	218	2.1
All	280	Fe _{diss} = 31.1×DOC - 192	0.58	204	1.6

^a All equations and coefficients significant at p < 0.0001.

^b Root mean square error

^c Standard error of slope.

<https://doi.org/10.1371/journal.pone.0211979.t002>

set for Fe_{diss} vs. a₄₄₀ (0.67) was lower than for the individual years. Overall, however, the results are consistent with century-old [1] and more recent studies [4] associating CDOM and Fe concentrations in lakes.

DOC is a stronger a₄₄₀ predictor than Fe_{diss}

Although many catchment and water quality conditions affect lake CDOM levels, we are most interested here in the relative effects of DOC and Fe_{diss} on a₄₄₀ because a₄₄₀ is used to quantify CDOM and often used to predict DOC, e.g., [19]. As shown below, both DOC and Fe_{diss} are strong predictors of a₄₄₀, but DOC is stronger. We performed simple and multiple regression analyses with a₄₄₀ as predicted variable and DOC and Fe_{diss} as predictor variables (Table 3). Regressions were performed using the entire a₄₄₀ range and just for sites with a₄₄₀ > 3.0 m⁻¹ because related work [19] showed a break in the DOC-a₄₄₀ relationship around a₄₄₀ = 3.0 m⁻¹. A tight fit between the two variables was found above this value, but much more scatter and a higher slope were found below. Griffin et al. [19] interpreted this finding to indicate that low-color DOM from autochthonous and anthropogenic sources was an important, but variable DOC contributor in waters with a₄₄₀ < 3.0 m⁻¹, and these sources were less important in high-CDOM waters dominated by allochthonous (humic-like) DOM.

Table 3. Simple and multiple regression relationships for a₄₄₀ vs. DOC and Fe_{diss}.

Data range	Best fit equation ^a	N	R ²	RMSE ^b	SE ^c
All data	a ₄₄₀ = 0.868×DOC - 4.89	434	0.90	2.15	0.015
	a ₄₄₀ = 0.0183×Fe _{diss} + 2.47	283	0.67	4.08	0.0008
	a ₄₄₀ = 0.746×DOC + 0.0046×Fe _{diss} - 4.15	277	0.93	1.89	0.023, 0.00056
a ₄₄₀ > 3.0 m ⁻¹	a ₄₄₀ = 0.967×DOC - 6.33	159	0.90	2.24	0.025
	a ₄₄₀ = 0.0141×Fe _{diss} + 5.77	136	0.52	5.10	0.0012
	a ₄₄₀ = 0.842×DOC + 0.0032×Fe _{diss} - 5.29	134	0.91	2.20	0.035, 0.00068

^a Units for variables: a₄₄₀ in m⁻¹; DOC in mg/L; Fe_{diss} in µg/L. All equations and coefficients significant at p < 0.0001.

^b Root mean square error.

^c SE = standard error for independent variable terms (slopes for simple regressions).

<https://doi.org/10.1371/journal.pone.0211979.t003>

DOC exhibited stronger relationships with a_{440} than did Fe_{diss} , but both variables were significant in multiple regressions (Table 3). Addition of Fe_{diss} as a second variable increased R^2 by only 0.03 for the entire a_{440} data range and 0.01 for $a_{440} > 3.0 \text{ m}^{-1}$. Similar results were found using log-transformed data (S4 Table), except that R^2 increased more when adding Fe_{diss} as a second variable (0.09 for all data, 0.03 for $a_{440} > 3.0 \text{ m}^{-1}$). De-trending to remove the influence of Fe_{diss} on the a_{440} -DOC relationship yielded an R^2 of 0.32. Removal of the influence of DOC on the a_{440} - Fe_{diss} relationship yielded an even lower R^2 of 0.12.

Weyhenmeyer et al. [14] similarly found that a least squares model using \ln DOC and $\ln \text{Fe}_{\text{diss}}$ explained 86% of the variance in $\ln a_{420}$. Linear de-trending of their data showed that DOC explained 38% of the variance when the Fe signal had been removed, and Fe explained 25% of the variance when the DOC signal was removed. Comparable de-trended values for \ln - \ln relationships of our data are 25% for DOC with the Fe_{diss} signal removed and 12% for Fe_{diss} when the DOC signal was removed. Although numerical values of Weyhenmeyer et al.'s original and de-trended R^2 results differ from ours, the overall outcomes of the analyses are similar: de-trending caused a large decrease in fit for a_{λ} -DOC relationships and even a larger decrease for a_{λ} - Fe_{diss} relationships. Together, these findings indicate that DOC is the more important explanatory variable statistically, but Fe_{diss} does explain some variance in a_{440} beyond that produced by the correlation between DOC and Fe_{diss} .

As noted above, DOC and Fe_{diss} , the two main chemical determinants of a_{440} , are themselves moderately correlated for the complete data set (Fig 3B) and within each year. In each case, R^2 for the Fe_{diss} -DOC relationship was lower than that for the corresponding Fe_{diss} - a_{440} relationship (Table 2), and 2016 values were more scattered than those for the previous years; R^2 for the total data set was only 0.58. Regression equations between Fe_{diss} and a_{440} (Table 2, Fig 3A) had x-intercepts of $a_{440} < 1 \text{ m}^{-1}$. In contrast, best-fit lines for linear regressions of Fe_{diss} vs. DOC had x-intercepts of 5–7 mg/L DOC (Table 2, Fig 3B). Together, these findings suggest that (i) Fe_{diss} is associated with the colored component of DOM and (ii) on average across all sites $\sim 6 \text{ mg/L}$ of DOC is not associated with Fe_{diss} . This likely represents low-color DOM with a low abundance of Fe-binding ligand groups, probably of autochthonous or anthropogenic origin. Photo-degradation of CDOM also could contribute to the low-color DOM pool, but it is uncertain whether CDOM photo-degradation reduces Fe binding capacity.

Weyhenmeyer et al. [14] found a curvilinear relationship ($R^2 = 0.49$) between the ratio a_{λ}/DOC and Fe_{diss} that might be interpreted as a measure of the effect of Fe_{diss} on the fraction of DOC that is colored. We found a similar relationship (Fig 4A) for our data; $R^2 = 0.64$ for a_{440}/DOC vs. $\ln \text{Fe}_{\text{diss}}$. As discussed above, however, the nature of DOM in low-CDOM waters ($a_{440} < 3.0 \text{ m}^{-1}$) likely differs from that in high-CDOM waters. The latter consists primarily of allochthonous, humic-like DOM; the former derives from various sources with generally lower color intensity and probably fewer binding sites for Fe_{diss} . Consequently, trends in a_{440}/DOC vs. Fe_{diss} may simply reflect changes in the nature of DOM as a_{440}/DOC increases. A plot of the relationship for sites dominated by allochthonous DOM (those with $a_{440} > 3.0 \text{ m}^{-1}$; Fig 4B) yielded an R^2 of only 0.46, and there was little trend in the ratio for $\text{Fe}_{\text{diss}} > 300 \text{ } \mu\text{g/L}$. Overall, the close fit between a_{440} and DOC for waters with $a_{440} > 3.0 \text{ m}^{-1}$ (Table 3) suggests that the DOM for these sites was dominated by humic-colored DOM.

Fe_{diss} had minor effects on other CDOM optical properties

The above results show that Fe_{diss} should be considered when evaluating a_{440} . Thus, it is worthwhile to assess whether other common optical measurements also are affected by Fe_{diss} . Results similar to those for a_{440}/DOC were obtained for SUVA_{254} , a more common DOC-

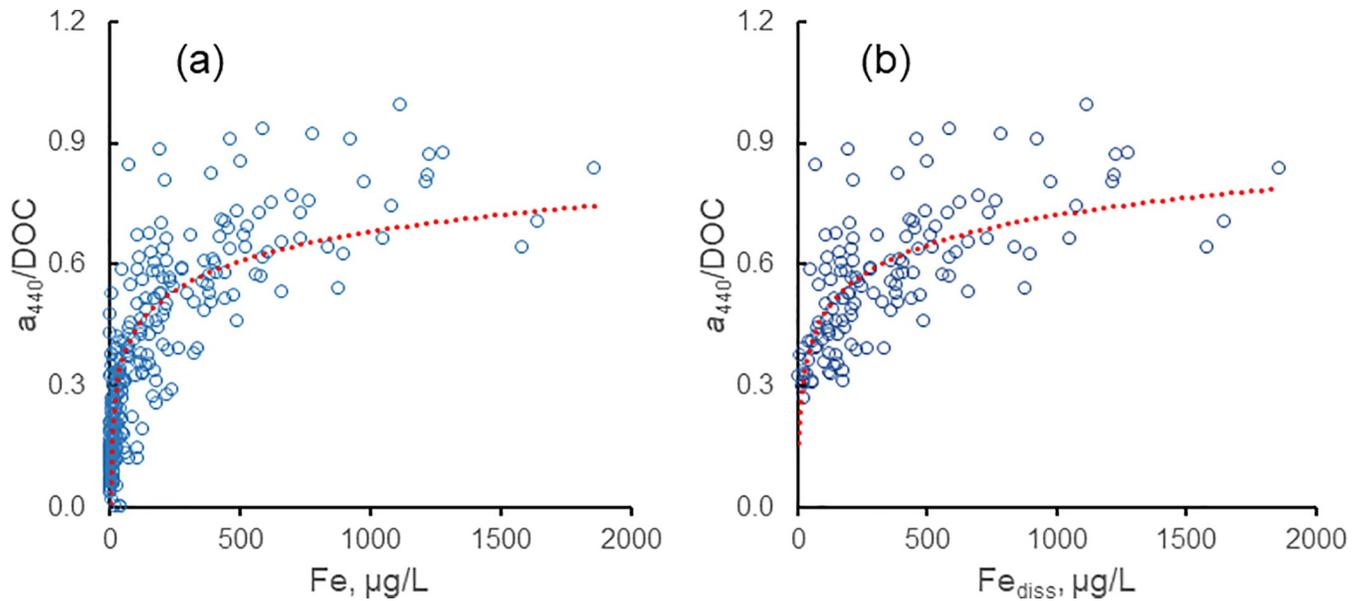


Fig 4. (a) a_{440}/DOC vs. Fe_{diss} for all data, $R^2 = 0.64$; (b) a_{440}/DOC vs. Fe_{diss} for sites with $a_{440} > 3.0 \text{ m}^{-1}$, $R^2 = 0.46$. (R^2 are for fit of a_{440}/DOC to $\ln(\text{Fe}_{\text{diss}})$).

<https://doi.org/10.1371/journal.pone.0211979.g004>

normalized optical measure. For the whole data set, a moderate fit was found for both uncorrected SUVA_{254} vs. $\ln \text{Fe}_{\text{diss}}$ ($R^2 = 0.67$) and for $\text{SUVA}_{254,\text{DOM}}$ vs. $\ln \text{Fe}_{\text{diss}}$ ($R^2 = 0.64$). For samples dominated by allochthonous DOM ($a_{440} > 3.0 \text{ m}^{-1}$), both relationships had lower R^2 (0.50 and 0.45, respectively, for uncorrected and Fe-corrected SUVA_{254}), with little trend above $\text{Fe}_{\text{diss}} = 300 \text{ µg/L}$ (Fig 5A).

Fe_{diss} contributions to SUVA_{254} , calculated according to [15], were small (mean = 1.9%, std. dev. = 1.9%, $n = 271$); on average, across all UGLS samples with Fe_{diss} and SUVA_{254} data,

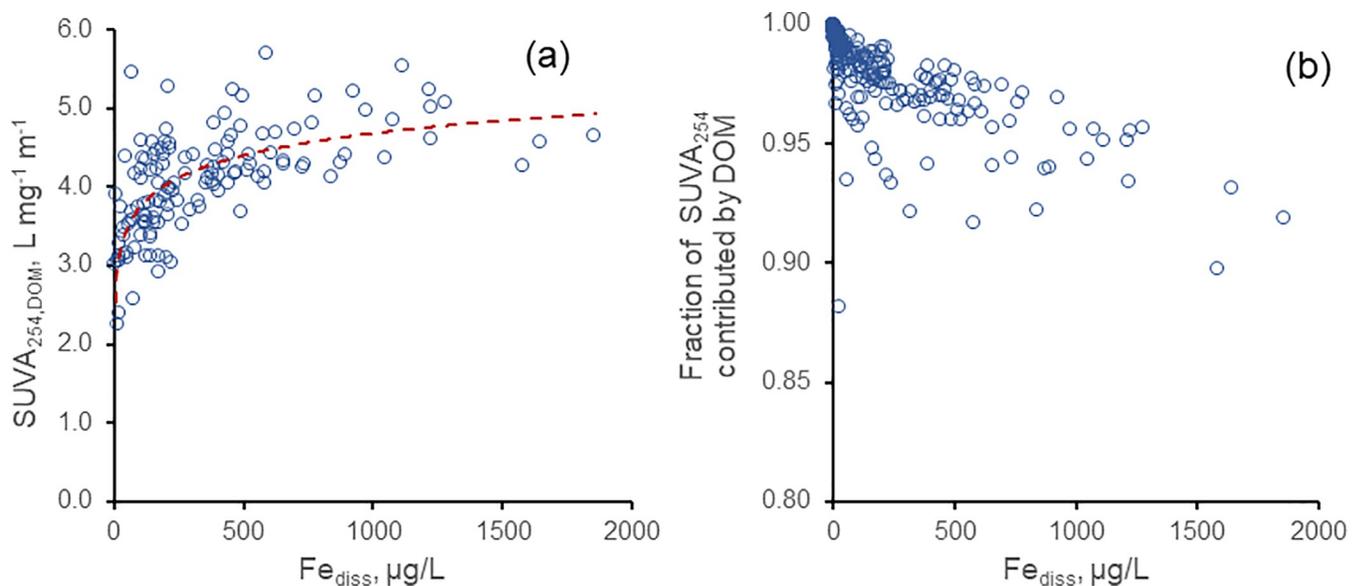


Fig 5. (a) $\text{SUVA}_{254,\text{DOM}}$ vs. Fe_{diss} for sites with $a_{440} > 3.0 \text{ m}^{-1}$, $R^2 = 0.45$ (R^2 is for fit to $\ln(\text{Fe}_{\text{diss}})$); (b) fraction of SUVA_{254} for all sites caused by DOM (i.e., corrected for Fe^{III} contribution) vs. Fe_{diss} .

<https://doi.org/10.1371/journal.pone.0211979.g005>

more than 98% of the SUVA₂₅₄ signal thus could be attributed to DOM. A small number of samples, however, had larger Fe_{diss} contributions to SUVA₂₅₄ (Fig 5B); 18 had Fe_{diss} contributions > 5%, and two had contributions > 10%. The lake with the largest contribution (11.9%), Crystal Lake (WI), is an ultra-clear oligotrophic seepage lake with low DOC (2.5 mg/L) and a SUVA₂₅₄ of only 0.64 L mg⁻¹ m⁻¹; it is an outlier relative to most lakes in the region. More relevant here are samples with higher Fe_{diss} and DOC. Seven samples with Fe_{diss} of 1000–1500 µg/L, had Fe_{diss} contributions to SUVA₂₅₄ of 4.4–6.7%, and Fe_{diss} contributions for three samples with Fe_{diss} > 1500 µg/L were 6.9–10.3%.

SUVA₂₅₄ values corrected for Fe_{diss} were slightly lower than uncorrected values, but of the 15 samples with original SUVA₂₅₄ > 5.0 L mg⁻¹ m⁻¹, 10 still had values > 5.0 after correction. The common upper limit for DOM-caused SUVA in natural waters is 5.0 L mg⁻¹ m⁻¹ [15]. Average SUVA₂₅₄ values for the 15 samples before and after correction (S5 Table) were 5.33 and 5.15 L mg⁻¹ m⁻¹, respectively. A third of these samples were from Johnson Lake, Minnesota (Itasca County), a small bog lake that generally had the highest CDOM and SUVA₂₅₄ levels in our studies. The average SUVA₂₅₄ before Fe-correction for Johnson Lake of 5.41 L mg⁻¹ m⁻¹ decreased to 5.23 L mg⁻¹ m⁻¹ after correction. Although high nitrate/nitrite concentrations (tens of mg/L range) may affect levels of SUVA₂₅₄ [25], concentrations of these ions were very low (few µg/L) in Johnson Lake and the other lakes we studied.

Spectral slopes, a measure of DOM composition, are also influenced by Fe_{diss} [16,17]. Plots of the spectral slopes S₃₅₀₋₄₀₀ and S₄₀₀₋₄₆₀ versus the ratio Fe_{diss}/a₄₄₀ showed no trends, but S₂₇₅₋₂₉₅ had a trend of smaller slopes with increasing Fe_{diss}/a₄₄₀, albeit with considerable scatter. S₂₇₅₋₂₉₅ values > 0.020 generally were from sites with a₄₄₀ < 3.0 m⁻¹, where low-colored autochthonous and anthropogenic DOM was dominant. Sites dominated by allochthonous DOM (a₄₄₀ > 3.0 m⁻¹) had lower scatter, but the trend explained little variance in S₂₇₅₋₂₉₅ (R² = 0.13). The trend in S₂₇₅₋₂₉₅ generally agrees with findings of others [16,17], who reported that Fe_{diss} decreased spectral slopes. The lack of trends in S₃₅₀₋₄₀₀ and S₄₀₀₋₄₆₀, however, reinforces the conclusion that Fe_{diss} at levels found in UGLS lakes does not strongly influence absorbance in the UV-A and visible regions.

Addition of Fe_{diss} had minor effects on a₄₄₀

To measure effects of Fe_{diss} on a₄₄₀ directly, we added known amounts of an acidified Fe^{III} solution to six lake waters with a range of a₄₄₀, DOC, and Fe_{diss} (Table 4). Although a₄₄₀ increased linearly with added Fe^{III} after readjusting the pH to the original value (Fig 6), the rate was small. The average rate of increase, 0.242 m⁻¹ per 100 µg/L of added Fe^{III}, was within the range observed by others: 0.19 and 0.29 m⁻¹ per 100 µg/L of added Fe^{III} ([3,17], respectively). The weak response to Fe^{III} additions indicates that changes in Fe_{diss} have only small effects on a₄₄₀, and inspection of absorbance spectra over the range 250–500 nm showed no

Table 4. Chemical characteristics of lakes and results for iron addition experiment.

Lake	a ₄₄₀ m ⁻¹	a ₂₅₄ m ⁻¹	DOC mg/L	SUVA L m ² /g	Fe _{diss} µg/L	Fe _{diss} /DOC µg/mg	Slope, a ₄₄₀ (m ⁻¹) per 100 µg/L of added Fe _{diss}	R ² , a ₄₄₀ vs. added Fe _{diss}
Bear Island	3.7	87	12.7	3.4	42	3.3	0.15	0.92
So. Sturgeon	15.4	253	24.0	4.6	396	16.5	0.22	0.84
Section 11	16.6	256	24.4	4.5	387	15.9	0.30	0.71
Johnson	21.6	310	27.4	4.9	492	17.9	0.26	0.63
Big Sandy River Lake	23.0	354	29.0	5.3	1217	41.9	0.17	0.84
Thomson Reservoir	23.5	363	33.2	4.7	577	17.3	0.35	0.98

<https://doi.org/10.1371/journal.pone.0211979.t004>

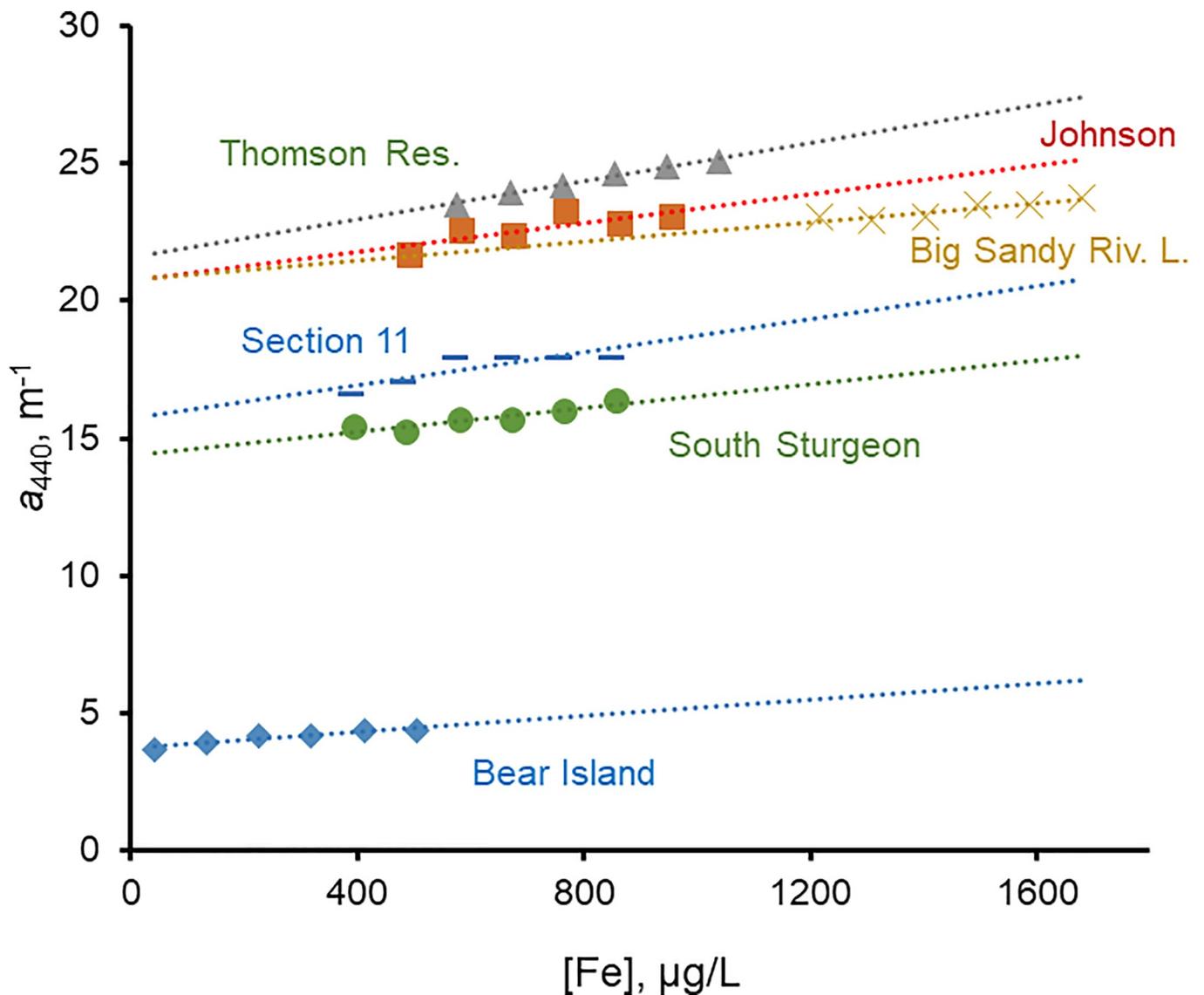


Fig 6. a_{440} vs. Fe_{diss} for six waters in iron addition experiment; slopes and statistical information on best-fit lines in Table 4.

<https://doi.org/10.1371/journal.pone.0211979.g006>

changes in shapes of the spectra. Addition of 3.0 mL of acidified stock Fe^{III} solution to 500 mL of deionized water or to 500 mL of 0.01 M EDTA at pH 6.5 yielded no measurable increases in a_{440} .

If the relationships in Fig 6 apply across the entire Fe^{III} range, a substantial fraction of the ambient a_{440} remains at $Fe_{diss} = 0$; extrapolating the best-fit lines in Fig 6 to the ordinate yielded a_{440} values in the range 90.3–99.7% (mean of 95.5%) of the ambient a_{440} . On average for the six lakes, ~ 95% of measured a_{440} thus can be attributed to DOM and only ~ 5% to enhanced absorptivity from Fe_{diss} or Fe_{diss} -DOM complexes. Xiao et al. [17] reported that Fe_{diss} (or Fe-DOM complexes) was responsible for up to 56% of a_{410} in 13 natural waters. Their highest value, however, was from a Finnish groundwater spring with very low a_{410} ($0.2\ m^{-1}$) and $Fe_{diss} = 42\ \mu\text{g/L}$. DOC was not reported but likely very low. It was an outlier among the waters, and Fe contributions to a_{410} for the 12 other samples were 0.6–8.7% (mean = 2.8%).

The fact that a_{440} did not increase when Fe^{III} was added to DI water or to an EDTA solution but increased by small amounts when added to CDOM-rich waters indicates that the increase is caused by interaction of Fe^{III} with DOM molecules and not Fe^{III} absorbance itself. The chemical nature of Fe interactions with DOM is complicated [26], and how they may affect absorption of visible light (e.g., at 440 nm) still is not well understood. Fe^{III} -complexes with carboxylate groups in humic substances can undergo photochemical reduction to Fe^{II} [27,28], and the occurrence of Fe^{II} -humic complexes in oxic waters thus cannot be ruled out. There also is evidence that some Fe associated with aquatic humic substances is bound irreversibly, apparently not as conventional metal-ligand complexes [29–31]. The literature has conflicting information on Fe^{II} stability in the presence of humic substances. Complexation by humic substances inhibited Fe^{II} autoxidation rates (oxidation by O_2) [32], but fulvic acid accelerated Fe^{II} oxidation by hydrogen peroxide (an intermediate in O_2 reduction to H_2O) [33]. Nonetheless, several studies [34,35] reported that Fe^{III} forms stronger complexes with DOM than Fe^{II} and probably is the predominant Fe-DOM form in oxic waters. Stability constants (K_f) for Fe^{III} and Fe^{II} with 12 DOM sources [35] were 10^2 – 10^4 higher for Fe^{III} than for Fe^{II} although stability constants varied widely among the DOM sources. Overall, our results indicate that Fe^{III} complexation by DOM has very small effects on CDOM chromophoric groups.

Application of experimental results to field data

We applied the experimental results to our field data to further evaluate Fe_{diss} effects on a_{440} , which is critical to know before attempting to use a_{440} to predict DOC. For example, the Fe_{diss} - a_{440} regression of the 2015 data (Fig 7) showed that some data points were far from the regression line. For the largest outliers (six high and eight low), we estimated the change in a_{440} that would occur if Fe_{diss} were adjusted to the “best fit” values of the regression relationship. The difference between measured and best-fit Fe_{diss} , multiplied by 0.242 m^{-1} per $100 \mu\text{g/L}$ of Fe_{diss} (average slope of the a_{440} - Fe^{III} relationship, Fig 6), provided estimates of the a_{440} change caused by the Fe_{diss} change. The results showed small a_{440} changes even for waters with large differences between measured and best-fit Fe_{diss} . For example, Blueberry Lake had the highest measured Fe_{diss} ($1224 \mu\text{g/L}$), and the best-fit Fe_{diss} for its measured a_{440} (19.6 m^{-1}) is $690 \mu\text{g/L}$. If the latter value represented the Fe_{diss} in this lake, a_{440} would be 18.3 m^{-1} , a decrease of 1.3 m^{-1} (a 6.6% change). Similar changes were found for the other waters with large differences between measured and best-fit Fe_{diss} (S6 Table); the average a_{440} change for the six high outliers was -3.7% (range -1.8 to -6.6%), and the average for the eight low outliers was $+2.2\%$ (range 1.2 to 2.9%).

“Iron-corrected” a_{440} values for samples with Fe_{diss} data are estimates of the a_{440} attributable to DOM alone ($a_{440,\text{OM}}$). These were obtained by multiplying measured Fe_{diss} values by 0.242 (average slope of the a_{440} - Fe^{III} relationships in Fig 6) and subtracting the result from measured a_{440} . For 136 waters with $a_{440} > 3.0 \text{ m}^{-1}$, $a_{440,\text{OM}}$ was $92.3 \pm 5.0\%$ (range of 71.3 – 99.7%) of measured a_{440} . Fe_{diss} accounted for $< 10\%$ of a_{440} in most lakes (102, or 75%), but in seven lakes it accounted for 15–30% of measured a_{440} (S7 Table). These lakes generally were river-influenced systems with relatively high Fe_{diss} concentrations and/or high $\text{Fe}_{\text{diss}}/\text{DOC}$ ratios, and all were samples collected in the high rainfall year 2016. Of the six lakes that also had data for 2014 or 2015, only Big Sandy Lake and Big Sandy River Lake had Fe_{diss} contributions to $a_{440} > 10\%$ in those years. We conclude that high rainfall promotes Fe export to lakes, resulting in higher Fe_{diss} contributions to a_{440} , and that lakes influenced by rivers with high CDOM and Fe are more likely to have relatively high Fe_{diss} contributions to a_{440} .

DOC can be predicted from a_{440} without correcting for Fe_{diss}

It is now possible to assess whether correction for the presence of Fe_{diss} is needed to allow accurate prediction of DOC from a_{440} . Regressions of measured DOC versus $a_{440,\text{OM}}$ and

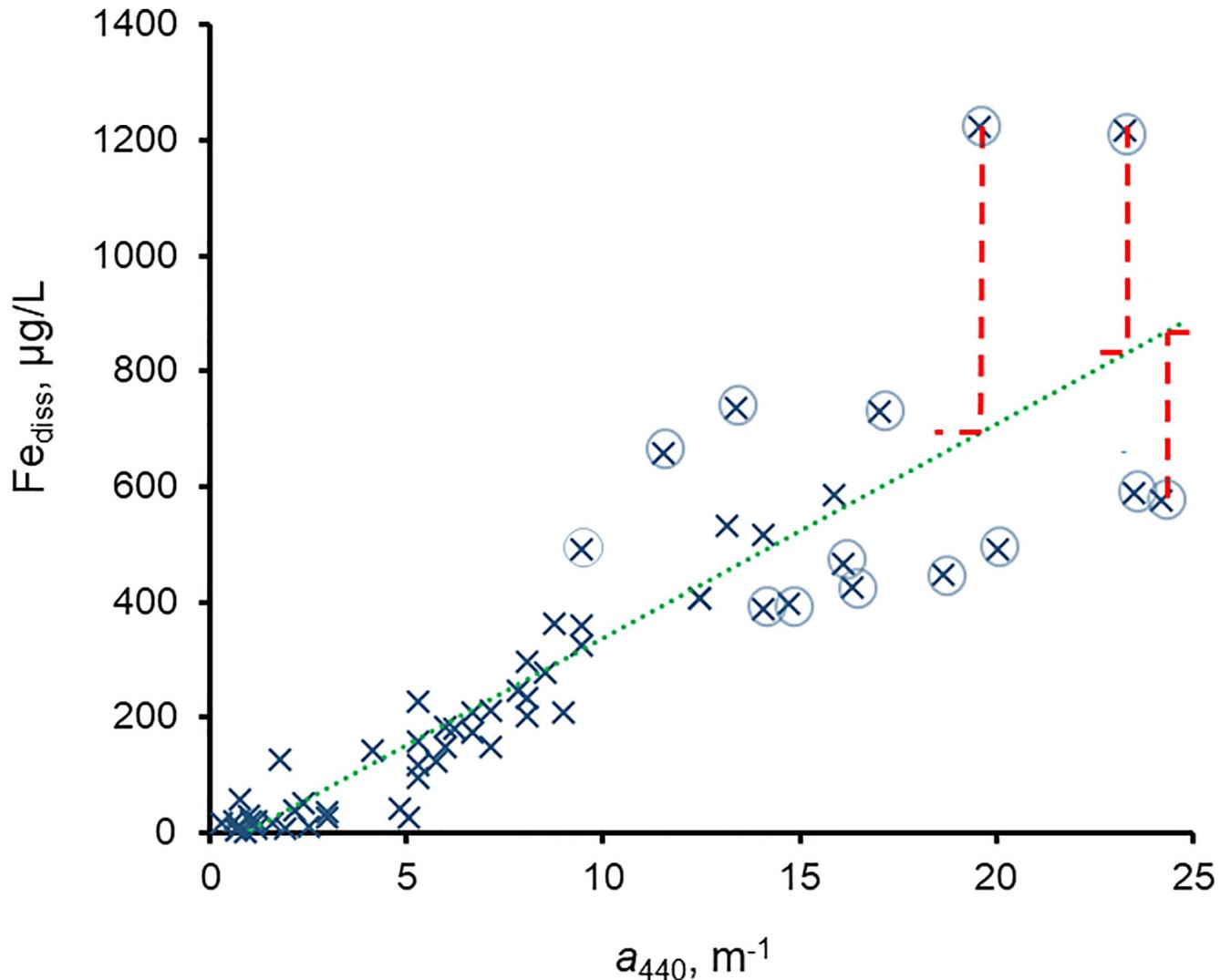


Fig 7. Fe_{diss} vs. a_{440} for 2015 data; best-fit regression line and statistics given in Table 2. Circled data: outliers examined for effects of Fe_{diss} on a_{440} (S6 Table). Dotted line: best-fit linear relationship: $Fe_{diss} = 36.8a_{440} - 30.7$; $R^2 = 0.79$. Dashed lines illustrate change in a_{440} for some outliers when Fe_{diss} was changed to “best-fit” value.

<https://doi.org/10.1371/journal.pone.0211979.g007>

measured DOC versus measured a_{440} yielded very similar relationships for the same data set ($a_{440} > 3.0 \text{ m}^{-1}$; $N = 134$):

$$\ln(\text{DOC}_{meas}) = 0.587 \times \ln(a_{440,OM}) + 1.57; \quad R^2 = 0.86, \text{RMSE} = 0.146, \text{slope SE} = 0.020, p < 0.0001 \quad (3)$$

$$\ln(\text{DOC}_{meas}) = 0.573 \times \ln(a_{440,meas}) + 1.55; \quad R^2 = 0.85, \text{RMSE} = 0.154, \text{slope SE} = 0.020, p < 0.0001 \quad (4)$$

The DOC predicted by Eq 3 from the estimated $a_{440,OM}$ for each site was compared to the DOC predicted from the measured a_{440} -DOC relationship (Eq 4); the relationship was almost exactly 1:1 (slope = 0.997; $R^2 = 0.99$) (S2 Fig). Consequently, we conclude that DOC predictions from measured a_{440} are just as accurate for our study sites as DOC predictions from a_{440}

corrected to remove the influence of Fe_{diss} . Moreover, for waters where $a_{440} > 3.0 \text{ m}^{-1}$, a_{440} is a good predictor of DOC (Table 3).

Long-term trends in CDOM and the role of Fe_{diss} in UGLS lakes

Long-term data across UGLS surface waters for CDOM [18] and Fe_{diss} are scarce. The extent of regional increases in CDOM, or whether such increases could be attributed to increases in Fe_{diss} , thus is unknown. Smaller-scale analyses suggest, however, that regional CDOM trends are more complicated than observed in Scandinavia and likely driven by climatic and hydrologic variations. For example, substantial intra- and inter-annual variations but no monotonic trends were found in a_{440} and DOC for 20 small lakes in Upper Michigan over six years [36]. Climatic conditions that affected carbon loadings from upland forests and wetlands were considered the drivers of these variations. Similarly, Brezonik et al. [18] found large a_{440} variations in seven lakes of the northern Wisconsin LTER program (all within a radius of 10 km, [37]) for the period 1990–2012, but only colored Crystal Bog had increasing a_{440} over the whole period. In a study on optical properties of the LTER lakes, Jane et al. [37] found inconsistent trends in DOC since 1990, with increases in two (including Crystal Bog), decreases in four, and no trend in one. Optical properties related to DOM chemical characteristics varied more with climatic conditions than DOC concentrations.

Björnerås et al. [38] recently reported temporal increases in Fe on broad scales in European and North American freshwaters. Their data overlap our region at only one site (in north-central Wisconsin). Additional data from the Wisconsin LTER lakes, which were not included in the Björnerås et al. study, showed that Fe increased in Crystal Bog, decreased in Trout Bog, and had no trends in the other five lakes since the 1980s (Mann-Kendall test; N. Lottig, Univ. Wisconsin, pers. comm., 2017). As noted above, Crystal Bog is the only LTER lake with increasing CDOM during the same period. The intra- and inter-annual variability in Fe and CDOM was high in all the lakes. In Crystal Bog, Fe_{diss} averaged 200 $\mu\text{g/L}$ for six pre-1990 measurements and 330 $\mu\text{g/L}$ for eight post-2010 measurements; the average increase of 130 $\mu\text{g/L}$ could account for an a_{440} increase of only $\sim 0.3 \text{ m}^{-1}$ (based on the average slope in Fig 6), but a_{440} in Crystal Bog actually increased by $\sim 5.5 \text{ m}^{-1}$ over this time [18]. Moreover, pH data for Crystal Bog showed no trends over the period of record (1981–2016) (S3 Fig). The increases in DOC [37] and a_{440} [18] in Crystal Bog thus cannot be explained by declining acidity, and the lack of similar trends in the other LTER lakes suggests that long-term climatic changes also are not responsible for the trends.

Conclusions

Our examination of the role of dissolved iron in optical properties of DOM in lakes supports three main conclusions. First, a_{440} and Fe_{diss} are well correlated in surface waters of the UGLS, with R^2 values of 0.73–0.77 for individual years, as has been shown in studies elsewhere. Second, experimental data show that iron has small effects on CDOM measured as a_{440} , but it is not the dominant factor for a_{440} or SUVA_{254} variations in UGLS lakes. The average increase in a_{440} with added Fe_{diss} (0.242 m^{-1} per 100 $\mu\text{g/L}$) means that increasing Fe_{diss} by 400 $\mu\text{g/L}$ would increase a_{440} by only $\sim 1.0 \text{ m}^{-1}$. Even this level of Fe_{diss} variation leads to an a_{440} change less than the expected error in using a_{440} as a proxy for DOC (RMSE of 1.3–1.8 m^{-1} , Table 3). Third, estimates of DOC based on measured a_{440} and $a_{440,\text{DOM}}$ (i.e., a_{440} corrected for Fe_{diss}) were essentially the same. Consequently, our data indicate that ambient levels of Fe_{diss} have only a minor influence on CDOM optical properties (a_{440} and SUVA_{254}) and do not affect DOC estimates based on a_{440} in lakes of our study area.

Supporting information

S1 Fig. Histograms of data distributions for a_{440} (CDOM), DOC, Fe_{diss} , and $SUVA_{254}$. Upper plots: untransformed data; lower plots: log-transformed (ln) values. (DOCX)

S2 Fig. DOC predicted from Fe-corrected a_{440} ($a_{440,OM}$) (Eq 3) vs. DOC predicted from measured a_{440} (Eq 4). Best-fit line: $a_{440,OM} = 0.997a_{440} + 0.032$; $R^2 = 0.99$; RMSE = 0.685, slope SE = 0.0086, $p < 0.0001$. (DOCX)

S3 Fig. Time trend for pH in Crystal Bog, Vilas County, Wisconsin, 1981–2016; data from the North Temperate Lakes Long Term Ecological Research (LTER) program (<http://lter.limnology.wisc.edu>). (DOCX)

S1 Table. Vertical profile data for three NLF lakes with a wide range of surface CDOM (a_{440}) values. (DOCX)

S2 Table. Fe_T , Fe_{diss} , and % Fe_{diss} for 2018 lake and associated river samples from the NLF ecoregion. (DOCX)

S3 Table. Fe_{diss} - a_{440} and Fe_{diss} -DOC relationships for log-transformed data. (DOCX)

S4 Table. Log-transformed regression relationships for a_{440} vs. DOC and Fe_{diss} . (DOCX)

S5 Table. $SUVA_{254}$ values for samples with measured $SUVA_{254} > 5.0$ before and after Fe_{diss} correction. (DOCX)

S6 Table. Changes in a_{440} for 2015 waters that had large differences between measured and best-fit Fe_{diss} after Fe_{diss} was changed to the best-fit value. (DOCX)

S7 Table. Samples with measured $a_{440} > 3.0 \text{ m}^{-1}$ having 15–30% of a_{440} caused by Fe_{diss} and 70–85% caused by colored DOM. (DOCX)

Acknowledgments

We gratefully acknowledge support from the National Science Foundation, the Minnesota Environmental and Natural Resources Trust fund, as recommended by the Legislative-Citizen Commission on Minnesota Resources, and Univ. of Minnesota's Office of the VP for Research and Retirees Association, U-Spatial Program, Sea Grant Program, and Agricultural Experiment Station. We thank numerous collaborators and student workers for assistance in sample collection and analysis. The senior author gratefully acknowledges mentoring early in his career by G.F. Lee, himself an iron researcher.

Author Contributions

Conceptualization: Patrick L. Brezonik.

Data curation: Patrick L. Brezonik, Claire G. Griffin.

Formal analysis: Patrick L. Brezonik.

Funding acquisition: Patrick L. Brezonik, Jacques C. Finlay, Raymond M. Hozalski.

Investigation: Patrick L. Brezonik, Jacques C. Finlay, Claire G. Griffin, Evelyn H. Boardman, Noah Germolus, Raymond M. Hozalski, Leif G. Olmanson.

Methodology: Patrick L. Brezonik, Claire G. Griffin, William A. Arnold, Evelyn H. Boardman, Noah Germolus.

Project administration: Patrick L. Brezonik, Jacques C. Finlay, Raymond M. Hozalski.

Resources: Jacques C. Finlay, William A. Arnold.

Supervision: Patrick L. Brezonik, Jacques C. Finlay, Claire G. Griffin, William A. Arnold, Raymond M. Hozalski.

Visualization: Patrick L. Brezonik, Claire G. Griffin, Leif G. Olmanson.

Writing – original draft: Patrick L. Brezonik.

Writing – review & editing: Patrick L. Brezonik, Jacques C. Finlay, Claire G. Griffin, William A. Arnold, Evelyn H. Boardman, Noah Germolus, Raymond M. Hozalski, Leif G. Olmanson.

References

1. Aschan O. Soluble humus material of northern fresh waters. *J Prakt Chemie* 1908; 77: 172.
2. Lohammar G. Wasserchemie und höhere Vegetation schwedischer Seen. *Symb. Bot. Upsaliens* 1938; 3: 1–252.
3. Kritzberg ES, Ekström SM. Increasing iron concentrations in surface waters—a factor behind brownification? *Biogeosciences* 2012; 9: 1465–1478. <https://doi.org/10.5194/bg-9-1465-2012>
4. Weyhenmeyer GA, Prairie YT, Tranvik LJ. Browning of boreal freshwaters coupled to carbon-iron interactions along the aquatic continuum. *PLoS ONE* 2014; 9: (2) e88104. <https://doi.org/10.1371/journal.pone.0088104> PMID: 24505396
5. Monteith DT, Stoddard JL, Evans CD, de Wit HA, Forsius M, Hørgåsen T, et al. Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. *Nature* 2007; 450: 537–539. <https://doi.org/10.1038/nature06316> PMID: 18033294
6. Haaland S, Hongve D, Laudon H, Riise G, Vogt RD. Quantifying the drivers of the increasing colored organic matter in Boreal surface waters. *Environ. Sci. Technol.* 2010; 44: 2975–2980. <https://doi.org/10.1021/es903179j> PMID: 20329770
7. Roulet N, Moore TR. Environmental chemistry—Browning the waters. *Nature* 2006; 444: 283–284. <https://doi.org/10.1038/444283a> PMID: 17108948
8. Ekström SM, Regnell O, Reader HE, Nilsson PA, Löfgren S, Kritzberg ES. Increasing concentrations of iron in surface waters as a consequence of reducing conditions in the catchment area. *J. Geophys. Res. Biogeosci.* 2016; 121: 479–493. <https://doi.org/10.1002/2015JG003141>
9. Kutser T, Pierson DC, Kallio KY, Reinart A, Sobek S. Mapping lake CDOM by satellite remote sensing. *Remote Sens. Environ.* 2005; 94: 535–540.
10. Olmanson LG, Brezonik PL, Finlay JC, Bauer ME. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sens. Environ.* 2016; 185: 119–128. <https://doi.org/10.1016/j.rse.2016.01.007>
11. Köhler SJ, Kothawala D, Futter MN, Liungman O, Tranvik LJ. In-lake processes offset increased terrestrial inputs of dissolved organic carbon and color to lakes. *PLoS ONE* 2013; 8: e70598. <https://doi.org/10.1371/journal.pone.0070598> PMID: 23976946
12. Tranvik LJ, Jansson M. Climate change (Communication arising): Terrestrial export of organic carbon. *Nature* 2002; 415: 861–862. <https://doi.org/10.1038/415861b>

13. Garmo OA, Skjelkvale BL, De Wit HA, Colombo L, Curtis C, Fölster J, et al. Trends in surface water chemistry in acidified areas in Europe and North America from 1990 to 2008. *Water Air Soil Pollut.* 2014; 225: 1–14.
14. Weishaar JL, Aiken GR, Bergamaschi BA, Fram MS, Fujii R, Mopper K. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* 2003; 37: 4702–4708. PMID: [14594381](https://pubmed.ncbi.nlm.nih.gov/14594381/)
15. Poulin BA, Ryan JR, Aiken GR. Effects of iron on optical properties of dissolved organic matter. *Environ. Sci. Technol.* 2014; 48, 10098–10106. <https://doi.org/10.1021/es502670r> PMID: [25084347](https://pubmed.ncbi.nlm.nih.gov/25084347/)
16. Maloney KO, Morris DP, Moses CO, Osburn CL. The role of iron and dissolved organic carbon in the absorption of ultraviolet radiation in humic lake water. *Biogeochemistry* 2005; 75: 393–407.
17. Xiao YH, Sara-Aho T, Hartikainen H, Vähätalo AV. Contribution of ferric iron to light absorption by chromophoric dissolved organic matter. *Limnol. Oceanogr.* 2013; 58: 653–662. <https://doi.org/10.4319/lo.2013.58.2.0653>
18. Brezonik PL, Olmanson LG, Finlay JC, Bauer ME. Factors affecting the measurement of CDOM in optically complex inland waters. *Remote Sens. Environ.* 2015; 157: 199–215. <https://doi.org/10.1016/j.rse.2014.04.033>
19. Griffin CG, Finlay JC, Brezonik PL, Olmanson LG, Hozalski RM. Limitations on using CDOM as a proxy for DOC in temperate lakes. *Water Res.* 2018; 144: 719–727. <https://doi.org/10.1016/j.watres.2018.08.007> PMID: [30099300](https://pubmed.ncbi.nlm.nih.gov/30099300/)
20. Peacock M, Freeman C, Gauci V, Lebron I, Evans CD. Investigations of freezing and cold storage for the analysis of peatland dissolved organic carbon (DOC) and absorbance properties. *Env. Sci.: Processes Impacts* 2015; 17: 1290–1301.
21. Shapiro J. Effect of yellow organic acids on iron and other metals in water. *J. Amer. Water Works Assoc.* 1964; 56: 1062–1082.
22. Garcia-Mina JM. Stability, solubility and maximum metal binding capacity in metal-humic complexes involving humic substances extracted from peat and organic compost. *Org. Geochem.* 2006; 37:1960–1972.
23. Singley JE, Harris RH, Maulding JS. Correction of color measurements to standard conditions. *J. Amer. Water Wks. Assoc.* 1966; 58: 455–457.
24. Brezonik PL, Bouchard RW Jr, Finlay JC, Griffin CG, Olmanson LG, Anderson JP, et al. Color, chlorophyll *a* and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. *Ecol. Appl.* 2019; 29: (in press).
25. Dilling J, Kaiser K. Estimation of the hydrophobic fraction of dissolved organic matter in water samples using UV photometry. *Water Res.* 2002; 36: 5037–5044. PMID: [12448552](https://pubmed.ncbi.nlm.nih.gov/12448552/)
26. Brezonik PL, Arnold WA. *Water chemistry.* New York, Oxford: Oxford University Press; 2011.
27. Miles CJ, Brezonik PL. Oxygen consumption by a photochemical ferrous-ferric catalytic cycle. *Environ. Sci. Technol.* 1981; 15: 1089–1095. <https://doi.org/10.1021/es00091a010> PMID: [22284115](https://pubmed.ncbi.nlm.nih.gov/22284115/)
28. Xie H, Zafiriou OC, Cai WJ, Zepp RG, Wang Y. Photooxidation and its effects on the carboxyl content of dissolved organic matter in two coastal rivers in the southeastern United States. *Environ. Sci. Technol.* 2004; 38: 4113–4119. PMID: [15352449](https://pubmed.ncbi.nlm.nih.gov/15352449/)
29. Hutchinson GE. *A Treatise on limnology, vol. 1.* New York: Wiley-Interscience; 1957.
30. Senesi N, Griffith SM, Schnitzer M, Townsend MG. Binding of Fe³⁺ by humic materials. *Geochim. Cosmochim. Acta* 1977; 41: 969–976.
31. Sedlacek J, Gjessing E, Rambaek JP. Isotope exchange between inorganic iron and iron naturally complexed by aquatic humus. *Sci. Tot. Env.* 1987; 62: 275–279.
32. Theis TL, Singer PC. Complexation of iron(II) by organic matter and its effect on iron(II) oxygenation. *Environ. Sci. Technol.* 1974; 8: 569–573.
33. Voelker BM, Sulzberger B. Effects of fulvic acid on Fe(II) oxidation by hydrogen peroxide. *Environ. Sci. Technol.* 1996; 30: 1106–1114.
34. Emmenegger L, Sigg L, Sulzberger B. Light-induced redox cycling of iron in circumneutral lakes. *Limnol. Oceanogr.* 2001; 46: 49–61.
35. Rose AL, Waite TD. Kinetics of iron complexation by dissolved natural organic matter in coastal waters. *Mar. Chem.* 2003; 84: 85–103.
36. Pace ML, Cole JJ. Synchronous variation of dissolved organic carbon and color in lakes. *Limnol. Oceanogr.* 2002; 47: 333–342.
37. Jane SF, Winslow LA, Remucal CK, Rose KC. Long-term trends and synchrony in dissolved organic matter characteristics in Wisconsin, USA, lakes: Quality, not quantity, is highly sensitive to climate. *J. Geophys. Res. Biogeosci.* 2017; 122: 546–561. <https://doi.org/10.1002/2016JG003630>

38. Björnerås C, Weyhenmeyer GA, Evans CD, Gessner MO, Grossart H-P, Kangur K, et al. Widespread increases in iron concentration in European and North American freshwaters. *Global Biogeochem. Cycles* 2017; 31: 1488–1500. <https://doi.org/10.1002/2017GB005749>



Satellite Remote Sensing

Protecting the quality of lakes and rivers is an important societal goal, but because of high costs, only a small fraction of these resources can be monitored by ground-based methods. The technology of satellite imagery has improved greatly in the past two decades, and key water quality features now can be measured at low cost by satellite sensors on virtually all lakes. These properties – chlorophyll, turbidity, colored dissolved organic matter (CDOM) and water clarity (Secchi depth) – control many important biological, chemical and physical characteristics of lakes.

Use of satellite technology has greatly expanded information available from ground-based monitoring, improved our understanding of the factors driving changes in surface waters, and helped improve data-driven resource management. This brochure describes current capabilities to measure key water quality characteristics of Minnesota's lakes using satellite imagery, as well as on-going studies by our research group to improve the effectiveness and use of these techniques.

Future prospects

With the expanded availability of optical and radar data, remote sensing of water quality and other characteristics (e.g. lake snow/ice phenology) will become widely used over the next decade and will greatly expand our knowledge on the status of water quality and thermal characteristics in Minnesota lakes. Our research group is developing a state-of-the-art system that will automatically download Sentinel and Landsat imagery for Minnesota and use high-performance computing technology to provide water quality information more quickly and more completely than is possible with conventional image processing.

For more information:

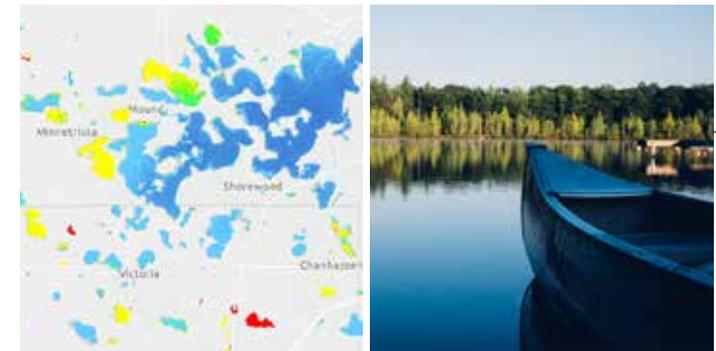
Marvin Bauer, PROFESSOR EMERITUS OR

Leif Olmanson, RESEARCH SCIENTIST
DEPARTMENT OF FOREST RESOURCES AND REMOTE SENSING AND
GEOSPATIAL ANALYSIS LABORATORY

mbauer@umn.edu and olman002@umn.edu
or visit <https://water.rs.umn.edu>

Support to develop this brochure was provided from National Science Foundation grant CBET 1510332, citizen science supplement, to Professor Ray Hozalski and Professor Emeritus Patrick Brezonik, Department of Civil, Environmental, and Geo- Engineering.

<https://water.rs.umn.edu>



Remote Sensing
for Water Quality



CFANS
COLLEGE OF FOOD, AGRICULTURAL
AND NATURAL RESOURCE SCIENCES

**WATER RESOURCES
CENTER**



CFANS
COLLEGE OF FOOD, AGRICULTURAL
AND NATURAL RESOURCE SCIENCES

**WATER RESOURCES
CENTER**



A harmonized image processing workflow using Sentinel-2/MSI and Landsat-8/OLI for mapping water clarity in optically variable lake systems



Benjamin P. Page^{a,*}, Leif G. Olmanson^b, Deepak R. Mishra^c

^a Water Resources Center, University of Minnesota, St. Paul, MN, USA

^b Remote Sensing and Geospatial Analysis Laboratory, Department of Forest Resources, University of Minnesota, St. Paul, MN, USA

^c Center for Geospatial Research, Department of Geography, University of Georgia, Athens, GA, USA

ARTICLE INFO

Keywords:

Water clarity
Harmonized
Atmospheric correction
Landsat-8
Sentinel-2
Lake monitoring
Lake management
Remote sensing
Validation

ABSTRACT

This study demonstrates the applicability of harmonizing Sentinel-2 MultiSpectral Imager (MSI) and Landsat-8 Operational Land Imager (OLI) satellite imagery products to enable the monitoring of inland lake water clarity in the Google Earth Engine (GEE) environment. Processing steps include (1) atmospheric correction and masking of MSI and OLI imagery, and (2) generating scene-based water clarity maps in terms of Secchi depth (SD). We adopted ocean-color based atmospheric correction theory for MSI and OLI sensors modified with associated scene-specific metadata and auxiliary datasets available in GEE to generate uniform remote sensing reflectances (R_{rs}) products over optically variable freshwater lake surfaces. MSI- R_{rs} products derived from the atmospheric correction were used as input predictors in a bootstrap forest to determine significant band combinations to predict water clarity. A SD model for MSI (SD_{MSI}) was then developed using a calibration dataset consisting of log-transformed $SD_{in situ}$ measurements ($\ln SD_{in situ}$) from 79 optically variable freshwater inland lakes collected within ± 1 day of satellite overpass on 23-Aug 2017 (MAE = 0.53 m) and validated with 276 samples collected within ± 1 day of a 12-Sep 2017 image (MAE = 0.66 m) across three ecoregions in Minnesota, USA. A separate SD model for MSI was also developed using similar spectral bands present on the OLI sensor ($SD_{s,OLI}$) where cross-sensor performance can be evaluated during coincident overpass events. $SD_{s,OLI}$ applied to both MSI and OLI (SD_{OLI}) models were further validated using two coincident overpass sets of imagery on 27-Sep 2017 ($n = 18$) and 13-Aug 2018 ($n = 43$), yielding a range of error from 0.25 to 0.67 m. Potential sources of model errors and limitations are discussed. Data derived from this multi-sensor methodology is anticipated to be used by researchers, lake resource managers, and citizens to expedite the pre-processing steps so that actionable information can be retrieved for decision making.

1. Introduction

The abundant surface waters in Minnesota face multiple threats from land-use change, eutrophication, invasive species, and warming temperatures (Bossenbroek et al., 2001; O'Reilly, 2015; Smith, 2016). Protecting water quality is critically important to lake-rich states because of the ecological and economic importance of water activities and tourism. To understand and ensure the sustainability of these aquatic ecosystems on a statewide scale, adoption of publicly available satellite Earth observation data will be necessary for effective management. While traditional *in situ* sampling methodologies can provide rapid, accurate information about targeted lakes, sampling more than a fraction of > 10,000 is laboriously challenging. On the other hand, research efforts using Landsat data to model water quality parameters such as turbidity and algal pigment concentration across inland lakes date as

far back as 1978 (Carpenter and Carpenter, 1983). Two of the most relevant for Minnesota lakes is the water clarity image processing protocol developed by Olmanson et al. (2001) and Kloiber et al. (2002a) using Landsat Thematic Mapper (TM) and Multi-spectral Scanner (MSS) sensor data. Since then, the launch of Landsat-8 (13-Feb 2013) carrying the Operational Land Imager (OLI) and the European Space Agency's (ESA) Sentinel-2A (23-Jun 2015) and 2B (7-Mar 2017) MultiSpectral Imager (MSI) constellation have advanced the capabilities of water quality products that can be derived from remote sensing systems (Tyler et al., 2016; Pahlevan et al., 2017a; Pahlevan et al., 2017b).

While interest has increased since earlier studies (Kloiber et al., 2002a, 2002b), routine monitoring of lake water quality using satellite remote sensing is still not a common practice by resource management agencies; particularly, at a time when moderate resolution imagery is

* Corresponding author.

E-mail address: pagex135@umn.edu (B.P. Page).

<https://doi.org/10.1016/j.rse.2019.111284>

Received 4 September 2018; Received in revised form 24 June 2019; Accepted 25 June 2019

0034-4257/ © 2019 Elsevier Inc. All rights reserved.

readily available from repository platforms like ESA's Copernicus Open Access Hub, Amazon Web Services, and the U.S. Geological Survey (USGS) Earth-Explorer at no cost. The absence of satellite remote sensing strategies for routine monitoring of lake water quality parameters on a statewide basis likely is attributable to the computational demands required for managing and analyzing the large volume of statewide imagery using common processing software. To this end, high performance and cloud-computing infrastructures will aid in image processing experimentation and efficiency. For example, Google's Earth Engine (GEE) cloud computational platform houses petabytes of imagery and other auxiliary geospatial datasets that can be accessed through a JavaScript application programming interface (Gorelick et al., 2017), and dense temporal MSI and OLI datasets can be seamlessly searched, handled and processed at a significantly faster rate.

The overall objectives of this research were to (1) implement modified atmospheric correction formulas into the GEE platform to produce harmonized remote sensing reflectance (R_{rs}) products between MSI and OLI imagery, and (2) demonstrate the applicability of the validated R_{rs} products by mapping water clarity (SD) across Minnesota's optically variable and temporarily dynamic inland lakes. Recently, the capability of both MSI and OLI to yield similar reflectance values has been shown to be feasible for both water (Vanhellemont and Ruddick, 2016; Page et al., 2018; Pahlevan et al., 2019) and land (Claverie et al., 2018), and investigating further into these capabilities in the GEE platform allow for faster cross-sensor experimentation than using conventional image processing techniques. In terms of water clarity model development, the Secchi depth (SD) relies heavily on multispectral channels centered in the blue to red region (typically 443 nm to 670 nm) of the electromagnetic spectrum (Harrington and Schiebe, 1992; Giardino et al., 2001; Olmanson et al., 2013; Lee et al., 2016), where the combination of ozone (absorbing) and Rayleigh (scattering) effects may constitute as much as 90% of the total top-of-atmosphere reflectance (ρ TOA) received by the satellite sensor (Gordon et al., 1997; Mishra et al., 2005). If we are to take advantage of available imagery for near real-time monitoring at higher frequencies, a consistent correction to compensate for the temporal variations in atmospheric properties in these wavelengths is necessary, and anticipated to strengthen future and existing models of water clarity. However, the consistency of any atmospheric correction over time across optically variable inland lakes is not yet fully understood. To address this matter, we processed MSI and OLI imagery using our Modified Atmospheric correction for Inland waters (MAIN) (described in Section 2.3) within the GEE environment and compared the generated R_{rs} values against the R_{rs} values converted from the USGS L8 Surface Reflectance Product (OLI-SR) as a reference. Further, an external image processing software (ACOLITE, Vanhellemont and Ruddick, 2015, version 20,190,326) also equipped to generate R_{rs} products from MSI and OLI imagery was included in the comparison to provide a secondary means of performance.

If the long-term goal is to develop reliable and cost-effective approaches to regional measurements of major indicators of water quality that can be used by management agencies to extend ground-based measurements (Olmanson et al., 2008; Olmanson et al., 2001; Kloiber et al., 2002b) then development of automated approaches that can take advantage of the improved spectral, spatial, radiometric and temporal resolution of the MSI and OLI systems are needed for improved water quality monitoring and fisheries management. Here, a SD model is developed from atmospherically corrected MSI imagery using a robust *in situ* SD dataset ($SD_{in situ}$) collected from the Citizen Lake Monitoring Program (CLMP), coordinated by the Minnesota Pollution Control Agency (MPCA, www.pca.state.mn.us/water/resources-volunteers) across three ecoregions in Minnesota, USA. Significant band-ratios in predicting SD were chosen based on a bootstrap forest technique (Breiman, 1996). Considering the spectral and spatial similarities between five MSI and OLI sensor bands in the visible and NIR portion of the electromagnetic spectrum (Table 1), a more restricted SD model was developed for MSI using comparable OLI bands (SD_{sOLI}) so that the

model may be extrapolated to OLI data to increase overall temporal resolution when using both sensors for multi-platform water clarity assessments. The data derived from this consistent multi-sensor methodology are anticipated to be used by lake resource managers and agencies, researchers and citizens to eliminate the pre-processing steps for satellite imaging applications so that actionable information can be readily retrieved for decision-making.

2. Methodology

2.1. Study area

The study area was intended to target as many of the > 12,000 inland surface waters across the state of Minnesota, USA as possible using Level-1 MSI and OLI imagery provided in the GEE repository. Positioned in the upper Midwest United States at 43.4–49.4° N, 89.4–96.8° W, Minnesota's vast collection of optically variable water bodies is spread across seven ecoregions (Fig. 1), geographical areas where the land cover (agriculture, forest, prairie, etc.), underlying geology, soils, and potential native plant community are relatively similar (Omernik, 1987). Four of these ecoregions include ~96% of Minnesota's lakes. The Northern Lakes and Forest Ecoregion (NLF), with 46% of the state's lakes, has a higher concentration of optically clearer waters (lower chlorophyll) and lakes with higher colored dissolved organic matter (CDOM) (Brezonik et al., 2019). The North Central Hardwood Forests Ecoregion (NCHF), with roughly 38% of the state's lakes, has a wide range of water clarity. Lakes in the Western Corn Belt Plains Ecoregion (WCBP), which has 7% of the state's lakes, generally are more eutrophic and have low water clarity. The Northern Glaciated Plains (NGP) Ecoregion, with 6% of the lakes, also has low water clarity. Statewide assessments for > 10,500 lakes using Landsat data revealed that water clarity has remained stable between 1985 and 2005 in the NLF and NCHF ecoregions but declined slightly in the WCBP and the NGP (Olmanson et al., 2013).

2.2. Satellite data and image pre-processing

Different data sets of S2A/MSI and L8/OLI imagery were searched from the GEE repository and allocated for either atmospheric correction assessment and/or SD model calibration/validation (Table 2). For SD model calibration and validation we targeted mostly clear imagery acquired from a late-summer index period (July 15–September 15) when short-term variability and water clarity are at a seasonal minimum (Stadelmann et al., 2001). Before any statistical evaluation or performance assessment, all L8/OLI Level-1 imagery were first converted from the digital number (DN) format into TOA reflectance (ρ TOA):

$$\rho_{TOA, OLI} = M_L(\lambda_i) \times DN_{OLI}(\lambda_i) + A_L(\lambda_i) \quad (1)$$

where ρ TOA is the top-of-atmosphere spectral reflectance measured by the OLI sensor at wavelength λ_i , and M_L and A_L are the band multiplicative and additive coefficients found in the image metadata (Landsat-8 Data Users Handbook, V2.0). For S2/MSI, Level-1C ρ TOA is achieved by multiplying the imagery by the scaling factor (0.0001).

Next, surface water bodies were isolated by masking out surrounding terrestrial features using a threshold technique in the SWIR portion of the spectrum. A simple threshold value (0.03) was assigned to a mosaicked image which consisted of a statewide median value from the SWIR band (B7 at 2201 nm) from the entire collection of Tier-1 L8 Surface Reflectance Product (OLI-SR) with < 2% cloud cover over Minnesota between 2013 and 2018 ($n = 172$). Imagery between June and October were used for masking to avoid snow and lake-ice contaminated pixels. A masking threshold of 0.03 worked well to separate the high absorbing water features from the surrounding landscape in this area, but this value may need some adjustment for other geographical regions. Once the pre-processing steps were finalized, we

Table 1

Sensor characteristics for L8/OLI and S2/MSI, including bandcenter, bandwidth, spatial resolution and signal-to-noise ratio (SNR). SNR values have been scaled for radiances observed over clear coastal waters.

(Adopted from Pahlevan et al., 2017b.)

Landsat-8/OLI											
Band ID	B1	B2	B3	B4	–	–	–	–	B5	B6	B7
Band center (nm)	443	482	561	655	–	–	–	–	865	1609	2201
Bandwidth (nm)	20	65	60	40	–	–	–	–	30	85	190
Resolution (m)	30	30	30	30	–	–	–	–	30	30	30
Signal-to-Noise Ratio (SNR)	284	321	223	113	–	–	–	–	45	10.1	7.4

Sentinel-2/MSI											
Band ID	B1	B2	B3	B4	B5	B6	B7	B8	B8A	B11	B12
Band center (nm)	444	497	560	664	705	740	783	842	865	1610	2190
Bandwidth (nm)	20	55	35	30	15	15	15	15	20	9	175
Resolution (m)	60	10	10	10	20	20	20	10	20	20	20
Signal-to-Noise Ratio (SNR)	439	102	79	45	45	34	26	20	16	2.8	2.2

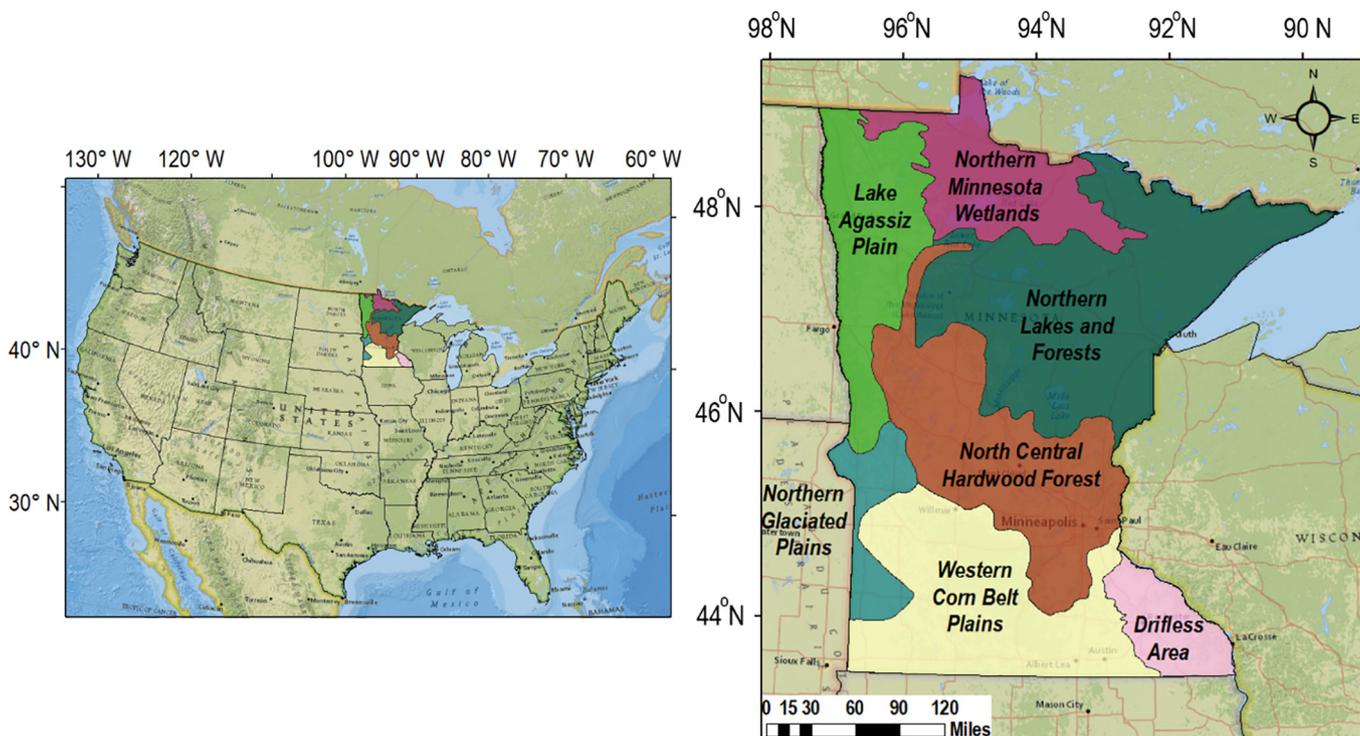


Fig. 1. Study area. Minnesota, USA and its corresponding ecoregions. Basemap source: National Geographic, Esri, Garmin, HERE, UNEP-WCMC, USGS, NASA, ESA, METI, NRCAN, GEBCO, NOAA, INCREMENT P.

Table 2

Image dates and ID(s) of S2/MSI and L8/OLI scenes used in the study along with the allocated application.

Date	GEE Scene ID	Application	n
23-Aug-17	COPERNICUS/S2/20170823T170849_20170823T171828_T15	SD model calibration	79
12-Sep-17	COPERNICUS/S2/20170912T170949_20170912T171451_T15	SD model validation	276
27-Sep-17	COPERNICUS/S2/20170927T172111_20170927T172106_T15TVK LANDSAT/LC08/C01/T1/LC08_027029_20170927	Atmospheric correction assessment, SD model validation	18
13-Aug-18	COPERNICUS/S2/20180813T170851_20180813T172023_T15TVK LANDSAT/LC08/C01/T1/LC08_027029_20180813	SD model validation	43

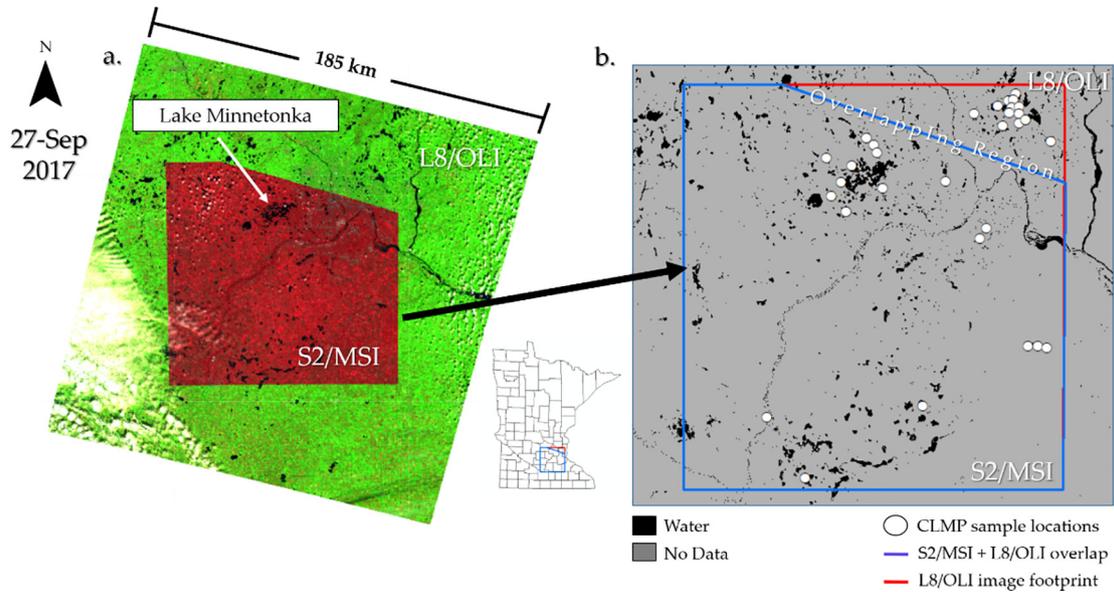


Fig. 2. False color composites of the coincident overpass imagery between L8-L1T Path 27, Row 29 (ID: LANDSAT/LC08/T1/LC08_027029_20170927) at 11:59 am CST (RGB: B7/B5/B2) and S2-L1C (ID: COPERNICUS/S2/20170927T172111) sub-track of T15TVK at 12:21 pm CST on 27-Sep 2017 (RGB: B8A/B4/B3) (a). Sampled locations by CLMP (white circles) outside the overlapping region between MSI and OLI (blue line) were excluded (b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

applied atmospheric correction (Section 2.3) to the collection of targeted MSI and OLI imagery for further analysis (Section 2.4).

2.3. Modified atmospheric correction for INland waters (MAIN)

We adopted atmospheric correction theory from traditional ocean-color techniques (Hu et al., 2000; Wang et al., 2009; Werdell et al., 2010; Dash et al., 2012; Vanhellemont and Ruddick, 2015), where the desired water-leaving reflectance (ρ_w) and subsequent remote-sensing reflectance (R_{rs}) (an apparent optical property) is derived (Gordon et al., 1983):

$$\rho_w(\lambda_i) = \rho_{rc}(\lambda_i) - \rho_{am}(\lambda_i)/t(\lambda_i) \quad (2)$$

$$R_{rs}(\lambda_i) = \rho_w/\pi \quad (3)$$

where $t(\lambda_i)$ is the diffuse transmittance from the water surface to the satellite (Hu et al., 2004). ρ_{rc} is the Rayleigh corrected reflectance (Gordon and Wang, 1994; Dash et al., 2012) and includes other correction factors including the Rayleigh scattering phase function (Doerffer, 1992) Fresnel correction (Gordon and Wang, 1994) as well as the ozone adjustment (Mishra et al., 2005; Dash et al., 2012). The ozone absorption coefficient for each spectral band was taken from the Aerosol Optical Depth Value-Added Product (Koontz et al., 2013) and the daily measured ozone concentration obtained from the merged products of Total Ozone Mapping Spectrometer (TOMS) Earth Probe, TOMS/Nimbus-7, TOMS/Meteor-3, and the Ozone Monitoring Instrument (OMI) available in the GEE repository (collection ID: TOMS/MERGED). Additionally, a digital elevation model (DEM) from the Shuttle Radar Topography Missions (SRTM, 30 m) (Farr et al., 2007) was used to calculate the Rayleigh optical thickness (Hansen and Travis, 1974) on a pixel-by-pixel basis.

The strong impact of the aerosol path reflectance (ρ_a) in the visible and NIR spectral range can be difficult to correct as complex scattering and absorbing properties of aerosols vary spectrally and with aerosol size, shape, chemistry and density (Vermeote et al., 2016). Previous research has demonstrated that observations of optically turbid water pixels within the Rayleigh-corrected shortwave infrared (SWIR) channels have comparable signal responses to that of clear water pixels (Wang et al., 2009; Werdell et al., 2010; Vanhellemont and Ruddick,

2015). Aerosol path radiance reflectance has been expressed as (Gordon and Wang, 1974):

$$\rho_a(\lambda_{NIR}) = ke^{(-c\lambda)} \quad (4)$$

where k and c are constants. Assuming negligible signal in the SWIR wavelengths even in the most optically complex waters (Vanhellemont and Ruddick, 2015), the two Rayleigh-corrected SWIR bands available on the MSI and OLI ($\rho_{rc}(\lambda_{SWIR-1,2})$) were used for aerosol determination rather than the NIR band, where the optically active constituents (OACs) in meso- to hyper-eutrophic inland lakes usually interfere with the NIR signal:

$$\rho_{rc}(\lambda_{SWIR-1}) = ke^{(-c\lambda)} = \rho_{rc}(\lambda_{SWIR-2}) \quad (5)$$

Aerosol type ϵ was then determined for each pixel as the negative of the slope of the straight line (Hu et al., 2000; Dash et al., 2012) between $\Delta\lambda_{SWIR-1,2}$ and $\Delta\ln(\rho_{rc}(\lambda_{SWIR-1,2}))$ as:

$$(\ln(\rho_{rc}(\lambda_{SWIR-2})) - \ln(\rho_{rc}(\lambda_{SWIR-1})))/(\lambda_{SWIR-2} - \lambda_{SWIR-1}) = -\epsilon \quad (6)$$

The output returns a raster image of ϵ which was extrapolated to the visible and NIR bands:

$$\rho_{am}(\lambda_{VIS-NIR}) = \rho_{rc}(\lambda_{SWIR-2}) \times (F_0'/F_0'(\lambda_{SWIR-2}))e^{(-\epsilon \times (\lambda_i/\lambda_{SWIR-2}))} \quad (7)$$

where F_0' is the instantaneous extraterrestrial solar irradiance adjusted for Earth-sun distance (Dash et al., 2012) and ρ_{am} is the aerosol reflectance map needed to quantify the remaining contributions from aerosols across each spectral band. Although the MSI and OLI sensors were not developed specifically for inland aquatic applications, the provided variables within the image metadata allow to fulfill the necessary equations to derive R_{rs} in the GEE environment.

2.4. Evaluation of MSI and OLI R_{rs} products

First, we assessed MAIN derived R_{rs} values using imagery acquired during a coincident overpass between MSI and OLI on 27-Sep 2017 (Table 2) over a region with a wide range of water clarity, the Lake Minnetonka area in east-central Minnesota (Fig. 2a). Within the overlapping region of the MSI and OLI footprint, Corresponding SD measurements from 18 optically variable inland lakes were sampled by participants at the Citizen Lake Monitoring Program (CLMP)

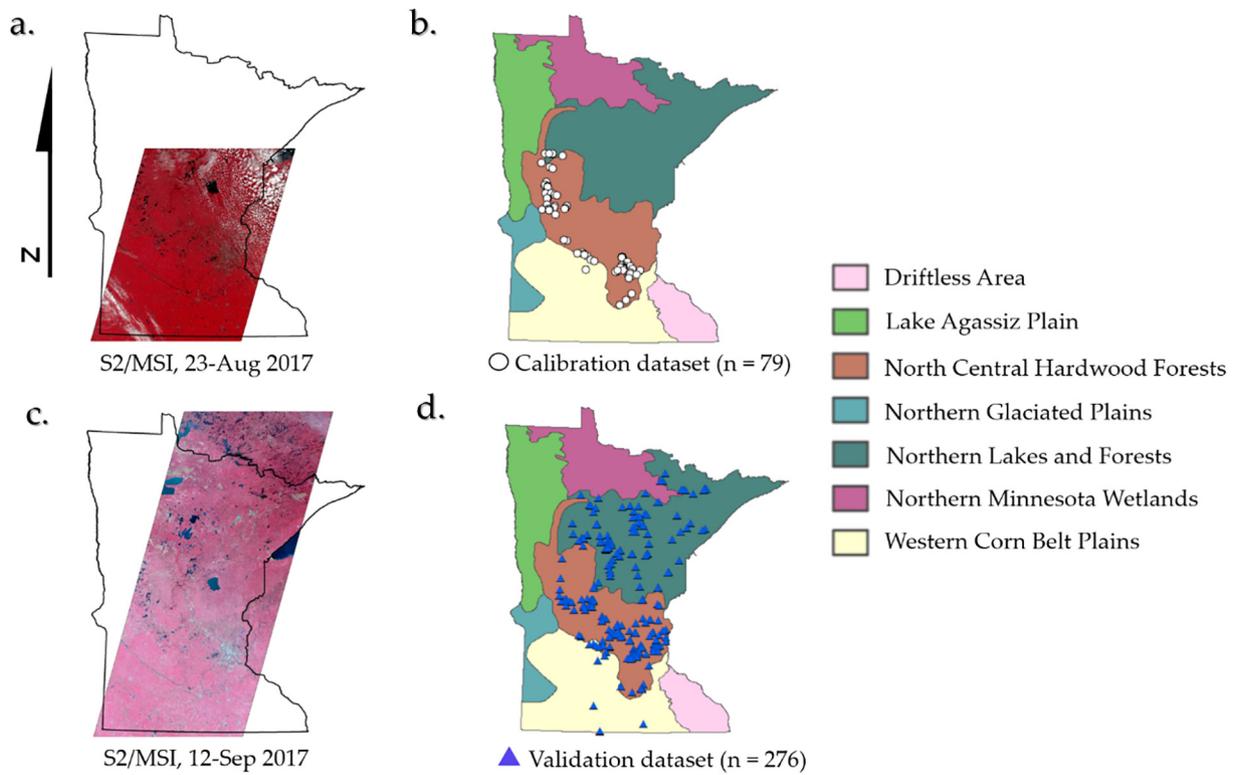


Fig. 3. False color MSI image composite (RGB: B8A/B4/B3) over Minnesota on 23-Aug 2017 (a) and on 12-Sep 2017 (c). Time-window qualified $SD_{in situ}$ sample locations collected by CLMP (± 1 day within satellite overpass) across the WCBP, NCHF, and NLF ecoregions are represented as white circles for the SD model calibration (b) and blue triangles for model validation dataset (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

within ± 1 day of the coincident overpass event (Fig. 2b). Any samples that fell beyond the overlapping footprints were excluded, because MSI imagery during 2017 was limited over the USA and for that date only covered a portion of southern Minnesota. Sampled point location data were uploaded to a GIS and transformed to a 50 m circular buffer around the centroid. Corresponding mean satellite ρ_{TOA} and R_{rs} values of pixel regions within the polygons were then extracted and tabulated. Prior to pixel extraction, MSI pixels were resampled to 30 m and registered to match the OLI georeferenced image prior to comparison (Storey et al., 2016).

In the absence of radiometric reference measurements such as buoy or other suitable matchup data for the Minnesota water bodies, several researchers have demonstrated that the inter-comparison of other atmospheric corrections represents an alternative option (Suresh et al., 2006; Vanhellemont and Ruddick, 2014; Bernardo et al., 2017). To this end, we evaluated MAIN derived MSI and OLI R_{rs} image products by evaluating the statistical closeness with R_{rs} values converted from the USGS Landsat-8 Surface Reflectance Product (OLI-SR), readily available in the GEE repository. Previous evaluations of MSI and OLI derived R_{rs} products have included the ocean color component of the Aerosol Robotic Network (AERONET-OC), inter-comparisons and cross-calibrations against other ocean color products, and optimizations of vicarious calibration gains (Pahlevan et al., 2017a, Pahlevan et al., 2017b, Ilori et al., 2019). In this study, lacking AERONET sites, OLI-SR spectra were defined as the closest representation of reference spectra considering previous successful reports on using OLI-SR with corresponding *in situ* water quality data for inland aquatic applications (Kuhn et al., 2019; Slonecker et al., 2016; Bernardo et al., 2017; Markert et al., 2018). We evaluated MAIN and ACOLITE derived R_{rs} values by calculating the root mean square difference (RMSD) across each multispectral band (λ_i) against the OLI-SR product:

$$RMSD(\lambda_i) = \frac{\sqrt{\sum_{i=1}^n (x_{Rrs}(\lambda_i) - x_{SR}(\lambda_i))^2}}{n} (\text{unit:sr} - 1) \quad (8)$$

where x_{Rrs} are the mean R_{rs} values from the sampled pixel regions using either the MAIN or ACOLITE method and x_{SR} is the OLI-SR reference spectra. Further, the same OLI and MSI coinciding overpass imagery were processed outside of GEE using ACOLITE (Vanhellemont and Ruddick, 2015) as a secondary comparison to provide a relative base on how well MAIN derived MSI and OLI R_{rs} values were performing over aquatic surfaces.

Next, we assessed the relative signal responses across comparable wavelengths (Table 1) using same MSI and OLI coincident imagery before and after atmospheric correction through mean absolute percent difference (MAPD%) in addition to the coefficient of determination (R^2) to evaluate cross-sensor consistency:

$$MAPD(\lambda_i) = \frac{\sum |x_{OLI(\lambda_i)} - x_{MSI(\lambda_i)}|}{x_{OLI(\lambda_i)}} \times 100 \quad (9)$$

where x_{MSI} and x_{OLI} are either the mean ρ_{TOA} or R_{rs} values for MSI and OLI at wavelength i , respectively.

2.5. Water clarity model calibration

SD is the most commonly measured water quality variable and has been shown to be strongly correlated with Landsat blue and red spectral bands (Kloiber et al., 2002a, 2002b; Olmanson et al., 2008). Corresponding blue and red (B2 and B4, respectively) MSI bands were hypothesized to contribute the most significance for SD model calibration purposes. Previous empirical methodologies for developing water clarity models involved direct stepwise linear regression between log-transformed $SD_{in situ}$ ($\ln SD_{in situ}$) and Landsat derived reflectances (Olmanson et al., 2001; Olmanson et al., 2013; Kloiber et al., 2002a, 2002b; Lillesand et al., 1983). However, the MSI sensor has three

additional red-edge bands (Table 1), one of which (B5, centered at ~705 nm) has been demonstrated to improve chlorophyll measurements (Gitelson et al., 2007; Gitelson et al., 2009; Mishra et al., 2013; Olmanson et al., 2015) thus providing the potential of yielding more reliable water clarity estimates. To explore the potential of these bands in predicting water clarity, we implemented a bootstrap forest technique within the JMP Pro 14 software (JMP®, Version 14. SAS Institute Inc., Cary, NC, 1989–2007) that informs the most significant bands and band-ratio combinations to model $\ln SD_{in situ}$.

The calibration dataset consisted of 79 $\ln SD_{in situ}$ measurements obtained by the Citizen Lake Monitoring Program (CLMP) collected within ± 1 day of the clear portions of the 23-Aug 2017 imagery (Fig. 3a-b) as the dependent variable and MAIN derived mean $MSI-R_{rs}$ values from bands B1-B8A (443–865 nm) and all band ratio permutations as independent input parameters (47 total terms). The bootstrap forest technique uses many decision trees to associate input terms with calibration/validation data, chosen in part randomly to determine the most significant terms that predict a response variable ($\ln SD_{in situ}$) based on the highest total sum of squares (SSTO) (Hastie et al., 2009). Prediction consistency of the bootstrap decision for each term was evaluated by splitting the samples into training (70%) and validation (30%) datasets and run for 10,000 iterations. Here, only two of the most contributing terms that produced highest coefficient of determination (R^2) with 79 $\ln SD_{in situ}$ measurements collected across the WCBP, NCHF, and NLF ecoregions of Minnesota were used for SD model development. We restricted to two-terms only so that they are consistent with water clarity models developed in the past (Kloiber 2002a, 2002b; Olmanson et al., 2008). Further, a second model was developed for MSI except the three red-edge (B5-B7) bands and one NIR band (B8) which are absent on the OLI sensor were excluded for consideration in order to establish a robust water clarity model for both platforms (SD_{sOLLI}). Overall model accuracy was then assessed on how well the calibrated SD_{MSI} (and SD_{sOLLI}) model forecasted $SD_{in situ}$ and was evaluated using the mean absolute error (MAE) as they are less sensitive to outliers (Seegers et al., 2018):

$$MAE = \frac{\sum_{i=1}^n |SD_{sensor} - SD_{in situ}|}{n} \quad (10)$$

where SD_{sensor} is either SD_{MSI} or SD_{sOLLI} , and a value of 0 is desired.

2.6. Water clarity model validation

For SD model validation, we first applied our calibrated SD_{MSI} model on an S2A/MSI image acquired on 12-Sep 2017 (Table 2) against 276 corresponding SD measurements collected by the CLMP ± 1 day of satellite overpass (Fig. 3c-d). Accuracy of all SD models were determined through MAE with corresponding $SD_{in situ}$ data. To evaluate the consistency of the SD_{MSI} and SD_{OLLI} models on a more temporal scale, two coinciding overpass images at different dates were used as secondary and tertiary SD model validation datasets. In addition to the 18 $SD_{in situ}$ measurements corresponding with the coincident MSI/OLI overpass imagery acquired on 27-Sep 2017 used to evaluate atmospheric correction (Fig. 2b), a second image pair acquired on 13-Aug 2018 over the same region (not displayed) was also included in the validation process and introduced another unique 43 $SD_{in situ}$ measurements (Table 2). It is important to note that the ranges of the secondary and tertiary 18 and 43 $SD_{in situ}$ measurements collected on 27-Sep 2017 (2^0) and 13-Aug 2018 (3^0) are 0.30–7.00 m and 0.45–6.70 m, respectively. These are comparable ranges used in the MAIN-derived SD_{MSI} (and SD_{sOLLI}) calibration (0.20–6.70 m) on 23-Aug 2017 and 1^0 validation (0.43–6.7 m) dataset on 12-Sep 2017 (Fig. 4), thus capturing the representational range of water clarity seen in Minnesota on a statewide scale in a single tile.

3. Results and discussion

3.1. Atmospheric correction

With the absence of *in situ* radiometric data in the Minnesota region for spectral comparison with MAIN derived $MSI-R_{rs}$ and $OLI-R_{rs}$ values, R_{rs} values converted from the USGS OLI-SR product ($R_{rs} = OLI-SR / \pi$) were used as reference (Section 2.4). Overall, both MAIN and ACOLITE methods yielded R_{rs} values comparable to the USGS OLI-SR product for both MSI and OLI from the 18 lakes sampled for atmospheric correction (Fig. 5).

The higher RMSD in the $MSI-R_{rs}$ spectra relative to $OLI-R_{rs}$ was expected due to the differences in the signal-to-noise ratio (SNR) (Table 1) of the MSI sensor rather than the direct OLI to OLI-SR comparison. Interestingly, R_{rs} values in the blue, green and red bands (B2-B4) resulting from both MAIN and ACOLITE yielded the lowest RMSD across comparable bands relative the reference spectra. This is important as these bands have been shown to contribute the highest significance in estimating SD for sensors used in previous research (Olmanson et al., 2013; Olmanson et al., 2001; Kloiber, 2002), and were considered as target variables for developing water clarity models (Section 2.5). Regardless of the higher RMSD, the mean R_{rs} spectra derived from ACOLITE share similar shape and magnitude relative to both MAIN and OLI-SR datasets from the 18 sampled lakes (Fig. 6). This provided confidence in the MAIN R_{rs} products considering the reasonable matchup with ACOLITE over aquatic surfaces and especially the OLI-SR product which was not originally intended for aquatic applications.

We also analyzed the relative signal response between MAIN and ACOLITE derived MSI and OLI reflectance values before and after atmospheric correction to evaluate cross-sensor performance. First, $MSI-R_{rs}$ and $OLI-R_{rs}$ values produced from MAIN outlined in Section 2.3 most notably generated positive (non-negative) values across all comparable spectral bands (Fig. 7), where negatives values have been commonly observed in previous studies due to inconsistent atmospheric correction in optically variable regions, resulting in masked pixels values with no data (Werdell et al., 2010; Bailey et al., 2010; Dash et al., 2012).

The mean absolute percent difference (MAPD%) in ρTOA values between MSI and OLI prior to correction were all $< 10\%$ in the visible bands but $\sim 14\%$ in the NIR band (Fig. 8). $\rho TOA R^2$ values were slightly lower between the MSI and OLI coastal band ($R^2 = 0.61$) and even less in the NIR band ($R^2 = 0.36$), however, the signal response between MSI and OLI ρTOA blue, green and red bands exhibit nearly synchronized readings, with $R^2 = 0.92, 0.98,$ and 0.96 respectively (Fig. 9a). Correlation differences observed in the coastal band were likely due to downscaling from the native 60 m spatial resolution to match OLI at 30 m (Mandanici and Bitelli, 2016), while the NIR band could again be attributed to the SNR difference between the two sensors or other inherent signal characteristics (Pahlevan et al., 2017b) (Table 1). After atmospheric corrections, band-by-band PD between MSI and OLI R_{rs} values mostly remained low as seen in the ρTOA comparison (Fig. 8). After MAIN processing, band-by-band R^2 values were conserved (Fig. 9b), and a similar 1:1 signal response in the band-by-band comparison after ACOLITE processing (Fig. 9c) suggested the suitability of using MAIN for aquatic applications. Harmonized R_{rs} products between the MSI and OLI were seen again after MAIN processing using another coincident overpass date acquired on 13-Aug 2018 (Fig. 10), again showing highest MAPD% in the coastal (B1) and NIR (B5) bands. These results show that the conservation in the R^2 values of the relative signal response between MSI and OLI and the consistently low relative MAPD after MAIN processing allow for relatively consistent R_{rs} retrievals from both sensors on a temporal scale regardless of atmospheric effects. More importantly, it allows for the development of consistent water clarity (and other water quality) models across both MSI and OLI sensors so that both datasets can be incorporated in time series analysis. From here, efforts can be made by lake management practices to improve the

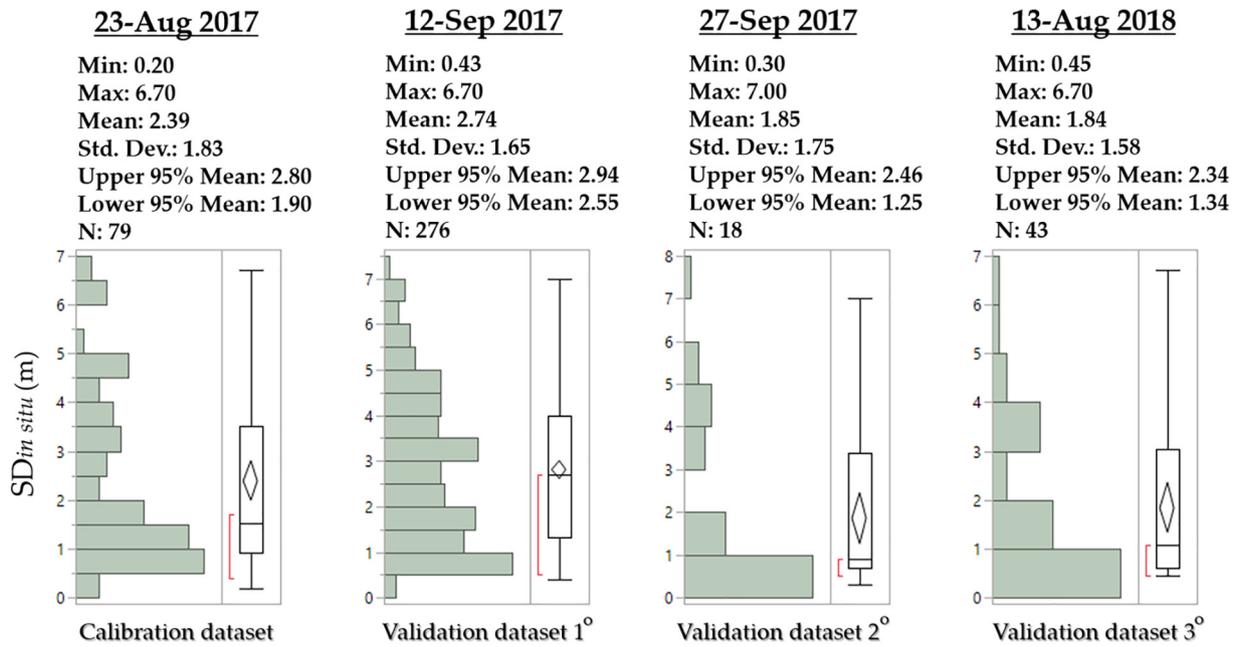


Fig. 4. Comparable statistics and corresponding histogram plots of *in situ* water clarity ($SD_{in\ situ}$) distribution from the calibration and three validation datasets.

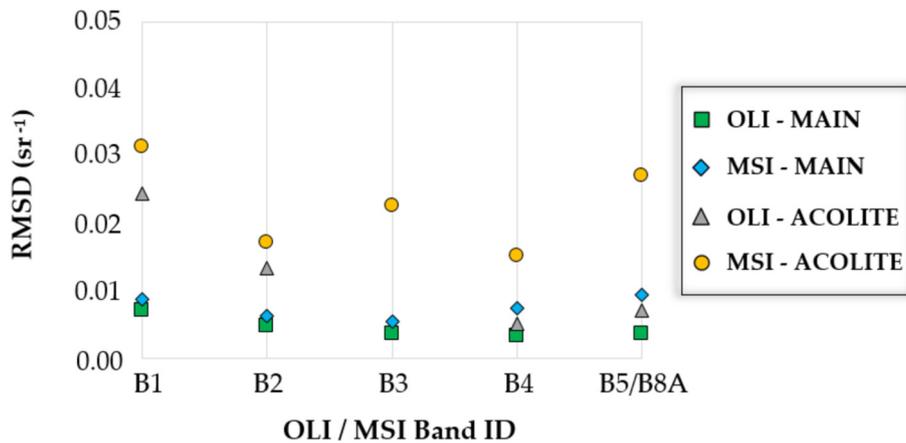


Fig. 5. Comparable RMSD (unit: sr^{-1}) values between MAIN and ACOLITE corrected MSI and OLI R_{rs} bands against the USGS OLI-SR reference spectra during a MSI/OLI coinciding overpass on 27-Sep 2017 ($n = 18$).

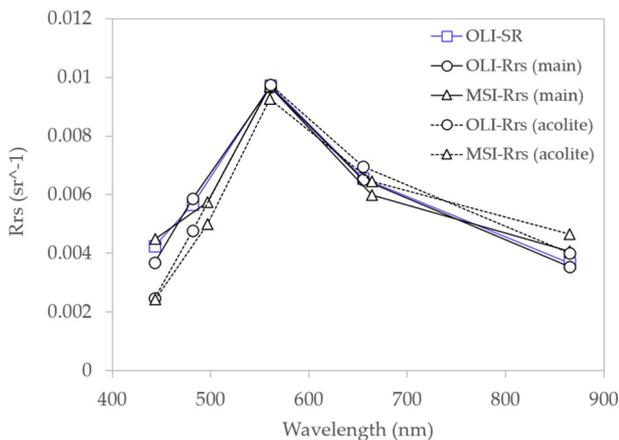


Fig. 6. Comparable MSI and OLI spectra derived from MAIN (solid line) and ACOLITE (dashed line) against the OLI-SR product (blue line) using the mean R_{rs} values from the 18 locations from the 27-Sep 2017 coincident overpass imagery. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

overall accuracy of the derived R_{rs} products for select study sites for continuous monitoring purposes. The overall accuracy of the generated MSI- R_{rs} and OLI- R_{rs} values would improve with frequent, routine *in situ* radiometric and water quality parameter measurements in a more systematic manner to fully understand the relationships between R_{rs} and the OACs for a particular optically variable water body.

3.2. Water clarity model calibration

Our next objective was to evaluate the applicability of the generated MSI and OLI R_{rs} products from MAIN in mapping water clarity (in terms of SD) (Section 2.5). Previous research on mapping water clarity in Minnesota used mainly a consistent blue/red, and blue 2-term model (Kloiber et al., 2002a, 2002b) calibrated through multiple linear regression analysis using Landsat satellite spectral data and corresponding $SD_{in\ situ}$ measurements (Olmanson et al., 2008; Olmanson et al., 2016). Here, we used a bootstrap forest method to obtain model-independent indicators of the most important predictors of water clarity using the 23-Aug 2017 calibration dataset (Section 2.5). The bootstrap forest technique was necessary in this case as the additional spectral bands of MSI (relative to OLI) could improve water clarity estimates. Of

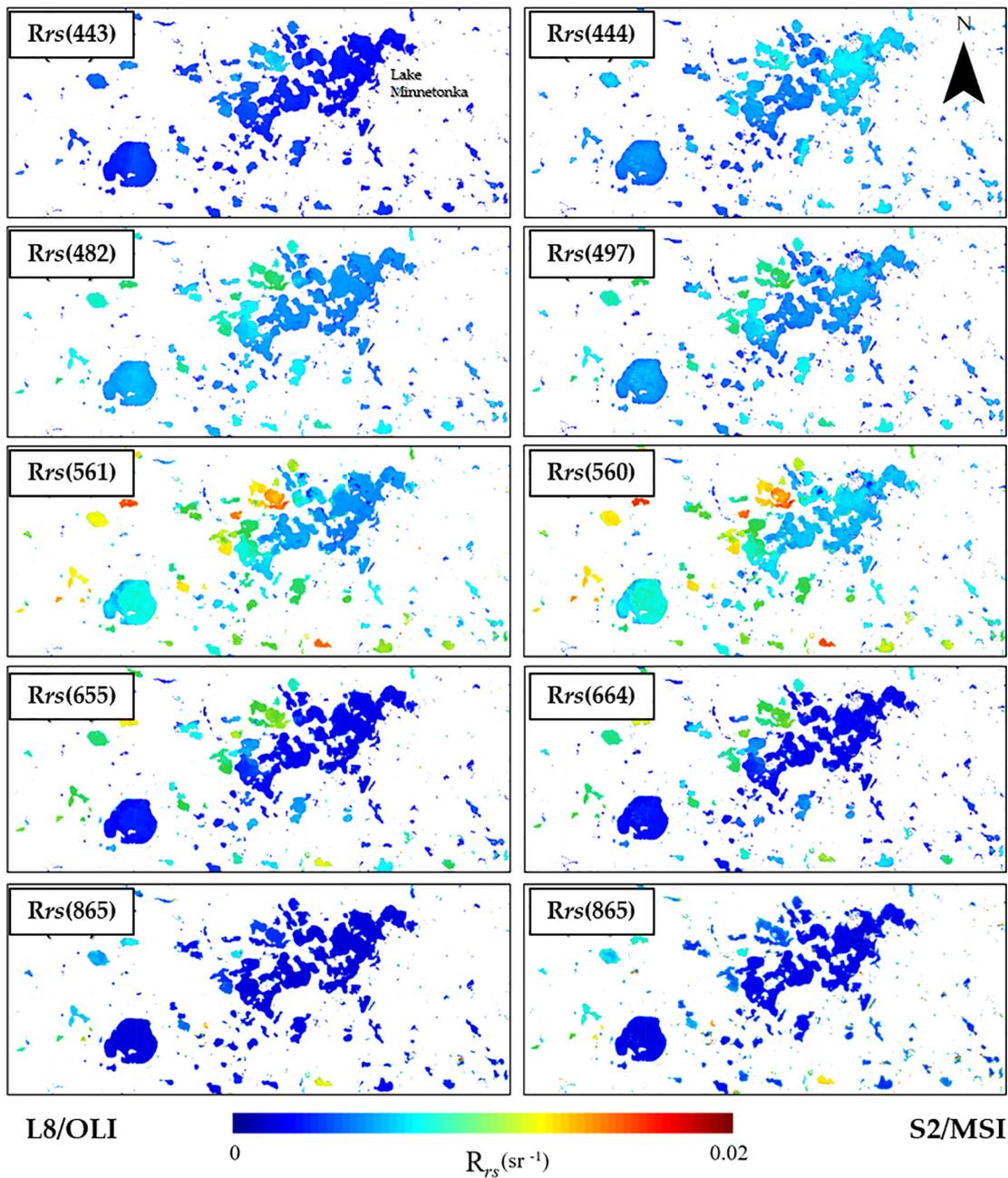


Fig. 7. Pseudocolor maps of comparable OLI- R_{rs} (left) and MSI- R_{rs} (right) products over Lake Minnetonka, MN and surrounding water bodies during the coinciding overpass on 27-Sep 2017.

the 47 terms tested as input predictors for $\ln SD_{in situ}$, the two terms yielding the highest sum of squares (SSTO) from the bootstrap forest were $MSI-R_{rs}(B2/B4)$ and $MSI-R_{rs}(B5 \times B4)$. The $MSI-R_{rs}(B2/B4)$ ratio was expected to be a contributing candidate in predicting $\ln SD_{in situ}$ as previous successes using prior Landsat satellites have been documented (Olmanson et al., 2008; Olmanson et al., 2016). These two terms resulting from the bootstrap forest generated the most favorable linear regression ($R^2 = 0.88$) with $\ln SD_{in situ}$ and took the form:

$$\ln SD_{MSI} = a_{MSI} (R_{rs}(B2)/R_{rs}(B4)) + b_{MSI} (R_{rs}(B5) \times R_{rs}(B4)) + c_{MSI} \quad (11)$$

where coefficients a_{MSI} (2.4367945), b_{MSI} (-2717.821), and c_{MSI} (-2.468818) were fit to the calibration data and $\ln SD_{MSI}$ is the log-transformation of the desired MSI derived SD (SD_{MSI}) for a given pixel (Fig. 11a). From these results it is clear that the $MSI-R_{rs}(B5)$ is a major contributor to water clarity prediction, a considerable advantage over

the OLI sensor. The capability of the $MSI-R_{rs}(B5)$ band to be used as a predictor for water clarity makes sense as chlorophyll-a exhibits strong reflectance in the 705–708 nm spectral region, and is often used in satellite-derived algal indices (Mishra et al., 2013; Augusto-Silva et al., 2014; Watanabe et al., 2015; Ogashawara et al., 2017).

An additional $\ln SD$ model was developed using the 23-Aug 2017 MSI imagery where only the comparable OLI bands were used as the independent variables ($\ln SD_{sOLI}$). Here, the $MSI-R_{rs}(B2)/R_{rs}(B4)$ ratio again resulted as the highest contributing candidate (in terms of SSTO) for predicting $\ln SD_{in situ}$, followed by the green band (B3). The $\ln SD_{sOLI}$ model using these two variables generated an R^2 of 0.85 with the *in situ* data in this case, slightly less but comparable to the $\ln SD_{MSI}$ relationship (Fig. 11b):

$$\ln SD_{sOLI} = a_{sOLI} (R_{rs}(B2)/R_{rs}(B4)) + b_{sOLI} (R_{rs}(B3)) + c_{sOLI} \quad (12)$$

where coefficients a_{sOLI} (2.6758384), b_{sOLI} (-29.49688), and c_{sOLI}

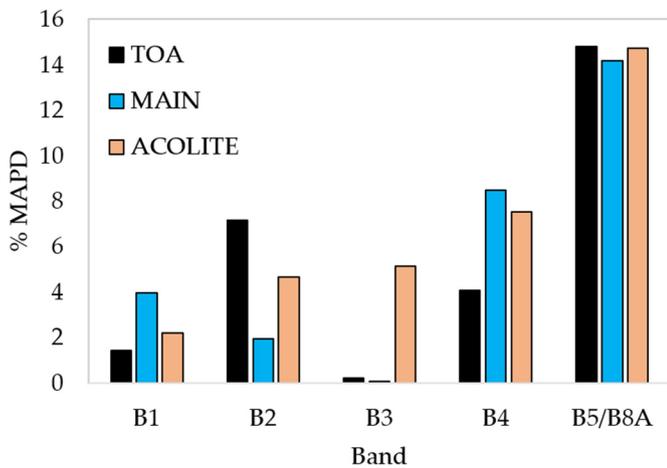


Fig. 8. Mean absolute percent difference (MAPD%) between MSI and OLI reflectance signals before atmospheric correction (TOA) and after MAIN and ACOLITE processing using the coincident 27-Sep 2017 imagery ($n = 18$).

(-2.468818) were fit to $\ln SD_{sOLI}$ data, and is mainly to be used for MSI in tandem with the OLI sensor after resampling the MSI pixels to match OLI 30 m spatial resolution for future cross-sensor comparisons. Finally, the converted SD_{MSI} and SD_{sOLI} values generated mean absolute errors

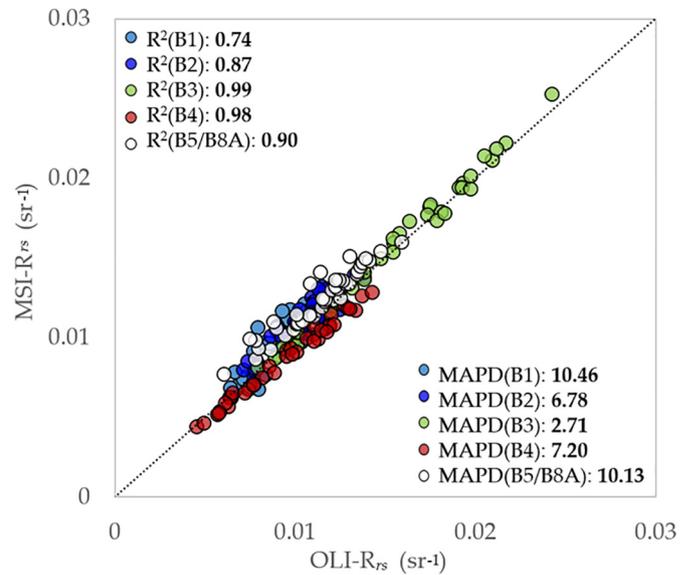


Fig. 10. Harmonized R_{rs} pixel values between MSI (y-axis) and OLI (x-axis) from the coincident overpass imagery acquired on 13-Aug 2018 ($n = 43$) after MAIN processing.

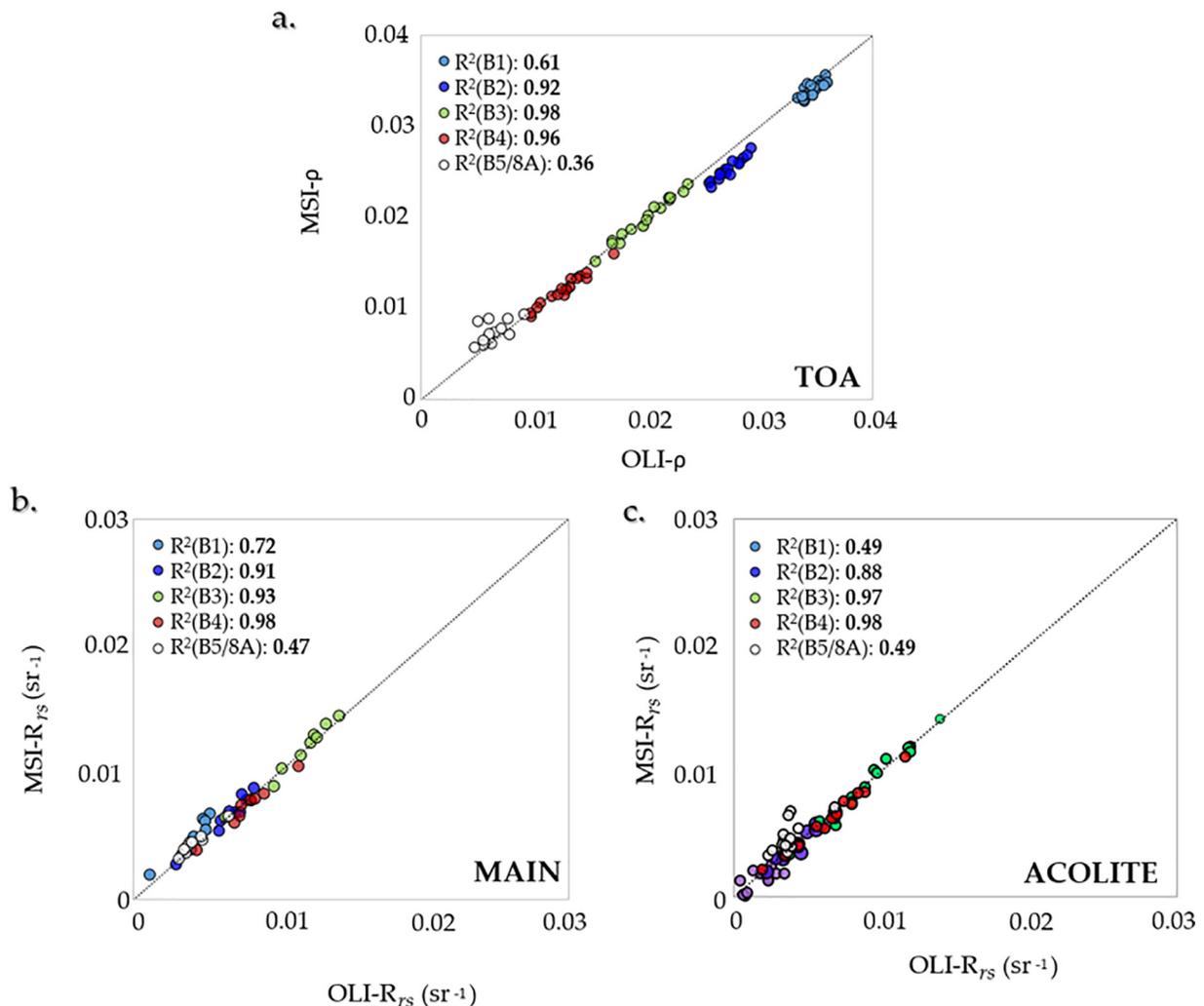


Fig. 9. Band-by-band relationships between comparable MSI (y-axis) and OLI (x-axis) spectral bands before atmospheric correction (TOA) (a) after MAIN (b) and ACOLITE (c) processing using the 27-Sep 2017 coincident overpass imagery ($n = 18$).

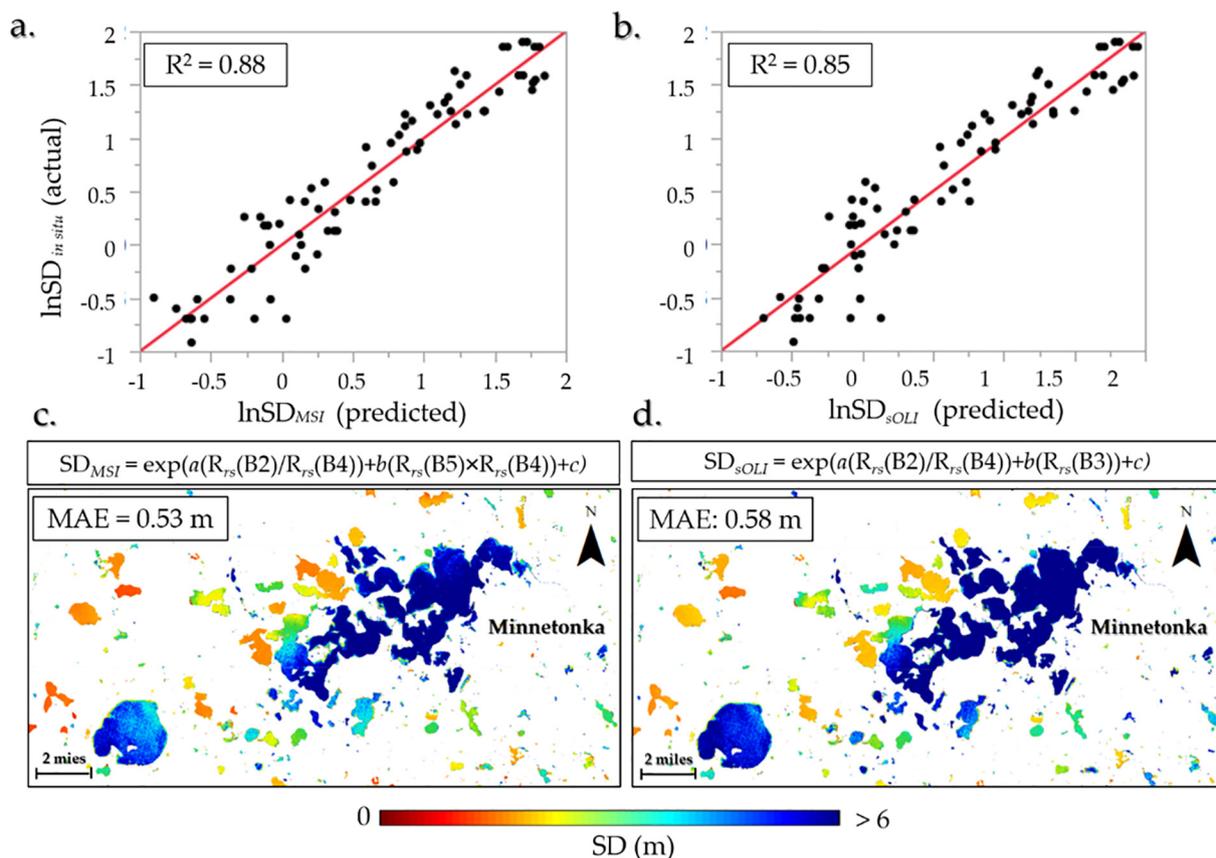


Fig. 11. Performance plots of predicted $\ln SD_{MSI}$ (a) and $\ln SD_{OLI}$ (b) models against the $\ln SD_{in situ}$ calibration dataset from 23-Aug 2017 ($n = 79$). Corresponding SD maps applied to the 23-Aug 2017 imagery in the GEE API (c).

(MAE) of 0.53 m and 0.58 m against the 23-Aug 2017 calibration dataset, respectively (Fig. 11c and d). This error (in terms of lake management) demonstrates the ability to derive a relatively accurate scene specific map of estimated SD from a MSI or OLI satellite image. For perspective, limnologists usually consider an SD < 2 m as indicative of eutrophic conditions, and SD < 1 m as indicative of hypereutrophy. Next, further validation datasets were used to evaluate whether this MAE varies across different dates of imagery as well as the deviation of the MAE across both sensors (Section 3.3).

3.3. Water clarity model validation

For SD model performance on a temporal scale, we compared three SD datasets corresponding with three dates of imagery: 276 $SD_{in situ}$ samples collected by the CLMP corresponding with ± 1 day within a S2A/MSI overpass on 12-Sep 2017 as a primary (1^0) validation dataset (considering the large sample size), 18 samples from the coincident overpass imagery on 27-Sep 2017 (2^0), and 43 samples from a tertiary (3^0) coincident overpass event between MSI and OLI on 13-Aug 2018 described in Section 2.6 (Table 2). The 1^0 validation dataset was mainly used for SD_{MSI} and SD_{OLI} model consistency for the MSI based models whereas the 2^0 and 3^0 datasets were used to evaluate cross sensor performance between SD_{OLI} and SD_{OLI} in addition to model validation.

The 1^0 validation dataset (12-Sep 2017) resulted in a MAE only 0.13 m greater than the calibration dataset between SD_{MSI} and $SD_{in situ}$ with an MAE of 0.66 m using Eq. (11) (Table 3). Similarly, a comparable MAE of 0.67 m resulted when using the SD_{OLI} model (Eq. (12)) (a 0.09 m increase from the calibration dataset). The slight increase of MAE using the 1^0 dataset may be a result of the increased sample size compared to the calibration dataset ($n = 276$) and may indicate a more robust representation of the model error. On the other hand, lower

Table 3

Mean absolute error (MAE) of satellite derived SD estimates from the three validation datasets.

Date	Model	MAE (m)	n
12-Sep-17	SD_{MSI}	0.66	276
	SD_{OLI}	0.67	
27-Sep-17	SD_{MSI}	0.25	18
	SD_{OLI}	0.33	
	SD_{OLI}	0.36	
13-Aug-18	SD_{MSI}	0.38	43
	SD_{OLI}	0.62	
	SD_{OLI}	0.44	

errors resulted when using the cross-sensor model (Eq. (12)) on the 2^0 and 3^0 validation datasets (Table 3). For the 2^0 validation dataset ($n = 18$), a consistent estimate of SD was generated for both sensors, with MAE of 0.33 m for SD_{OLI} and 0.36 m for SD_{OLI} , as expected due to the harmonized R_{rs} input bands. When applying Eq. (11) to the MSI imagery, a lower MAE of 0.25 m was generated, reassuring the improvement of water clarity estimation using the 705 nm band. A lower MAE of 0.38 m also resulted in the 3^0 validation dataset when using the SD_{MSI} model opposed to 0.62 m when using SD_{OLI} . In summary, the developed SD models had a MAE range of 0.25–0.67 m.

Satellite and field measurements can never exactly match. Any disagreements between the two could originate from many sources: difference in spatial coverage (20–30 m pixels vs. a single Secchi diameter), error in field measurements, error in the satellite atmospheric correction, and errors in the SD model. Maybe it is more important to look at the consistency of satellite-based values as if they were the standard way of measuring water clarity (e.g., in terms of consistency between sensors, etc.) rather than by comparing to ground-based SD

values as though they represent the “true values” of water clarity. The relative consistency of MAE values across all three comparisons (1°, 2°, 3°) may be saying something more important than the actual comparisons with the ground data.

4. Conclusions

This study aimed to provide a multi-sensor processing methodology for MSI and OLI imagery to map water clarity in Minnesota so that regional water quality assessments may be carried out in a fast, routine manner. The performance of MAIN was successfully demonstrated as a viable alternative to derive realistic R_{rs} values for lake water quality applications, and the advantage of GEE allows for quick and vigorous testing of the proposed methodologies for both atmospheric correction and water clarity model validation in either different geographical regions or scales.

Implementing the strategies demonstrated in this study in a high performance computing environment will allow the processes to be automated for generating near real-time water clarity maps for Minnesota's inland water bodies as soon as the image products become available for download. The cross-sensor image processing methodology in its current form could have a large impact on the routine monitoring protocols conducted by lake management and other resource agencies. For example, the exported water clarity maps from GEE can readily be implemented into a GIS or other web map service, and could help educate resource managers regarding areas susceptible to eutrophication either in person or online. Further, information derived from these maps could aid in characterizing the phenology of water clarity patterns on a regional scale. Being able to prioritize sampling efforts toward the more affected water bodies without extensive field sampling could become a capacity building exercise for a routine monitoring practice standard, and ultimately reduce time and financial constraints.

Acknowledgements

Support for this research was provided by the University of Minnesota Water Resources Center in cooperation with the U.S. Geological Survey Water Resources Research Act Program and the Minnesota Environment and Natural Resources Trust Fund (ENRTF) through the Legislative-Citizen Commission on Minnesota Resources (LCCMR). Partial support was provided by the National Science Foundation, United States (NSF) award no. 1442672. Sentinel-2 and Landsat-8 data were provided by the European Space Agency and USGS, respectively, acquired through the Google Earth Engine JavaScript API. We also thank Pamela Anderson Supervisor of the Water Quality Monitoring Unit, MPCA and the CLMP volunteers for the SD data.

References

Augusto-Silva, P., Ogashawara, I., Barbosa, C., De Carvalho, L., Jorge, D., Fornari, C., Stech, J., 2014. Analysis of MERIS reflectance algorithms for estimating chlorophyll-a concentration in a Brazilian reservoir. *Remote Sens.* 6, 11689–11707. <https://doi.org/10.3390/rs61211689>.

Bailey, S., Franz, B., Werdell, J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Opt. Express* 18, 7521–7527. <https://doi.org/10.1364/OE.18.997521>.

Bernardo, N., Fernanda, W., Rodrigues, T., Alcantara, E., 2017. Atmospheric correction issues for retrieving total suspended concentrations in inland waters using OLI/Landsat-8 image. *Adv. Space Res.* 59, 2335–2348. <https://doi.org/10.1016/j.asr.2017.02.017>.

Bossenbroek, J., Kraft, C., Nekola, J., 2001. Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecol. Appl.* 11, 1778–1788. [https://doi.org/10.1890/1051-0761\(2001\)011\[1778:POLDDUJ2.0.CO;2](https://doi.org/10.1890/1051-0761(2001)011[1778:POLDDUJ2.0.CO;2).

Breiman, L., 1996. Bagging predictors. *Mach. Learn.* 24 (2), 123–140 (f).

Brezonik, P., Bouchard Jr., R., Finlay, J., Griffin, C., Anderson, J., Arnold, W., Hozalski, R., 2019. Color, chlorophyll a, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. *Ecol. Appl.* 29 (3), e01871 (10/1002/eap.1871).

Carpenter, D., and Carpenter, S. 1983. Modeling inland water quality using Landsat data. *Remote Sens. Environ.* 13, 345–252. [https://doi.org/10.1016/0034-4257\(83\)90035-4](https://doi.org/10.1016/0034-4257(83)90035-4).

Claverie, M., Ju, J., Masek, J.G., Dungan, J., Vermote, E., Roger, J.-C., Skakun, S., Justice, C., 2018. The harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* 219, 145–161.

Dash, P., Walker, N., Mishra, D., Eurico, D., Ladner, S., 2012. Atmospheric correction and vicarious calibration of Oceansat-1 Ocean Color Monitor (OCM) data in coastal case 2 waters. *Remote Sens.* 4, 1716–1740. <https://doi.org/10.3390/rs4061716>.

Doerffer, R., 1992. Imaging Spectroscopy for Detection of Chlorophyll and Suspended Matter. GKSS-Forschungszentrum.

Farr, T., et al., 2007. The shuttle radar topography mission. *Rev. Geophys.* 45. <https://doi.org/10.1029/2005RG000183>.

Giardino, C., Pepe, M., Alessandro Brivio, P., Ghezzi, P., Zilioli, E., 2001. Detecting chlorophyll, secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *Sci. Total Environ.* 268, 19–29. [https://doi.org/10.1016/S0048-9697\(00\)00692-6](https://doi.org/10.1016/S0048-9697(00)00692-6).

Gitelson, A., Schalles, J., Hladik, C., 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sens. Environ.* 109, 464–472. <https://doi.org/10.1016/j.rse.2007.01.016>.

Gitelson, A., Gurlin, D., Moses, W., and Barrow, T., 2009. A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters. *Environ. Res. Lett.* 4 (2009) 045003 (5pp). <https://doi.org/10.1088/1748-9326/4/4/045003>.

Gordon, H., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Appl. Opt.* 33, 443–452. <https://doi.org/10.1364/AO.33.000443>.

Gordon, H., Clark, D., Brown, J., Brown, O., Evans, R., Broenkow, W., 1983. Phytoplankton pigment concentrations in the middle Atlantic bight: comparison of ship determinations and CZCS estimates. *Appl. Opt.* 22, 20–36. <https://doi.org/10.1364/AO.22.000020>.

Gordon, H., Du, T., Zhang, T., 1997. Atmospheric correction of ocean color sensors: analysis of the effects of residual instrument polarization sensitivity. *Appl. Opt.* 36, 6938–6948. <https://doi.org/10.1364/AO.36.006938>.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.

Hansen, J., Travis, D., 1974. Light scattering in planetary atmospheres. *Space Sci. Rev.* 16, 527–610. <https://doi.org/10.1007/BF00168069>.

Harrington, J., Schiebe, F., 1992. Remote sensing of Lake Chicot, Arkansas: monitoring suspended sediments, turbidity, and secchi depth with Landsat MSS data. *Remote Sens. Environ.* 39, 15–27. [https://doi.org/10.1016/0034-4257\(92\)90137-9](https://doi.org/10.1016/0034-4257(92)90137-9).

Hastie, T.J., Tibshirani, R.J., Friedman, J.H., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer-Verlag, New York.

Hu, C., Carder, K., Muller-Karger, F., 2000. Atmospheric correction of SeaWiFS imagery over turbid coastal waters: a practical method. *Remote Sens. Environ.* 7, 195–206. [https://doi.org/10.1016/S0034-4257\(00\)0080-8](https://doi.org/10.1016/S0034-4257(00)0080-8).

Hu, C., Chen, Z., Clayton, T.D., Swarzenski, P., Brock, J.C., Muller-Karger, F.E., 2004. Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: initial results from Tampa Bay, FL. *Remote Sens. Environ.* 93, 423–441. <https://doi.org/10.1016/j.rse.2004.08.007>.

Ilori, C.O., Pahlevan, N., Knudby, A., 2019... Analyzing Performances of Different Atmospheric Correction Techniques for Landsat 8: Application for Coastal Remote Sensing. *Remote Sens.* 11 (4), 469.

Kloiber, S., Brezonik, P., Olmanson, L., Bauer, M., 2002a. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sens. Environ.* 82, 38–47.

Kloiber, S., Brezonik, P., Bauer, M., 2002b. Application of Landsat imagery to regional-scale assessments of lake clarity. *Water Res.* 36, 4330–4340. [https://doi.org/10.1016/S0043-1354\(02\)00146-X](https://doi.org/10.1016/S0043-1354(02)00146-X).

Koontz, A., Flynn, C., Hodges, G., Michalsky, J., Barnard, J., 2013. *Aerosol Optical Depth Value-Added Product*. U.S. Department of Energy, pp. 32.

Kuhn, C., Valerio, A., Ward, N., Loken, L., Sawakuchi, H.O., Kampel, M., Richey, J., Stadler, P., Crawford, J., Striegl, R., Vermote, E., Pahlevan, N., Butman, D., 2019. Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity. *Remote Sens. Environ.* 224, 104–118. <https://doi.org/10.1016/j.rse.2019.01.023>.

Landsat 8 Data User's Manual. Version 2.0. <https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf>

Lee, Z., Shang, S., Hu, C., Du, K., Weidemann, A., Hou, W., Lin, J., Lin, G., 2016. Secchi disk depth: a new theory and mechanistic model for underwater visibility. *Remote Sens. Environ.* 169, 139–149. <https://doi.org/10.1016/j.rse.2015.08.002>.

Lillesand, T.M., 1983. Issues surrounding the commercialization of civil land remote sensing from space. *Photogramm. Eng. Remote Sens.*

Mandani, E., and Bitelli, G., 2016. Preliminary comparison of Sentinel-2 and Landsat-8 imagery for a combined use. *Remote Sens.*, 1014. <https://doi.org/10.3390/rs8121014>.

Markert, K., et al., 2018. Historical and operational monitoring of surface sediments in the Lower Mekong Basin using Landsat and Google Earth Engine cloud computing. *Remote Sens.* 10, 909.

Mishra, D., Narumalani, S., Rundquist, D., Lawson, M., 2005. Characterizing the vertical diffuse attenuation coefficient for downwelling irradiance in coastal waters: implications for water penetration by high resolution satellite data. *ISPRS J. Photogramm. Remote Sens.* 60, 48–64. <https://doi.org/10.1016/j.isprs.2005.09.003>.

Mishra, D., Mishra, R., Lee, Z., and Tucker, C., 2013. Quantifying cyanobacterial phycocyanin concentration in turbid productive waters: a quasi-analytical approach. *Remote Sens. Environ.* 113, 141–151. <https://doi.org/10.1016/j.rse.2013>.

- 02.004.
- Ogashawara, I., Mishra, D., Gitelson, A., 2017. Remote sensing of inland waters: Background and current state-of-the-art. In: Mishra, D.R., Ogashawara, I., Gitelson, A.A. (Eds.), *Bio-Optical Modeling and Remote Sensing of Inland Waters*, 1st edition. Elsevier, Waltham, MA, 9780128046449, pp. 332. <https://doi.org/10.1016/B978-0-12-804644-9.00001-X>.
- Olmanson, L., Kloiber, S., Bauer, M., Brezonik, P., 2001. *Image Processing Protocol for Regional Assessments of Lake Water Quality*. Water Resources Center and Remote Sensing Laboratory, University of Minnesota, St. Paul, MN.
- Olmanson, L., Bauer, M., Brezonik, P., 2008. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sens. Environ.* 112, 4086–4097. <https://doi.org/10.1016/j.rse.2007.12.013>.
- Olmanson, L., Brezonik, P., Bauer, M., 2013. Geospatial and temporal analysis of a 20-year record of Landsat-based water clarity in Minnesota's 10,000 lakes. *Journal of American Water Resources Association* 1–14. <https://doi.org/10.1111/jawr.12138>.
- Olmanson, L.G., Brezonik, P.L., Finlay, J.C., Bauer, M.E., 2016. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sens. Environ.* 185, 119–128.
- Olmanson, L., Brezonik, P., and Bauer, M., 2015. Remote sensing for regional lake water quality assessment: capabilities and limitations of current and upcoming satellite systems. *In Advances in Watershed Science and Assessment*, 111–140. [doi.org/https://doi.org/10.1007/978-3-319-14212-8_5](https://doi.org/10.1007/978-3-319-14212-8_5).
- Omernik, J., 1987. Ecoregions of the conterminous United States. *Ann. Assoc. Am. Geogr.* 77 (1), 118–125.
- O'Reilly, C., et. al., 2015. Rapid and highly variable warming of lake surface waters around the globe. *AGU Geophysical Research Letters* 42, 24. [doi:https://doi.org/10.1002/2015GL066235](https://doi.org/10.1002/2015GL066235).
- Page, B., Kumar, A., and Mishra, D., 2018. A novel cross-satellite based assessment of the spatio-temporal development of a cyanobacterial harmful algal bloom. *Int. J. Appl. Earth Obs. Geoinf.* 66, 68–81. [doi.org/https://doi.org/10.1016/j.jag.2017.11.003](https://doi.org/10.1016/j.jag.2017.11.003).
- Pahlevan, N., Schott, J., Franz, B.A., Zibordi, G., Markham, B., Bailey, S., Schaaf, C., Ondrusek, M., Greb, S., Strait, C., 2017a. Landsat 8 remote sensing reflectance (R_{rs}) products: evaluations, intercomparisons and enhancements. *Remote Sens. Environ.* 190 (1), 289–301. <https://doi.org/10.1016/j.rse.2016.12.030>.
- Pahlevan, N., Sarkar, S., Franz, B.A., Balasubramanian, S.V., and He, J., 2017b. Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: demonstrations and validations. *Remote Sens. Environ.* 201, 47–56. [doi.org/https://doi.org/10.1016/j.rse.2017.08.033](https://doi.org/10.1016/j.rse.2017.08.033).
- Pahlevan, N., Chittimalli, S., Balasubramanian, S., Vellucci, V., 2019. Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. *Remote Sens. Environ.* 220, 19–29. <https://doi.org/10.1016/j.rse.2018.10.027>.
- Seegers, B., Stumpf, R., Schaeffer, B., Loftin, K., Werdell, J., 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. *Opt. Express* 26 (6), 7404–7422. <https://doi.org/10.1364/OE.26007404>.
- Slonecker, E. T., Jones, D., and Pellerin, B., 2016. The new Landsat 8 potential for remote sensing of colored dissolved organic matter (CDOM). *Mar. Pollut. Bull.* 107, 2, 518–527. [doi.org/https://doi.org/10.1016/j.marpolbul.2016.02.076](https://doi.org/10.1016/j.marpolbul.2016.02.076).
- Smith, V., 2016. Effects of eutrophication on maximum algal biomass in lake and river ecosystems. *Inland Waters* 6, 147–154. doi.org/10.5268/IW-6.2.937.
- Stadelmann, T., Brezonik, P., and Kloiber, S., 2001. Seasonal patterns of chlorophyll-a and Secchi disk transparency in lakes of east-central Minnesota: implications for design of ground- and satellite-based monitoring programs. *Lake and Reservoir Management* 17, 299–314. [doi.org/https://doi.org/10.1080/07438140109354137](https://doi.org/10.1080/07438140109354137).
- Storey, J., Roy, D., Masek, J., Gascon, F., Dwyer, J., and Choate, M., 2016. A note on the temporary misregistration of Landsat-8 operational land imager (OLI) and Sentinel-2 multi spectral instrument (MSI) imagery. *Remote Sens. Environ.* 186, 121–122. [doi.org/https://doi.org/10.1016/j.rse.2016.08.025](https://doi.org/10.1016/j.rse.2016.08.025).
- Suresh, T., Desa, E., Mascarenhas, A., Matondkar, S., Naik, P., Nayak, S., 2006. Cross-calibration of IRS-P4 OCM satellite sensor. *Proc. SPIE* 6404, 1–9. <https://doi.org/10.1117/12.694019>.
- Tyler, A., Hunter, P., Spyarakos, E., Groom, S., Maria Constantinescu, A., and Kitchen, J., 2016. Developments in Earth observation for the assessment and monitoring of inland, transitional, coastal and shelf-sea waters. *Sci. Total Environ.* 572, 1307–1321. [doi.org/https://doi.org/10.1016/j.scitotenv.2016.01.020](https://doi.org/10.1016/j.scitotenv.2016.01.020).
- Vanhellemont, Q., and Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. *Remote Sens. Environ.* 145, 105–115. [doi.org/https://doi.org/10.1016/j.rse.2014.01.009](https://doi.org/10.1016/j.rse.2014.01.009).
- Vanhellemont, Q., and Ruddick, K., 2015. Advantages of high quality SWIR bands for ocean colour processing: examples from Landsat-8. *Remote Sens. Environ.* 161, 89–106. [doi.org/https://doi.org/10.1016/j.rse.2015.02.007](https://doi.org/10.1016/j.rse.2015.02.007).
- Vanhellemont, Q., Ruddick, K., 2016. Acolite for Sentinel-2: Aquatic applications of MSI imagery. In: *Proceedings of the 2016 ESA Living Planet Symposium*. ESA Special Publication, SP, pp. 740.
- Vermote, E., Justice, C., Claverie, M., and Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* 185, 46–56. [doi.org/https://doi.org/10.1016/j.rse.2016.04.008](https://doi.org/10.1016/j.rse.2016.04.008).
- Wang, M., Son, S., Shi, W., 2009. Evaluation of MODIS SWIR and NIR-SWIR atmospheric correction algorithms using SeaWiFS data. *Remote Sens. Environ.* 113, 635–644.
- Watanabe, F., Alcantara, E., Rodrigues, T., Imai, N., Barbosa, C., and da Silva Rotta, L., 2015. Estimation of chlorophyll-a-concentration and the trophic state of the Barra Bonita hydroelectric reservoir using OLI/Landsat-8 images. *Int. J. Environ. Res. Public Health*, 12, 10391–10417. [doi.org/https://doi.org/10.3390/ijerph120910391](https://doi.org/10.3390/ijerph120910391).
- Werdell, J., Franz, B., Bailey, S., 2010. Evaluation of shortwave infrared atmospheric correction for ocean color remote sensing of Chesapeake Bay. *Remote Sens. Environ.* 114, 2238–2247. <https://doi.org/10.1016/j.rse.2010.04.027>.



Limitations on using CDOM as a proxy for DOC in temperate lakes

Claire G. Griffin^{a,*}, Jacques C. Finlay^a, Patrick L. Brezonik^b, Leif Olmanson^c,
Raymond M. Hozalski^b

^a Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN, United States

^b Department of Civil, Environmental, and Geo-Engineering, University of Minnesota, Minneapolis, MN, United States

^c Remote Sensing and Geospatial Analysis Laboratory, Department of Forest Resources, University of Minnesota, Saint Paul, MN, United States

ARTICLE INFO

Article history:

Received 24 May 2018

Received in revised form

30 July 2018

Accepted 4 August 2018

Available online 6 August 2018

Keywords:

Colored dissolved organic matter

Dissolved organic carbon

Lakes

Aquatic carbon cycle

ABSTRACT

Colored dissolved organic matter (CDOM) has been widely studied as part of efforts to improve understanding of the aquatic carbon cycle, by laboratory, *in situ*, and remote sensing methods. We studied ecoregion-scale differences in CDOM and dissolved organic carbon (DOC) to understand variability in organic matter composition and the use of CDOM as a proxy for DOC. Data from 299 lakes across the U.S. Upper Midwest showed that CDOM, measured as absorptivity at 440 nm (a_{440}), correlated strongly with DOC ($R^2 = 0.81$, $n = 412$). Colored lakes in the Northern Lakes and Forests (NLF) ecoregion drove this relationship. Lakes in the North Central Hardwood Forests (NCHF) had low color (most had $a_{440} < 3 \text{ m}^{-1}$) and weaker CDOM-DOC relationships ($R^2 = 0.47$). Spectral slopes and specific ultraviolet absorbance (SUVA), indicated relatively low aromaticity and non-terrestrial DOM sources in low color lakes. Multiple regression analyses that included total dissolved nitrogen (TDN) and CDOM, but not chlorophyll *a*, improved DOC estimates in low color lakes, suggesting a dominant contribution of non-planktonic sources of low color DOM in these lakes. Our results show that CDOM is a reliable, regional proxy for DOC in lakes where forests and wetlands dominate the landscape and the DOM is primarily terrestrial in origin. Mapping of lake DOC at broad spatial scales by satellite-derived CDOM has lower accuracy in low color lakes.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Optical properties of dissolved organic matter (DOM) are widely used to estimate the composition and quantity of dissolved organic carbon (DOC) in aquatic systems (Helms et al., 2008; Massicotte et al., 2017). In particular, chromophoric (or colored) dissolved organic matter (CDOM), measured as light absorption at a specific wavelength (e.g., 350, 375, 420, or 440 nm), often correlates strongly with bulk DOC in a variety of freshwater and estuarine systems (Osburn and Stedmon, 2011; Spencer et al., 2012). Thus, CDOM has been used as an inexpensive and convenient proxy for DOC in applications such as satellite remote sensing across broad spatial scales (Del Castillo and Miller, 2008; Griffin et al., 2018; Kutser et al., 2015; Olmanson et al., 2016) and *in situ* sensors in streams (Pellerin et al., 2012; Sobczak and Raymond, 2015).

DOM originates from a variety of sources, including natural

autochthonous and allochthonous sources, as well as anthropogenic sources such as wastewater effluent and agricultural and urban runoff. CDOM largely derives from leachate of decayed terrestrial and aquatic vegetation; in waters with relatively low CDOM levels autochthonous production of organic matter by phytoplankton and aquatic macrophytes is an important CDOM source. DOM consists of many diverse molecules (Sipler et al., 2017), and aromatic compounds are important components of CDOM (Hur et al., 2009; Yang et al., 2015), which is strongly related to the fraction of hydrophobic organic acids. By absorbing light in ultraviolet and visible wavelengths, CDOM plays a major role in lake functioning by reducing the depth at which photosynthesis can occur (Karlsson et al., 2015; Koizumi et al., 2018; Thrane et al., 2014), regulating temperature (Caplanne and Laurion, 2008; Houser, 2006), and shielding organisms from harmful UV irradiation (Leavitt et al., 2003; Sommaruga, 2001). Additionally, CDOM and associated humic substances can negatively affect water treatment processes and treated water quality, including increasing coagulant demand, fouling filtration membranes, and reacting with chlorine to form potentially toxic disinfection byproducts (Edzwald

* Corresponding author.

E-mail address: griff356@umn.edu (C.G. Griffin).

and Tobiason, 1999; Lee et al., 2004; Liang and Singer, 2003; Stevens et al., 1976). CDOM has further been associated with mobilization of metals (McKnight and Bencala, 1990) and hydrophobic organic compounds. Modification of landscapes, such as draining of wetlands for agriculture and urban land uses, has strong effects on the chemical composition of DOM and CDOM in rivers and lakes (Schindler, 2009; Wilson and Xenopoulos, 2009). In light of rapidly changing climate and land use in many temperate watersheds, an improved grasp of CDOM and DOC dynamics in lakes is essential to understanding whole-ecosystem functioning.

Over the past ~15 years, many studies have demonstrated that satellite remote sensing can be used to conduct synoptic “sampling” of CDOM across broad spatial scales, in both freshwater (e.g., Zhu et al., 2014; Olmanson et al., 2016b; Chen et al., 2017; Spyarakos et al., 2017) and marine ecosystems (e.g., Del Castillo and Miller, 2008; Xie et al., 2012; Fichot et al., 2016). Regression-based models relating laboratory-measured CDOM to satellite reflectance can result in large datasets that track CDOM through time and space (Kutser, 2012). These empirical models rely upon relationships between reflectance in the visible or near infrared spectrum to absorption at UV or blue wavelengths, which are influenced by both the quantity and composition of DOM. As ongoing work in our lab shows, remote sensing CDOM models generally must be calibrated on regional to local scales, as differences in atmospheric corrections and water quality parameters (such as suspended solids mineralogy, particle size, and algal communities) make it difficult to transfer models between regions (Brezonik et al., 2015; Griffin et al., 2018). Although CDOM itself is an inherently important variable due to its effects on light and photochemistry (Karlsson et al., 2009), CDOM is widely used to estimate DOC for study of freshwater carbon cycling (Chen et al., 2017; Spencer et al., 2012; Tehrani et al., 2013). In many riverine systems, estimation of DOC from CDOM works well (Stedmon et al., 2011; Yamashita et al., 2011), although important exceptions occur especially in large, impounded rivers (e.g., the St. Lawrence, Columbia, and Colorado Rivers; Spencer et al., 2012). Lakes and lake-influenced rivers may have weaker CDOM-DOC relationships owing to production of low-color, autochthonous DOM from phytoplankton and macrophytes. As well, inputs of low-color DOM from wastewater effluent or agricultural run-off to surface waters may shift the CDOM-DOC relationship from that found in forested watersheds.

Here, we examine ecoregional variability in CDOM properties using a multi-year, multi-season dataset of lake samples across two north temperate ecoregions in the U.S. Upper Midwest: the North Central Hardwood Forests (NCHF) and Northern Lakes and Forests (NLF). We aimed to determine general conditions in which CDOM may be used as a reliable proxy for DOC on broad, regional scales. Our results indicate that CDOM and DOC are strongly related in highly colored lakes and the NLF generally, but DOC in low color lakes cannot be predicted reliably from CDOM alone.

2. Materials and methods

2.1. Water sample collection and processing

We collected 412 samples from 2014 to 2016 from 299 lake sites in the NLF and NCHF, which span the Upper Midwest states of Minnesota, Wisconsin, and Michigan (USA; Fig. 1). The NLF is heavily forested (49%), with many wetlands (27%) and areas of open water (5%) as calculated from the 2011 National Land Cover Data set (Homer et al., 2015). Developed land accounts for 4% of the NLF, and agriculture covers another 7% of the land surface. In contrast, 48% of the NCHF is used for agricultural purposes and 9% is urban, while wetlands only account for 10% of land cover and forests account for 26%. Many lakes were sampled repeatedly, 2–6 times, over the

course of three years. Water was collected from ~0.25 m below the lake surface using acid-washed and triple-rinsed polycarbonate or high-density polyethylene (HDPE) bottles. The bulk of the lake water samples (301 out of 412) were collected by University of Minnesota personnel with the remainder collected by partner organizations. Samples for chlorophyll and dissolved constituents were filtered within 24 h of collection. The dissolved fraction was obtained by filtering the water through 0.45 μm Geotech High Capacity capsule filters. Samples for DOC and total dissolved nitrogen (TDN) were acidified using 2 M HCl and stored in pre-ashed 20 mL glass bottles at 4 °C. Filtered water for measurement of CDOM absorbance spectra was stored in pre-ashed 40 mL amber glass bottles, with no headspace. Samples for dissolved inorganic carbon (DIC) were filtered and stored in pre-ashed 20 mL glass bottles with no headspace. Chlorophyll-*a* was isolated from the water by vacuum filtering onto 0.22 μm cellulose nitrate filters and stored frozen until analysis. For the sampling done by partner organizations, water samples were filtered in the field with syringe-mounted Whatman GF/F filters, and DOC and TDN samples were placed in HDPE bottles and then frozen for storage. Other dissolved constituents were stored in the same manner as the UMN samples as described above. Samples collected by partner organizations were shipped overnight to the University of Minnesota for subsequent analyses. Chlorophyll-*a* was not collected by the partner organizations.

2.2. Sample analysis

DOC and TDN were measured on a Shimadzu TOC L CSN analyzer, after acidification and sparging. DIC was calculated as the difference in carbon between DOC and non-acidified water samples analyzed on the Shimadzu. UV-visible light absorbance of filtered water samples ($\lambda = 200\text{--}800\text{ nm}$) was measured using a Shimadzu 1601UV-PC dual beam spectrophotometer through 1 or 5 cm quartz cuvettes, against a nanopure water blank. Absorbance values were converted to absorptivity (absorption coefficients) using:

$$a(\lambda) = 2.303A(\lambda) / l \quad (1)$$

where a is the absorption coefficient at a given wavelength (λ), A is absorbance at wavelength λ , and l is the cell path length (m). Absorbance scans were blank-corrected before conversion. Specific UV absorbance (SUVA) was calculated by dividing UV absorbance (A) at 254 nm by DOC concentration (in mg/L), after correcting for the cell path length. Spectral slope parameters (S) were calculated using a nonlinear fit of an exponential function to absorption in the ranges 275–295 nm, 350–400 nm, and 400–460 nm, as in the following equation:

$$a(\lambda) = a(\lambda_{ref})e^{-S(\lambda - \lambda_{ref})} \quad (2)$$

where a is the absorption coefficient at a given wavelength (λ), λ_{ref} is a reference wavelength, and S is the slope fitting parameter. S_r was calculated as the ratio of $S_{275-295}$ to $S_{350-400}$. Chlorophyll-*a* (chl-*a*) was measured using standard fluorometric methods after 90% acetone extraction. A subset of samples was analyzed for nitrate/nitrite ($\text{NO}_3^-/\text{NO}_2^-$) and ammonium using a Lachat Quickchem FIA (Hach Company) with a detection limit of 10 μg/L ($n = 36$).

2.3. Data analysis

Data analyses were performed in R using base code and packages ‘caret’, ‘segmented’, and ‘Metrics’. DOC and a_{440} failed Shapiro-Wilks normality tests, and were thus natural log-transformed for regression analyses. Figures were made in R using ‘ggplot2’ and ‘gridExtra’, except for Fig. 1 which was created in ArcGIS v10.5.1.

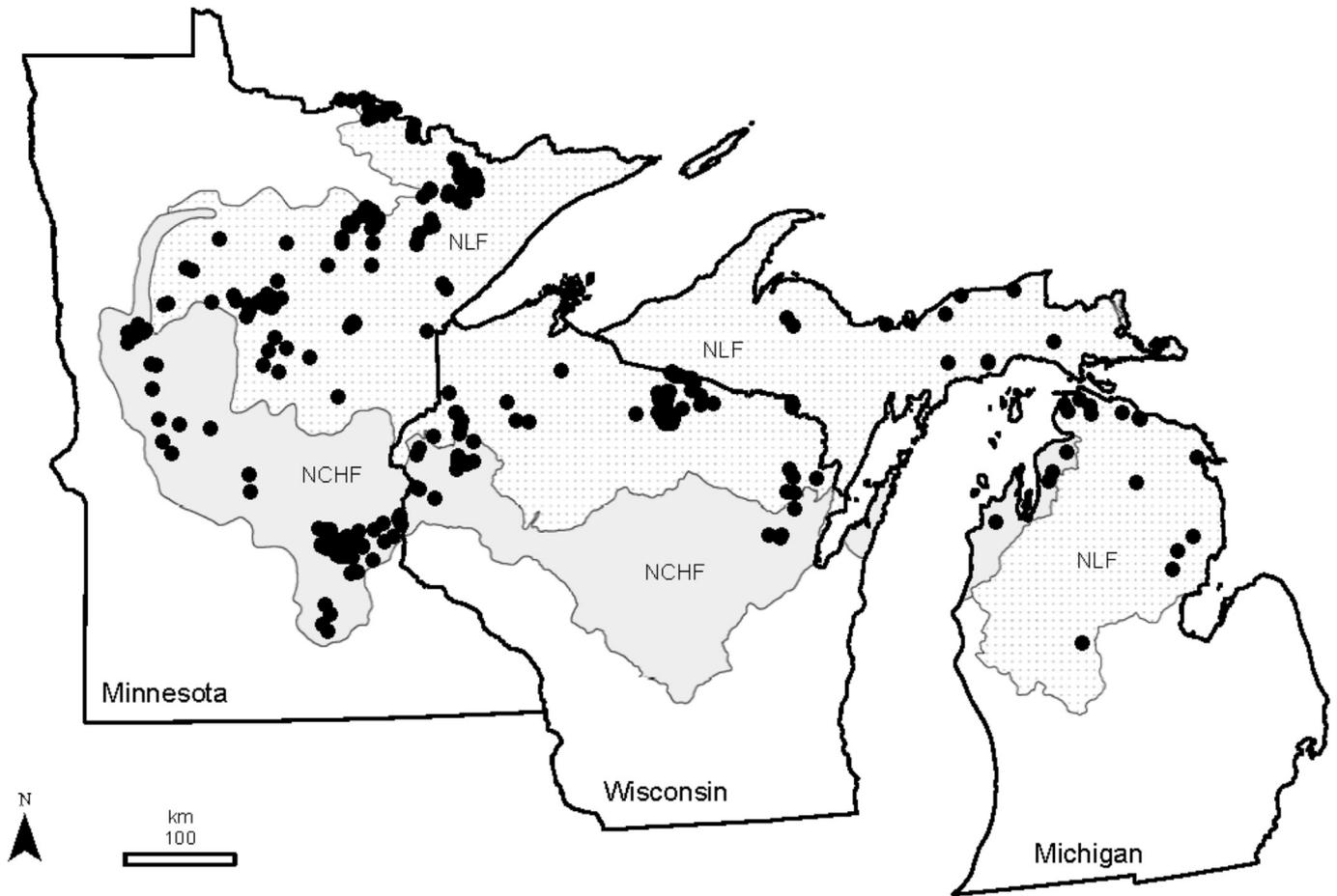


Fig. 1. Sampling sites across three Upper Midwest states from 2014 to 2016. The Northern Lakes and Forests (NLF) is represented by dotted fill, and NCHF by solid grey fill, with sampling locations indicated by filled circles.

3. Results

Lakes in the NLF generally had higher CDOM values and greater variability in CDOM quantity and compositional parameters than NCHF lakes (Table 1). For example, the mean (\pm standard deviation) a_{440} for NLF lakes was $6.05 \pm 7.03 \text{ m}^{-1}$ ($n = 313$) compared with $1.45 \pm 1.08 \text{ m}^{-1}$ ($n = 107$) for NCHF lakes. Similarly, SUVA, an indicator of aromaticity, was higher and more variable in the NLF lakes ($3.09 \pm 1.35 \text{ L mg C}^{-1} \text{ m}^{-1}$) than in the NCHF lakes ($2.09 \pm 0.69 \text{ L mg C}^{-1} \text{ m}^{-1}$). NCHF lakes generally contained higher chl-*a* concentrations (Table 1) and a greater range of TDN concentrations than NLF lakes. These results summarize all the individual measurements,

including those for 58 lakes sampled multiple times across years and/or seasons.

Log-transformed CDOM (a_{440}) was strongly and linearly related to log-transformed DOC across the entire dataset ($R^2 = 0.81$; Table 2; Fig. 2):

$$\ln(\text{DOC}) = 1.946 + 0.388 * \ln a_{440} \quad (3)$$

High variability was observed in low color waters ($a_{440} < 3 \text{ m}^{-1}$). An a_{440} of 3 m^{-1} was used to divide waters into “low” and “high” colored groups based on visual inspection of the DOC- a_{440} relationship; this threshold is approximately where waters become

Table 1
Summary statistics of variation in a_{440} (m^{-1}), DOC (mg/L), SUVA (mg/L m^{-1}), S_r by ecoregion.

	a_{440}	DOC	SUVA	$S_{275-295}$	S_r	Chl- <i>a</i>	TDN
NCHF							
mean (st dev)	1.45 (1.08)	8.07 (2.73)	2.09 (0.69)	23.23 (4.17)	1.251 (0.24)	16.50 (21.26)	0.61 (0.29)
median							
min	0.09	3.07	0.79	33.51	0.68	1.22	0.08
max	5.30	17.79	4.42	14.59	1.79	98.71	2.23
n	107	108	105	107	107	78	96
NLF							
mean (st dev)	6.05 (7.23)	12.18 (7.88)	3.09 (1.35)	19.40 (5.24)	1.08 (0.31)	4.74 (3.55)	0.49 (0.11)
median							
min	0	2.46	0.39	39.81	0.511	0.02	0.50
max	32.47	36.15	5.84	12.57	2.58	24.89	0.95
n	313	309	305	309	309	194	274

Table 2

Regression equations to predict log-transformed DOC from log-transformed a_{440} for lakes in three Upper Midwest states. Regressions using un-transformed, raw CDOM and DOC values are found in Table S1. All regressions were highly significant ($p < 0.0001$).

	n	Intercept	Slope	R^2	RMSE (mg/L)
ALL	412	1.95	0.39	0.81	2.52
<3 m ⁻¹	257	1.95	0.29	0.44	1.82
>3 m ⁻¹	154	1.56	0.57	0.9	2.00
NLF	307	1.91	0.40	0.85	2.54
NCHF	105	2.01	0.30	0.47	2.00
NCHF, <3 m ⁻¹	95	2.01	0.30	0.42	5.67
NLF, <3 m ⁻¹	162	1.86	0.27	0.45	5.22
NLF, >3 m ⁻¹	145	1.55	0.57	0.91	12.32

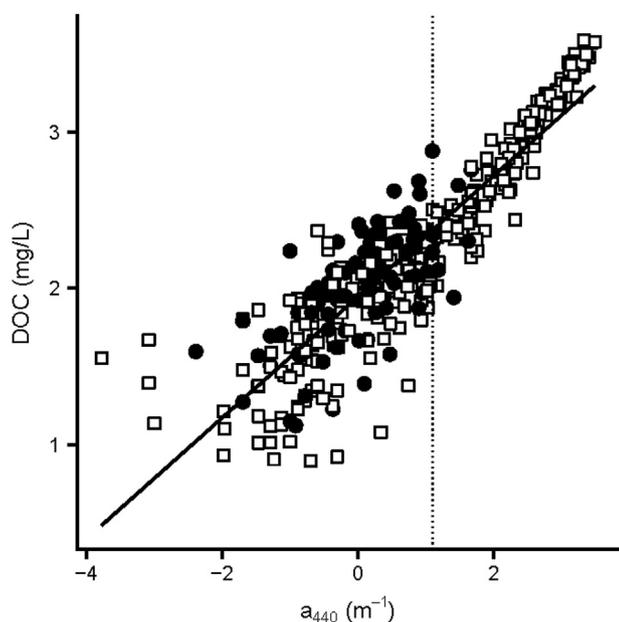


Fig. 2. DOC vs. CDOM for NCHF (filled circles) and NLF (open squares) ecoregions in MN, WI, and MI. Solid line indicates the regression line presented in equation (3). The dotted line is at $a_{440} = 3 \text{ m}^{-1}$, indicating the separation between low and high colored lakes.

visibly brown or tea colored to the human eye. Although a piecewise regression analysis did not yield a significant “breakpoint” where the slope of the DOC- a_{440} relationship changed, the two groups ($a_{440} < 3$ and $>3 \text{ m}^{-1}$) nonetheless showed strikingly different relationships between DOC and a_{440} . In addition, similar piecewise regression analyses of $S_{275-295}$ and SUVA vs. a_{440} showed significant inflection points at 2.67 and 3.44 m^{-1} , respectively. When only lakes with a_{440} exceeding 3 m^{-1} were considered, R^2 increased to 0.90 and the slope increased from 0.388 in Equation (3) to 0.569. In contrast, for $a_{440} < 3 \text{ m}^{-1}$, DOC and a_{440} were only moderately correlated ($R^2 = 0.44$) with a lower slope (0.285) than either of the above regressions (Table 2; Fig. 2). NCHF lakes were predominantly low in color, with 95 samples having $a_{440} < 3 \text{ m}^{-1}$ and only 10 samples with $a_{440} > 3 \text{ m}^{-1}$ (maximum a_{440} of 5.3 m^{-1}). Further analysis of differences between ecoregions within low color samples showed differences in intercept and RMSE, but the slopes of DOC versus CDOM relationships were not significantly different between the two ecoregions (ANCOVA, $p = 0.4646$).

Additional optical characteristics, such as SUVA and spectral slopes, also showed differences in organic matter quality between low and high color lakes. Optical characteristics within the UV-B

and UV-C regions correlated more closely with a_{440} than with DOC (Fig. 3). SUVA was correlated more strongly with natural log-transformed a_{440} ($R^2 = 0.89$) than with log-transformed DOC ($R^2 = 0.67$). $S_{275-295}$ similarly showed a more robust relationship with log-transformed a_{440} ($R^2 = 0.85$) than with log-transformed DOC ($R^2 = 0.57$; Fig. 3). In contrast, spectral slopes for UV-A and visible wavelengths ($S_{350-400}$ and $S_{400-460}$) had low R^2 values for both log-transformed a_{440} and DOC, ranging from 0.03 to 0.14. S_p , the ratio of $S_{275-295}$ and $S_{350-400}$, moderately correlated with a_{440} and DOC ($R^2 = 0.62$ and 0.43, respectively), as might be expected for a parameter that relies on both UV-B and UV-A wavelengths (Fig. S1). Low color samples ranged widely in spectral slope values, particularly for slopes at longer wavelengths (Fig. 2; S1). Above $\sim 3 \text{ m}^{-1}$, however, most of these longer-wavelength optical characteristics were stable or showed small, positive increases.

We also explored whether other limnological properties, such as DIC, chl-*a*, and TDN, provided useful information on DOM properties. DIC did not correlate with CDOM or DOC, and was not a significant parameter in multiple regressions to predict DOC with other variables. Chl-*a* was not strongly related to CDOM, DOC, or SUVA when all samples were considered together, but separating the NLF and NCHF samples produced clear differences in chl-*a* patterns with the three parameters (Fig. 4). For NLF sites, chl-*a* did not correlate with CDOM, DOC, and SUVA ($R^2 = 0.05$, 0.05, 0.12, respectively), and chl-*a* concentrations changed little across the ranges of these organic matter metrics. Chl-*a* varied widely across NCHF sites, but regressions between organic matter spectral characteristics and chl-*a* remained weak. For example, a_{440} had the weakest relationship with chl-*a* for NCHF samples ($R^2 = 0.07$, $p = 0.014$), while that for DOC was slightly higher ($R^2 = 0.13$, $p = 0.001$).

Because DOC can be a product of both allochthonous (usually colored) and autochthonous (low-colored) DOM sources, we considered whether a multiple regression using both chl-*a* and CDOM would improve DOC predictions compared to Equation (3) (Table 3). For 264 samples where chl-*a*, a_{440} and DOC were all available, chl-*a* was only a weakly significant parameter in multiple regression ($p = 0.0339$). Compared to equation (3), including chlorophyll-*a* with CDOM to predict DOC increased R^2 marginally, from 0.81 to 0.84. In low colored lakes, chl-*a* was significant ($p = 0.0001$), but the R^2 did not change with its inclusion in a multiple regression. Chl-*a* was not a significant predictor of DOC in highly colored lakes ($p = 0.522$).

Nitrate and ammonium concentrations were low in the subset analyzed, with many samples falling below the detection limit of 10 $\mu\text{g/L}$ (72% and 69%, respectively). These results show that TDN was dominated by dissolved organic nitrogen, as expected during summer conditions in lake surface waters (Finlay et al., 2013). Overall concentrations of TDN were moderate to low (<1.5 mg/L). As with chl-*a*, the TDN- a_{440} relationship was different for NLF and NCHF sites. In the NLF, TDN increased with a_{440} (Fig. 4c) ($R^2 = 0.68$, slope = 0.0267, $p < 0.0001$). The a_{440} -TDN relationship explained less of the TDN variance in the NCHF and had a much higher slope (Fig. 4d) ($R^2 = 0.29$, slope = 0.151, $p < 0.0001$). When TDN and CDOM were used in a multiple regression analyses for DOC, R^2 improved to 0.92 for the whole data set, and TDN was a highly significant parameter ($p < 0.0001$; Table 3). When only sites with $a_{440} < 3 \text{ m}^{-1}$ were used, a large increase in R^2 (from 0.44 to 0.77) was found when TDN was included. Highly colored sites showed only slight improvement in accuracy with the inclusion of TDN, in part because of the already strong relationship between a_{440} and DOC.

Two-term and three-term models have also been used to predict DOC in freshwater and marine environments. A variety of parameters have been used including absorbance at 254, 270 and/or

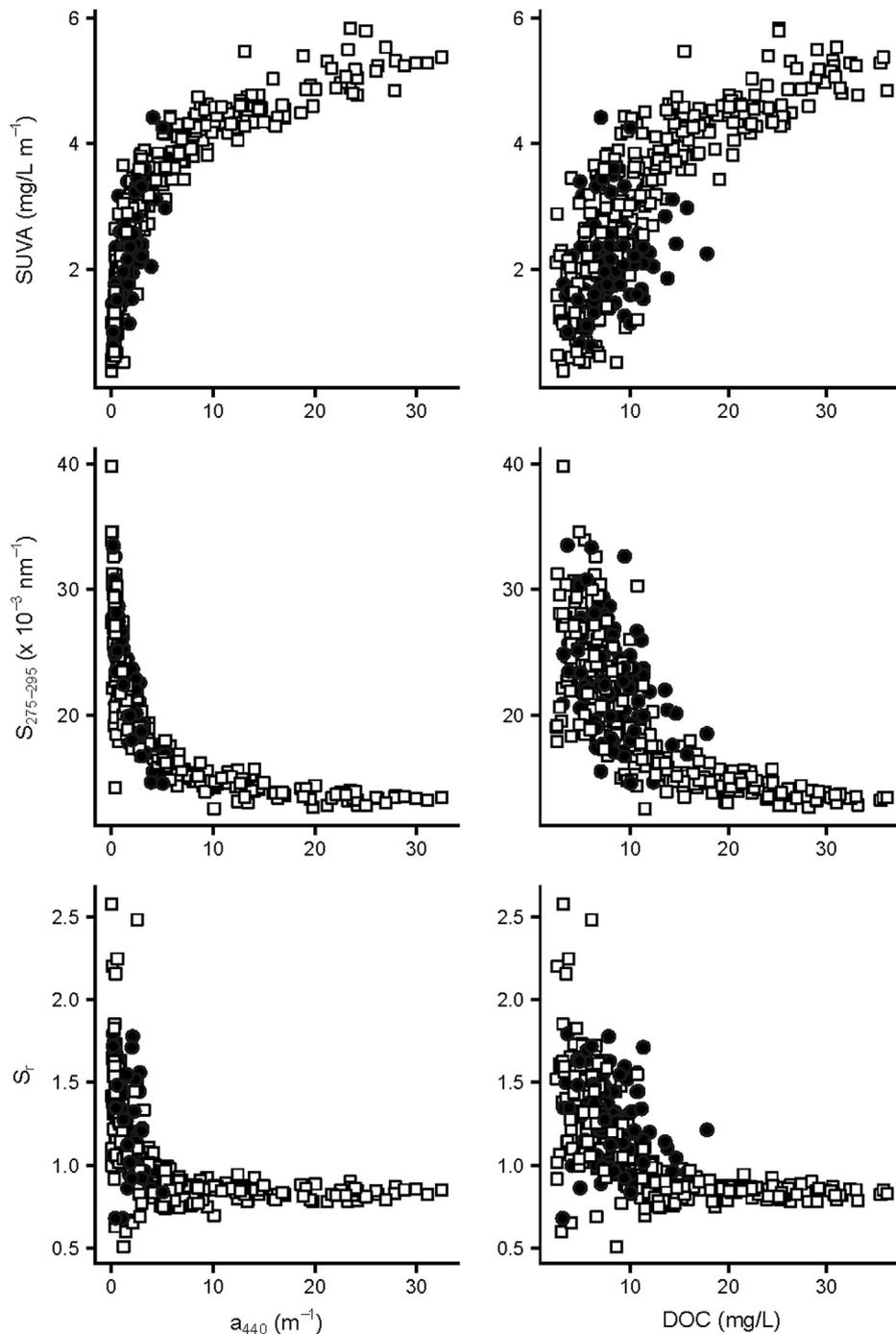


Fig. 3. Plots of SUVA, spectral slope for 275–295 nm, and the slope ratio, S_r vs. a_{440} (left column) and DOC (right column).

350 nm (Carter et al., 2012; Tipping et al., 2009), absorptivity at 275 and 295 nm (Fichot and Benner, 2011), and inclusion of spectral slopes (Asmala et al., 2012). We used a_{254} in addition to a_{440} because wavelengths in the UV-C spectrum resulted in the most robust models in previous studies. The two-term absorptivity models increased R^2 by 0.05–0.09 relative to a single term model for groups of low-color, high-color, and all lakes (Table 3). Nonetheless, two-term models using TDN and a_{440} were more effective at estimating DOC than using absorption at two different wavelengths, particularly for low-color lakes.

4. Discussion

The central goal of this study was to assess the regional variability in CDOM-DOC relationships across a wide range of lakes and examine factors that control such variability. Overall, CDOM and DOC were highly correlated, particularly in lakes with $a_{440} > 3 \text{ m}^{-1}$. In less colored lakes, particularly in the mixed land-cover conditions of the NCHF, the composition of DOM varied more widely leading to weaker relationships between CDOM and DOC. CDOM has become widely used as a proxy for DOC in freshwater systems (Brezonik et al., 2015; Spencer et al., 2012), but its validity has yet to

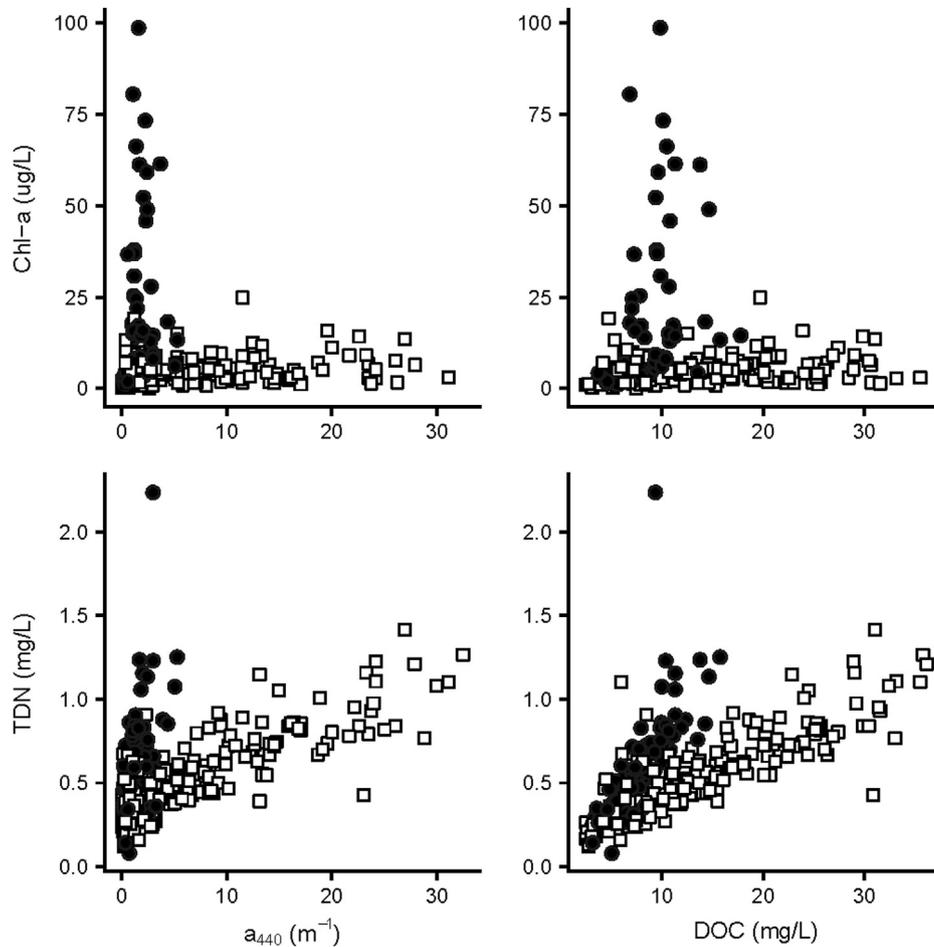


Fig. 4. (Upper) Plots of chl-*a* vs. a_{440} (left) and DOC (right), and (lower) plots of TDN vs. a_{440} (left) and DOC (right) for NCHF sites (filled circles) and NLF (open squares).

Table 3
Predictions of DOC using a_{440} with chl-*a*, TDN, or a_{254} . Coefficient a_0 is the intercept and a_2 is chl-*a*, TDN, or a_{254} . All regressions were performed using natural log-transformed data.

	n	a_0		a_{440}		a_2		R^2	p
		coefficient	p	coefficient	p	coefficient	p		
+ chl- <i>a</i>	264	1.95	<0.0001	0.37	<0.0001	0.03	0.0339	0.84	<0.0001
+ chl- <i>a</i> , < 3 m ⁻¹	151	1.88	<0.0001	0.24	<0.0001	0.07	0.0001	0.48	<0.0001
+ chl- <i>a</i> , > 3 m ⁻¹	111	1.56	<0.0001	0.56	<0.0001	0.01	0.522	0.89	<0.0001
+ TDN	369	2.36	<0.0001	0.30	<0.0001	0.44	<0.0001	0.92	<0.0001
+ TDN, < 3 m ⁻¹	243	2.34	<0.0001	0.22	<0.0001	0.46	<0.0001	0.77	<0.0001
+ TDN, > 3 m ⁻¹	125	1.88	<0.0001	0.48	<0.0001	0.21	<0.0001	0.94	<0.0001
+ a_{254}	368	2.25	<0.0001	0.55	<0.0001	-0.44	<0.0001	0.87	<0.0001
+ a_{254} , <3 m ⁻¹	242	2.15	<0.0001	0.44	<0.0001	-0.31	<0.0001	0.53	<0.0001
+ a_{254} , >3 m ⁻¹	125	2.28	<0.0001	0.74	<0.0001	-0.76	<0.0001	0.95	<0.0001

be widely evaluated in lakes (Li et al., 2016; Massicotte et al., 2017; Wilkinson et al., 2013). Using a large database of lake measurements from three Upper Midwest states, we found clear differences in DOM composition associated with CDOM levels and two large ecoregions. Low-color lakes in both ecoregions show a greater diversity of spectral characteristics and relatively weak relationships between CDOM and DOC than highly-colored lakes (Table 2). Thus, mapping lake carbon pools using satellite remote sensing of CDOM is most reliable for lakes of $a_{440} > 3 \text{ m}^{-1}$.

Low color DOM in lakes can be sourced from both anthropogenic landscapes and autochthonous production. Widespread wetland drainage (Dahl, 1990), land use change, and hydrological

management practices (e.g., tile drainage) leads to decreased inputs of highly colored DOM to aquatic systems from the landscape (Dalzell et al., 2011; Giling et al., 2014; Li et al., 2018). Allochthonous organic matter inputs from agricultural or urban landscapes may be substantial (Hosen et al., 2014; Xenopoulos et al., 2003), but overall the organic matter in such systems lacks abundant aromatic, light-absorbing compounds (Tsui and Finlay, 2011). Landscape modification has also led to increasing eutrophication of Midwestern lakes, with substantial autochthonous production of low color DOM from both phytoplankton and macrophytes (Meili, 1992; Sommaruga and Augustin, 2006; Zhang et al., 2009). Only 12 colored lakes had chl-*a* concentrations exceeding 10 $\mu\text{g/L}$, reflecting

low standing stocks of algal biomass and productivity typical of oligotrophic and mesotrophic lakes. In contrast, 30% of low color lakes had chl-*a* levels exceeding 10 µg/L (Fig. 4), indicating potentially substantial production of low-color, autochthonous DOM through processes such as cell lysis, zooplankton grazing, and microbial exudates. In the lakes studied here, TDN is dominated by dissolved organic nitrogen (DON) and is strongly related to DOC concentrations (Fig. 4). In low color lakes, the DOC:TDN ratio averaged 17.3, in line with the C:N ratio of autochthonous DOM from both phytoplankton and macrophytes (Qu et al., 2013; Zhou et al., 2016). The higher average C:N ratio of 28.0 for colored lakes ($a_{440} > 3 \text{ m}^{-1}$) is indicative of greater terrestrial DOM inputs.

In addition to phytoplankton, abundant macrophytes in many lakes could have produced substantial amounts of low-color, autochthonous DOM (Ginger et al., 2017; Lapierre and Frenette, 2009; Mann and Wetzel, 1996). Chl-*a*, primarily an indicator of phytoplankton biomass, thus by itself cannot explain the variability in CDOM-DOC relationships introduced by macrophytic DOM. The low C:N ratio of DOM from both phytoplankton and macrophytes contributes to the power of TDN in multiple regression with CDOM to predict DOC (Table 3).

Previous studies accounted for the presence of low-color DOC in natural waters by using two-term models, often absorbance at 254 nm and 350 nm (Carter et al., 2012) although other wavelengths could be used (Fichot and Benner, 2011). As with spectral slope values or ratios (e.g., $a_{250}:a_{365}$, as in Peuravuori and Pihlaja, 1997), two-term models allow for variations in absorbance to reflect both the total amount and composition of organic matter. Compositional changes, even as DOC remains constant, would lead to shifts in absorbance at one wavelength, relative to another (Adams et al., 2018). Such two term models generally are most effective when using absorption in UV wavelengths, which limits their use in remote sensing contexts. Both field observations and experimental mesocosm results (Adams et al., 2018) have shown that these two-term models can be effective when phytoplankton are the major source of DOC. The limited significance of chl-*a* in our own models suggests, however, that additional sources of low-color DOM are important from either macrophytes or anthropogenic run-off. The optical characteristics of these sources have not yet been as thoroughly characterized as more humic end-members or phytoplankton-derived DOM. Further, forest type is an important factor in determining CDOM-DOC relationships that cannot be explained by chl-*a*. For instance, Li et al. (2018) found that DOC:CDOM ratios were consistent in evergreen leaf leachate, but could vary widely in leachate from deciduous plants. The transition from mixed broadleaf forest to evergreen forests between the NCHF and NLF could thus also drive the differences in CDOM-DOC relationships between the two ecoregions. Fe previously has also been implicated in increasing absorptivity in boreal waters (e.g., Weyhenmeyer et al., 2014), but such interference does explain CDOM-DOC variance in low color lakes of the Upper Midwest, where Fe is also typically low (Eilers et al., 1988; Poulin et al., 2014; unpublished data). Additionally, Asmala et al. (2012) demonstrated Fe was only important in two-term DOC predictive models if absorption at 520 nm or more was used as the second term.

Microbial transformation and removal of DOM also may influence CDOM-DOC relationships (Mann et al., 2012; Ostapenia et al., 2009). Compositional characteristics of DOM in the NCHF, such as low C:N ratios, high $S_{275-295}$, and low SUVA, suggest that this DOM is labile and can be quickly removed from the waterscape (Frey et al., 2016). Such preferential removal of low color DOM can shift CDOM-DOC relationships, as recalcitrant humic material tends to remain in the water column and undergoes further microbial degradation or photochemical processing only slowly, as described in many river systems (Creed et al., 2015; Stackpole et al., 2017).

Microbial degradation may have driven some of the large variability in DOM composition within low color lakes, where labile and semi-labile DOM may have been consumed at varying rates. Evidence from the Upper Mississippi River, however, indicates that terrestrially-derived DOC is preserved even as the DOC is transported downstream (Voss et al., 2017) in a drainage basin that overlaps with both the NCHF and NLF.

Photochemical processes often play important roles in transforming DOM from both terrestrial and microbial sources in lake-rich environments. In contrast to microbial processing, which leads to the loss of low-color DOM, photobleaching leads to preferential loss of CDOM relative to overall DOC concentrations in lake epilimnia (Berto et al., 2013; Osburn et al., 2009). The highly connected waterscape of the NLF may lead to long exposure of DOM to solar radiation. The importance of photodegradation to DOM loss on a watershed or ecoregional scales has yet to be quantified for the Upper Great Lakes region. $S_{275-295}$ has been previously used as an indicator of photobleaching (Helms et al., 2008). If photobleaching were a strong driver of DOM composition, we would expect marked variability in $S_{275-295}$ relative to DOC in colored lakes, and, indeed, $S_{275-295}$ had a weaker relationship with DOC than with CDOM in waters with $a_{440} > 3 \text{ m}^{-1}$, indicating that photochemical reactions might influence DOM composition. Despite this, $S_{275-295}$ and Sr both remain unchanged where a_{440} exceeds 3 m^{-1} , contrary to what would be expected if photobleaching was a significant driver of changes in DOM composition in highly colored lakes. Additionally, consistency of the relationship between DOC and CDOM above 3 m^{-1} across our dataset suggests that photobleaching had limited influence on colored lakes. Solar irradiation can also lead to the formation of humic-like substances from proteinaceous material in clear, shallow lakes where UV-B is not absorbed by other material (Berto et al., 2013; Bianco et al., 2014). Such processes may be important in oligotrophic groundwater-fed lakes like (Dean and Schwalb, 2002), but are unlikely to be major factors controlling CDOM in most lakes within our study region.

The composition of DOM (and thus a_{440} -DOC relationships) also has temporal components, such as the introduction of photo-labile, humic, microbial DOM to the epilimnion during spring turnover (Osburn et al., 2001). Indeed, our sampling included the unusual hydrological year of 2016, where record-breaking precipitation throughout Minnesota led to extensive flooding and mobilization of sediments and nutrients. Our intra-annual sampling of Minnesota lakes showed that a_{440} -DOC relationships remained mostly stable (Table S2), reflecting the idea that the drivers of CDOM levels act on longer timescales than the drivers of suspended solids or chlorophyll-*a* levels in lakes (Stadelmann et al., 2001). A few sites did show large changes in DOM composition over the course of the summer and fall. Staring Lake, in suburban Minneapolis, showed large deviations from the DOC expected based on Equation (3), with residuals as much as 6 mg/L. Staring Lake has undergone significant restoration efforts, including carp removal, which may have contributed to changing patterns of autochthonous production. In most lakes, however, storage effects and long residence times mute the hydrological and biogeochemical signals that might be expected in response to storm or snow-melt events (Strock et al., 2016).

5. Conclusions: can CDOM be used as a proxy for DOC in temperate lakes?

CDOM-DOC relationships in lakes are products of variability in watershed land use, hydrological connectivity, and autochthonous production. Our results show that autochthonous production and anthropogenic influences on landscapes lead to large variations in DOC relative to CDOM in low color, temperate lakes. DOC in low

colored lakes showed only moderate correlations with a_{440} in both the NLF and NCHF ecoregions. Incorporating TDN data into predictive models for low color lakes improved estimates of DOC in both ecoregions, although chl-*a* was not predictive. Low-color DOC may thus be derived from macrophytes and phytoplankton within a lake or external, anthropogenic sources, such as wastewater and urban run-off. Real-time mapping of DOC from CDOM in low color lakes is thus limited without ancillary water quality data. In visibly colored lakes ($a_{440} > 3 \text{ m}^{-1}$), however, CDOM can be used as a reliable proxy for DOC throughout the U.S. Upper Midwest.

Acknowledgments

The authors thank the United States National Science Foundation (Division of Chemical, Bioengineering, Environmental, and Transport Systems #1510332); the Minnesota Environment and Natural Resources Trust Fund, United States; the University of Minnesota Office of the Vice President of Research, United States; the University of Minnesota Retirees Association, United States; and the University of Minnesota U-Spatial Program, United States for funding this work. We thank Michelle Rorer and many other members of the Finlay lab for their help collecting and analyzing water quality samples. Thanks to Drs. Marvin Bauer, William Arnold, and Yiling Chen whose feedback and input to our work examining CDOM in Midwestern lakes has been valuable.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.watres.2018.08.007>.

References

- Adams, J., Tipping, E., Feuchtmayr, H., Carter, H.T., Keenan, P., 2018. The contribution of algae to freshwater dissolved organic matter: implications for UV spectroscopic analysis. *Int. Waters* 10–21. <https://doi.org/10.1080/20442041.2017.1415032>.
- Asmla, E., Stedmon, C.A., Thomas, D.N., 2012. Linking CDOM spectral absorption to dissolved organic carbon concentrations and loadings in boreal estuaries. *Estuar. Coast Shelf Sci.* 111, 107–117. <https://doi.org/10.1016/j.ecss.2012.06.015>.
- Berto, S., Isaia, M., Sur, B., De Laurentiis, E., Barsotti, F., Buscaino, R., Maurino, V., Minerio, C., Vione, D., 2013. UV-vis spectral modifications of water samples under irradiation: lake vs. subterranean water. *J. Photochem. Photobiol. A Chem.* 251, 85–93. <https://doi.org/10.1016/j.jphotochem.2012.10.019>.
- Bianco, A., Minella, M., De Laurentiis, E., Maurino, V., Minerio, C., Vione, D., 2014. Photochemical generation of photoactive compounds with fulvic-like and humic-like fluorescence in aqueous solution. *Chemosphere* 111, 529–536. <https://doi.org/10.1016/j.chemosphere.2014.04.035>.
- Brezonik, P.L., Olmanson, L.G., Finlay, J.C., Bauer, M.E., 2015. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sens. Environ.* 157, 199–215. <https://doi.org/10.1016/j.rse.2014.04.033>.
- Caplanne, S., Laurion, I., 2008. Effect of chromophoric dissolved organic matter on epilimnetic stratification in lakes. *Aquat. Sci.* 70, 123–133. <https://doi.org/10.1007/s00027-007-7006-0>.
- Carter, H.T., Tipping, E., Koprivnjak, J.F., Miller, M.P., Cookson, B., Hamilton-Taylor, J., 2012. Freshwater DOM quantity and quality from a two-component model of UV absorbance. *Water Res.* 46, 4532–4542. <https://doi.org/10.1016/j.watres.2012.05.021>.
- Chen, J., Zhu, W.N., Tian, Y.Q., Yu, Q., 2017. Estimation of colored dissolved organic matter from Landsat-8 imagery for complex inland water: case study of Lake Huron. *IEEE Trans. Geosci. Rem. Sens.* 55, 1–12. <https://doi.org/10.1109/TGRS.2016.2638828>.
- Creed, I.F., McKnight, D.M., Pellerin, B.A., Green, M.B., Bergamaschi, B.A., Aiken, G.R., Burns, D.A., Findlay, S.E.G., Shanley, J.B., Striegl, R.G., Aulenbach, B.T., Clow, D.W., Laudon, H., McGlynn, B.L., McGuire, K.J., Smith, R.A., Stackpoole, S.M., 2015. The river as a chemostat: fresh perspectives on dissolved organic matter flowing down the river continuum. *Can. J. Fish. Aquat. Sci.* 72, 1272–1285. <https://doi.org/10.1139/cjfas-2014-0400>.
- Dahl, T.E., 1990. Wetlands Losses in the United States 1780s to 1980s. US Fish and Wildlife Service. <https://doi.org/10.2144/000113917>.
- Dalzell, B.J., King, J.Y., Mulla, D.J., Finlay, J.C., Sands, G.R., 2011. Influence of subsurface drainage on quantity and quality of dissolved organic matter export from agricultural landscapes. *J. Geophys. Res. Biogeosci.* 116, 1–13. <https://doi.org/10.1029/2010JG001540>.
- Dean, W.E., Schwab, A., 2002. The lacustrine carbon cycle as illuminated by the waters and sediments of two hydrologically distinct headwater lakes in north-central Minnesota, U.S.A. *J. Sediment. Res.* 72, 416–431. <https://doi.org/10.1306/101801720416>.
- Del Castillo, C.E., Miller, R.L., 2008. On the use of ocean color remote sensing to measure the transport of dissolved organic carbon by the Mississippi River Plume. *Remote Sens. Environ.* 112, 836–844. <https://doi.org/10.1016/j.rse.2007.06.015>.
- Edzwald, J.K., Tobiasson, J.E., 1999. Enhanced coagulation: US requirements and a broader view. *Water Sci. Technol.* 40, 63–70. [https://doi.org/10.1016/S0273-1223\(99\)00641-1](https://doi.org/10.1016/S0273-1223(99)00641-1).
- Eilers, J.M., Brakke, D.F., Landers, D.H., 1988. Chemical and physical characteristics of lakes in the Upper Midwest, United States. *Environ. Sci. Technol.* 22, 164–172. <https://doi.org/10.1021/es00167a005>.
- Fichot, C.G., Benner, R., 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. *Geophys. Res. Lett.* 38, 1–5. <https://doi.org/10.1029/2010GL046152>.
- Fichot, C.G., Downing, B.D., Bergamaschi, B.A., Windham-Myers, L., Marvin-Dipasquale, M., Thompson, D.R., Gierach, M.M., 2016. High-resolution remote sensing of water quality in the San Francisco Bay-Delta Estuary. *Environ. Sci. Technol.* 50, 573–583. <https://doi.org/10.1021/acs.est.5b03518>.
- Finlay, J.C., Small, G.E., Sterner, R.W., 2013. Human Influences on Nitrogen. *Science* (80-.) 342, 247–250. <https://doi.org/10.1126/science.1242575>.
- Frey, K.E., Sobczak, W.V., Mann, P.J., Holmes, R.M., 2016. Optical properties and bioavailability of dissolved organic matter along a flow-path continuum from soil pore waters to the Kolyma River mainstem, East Siberia. *Biogeosciences* 13, 2279–2290. <https://doi.org/10.5194/bg-13-2279-2016>.
- Giling, D.P., Grace, M.R., Thomson, J.R., MacNally, R., Thompson, R., 2014. Effect of native vegetation loss on stream ecosystem processes: Dissolved organic matter composition and export in agricultural landscapes. *Ecosystems* 17, 82–95. <https://doi.org/10.1007/s10021-013-9-97>.
- Ginger, L.J., Zimmer, K.D., Herwig, B.R., Hanson, M.A., Hobbs, W.O., Small, G.E., Cotner, J.B., 2017. Watershed vs. within-lake drivers of nitrogen: Phosphorus dynamics in shallow lakes: Phosphorus. *Ecol. Appl.* 27, 2155–2169. <https://doi.org/10.1002/eap.1599>.
- Griffin, C.G., McClelland, J.W., Frey, K.E., Fiske, G., Holmes, R.M., 2018. Quantifying CDOM and DOC in major Arctic rivers during ice-free conditions using Landsat TM and ETM+ data. *Remote Sens. Environ.* 209, 395–409. <https://doi.org/10.1016/j.rse.2018.02.060>.
- Helms, J.R., Stubbins, A., Ritchie, J.D., Minor, E.C., Kieber, D.J., Mopper, K., 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* 53, 955–969. <https://doi.org/10.4319/lo.2008.53.3.0955>.
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., Megown, K., 2015. Completion of the 2011 National Land Cover Cover Database for the conterminous United States - representing a decade of land cover change information. *Photogramm. Eng. Rem. Sens.* 81.
- Hosen, J.D., McDonough, O.T., Febria, C.M., Palmer, M.A., 2014. Dissolved organic matter quality and bioavailability changes across an urbanization gradient in headwater streams. *Environ. Sci. Technol.* 48, 7817–7824. <https://doi.org/10.1021/es501422z>.
- Houser, J.N., 2006. Water color affects the stratification, surface temperature, heat content, and mean epilimnetic irradiance of small lakes. *Can. J. Fish. Aquat. Sci.* 63, 2447–2455. <https://doi.org/10.1139/f06-131>.
- Hur, J., Park, M.-H., Schlautman, M.A., 2009. Microbial transformation of dissolved leaf litter organic matter and its effects on selected organic matter operational descriptors. *Environ. Sci. Technol.* 43, 2315–2321. <https://doi.org/10.1021/es802773b>.
- Karlsson, J., Bergström, A.K., Byström, P., Gudas, C., Rodríguez, P., Hein, C., 2015. Terrestrial organic matter input suppresses biomass production in lake ecosystems. *Ecology* 96, 2870–2876. <https://doi.org/10.1890/15-0515.1sm>.
- Karlsson, J., Byström, P., Ask, J., Ask, P., Persson, L., Jansson, M., 2009. Light limitation of nutrient-poor lake ecosystems. *Nature* 460, 506–509. <https://doi.org/10.1038/nature08179>.
- Koizumi, S., Craig, N., Zwart, J.A., Kelly, P.T., Weidel, B.C., Jones, S.E., Solomon, C.T., 2018. Experimental whole-lake dissolved organic carbon increase alters fish diet and density but not growth or productivity. *Can. J. Fish. Aquat. Sci.* <https://doi.org/10.1139/cjfas-2017-0283>.
- Kutser, T., 2012. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sens. Environ.* 123, 334–338. <https://doi.org/10.1016/j.rse.2012.04.004>.
- Kutser, T., Verpoorter, C., Paavel, B., Tranvik, L.J., 2015. Estimating lake carbon fractions from remote sensing data. *Remote Sens. Environ.* 157, 138–146. <https://doi.org/10.1016/j.rse.2014.05.020>.
- Lapierre, J.F., Frenette, J.J., 2009. Effects of macrophytes and terrestrial inputs on fluorescent dissolved organic matter in a large river system. *Aquat. Sci.* 71, 15–24. <https://doi.org/10.1007/s00027-009-9133-2>.
- Leavitt, P.R., Cumming, B.F., Smol, J.P., Reasoner, M., Pienitz, R., Hodgson, D. a., 2003. Climatic control of ultraviolet radiation effects on lakes. *Limnol. Oceanogr.* 48, 2062–2069. <https://doi.org/10.4319/lo.2003.48.5.2062>.
- Lee, N., Amy, G., Croué, J.P., Buisson, H., 2004. Identification and understanding of fouling in low-pressure membrane (MF/UF) filtration by natural organic matter (NOM). *Water Res.* 38, 4511–4523. <https://doi.org/10.1016/j.watres.2004.08.013>.
- Li, J., Yu, Q., Tian, Y.Q., Boutt, D., 2018. Effects of landcover, soil property, and temperature on covariations of DOC and CDOM in inland waters. *J. Geophys.*

- Res. Biogeosci. 123, 1–14. <https://doi.org/10.1002/2017JG004179>.
- Li, S., Zhang, J., Mu, G., Ha, S., Sun, C., Ju, H., Zhang, F., Chen, Y., Ma, Q., 2016. Optical properties of chromophoric dissolved organic matter in the yinma river watershed and drinking water resource of northeast China. *Pol. J. Environ. Stud.* 25, 1061–1073. <https://doi.org/10.15244/pjoes/61669>.
- Liang, L., Singer, P.C., 2003. Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. *Environ. Sci. Technol.* 37, 2920–2928. <https://doi.org/10.1021/es026230q>.
- Mann, C.J., Wetzel, R.G., 1996. Loading and utilization of dissolved organic carbon from emergent macrophytes. *Aquat. Bot.* 53, 61–72. [https://doi.org/10.1016/0304-3770\(95\)01012-2](https://doi.org/10.1016/0304-3770(95)01012-2).
- Mann, P.J., Davydova, A., Zimov, N., Spencer, R.G.M., Davydov, S., Bulygina, E., Zimov, S., Holmes, R.M., 2012. Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin. *J. Geophys. Res. Biogeosci.* 117, 1–15. <https://doi.org/10.1029/2011JG001798>.
- Massicotte, P., Asmala, E., Stedmon, C., Markager, S., 2017. Global distribution of dissolved organic matter along the aquatic continuum: Across rivers, lakes and oceans. *Sci. Total Environ.* 609, 180–191. <https://doi.org/10.1016/j.scitotenv.2017.07.076>.
- McKnight, D.M., Benca, K.E., 1990. The chemistry of iron, aluminum, and dissolved organic material in three acidic, metal-enriched, mountain streams, as controlled by watershed and in-stream processes. *Water Resour. Res.* 26, 3087–3100. <https://doi.org/10.1029/WR026i012p03087>.
- Meili, M., 1992. Sources, concentrations and characteristics of organic matter in softwater lakes and streams of the Swedish forest region. *Hydrobiologia* 229, 23–41. <https://doi.org/10.1007/BF00006988>.
- Olmanson, L.G., Brezonik, P.L., Finlay, J.C., Bauer, M.E., 2016a. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sens. Environ.* 185, 119–128. <https://doi.org/10.1016/j.rse.2016.01.007>.
- Osburn, C.L., Morris, D.P., Thorn, K., Moeller, R.E., 2001. Chemical and optical changes in freshwater dissolved organic matter exposed to solar radiation. *Biogeochemistry* 54, 251–278. <https://doi.org/10.1023/A:101065742>.
- Osburn, C.L., Retamal, L., Vincent, W.F., 2009. Photoreactivity of chromophoric dissolved organic matter transported by the Mackenzie River to the Beaufort Sea. *Mar. Chem.* 115, 10–20. <https://doi.org/10.1016/j.marchem.2009.05.003>.
- Osburn, C.L., Stedmon, C.A., 2011. Linking the chemical and optical properties of dissolved organic matter in the Baltic – North Sea transition zone to differentiate three allochthonous inputs. *Mar. Chem.* 126, 281–294. <https://doi.org/10.1016/j.marchem.2011.06.007>.
- Ostapenia, A.P., Parparov, A., Berman, T., 2009. Lability of organic carbon in lakes of different trophic status. *Freshw. Biol.* 54, 1312–1323. <https://doi.org/10.1111/j.1365-2427.2009.02183.x>.
- Pellerin, B.A., Saraceno, J.F., Shanley, J.B., Sebestyen, S.D., Aiken, G.R., Wollheim, W.M., Bergamaschi, B.A., 2012. Taking the pulse of snowmelt: In situ sensors reveal seasonal, event and diurnal patterns of nitrate and dissolved organic matter variability in an upland forest stream. *Biogeochemistry* 108, 183–198. <https://doi.org/10.1007/s10533-011-9589-8>.
- Peuravuori, J., Pihlaja, K., 1997. Molecular size distribution and spectroscopic properties of aquatic humic substances. *Anal. Chim. Acta* 337, 133–149. [https://doi.org/10.1016/S0003-2670\(96\)00412-6](https://doi.org/10.1016/S0003-2670(96)00412-6).
- Poulin, B.A., Ryan, J.N., Aiken, G.R., 2014. Effects of iron on optical properties of dissolved organic matter. *Environ. Sci. Technol.* 48, 10098–10106. <https://doi.org/10.1021/es502670r>.
- Qu, X., Xie, L., Lin, Y., Bai, Y., Zhu, Y., Xie, F., Giesy, J.P., Wu, F., 2013. Quantitative and qualitative characteristics of dissolved organic matter from eight dominant aquatic macrophytes in Lake Dianchi, China. *Environ. Sci. Pollut. Res.* 20, 7413–7423. <https://doi.org/10.1007/s11356-013-1761-3>.
- Schindler, D.W., 2009. Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnol. Oceanogr.* 54, 2349–2358. https://doi.org/10.4319/lo.2009.54.6_part_2.2349.
- Sipler, R.E., Kellogg, C., Yager, P., Sipler, R.E., Kellogg, C.T.E., Connelly, T.L., Roberts, Q.N., 2017. Microbial community response to terrestrially derived dissolved organic matter in the coastal Arctic. *Front. Microbiol.* 8. <https://doi.org/10.3389/fmicb.2017.01018>.
- Sobczak, W.V., Raymond, P.A., 2015. Watershed hydrology and dissolved organic matter export across time scales: minute to millennium. *Freshw. Sci.* 34, 392–398. <https://doi.org/10.1086/679747>.
- Sommaruga, R., 2001. The role of solar UV radiation in the ecology of alpine lakes. *J. Photochem. Photobiol. B Biol.* 62, 35–42. [https://doi.org/10.1016/S1011-1344\(01\)00154-3](https://doi.org/10.1016/S1011-1344(01)00154-3).
- Sommaruga, R., Augustin, G., 2006. Seasonality in UV transparency of an alpine lake is associated to changes in phytoplankton biomass. *Aquat. Sci.* 68, 129–141. <https://doi.org/10.1007/s00027-006-0836-3>.
- Spencer, R.G.M., Butler, K.D., Aiken, G.R., 2012. Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA. *J. Geophys. Res. Biogeosci.* 117. <https://doi.org/10.1029/2011JG001928>.
- Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G.H., Neil, C., Barbosa, C.C.F., Binding, C.E., Bradt, S., Bresciani, M., Dall'Omo, G., Giardino, C., Gitelson, A.A., Kutser, T., Li, L., Matsushita, B., Martinez-Vicente, V., Matthews, M.W., Ogashawara, I., Ruiz-Verdú, A., Schalles, J.F., Tebbs, E., Zhang, Y., Tyler, A.N., 2017. Optical types of inland and coastal waters. *Limnol. Oceanogr.* 63, 846–870. <https://doi.org/10.1002/lno.10674>.
- Stackpole, S.M., Stets, E.G., Clow, D.W., Burns, D.A., Aiken, G.R., Aulenbach, B.T., Creed, I.F., Hirsch, R.M., Laudon, H., Pellerin, B.A., Striegl, R.G., 2017. Spatial and temporal patterns of dissolved organic matter quantity and quality in the Mississippi River Basin, 1997–2013. *Hydrol. Process.* 31, 902–915. <https://doi.org/10.1002/hyp.11072>.
- Stadelmann, T.H., Brezonik, P.L., Kloiber, S., 2001. Seasonal patterns of chlorophyll a and secchi disk transparency in lakes of east-Central Minnesota: Implications for design of ground- and satellite-based monitoring programs. *Lake Reserv. Manag.* 17, 299–314. <https://doi.org/10.1080/07438140109354137>.
- Stedmon, C.A., Amon, R.M.W., Rinehart, A.J., Walker, S.A., 2011. The supply and characteristics of colored dissolved organic matter (CDOM) in the Arctic Ocean: Pan Arctic trends and differences. *Mar. Chem.* 124, 108–118. <https://doi.org/10.1016/j.marchem.2010.12.007>.
- Stevens, A.A., Slocum, C.J., Seeger, D.R., Robeck, G., 1976. Chlorination of organics in drinking water. *J. Am. Water Works Assoc.* 68, 615–620. <https://doi.org/10.1002/j.1551-8833.1976.tb02506.x>.
- Strock, K.E., Saros, J.E., Nelson, S.J., Birkel, S.D., Kahl, J.S., McDowell, W.H., 2016. Extreme weather years drive episodic changes in lake chemistry: implications for recovery from sulfate deposition and long-term trends in dissolved organic carbon. *Biogeochemistry* 127, 353–365. <https://doi.org/10.1007/s10533-016-0185-9>.
- Tehrani, N.C., D'Sa, E.J., Osburn, C.L., Bianchi, T.S., Schaeffer, B.A., 2013. Chromophoric dissolved organic matter and dissolved organic carbon from sea-viewing wide field-of-view sensor (seawifs), moderate resolution imaging spectroradiometer (modis) and meris sensors: Case study for the northern gulf of Mexico. *Rem. Sens.* 5, 1439–1464. <https://doi.org/10.3390/rs5031439>.
- Thrane, J.-E., Hessen, D.O., Andersen, T., 2014. The absorption of light in lakes: negative impact of dissolved organic carbon on primary productivity. *Ecosystems* 17, 1040–1052. <https://doi.org/10.1007/s10021-014-9776-2>.
- Tipping, E., Corbishley, H.T., Koprivnjak, J.F., Lapworth, D.J., Miller, M.P., Vincent, C.D., Hamilton-Taylor, J., 2009. Quantification of natural DOM from UV absorption at two wavelengths. *Environ. Chem.* 6, 472–476. <https://doi.org/10.1071/EN09090>.
- Tsui, M.T.K., Finlay, J.C., 2011. Influence of dissolved organic carbon on methylmercury bioavailability across Minnesota stream ecosystems. *Environ. Sci. Technol.* 45, 5981–5987. <https://doi.org/10.1021/es200332f>.
- Voss, B.M., Wickland, K.P., Aiken, G.R., Striegl, R.G., 2017. Biological and land use controls on the isotopic composition of aquatic carbon in the Upper Mississippi River Basin. *Global Biogeochem. Cycles* 31, 1271–1288. <https://doi.org/10.1002/2017GB005699>.
- Weyhenmeyer, G.A., Prairie, Y.T., Tranvik, L.J., 2014. Browning of boreal freshwaters coupled to carbon-iron interactions along the aquatic continuum. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0088104>.
- Wilkinson, G.M., Pace, M.L., Cole, J.J., 2013. Terrestrial dominance of organic matter in north temperate lakes. *Global Biogeochem. Cycles* 27, 43–51. <https://doi.org/10.1029/2012GB004453>.
- Wilson, H.F., Xenopoulos, M.A., 2009. Effects of agricultural land use on the composition of fluvial dissolved organic matter. *Nat. Geosci.* 2, 37–41. <https://doi.org/10.1038/ngeo391>.
- Xenopoulos, M.A., Lodge, D.M., Frenness, J., Kreps, T.A., Bridgman, S.D., Grossman, E., Jackson, C.J., 2003. Regional comparisons of watershed determinants of dissolved organic carbon in temperate lakes from the Upper Great Lakes region and selected regions globally. *Limnol. Oceanogr.* 48, 2321–2334. <https://doi.org/10.4319/lo.2003.48.6.2321>.
- Xie, H., Aubry, C., Bélanger, S., Song, G., 2012. The dynamics of absorption coefficients of CDOM and particles in the St. Lawrence estuarine system: Biogeochemical and physical implications. *Mar. Chem.* 128–129, 44–56. <https://doi.org/10.1016/j.marchem.2011.10.001>.
- Yamashita, Y., Kloeppel, B.D., Knoepp, J., Zausen, G.L., Jaffé, R., 2011. Effects of watershed history on dissolved organic matter characteristics in headwater streams. *Ecosystems* 14, 1110–1122. <https://doi.org/10.1007/s10021-011-9469-z>.
- Yang, L., Chang, S.W., Shin, H.S., Hur, J., 2015. Tracking the evolution of stream DOM source during storm events using end member mixing analysis based on DOM quality. *J. Hydrol.* 523, 333–341. <https://doi.org/10.1016/j.jhydrol.2015.01.074>.
- Zhang, Y., van Dijk, M.A., Liu, M., Zhu, G., Qin, B., 2009. The contribution of phytoplankton degradation to chromophoric dissolved organic matter (CDOM) in eutrophic shallow lakes: Field and experimental evidence. *Water Res.* 43, 4685–4697. <https://doi.org/10.1016/j.watres.2009.07.024>.
- Zhou, Y., Zhou, J., Jeppesen, E., Zhang, Y., Qin, B., Shi, K., Tang, X., Han, X., 2016. Will enhanced turbulence in inland waters result in elevated production of autochthonous dissolved organic matter? *Sci. Total Environ.* 543, 405–415. <https://doi.org/10.1016/j.scitotenv.2015.11.051>.
- Zhu, W., Yu, Q., Tian, Y.Q., Becker, B.L., Zheng, T., Carrick, H.J., 2014. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. *Remote Sens. Environ.* 140, 766–778. <https://doi.org/10.1016/j.rse.2013.10.015>.



Assessment of the chlorine demand and disinfection byproduct formation potential of surface waters via satellite remote sensing

Yiling Chen ^a, William A. Arnold ^a, Claire G. Griffin ^{b,1}, Leif G. Olmanson ^c,
Patrick L. Brezonik ^a, Raymond M. Hozalski ^{a,*}

^a Department of Civil, Environmental, and Geo- Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN, 55455-0116, United States

^b Department of Ecology, Evolution, and Behavior, University of Minnesota, 1987 Upper Buford Circle, St. Paul, MN, 55108-6097, United States

^c Department of Forest Resources, University of Minnesota, 1530 Cleveland Avenue North, St. Paul, MN, 55108-6112, United States

ARTICLE INFO

Article history:

Received 17 June 2019

Received in revised form

14 August 2019

Accepted 18 August 2019

Available online 21 August 2019

Keywords:

Colored dissolved organic matter

Disinfection byproducts

Chlorine demand

Satellite remote sensing

Surface water quality

Drinking water source

ABSTRACT

The ability of satellites to assess surface water quality indicators such as colored dissolved organic matter (CDOM) suggests that remote sensing could be a useful tool for evaluating water treatability metrics in considering potential drinking water supplies. To explore this possibility, 24 surface water samples were collected throughout Minnesota, USA with wide ranging values of CDOM (a_{440} ; 0.41–27.9 m^{-1}), dissolved organic carbon (DOC; 5.5–47.6 mg/L) and specific ultraviolet absorbance at 254 nm (SUVA₂₅₄; 1.3–5.1 L/mg·M). Laboratory experiments were performed to quantify chlorine demand and the formation of two classes of halogenated disinfection byproducts (DBPs), trihalomethanes (THMs) and haloacetic acids (HAAs), using the uniform formation conditions (UFC) test. Chlorine demand and THM_{UFC} were linearly correlated with CDOM ($R^2 = 0.97$ and 0.91, respectively), indicating that CDOM is a useful predictor of these parameters. On the other hand, data comparing di- and tri-HAA_{UFC} with CDOM were better fit by a logarithmic relationship ($R^2 = 0.73$ and 0.87, respectively), while mono-HAA_{UFC} was linearly correlated with CDOM ($R^2 = 0.46$) but only for low-to moderately-colored waters ($a_{440} \leq 11 m^{-1}$). The correlations relating chlorine demand and DBP_{UFC} values with CDOM were coupled with satellite CDOM assessments to estimate chlorine demand and DBP_{UFC} values for all surface waters larger than 0.05 km² in the state of Minnesota, USA. The resulting maps suggest that only 21.8% of Minnesota lakes would meet both the THM and HAA maximum contaminant levels, but only when pre-disinfection treatment removes 75% of DBP precursors. There are limitations to determining CDOM using satellites for high color surface waters ($a_{440} > 11 m^{-1}$), however, leading to underpredicted values for CDOM, chlorine demand, and DBP_{UFC}. Overall, the results demonstrate the potential benefits of satellite remote sensing for assessing potential drinking water sources and water treatability metrics.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Dissolved organic matter (DOM) originates primarily from two natural sources, production by aquatic organisms (autochthonous) and leaching of decaying terrestrial plant matter (allochthonous) (Thurman, 2012). DOM is also derived from anthropogenic sources such as wastewater effluent and agricultural and urban runoff. Colored dissolved organic matter (CDOM) is the portion of DOM

that absorbs light at a specific wavelength, commonly 420 or 440 nm for freshwaters and 412 nm for marine waters (Brezonik et al., 2015). CDOM is primarily composed of humic and fulvic acids that contain abundant unsaturated carbon-carbon bonds, including aromatic moieties, which are known to absorb light in the ultraviolet (UV) and visible wavelength ranges (Del Castillo and Miller, 2008; Shank et al., 2005). CDOM is a major water quality driver because of its ability to mobilize metals (McKnight and Benca, 1990) and hydrophobic chemicals, act as a photosensitizer for aquatic photochemistry (Gerecke et al., 2001), and control aquatic ecosystem processes (e.g., decrease light penetration) (Houser, 2006; Sommaruga, 2001; Thrane et al., 2014). Although not directly harmful to human health, DOM in source waters, including the CDOM fraction, can have negative effects on the cost

* Corresponding author.

E-mail address: hozalski@umn.edu (R.M. Hozalski).

¹ Current address: Department of Environmental Sciences, University of Virginia, Clark Hall, 291 McCormick Road, Charlottesville, VA, 22905, United States.

and effectiveness of water treatment operations and the finished water quality. It increases the consumption of water treatment chemicals such as coagulants (Sharp et al., 2006) and fouls membranes used for filtration (Zularisam et al., 2006). Furthermore, it reacts with chemical disinfectants such as free chlorine leading to the formation of potentially toxic disinfection byproducts (DBPs) (Beggs et al., 2009; Bellar et al., 1974; Eikebrokk et al., 2004; Herzsprung et al., 2012; Ledesma et al., 2012; Sedlak and von Gunten, 2011). Thus, CDOM plays a critical role in affecting both the chemistry of surface waters and their potential use as drinking water supplies.

The DOM content of surface waters is usually obtained by grab sampling and subsequent laboratory measurement of the dissolved organic carbon (DOC) concentration using a carbon analyzer or the UV/visible light absorbance using a spectrophotometer. Using such a sampling methodology to monitor DOM in large numbers of surface water bodies, on a regular (e.g., monthly) basis would be tedious and cost-prohibitive. Fortunately, investigations over the past 15 years have demonstrated the feasibility and practicality of using satellite remote sensing to measure CDOM in surface waters (Chen et al., 2017; Del Castillo and Miller, 2008; Fichot et al., 2015; Olmanson et al., 2016; Spyrakos et al., 2018; Xie et al., 2012; Zhu et al., 2014). By correlating measured CDOM to satellite reflectance, models can be developed for retrieving CDOM from publicly available satellite imagery. Such imagery, from satellite sensors such as the Landsat series, provide global coverage going back to 1972, potentially permitting researchers to quantify CDOM distributions and variability across long time periods and broad regional scales (Kutser, 2012). Moreover, CDOM typically is closely correlated with DOC in lakes and rivers (Griffin et al., 2018), and thus, the use of CDOM as a proxy of DOC for assessment of water treatability has important practical advantages over other water quality parameters because CDOM can be obtained via satellite remote sensing and is associated with humic substances that are known to have significant impacts on drinking water processes.

Given the importance of DOM in terms of water treatability, many researchers have developed empirical models that relate DOM content as DOC, UV absorbance at 254 nm (UV_{254}), or specific UV absorbance at 254 nm ($SUVA_{254}$) to various measures of water treatability, including chlorine demand (Abdullah et al., 2009; Yee et al., 2006), DBP formation (Amy et al., 1998; Chang et al., 1996; Rodriguez et al., 2000; Sohn et al., 2004; Watson et al., 1993), coagulant consumption (Van Leeuwen et al., 2005; Yang et al., 2010), and membrane fouling (Li and Chen, 2004; Liang et al., 2006). Such models are useful for process control because they permit important treatment concerns, such as water treatment chemical demand and DBP formation, to be predicted using parameters such as DOC and UV_{254} that can be measured in near-real time at a given location (Brinkman and Hozalski, 2011). We are unaware of any water treatability models that consider CDOM as a predictive variable. Developing models that consider CDOM as a variable instead of DOC or UV_{254} would enable the use of satellite remote sensing to assess potential water supplies and monitor the treatability of existing surface water supplies.

The objective of this study was to develop relationships between CDOM and two measures of water treatability, chlorine demand and DBP formation. A focus of this study was to compare trends for different classes of DBPs. Surface water samples were collected from 24 locations (in 20 lakes and 1 river) throughout Minnesota, and the formation of two regulated DBP classes (THMs and HAAs) upon chlorination was measured following the uniform formation conditions (UFC) test (Summers et al., 1996). The relationships between CDOM, DOC, UV-vis absorbance, and $SUVA$ values for the water samples and their measured chlorine demand and DBP formation potentials (THMs, mono-, di- and tri-HAAs)

were determined. The resulting models were coupled with Landsat-derived CDOM measurements to create statewide chlorine demand and DBP_{UFC} maps of Minnesota's lakes and rivers to demonstrate the possible use of satellite remote sensing for rapid large-scale assessments of water treatability.

2. Methods

2.1. Chemicals and reagents

A table of the chemicals used and their purities and suppliers is provided in the Supplementary Material (Table S1). All solutions were prepared using ultrapure water (18.2 M Ω -cm resistivity; Millipore).

2.2. Lake water sampling

Surface water samples were collected from 24 locations (20 lakes and 1 river; Lake Vermilion, Lake of the Woods, and Upper Red Lake were sampled twice in different locations) in four large ecoregions of Minnesota, USA – the Northern Lakes and Forests (NLF; 15 lakes), the North Central Hardwood Forests (NCHF; 4 lakes), the Northern Minnesota Wetlands (NMW; 4 lakes), and the Western Corn Belt Plains (WCBP; 1 lake) (Fig. S1). The adjacent NCHF and WCBP ecoregions were combined into a single category (NCHF + WCBP) because there was only one lake representing the WCBP and very little of its watershed is within the WCBP. The water samples were collected during the summer of 2016 and represent a subset of 412 samples collected from 299 locations (mostly lakes) across the Upper Midwest states of Minnesota, Wisconsin, and Michigan between 2014 and 2016 for analysis of CDOM and other water quality parameters (Griffin et al., 2018). The 24 water samples included in this investigation were selected to provide geographic diversity and a wide range of CDOM. Water samples typically were collected from a boat, canoe, or kayak at least 100 m from shore and within 1 m of the water surface by submerging acid-washed and triple-rinsed polycarbonate or high-density polyethylene (HDPE) containers to arm's length and then removing the cap. GPS coordinates of the sampling locations are provided in Table S2. Water samples for subsequent chlorine demand and DBP testing were filtered within a few hours of collection by pumping from the collection containers through 0.45 μ m Geotech High Capacity capsule filters into clean containers. Aliquots of the filtered water for subsequent UV-visible light scans and DOC analysis were placed into separate pre-ashed 40 mL amber glass vials with the latter acidified to pH < 2 using 2 M HCl. All water samples were stored in the dark under refrigeration until further use. Chlorophyll-*a* was isolated from lake water samples by vacuum filtering onto 0.22 μ m cellulose nitrate filters and stored frozen until analysis.

2.3. Chlorine demand and DBP formation testing

The chlorine demand of each water sample was determined by adding a stock solution of sodium hypochlorite (NaOCl) with 10% available chlorine at three different doses (Cl_2 :DOC ratios of 1:1, 1.7:1 and 2:1) and measuring the residual concentration after incubating for 24 h at room temperature. The chlorine dose for subsequent UFC testing was selected to yield a free chlorine residual of 1.0 ± 0.4 mg Cl_2 /L at the end of incubation. All glassware used in the chlorine demand and UFC experiments was cleaned by soaking in a concentrated hydrochloric acid bath (10%) for 24 h, rinsing with ultrapure water, and then baking at 550 °C for at least 6 h.

For UFC testing, triplicate aliquots of each water sample were

dosed with 2 mL/L borate buffer (1.0 M boric acid and 0.26 M sodium hydroxide, pH = 8.0) and adjusted to pH = 8.0 ± 0.2 using 1.0 M sulfuric acid or 1.0 M sodium hydroxide. After an incubation time of 24 h in the dark at 20 ± 1 °C, the chlorine residual was determined to ensure that it was within the required range. Samples outside the desired range were discarded and the test repeated with the chlorine dose adjusted as necessary. The chlorine residual was then quenched with Na₂SO₃ prior to extraction and analysis of DBPs.

2.4. Analytical methods

DOC was measured using a Shimadzu TOC L CSN analyzer after sparging the acidified samples with nitrogen gas. UV–Vis absorbance (*A*) scans from 200 to 800 nm were obtained using a Shimadzu 1601UV-PC dual beam spectrophotometer and 1 or 5 cm quartz cells. CDOM was calculated as the Napierian absorption coefficient at 440 nm (*a*₄₄₀) using the following equation:

$$a_{440} \left(\text{m}^{-1} \right) = 2.303 * A_{440} / l \quad (1)$$

where: *A*₄₄₀ is the absorbance at 440 nm and *l* is the path length (m). Specific ultraviolet absorbance at 254 nm (SUVA₂₅₄) was calculated by normalizing the absorbance at 254 nm (UV₂₅₄) to the DOC concentration. Spectral slopes *S*_{275–295} and *S*_{350–400} were computed by fitting a linear regression to the natural logarithm of absorbance between the wavelength 275–295 and 350–400 nm, respectively. Chlorophyll *a* (chl *a*) was determined using a standard fluorometric method (Method 10200H) (Eaton et al., 2005). Total suspended solids (TSS) was obtained by filtering lake water through a pre-ashed 47 mm Whatman GF/F filter and normalizing the mass of dry residue by the volume of lake water filtered. Bromide concentration was determined via ion chromatography. Free chlorine was measured using the *N,N*-diethyl-*p*-phenylenediamine (DPD) method and a Hach DR/890 portable colorimeter.

The concentrations of the THM₄ compounds (chloroform, TCM; bromodichloromethane, BDCM; dibromochloromethane, DBCM, and bromoform, TBM) and HAA₉ compounds (monochloroacetic acid, MCAA; dichloroacetic acid, DCAA; trichloroacetic acid, TCAA; monobromoacetic acid, MBAA; dibromoacetic acid, DBAA; tri-bromoacetic acid, TBAA; chlorodibromoacetic acid, CDBAA; bromochloroacetic acid, BCAA; and bromodichloroacetic acid, BDCAA) were determined using gas chromatography-microelectron capture detection (GC-μECD) according to modified USEPA Methods 551.1 and 552.3, respectively. Method details are provided in the SI. Specific chlorine demand and specific THM and HAA formation potentials (Specific THM_{UFC} and HAA_{UFC}) were calculated by normalizing chlorine demand, THM_{UFC}, or HAA_{UFC} to the DOC concentration.

2.5. Data analysis and mapping

CDOM maps of Minnesota were created using an empirical model developed from Landsat 8 OLI Surface Reflectance (SR) Level-2 products from the Earth Resources Observation and Science (EROS) Center. The model was derived by step-wise regression to relate CDOM (*a*₄₄₀) data from 194 *in situ* measurements with corresponding spectral-radiometric data from four paths of clear Landsat 8 imagery from 2015 to 2016. The two-variable model used two band ratios: OLI3/OLI4 (green/red) and OLI4/OLI5 (red/near infrared):

$$\ln a_{440} = -5.478 * (\text{OLI3/OLI4}) - 0.633 * (\text{OLI4/OLI5}) + 8.135 \quad (2)$$

and had an *R*² of 0.77 and root-mean square error (RMSE) of 0.624 m⁻¹. This model was applied to 2016 clear Landsat 8 imagery with a few 2017 images to fill in areas not available in the 2016 imagery due to cloud cover. These areas were in western Minnesota and included portions of the Lake of the Woods and the Red Lakes, which are among Minnesota's largest lakes. Image processing was conducted using ERDAS Imagine software. Maps produced from multiple Landsat images were mosaicked to create the final state-wide CDOM map for Minnesota. The resulting pixel level CDOM maps together with experimentally obtained correlations (either Pearson for linear or LINEST for non-linear using Microsoft Excel) between chlorine demand and *a*₄₄₀, THM_{UFC} and *a*₄₄₀, and HAA_{UFC} and *a*₄₄₀ were used to create state-wide maps for chlorine demand, THM_{UFC}, and HAA_{UFC}.

3. Results

3.1. Characteristics of lake waters

The water quality parameter values for the 24 water samples are summarized in Table S2. Observed ranges for key parameters are: CDOM (*a*₄₄₀): 0.41–27.9 m⁻¹, DOC: 5.5–47.6 mg/L, UV₂₅₄: 8.7–200 m⁻¹, SUVA₂₅₄: 1.3–5.1 L/mg-m, chlorophyll *a*: 1.1–14.4 μg/L, TSS: 0.7–13.9 mg/L, and pH: 7.1–8.8. Bromide concentrations were generally less than the limit of detection (0.01 mg/L) except for five samples with values ranging from 0.04 to 0.07 mg/L.

Relationships among the measured DOM parameters (*a*₄₄₀, DOC, UV₂₅₄ and SUVA₂₅₄) are shown in Fig. 1. DOC and UV₂₅₄ correlated linearly with *a*₄₄₀ across the entire data range (*R*² = 0.93 and 0.98, respectively). The slopes and *R*² values for the Pearson correlations were similar when excluding three sampling sites with very high *a*₄₄₀ values (South Sturgeon Lake, Lake Vermilion (Pike Bay)), and Upper Red Lake #2; data not shown). For SUVA₂₅₄, the greatest correlation coefficient values were obtained for linear regressions with log *a*₄₄₀ (*R*² = 0.88) and log DOC (*R*² = 0.78). *a*₄₄₀ did not correlate with other water quality indicators (*R*² < 0.18, and *p* > 0.05) including pH, TSS, chlorophyll *a*, and bromide (Table S2). Spectral slopes were computed from plots of the natural logarithm of absorbance vs. wavelength over the range 250–450 nm for each water sample (Fig. S2). Spectral slopes (*S*_{275–295}) among the water samples ranged from 0.013 to 0.028 nm⁻¹, but samples with visible color (i.e., *a*₄₄₀ > 3 m⁻¹) had a narrow range of *S*_{275–295}, 0.013–0.018 nm⁻¹. *S*_{350–400} also showed small variation ranging from 0.015 to 0.020 nm⁻¹, except for Mille Lacs Lake (low *a*₄₄₀) which had a *S*_{350–400} of 0.009 nm⁻¹. The results indicated that the optical properties of *a*₄₄₀ from moderate to high *a*₄₄₀ waters are consistent across a large geographic range.

3.2. Chlorine demand

Chlorine demand values (mean ± standard deviation) for the 24 samples (Table S3) ranged from 3.68 ± 0.10 to 72.0 ± 0.17 mg Cl₂/L, with a mean of 14.9 mg Cl₂/L and a median of 8.58 mg Cl₂/L. In addition, the specific chlorine demand ranged from 0.56 ± 0.01 to 1.70 ± 0.01 mg Cl₂/mg C, and the mean and median specific chlorine demand were 0.97 and 0.85 mg Cl₂/mg C, respectively.

Chlorine demand exhibited a strong linear correlation with *a*₄₄₀ (*R*² = 0.97; slope = 2.18 m-mg Cl₂/L), DOC (*R*² = 0.98; slope = 1.75 mg Cl₂/mg C), and UV₂₅₄ (*R*² = 0.99; slope = 0.35 m-mg Cl₂/L) (Fig. 2), indicating that all three parameters are useful predictors of chlorine demand for the lakes. The slopes and *R*² values remained similar when the three lakes with very high *a*₄₄₀ values were excluded from the regressions (data not shown). In addition, the organic carbon normalized specific chlorine demand correlated linearly with SUVA₂₅₄ (*R*² = 0.84; slope = 0.28 m-mg Cl₂/L).

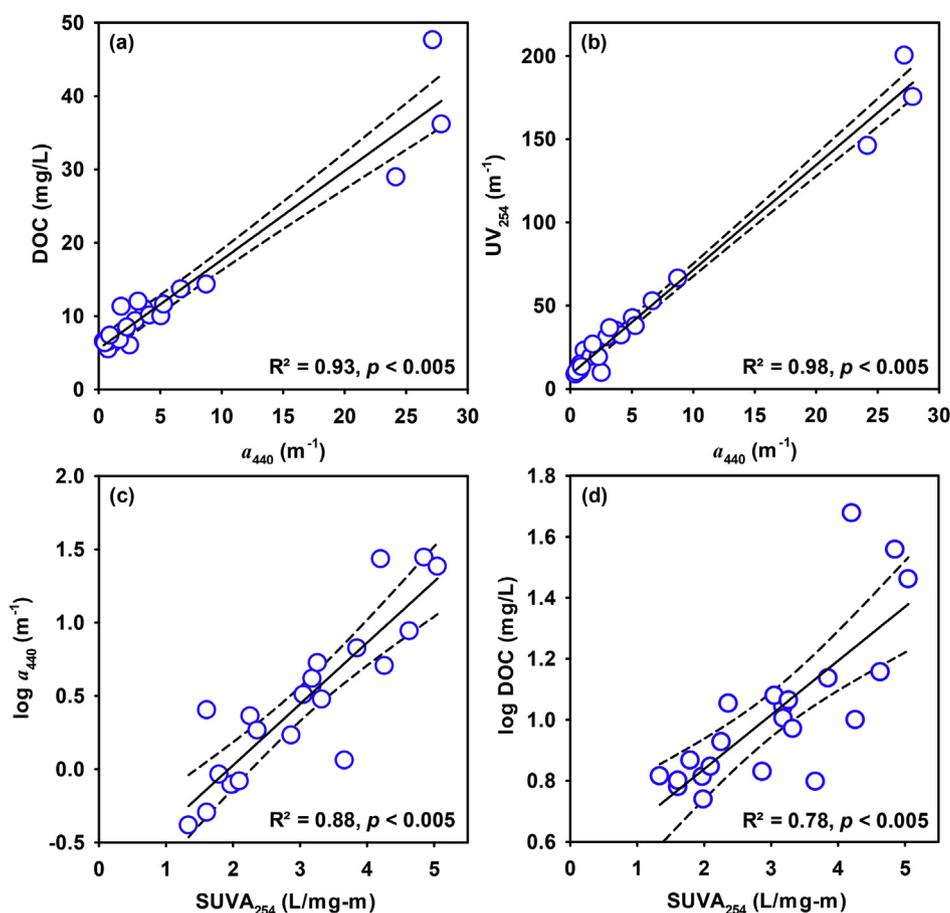


Fig. 1. Correlations among measured DOM parameters for selected Minnesota surface water samples (Table S1). The equations for the regressions shown are:

(a) $\text{DOC (mg/L)} = 1.21 (\pm 0.07) \cdot a_{440} (\text{m}^{-1}) + 5.54 (\pm 0.76)$;

(b) $\text{UV}_{254} (\text{m}^{-1}) = 6.30 (\pm 0.20) \cdot a_{440} (\text{m}^{-1}) + 8.57 (\pm 2.03)$;

(c) $\log a_{440} (\text{m}^{-1}) = 0.42 (\pm 0.05) \cdot \text{SUVA}_{254} (\text{L/mg-m}) - 0.81 (\pm 0.16)$;

(d) $\log \text{DOC (mg/L)} = 0.18 (\pm 0.03) \cdot \text{SUVA}_{254} (\text{L/mg-m}) + 0.49 (\pm 0.10)$.

* Dashed lines represent 95% confidence intervals about the solid regression lines.

3.3. DBP formation

A wide range of DBP_{UFC} values was found for the 24 waters (Table S3). South Sturgeon Lake, Lake Vermilion (Pike Bay), and Lake of the Woods (Muskeg Bay), which had the highest DOC levels, generated the maximum values for both THM_{UFC} and HAA_{UFC} . The only THMs observed as chlorination products were TCM and BDCM (Fig. S3a). Although a few lakes contained bromide at concentrations exceeding the LOD, TCM was the dominant THM species formed in all samples, with mass concentration corresponding to >80% of the total THMs. The observed THM_{UFC} values ranged from 104 ± 16.9 to $2019 \pm 3.73 \mu\text{g/L}$, and the mean THM_{UFC} ($421 \mu\text{g/L}$) was much greater than the median THM_{UFC} ($264 \mu\text{g/L}$). In contrast, specific THM_{UFC} values ranged from 15.9 ± 0.95 to $55.9 \pm 0.10 \mu\text{g/mg C}$ with comparable mean and median values (29.8 and $27.5 \mu\text{g/mg C}$, respectively).

Five HAA species (MCAA, MBAA, DCAA, TCAA, BDCAA) were detected after chlorine dosing (Fig. S3b). In general, the total HAA_{UFC} values were comparable to the total THM_{UFC} values (Table S3). Although South Sturgeon Lake had the maximum HAA_{UFC} ($1366 \pm 103 \mu\text{g/L}$), Lake of the Woods (Fourmile Bay) had the highest specific HAA_{UFC} ($76.0 \pm 2.19 \mu\text{g/mg C}$). The lowest HAA_{UFC} ($47.0 \pm 19.4 \mu\text{g/L}$) and specific HAA_{UFC} ($7.22 \pm 2.98 \mu\text{g/mg C}$) were obtained for Island Lake. The mean HAA_{UFC} was $469 \mu\text{g/L}$, which is greater than the median value ($363 \mu\text{g/L}$), while the

average specific HAA_{UFC} ($35.1 \mu\text{g/mg C}$) was similar to the median value ($32.0 \mu\text{g/mg C}$). Distributions of HAAs were sorted into mono-, di- and tri-HAA species (Table S4) because each group apparently is formed from distinct precursors and chemical pathways (Chellam and Krasner, 2001). Consistent with many previous studies (Andrews et al., 2005; Krasner et al., 2006; Nuckols et al., 2001), the dominant HAA species were DCAA and TCAA, accounting for $74.6 \pm 11.1\%$ of the HAA_5 pool by mass, followed by MCAA. MBAA and BDCAA were detected but never exceeded 1% of the total HAA_5 pool (data not shown).

3.4. Correlations between DBP formation and CDOM

THM_{UFC} exhibited a strong linear correlation with a_{440} ($R^2 = 0.91$; slope = $52.5 \text{ m-}\mu\text{g/L}$; Fig. 3a), indicating that CDOM (a_{440}) is a useful predictor of THM yields. Considering only the 21 waters with a_{440} below 9 m^{-1} , the R^2 decreased slightly ($R^2 = 0.84$), but the slope increased to $62.5 \text{ m-}\mu\text{g/L}$.

Regression analyses on the formation potentials of tri-, di- and mono-HAAs (Fig. 3b–d) in relation to a_{440} showed different trends than the $\text{THM}_{\text{UFC}}-a_{440}$ relationship. Tri- HAA_{UFC} correlated linearly with a_{440} ($R^2 = 0.87$) for low-to moderately-colored waters ($a_{440} \leq 11 \text{ m}^{-1}$), but logarithmically when the entire data set was considered. The relationship between di- HAA_{UFC} and a_{440} was similar, but with a lower slope in the low-to moderately-colored

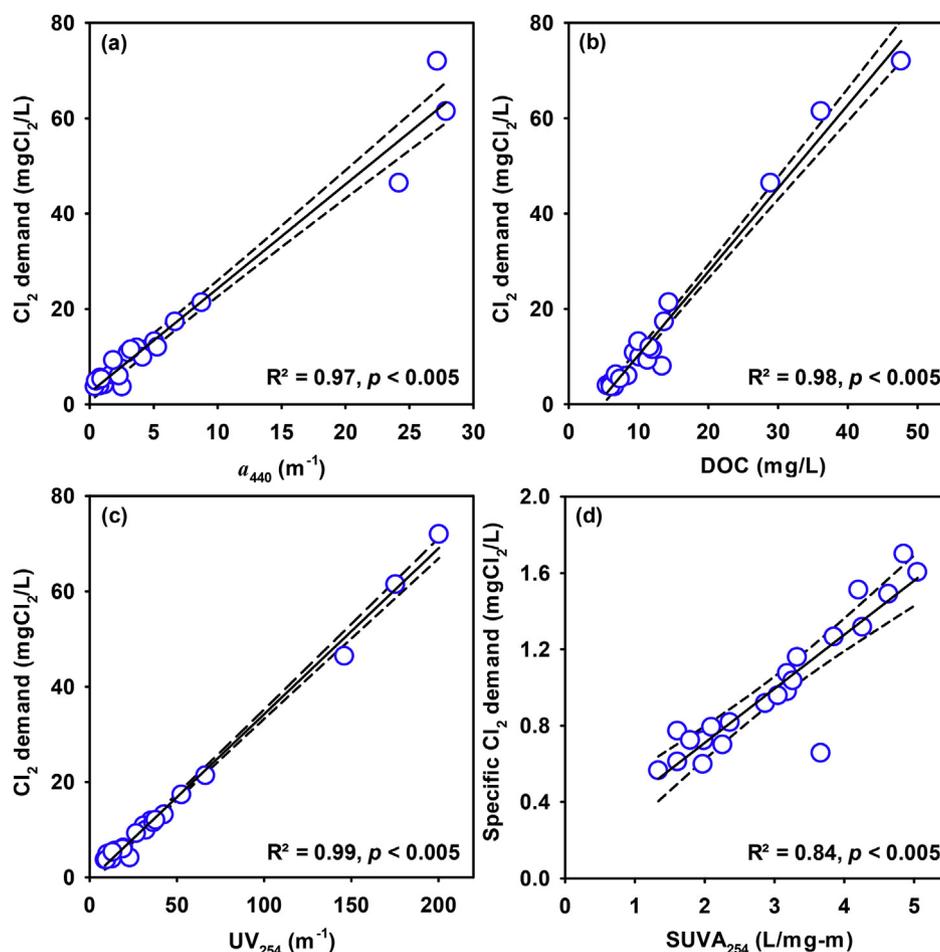


Fig. 2. Correlations of chlorine demand with CDOM (a_{440}), DOC, and UV_{254} (panels a–c) and specific chlorine demand with $SUVA_{254}$ (panel d). The regression equations are:

- (a) Cl_2 demand ($mgCl_2/L$) = $2.18 (\pm 0.09) \cdot a_{440} (m^{-1}) + 2.47 (\pm 0.90)$;
 (b) Cl_2 demand ($mgCl_2/L$) = $1.75 (\pm 0.06) \cdot DOC (mg/L) - 7.17 (\pm 0.92)$;
 (c) Cl_2 demand ($mgCl_2/L$) = $0.35 (\pm 0.01) \cdot UV_{254} (m^{-1}) - 0.55 (\pm 0.44)$;
 (d) specific Cl_2 demand ($mgCl_2/mgC$) = $0.28 (\pm 0.03) \cdot SUVA_{254} (L/mg-m) + 0.14 (\pm 0.09)$.

* Dashed lines represent 95% confidence intervals about the solid regression lines.

lake waters ($R^2 = 0.73$; slope = $25.7 m-\mu g/L$). Mono-HAA_{UFC}, on the other hand, exhibited a relatively weak correlation with a_{440} , with a linear correlation only observed for low-to moderately-colored waters ($R^2 = 0.46$; slope = $16.7 m-\mu g/L$). Overall, these results agree with previous reports (Bond et al., 2009; Bull et al., 2006; Chang et al., 2006; Dickenson et al., 2008; Hua and Reckhow, 2007; Kanokkantapong et al., 2006; Kraus et al., 2010; Liang and Singer, 2003) that suggest the three classes of HAA species are formed from different precursors (see below).

3.5. Correlations between DBP formation and DOC or UV absorbance

DOC and UV_{254} have been widely used as surrogate parameters to predict DBP formation (Chowdhury et al., 2009; Sadiq and Rodriguez, 2004). Given the linear relationships between UV_{254} and a_{440} and between DOC and a_{440} (Fig. 1), it is not surprising that the strength of correlations between THM_{UFC}, di- and tri-HAA_{UFC} and DOC (Fig. S4) or UV_{254} (Fig. S5) were similar to those for a_{440} . THM_{UFC} exhibited a relatively strong linear relationship with DOC ($R^2 = 0.84$; slope = $40.3 \mu g/mgC$) and UV_{254} ($R^2 = 0.91$; slope = $8.21 m-\mu g/L$). Di- and tri-HAA_{UFC} relationships with DOC and UV_{254} , however, were best fit by logarithmic relationships

($R^2 = 0.87$ and 0.73 , respectively, for DOC and $R^2 = 0.91$ and 0.82 , respectively, for UV_{254}). When considering only the low-to moderately-colored lake waters, di- and tri-HAA_{UFC} increased linearly with increasing DOC ($R^2 = 0.67$ and 0.66 , respectively) and UV_{254} ($R^2 = 0.90$ and 0.83 , respectively). These results support the conclusion from previous studies that UV_{254} serves as a slightly better predictor of DBP formation than DOC concentration (Edzwald et al., 1985; Najm et al., 1994).

Mono-HAA_{UFC} correlated poorly with DOC and UV_{254} when all data were considered, but a stronger correlation was found for low-to moderately-colored waters ($R^2 = 0.69$, slope = $15.8 \mu g/mg C$ for DOC, Fig. S4d; $R^2 = 0.58$, slope = $2.73 m-\mu g/L$, Fig. S5d for UV_{254}).

3.6. Correlations between specific DBP formation and SUVA

SUVA is often used as a proxy for the hydrophobicity, molecular weight (MW) and aromaticity of DOM (Dickenson et al., 2008; Hua and Reckhow, 2007; Singer et al., 2002). DOC-normalized or specific THM_{UFC} and tri-HAA_{UFC} exhibited relatively strong linear correlations with SUVA ($R^2 = 0.67$ and 0.65 , respectively; slope = 7.57 and $8.17 m-\mu g/L$, respectively; Fig. S6) suggesting that these DBPs are formed from UV absorbing organic matter (i.e., humic substances). Conversely, relatively weak correlations were observed

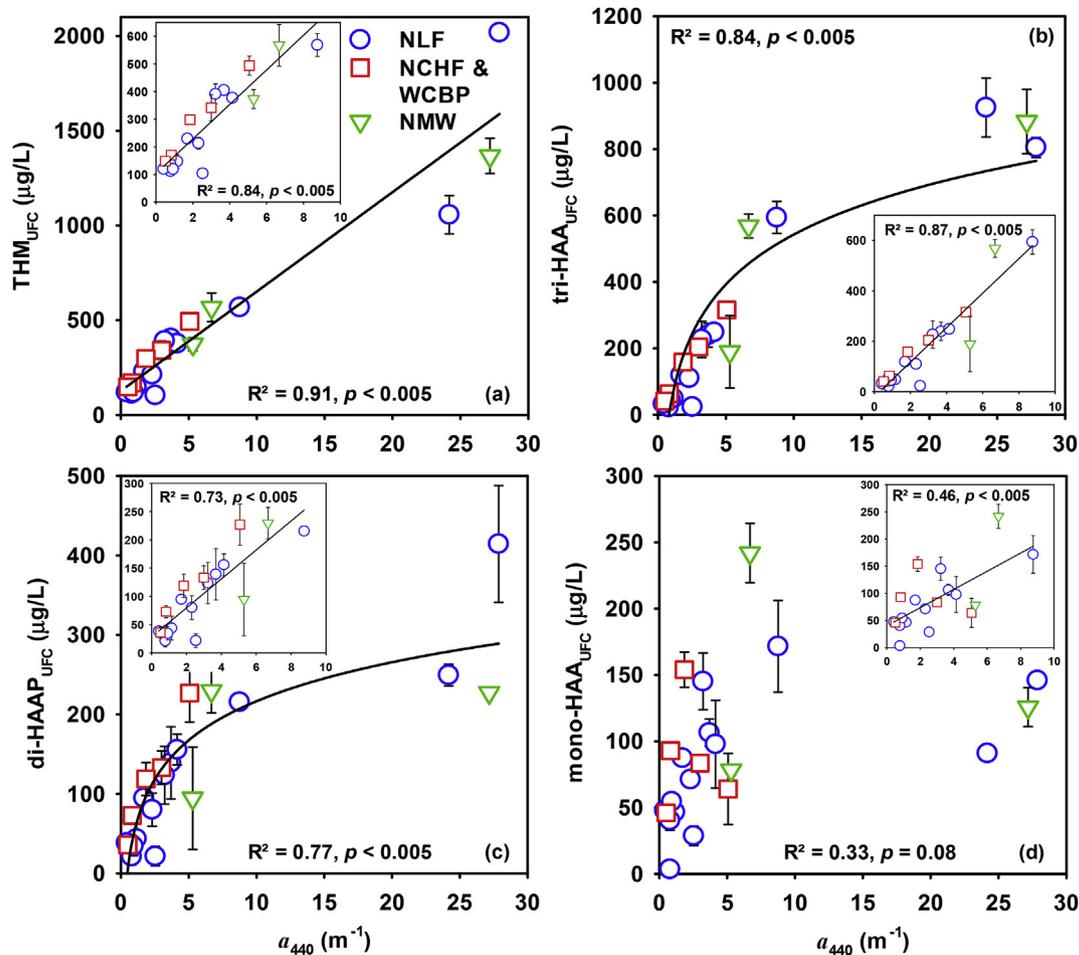


Fig. 3. Correlations of THM_{UFC}, tri-HAA_{UFC}, di-HAA_{UFC}, and mono-HAA_{UFC} with CDOM (a_{440}). The equations for the regressions are:
 (a) THM_{UFC} ($\mu\text{g/L}$) = $52.5 (\pm 3.67) \cdot a_{440} (\text{m}^{-1}) + 127 (\pm 37.9)$; THM_{UFC} ($\mu\text{g/L}$) for low a_{440} = $62.5 (\pm 6.59) \cdot a_{440} (\text{m}^{-1}) + 104 (\pm 23.8)$;
 (b) tri-HAA_{UFC} ($\mu\text{g/L}$) = $217 (\pm 21.0) \cdot \ln a_{440} (\text{m}^{-1}) + 43.4 (\pm 33.4)$; tri-HAA_{UFC} ($\mu\text{g/L}$) for low a_{440} = $68.3 (\pm 6.34) \cdot a_{440} (\text{m}^{-1}) - 19.0 (\pm 22.9)$;
 (c) di-HAAP_{UFC} ($\mu\text{g/L}$) = $70.7 (\pm 8.64) \cdot \ln a_{440} (\text{m}^{-1}) + 53.8 (\pm 13.8)$; di-HAAP_{UFC} ($\mu\text{g/L}$) for low a_{440} = $25.7 (\pm 3.76) \cdot a_{440} (\text{m}^{-1}) + 28.2 (\pm 13.6)$;
 (d) mono-HAA_{UFC} ($\mu\text{g/L}$) for low a_{440} = $16.7 (\pm 4.41) \cdot a_{440} (\text{m}^{-1}) + 40.3 (\pm 15.9)$.

between SUVA and specific mono- and di-HAA_{UFC}. In addition, when lakes with SUVA exceeding 3 L/mg-m were considered, the mono-HAA yields started to decrease with increasing SUVA (Fig. S6d). These results imply that low UV absorbing organic matter (i.e., non-humic substances) contributed substantially to formation of mono- and di-HAA species.

3.7. Chlorine demand and DBP_{UFC} mapping

Using a_{440} data from satellite imagery and the correlations between chlorine demand and a_{440} (Fig. 2), THM_{UFC} and a_{440} (Fig. 3) and HAA_{UFC} and a_{440} (Fig. S7), state-wide maps for chlorine demand and DBP formation potentials were produced. For the map for THM_{UFC} (Fig. 4) shows relatively high levels of THM_{UFC} (>400 $\mu\text{g/L}$) in lakes of northeastern Minnesota, especially within the NLF ecoregion. This result is not surprising as most Minnesota lakes with $a_{440} > 3 \text{ m}^{-1}$ are located in the NLF ecoregion. Lower THM_{UFC} values are predicted for the NCHF, NMW and WCBP ecoregions, and most lakes in these three ecoregions are relatively low in color. Consistent with the THM_{UFC} results, NLF lakes generally have higher levels of HAA_{UFC} (>480 $\mu\text{g/L}$) (Fig. S8) and chlorine demand (>5 mgCl_2/L) (Fig. S9). The distribution of DBP_{UFC} levels in the four ecoregions clearly reflected their current distributions of forested, urban and agricultural land uses. For example, the NLF ecoregion is extensively forested,

with large areas of wetlands and open water. Undeveloped land covers 89% of the NLF ecoregion. In contrast, almost half of the NCHF ecoregion is cultivated, with forests and wetlands only making up 35% of land cover (Brezonik et al., 2019). The WCBP ecoregion is dominated by agricultural use (83% of land cover) but also contains some open grassland (Brezonik et al., 2019).

Comparisons of values of CDOM, chlorine demand, THM_{UFC}, and HAA_{UFC} estimated from satellite imagery versus experimentally measured values are shown in Fig. 5. The estimated values represent an average of the grid cell containing the physical sampling location and the neighboring cells. For lakes with CDOM values less than 10 m^{-1} , the plotted values for all parameters lie close to the 1:1 line (Fig. 5a), indicating that satellite remote sensing is capable of accurate assessing CDOM and CDOM-associated parameters such as THM_{UFC} in low to moderately-colored waters. The agreement was poor, however, for two lakes with very high CDOM levels ($a_{440} > 20 \text{ m}^{-1}$) (i.e., South Sturgeon Lake and Lake Vermilion (Pike Bay)), as the estimated CDOM, chlorine demand and DBP_{UFC} values were substantially less than the corresponding experimentally-measured values (Fig. 5a–d). In addition, negative HAA_{UFC} values were obtained for six low CDOM ($a_{440} < 2.5 \text{ m}^{-1}$) sampling sites (Fig. 5d). The latter results suggest that estimation of HAA_{UFC} using satellite remote sensing can be problematic for waters with very low color.

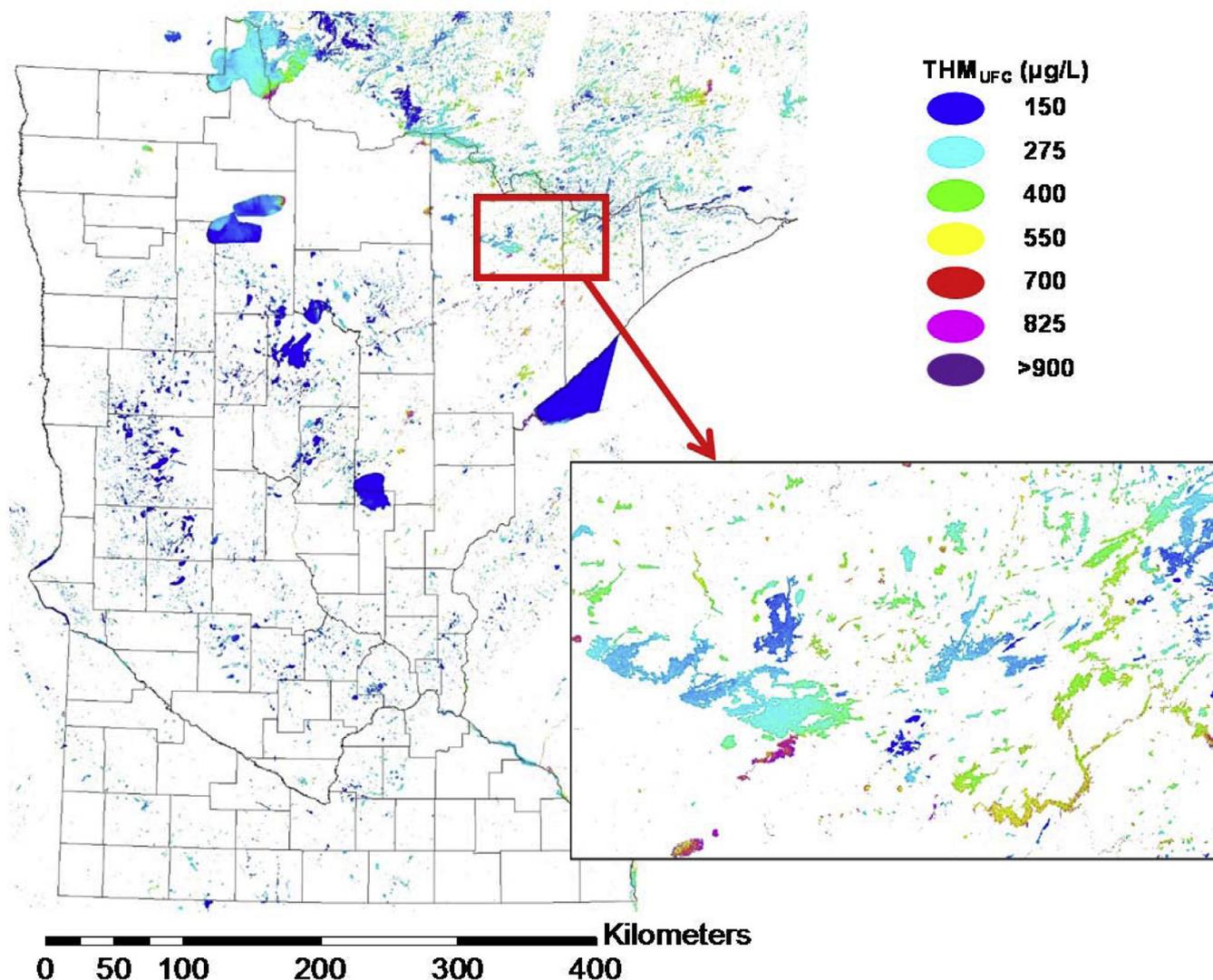


Fig. 4. Distribution of estimated lake THM_{UFC} concentrations in Minnesota lakes based on 2016 and 2017 Landsat 8 images. A zoom in of the Ely area in northeast Minnesota is shown in the inset figure.

Based on the USEPA Stage 1 Disinfectants and Disinfection Byproducts Rule, maximum contaminant levels for THM₄ and HAA₅ are 80 µg/L and 60 µg/L, respectively. As our DBP_{UFC} values are for raw lake water, lake-averaged CDOM values from the statewide CDOM mapping were used to estimate the percentage of lakes that would violate either or both abovementioned MCLs by assuming three potential levels of DBP precursor removal (25% (poor removal), 50% (conventional treatment), and 75% (enhanced coagulation)); Fig. 6 and Table S7). Consequently, THM₄ concentrations in raw waters would be limited to 107, 160, and 320 µg/L, respectively; while for HAA₅ the limits would be 80, 120, and 240 µg/L, respectively. None of Minnesota's lakes ($n = 11,690$) would be expected to meet the THM₄ MCL assuming 25% precursor removal, but 62.4% of Minnesota lakes would meet the MCL assuming 75% precursor removal. The percentage of Minnesota lakes that would be expected to meet the HAA₅ MCL assuming 25% precursor removal was only 11.3% but increased to 21.8% for 75% removal. Finally, the percentage of Minnesota lakes that would be expected to meet both DBP MCLs was only 21.8% for 75% precursor removal. It is important to note, however, that this estimation considered only the number of lakes and not lake area or volume.

Referring to the maps showing the distribution of DBP_{UFC} levels (Fig. 4 and Fig. S9), most lakes in some regions of Minnesota, such as the northeast, are not likely to be useful as drinking water sources because of their high color and potential to form high levels of DBPs even when reasonable precursor removals are assumed.

4. Discussion

4.1. CDOM as a predictor of chlorine demand DBP formation

Historically, predictive models for DBP formation include parameters such as DOC, UV₂₅₄, and SUVA₂₅₄, which are commonly used surrogates for DOM quantity and composition (Chowdhury et al., 2009; Sadiq and Rodriguez, 2004). Other parameters include pH, water temperature, chlorine dose, and reaction time. Most modeling efforts have focused on THM formation with considerably less work on HAA formation (Chowdhury et al., 2009; Sadiq and Rodriguez, 2004). For example (Singer and Chang, 1989), developed linear relationships between THM formation and UV₂₅₄. Other researchers have developed multivariate models to relate DBP concentrations to a combination of DOC and other variables

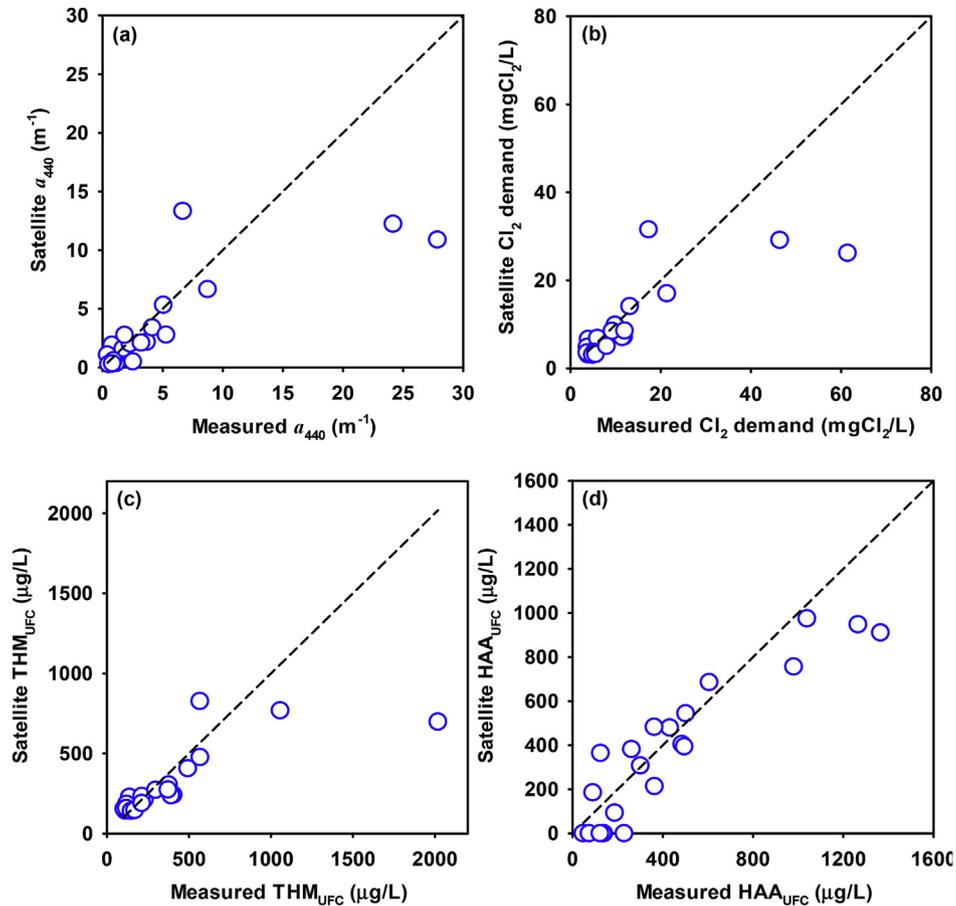


Fig. 5. Comparison of estimated (satellite) and measured values of (a) a_{440} , (b) chlorine demand; (c) THM formation potential (THM_{UFC}); and (d) HAA formation potential (HAA_{UFC}). The dashed line is a 1:1 line. The results for the Mississippi River and Upper Red #2 (at the mouth of the Tamarac River) were excluded from the plots because the riverine flows create highly dynamic conditions not suitable for comparisons with relatively infrequent satellite imagery. Negative HAA_{UFC} values were set to 0.

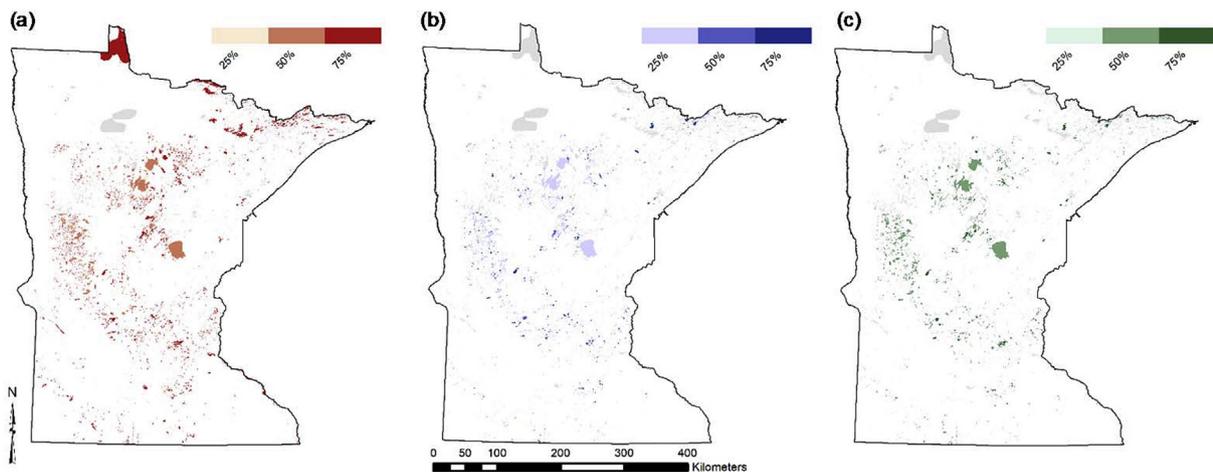


Fig. 6. Distribution of Minnesota lakes expected to meet the (a) THM_4 MCL; (b) HAA_5 MCL; (c) THM_4 and HAA_5 MCL at three potential levels of DBP precursor removal (25%, 50%, and 75%). The grey color shows the rest of lakes that violate either or both MCLs. Calculations were made based on the average of 2015 and 2016 Landsat derived a_{440} measurements, and the developed models correlating a_{440} and DBP formation potentials (THM_{UFC} and HAA_{UFC}).

including water temperature, bromide level, chlorine dose, and contact time (Amy et al., 1987; Sérodes et al., 2003). Such empirical models often have limited applicability given the narrow range of water qualities used for model development. DOC is a broad indicator of the concentration of DOM whereas UV_{254} represents

specific structural characteristics and functional groups (Croue et al., 2000; Edzwald et al., 1985). $SUVA_{254}$ provides a relative proxy for the aromaticity of DOM, and as a result, its feasibility at predicting DBP formation highly depends on the relative importance of aromatic DOM in reacting with chlorine to form DBPs.

Unfortunately, neither DOC nor UV_{254} can be obtained directly from satellite observations. Although UV_{254} is an optical property that provides a sensitive measure of DOM content, UV irradiation is so strongly absorbed by the Earth's atmosphere that very little is reflected back to orbiting satellites.

CDOM (a_{440}), in contrast, is readily obtained from satellite observations because light in the visible and longer wavelength ranges is not as strongly absorbed by the atmosphere yet is sufficiently absorbed by CDOM to result in predictable shifts in the reflectance spectra. CDOM in the limnology literature refers to the color or light absorptivity of dissolved substances in water, which has implications for growth of phytoplankton and submerged vegetation (Gallegos, 2001; Geider, 1987), lake trophic status (Brezonik et al., 2005), and biogeochemical cycling of elements (Zepp et al., 2007). As such, CDOM is an important variable affecting lake and river ecology. The color of drinking water is important as an aesthetic parameter with a color maximum of 15 mg Pt/L (as measured by the cobalt-platinum method) in the USA (EPA, 2018) and generally correlates with light absorptivity.

We are unaware, however, of any reports in the literature regarding models for predicting DBP formation that use CDOM (a_{440}) as a proxy. The results of this study demonstrated strong correlations between a_{440} and THM_{UFC} , di- HAA_{UFC} and tri- HAA_{UFC} ; indicating that CDOM can serve as a predictor of DBP formation. This has important implications for the use of satellite remote sensing to assess the suitability and treatability of surface waters as potential drinking water sources.

Unfortunately, mono- HAA_{UFC} did not correlate strongly with a_{440} , indicating that a_{440} alone may not be adequate for predicting mono-HAA formation in source waters. Total HAA_5 , however, is reasonably well predicted using a_{440} ($R^2 = 0.85$; Fig. S7a) because mono- HAA_{UFC} accounted for less than 26% of the total HAA_5 pool. Thus, using this correlation between a_{440} and HAA_5 for HAA_{UFC} mapping is possible. Despite the issues with mono-HAA, CDOM is a reasonably good predictor of the formation of regulated DBPs upon chlorination.

4.2. DBP formation and correlations with UV_{254} and DOC

Reported values for THM_{UFC} and HAA_{UFC} typically range from 3 to 388 $\mu\text{g/L}$, but values as high as 1300 $\mu\text{g/L}$ have been observed in a few cases (White et al., 2003; Zeng and Arnold, 2013). Additionally, values for THM and HAA formation potentials (THM_{FP} and HAA_{FP}) ranging from 20 to 450 $\mu\text{g/L}$ have been reported (Bougeard et al., 2010; Gang et al., 2002; Hua and Reckhow, 2007; Lee et al., 2007; Summers et al., 1996; Zhao et al., 2016). DBP_{FP} is expected to be greater than DBP_{UFC} due to the higher chlorine doses and longer contact times in the formation potential test. Reported chlorine demand values also exhibit a wide range (2.0–108 $\text{mg Cl}_2/\text{L}$) (Gang et al., 2002; Zeng and Arnold, 2013; Zhao et al., 2016). The chlorine demand and DBP_{UFC} values determined in the present study fall into the above reported ranges, with the exception of a few waters containing DOC concentrations that were much greater than surface waters routinely used as drinking water supplies, including South Sturgeon Lake, Lake Vermilion (Pike Bay) and Lake of the Woods (Muskeg Bay). When the results were normalized to DOC to obtain specific chlorine demand and specific DBP_{UFC} , values from the present study (specific chlorine demand: 0.56 ± 0.01 to 1.70 ± 0.01 $\text{mg Cl}_2/\text{mg C}$, specific THM_{UFC} : 15.9 ± 0.95 to 55.9 ± 0.10 $\mu\text{g}/\text{mg C}$, specific HAA_{UFC} : 7.22 ± 2.98 to 76.0 ± 2.19 $\mu\text{g}/\text{mg C}$) were comparable with previously published values (specific chlorine demand: 0.5–1.8 $\text{mg Cl}_2/\text{mg C}$, specific DBP_{UFC} : 2.5–106 $\mu\text{g}/\text{mg C}$).

Previously reported correlations of chlorine demand and DBP formation potential with DOC, UV_{254} , and $SUVA_{254}$ were

summarized (Table S6) to allow comparison with our findings. The slope of chlorine demand versus DOC in the present study (1.75 ± 0.06 $\text{mg Cl}_2/\text{mg C}$) was greater than previously published values (0.56–1.05 $\text{mg Cl}_2/\text{mg C}$) (Gang et al., 2002; Zeng and Arnold, 2013). The slope of the THM_{UFC} versus DOC plot in the present study (52.5 ± 3.67 $\mu\text{g}/\text{mg C}$) was also somewhat greater than values reported for raw waters (30–36 $\mu\text{g}/\text{mg C}$) (Gang et al., 2002; Summers et al., 1996), but much greater than that for treated (coagulation/filtration/ozone/GAC) waters (0.038 $\mu\text{g}/\text{mg C}$) (Bougeard et al., 2010). The slope of HAA_{UFC} versus UV_{254} observed in the present study was similar to previously reported values. The slope of HAA_{UFC} versus DOC in the present study, however, was greater than previously reported values, which is likely due to the different DBP precursors present in the tested waters. Thus, it appears that organic matter in the lakes of this study was often more reactive with free chlorine than those in previous investigations. Nevertheless, other differences in experimental conditions, such as chlorine contact time, pH, or water temperature, also could have played a role in the observed differences in specific chlorine demand and specific DBP_{UFC} .

4.3. Factors affecting DBP formation

Previous research has suggested that the hydrophobic or aromatic fractions of DOM are primarily responsible for the formation of THMs and HAAs (Bond et al., 2012; Kitis et al., 2002; Pan and Zhang, 2013; Zhai and Zhang, 2011). For example, humic substances (i.e., fulvic and humic acids), generally representing the major fraction of DOM in surface waters, are the main THM and HAA precursors (Leenheer and Croué, 2003; Reckhow and Singer, 1985). Specifically, 1,3-dihydroxybenzene structures, as phenolic aromatic moieties, preferentially generate THMs (Rook, 1977). Phenolic groups also are important TCAA precursors (Bond et al., 2009; Dickenson et al., 2008). Earlier research on chlorination of DOM fractions isolated from whole water samples indicated that DCAA precursors are more hydrophilic and lower in MW than TCAA precursors (Hua and Reckhow, 2007; Kanokkantapong et al., 2006; Liang and Singer, 2003), and aliphatic β -dicarbonyl moieties are key precursors in DCAA formation (Bond et al., 2009; Bull et al., 2006; Chang et al., 2006; Dickenson et al., 2008). Likewise, SUVA was an effective predictor for specific THMs and tri- HAA_{UFC} in our study suggesting that aromatic structures, such as phenolic groups, are the primary precursor of THMs and tri-HAAs (Singer et al., 2002). The weaker relationship observed between SUVA and specific di- HAA_{UFC} indicated that the di-HAA precursors are less aromatic, and aliphatic carbon may play a more important role in their formation than in that of tri-HAAs. SUVA was shown to be a poor indicator of mono-HAA formation, suggesting that non-humic substances are important precursors of mono-HAAs. Our findings lend further support to the finding that the mono-, di-, and tri-HAAs have different precursors and formation mechanisms upon chlorination.

Besides the nature and concentrations of DOM, the underlying causes of differences in speciation of DBPs in different waters could also include bromide ion concentration, chlorine dose and chlorination conditions (e.g., pH, contact time) (Panyapinyopol et al., 2005; Symons et al., 1996; Xie et al., 2006; Zhao et al., 2006). Bromide concentrations above 0.1 mg/L trigger concern about brominated DBP formation (Zhang et al., 2011) due to the formation of HOBr and its ability to react with DBP precursors. In our study, bromide was < LOD in most lakes, with a maximum of 0.066 mg/L in Lake Vermilion (Pike Bay), effectively limiting the generation of brominated DBPs. In addition, previous publications have shown that pH affects HAA and THM formation in opposite ways; THM formation increases with increasing pH, but HAA formation decreases with increasing pH (Singer et al., 1995).

4.4. Implications for use of satellite remote sensing for assessment of water treatability

The generally strong correlations between CDOM and chlorine demand or DBP_{UFC} suggests that satellite remote sensing is a useful technology for assessing the treatability of low-to moderately-colored surface waters with regard to chlorine demand and DBP formation. Furthermore, it might be possible to elucidate long-term trends in surface water treatability measures impacted by CDOM, such as chlorine demand, using historical Landsat imagery. Satellite remote sensing could also prove valuable in assessing other water treatability metrics, such as coagulant demand and membrane fouling, if the metrics are strongly linked to CDOM levels.

DBP_{UFC} values for two highly colored waters computed from satellite assessed CDOM values, however, were severely underestimated. High CDOM ($a_{440} > 11 \text{ m}^{-1}$) results in strong light absorption with very little reflected light reaching the satellite sensors. The resulting weak signal leads to substantial errors in predicted CDOM. To date, most investigations on satellite remote sensing have focused on CDOM levels up to 10 m^{-1} (Kutser et al., 2005a, 2005b). Olmanson et al. (2016), however, reported models that considered lakes with high CDOM levels (a_{440} up to 18 m^{-1}). The models showed an increase in uncertainty at higher a_{440} , which is consistent with our results. In addition, other factors such as interference from suspended solids could also affect satellite CDOM values. Hence, our approach for assessing DBP_{UFC} from satellite imagery should be used with caution for high CDOM and optically complex waters.

The chlorine demand and DBP formation experiments in this study were performed on raw lake water samples and the maps created reflect these raw water values. Certainly, the characteristics of organic matter and specific precursors present in the raw waters in this study will differ from treated waters that have undergone organic matter removal via coagulation/flocculation. Coagulation processes can remove 50% or more of the DOC with even greater reductions in UV_{254} , suggesting that coagulation preferentially removes higher MW and more aromatic DOM (Chowdhury et al., 2009). Further, the removal of DOM increases the bromide to organic carbon ratio, which can shift the distribution of DBPs toward more brominated species (Chowdhury et al., 2009). Filtered, but untreated, surface waters were used in our UFC tests because models have been developed to relate CDOM to light reflectance values from lake surfaces obtained by satellites. Thus, maps showing predictions of Minnesota lakes likely to meet DBP MCLs were prepared assuming various levels of precursor removal (25%, 50%, and 75%) that could be obtained with conventional treatment or enhanced coagulation. The resulting state-wide assessments suggested that a very low percentage of Minnesota lakes, especially in the heavily forested NLF ecoregion, would meet both THM and HAA MCLs unless aggressive pre-disinfection treatment was done to remove at least 75% of DBP precursors.

Finally, another potential limitation of this work is that satellites are measuring near surface parameters while water utilities often withdraw water from greater depths, such as below the thermocline in stratified lakes. The effect of this discrepancy in where measurements are made in the water body and where the water is withdrawn on predictions of CDOM and CDOM-related parameters is uncertain, but believed to be relatively minor, compared to other water quality parameters that exhibit strong depth dependence such as chlorophyll *a*. Vertical profiles of CDOM in a few colored Minnesota lakes showed small to moderate changes with depth (Brezonik et al., 2019).

5. Conclusions

Twenty-four surface water samples were collected throughout

Minnesota, USA with wide ranging values of CDOM (a_{440} ; $0.41\text{--}27.9 \text{ m}^{-1}$), dissolved organic carbon (DOC; $5.5\text{--}47.6 \text{ mg/L}$) and specific ultraviolet absorbance at 254 nm ($SUVA_{254}$; $1.3\text{--}5.1 \text{ L/mg-M}$). Laboratory experiments were performed to quantify chlorine demand and the formation of two classes of halogenated disinfection byproducts (DBPs), trihalomethanes (THMs) and haloacetic acids (HAAs), using the uniform formation conditions (UFC) test. The correlations relating chlorine demand and DBP_{UFC} values with CDOM were coupled with satellite CDOM assessments to estimate chlorine demand, THM_{UFC} , and HAA_{UFC} values for all surface waters larger than 0.05 km^2 in the state of Minnesota, USA. The main conclusions from this work are as follows:

- Chlorine demand and THM_{UFC} correlated linearly with CDOM (a_{440}) and di- and tri- HAA_{UFC} correlated logarithmically with CDOM when all sampled waters were considered. For low-to moderately-colored waters ($a_{440} \leq 11 \text{ m}^{-1}$), a linear relationship provided a good fit ($R^2 \geq 0.73$) for all of these cases and a moderate fit ($R^2 = 0.46$) for mono- HAA_{UFC} versus CDOM. Thus, CDOM (a_{440}) appears to be an effective predictor of chlorine demand, THM_{UFC} , and HAA_{UFC} .
- CDOM (a_{440}), chlorine demand, THM_{UFC} , and HAA_{UFC} did not correlate with conventional water quality indicators including TSS and chlorophyll *a*.
- For highly colored surface waters ($a_{440} > 11 \text{ m}^{-1}$), CDOM and DBP_{UFC} values were substantially underestimated from the satellite data due to various factors, including strong absorption of incident light resulting in a weak signal received by the satellite sensors.
- Overall, the results demonstrate that satellite remote sensing could be a useful tool for assessing water treatability metrics like chlorine demand and DBP formation, especially for low-to moderately-colored lakes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge funding from the State of Minnesota's Environment and Natural Resources Trust Fund (ENRTF) and National Science Foundation (CBET-1510332). The authors thank Noah Germolus for assistance with bromide concentration measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.watres.2019.115001>.

References

- Abdullah, M.P., Yee, L.F., Ata, S., Abdullah, A., Ishak, B., Abidin, K.N.Z., 2009. The study of interrelationship between raw water quality parameters, chlorine demand and the formation of disinfection by-products. *Phys. Chem. Earth* 34, 806–811. <https://doi.org/10.1016/j.pce.2009.06.014>.
- Amy, G.L., Chadik, P.A., Chowdhury, Z.K., 1987. Developing models for predicting trihalomethane formation potential and kinetics. *Journal-American Water Work. Assoc.* 79, 89–97.
- Amy, G.L., Siddiqui, M., Ozekin, K., Zhu, H.W.H.W., Wang, C., 1998. Empirical Based Models for Predicting Chlorination and Ozonation Byproducts: Haloacetic Acids, Chloral Hydrate, and Bromate. Report CX 819579. U.S. Environmental Protection Agency, Cincinnati, Ohio.
- Andrews, R.C., Alam, Z., Hofmann, R., Lachuta, L., Cantwell, R., Andrews, S., Moffet, E., Gagnon, G.A., Rand, J., Chauret, C., 2005. Impact of Chlorine Dioxide

- on Transmission, Treatment, and Distribution System Performance. American Water Works Association.
- Beggs, K.M.H., Summers, R.S., McKnight, D.M., 2009. Characterizing chlorine oxidation of dissolved organic matter and disinfection by-product formation with fluorescence spectroscopy and parallel factor analysis. *J. Geophys. Res. Biogeosciences* 114.
- Bellar, T.A., Lichtenberg, J.J., Kroner, R.C., 1974. The occurrence of organohalides in chlorinated drinking waters. *Journal-American Water Work. Assoc.* 66, 703–706.
- Bond, T., Goslan, E.H., Parsons, S.A., Jefferson, B., 2012. A critical review of trihalomethane and haloacetic acid formation from natural organic matter surrogates. *Environ. Technol. Rev.* 1, 93–113. <https://doi.org/10.1080/09593330.2012.705895>.
- Bond, T., Henriot, O., Goslan, E.H., Parsons, S.A., Jefferson, B., 2009. Disinfection byproduct formation and fractionation behavior of natural organic matter surrogates. *Environ. Sci. Technol.* 43, 5982–5989.
- Bougeard, C.M.M., Goslan, E.H., Jefferson, B., Parsons, S.A., 2010. Comparison of the disinfection by-product formation potential of treated waters exposed to chlorine and monochloramine. *Water Res.* 44, 729–740.
- Brezonik, P., Menken, K.D., Bauer, M., 2005. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). *Lake Reservoir Manag.* 21, 373–382. <https://doi.org/10.1080/07438140509354442>.
- Brezonik, P.L., Bouchard, R.W., Finlay, J.C., Griffin, C.G., Olmanson, L.G., Anderson, J.P., Arnold, W.A., Hozalski, R.M., 2019. Color, chlorophyll a, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. *Ecol. Appl.* 29, e01871. <https://doi.org/10.1002/eap.1871>.
- Brezonik, P.L., Olmanson, L.G., Finlay, J.C., Bauer, M.E., 2015. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sens. Environ.* 157, 199–215. <https://doi.org/10.1016/j.rse.2014.04.033>.
- Brinkman, B.M., Hozalski, R.M., 2011. Temporal variation of NOM and its effects on membrane treatment. *Journal-American Water Work. Assoc.* 103, 98–106.
- Bull, R., Reckhow, D., Rotello, V., Bull, O.M., Kim, J., 2006. Use of Toxicological and Chemical Models to Prioritize DBP Research, Water Intelligence Online. <https://doi.org/10.1016/j.jns.2013.07.1872>.
- Chang, E.E., Chao, S., Chiang, P., Lee, J., 1996. Effects of chlorination on THMs formation in raw water. *Toxicol. Environ. Chem.* 56, 211–225.
- Chang, E.E., Chiang, P.C., Chao, S.H., Lin, Y.L., 2006. Relationship between chlorine consumption and chlorination by-products formation for model compounds. *Chemosphere* 64, 1196–1203.
- Chellam, S., Krasner, S.W., 2001. Disinfection byproduct relationships and speciation in chlorinated nanofiltered waters. *Environ. Sci. Technol.* 35, 3988–3999. <https://doi.org/10.1021/es010775n>.
- Chen, J., Zhu, W.N., Tian, Y.Q., Yu, Q., 2017. Estimation of colored dissolved organic matter from Landsat-8 imagery for complex inland water: case study of Lake Huron. *IEEE Trans. Geosci. Remote Sens.* 55, 2201–2212. <https://doi.org/10.1109/TGRS.2016.2638828>.
- Chowdhury, S., Champagne, P., McLellan, P.J., 2009. Models for predicting disinfection byproduct (DBP) formation in drinking waters: a chronological review. *Sci. Total Environ.* 407, 4189–4206. <https://doi.org/10.1016/j.scitotenv.2009.04.006>.
- Croue, J.-P., Korshin, G.V., Benjamin, M.M., 2000. Characterization of Natural Organic Matter in Drinking Water. American Water Works Association.
- Del Castillo, C.E., Miller, R.L., 2008. On the use of ocean color remote sensing to measure the transport of dissolved organic carbon by the Mississippi River Plume. *Remote Sens. Environ.* 112, 836–844. <https://doi.org/10.1016/j.rse.2007.06.015>.
- Dickenson, E.R.V., Summers, R.S., Croué, J.P., Gallard, H., 2008. Haloacetic acid and trihalomethane formation from the chlorination and bromination of aliphatic β -dicarbonyl acid model compounds. *Environ. Sci. Technol.* 42, 3226–3233. <https://doi.org/10.1021/es0711866>.
- Eaton, A.D., Clesceri, L.S., Rice, E.W., Greenberg, A.E., Franson, M., 2005. APHA: Standard Methods for the Examination of Water and Wastewater. Centen. In: APHA, AWWA (Eds.). WEF, Washington, DC.
- Edzwald, J.K., Becker, W.C., Wattier, K.L., 1985. Surrogate parameters for monitoring organic matter and THM precursors. *Journal-American Water Work. Assoc.* 77, 122–132.
- Eikebrokk, B., Vogt, R.D., Liltved, H., 2004. NOM increase in Northern European source waters: discussion of possible causes and impacts on coagulation/contact filtration processes. *Water Sci. Technol. Water Supply* 4, 47–54.
- EPA, 2018. 2018 Drinking Water Standards and Advisory.
- Fichot, C.G., Downing, B.D., Bergamaschi, B.A., Windham-Myers, L., Marvin-DiPasquale, M., Thompson, D.R., Gierach, M.M., 2015. High-resolution remote sensing of water quality in the San Francisco Bay–Delta estuary. *Environ. Sci. Technol.* 50, 573–583.
- Gallegos, C.L., 2001. Calculating optical water quality targets to restore and protect submersed aquatic vegetation: overcoming problems in partitioning the diffuse attenuation coefficient for photosynthetically active radiation. *Estuaries* 24, 381–397.
- Gang, D.D., Segar Jr., R.L., Clevenger, T.E., Banerji, S.K., 2002. Using chlorine demand to predict TTHM and HAA9 formation. *Journal-American Water Work. Assoc.* 94, 76–86.
- Geider, R.J., 1987. Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: implications for physiology and growth of phytoplankton. *New Phytol.* 106, 1–34.
- Gerecke, A.C., Canonica, S., Muller, S.R., Scharer, M., Schwarzenbach, R.P., 2001. Quantification of dissolved natural organic matter (DOM) mediated photo-transformation of phenylurea herbicides in lakes. *Environ. Sci. Technol.* 35, 3915–3923. <https://doi.org/10.1021/es010103x>.
- Griffin, C.G., Finlay, J.C., Brezonik, P.L., Olmanson, L., Hozalski, R.M., 2018. Limitations on using CDOM as a proxy for DOC in temperate lakes. *Water Res.* 144, 719–727. <https://doi.org/10.1016/j.watres.2018.08.007>.
- Herzprung, P., von Tümpling, W., Hertkorn, N., Harir, M., Büttner, O., Bravidor, J., Friese, K., Schmitt-Kopplin, P., 2012. Variations of DOM quality in inflows of a drinking water reservoir: linking of van Krevelen diagrams with EEMF spectra by rank correlation. *Environ. Sci. Technol.* 46, 5511–5518.
- Houser, J.N., 2006. Water color affects the stratification, surface temperature, heat content, and mean epilimnetic irradiance of small lakes. *Can. J. Fish. Aquat. Sci.* 63, 2447–2455.
- Hua, G., Reckhow, D.A., 2007. Characterization of disinfection byproduct precursors based on hydrophobicity and molecular size. *Environ. Sci. Technol.* 41, 3309–3315.
- Kanokkantapong, V., Marhaba, T.F., Panyapinyophol, B., Pavasant, P., 2006. FTIR evaluation of functional groups involved in the formation of haloacetic acids during the chlorination of raw water. *J. Hazard Mater.* 136, 188–196.
- Kitis, M., Karanfil, T., Wigton, A., Kilduff, J.E., 2002. Probing reactivity of dissolved organic matter for disinfection by-product formation using XAD-8 resin adsorption and ultrafiltration fractionation. *Water Res.* 36, 3834–3848.
- Krasner, S.W., Weinberg, H.S., Richardson, S.D., Pastor, S.J., Chinn, R., Scimanti, M.J., Onstad, G.D., Thurston, A.D., 2006. Occurrence of a new generation of disinfection byproducts. *Environ. Sci. Technol.* 40, 7175–7185.
- Kraus, T.E.C., Anderson, C.A., Morgenstern, K., Downing, B.D., Pellerin, B.A., Bergamaschi, B.A., 2010. Determining sources of dissolved organic carbon and disinfection byproduct precursors to the McKenzie River, Oregon. *J. Environ. Qual.* 39, 2100–2112.
- Kutser, T., 2012. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sens. Environ.* 123, 334–338.
- Kutser, T., Pierson, D.C., Kallio, K.Y., Reinart, A., Sobek, S., 2005a. Mapping lake CDOM by satellite remote sensing. *Remote Sens. Environ.* 94, 535–540. <https://doi.org/10.1016/j.rse.2004.11.009>.
- Kutser, T., Pierson, D.C., Tranvik, L., Reinart, A., Sobek, S., Kallio, K., 2005b. Using satellite remote sensing to estimate the colored dissolved organic matter absorption coefficient in lakes. *Ecosystems* 8, 709–720. <https://doi.org/10.1007/s10021-003-0148-6>.
- Ledesma, J.L.J., Köhler, S.J., Futter, M.N., 2012. Long-term dynamics of dissolved organic carbon: implications for drinking water supply. *Sci. Total Environ.* 432, 1–11. <https://doi.org/10.1016/j.scitotenv.2012.05.071>.
- Lee, W., Westerhoff, P., Croué, J.-P., 2007. Dissolved organic nitrogen as a precursor for chloroform, dichloroacetonitrile, N-nitrosodimethylamine, and trichloronitromethane. *Environ. Sci. Technol.* 41, 5485–5490.
- Leenheer, J.A., Croué, J.-P., 2003. Characterizing aquatic dissolved organic matter. *Environ. Sci. Technol.* 37 (1), 19A–26A. <https://doi.org/10.1021/es032333c>.
- Li, C.-W., Chen, Y.-S., 2004. Fouling of UF membrane by humic substance: effects of molecular weight and powder-activated carbon (PAC) pre-treatment. *Desalination* 170, 59–67.
- Liang, L., Singer, P.C., 2003. Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. *Environ. Sci. Technol.* 37, 2920–2928.
- Liang, S., Song, L., Tao, G., Kekre, K.A., Seah, H., 2006. A modeling study of fouling development in membrane bioreactors for wastewater treatment. *Water Environ. Res.* 78, 857–864.
- McKnight, D.M., Bencala, K.E., 1990. The chemistry of iron, aluminum, and dissolved organic material in three acidic, metal-enriched, mountain streams, as controlled by watershed and in-stream processes. *Water Resour. Res.* 26, 3087–3100.
- Najm, I.N., Patania, N.L., Jacangelo, J.G., Krasner, S.W., 1994. Evaluating surrogates for disinfection by-products. *J. Am. Water Work. Assoc.* 86, 98–106.
- Nuckols, J., Rossman, L.A., Singer, P.C., 2001. Development of Exposure Assessment Methods for THM and HAA in Water Distribution Systems. American Water Works Association.
- Olmanson, L.G., Brezonik, P.L., Finlay, J.C., Bauer, M.E., 2016. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sens. Environ.* 185, 119–128. <https://doi.org/10.1016/j.rse.2016.01.007>.
- Pan, Y., Zhang, X., 2013. Four groups of new aromatic halogenated disinfection byproducts: effect of bromide concentration on their formation and speciation in chlorinated drinking water. *Environ. Sci. Technol.* 47, 1265–1273.
- Panyapinyophol, B., Marhaba, T.F., Kanokkantapong, V., Pavasant, P., 2005. Characterization of precursors to trihalomethanes formation in Bangkok source water. *J. Hazard Mater.* 120, 229–236.
- Reckhow, D.A., Singer, P.C., 1985. Mechanisms of organic halide formation during fulvic acid chlorination and implications with respect to preozonation. *Water Chlorination Chem. Environ. Impact Heal. Eff.* 5, 1229–1257.
- Rodriguez, M.J., Sérodes, J., Morin, M., 2000. Estimation of water utility compliance with trihalomethane regulations using a modelling approach. *J. Water Supply Res. Technol.* 49, 57–73.
- Rook, J.J., 1977. Chlorination reactions of fulvic acids in natural waters. *Environ. Sci. Technol.* 11, 478–482.
- Sadiq, R., Rodriguez, M.J., 2004. Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review. *Sci. Total Environ.* 321,

- 21–46.
- Sedlak, D.L., von Gunten, U., 2011. The chlorine dilemma. *Science* 331 (80), 42–43.
- Sérodes, J.-B., Rodriguez, M.J., Li, H., Bouchard, C., 2003. Occurrence of THMs and HAAs in experimental chlorinated waters of the Quebec City area (Canada). *Chemosphere* 51, 253–263.
- Shank, G.C., Zepp, R.G., Whitehead, R.F., Moran, M.A., 2005. Variations in the spectral properties of freshwater and estuarine CDOM caused by partitioning onto river and estuarine sediments. *Estuar. Coast Shelf Sci.* 65, 289–301.
- Sharp, E.L., Parsons, S.A., Jefferson, B., 2006. Seasonal variations in natural organic matter and its impact on coagulation in water treatment. *Sci. Total Environ.* 363, 183–194.
- Singer, P.C., Chang, S.D., 1989. Correlations between trihalomethanes and total organic halides formed during water treatment. *Journal-American Water Work. Assoc.* 81, 61–65.
- Singer, P.C., Obolensky, A., Greiner, A., 1995. DBPs in chlorinated North Carolina drinking waters. *J. Am. Water Work. Assoc.* 87, 83–92.
- Singer, P.C., Weinberg, H.S., Brophy, K., 2002. Relative Dominance of Haloacetic Acids and Trihalomethanes in Treated Drinking Water. American Water Works Association.
- Sohn, J., Amy, G., Cho, J., Lee, Y., Yoon, Y., 2004. Disinfectant decay and disinfection by-products formation model development: chlorination and ozonation by-products. *Water Res.* 38, 2461–2478.
- Sommaruga, R., 2001. The role of solar UV radiation in the ecology of alpine lakes. *J. Photochem. Photobiol. B Biol.* 62, 35–42.
- Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G.H., Neil, C., Barbosa, C.C.F., Binding, C.E., Bradt, S., 2018. Optical types of inland and coastal waters. *Limnol. Oceanogr.* 63, 846–870.
- Summers, R.S., Hooper, S.M., Shukairy, H.M., Solarik, G., Owen, D., 1996. Assessing DBP yield: uniform formation conditions. *Journal-American Water Work. Assoc.* 88, 80–93.
- Symons, J.M., Krasner, S.W., Scimmenti, M.J., Simms, L.A., Sorensen Jr., H.W., Spietel Jr., G.E., Diehl, A.C., 1996. Influence of bromide ion on trihalomethane and haloacetic acid formation. *Disinfect. by-products water Treat. Chem. their Form. Control* 91–130.
- Thrane, J.-E., Hessen, D.O., Andersen, T., 2014. The absorption of light in lakes: negative impact of dissolved organic carbon on primary productivity. *Ecosystems* 17, 1040–1052.
- Thurman, E.M., 2012. *Organic Geochemistry of Natural Waters*. Springer Science & Business Media.
- Van Leeuwen, J., Daly, R., Holmes, M., 2005. Modeling the treatment of drinking water to maximize dissolved organic matter removal and minimize disinfection by-product formation. *Desalination* 176, 81–89.
- Watson, M., American Water Works Association, Water Industry Technical Action Fund (U.S.), 1993. *Mathematical Modeling of the Formation of THMs and HAAs in Chlorinated Natural Waters, Final Report*. Water Industry Technical Action Fund. American Water Works Association.
- White, D.M., Garland, D.S., Narr, J., Woolard, C.R., 2003. Natural organic matter and DBP formation potential in Alaskan water supplies. *Water Res.* 37, 939–947.
- Xie, H., Aubry, C., Bélanger, S., Song, G., 2012. The dynamics of absorption coefficients of CDOM and particles in the St. Lawrence estuarine system: biogeochemical and physical implications. *Mar. Chem.* 128, 44–56.
- Xie, S., Wen, D., Shi, D., Tang, X., 2006. Reduction of precursors of chlorination by-products in drinking water using fluidized-bed biofilm reactor at low temperature. *Biomed. Environ. Sci.* 19, 360.
- Yang, Z., Gao, B., Yue, Q., 2010. Coagulation performance and residual aluminum speciation of Al₂(SO₄)₃ and polyaluminum chloride (PAC) in Yellow River water treatment. *Chem. Eng. J.* 165, 122–132.
- Yee, L.F., Abdullah, M.P., Ata, S., Ishak, B., 2006. Dissolved organic matter and its impact on the chlorine demand of treated water. *Malaysian J. Anal. Sci.* 10, 243–250.
- Zeng, T., Arnold, W.A., 2013. Clustering chlorine reactivity of haloacetic acid precursors in inland lakes. *Environ. Sci. Technol.* 48, 139–148.
- Zepp, R.G., Erickson Iii, D.J., Paul, N.D., Sulzberger, B., 2007. Interactive effects of solar UV radiation and climate change on biogeochemical cycling. *Photochem. Photobiol. Sci.* 6, 286–300.
- Zhai, H., Zhang, X., 2011. Formation and decomposition of new and unknown polar brominated disinfection byproducts during chlorination. *Environ. Sci. Technol.* 45, 2194–2201.
- Zhang, J., Yu, J., An, W., Liu, J., Wang, Y., Chen, Y., Tai, J., Yang, M., 2011. Characterization of disinfection byproduct formation potential in 13 source waters in China. *J. Environ. Sci.* 23, 183–188.
- Zhao, Y., Yang, H. wei, Liu, S. ting, Tang, S., Wang, X. mao, Xie, Y.F., 2016. Effects of metal ions on disinfection byproduct formation during chlorination of natural organic matter and surrogates. *Chemosphere* 144, 1074–1082. <https://doi.org/10.1016/j.chemosphere.2015.09.095>.
- Zhao, Z.-Y., Gu, J.-D., Fan, X.-J., Li, H.-B., 2006. Molecular size distribution of dissolved organic matter in water of the Pearl River and trihalomethane formation characteristics with chlorine and chlorine dioxide treatments. *J. Hazard Mater.* 134, 60–66.
- Zhu, W., Yu, Q., Tian, Y.Q., Becker, B.L., Zheng, T., Carrick, H.J., 2014. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. *Remote Sens. Environ.* 140, 766–778.
- Zularisam, A.W., Ismail, A.F., Salim, R., 2006. Behaviours of natural organic matter in membrane filtration for surface water treatment—a review. *Desalination* 194, 211–231.