Regional Measurements and Spatial/Temporal Analysis of CDOM in 10,000+ Optically Variable Minnesota Lakes using Landsat 8 Imagery

Leif G. Olmanson[1*], Benjamin P. Page[2], Jacques C. Finlay[3], Patrick L. Brezonik[4], Marvin E. Bauer[1], Claire G. Griffin[3,5], Raymond M. Hozalski[4]

[1]*Department of Forest Resources, University of Minnesota, St. Paul, Minnesota 55108 USA*

[2]*Water Resources Center, University of Minnesota, St. Paul, Minnesota 55108 USA*

[3]*Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, Minnesota 55108 USA*

[4]*Department of Civil, Environmental, and Geo- Engineering, University of Minnesota, Minneapolis, Minnesota 55455 USA*

[5]*Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia 22905 USA*

[*]Corresponding author. E-mail: olman002@umn.edu

Abstract

Information on colored dissolved organic matter (CDOM) is essential for understanding and managing lakes but is often not available, especially in lake-rich regions where concentrations are often highly variable in time and space. We developed remote sensing methods that can use both Landsat and Sentinel satellite imagery to provide census-level CDOM measurements across the state of Minnesota, USA, a lake-rich landscape with highly varied lake, watershed, and climatic conditions. We evaluated the error of satellite derived CDOM resulting from two atmospheric correction methods with in situ data, and found that both provided substantial improvements over previous methods. We applied CDOM models to 2015 and 2016 Landsat 8 OLI imagery to create 2015 and 2016 Minnesota statewide CDOM maps (reported as absorption coefficients at 440 nm, a_{440}) and used those maps to conduct a geospatial analysis at the ecoregion level. Large differences in a_{440} among ecoregions were related to predominant land cover/use; lakes in ecoregions with large areas of wetland and forest had significantly higher CDOM levels than lakes in agricultural ecoregions. We compared regional lake CDOM levels between two years with strongly contrasting precipitation (close-to-normal precipitation year in 2015 and much wetter conditions with large storm events in 2016). CDOM levels of lakes in agricultural ecoregions tended to decrease between 2015 and 2016, probably because of dilution by rainfall, and 7% of lakes in these areas decreased in a_{440} by $\geq$ 3 m^{-1}. In two ecoregions with high forest and wetlands cover, a_{440} increased by more than 3 m^{-1} in 28 and 31% of the lakes, probably due to enhanced transport of CDOM from forested wetlands. With appropriate model tuning and validation, the approach we describe could be extended to other regions, providing a method for frequent and comprehensive measurements of CDOM, a dynamic and important variable in surface waters.

Keywords: Satellite remote sensing**,** water color, atmospheric correction, water quality monitoring, lake management, inland waters

1. Introduction

Research in recent decades has revealed a central role for colored (or chromophoric) dissolved organic matter (CDOM) in regulating major physical, chemical and biological processes in lakes and rivers (e.g., reviewed in Solomon et al. 2015, Williamson et al. 1999, Creed et al. 2018, and elsewhere). We now know that CDOM functions as one of a small number of "master variables," similar to phosphorus, pH and redox potential, that control important aspects of the composition and functioning of aquatic ecosystems and regulate their responses to environmental change (Williamson et al. 1999, Creed et al. 2018). Recent studies show that CDOM levels strongly influence: (a) light and thermal regimes in lakes (e.g., Houser 2006, Ask et al. 2009, Thrane et al. 2014, Pilla et al. 2018, Snucins and Gunn 2000), (b) biogeochemical cycles (e.g., Knoll et al. 2018, Corman et al. 2018), (c) food web processes and interactions (e.g., Karlsson et al. 2009, Solomon et al. 2015), (d) contaminant bioavailability (e.g., Tsui and Finlay 2011), and (e) water clarity (e.g., Brezonik et al. 2019a). Knowledge of the sources, levels, and cycling of CDOM in freshwaters thus is important for aquatic resource management and for predicting the outcomes of environmental change.

Moderate to high levels of CDOM in freshwaters are determined largely by rates of transport from soils and wetlands in surrounding watersheds and thus are affected by a combination of factors related to vegetation and hydrology. The dependency of aquatic CDOM on dynamic external sources, combined with internal production and loss processes in aquatic systems, can lead to high variability of CDOM levels across landscapes and within lakes at time scales of seasons to years (Brezonik et al. 2015, Williamson, et al. 2015). Human-driven changes in temperature, atmospheric chemistry, land use and watershed hydrology also can have strong effects on CDOM (Creed et al. 2018, Finstad et al. 2016, Kritzberg 2017, Stanley et al. 2012, de Wit 2018).

Although CDOM is easily measured in the laboratory, the availability of in situ CDOM data is surprisingly limited relative to its importance, even in states like Minnesota, where monitoring of its > 10,000 surface waters is a major focus of many state, tribal and local agencies. Several recent, large-scale assessments of regional U.S. lake monitoring efforts (Stanley et al. 2012; Ross et al. 2019) showed that far fewer data were available for CDOM and related variables such as DOC compared to nutrients, chlorophyll, and water clarity, despite the strong effects of CDOM on those and other physicochemical variables. The spatial and temporal variation in CDOM in surface waters suggests the need for more CDOM data to improve understanding of drivers and better predict lake responses to stresses ranging from local land-cover changes to global climate change. Some countries with large numbers of CDOM-rich lakes have incorporated routine monitoring of CDOM or a related parameter such as DOC (e.g., Sobek et al. 2007). The relative lack of CDOM data for U.S. lakes (Stanley et al. 2019) may stem from the fact that many monitoring programs initially started in relatively low-CDOM regions but also from the fact that the importance of CDOM as a driver of ecological conditions has been appreciated only recently.

Whatever the cause, the availability of CDOM data remains deficient compared to its importance. Remote sensing using satellite-based sensors could play an important role in providing CDOM data at high temporal and spatial resolution. Recent studies show that the

Landsat sensors (Kutser et al. 2005, Brezonik et al. 2005, Kutser et al. 2009, Olmanson et al. 2016a), and Sentinel-2/MSI sensors (Toming et al. 2016, Chen et al. 2017) can provide such data at scales relevant for inland lakes as small as 4 hectares (ha).

Recent improvements in Earth-observing satellite sensors have expanded the capabilities to measure optically-related water quality characteristics, including CDOM, in lakes (Olmanson et al. 2016a; Tyler et al. 2016, Pahlevan et al. 2019, Page et al. 2019). Specifically, the Landsat 8 Operational Land Imager (L8/OLI) and the European Space Agency (ESA) Sentinel-2 MultiSpectral Imager (S2/MSI) have improved spatial, spectral, radiometric and temporal resolution compared with earlier sensors. With the L8/OLI and S2/MSI constellation collecting imagery every 3 to 5 days, frequent satellite-based measurements of a variety of key water quality variables on lakes are now possible.

The use of satellite imagery to measure CDOM at large regional scales and over multiple time periods requires analysis of multiple images. Unless ground-based data are available to calibrate each image (a requirement difficult to achieve), accurate methods are needed for atmospheric correction of images to produce surface reflectance data directly representative of optical signals from waterbodies. Although various approaches have been reported to accomplish this (e.g., Pahlevan et al. 2017a,b and Vanhellemont and Ruddick 2015, 2016), we have found that many of them yield unreliable results for inland lakes (Olmanson et al. 2011, Page et al. 2019). The recent availability of surface reflectance products from the EROS Center appears to have overcome this obstacle for Landsat 8 imagery (Kuhn et al. 2019), and Page et al. (2019) described a workflow process to atmospherically correct and harmonize S2/MSI and L8/OLI satellite imagery in Google's Earth Engine (GEE) (Gorelick et al. 2017).

This paper describes application of these advances to measure CDOM on all waterbodies larger than 4 ha across a large geographic region (the state of Minnesota) that encompasses more than 226,000 km^2 and contains officially 11,842 lakes 4 ha or larger in area (https://www.dnr.state.mn.us/faq/mnfacts/water.html). The paper describes a robust semi-empirical approach for routine monitoring of CDOM using L8/OLI imagery. We demonstrate the consistency and reliability of two atmospheric correction methods to generate remote sensing reflectance (R_{rs}) products and use these products to assemble a CDOM database on more than 10,500 lakes for both 2015 and 2016. We assess the accuracy of retrieved CDOM data for both low- and high-CDOM waters and summarize distributions of CDOM in Minnesota lakes at the ecoregion level.

2. Methods

2.1 Study area

Minnesota, a large, lake-rich state in the Upper Midwest of the U.S., comprises parts of seven ecoregions (Omernik and Griffith 2014) that differ in land-cover, geology, soils, vegetation and hydrologic conditions (Figure 1). Known popularly as "the land of 10,000 lakes," Minnesota actually has approximately 12,000 waterbodies with surface areas ≥ 4 ha (Olmanson et al. 2014) and many more that are smaller than that. The lakes are distributed broadly (but not uniformly) across the ecoregions. Two ecoregions, the Northern Lakes and Forests (NLF) and North Central Hardwood Forest (NCHF), together comprise 49% of the state's area and contain 84% of the state's lakes (47% and 37%, respectively). According to Olmanson et al. (2014), about one-fourth of the heavily forested NLF (mixed conifers and hardwoods) is wetlands and lakes; only 4% is urban and 7% agricultural land. The high proportion of forest (66%) and wetlands (14%) leads to high CDOM levels in many NLF surface waters (Griffin et al. 2018;

Brezonik et al. 2019a,b). In contrast, half of the NCHF is agricultural land, and about 10% is urban or suburban; forests account for only about 17% of the ecoregion, and wetlands constitute 11% of the landscape.

The Western Corn Belt Plain (WCBP) occupies most of southern Minnesota and is dominated (~ 77%) by row-crop agriculture (mainly corn and soybean); its land-cover is only ~ 7% forested. The Northern Glaciated Plains (NGP) ecoregion occupies a small region of southwest Minnesota and is similar to the WCBP in agricultural land cover (74%) but has a higher percentage of grassland (9%). Together, the WCBP and NGP contain 12% of the state's lakes. The Lake Agassiz Plain (LAP) ecoregion (Omernik and Griffith 2014), formerly called the Red River Valley ecoregion (Omernik 1987), has the highest percentage (84%) of agricultural land among the state's ecoregions, and the flat land is a remnant of glacial Lake Agassiz. This ecoregion has only 215 lakes (2% of the state's total). The Northern Minnesota Wetlands (NMW) ecoregion is contiguous to the NLF and is similarly heavily forested (52%). The NMW has more wetlands (19%), however, and its flat landscape contains few lakes, although three of the state's largest lakes, Upper and Lower Red Lake and Lake of the Woods, are in the NMW. The non-glaciated Driftless Area in southeastern Minnesota has only a few small manmade ponds and reservoirs and backwater areas of the Mississippi River.

2.2 Calibration data

A dataset of ground-based CDOM levels for satellite imagery calibration was developed from our ongoing CDOM studies (e.g., Griffin et al. 2018, Brezonik et al. 2019a,b) and includes data from the Minnesota Pollution Control Agency (MPCA) and several other agencies and collaborators. Sampling in 2015 was focused in the NLF and NCHF in northern Minnesota

(Figure 1) and was expanded to include the NMW ecoregion in 2016 and the WCBP, NGP, and LAP ecoregions in 2017. Most lakes were sampled only once, but a selection of lakes were sampled once each year and a few were sampled approximately monthly in 2016 or 2017. Details of sampling were provided previously (Griffin et al. 2018; Brezonik et al. 2019). All observations (site-date combinations) were treated separately; i.e., multiple samples from a lake were not averaged. A total of 1,586 CDOM measurements were collected over 2015-2018, many from routine monitoring efforts by collaborators (Brezonik et al. 2019a). These efforts provided a large dataset of field measurements for calibration and validation.

Sampling procedures and field and laboratory analyses followed standard limnological practices. Detailed methods were described by Griffin et al. (2018). In brief, most water samples were collected from ~ 0.25 m below the lake surface; the MPCA samples were a 0-2 m integrated sample of the epilimnion. 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. Samples for DOC were acidified using 2 M HCl and stored in pre-ashed 20 mL glass bottles at 4°C. Other samples were stored in acid-washed and triple-rinsed polycarbonate or high-density polyethylene bottles and filtered for analysis of various dissolved constituents within 24 h of collection.

CDOM was determined from absorbance measurements at 440 nm, using a Shimadzu 1601UV-PC dual beam spectrophotometer through 1 or 5 cm quartz cuvettes against a nanopure water blank. Absorbance was converted to Napierian absorption coefficients (Kirk 2010) using:

$$a_{440} = 2.303 \mathrm{A}_{440}/l \qquad (1)$$

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} (m^{-1}).

2.3 Image acquisition and processing

A critical component of image processing for aquatic environments is a consistent atmospheric correction (AC) method that can yield reliable estimates of the surface water-leaving reflectance (ρ_w), an optically active input parameter for various satellite-based water quality models (Gordon and Wang 1994). We evaluated atmospherically corrected L8/OLI remote sensing reflectance ($R_{rs} = \rho_w/\pi$) products derived from the **M**odified **A**tmospheric Correction for **IN**land waters (MAIN) (Page et al. 2019) method implemented in Google Earth Engine (GEE) (Gorelick et al., 2017) to map CDOM in Minnesota lakes. Mean R_{rs} values were extracted from a 50-m buffer around each sample location within the open water area of each lake using a collection of clear imagery from L8/OLI to develop a CDOM retrieval algorithm. Paths of clear L8/OLI imagery with coincident field data from 2015 and 2016 were used for model calibration, and coincident L8/OLI and S2/MSI imagery from 2018 were used with corresponding field data for independent validation of the results. Finally, R_{rs} values from the U.S. Geological Survey Surface Reflectance Product (OLI-SR version 1.3.0) also were evaluated for cross-model comparisons.

2.4 CDOM modeling approach

Because CDOM concentrations in most lakes are stable on at least a short-term basis (days to weeks) (e.g., Brezonik et al. 2015), we used calibration/validation data that had been collected within 30 days of imagery. This yielded 250 calibration measurements corresponding to five clear paths of L8/OLI imagery in 2015 and 2016 (Table 1). An additional 157 measurements from MAIN-processed coincident Landsat 8 and Sentinel-2 imagery for August 13, 2018 were used for independent validation and harmonization of the L8/OLI and S2/MSI sensors (Table 1); 62 of these measurements corresponded with clear L8/OLI imagery and 95 corresponded with clear S2/MSI imagery. The calibration set included lakes distributed across the state with a wide range of CDOM (a_{440} = 0.2-32.5 m^{-1}). The CDOM range in the validation set fit closely with the calibration set at low to moderate CDOM levels (up to a_{440} ~ 10 m^{-1}) but lacked higher values (Table A1) because wildfire smoke (originating in California USA and Canada) caused haze interference in northern Minnesota, where the high CDOM lakes occur, for the August 13, 2018 validation imagery.

To explore the potential of all available OLI bands and band ratios to predict CDOM, modeled as $\ln(a_{440})$, we used the bootstrap forest technique in JMP Pro 14 SAS Institute (2018) and evaluated the most significant combinations. The calibration dataset of measured a_{440} values corresponding with the five clear L8/OLI image paths was used as the dependent variable (Tables 1 and A1), and MAIN-derived (and OLI-SR) mean R_{rs} values for L8/OLI bands B1-B5 and all band-ratio permutations were the independent input variables (26 total terms). The two highest-contributing terms that produced the highest coefficient of determination (R^2) and lowest root mean square error (RMSE) with measured data were identified using step-wise regression and were used to develop the models.

To evaluate model predictive capability, the data were divided into four randomized groups. For each possible combination, three groups were used as a training set to develop a correlation, and the remaining group was used as a confirmation set. Performance of the models generated from the four randomly selected calibration/confirmation datasets was evaluated from the coefficient of determination (R^2) and root mean square error (RMSE) for model-predicted vs. measured a_{440}, and the average and range of performance of the four datasets were calculated.

As an additional check on the consistency of the model over a broader temporal scale and MAIN harmonization of L8/OLI and S2/MSI R_{rs} values, we applied the model derived from L8/OLI imagery to the independent validation datasets described above (Table 1). Accuracy was compared against measured a_{440} for each validation image using mean absolute error (MAE)

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |a_{440,\text{sensor}} - a_{440,\text{in situ}}|}{n} \qquad (2)$$

where $a_{440,sensor}$ is either $a_{440,\mathrm{MSI}}$ or $a_{440,\mathrm{OLI}}$. MAE = 0 indicates a perfect fit.

2.5 Statewide CDOM database

To create the 2015 statewide CDOM map, we used five clear paths (i.e., images from the same path and date but from multiple rows (two to five) to cover the state) of L8/OLI imagery (Table 1). For the 2016 map there were five mostly clear paths from 2015, but because a few areas in western Minnesota did not have any clear imagery in 2016, we also used two clear paths of 2017 L8/OLI imagery to fill in missing areas to complete the 2016 map (Table 1). To produce maps, the validated CDOM model was applied to the corresponding selected MAIN-derived R_{rs} bands in the GEE application program interface (API) (Page et al. 2019) for each path of imagery (Table 1) used for the 2015 and 2016 CDOM maps and exported in GeoTIFF format. The paths were mosaicked into statewide maps using ERDAS Imagine to create 2015 and 2016 pixel-level CDOM maps for Minnesota. To create a lake-level database, we used a polygon layer previously constructed (Olmanson et al. 2008) to include all Minnesota lakes, reservoirs and open-water wetlands ≥ 4 ha and the signature editor in ERDAS Imagine to extract a_{440} data for all lakes in the images using the lake polygon layer. The GetHist program (Olmanson et al. 2008) was used to calculate the mean a_{440} values from the middle 70 percent and linked to each lake polygon to create lake-level maps for 2015 and 2016.

To compile the data for analysis of CDOM at the ecoregion level, we used Esri ArcMap 10.5.1 to link each lake polygon to its respective ecoregion, and JMP Pro 14 to calculate CDOM distributions for each Minnesota ecoregion.

3. Results and discussion

3.1 CDOM model results

After exploration of various two-term regression models using L8/OLI, we identified the best model as having the form:

$$\ln(a_{440}) = a(R_{rs}(B4)/R_{rs}(B3)) + b(R_{rs}(B5)/R_{rs}(B3))) + c \qquad (3)$$

where coefficients, a, b, and c were fit to the calibration data by regression analysis, $\ln(a_{440})$ is the natural logarithm of the L8/OLI-derived a_{440} for a given sample location and B represents the corresponding L8/OLI spectral band. From the combined L8/OLI dataset, the $\ln(a_{440})$ prediction model generated a strong fit with $R^2 = 0.85$ and RMSE = 0.49 for MAIN, and $R^2 = 0.83$ and RMSE = 0.52 for OLI-SR (Table A2, Figure 2). MAIN-based results also fit closer to the 1:1 line than OLI-SR results, but both methods provided a better fit in the lower and higher ranges than our previous efforts (Olmanson et al. 2016a, b).

To evaluate model performance in different CDOM ranges, we split the data into low, medium and high sets (a_{440} = 0.2-3.0, 3-10 and 10-32.5 m^{-1}, respectively) and calculated MAE (Table 2a). In all ranges, MAIN-corrected imagery had lower MAE values than OLI-SR-corrected imagery, and although the MAE increased with a_{440}, the values were a relatively small fraction of the median a_{440} for the range. We also plotted measured a_{440} from low to high with model predicted a_{440} for MAIN and OLI-SR (Figure 3). MAIN-based results fit closer to the line for field measured a_{440} than OLI-SR results and deviation from the line for field measured a_{440} increased with increasing CDOM.

The use of MAIN or OLI-SR image correction together with the best-fit model resulted in substantial improvements in CDOM estimation compared to previous methods, largely due to improved atmospheric correction and a relatively large and varied in situ dataset (Figure 2). In comparison with other models in the literature, the green/red model of Kutser et al. (2005a) and red/green model of Menken et al. (2006) when applied to the combined L8/OLI dataset generated comparatively weak linear regressions with ln(a_{440}): R^2 values of 0.46 and 0.51, respectively, and higher RMSE values of 0.93 and 0.88, respectively (Table A2). The green/blue, red model of Griffin et al. (2011), which uses the blue band, where CDOM absorption is much stronger, generated no convincing relationship (average R^2 = 0.04, RMSE = 1.24), which indicates interference from other optically active constituents (Table A2). Compared against previous models, our approach offered substantial improvements in a_{440} measurements especially in the higher and lower ranges.

3.2 CDOM model validation

The semi-empirical model developed here was applied to some 2015, 2016 and 2017 L8/OLI images that were not used for model development to complete the 2015 and 2016 CDOM maps for Minnesota. Because these data do not have in situ validation data it is important to use an independent validation dataset to determine the accuracy that can be expected when the model is used on images not included in the calibration dataset. The validation dataset consists of overlapping L8/OLI and S2/MSI images acquired on August 13, 2018 that were mostly clear but had visible wildfire smoke in northern Minnesota. The L8/OLI validation data for the low and medium CDOM ranges resulted in higher MAE values (1.46 and 2.26 m^{-1}, respectively) than found for the corresponding calibration results (MAE = 0.42 and 1.79 m^{-1}, respectively) using Eq. (2) (Table 2b). The MAE of 1.63 m^{-1} for the whole validation dataset is comparable to that for the calibration dataset with a MAE of 1.61 m^{-1}, likely because of the lack of high CDOM values in the validation data (because the haze problem in northern Minnesota imagery). Despite the lack of high CDOM lakes, the validation data range still represented a large majority (> 92%) of surface waterbodies in Minnesota; CDOM values > 10 m^{-1} occurred in only 8% of the state's surface waters. If we consider only lakes and reservoirs and exclude open-water wetlands (i.e. shallower waterbodies that have abundant aquatic vegetation but include open-water areas where CDOM measurements can be extracted), CDOM > 10 m^{-1} occurred in only 6% of the lakes. The S2/MSI validation dataset yielded larger MAE values of 1.58 and 2.90 m^{-1} (Table 2b) for the low and medium CDOM ranges than corresponding values for the calibration data (0.43 and 2.05 m^{-1}, Table 2a). The larger errors could indicate that the validation imagery is less than ideal, especially for the lower CDOM values, because smoke effects may have been more widespread than what was obvious for northern Minnesota. Nevertheless, the MAE values indicate acceptable confidence in the resulting maps.

3.3 Geospatial analysis of statewide CDOM database

For geospatial analysis of CDOM at the ecoregion level, we calculated the mean CDOM value for each waterbody (i.e. lakes, reservoirs and open-water wetlands) using the pixel-level maps for 2015 and 2016 (Figures A1 and A2, respectively). These maps are also available in an online LakeBrowser at https://lakes.rs.umn.edu/. Satellite-derived a_{440} values encompassed broad ranges – from near undetectable (0.1 m^{-1}) to ~ 25.5 m^{-1} in both years. Standard deviations across all waterbodies for both years were larger than the mean values, and median values were less than the mean values (Table 3) indicating skewed distributions, with many more low-CDOM waterbodies than high ones. Large differences in means, medians and statistical distributions were found between the ecoregions, with high CDOM waters concentrated mainly in the NLF and NMW. Nonetheless, a few waterbodies had high CDOM levels in all ecoregions in both years. Standard deviations for a_{440} within all ecoregions were close to or larger than the mean values, consistently indicating skewed distributions. Mean a_{440} and distributional statistics were similar for the four southern and western ecoregions (NCHF, WCBP, NGP, LAP), and in all cases 90% of their waterbodies had a_{440} < ~ 6 m^{-1}.

Using the individual waterbody data for both years, we calculated the 2015-2016 mean value for each waterbody and created a “lake-level” map (Figure 4). The associated statistical distributions by ecoregion (Figure A3 and Table 4a) are similar to those described above for the individual years. The mean a_{440} values for the two most northern ecoregions (NLF and NMW) were higher than the means for the other four ecoregions in both years and for the average over the time period, and the differences were even more pronounced for the 75% and 90% quantile values. For example, 10% of the waterbodies in the NLF and NMW had average a_{440} values >

17.5 m^{-1} in 2015-2016, but the 90% quantile values for the other four ecoregions were only 4.4-5.4 m^{-1} (Table 4a).

Waterbodies with high CDOM tend to have watersheds dominated by forests and wetlands, but further inspection of high CDOM waterbodies in agricultural ecoregions (e.g., WCBP, NGP) indicated that they were mainly open-water wetlands with abundant aquatic vegetation, where vegetation and bottom effects could affect R_{rs} and provide erroneous results with satellite imagery methods. Ideally, pixels affected by aquatic vegetation or bottom sediment would be masked because they are unsuitable for remotely sensed estimates of water quality. Open-water wetlands were not well represented in the calibration dataset, however, because they typically are ringed with emergent vegetation and are difficult to access. Because masking all affected pixels is not always possible in large regional assessments, it is important to know the limitations of the analysis and whether the satellite-based measurements are realistic for the waterbodies that are being studied. Open-water wetlands tend to have high DOM concentrations, which suggests that the satellite-based measurements are correct, but this issue needs further investigation in future studies.

To minimize the effects of shallower open-water wetlands on CDOM statistical distributions, we removed these waterbodies from the dataset and found distributions (Figure A4 and Table 4b) similar to those in Table 4a but with fewer high CDOM waters in the agricultural ecoregions. Overall, mean a_{440} values and distributional statistics (except for maximum values) were slightly lower in all ecoregions for the subset without open-water wetlands. For example, for the four ecoregions with low average CDOM levels, the 90% quantile values were ~80% of the corresponding values for the dataset that includes the shallow open-water wetlands (Table

4a), suggesting that on average, open-water wetlands tend to have slightly higher CDOM levels than lakes and reservoirs.

3.4 Potential sources of error

Considering error levels indicated by MAE, atmospheric correction by MAIN resulted in lower error than using OLI-SR (Table 2a, Figure 3), with overall MAE averages of 1.61 and 1.82 m^{-1}, respectively. MAE values for both correction methods increased across the three CDOM ranges (low, medium, high) with MAIN and OLI-SR, and they represented ~25-30% of the midpoint a_{440} values of each range. Although the model developed using L8/OLI imagery worked reasonably well with our validation S2/MSI imagery, MAE values for the validation set were consistently lower for L8/OLI than for S2/MSI. Further research with a larger dataset would help to determine whether a separate S2/MSI model could improve the relationship with measured data.

Although Brezonik et al. (2015) concluded that CDOM is generally stable on intra-seasonal time scales, we found large fluctuations in CDOM in some highly colored lakes in flowage systems (i.e., with large watersheds relative to lake areas) following large storm events in summer of 2016. For this study, we used CDOM data within 30 days of image acquisition, but because numerous storm events occurred in the state during summer of 2016, this could have been too large a window for some highly colored flowage lakes and could account for some of the overall error. The low R_{rs} signals from high-CDOM, low-suspended sediment water and potential errors in atmospheric correction of such waters also could be contributing factors.

Differences between satellite and field measurements could originate from many sources including (1) differences in spatial coverage (20-30 m pixels vs. a single grab sample), (2) temporal variations in CDOM between the time of satellite overpass and sample collection, (3) errors in collection and laboratory analyses, (4) differences that may arise in predicting measured a_{440} from any retrieval model, and (5) satellite atmospheric correction errors. The latter potentially may have been exacerbated by haze differences due to smoke in the validation vs. the calibration dataset in this study. Given these issues and some uncertainties associated with the representativeness of field data, it may be better simply to regard satellite-based methods as the standard values for census-level CDOM data at regional scales. Ground-based measurements are simply infeasible to gather at such spatial scales and short timescales. Of course, use of clear imagery and appropriately calibrated models is essential for accurate results.

3.5 Applications to research and management

CDOM data for thousands of lakes measured at seasonal to annual time scales with the satellite imagery methods described here are invaluable for lake management and research. CDOM directly affects many important characteristics of lakes, such as temperature and light regimes, primary production, and carbon cycling. It also affects many variables relevant to lake management, including fisheries production and contaminant concentrations and reactivity. Despite its important role, in situ data for CDOM are much more limited compared to other key variables, such as chlorophyll and phosphorus (Stanley et al. 2019). Thus, frequent measurement of CDOM at regional scales represents an important resource for research and management.

To illustrate the use of large-scale CDOM measurements, we examined the changes in CDOM levels between two consecutive years with contrasting rainfall. Using the lake subset (Table A3), we analyzed the change in a_{440} between 2015 and 2016. Comparison of precipitation ranking maps for 2015 and 2016 shows major contrasts in hydrologic regimes between the years, with 2015 fairly typical for most areas and 2016 unusually wet for most of the state, including the NMW and NLF ecoregions (Figure A5). Comparing CDOM levels between years 2015 and 2016 (Table A4), we found that levels decreased by at least 3 m^{-1} in about 7% of the lakes in agricultural ecoregions (LAP, NGP and WCBP), but levels increased in the ecoregions with more forest and wetlands (Figure 5). Within the NMW and NLF ecoregions, 31% and 28% of the lakes, respectively, had changes in $a_{440} \geq 3$ m^{-1}, but only 5% of the lakes in the NCHF (a transition ecoregion) changed more than 3 m^{-1}. It also is interesting to note that the mean and median a_{440} values for the two high-CDOM ecoregions (NLF and NMW) increased substantially from 2015 to 2016 (Table A3). In contrast, in almost all cases these statistics decreased in the ecoregions with more agricultural and less forest/wetland land cover, apparently because of dilution by increased precipitation. Although CDOM is generally stable at timescales of weeks to months for many lakes, this analysis suggests that lakes in watersheds with large CDOM source areas (i.e. forested wetlands) can exhibit substantial precipitation-driven variability. De Wit et al. (2016) made similar conclusions based on analysis of long-term precipitation and CDOM records in Scandinavia, and our calibration database also supports this conclusion. This example provides an illustration of the utility of remote sensing methods to quantify CDOM changes in response to environmental drivers such as precipitation, temperature and land cover changes.

4. Conclusions

This paper demonstrates that remote sensing using satellite-based sensors can play an important role in providing census-level CDOM data over large areas at high temporal and spatial resolution. The constellations of L8/OLI, upcoming Landsat 9/OLI and Sentinel 2/MSI will greatly expand the capabilities to measure several optically-related water quality characteristics, including CDOM.

Strong relationships for CDOM (a_{440}) were found using both MAIN and OLI-SR atmospheric correction methods. Atmospheric correction using MAIN substantially improved model performance, and has the advantage of being able to harmonize the R_{rs} values of L8/OLI and S2/MSI, which will be important for automated image processing and near real-time monitoring. The range of a_{440} values in our calibration dataset (0.2-32.5 m^{-1}) likely represents the general distribution of CDOM throughout Minnesota.

Although further investigation of CDOM levels in shallow open-water wetlands of agricultural areas should be undertaken, our results indicate that assessment of CDOM at regional (statewide) scales is feasible using Landsat and Sentinel data. Such assessments can provide the basis for numerous regional-scale analyses related to CDOM, such as (a) change detection, as discussed above, (b) evaluating water clarity issues (e.g., Brezonik et al. 2019a), (c) quantifying patterns of temperature structure, (d) estimating carbon storage and mercury levels in lakes and wetlands, (e) predicting photochemical reaction rates in surface waters, and (f) assessing water treatability metrics, such as chlorine demand and disinfection byproduct formation (Chen et al. 2019). This approach could be extended to other regions, providing similar results with appropriate model tuning and validation.

Funding: This work was supported in part by National Science Foundation grant (CBET 1510332), Minnesota Environmental and Natural Resources Trust fund, Minnesota Agricultural Experiment Station, and University of Minnesota's U-Spatial Program, Sea Grant Program and Office of the VP for Research and Retirees Association.

Acknowledgments: We thank the Minnesota Pollution Control Agency's lake water quality assessment program and numerous collaborators, research staff, and students for assistance in sample collection and analysis.

Author Contributions: Conceptualization of the study that led to this paper P.L.B., L.G.O., and J.C.F., Conceptualization and methodology of remote sensing conducted by L.G.O., and B.P.P. Formal statistical analysis done by L.G.O., and P.L.B. Field data planning analysis and collection C.G.G., P.L.B., J.C.F., L.G.O., R.M.H. Writing and draft preparation done by L.G.O., P.L.B., J.C.F., B.P.P. Review and editing done by L.G.O., P.L.B., J.C.F., B.P.P., M.E.B., R.M.H. and C.G.G. All authors contributed substantially to the produced work. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Supplemental Material:

References

Ask, J., J. Karlsson, L. Persson, P. Ask, P. Byström, and M. Jansson. 2009. Terrestrial organic matter and light penetration: Effects on bacterial and primary production in lakes. Limnology and Oceanography **54**:2034-2040. https://doi.org/10.4319/lo.2009.54.6.2034.

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. 2019a. Color, chlorophyll a, and suspended solids effects on

Secchi depth in lakes: implications for trophic state assessment. Ecological Applications **29**:e01871. https://doi.org/10.1002/eap.1871.

Brezonik, P. L., J. C. Finlay, C. G. Griffin, W. A. Arnold, E. H. Boardman, N. Germolus, R. M. Hozalski, and L. G. Olmanson. 2019b. 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.

Brezonik, P. L., K. Menken, and M. E. Bauer. 2005. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). Lake Reserv. Manage. 21(4): 373-382. https://doi.org/10.1080/07438140509354442.

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. http://dx.doi.org/10.1016/j.rse.2014.04.033.

Chen, Y., W. A. Arnold, C. G. Griffin, L G. Olmanson, P. L. Brezonik, and R. M. Hozalski. 2019. Assessment of the chlorine demand and disinfection byproduct formation potential of surface waters via satellite remote sensing. Water Research 165: 115001. https://doi.org/10.1016/j.watres.2019.115001.

Chen, J., W. Zhu, Y. Q. Tian, Q. Yu, Y. Zheng, and L. Huang. 2017. Remote estimation of colored dissolved organic matter and chlorophyll-a in Lake Huron using Sentinel-2 measurements. J. Appl. Remote Sens.11(3), 036007. https://doi.org/10.1117/1.JRS.11.036007.

Corman, J. R., B. L. Bertolet, N. J. Casson, S. D. Sebestyen, R. K. Kolka, and E. H. Stanley. 2018. Nitrogen and Phosphorus Loads to Temperate Seepage Lakes Associated With

Allochthonous Dissolved Organic Carbon Loads. Geophysical Research Letters 45:5481-5490. https://doi.org/10.1029/2018GL077219.

Creed, I. F., A.-K. Bergström, C. G. Trick, N. B. Grimm, D. O. Hessen, J. Karlsson, K. A. Kidd, E. Kritzberg, D. M. McKnight, E. C. Freeman, O. E. Senar, A. Andersson, J. Ask, M. Berggren, M. Cherif, R. Giesler, E. R. Hotchkiss, P. Kortelainen, M. M. Palta, T. Vrede, and G. A. Weyhenmeyer. 2018. Global change-driven effects on dissolved organic matter composition: Implications for food webs of northern lakes. Global Change Biology 24:3692-3714. http://dx.doi.org/10.1111/gcb.14129.

de Wit, H. A., S. Valinia, G. A. Weyhenmeyer, M. N. Futter, P. Kortelainen, K. Austnes, D. O. Hessen, A. Räike, H. Laudon, and J. Vuorenmaa. 2016. Current Browning of Surface Waters Will Be Further Promoted by Wetter Climate. Environ. Sci. Technol. Lett. 2016, 3, 43.0-435. https://doi.org/10.1021/acs.estlett.6b00396.

Finstad, A. G., T. Andersen, S. Larsen, K. Tominaga, S. Blumentrath, H. A. de Wit, H. Tømmervik, and D. O. Hessen. 2016. From greening to browning: Catchment vegetation development and reduced S-deposition promote organic carbon load on decadal time scales in Nordic lakes. Scientific Reports 6:31944. https://doi.org/10.1038/srep31944.

Gordon, H.R. and M. Wang, 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.

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.

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. https://doi.org/10.1016/j.watres.2018.08.007.

Griffin, C. G., Frey, K. E., Rogan, J., and Holmes, R.M. 2011. Spatial and interannual variability of dissolved organic matter in the Kolyma River, East Siberia, observed using satellite imagery. Journal of Geophysical Research – Biogeosciences, 116(12), G03018. http://dx.doi.org/10.1029/2010JG001634.

Houser, N. J. 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.

Karlsson, J., P. Bystrom, J. Ask, P. Ask, L. Persson, and M. Jansson. 2009. Light limitation of nutrient-poor lake ecosystems. Nature 460:506-508. https://doi.org/10.1038/nature08179.

Kirk, J. T. O. 2010. Light and photosynthesis in aquatic ecosystems, 3rd edition. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9781139168212

Knoll, L. B., C. E. Williamson, R. M. Pilla, T. H. Leach, J. A. Brentrup, and T. J. Fisher. 2018. Browning-related oxygen depletion in an oligotrophic lake. Inland Waters 8:255-263. https://doi.org/10.1080/20442041.2018.1452355.

Kritzberg, E. S. 2017. Centennial-long trends of lake browning show major effect of afforestation. Limnology and Oceanography Letters 2:105-112. https://doi.org/10.1002/lol2.10041.

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. of Environ. 224, 104–118. https://doi.org/10.1016/j.rse.2019.01.023

Kutser, T., D. C. Pierson, K. Y. Kallio, A. Reinart, and S. Sobek. 2005. Mapping lake CDOM by satellite remote sensing. Remote Sensing of Environment 94:535-540. https://doi.org/10.1016/j.rse.2004.11.009.

Kutser, T., L. Tranvik, and D. C. Pierson. 2009. Variations in colored dissolved organic matter between boreal lakes studied by satellite remote sensing. Journal of Applied Remote Sensing 3(1):033538. https://doi.org/10.1117/1.3184437.

Menken, K., Brezonik, P. L., and Bauer, M. E. (2006). Influence of chlorophyll and humic color on reflectance spectra of lakes: Implications for measurement of lake-water properties by remote sensing. Lake and Reservoir Management, 22(3):179-190. https://doi.org/10.1080/07438140609353895.

Olmanson, L. G., M. E. Bauer and P. L. Brezonik. 2008. Development and analysis of a 20-year Landsat water clarity census of Minnesota's 10,000 lakes. Remote Sens. Environ. 112:4086-97. https://doi.org/10.1016/j.rse.2007.12.013.

Olmanson, L.G., Bauer, M.E., and Brezonik, P.L. 2011. Evaluation of Medium to Low Resolution Satellite Imagery for Regional Lake Water Quality Assessment. Water Resources Research. 47:W09515. http://dx.doi.org/10.1029/2011WR011005.

Olmanson, L. G., P. L. Brezonik, J. C. Finlay, and M. E. Bauer. 2016a. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. Remote Sensing of Environment 185:119-128. https://doi.org/10.1016/j.rse.2016.01.007.

Olmanson, L. G., P. L. Brezonik, J. C. Finlay, and M. E. Bauer. 2016b. Regional-scale measurement of colored dissolved organic matter in freshwater lakes by satellite imagery.

European Space Agency, Living Planet Symposium 2016, Prague, Czech Republic. May 9-13, 2016. http://lps16.esa.int/posterfiles/paper2319/Olmanson_CDOM_poster_ESA_LPS_final.pdf. Accessed November 6, 2019.

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(3): 748-761. https://doi.org/10.1111/jawr.12138.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77(1):118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x.

Omernik, J. M. and G. E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54(6):1249–1266. https://doi.org/10.1007/s00267-014-0364-1.

Page, B.P. 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. 231:111284. https://doi.org/10.1016/j.rse.2019.111284.

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.

Pahlevan, N., Schott, J., Franz, B.A., Zibordi, G., Markham, B., Bailey, S., Schaaf, C., Ondrusek, M., Greb, S., and 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. https://doi.org/10.1016/j.rse.2017.08.033.

Pilla, R. M., C. E. Williamson, J. Zhang, R. L. Smyth, J. D. Lenters, J. A. Brentrup, L. B. Knoll, and T. J. Fisher. 2018. Browning-Related Decreases in Water Transparency Lead to Long-Term Increases in Surface Water Temperature and Thermal Stratification in Two Small Lakes. Journal of Geophysical Research: Biogeosciences 123(5):1651-1665. https://doi.org/10.1029/2017JG00432.

Rampi, L.P., Knight, J. F., Bauer, M.E. (2016). Minnesota Land Cover Classification and Impervious Surface Area by Landsat and Lidar: 2013 Update. Retrieved from the Data Repository for the University of Minnesota, http://doi.org/10.13020/D6JP4S.

Ross, M. R. V., S. N. Topp, A. P. Appling, X. Yang, C. Kuhn, D. Butman, M. Simard, and T. Pavelsy. 2019. AquaSat: a dataset to enable remote sensing of water quality for inland waters. *Water Res. Research* **55**: 10012-10025. doi: 10.1029/2019WR024883

Snucins, E. and J. Gunn. 2000. Interannual variation in the thermal structure of clear and colored lakes. Limnology and Oceanography 45:1639-1646. https://doi.org/10.4319/lo.2000.45.7.1639.

Sobek, S., L. J. Tranvik, Y. T. Prairie, P. Kortelainen, and J. J. Cole. 2007. Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. Limnology and Oceanography 52:1208-1219.

Solomon, C., S. Jones, B. Weidel, I. Buffam, M. Fork, J. Karlsson, S. Larsen, J. Lennon, J. Read, S. Sadro, and J. Saros. 2015. Ecosystem Consequences of Changing Inputs of Terrestrial

Dissolved Organic Matter to Lakes: Current Knowledge and Future Challenges. Ecosystems 18:376-389. DOI: 10.1007/s10021-015-9848-y.

Stanley, E. H., S. M. Powers, N. R. Lottig, I. Buffam, and J. T. Crawford. 2012. Contemporary changes in dissolved organic carbon (DOC) in human-dominated rivers: is there a role for DOC management? Freshwater Biology **57**:26-42. https://doi.org/10.1111/j.1365-2427.2011.02613.x.

Stanley, E. H., S. M. Collins, N. R. Lottig, S. K. Oliver, K. E. Webster, K. S. Cheruvelil, and P. A. Soranno. 2019. Biases in lake water quality sampling and implications for macroscale research. Limnology and Oceanography 64(4):1572-1585. https://doi.org/10.1002/lno.11136. doi.org/10.1002/lno.11136.

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. https://doi.org/10.1007/s10021-014-9776-2.

Toming, K., T. Kutser, A. Laas, M. Sepp, B. Paavel, and T. Nõges. 2016. First Experiences in Mapping Lake Water Quality Parameters with S2/MSI Imagery. Remote Sensing 8:640. https://doi.org/10.3390/rs8080640.

Tsui, M. T. K., and J. C. Finlay. 2011. Influence of Dissolved Organic Carbon on Methylmercury Bioavailability across Minnesota Stream Ecosystems. Environmental Science & Technology 45:5981-5987. https://doi.org/10.1021/es200332f.

Tyler, A., Hunter, P., Spyrakos, 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. https://doi.org/10.1016/j.scitotenv.2016.01.020.

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. https://doi.org/10.1016/j.rse.2015.02.007.

Vanhellemont, Q. and Ruddick, K., 2016. Acolite for Sentinel-2: Aquatic applications of MSI imagery. Proceedings of the 2016 ESA Living Planet Symposium, ESA Special Publication, SP 740.

Williamson, C. E., D. P. Morris, M. L. Pace, and A. G. Olson. 1999. Dissolved organic carbon and nutrients as regulators of lake ecosystems: Resurrection of a more integrated paradigm. Limnology and Oceanography 44:795-803. https://doi.org/10.4319/lo.1999.44.3_part_2.0795.

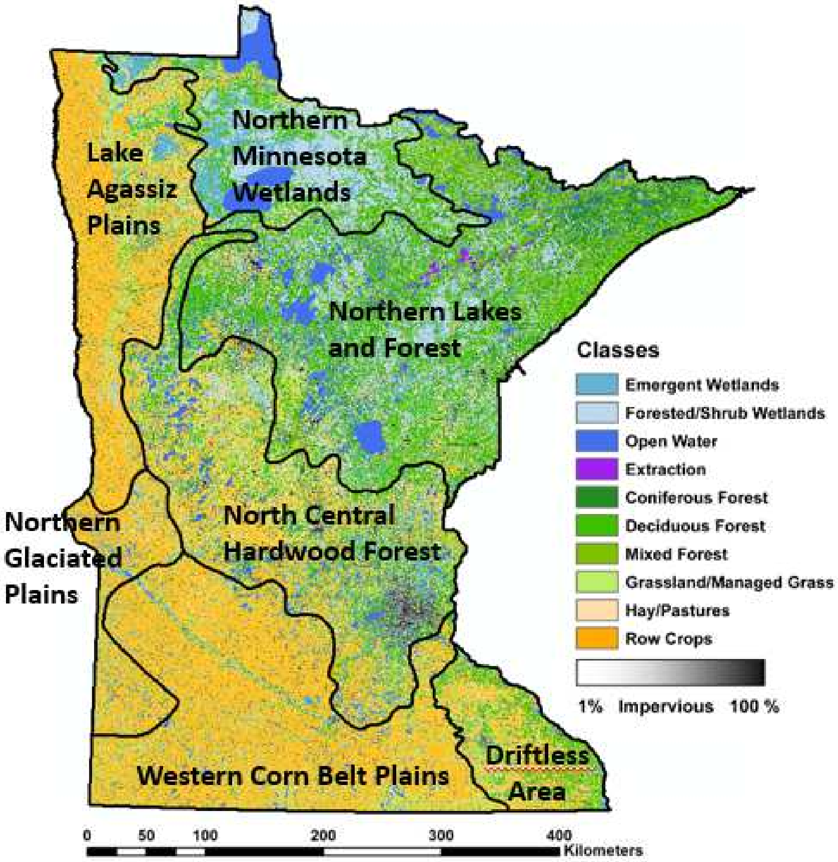


Figure 1. Minnesota 2013 land cover map (Rampi et al., 2016) with ecoregion boundaries (Omernik and Griffith, 2014).

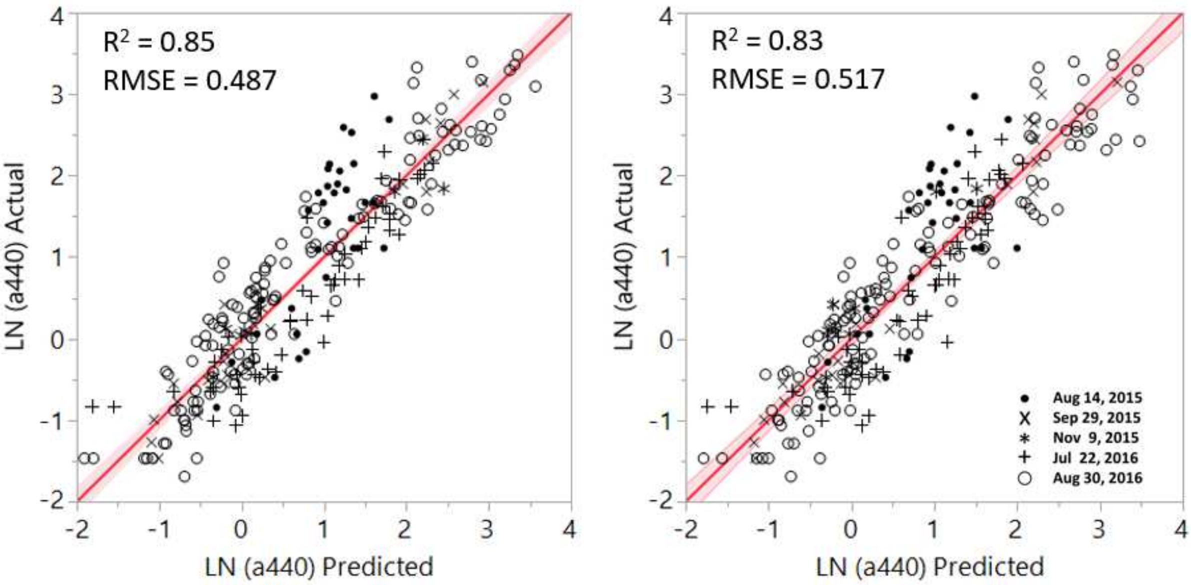


Figure 2. Landsat 8 CDOM models using MAIN (left) and SR (right) R_{rs} products.

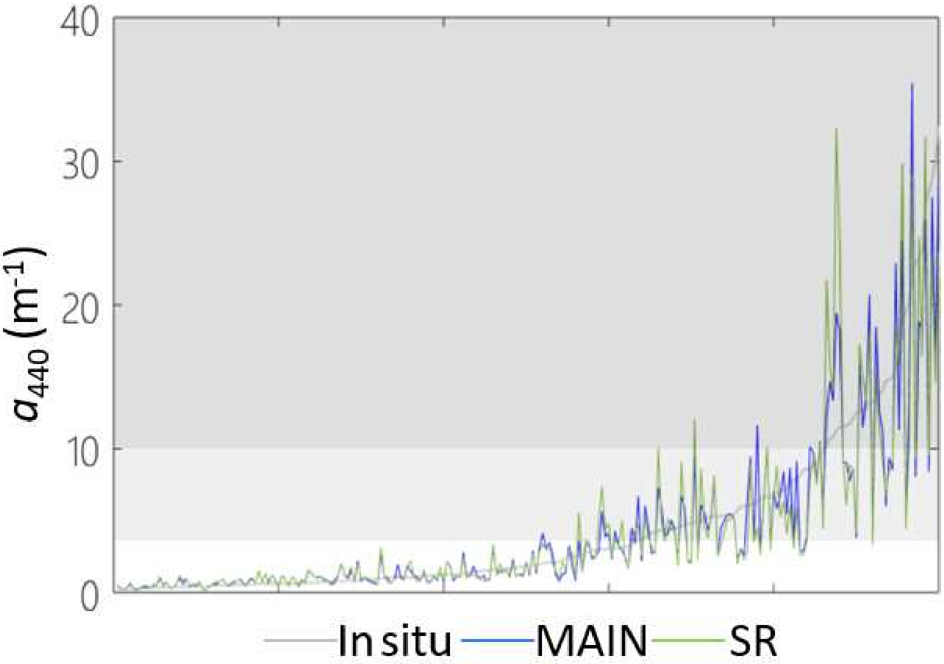


Figure 3. In situ a_{440} data sorted from low to high with resulting MAIN and SR model-derived a_{440} showing increasing divergence with increasing in situ $a_{440.}$ The shading represents low, medium, and high CDOM levels.

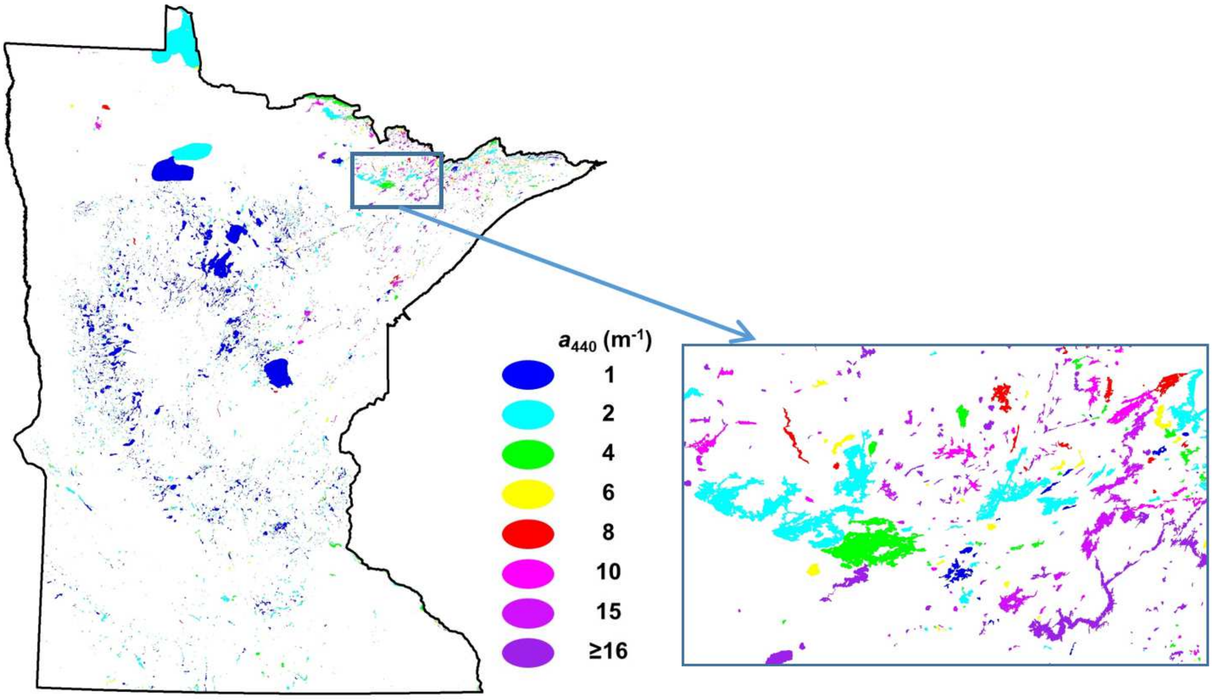


Figure 4. Mean 2015-2016 lake-level CDOM map with blowup of the Ely lakes area.

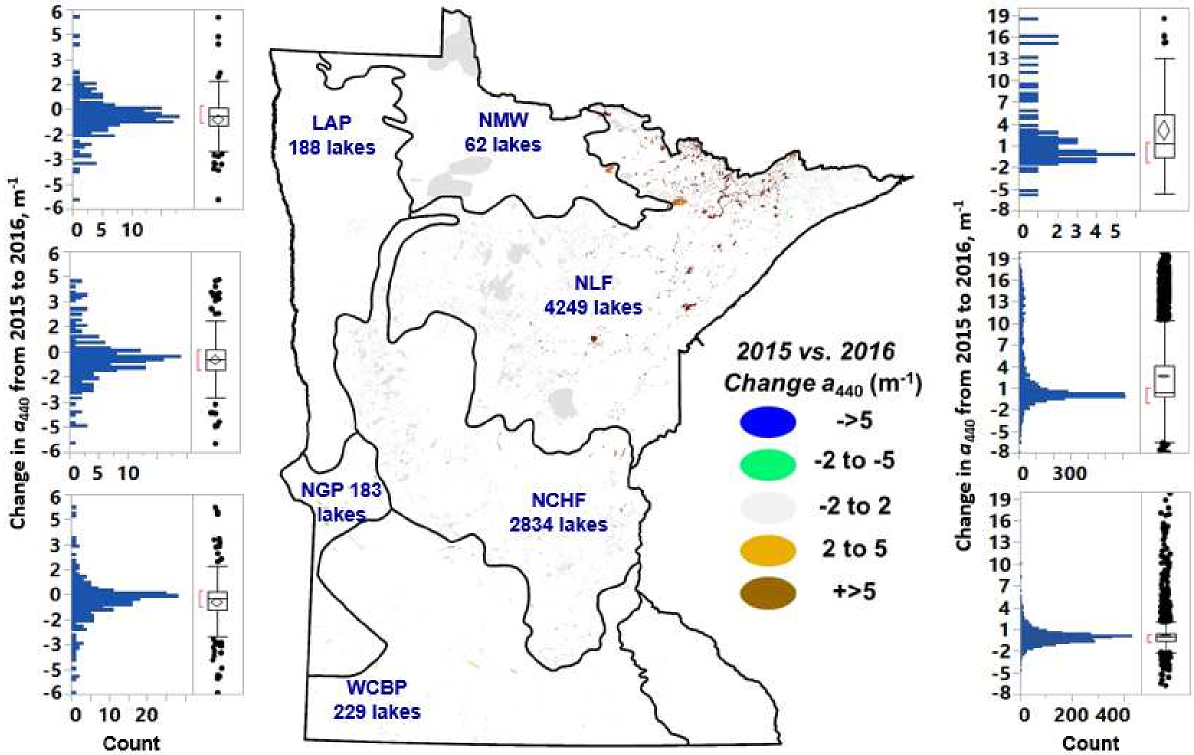


Figure 5. Percent change in *a*440 between 2015 compared to 2016 for each ecoregion. Increase of a440 between 2015 and 2016 due to increased precipitation in 2016 is focused in ecoregions with high coverage of forest and wetlands (NLF and NMW) while *a*440 decreases are in agricultural ecoregions (LAP, NGP and WCBP).

Table 1. Landsat 8 images used for calibration/validation and images used for 2015 and 2016-17 CDOM maps and associated number of ground-based (a_{440}) measurements.

Purpose	Sensor	Date	Path	Rows	N
calibration, 2015 map	L8/OLI	8/14/2015	26	27-28	33
2015 map	L8/OLI	9/20/2015	29	26-28	
calibration, 2015 map	L8/OLI	9/29/2015	28	26-30	24
2015 map	L8/OLI	11/7/2015	29	26-29	
calibration, 2015 map	L8/OLI	11/9/2015	27	26-30	9
calibration, 2016 map	L8/OLI	7/22/2016	27	26-29	53
calibration, 2016 map	L8/OLI	8/30/2016	28	26-28	131
2016 map	L8/OLI	11/4/2016	26	26-30	
2016 map	L8/OLI	11/9/2016	29	26-30	
2016 map	L8/OLI	11/11/2016	27	26-30	
2016 map	L8/OLI	5/13/2017	28	28-30	
2016 map	L8/OLI	9/9/2017	29	26-30	
Validation	S2/MSI	8/13/2018	MN_Middle	MN_N	95
Validation	L8/OLI	8/13/2018	27	28-30	62

Table 2. Error analysis for (a) L8 calibration dataset of MAIN and EROS SR CDOM models and (b) validation dataset for L8 and S2 models showing mean absoulte error (MAE) in three ranges of a_{440}.

(a) Calibration data	**CDOM range (a_{440}, m^{-1})**			
Model	**0-3**	**3-10**	**10-33**	**All**
MAE (MAIN), m^{-1}	0.42	1.79	6.07	1.61
MAE (EROS SR), m^{-1}	0.43	2.05	7.10	1.82
N*	147	67	36	250
(b) Validation data				
L8-OLI MAE, m^{-1}; (N)	1.46 (49)	2.26 (12)	2.43 (1)	1.63 (62)
S2-MSI MAE, m^{-1}; (N)	1.58 (79)	2.90 (15)	2.93 (1)	1.80 (95)

* N is the number of data points in each range.

Table 3. Summary statistics and quantile information for 2015 and 2016 CDOM (a_{440}, m^{-1}) in waterbodies of Minnesota's six main ecoregions.

a). All measured waterbodies: 2015

	Ecoregion						
Statistic	**All**	**NLF**	**NMW**	**NCHF**	**WCBP**	**NGP**	**LAP**
Mean	3.54	4.83	6.45	2.05	3.25	2.99	2.89
Std dev	4.28	5.37	5.89	1.96	3.29	2.92	2.99
Std err mean	0.04	0.07	0.64	0.03	0.14	0.14	0.15
Minimum	0.16	0.16	0.71	0.20	0.25	0.55	0.30
Quantiles:							
10%	0.76	0.69	1.18	0.79	1.02	1.10	0.95
25%	1.15	1.17	2.22	1.07	1.55	1.52	1.30
Median (50%)	1.91	2.52	4.60	1.57	2.29	2.08	1.87
75%	3.82	6.69	7.96	2.32	3.57	3.30	3.19
90%	8.62	12.83	17.27	3.46	6.05	5.56	5.77
Maximum	25.50	25.50	25.50	25.50	25.50	25.50	23.60
N	10,782	5,081	83	4196	583	407	402

b) All measured waterbodies: 2016

	Ecoregion						
Statistic	**All**	**NLF**	**NMW**	**NCHF**	**WCBP**	**NGP**	**LAP**
Mean	4.90	7.53	9.70	2.58	2.50	2.49	2.13
Std dev	6.72	8.40	7.91	3.57	2.32	2.91	2.65
Std err mean	0.06	0.11	0.83	0.05	0.08	0.14	0.13
Minimum	0.10	0.20	0.51	0.21	0.32	0.10	0.20
Quantiles:							
10%	0.67	0.70	1.23	0.64	0.79	0.72	0.51
25%	1.03	1.22	2.93	0.92	1.18	1.00	0.81
Median (50%)	1.93	3.26	6.99	1.48	1.81	1.62	1.37
75%	4.81	11.89	16.69	2.59	2.85	2.87	2.29
90%	17.03	23.59	23.44	5.02	4.84	5.18	4.30
Maximum	25.50	25.50	25.50	25.50	19.44	25.50	23.70
N	11,565	5,337	91	4,451	748	411	406

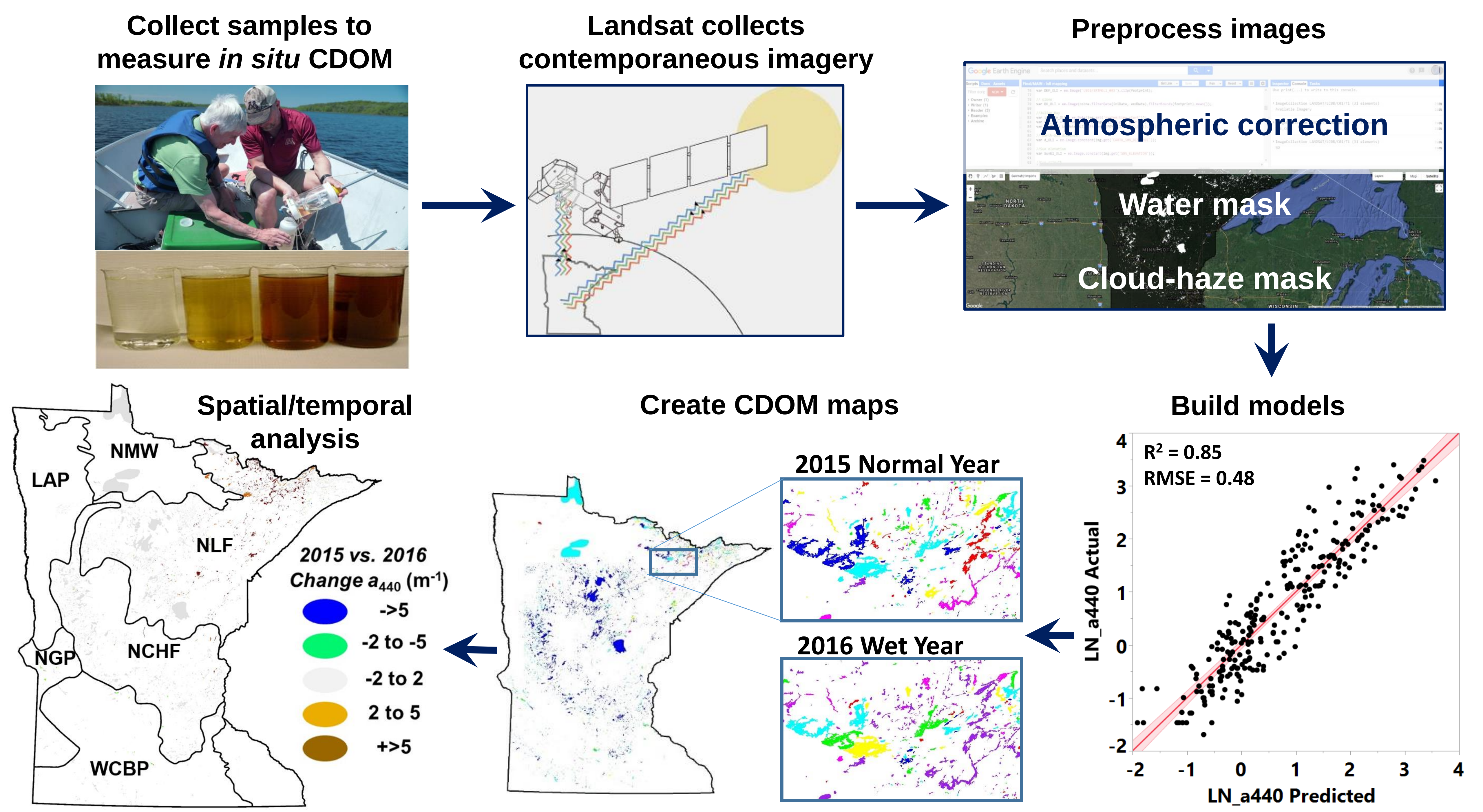
Table 4. Summary statistics and quantile information for 2015-2016 average CDOM (a_{440}, m^{-1}) for all measured waterbodies and lakes and reservoirs (without open-water wetlands) only in Minnesota and its six main ecoregions.

a) All measured waterbodies

		Ecoregion					
Statistic	**All**	**NLF**	**NMW**	**NCHF**	**WCBP**	**NGP**	**LAP**
Mean	4.34	6.31	8.47	2.46	2.87	2.81	2.56
Std dev	5.34	6.63	6.70	2.94	2.36	2.69	2.64
Std err mean	0.05	0.09	0.70	0.04	0.09	0.13	0.13
Minimum	0.10	0.10	0.70	0.24	0.34	0.53	0.25
Quantiles:							
10%	0.80	0.74	1.54	0.80	1.00	1.02	0.79
25%	1.19	1.29	2.64	1.08	1.46	1.39	1.13
Median (50%)	2.03	3.20	6.11	1.60	2.21	1.91	1.72
75%	4.63	9.76	13.54	2.52	3.34	3.22	2.77
90%	12.79	17.54	17.62	4.40	5.38	5.37	5.16
Maximum	25.50	25.50	25.50	25.50	15.82	25.50	22.14
N	11,625	5,378	91	4,462	753	411	408

b) Lakes and reservoirs only

		Ecoregion					
Statistic	**All**	**NLF**	**NMW**	**NCHF**	**WCBP**	**NGP**	**LAP**
Mean	4.21	5.98	7.30	1.92	2.28	2.37	2.11
Std dev	5.34	6.43	6.56	1.98	1.67	1.66	2.17
Std err mean	0.06	0.10	0.81	0.04	0.10	0.12	0.16
Minimum	0.10	0.10	0.70	0.31	0.49	0.53	0.25
Quantiles:							
10%	0.75	0.71	1.29	0.76	0.91	0.88	0.71
25%	1.08	1.20	2.15	0.97	1.22	1.27	1.06
Median (50%)	1.84	2.92	5.38	1.40	1.78	1.87	1.44
75%	4.50	9.15	10.43	2.11	2.69	3.08	2.32
90%	12.97	16.89	17.40	3.35	4.28	4.55	4.07
Maximum	25.50	25.50	25.16	25.50	13.52	10.39	20.94
N	8,182	4,461	65	2,911	279	183	188



Collect samples to measure in situ CDOM
Landsat collects contemporaneous imagery
Preprocess images
Atmospheric correction
Water mask
Cloud-haze mask
Build models
R2 = 0.85
RMSE = 0.48
LN_a440 Actual
LN_a440 Predicted
4
3
2
1
0
-1
-2
-2
-1
0
1
2
3
4
Create CDOM maps
2015 Normal Year
2016 Wet Year
Spatial/temporal analysis
NMW
LAP
NLF
NGP
NCHF
WCBP
2015 vs. 2016
Change a440 (m-1)
->5
-2 to -5
-2 to 2
2 to 5
+>5