Remote Sensing of Chlorophyll-A in Clear vs. Turbid Waters in Lakes

Forough Fendereski , Irena F. Creed * , Charles G. Trick

Posted Date: 14 August 2024

doi: 10.20944/preprints202408.0962.v1

Keywords: phytoplankton blooms; chlorophyll-a; optical properties; Landsat; Google Earth Engine; Lake Winnipeg.

 doi:10.20944/preprints202408.0962.v1

Article

Remote Sensing of Chlorophyll-*a* in Clear vs. Turbid Waters in Lakes

Forough Fendereski [1], Irena F. Creed [2,*] and Charles G. Trick [3]

[1] University of Saskatchewan, School of Environment and Sustainability, 117 Science Place, Saskatoon, Saskatchewan, Canada S7N 5C8; forough.fendereski@usask.ca
[2] University of Toronto, Dept. of Physical and Environmental Sciences, 1265 Military Trail, Toronto, Ontario, Canada M1C 1A4
[3] University of Toronto, Dept. of Health and Society, 1265 Military Trail, Toronto, Ontario, Canada M1C 1A4; charles.trick@utoronto.ca
* Correspondence: irena.creed@utoronto.ca

Abstract: Chlorophyll-*a* (Chl-*a*), a proxy for phytoplankton biomass, is one of the few biological water quality indices detectable using satellite observations. However, models for estimating Chl-*a* from satellite signals are currently unavailable for many lakes. The application of Chl-*a* prediction algorithms may be affected by the variance in optical complexity within lakes. Using Lake Winnipeg in Canada as a case study, we demonstrated that separating models by the lake's basins [north basin (NB) and south basin (SB)] can improve Chl-*a* predictions. By calibrating more than 40 commonly used Chl-*a* estimation models using Landsat data for Lake Winnipeg, we achieved higher correlations between *in situ* and predicted Chl-*a* when building models with separate Landsat-to-in-situ matchups from NB and SB (R^2 = 0.85 and 0.76, respectively; $p < 0.05$), compared to using matchups from the entire lake (R^2 = 0.38, $p < 0.05$). In the deeper, clearer waters of the NB, a green-to-blue band ratio provided better Chl-*a* predictions, while in the shallower, highly turbid SB, a red-to-green band ratio was more effective. Our approach can be used for rapid Chl-*a* modeling in large lakes using cloud-based platforms like Google Earth Engine with any available satellite or time series length.

Keywords: phytoplankton blooms; chlorophyll-a; optical properties; Landsat; Google Earth Engine; Lake Winnipeg

1. Introduction

Recent increases in the frequency, prevalence, and intensity of harmful phytoplankton blooms in lakes worldwide have raised significant environmental concerns [1–3]. These blooms contribute to issues such as toxin production, water odor, taste degradation, and aesthetic decline [4,5]. They also reduce biodiversity through habitat degradation and alter energy and nutrient fluxes within food webs [5–7]. Addressing these challenges requires a thorough understanding of past, current, and potential future occurrences of phytoplankton blooms, necessitating regular, lake-wide monitoring of phytoplankton biomass [8]. However, traditional *in situ* monitoring methods are costly, time-consuming, and limited in spatial and temporal coverage [9].

Satellite sensors offer a cost-effective solution for monitoring chlorophyll-*a* (Chl-*a*), a proxy for phytoplankton biomass, across lakes by mapping its concentration [10]. Estimating Chl-*a* from satellite spectral data relies on the absorption and scattering of light by phytoplankton at specific wavelengths [11]. In waters with straightforward optical properties, phytoplankton and related substances predominantly determine the color of the water and thus the signals received by satellite sensors [12]. However, in optically complex waters like many inland and coastal areas, non-phytoplankton optically active constituents (OACs) such as detritus and colored dissolved organic matter (CDOM) also influence the color of water, complicating Chl-*a* estimation from satellite data

 doi:10.20944/preprints202408.0962.v1

[12]. These OACs vary widely in type, quantity, and band ratios across different water bodies [13], posing challenges for developing universal Chl-*a* prediction models [14,15].

Successful Chl-*a* estimation from satellite data requires the development of region-specific models, known as regional models, which account for the optical characteristics of the target water body [11,16,17]. Efforts to create such models often rely on lake-wide approaches that overlook within-basin variations in water optical properties [18].

This study aims to propose a method for developing simple yet accurate Chl-*a* prediction models that address the optical heterogeneity of lakes. Using Lake Winnipeg (LW), Canada, as a case study, we hypothesize that improving Chl-*a* predictions can be achieved by separating the lake into its optically distinct basins. We compare "basin-specific" models against conventional "lake-wide" models to evaluate their efficacy. Enhanced Chl-*a* predictions using basin-specific models could facilitate Chl-*a* modeling on platforms like Google Earth Engine (GEE), which is crucial for processing and analyzing long-term satellite data at broad spatial scales using medium- to high-resolution imagery, often computationally intensive outside cloud-based environments.

2. Materials and Methods

2.1. Case Study Area

Lake Winnipeg, with a surface area of 23,750 km², is recognized as the world's 10th largest, North America's 7th largest, and Canada's 3rd largest freshwater lake [19,20]. It is a shallow, wind-mixed, polymictic lake, divided into two physically and biogeochemically different basins: the north basin (NB) and the south basin (SB) (Figure 1) [19]. Compared to the SB, the NB is larger, encompassing 74% of the total lake area, and deeper, with an average depth of 13.3 m and a maximum depth of 18.0 m. The water in the NB is relatively transparent. In contrast, the SB, encompasses 11% of the total lake area and is shallower, with an average depth of 9.0 m and a maximum depth of 12.0 m. The Red River transports a large amount of sediments to the SB, combined with sediment resuspension, making the SB highly turbid. A deep, narrow channel, known as the Narrows, separates the NB and SB. The Narrows is characterized by its relatively rapid water flow, with a maximum depth of approximately 60.0 m and a width of around 2.6 km [19].

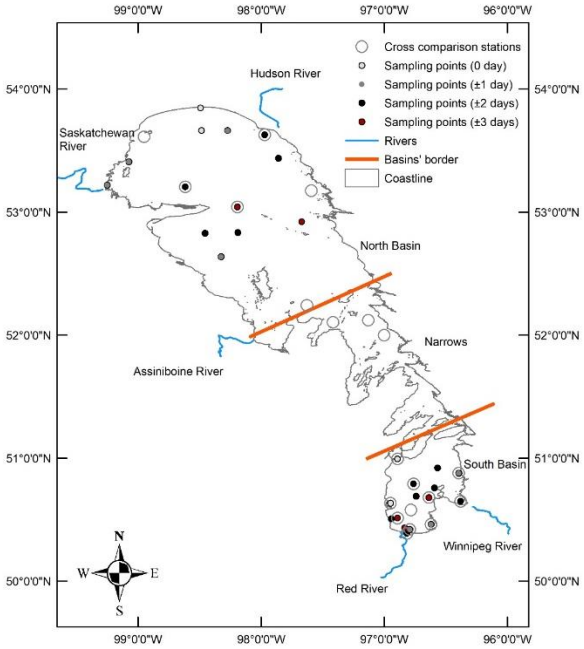


Figure 1. Map of Lake Winnipeg showing the location of the two basins and the narrows, as well as the location of the *in situ* Chl-*a* sampling points used for studying turbidity (all the circles shown on the map) and those used as Chl-*a* matchups from different temporal windows (0 to ±3 days shown as

 doi:10.20944/preprints202408.0962.v1

white, grey, black, and red circles, respectively). The empty circles represent stations used for cross comparing two Landsat sensors (ETM+ and OLI).

2.2. LW Turbidity

We used *in situ* turbidity data (Nephelometric Turbidity Units; NTU) measured throughout LW, provided by Manitoba Agriculture and Resource Development during the growing seasons (May to October) from 2002 to 2018 (Figure 1). A non-parametric, one-way analysis of variance by ranks (Kruskal–Wallis H test) was employed to determine whether the two LW basins (NB and SB) and the Narrows differed in their turbidity levels [21]. The choice of a non-parametric test was based on the non-normal distribution of the data (Shapiro-Wilcoxon test; $p > 0.05$). A significant result from the Kruskal-Wallis test ($p < 0.05$) indicated that at least one basin differed from the others. We then applied a multi-step *a posteriori* pairwise testing procedure based on studentized range statistics (non-parametric multiple comparison Dunn's procedure) to identify which basins were significantly different from each other [22].

2.3. Modeling Chl-a

2.3.1. In Situ Data Acquisition

In situ Chl-*a* measurements (μgL^{-1}) were provided by the Manitoba Sustainable Development, Water Quality Management Section in 2019. Chl-*a* data were collected from 32 stations distributed throughout the lake (14 stations in the NB, 15 in the SB, and three in the Narrows; Figure 1) during the peak phytoplankton biomass season (July to October) between 2002 and 2018 [19]. All Chl-*a* data were measured using spectrophotometric methods (American Public Health Association (APHA) 10200H [23]). Since satellite-derived Chl-*a* information is limited to the uppermost water layer, we only used samples taken from the surface of the lake (≤ 0.1 m depth). Although using Chl-*a* samples from deeper layers (≤ 0.5 m depth) could have substantially increased the number of *in situ* and satellite matchup points, it reduced the performance of Chl-*a* models (Figure S1). Our criteria for *in situ* Chl-*a* sample selection resulted in no samples from the Narrows, as all samples there were from deeper than 0.1 m. Due to the log-normal distribution of Chl-*a* in nature [24], all Chl-*a* data were logarithmically transformed (using natural log) prior to analysis.

2.3.2. Landsat Data

Landsat 4-5 Level 1 Collection 2 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images covering the entirety of LW were used in this study (https://www.usgs.gov/landsat-missions/landsat-collection-2). Both sensors (TM and ETM+) have a spatial resolution of 30 m. We used four visible bands: blue (B, 450–520 nm), green (G, 520–600 nm), red (R, 630–690 nm), and near-infrared (NIR, 770–900 nm). Matchups between Chl-*a* sample times and Landsat image acquisition times were extracted from 9 TM and 20 ETM+ images (6 TM and 6 ETM+ path/rows). Images with more than 90% cloud or cloud shadow cover were excluded.

2.3.3. Pre-Processing Landsat Level 1 Data

Partial atmospheric correction was applied to Landsat Level 1 data (Figure 2) because Level 2 products may introduce higher uncertainty in aerosol calculations over inland waters, leading to erroneous results [25,26]. All pre-processing steps for transforming Landsat data to partially corrected satellite top-of-atmosphere (TOA) reflectance were performed in GEE (Figure 2, Supplementary Material). Land pixels and those covered by cloud or cloud shadow during satellite image acquisition were masked out using a standard LW boundary shapefile and Pixel Quality Assessment bands provided with Landsat Level 2 products. The individual pixels were then pre-processed in three steps.

 doi:10.20944/preprints202408.0962.v1

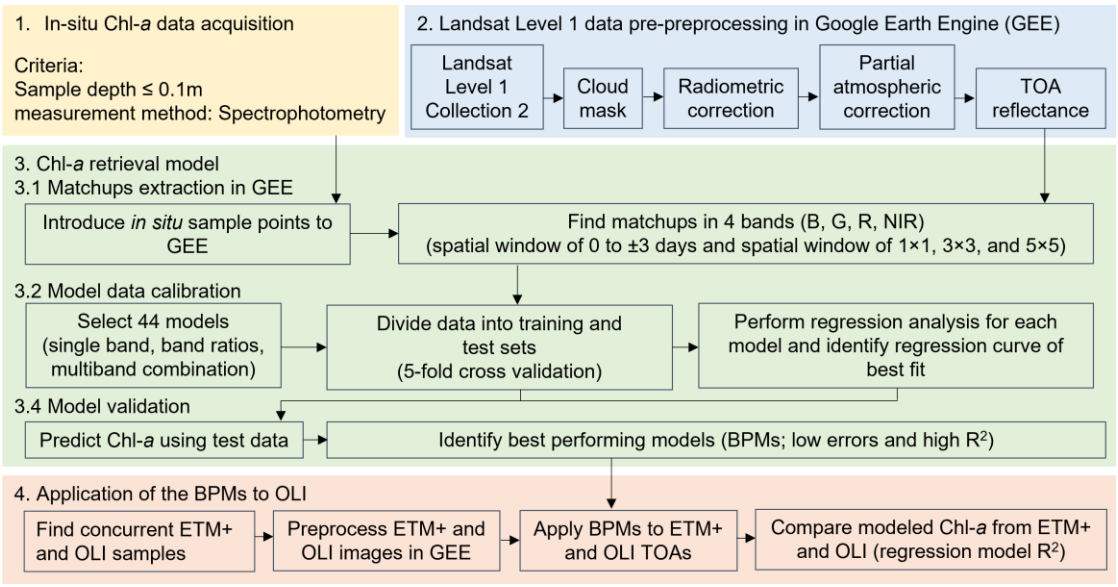


Figure 2. Flowchart of methods used in the study.

Step 1. Radiometric correction

Landsat Level-1 data are stored as digital numbers (DNs). During the radiometric correction, we recalibrated the DNs to the satellite TOA radiance ($L\lambda$) using the standard equation [27] (1):

$$L_\lambda = (DN_\lambda \times G_\lambda) + B_\lambda \quad (1)$$

Where:

L_λ is TOA radiance for band λ,

DN_λ is DN for band λ,

G_λ is the multiplicative rescaling factor for band λ, and

B_λ is the additive rescaling factor for band λ.

Step 2. Partial atmospheric correction:

L_λ values were corrected for Rayleigh scattering using a partial atmospheric correction algorithm [28]. We used an inverse algorithm based on a simplified radiative transfer model to calculate Rayleigh scattering radiance (or Rayleigh path radiance) for each band ($L_{r(\lambda)}$) [28] (2):

$$L_{r(\lambda)} = \left(\frac{ESUN_\lambda \times cos\theta_s \times P_r}{4\pi \times (cos\theta_s + cos\,\theta)}\right) \times \left(1 - exp\left(-\tau_r(\lambda) \times \left(\left(\frac{1}{cos\theta_s}\right) + \left(\frac{1}{cos\,\theta}\right)\right)\right)\right) \times t_{oz}\uparrow(\lambda) \times t_{oz}\downarrow(\lambda) \quad (2)$$

Where:

$ESUN$ is the mean solar exo-atmospheric irradiance,

P_r is the Rayleigh scattering phase function,

θ_s is the solar zenith angle (degrees),

θ is the satellite viewing angle (degrees) (set to 0°),

τ_r is the Rayleigh optical thickness, and

$t_{oz}\uparrow$ and $t_{oz}\downarrow$ are upward and downward ozone transmittance, respectively.

The phase function P_r was calculated using (3):

$$P_r = \frac{3}{4} \times \frac{1-\gamma}{1+2\gamma} \times (1 + cos^2\Theta) + \frac{3\gamma}{1+2\gamma} \quad (3)$$

Where:

Θ (the view zenith angle) is the scattering angle ($180° - \theta_s$),

$\gamma = \delta/(2- \delta)$ [29], and

δ is the depolarization factor [30], which denotes the polarization of anisotropic particles at right angles and is dependent on the wavelength, atmospheric pressure, and air mass (with the last two variables being constant values) [31].

 doi:10.20944/preprints202408.0962.v1

The Rayleigh optical thickness, τ_r, was calculated using [32,33] (4):

$$\tau_r = 0.008569\lambda^{-4} \times (1 + 0.0113\lambda^{-2} + 0.00013\lambda^{-4}) \quad (4)$$

Where:
λ is the band specific mid-wavelength value (nm).
$t_{oz}\uparrow$ and $t_{oz}\downarrow$ were calculated using [34] (5,6):

$$t_{oz} \uparrow = exp(-\tau_{oz}) \quad (5)$$

$$t_{oz} \downarrow = exp\left(\frac{-\tau_{oz}}{cos\theta_s}\right) \quad (6)$$

Where:
t_{oz} is the constant ozone optical thickness [35].
Finally, the Rayleigh-corrected radiance; $\hat{L}_\lambda$ was obtained by subtracting the Rayleigh scattering radiance (Lr) from the TOA radiance (L_λ) (7):

$$\hat{L}_{\lambda=} L_\lambda - L_r \quad (7)$$

Step 3. Converting radiance to reflectance:
TOA reflectance for each band ($\rho\lambda$), was calculated using (8):

$$\rho\lambda = \frac{\pi \times \hat{L}_\lambda \times d^2}{ESUN_\lambda \times cos\theta_s} \quad (8)$$

Where:
$\rho\lambda$ is TOA reflectance for wavelength range or band λ, and
d is the Earth-to-sun distance in astronomical units [27].

2.3.4. Extracting Matchups

We extracted remote sensing reflectance (R_{rs}) for each band (B, G, R, and NIR; Section 2.3.3) and paired it with the corresponding sampling points (Section 2.3.1) in GEE (See Supplementary Material). To increase the chance of matching satellite observations with sampled Chl-*a*, we considered Landsat images within ±3 days of the sampling dates. Larger temporal windows were not used due to the highly dynamic nature of blooms in LW [36]. Satellite data were extracted from corresponding pixels to sampling point locations (single pixels; spatial windows of 1 × 1). To mitigate potential errors in digitization [37], we also extracted R_{rs} using the median of pixels corresponding to sampling locations and their eight (spatial window of 3 × 3) and 24 neighboring pixels (spatial window of 5 × 5). Based on these criteria, a total of 42 samples (18 in the NB and 24 in the SB) from 28 stations (13 in the NB and 15 in the SB) were identified (Figure 1, Table 1). We found 9 Landsat images for the NB, 11 for the SB, and 20 for the entire LW matching the time and location of the sampling points, providing 23 matchup points in the NB, 30 in the SB, and 57 for the entire LW (Table 1). We created 36 sets of matchups with different temporal windows (0 to ±3 days), different spatial windows (1 × 1, 3 × 3, and 5 × 5), and in each LW basin (NB-specific and SB-specific) as well as for the entire lake (LW-specific).

Table 1. Matchups for sensitivity tests for the lake (NB-specific, SB-specific, and LW-specific) and temporal windows (0 to ±3 days). The values are the same for different spatial windows (1×1, 3×3, and 5×5).

Basin	Temporal window (days)	# of matchups	*In situ* Chl-*a* (μgL^{-1})	# of stations	# of samples	# of images	Landsat sensors	Landsat scenes (path/row)	Year	Month
NB	0	2	10.1 - 147	2	2	1	5	3322	2011	9
	±1	10	1.91 - 147	6	8	6	5, 7	3223, 3322/23	2010, 2011	7, 8, 9, 10

 doi:10.20944/preprints202408.0962.v1

	±2	17	1.91 - 147	11	13	8	5, 7	3223, 3322/23	2010, 2011	7, 8, 9, 10
	±3	23	1.91 - 147	13	18	9	5, 7	3223, 3322/23	2010, 2011	7, 8, 9, 10
SB	0	3	3.05 – 4.01	2	2	2	5, 7	3124/25	2010	7, 8
	±1	7	3.05 – 6.68	5	5	5	5, 7	3025, 3124/25	2010, 2011	7, 8
	±2	20	2.67 – 9.55	12	15	8	5, 7	3025, 3124/25	2010, 2011	7, 8, 10
	±3	30	2.67 - 147	15	24	11	5, 7	3025, 3124/25	2010, 2011	7, 8, 9, 10
LW	0	5	3.05 - 147	4	4	3	5, 7	3322, 3124/25	2010, 2011	7, 8, 9
	±1	17	1.91 - 147	11	13	11	5, 7	3223, 3025, 3124/25, 3322/23	2010, 2011	7, 8, 9, 10
	±2	27	1.91 - 147	23	28	16	5, 7	3223, 3025, 3124/25, 3322/23	2010, 2011	7, 8, 9, 10
	±3	57	1.91 - 147	28	42	20	5, 7	3223, 3025, 3124/25, 3322/23	2010, 2011	7, 8, 9, 10

2.3.5. Model Calibration

For each set of matchups (Section 2.3.4), we tested the performance of 44 models, including all possible bands and band ratios, as well as various commonly used multiband combinations (Supplementary Material, Table S1) to predict Chl-*a* in LW using regression analyses. Each model was fitted using the matchup natural log-transformed *in situ* Chl-*a* values against the partially corrected R_{rs} (with and without natural log-transformation) using simple linear regression equations (Table S1). We then identified the Best Performing Model (BPM) for each set of matchups as the model that exhibited the highest coefficient of determination (R^2) and was statistically significant ($p < 0.05$). Model calibration on the BPMs was performed using 5-fold cross-validation, chosen to balance bias and variance given the number of matchups available [38,39]. During model calibration, for each fold of the 5-fold cross-validation, the models were fitted to the training dataset (natural log-transformed *in situ* Chl-*a* values against partially corrected R_{rs}) using simple linear regression equations. We then estimated the average predictive performance of each model during the calibration stage using the average R^2 of the five cross-validation folds.

2.3.6. Model Validation

To assess the predictive capacity of the BPMs identified during the initial calibration phase for each set of matchups (Section 2.3.5), we validated each model by applying the slope and intercept from each calibration to its associated test matchup set in the cross-validation fold (k = 5). The predictive performance of the BPMs was measured the predictive performance of the BPMs using Root Mean Square Error (RMSE; μgL^{-1}) for each cross-validation fold as follows (9):

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n}\left(\frac{\left((y_o - y_p)^2\right)}{n}\right)} \quad (9)$$

Where:
y_o represents the observed Chl-*a* values, and
y_p represents the predicted Chl-*a* values.
We normalized the RMSE (NRMSE) for each cross-validation fold as follows (10):

$$\mathrm{NRMSE} = \frac{\sqrt{\sum_{i=1}^{n}\left(\frac{\left((y_o - y_p)^2\right)}{n}\right)}}{\overline{yo}} \quad (10)$$

Where:
$\overline{yo}$ is average observed Chl-*a*.
We calculated the Root Mean Squared Logarithmic Error (RMSLE; μgL^{-1}) as follows (11):

 doi:10.20944/preprints202408.0962.v1

$$\text{RMSLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(log[1+y_o] - log[1+y_p])^2} \quad (11)$$

We also calculated the Mean Absolute Error (MAE; μgL^{-1}) and Mean Absolute Percentage Error (MAPE; %) for the five cross-validation folds to measure the predictive performance of the models as follows (12,13):

$$\text{MAE} = \frac{\sum_{i=1}^{n}|y_o - y_p|}{n} \quad (12)$$

$$\text{MAPE} = 100 \times \text{median of}\left[\frac{|y_o - y_p|}{y_p}\right] for\ 1 = 1, \ldots, n. \quad (13)$$

Bias was measured as follows (14):

$$\text{Bias} = \frac{\sum_{i=1}^{n}(y_o - y_p)}{n} \quad (14)$$

Finally, we compared the performances of the models based on the average errors for all the cross-validation folds [40].

2.4. Application of the BPMs to Landsat OLI

The Landsat TM and ETM+ were launched in 1984 and 1999 (respectively), and the Landsat Operational Land Imager (OLI) was launched in 2013. In our study, all matchups since 2013 were discarded because the *in situ* Chl-*a* data in those years were taken from depths below 0.1 m, which did not meet our matchup selection criteria (Section 2.3.1). To test the applicability of our BPMs for predicting Chl-*a* using OLI, we applied the models to concurrent OLI and ETM+ samples (with 1 to 8 days of difference between the two sensors overpassing times) from 19 stations distributed throughout LW between 2013 to 2017 (Figure 1).

For each OLI matchup, Level 1 Collection 2 OLI data (https://www.usgs.gov/landsat-missions/landsat-collection-2) for the bands used in the BPMs (i.e., B (450–515 nm), G (595–600 nm), and Red (630–680 nm)) were pre-processed in GEE using the same approach as for Landsat TM and ETM+ (Section 2.3.3). After applying the BPMs to ETM+ and their concurrent OLI data, we compared the modeled Chl-*a* derived from each sensor (a total of 26 matching points) using a simple linear regression model.

2.5. Comparing Chl-a Predictions Using Basin-Scale vs. Lake-Scale BPMs

To compare Chl-*a* predictions using basin-specific versus lake-specific BPMs, we applied the BPMs from each approach to all Landsat images during the peak Chl-*a* period (July–October) from 1984 to 2023 in GEE. We then compared the ability of the predictions from the two approaches to capture spatial and temporal changes in peak Chl-*a* in LW. All Landsat images were pre-processed to partially corrected R_{rs} in GEE (section 2.3.3) prior to model application.

To study spatial changes in peak Chl-*a* in LW, we calculated the median of peak Chl-*a* for each pixel within each basin for *all* years from 1984 to 2023. To study temporal changes in peak Chl-*a* in LW, we calculated the median peak Chl-*a* for all pixels within each basin for each year from 1984 to 2003.

3. Results

3.1. Spatial Heterogeneity in LW Turbidity

Our results revealed significant differences in turbidity levels among different regions of LW during the growing season (May-October). Specifically, turbidity in the NB was consistently lower

 doi:10.20944/preprints202408.0962.v1

than in the SB and the Narrows ($p < 0.05$), whereas no significant difference was observed between the SB and the Narrows ($p \geq 0.05$; Table 2).

Table 2. Comparing turbidity (Nephelometric Turbidity Units; NTU) in the two basins and the narrows (median values); number of sampling data are added in the parenthesis next to the median values. The letters a and b on the values within each month show if the basins/narrows were significantly different from each other based on the Dunn's multiple comparison test results ($p < 0.05$).

All months	May	June	July	Aug	Sep	Oct	All months
NB	5.5[a] (723)	-	5.39[a] (205)	5.53[a] (176)	3.81[a] (95)	5.55[a] (207)	9.44[a] (41)
SB	15.00[b] (396)	10.60[a] (20)	8.06[b] (107)	19,00[b] (58)	14.90[b] (73)	19.75[b] (58)	20.50[a,b] (80)
Narrows	16.1[b] (243)	5.50[b] (2)	9.10[b] (78)	21.80[b] (62)	17.45[b] (70)	17.90[b] (23)	14.70[b] (8)

Source: Raw *in situ* data from Manitoba Agriculture and Resource Development; Kruskal-Wallis and Dunn's test ($p < 0.05$). Data were not normally distributed (Shapiro Wilcoxon test; $p > 0.05$).

3.2. Best Performng Models

The performance of basin-specific models in predicting Chl-*a* was superior to that of lake-specific models (Figure 3 and Table S1). The NB-specific and SB-specific BPMs achieved R^2 values of 0.85 ($n = 10$) and 0.76 ($n = 7$), respectively ($p < 0.05$; Figure 3 and Tables S2, S3). In contrast, the lake-specific BPM showed a lower R^2 of 0.38 ($p < 0.05$; $n = 17$; Figure 3 and Table S4).

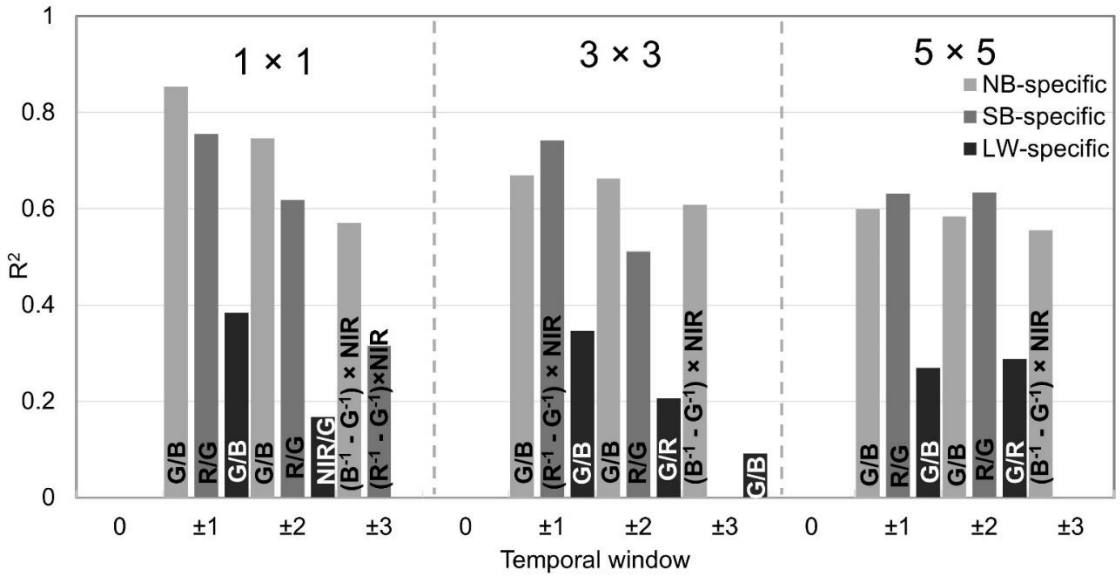


Figure 3. Correlation coefficients (R^2) of the best performing model among 44 tested algorithms for 36 sets of matchups with different temporal windows (0 to ±3 days) and different spatial windows (1 × 1, 3 × 3, and 5 × 5) in each Lake Winnipeg basin (i.e., NB- and SB-specific) and in the entire lake (i.e., LW-specific). No values on the plot refers to correlation coefficients with no significant relationship between *in situ* and satellite Chl-*a* ($p \geq 0.05$) or to the number of matchups below 10.

The NB-specific BPM ($R^2 = 0.85$) used matchups from ±1 day with a 1 × 1 spatial window (Figure 3 and Table S2). Increasing the spatial window to 3 × 3 and 5 × 5 decreased the R^2 to 0.70 and 0.60, respectively, while increasing the temporal window to ±2 and ±3 days decreased the R^2 to 0.75 and 0.57, respectively. Despite these changes, increasing the temporal window to ±2 days still provided a high R^2 of 0.75 with reduced errors (NRMSE decreased from 1.98 to 0.73; bias reduced from 9.39 to 2.04).

The SB-specific BPM ($R^2 = 0.76$) used matchups from ±1 day with a 1 × 1 spatial window (Figure 3 and Table S3). Increasing the spatial window to 3 x 3 and 5 x 5 decreased the R^2 to 0.74 and 0.63, respectively, while increasing the temporal window to ±2 and ±3 days decreased the R^2 to 0.61 and

 doi:10.20944/preprints202408.0962.v1

0.32, respectively. The temporal window of ±2 days provided 17 matchups with an R^2 of 0.61 and reduced bias (bias decreased from 0.25 to 0.23).

Like the basin-specific models, the LW-specific model (R^2 = 0.38, based on matchups from ±1 day with a 1 × 1 spatial window) exhibited decreasing R2 values with larger spatial and temporal windows (Figure 3 and Table S4).

3.3. Best Chl-a Prediction Models

From the BPMs in section 3.2, the basin-specific models (NB and SB) using a 1 × 1 spatial window and a temporal window of ±2 days were selected for their balanced performance (high R^2 and low errors) and increased confidence due to a larger number of matchups used.

In the NB, a 2-band model based on the R_{rs} of the band ratio of G to B showed the highest performance (R^2 = 0.74) (Table 3 and Figure 4a) (15):

$$\ln \text{Chl} - a = 8.153 \times [G/B] - 5.622 \quad (15)$$

In the SB, a 2-band model based on the R_{rs} of the band ratio of R to G performed best (R^2 = 0.62) (Table 3 and Figure 4b) (16):

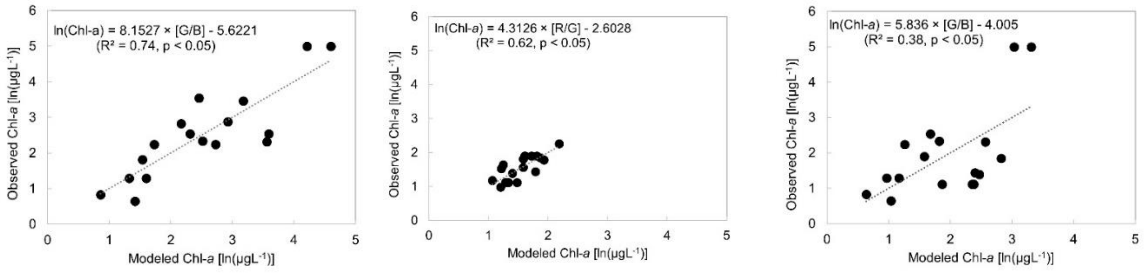
$$\ln \text{Chl} - a = 4.313 \times [R/G] - 2.603 \quad (16)$$

In the entire LW, a 2-band model based on the R_{rs} of the band ratio of G to B achieved the highest performance (R^2 = 0.38) (Table 3 and Figure 4c) (17):

$$\ln \text{Chl} - a = 5.836 \times [G/B] - 4.005 \quad (17)$$

Table 3. Calibration and validation results for the NB- and SB-specific BPMs (spatial window of 1× 1 and temporal window of ±2 days) and the LW-specific BPM (spatial window of 1× 1 and temporal window of ±1 days). All the calibrations R^2 are significant at $p < 0.05$.

Basin	Model	Calibration R^2	RMSE (μgL^{-1})	RMSLE (μgL^{-1})	NRMSE	MAE (μgL^{-1})	MAPE (%)
NB	G/B	0.74	20.53	0.65	0.88	14.51	55.43
SB	R/G	0.62	1.14	0.20	0.24	1	22.81
LW	G/B	0.38	21.57	0.87	1.29	13.57	64.27



(a) **(b)** **(c)**

Figure 4. Regression of observed ln Chl-*a* (y axis; μgL^{-1}) to modeled ln Chl-*a* (x axis; μgL^{-1}) using the (a) NB-, (b) SB-, and (c) LW-specific best performing models.

Validation tests using 5-fold cross-validation further supported the superiority of basin-specific models over the LW-specific model (Table 3).

3.4. Application of the Models to Landsat OLI

Cross-comparison between ETM+ and OLI sensors indicated a strong linear relationship between Chl-*a* predictions using OLI (dependent variable) and those using ETM+ (independent

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 August 2024 doi:10.20944/preprints202408.0962.v1

variable) ($R^2 = 0.87$, $p < 0.05$) (Figure 5). Notably, the correlation was stronger in the NB ($R^2 = 0.83$, $p < 0.05$) compared to the SB ($R^2 = 0.75$, $p < 0.05$) (Figure 5).

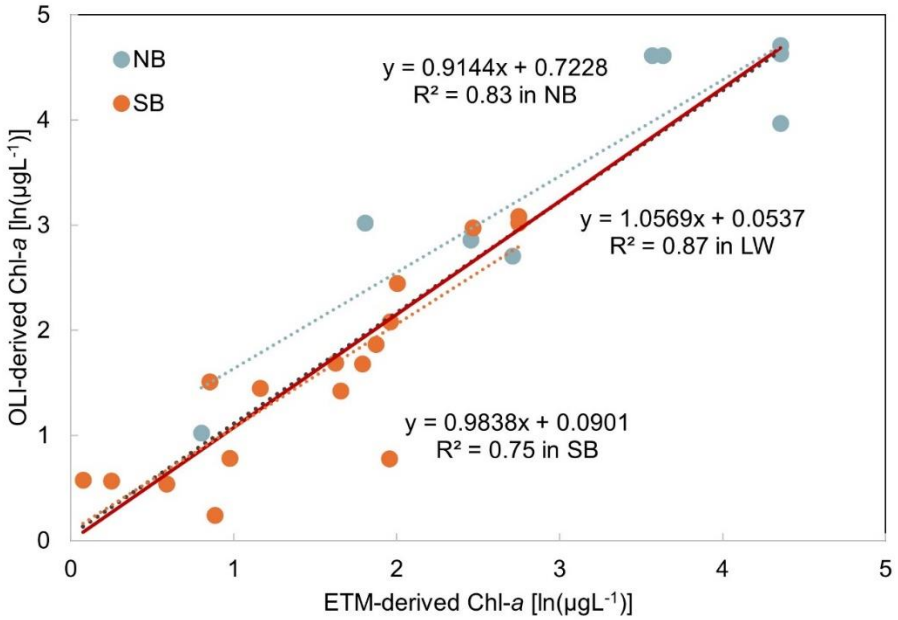


Figure 5. Chl-*a* predictions using OLI (y axis; μgL^{-1}) and ETM+ (x axis; μgL^{-1}) for NB (blue dots and line; $R^2 = 0.83$), SB (orange dots and line; $R^2 = 0.75$) and the entire Lake Winnipeg (LW) (black line; $R^2 = 0.87$). The red line shows the intercept ($p < 0.05$).

3.5. Comparing Chl-a Predictions Using the Basin-Scale vs Lake-Scale BPMs

Figure 6 illustrates the median (1984-2023) of natural log-transformed peak Chl-*a* (μgL^{-1}) in the NB and SB using basin-specific (Figure 6a) and lake-specific BPMs (Figure 6b). Basin-specific models predicted higher Chl-*a* concentrations in the NB compared to the SB (Figure 6a), while lake-specific models showed the opposite pattern (Figure 6b)

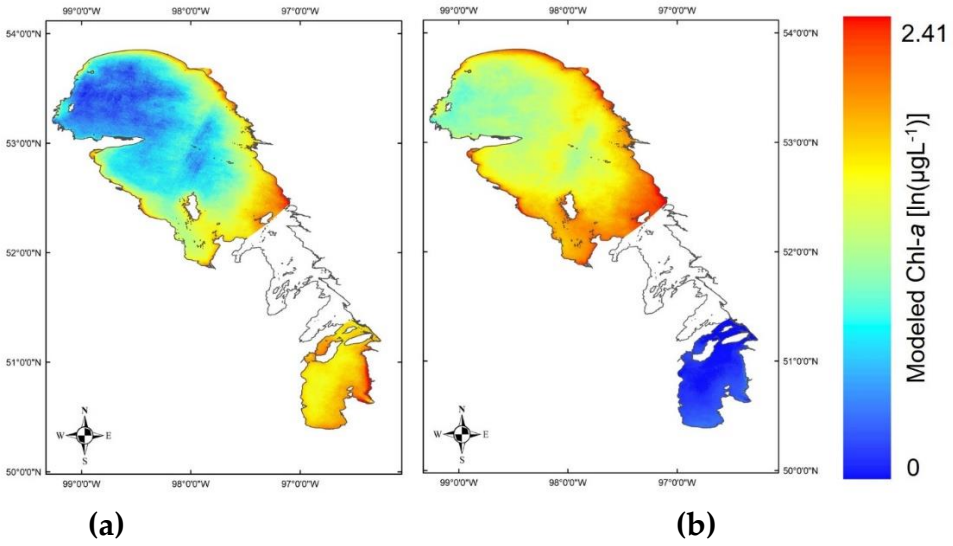


Figure 6. Chl-*a* predictions using OLI (y axis; μgL^{-1}) and ETM+ (x axis; μgL^{-1}) for NB (blue dots and line; $R^2 = 0.83$), SB (orange dots and line; $R^2 = 0.75$) and the entire Lake Winnipeg (LW) (black line; $R^2 = 0.87$). The red line shows the intercept ($p < 0.05$).

 doi:10.20944/preprints202408.0962.v1

Figure 7 illustrates the annual time series of peak Chl-*a* in the NB and SB derived from basin-specific and lake-specific BPMs. In the NB, lake-specific models consistently predicted lower Chl-*a* concentrations compared to basin-specific models (Figure 7). Conversely, in the SB, lake-specific models predicted higher Chl-*a* concentrations compared to basin-specific models (Figure 7). Temporal variations in peak Chl-*a* from 2002 to 2019 also exhibited distinct patterns between basin-specific and lake-specific models, reflecting different spatial dynamics within LW (Figure 7).

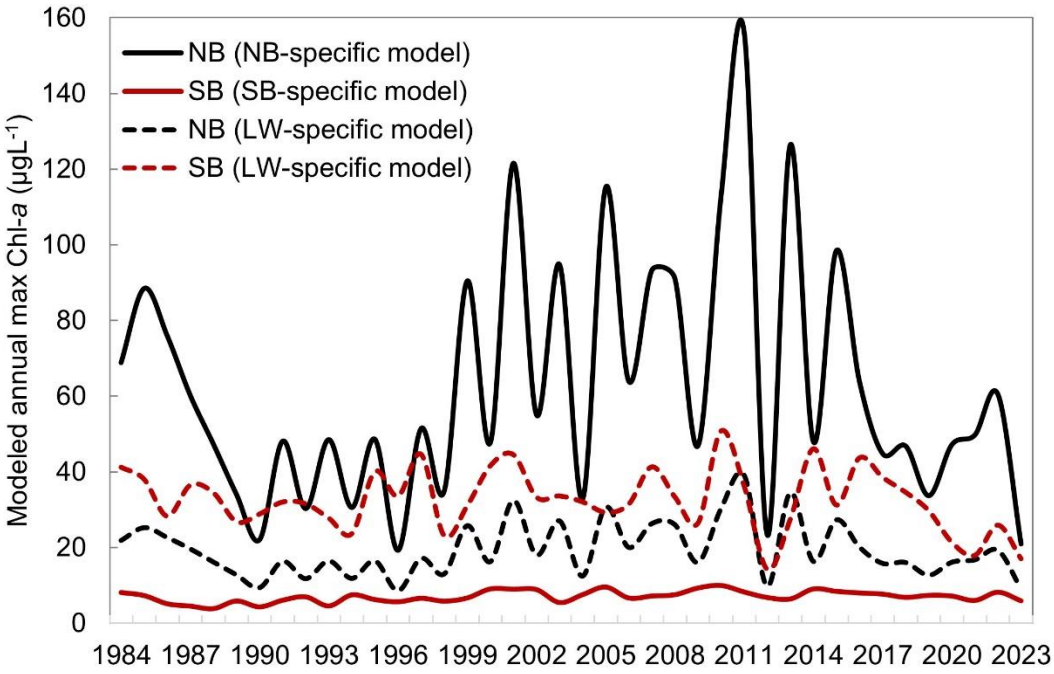


Figure 7. Annual time series (1984 to 2023) of peak Chl-*a* (μgL^{-1}) in the NB (black lines) and SB (red lines) using basin-specific (solid lines) and lake-specific (dashed lines) models.

4. Discussion

Remote sensing is a crucial tool for the rapid and consistent monitoring of spatial and temporal changes in surface water Chl-*a*. Limnologists often use global satellite-derived level-2 or level-3 Chl-*a* products designed for optically complex waters [36,41,42] without considering variations in water optical properties among lakes [40]. Chl-*a* prediction using remote sensing reflectance can be improved by developing models that account for potential within-lake variance in optical properties. Such models can be delineated based on optical properties or, as in our study, by morphometric characteristics that distinguish lake basins with unique optical properties. Our study demonstrates that by considering within-basin differences in optical properties and dividing the lake into optically distinct parts, Chl-*a* prediction can be improved.

4.1. Basin-Specific Chl-a Prediction Models

The physical and biogeochemical differences between LW's NB and SB [19,36] create different optical properties, favoring basin-specific models over lake-specific models in both basins. BPMs in both basins used band ratios. Compared to single bands, band ratios enhance Chl-*a* estimations by reducing the atmospheric, irradiance, and air-water surface effects on reflectance [43]. Band ratios also minimize the impact of non-phytoplankton OACs found in inland waters, which can alter the spectral profile observed by the sensor [44,45]. This is significant because the varying concentrations of these constituents can explain the lack of correlation between observed Chl-*a* and single bands in lakes [43], as observed in our study (Supplementary Material).

Many models based on band ratios have been proposed for estimating Chl-*a* in inland, coastal, and open ocean waters [11,46]. These models leverage the reflection and absorption patterns of

 doi:10.20944/preprints202408.0962.v1

sunlight by Chl-*a*. With increasing Chl-*a*, the relative energy between 520 and 600 nm (Landsat G band) increases, while the energy between 450 and 520 nm (Landsat B band) decreases [47]. In our study, while BPMs in both NB and SB included the G band, the second band in the ratio differed between the two basins.

In the NB, where the water is deeper and clearer [19], Chl-*a* reaches higher concentrations during bloom events [19,35]. In this basin, the B to G band ratio provided the highest correlation with observed Chl-*a* ($R^2 > 0.80$). The B to G band ratio is widely used in estimating Chl-*a* in lakes with low concentrations of inorganic suspended matter or in eutrophic and hypertrophic waters where phytoplankton are the predominant constituents and non-phytoplankton constituents have a smaller effect on reflectance [48–50].

In the SB, where the water is shallow, well-mixed, and highly turbid [19,36], the R to G band ratio provided the highest correlation with observed Chl-*a* ($R^2 > 0.70$). The SB creates optically complex water conditions, where non-phytoplankton OACs interfere with spectral bands – particularly the B band. The concentrations of these constituents do not always correlate with Chl-*a* [51]. This interference complicates Chl-*a* estimation in optically complex turbid waters using the B band [44,52,53]. To account for the overlaps of the absorption of non-phytoplankton particles, the R to G band ratio is commonly used for estimating Chl-*a* in highly turbid waters [43,54–56]. Many Chl-*a* prediction models developed for turbid, optically complex waters are also based on the red and near-infrared regions of the electromagnetic spectrum [43,57]. Like the red region, the near-infrared region (near 700 nm wavelength) shows less sensitivity to non-phytoplankton particles' interference compared to the blue region [58]. However, near-infrared-based models are successful with satellites having a band centered at the red- near-infrared edge, like MERIS [58,59]. Multispectral sensors like Landsat have broad NIR bands (770–900 nm in TM and ETM+ and 845–885 nm in OLI), which have been shown to be unsuitable for estimating Chl-*a* [43]. In our study, we could not improve the relationship between observed and modeled Chl-*a* using NIR in either basin-specific or lake-specific models.

4.2. Application of the Models to Landsat OLI

Our matchup selection criteria excluded *in situ* matchups concurrent with the period of Landsat OLI operation (2013 to present). OLI differs from TM/ETM+ in image quality and the number and width of spectral bands. Despite these differences, the similar spatial and spectral resolutions allow for intercalibration between these sensors [60]. Our cross-comparison analysis showed that the models developed in this study are applicable for predicting Chl-*a* using OLI. Other studies have also successfully applied models developed based on TM or ETM+ to OLI images or vice versa [3,61].

The matchup selection criteria in our study left out *in situ* matchups concurrent with the period of Landsat OLI operation (2013 to present). OLI is different from TM/ETM+ for image quality and number and width of spectral bands. Despite the differences between TM/ETM+ and OLI, the similar spatial and spectral resolutions allow intercalibration between these sensors [60]. Our cross-comparison analysis showed the models developed in this study are applicable to predict Chl-*a* using OLI. Similarly, other studies successfully applied models developed based on TM or ETM+ to OLI images or vis versa [e.g., 3,61].

4.3. Future Applications

Overall, the high correlation between modeled and observed Chl-*a* obtained using the basin-specific models suggests these models can reliably capture relative changes in Chl-*a* in LW. Therefore, the models can be used for predicting Chl-*a* in LW with Landsat data to study spatial and temporal variabilities of Chl-*a* [61,62]. However, the lack of *in situ* Chl-*a* data with concentrations above 10 µgL-1 in the SB due to our matchup selection criteria may affect the accuracy of absolute Chl-*a* predictions in this basin [61]. More *in situ* Chl-*a* data can also enable Chl-*a* predictions using an optical water types (OWTs) approach, which requires large *in situ* Chl-a data for training. Using this approach, lakes are classified into OWTs before prediction [40,63,64], with unique Chl-*a* prediction models developed for each water type. Our matchup selection criteria also excluded any samples from the

 doi:10.20944/preprints202408.0962.v1

Narrows. Like the SB, the Narrows is optically complex, where non-phytoplankton particles are primary determinants of water transparency [19]. However, applying the SB-specific model to the Narrows requires *in situ* Chl-*a* data to validate these models' performance in this part of the lake.

Applying models to Landsat imagery (available since 1984) in GEE (see Supplementary Material for the code) can provide the longest available Chl-*a* time series for most water bodies. The length of this time series (40 years) allows for the study of long-term climate trends or oscillations' impacts on changes in Chl-*a*. At least 30-40 years of continuous and consistent data on water temperature (a critical indicator of climate change) and Chl-*a* is required to separate climate change signals from natural variability [64–67]. This information will also provide ecosystem managers and scientists with a better understanding of the sensitivity of different regions of their study area to environmental perturbations. This can lead to more accurate predictions of future lake ecosystem responses to increasing anthropogenic pressures in the face of a changing climate. Ecosystem managers can use this knowledge to design contextual and more efficient management strategies towards eutrophication and phytoplankton blooms in lakes.

5. Conclusion

In this study, we predicted relative concentrations of Chl-*a* in a large lake using simple empirical equations applied to Landsat data. Our results highlight the importance of considering optical differences within lakes when developing Chl-*a* prediction models. By separating a lake into basins with distinct water optical properties, we significantly improved the relationships between *in situ* and modeled Chl-*a*. These models can be easily applied in cloud-based platforms like Google Earth Engine (GEE) for computationally intensive modeling of Chl-*a* on broad spatial and temporal scales. The user-friendly GEE code requiring minimal adjustments of the times and locations of *in situ* Chl-*a* samples that was developed for this study is freely available, facilitating the development of models for optically distinct lakes world-wide.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1: Best performing models' R^2 using *in situ* matchups from different depth, Table S1: Best calibration results for each tested model, Table S2: Calibration and validation results for the NB-specific Best Performing Models (BPMs), Table S3: Calibration and validation results for the SB-specific Best Performing Models (BPMs), Table S4: Calibration and validation results for the Lake-scale Best Performing Models (BPMs). Google Earth Engine codes for atmospheric correction of Landsat Level 1 data, masking land and low-quality pixels, extracting satellite matchups, applying Chl-*a* model to the Landsat time series, and exporting the modeled Chl-*a* images.

Author Contributions: Conceptualization, F.F., I.F.C., and C.G.T.; methodology, F.F., I.F.C., and C.G.T.; GEE code, F.F.; software, F.F.; validation, F.F.; formal analysis, F.F. and I.F.C.; investigation, F.F.; resources, F.F.; data curation, F.F.; writing—original draft preparation, F.F. and I.F.C.; writing—review and editing, F.F. and I.F.C.; visualization, F.F.; supervision, I.F.C. and C.G.T.; project administration, I.F.C.; funding acquisition, I.F.C.

Funding: The research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (05265-2019) and an Environment and Climate Change Canada - Climate Action and Awareness Fund (ECCC-CAAF) (EDF-CA-2021i023) grant awarded to I.F.C.

Data Availability Statement: This research used *in situ* Chl-*a* data provided by the Manitoba Sustainable Development, Water Quality Management Section 2019, and turbidity data from Manitoba Agriculture and Resource Development. These data are freely available from the sources upon request. Landsat Collection 2 data provided by the U.S. Geological Survey (USGS) are available in Google Earth Engine achieve and from https://www.usgs.gov/landsat-missions/landsat-collection-2.

Acknowledgments: The authors would like to thank Ian Mackenzie for his assistance in GEE scripting and David Aldred and Michael Dallosch for their technical support.

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

 doi:10.20944/preprints202408.0962.v1

References

1. Huisman, J.; Codd, G.A.; Paerl, H.W.; Ibelings, B.W.; Verspagen, J.M.; Visser, P.M. Cyanobacterial blooms. *Nat. Rev. Microbiol.* **2018**, 16(8), 471-483. doi: 10.1038/s41579-018-0040-1
2. Ho, J.C.; Michalak, A.M.; Pahlevan, N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature* **2019**, 574 (7780), 667–670. doi: 10.1038/s41586-019-1648-7.
3. Hou, X.; Feng, L.; Dai, Y.; Hu, C.; Gibson, L.; Tang, J.; Lee, Z.; Wang, Y.; Cai, X.; Liu, J.; Zheng, Y. Global mapping reveals increase in lacustrine algal blooms over the past decade. *Nat. Geosci.* **2022**, 15(2), 130–134. doi: 10.1038/s41561-021-00887-x
4. Randolph, K.; Wilson, J.; Tedesco, L.; Li, L.; Pascual, D. L.; Soyeux, E. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll-a and phycocyanin. *Remote. Sens. Environ.* **2008**, 112(11), 4009–4019. doi: 10.1016/J.RSE.2008.06.002.
5. Reid, J.L.; Bergman, J.N.; Kadykalo, A.N.; Taylor, J.J.; Twardek, W.; Rytwinski, T.; Chhor, A.D.; Frempong-Manso, A.; Martel, A.L.; Lapointe, N.W.R.; Bennett, J.R.; Nguyen, V.M.; Reid A.J.; Marty, J.; Robinson, S.A.; Drake, A.R.; Winegardner, A.K.; Gregory-Eaves, I.S.; Taylor, M.K.; Smol, J.P.; Creed, I.F.; O'Connor, C.M.; Cooke, S.J. Developing a national level evidence-based toolbox for addressing freshwater biodiversity threats. *Biol. Conserv*. **2022**, 269:109533. doi: https://doi.org/10.1016/j.biocon.2022.109533.
6. An, K.G.; Lee, J.Y.; Kumar, H.K.; Lee, S.J.; Hwang, S.J.; Kim, B.H.; Park, Y.S.; Shin, K.H.; Park, S.; Um, H.Y. Control of algal scum using top-down biomanipulation approaches and ecosystem health assessments for efficient reservoir management. *Water. Air. Soil. Pollut.* **2010**, 205(1–4), 3–24. doi: 10.1007/s11270-009-0053-5.
7. Weirich, C.A.; Miller, T.R. Freshwater harmful algal blooms: Toxins and children's health. *Curr. Probl. Pediatr. Adolesc. Health Care* **2014**, 44(1), 2–24. doi: 10.1016/j.cppeds.2013.10.007.
8. Binding, C.E.; Pizzolato, L.; Zeng, C. EOLakeWatch; delivering a comprehensive suite of remote sensing algal bloom indices for enhanced monitoring of Canadian eutrophic lakes. *Ecol. Indic.* **2021**, 121, p.106999. doi: 10.1016/j.ecolind.2020.106999.
9. Guo, Q.; Wu, X.; Bing, Q.; Pan, Y.; Wang, Z.; Fu, Y.; Wang, D.; Liu, J. Study on retrieval of chlorophyll-a concentration based on Landsat OLI Imagery in the Haihe River, China. *Sustainability* **2016**, 8(8), 1-15. doi: 10.3390/su8080758.
10. Kislik, C.; Dronova, I.; Grantham, T.E.; Kelly, M. Mapping algal bloom dynamics in small reservoirs using Sentinel-2 imagery in Google Earth Engine. *Ecol. Indic.* **2022**, 140, 1-12. doi: 10.1016/j.ecolind.2022.109041.
11. Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters, *Int. J. Remote Sens.* **2011**, 32(21), 6855–6899, doi:10.1080/01431161.2010.512947.
12. Sathyendranath, S.; Brewin, R.J.W.; Jackson, T.; Mélin, F.; Platt, T. Ocean-colour products for climate-change studies: What are their ideal characteristics? *Remote Sens. Environ.* **2017**, 203, 125–138. doi: 10.1016/j.rse.2017.04.017.
13. Kuhn, C.; de Matos 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. Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity, *Remote Sens. Environ.* **2019**, 224, 104–118, doi:10.1016/J.RSE.2019.01.023
14. Watanabe, F.; Alcantara, E.; Rodrigues, T.; Rotta, L.; Bernardo, N.; IMAI, N. Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil), *An. Acad. Bras. Cienc.* **2017**, 90(2 suppl 1), 1987–2000, doi:10.1590/0001-3765201720170125.
15. Tan, W.; Liu, P.; Liu, Y.; Yang, S.; Feng, S.; Tan, W.; Liu, P.; Liu, Y.; Yang, S.; Feng, S. A 30-year assessment of phytoplankton blooms in Erhai Lake using Landsat imagery: 1987 to 2016. *Remote Sens.* **2017**, 9(12), 1265–1280, doi:10.3390/rs9121265.
16. Sass, G. Z., Creed, I. F., Bayley, S. E. and Devito, K. J.: Understanding variation in trophic status of lakes on the Boreal Plain: A 20-year retrospective using Landsat TM imagery. *Remote Sens. Environ*. **2007**, 109(2), 127–141, doi: 10.1016/j.rse.2006.12.010.
17. Ho, J.C.; Stumpf, R.P.; Bridgeman, T.B; Michalak, A.M. Using Landsat to extend the historical record of lacustrine phytoplankton blooms: a Lake Erie case study. *Remote Sens. Environ.* **2017**, 191, 273–285. doi: 10.1016/j.rse.2016.12.013.
18. Sayers, M.J.; Bosse, K.R.; Shuchman, R.A.; Ruberg, S.A.; Fahnenstiel, G.L.; Leshkevich, G.A.; Stuart, D.G.; Johengen, T.H.; Burtner, A.M.; Palladino, D. Spatial and temporal variability of inherent and apparent optical properties in western Lake Erie: Implications for water quality remote sensing. *J. Great Lakes Res.* **2019**, 45(3), 490–507. doi: 10.1016/j.jglr.2019.03.011.

19. Environment Canada Manitoba Water Stewardship: State of Lake Winnipeg: 1999–2007, Manitoba Water Stewardship and Environment Canada, Winnipeg., **2011**.
20. Ulrich, A.E.; Malley, D.F.; Watts, P.D. Lake Winnipeg Basin: advocacy, challenges and progress for sustainable phosphorus and eutrophication control. *Sci. Total Environ*. **2016**, 542, 1030–1039. doi: 10.1016/j.scitotenv.2015.09.106.
21. Zar, J.H. (Ed.): Biostatistical Analysis, fourth ed. Prentice Hall, Upper Saddle River, NJ, **1999**.
22. Wheater, C.P.; Cook, P.A. (Eds.): Using statistics to understand the environment, The Taylor & Francis e-Library, London and New York, **2005**.
23. APHA (American Public Health Association), AWWA (American Water Works Association), and WPCF (Water Pollution Control Federation), **1998**. Standard Methods for the Examination of Water and Wastewater (20th Edition). APHA, New York, New York.
24. Campbell, J.W. The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res.* **1995**, 100, 13237-13254. doi: 10.1029/95JC00458.
25. Guanter L.; Ruiz-Verdu, A.; Odermatt, D.; Giardino, C.; Simis, S.; Heege, T.; Domínguez- Gómez, J.A.; Moreno, J. Atmospheric correction of ENVISAT/MERIS data over inland waters: validation for European Lakes. *Remote Sens. Environ*. **2008**, 114, 467–480. doi: 10.1016/j.rse.2009.10.004.
26. Lobo, F.L.; Costa, M.P.F.; Novo, E.M.L.M. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. *Remote Sens. Environ.* **2015**, 157, 170-184. doi:10.1016/j.rse.2014.04.030
27. Chander, G.; Markham, B.; Helder, D. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* **2009**, 113, 893–903. doi: 10.1016/j.rse.2009.01.007.
28. Gilabert, M.A; Conese, C.; Maselli, F. An atmospheric correction method for the automatic retrieval of surface reflectances from TM images. *Int. J. Remote Sens.* **1994**, 15, 2065–2086. doi: 10.1080/01431169408954228.
29. Bucholtz, A. Rayleigh-scattering calculations for the terrestrial atmosphere. *Appl. Opt.* **1995**, 34, 2765–2773. doi: 10.1364/AO.34.002765.
30. Young, A.T. Revised depolarization corrections for atmospheric extinction. *Appl. Opt.* **1980**, 19(20), 3427. doi:10.1364/AO.19.003427.
31. Vermote E.; Tanré, D.; Deuzé, J.L.; Herman, M.; Morcrette, J.J.; Kotchenova S.Y. **2006**, Second simulation of a satellite signal in the solar spectrum - Vector (6SV). MODIS land surface reflectance science computing facility, user manual part two.
32. Bodhaine, B.A.; Wood, N.B.; Dutton, E.G.; Slusser, J.R. On Rayleigh optical depth calculations, *J. Atmos. Ocean. Tech.* **1999**, 16, 1854–1861. doi: 10.1175/1520-0426(1999)016<1854:ORODC>2.0.CO;2.
33. Hansen, J.E.; Travis, L.D. Light scattering in planetary atmospheres. *Space Sci. Rev.* **1974**, 16, 527–610, doi:10.1007/BF00168069.
34. Sturm, B. **1981**. The atmospheric correction of remotely sensed data and the quantitative determination of suspended matter in marine water surface layers. In Remote Sensing in Meteorology, Oceanography and Hydrology, edited by A. P. Cracknell (Chichester: Ellis Horwood Limited), Chapter 11.
35. Jorge, D.S.; Barbosa, C.C.; De Carvalho, L.A.; Affonso, A.G.; Lobo, F.D.L.; Novo, E.M.D.M. SNR (signal-to-noise ratio) impact on water constituent retrieval from simulated images of optically complex Amazon lakes. *Remote Sens* **2017**, 9(7), 1–18. doi:10.3390/rs9070644.
36. Binding, C.E.; Greenberg, T.A.; McCullough, G.; Watson, S.B.; Page, E. An analysis of satellite-derived chlorophyll and algal bloom indices on Lake Winnipeg. *J. Great Lakes Res.* **2018**, 44(3), 436–446. doi: 10.1016/j.jglr.2018.04.001.
37. Laiolo, L.; Matear, R.; Soja-Woźniak, M.; Suggett, D.J.; Hughes, D.J.; Baird, M.E; Doblin, M.A. Modelling the impact of phytoplankton cell size and abundance on inherent optical properties (IOPs) and a remotely sensed chlorophyll-a product. *J. Mar. Syst.* **2021**, 213, 1–12. doi: 10.1016/j.jmarsys.2020.103460.
38. Hastie, T.; Tibshirani, R.; Friedman, J. John Lu, Z.Q., **2010**. The elements of statistical learning: data mining, inference, and prediction. 2nd Edition, Springer. 2017. NY. 745p.
39. Yates, L.A.; Aandahl, Z.; Richards, S.A.; Brook, B.W. Cross validation for model selection: a review with examples from ecology. *Ecol. Monogr.* **2023**, 93(1), p.e1557. doi: 10.1002/ecm.1557.
40. Dallosch, M.A.; Creed, I.F. Optimization of Landsat chl-a retrieval algorithms in Freshwater Lakes through classification of optical water types. J. *Remote Sens.* **2021**, 13(22), p.4607.

41. Ali, G.; English, C. Phytoplankton blooms in Lake Winnipeg linked to selective water-gatekeeper connectivity. *Scientific reports* **2019**, 9(1), p.8395. doi: 10.1038/s41598-019-44717-y.
42. Zhang, Y.; Ma, R.; Zhang, M.; Duan, H.; Loiselle, S.; Xu, J. Fourteen-year record (2000–2013) of the spatial and temporal dynamics of floating algae blooms in Lake Chaohu, observed from time series of MODIS images. *Remote Sens.* **2015**, 7(8), 10523–10542. doi: 10.3390/rs70810523.
43. Ha, N.T.T.; Koike, K.; Nhuan, M.T.; Canh, B.D.; Thao, N.T.P.; Parsons, M. Landsat 8/OLI two bands ratio algorithm for chlorophyll-a concentration mapping in hypertrophic waters: An application to West Lake in Hanoi (Vietnam). *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, 10(11), 4919–4929. doi: 10.1109/JSTARS.2017.2739184.
44. Keith, D.J.; Yoder, J.A.; Freeman, S.A. Spatial and Temporal Distribution of Coloured Dissolved Organic Matter (CDOM) in Narragansett Bay, Rhode Island: Implications for Phytoplankton on Coastal Waters. *Estuar. Coast Shelf Sci.* **2002**, 55, 705–717. doi:10.1006/ecss.2001.0922.
45. Vincent, R.K.; Qin, X.; McKay, R.M.L.; Miner, J.; Czajkowski, K.; Savino, J.; Bridgeman, T. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sens. Environ.* **2004**, 89(3), 381–392. doi: 10.1016/j.rse.2003.10.014.
46. Maeda, E.E.; Lisboa, F.; Kaikkonen, L.; Kallio, K.; Koponen, S.; Brotas, V.; Kuikka, S. Temporal patterns of phytoplankton phenology across high latitude lakes unveiled by long-term time series of satellite data, *Remote Sens. Environ.* **2018**, 221, 609–620, doi: 10.1016/j.rse.2018.12.006.
47. Nas, B.; Karabork, H.; Ekercin, S.; Berktay, A. Mapping chlorophyll-a through *in situ* measurements and Terra ASTER satellite data. *Environ. Monit. Assess.* **2009**, 157, 375–382. doi: 10.1007/s10661-008-0542-9.
48. Keith, D.; Rover, J.; Green, J.; Zalewsky, B.; Charpentier, M.; Thursby, G.; Bishop. J. Monitoring algal blooms in drinking water reservoirs using the Landsat-8 Operational Land Imager. *Int. J. Remote Sens.* **2018**, 39(9), 2818–2846. doi: 10.1080/01431161.2018.1430912.
49. Paltsev, A.; Creed, I.F. Multi-decadal changes in phytoplankton biomass in northern temperate lakes as seen through the prism of landscape properties. *Glob. Change Biol.* **2022**, 28:2272-2085. https://doi.org/10.1111/gcb.16079
50. Paltsev, A.; Creed, I.F. Are northern lakes in relatively intact temperate forests showing signs of increasing phytoplankton biomass? *Ecosyst.* **2022**, 25:727-755. https://doi.org/10.1007/s10021-021-00684-y
51. Palmer, Stephanie CJ, Tiit Kutser, and Peter D. Hunter. Remote sensing of inland waters: Challenges, progress and future directions. *Remote Sens. Environ.* **2015**, 157, 1–8. doi: 10.1016/j.rse.2014.09.021.
52. Ruddick, K.G.; Gons, H.J.; Rijkeboer, M.; Tilstone, G. Optical remote sensing of chlorophyll-a in case 2 waters by use of an adaptive two-band algorithm with optimal error properties. *Appl. Opt.* **2001**, 40(21), 3575–3585. doi: 10.1364/AO.40.003575.
53. Dierssen, H.M. Perspectives on empirical Approaches for ocean color remote sensing of chlorophyll in a changing climate. *Proc. Natl. Acad. Sci. USA* **2010**, 107(40), 17073–17078. doi:10.1073/pnas.0913800107.
54. Tzortziou, M.; Subramaniam, A.; Herman, J.R.; Gallegos, C.L.; Neale, P.J.; Harding Jr, L.W. Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuar. Coast Shelf Sci.* **2007**, 72(1-2), 16–32. doi: 10.1016/j.ecss.2006.09.018.
55. Le, C.; Hu, C.; English, D.; Cannizzaro, J; Kovach, C. Climate-driven chlorophyll-a changes in a turbid estuary: Observations from satellites and implications for management. *Remote Sens. Environ.* **2013**, 130, 11–24. doi: 10.1016/j.rse.2012.11.011.
56. Moradi, M.; Kabiri, K. Spatio-temporal variability of red-green chlorophyll-a index from MODIS data–Case study: Chabahar Bay, SE of Iran. *Cont. Shelf Res.* **2019**, 184, 1–9. doi: 10.1016/j.csr.2019.07.002.
57. Mishra, S.; Mishra, D.R. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sens. Environ.* **2012**, 117, 394-406. doi: 10.1016/j.rse.2011.10.016.
58. Gilerson, A.A.; Gitelson, A.A.; Zhou, J.; Gurlin, D.; Moses, W.; Ioannou, I.; Ahmed, S.A. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Opt. Express* **2010**, 18(23), 24109–24125. doi: 10.1364/OE.18.024109.
59. Gitelson, A.A.; Dall'Olmo, G.; Moses, W.; Rundquist, D.C.; Barrow, T.; Fisher, T.R.; Gurlin, D.; Holz, J. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sens. Environ.* **2008**, 112(9), 3582–3593. doi: 10.1016/j.rse.2008.04.015.
60. Teillet, P.M.; Ren, X. Spectral band difference effects on vegetation indices derived from multiple satellite sensor data. *Can. J. Remote Sens.* **2008**, 34, 159–173. doi: 10.5589/m08-025.

61. Boucher, J.; Weathers, K.C.; Norouzi, H.; Steele, B. Assessing the effectiveness of Landsat 8 chlorophyll a retrieval algorithm for regional freshwater monitoring. *Ecol. Appl.* **2018**, 28(4), 1044–1054. doi: 10.1002/eap.1708.
62. Thomalla, S.J.; Fauchereau, N.; Swart, S.; Monterio, P.M.S.: Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. *Biogeosciences* **2011**, 8, 2849–2866, doi:10.5194/bg-8-2849.
63. Neil, C.; Spyrakos, E.; Hunter, P.D.; Tyler, A.N. A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. *Remote Sens. Environ*. **2019**, 229, 159-178. doi: 10.1016/j.rse.2019.04.027.
64. Zhao, D.; Huang, J.; Li, Z.; Yu, G.; Shen, H. Dynamic monitoring and analysis of chlorophyll-a concentrations in global lakes using Sentinel-2 images in Google Earth Engine. *Sci. Total Environ.* **2024**, 912, p.169152. doi: 10.1016/j.scitotenv.2023.169152.
65. Henson, S.A.; Sarmiento, J.L.; Dunne, J.P.; Bopp, L.; Lima, I.; Doney, S.C.; John, J.; Beaulieu, C. Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences* **2010**, 7(2), 621–640. doi:10.5194/bg-7-621-2010.
66. Mirtl, M.T.; Borer, E.; Djukic, I.; Forsius, M.; Haubold, H.; Hugo, W.; Jourdan, J.; Lindenmayer, D.; McDowell, W.H.; Muraoka, H.; Orenstein, D.E.; Pauw, J.C.; Peterseil, J.; Shibata, H.; Wohner, C.; Yu, X.; Haase, P. Genesis, goals and achievements of long-term ecological research at the global scale: a critical review of ILTER and future directions, *Sci. Total Environ.* **2018**, 626, 1439–1462. doi: 10.1016/J.SCITOTENV.2017.12.001.
67. Salmaso, N.; Anneville, O.; Straile, D.; Viaroli, P. European large perialpine lakes under anthropogenic pressures and climate change: present status, research gaps and future challenges. *Hydrobiologia* **2018**, 824(1), 1–32. doi:10.1007/s10750-018-3758-x.
68. Salgado-Hernanz, P.M.; Racault, M.-F.; Font-Muñoz, J.S.; Basterretxea, G. Trends in phytoplankton phenology in the Mediterranean Sea based on ocean-colour remote sensing. *Remote Sens. Environ.* **2019**, 221, 50–64. doi: 10.1016/J.RSE.2018.10.036.