Article

Remote Sensing of Lake Water Clarity: Performance and Transferability of Both Historical Algorithms and Machine Learning

Hannah J. Rubin [1,2,*], David A. Lutz [3], Bethel G. Steele [4], Kathryn L. Cottingham [5], Kathleen C. Weathers [4], Mark J. Ducey [6], Michael Palace [7,8], Kenneth M. Johnson [9] and Jonathan W. Chipman [2]

1   Department of Civil and Environmental Engineering, University of Tennessee, Rm 411, John D. Tickle Building, Knoxville, TN 37920, USA
2   Department of Geography, Dartmouth College, 120 Fairchild Hall, Hanover, NH 03755, USA; jonathan.w.chipman@dartmouth.edu
3   Environmental Studies Program, Dartmouth College, 6182 Steele Hall, Hanover, NH 03755, USA; david.a.lutz@dartmouth.edu
4   Cary Institute of Ecosystem Studies, Millbrook, NY 12545, USA; steeleb@caryinstitute.org (B.G.S.); weathersk@caryinstitute.org (K.C.W.)
5   Department of Biological Sciences, Dartmouth College, 6044 Life Sciences Center, Hanover, NH 03755, USA; kathryn.l.cottingham@dartmouth.edu
6   Department of Natural Resources and the Environment, 114 James Hall, University of New Hampshire, James Hall, Durham, NH 03824, USA; mark.ducey@unh.edu
7   Earth Systems Research Center, University of New Hampshire, Morse Hall, Durham, NH 03824, USA; palace@guero.sr.unh.edu
8   Department of Earth Sciences, College of Engineering and Physical Sciences, University of New Hampshire, James Hall, Durham, NH 03824, USA
9   Department of Sociology and Carsey School of Public Policy, University of New Hampshire, Durham, NH 03824, USA; ken.johnson@unh.edu
*   Correspondence: hrubin@vols.utk.edu

Abstract: There has been little rigorous investigation of the transferability of existing empirical water clarity models developed at one location or time to other lakes and dates of imagery with differing conditions. Machine learning methods have not been widely adopted for analysis of lake optical properties such as water clarity, despite their successful use in many other applications of environmental remote sensing. This study compares model performance for a random forest (RF) machine learning algorithm and a simple 4-band linear model with 13 previously published empirical non-machine learning algorithms. We use Landsat surface reflectance product data aligned with spatially and temporally co-located in situ Secchi depth observations from northeastern USA lakes over a 34-year period in this analysis. To evaluate the transferability of models across space and time, we compare model fit using the complete dataset (all images and samples) to a single-date approach, in which separate models are developed for each date of Landsat imagery with more than 75 field samples. On average, the single-date models for all algorithms had lower mean absolute errors (MAE) and root mean squared errors (RMSE) than the models fit to the complete dataset. The RF model had the highest pseudo-R^2 for the single-date approach as well as the complete dataset, suggesting that an RF approach outperforms traditional linear regression-based algorithms when modeling lake water clarity using satellite imagery.

Keywords: Landsat; lakes; Secchi depth; machine learning; random forest; water clarity

1. Introduction

Water clarity is both an indicator and an influencer of lakes' function in the landscape. For example, water clarity influences lake ecosystem function [1,2] and modulates susceptibility to climate change [3]. At the same time, a change in clarity can affect human

perceptions of a lake's value [4–7] and thus the provisioning of ecosystem services to society [8]. Monitoring and observing trends in lake water clarity is therefore important for local and regional stakeholders.

Secchi depth is a reliable, easy, and affordable measurement of water clarity (or light attenuation within the water column) that is or has been used by scientists, lake managers, and community members [9–14] across wide geographic areas for decades [9,10,12–16]. A Secchi disk is a white or black-and-white disk [17] that is lowered into a lake, and the depth at which the observer can no longer see the disk and where it reappears again when it is raised is recorded [18]. Secchi depth is typically measured in the deeper areas of the lake, ideally at regular intervals throughout the ice-off season, and it is often measured in multiple locations within a lake when a lake monitoring program is present. Measurements can range from a few centimeters in very turbid lakes to more than 10 meters in deep, clear lakes [19]. Secchi depth depends on multiple factors such as suspended sediment, dissolved organic matter, and phytoplankton biomass [20–22]. In many lakes, ice-out is followed by a period of low water transparency due to turbidity from spring snow-melt and destratification conditions as well as the spring phytoplankton bloom [23]. This period is followed by the spring "clear-water phase" as zooplankton populations increase and consume phytoplankton. Once deep lakes stratify, phytoplankton populations are re-established and summer water clarity is determined by nutrient availability, weather conditions, and lake food web structure. Collections of in situ Secchi measurements over wide areas are labor-intensive and often logistically challenging, necessitating other methods to expand the temporal and spatial scope of water quality assessment.

One such method is satellite remote sensing, whose data have been used to evaluate water clarity in lakes for more than 40 years [24–32] and have been acknowledged as effective tools for monitoring local and regional trends in Secchi depth. The same optical water properties that influence attenuation of light in the water column (and thus in situ measurements of transparency) also determine spectral reflectance back to the satellite, such as turbidity due to suspended sediments, brown coloration resulting from dissolved organic compounds, and chlorophyll and other pigments used by phytoplankton to harvest light for photosynthesis [33]. The water-leaving radiance is then mediated by absorption and scattering in the atmosphere before being recorded by the sensor. As a result, the use of remotely sensed data for widespread monitoring of water quality is attractive and the focus of several federal, state, and local agencies and organizations. Previous remote sensing research on water clarity has generally focused on small sample sizes (<20 lakes) over short periods of time (<3 years) [34–39], although exceptions do exist [40–42]. The transferability of models among locations and times is critical to the operational application of remote sensing methods to lake monitoring because an ultimate goal is to expand spatial and temporal coverage and avoid relying on field data collection campaigns as the only source of water quality data. There are, however, some limitations to satellite remote sensing; for instance, poor atmospheric conditions can make images unusable, pixel size may be too large to pair with small waterbodies, or the temporal return period may be too long to effectively capture rapid changes in surface water conditions.

One tool that has been used to address algorithm transferability is machine learning, which harnesses computational power to learn and identify underlying patterns within data sets [43]. While machine learning techniques are rapidly being adopted in other environmental applications of remote sensing, they have been less frequently used for remote sensing of lake water quality [44–47]. When machine learning has been used, the papers have predominantly focused specifically on chlorophyll-A [44] or suspended sediment [48] rather than overall light attenuation. Here, we pioneer the use of random forest (RF) modeling with regression (hereafter, RF), a type of machine learning that selects class membership based on a network of classification and regression trees [49], for use with Secchi depth-specific algorithms.

This analysis aims to (1) investigate whether RF modeling can produce more accurate estimates of lake water clarity from satellite imagery than 13 previously published

regression-based algorithms (Tables 1 and 2) evaluate the transferability of both RF models and traditional regression models across a very large dataset. All of the algorithms considered here can be categorized as empirical, meaning that they begin with observations and attempt to find a model that best fits those observations. The main alternative approach, not included in this study, consists of physically-based or quasi-analytical algorithms that model the bio-optical properties of the water column based on the scattering and absorption of light [33,50–54]. Each of these approaches, empirical and physical, has its advantages. We focus here on empirical algorithms because they are in widespread use for monitoring Secchi depth, and because RF itself is an empirical method. In addition, many of the physical or quasi-analytical methods require narrower or differently placed spectral bands than those available on Landsat. To date, there has been limited investigation of the applicability of empirical models developed at one location or time to other lakes and other image dates with differing atmospheric conditions [41] or bio-optical properties; filling that gap is the focus of this study.

2. Materials and Methods

We gathered Landsat satellite imagery from 4 Landsat missions collected on 1962 dates over 34 years, along with 36,621 in situ Secchi depth observations across 397 lakes on 4745 dates for this analysis. First, we compared the performance of an RF algorithm trained on the full dataset to that of 13 previously published algorithms fit to the same data alongside a 4-band multivariate algorithm, as measured by mean absolute error (MAE), root mean squared error (RMSE), and a pseudo- coefficient of determination (pseudo- R^2, calculated as 1—(residual sum of squares (RSS)/total sum of squares (TSS))) [55]. Then, for all the algorithms, we compared the performance of models fit separately to each Landsat date to the model fit to the full dataset.

2.1. Study Area

This study focused on lakes with surface areas of at least 150 ha in the US states of Maine, New Hampshire, Vermont, and New York. Though there is a considerable body of literature on remote sensing of lake water clarity in the midwestern USA [14,24,40,56,57], this study area is primarily forested and contains a large number of clear water lakes with long-term in situ measurements [58] (Figure 1). Although water quality in this region is of tremendous environmental, recreational [59], and economic importance [4,5], few algorithms are available to estimate Secchi depth in clear water systems.
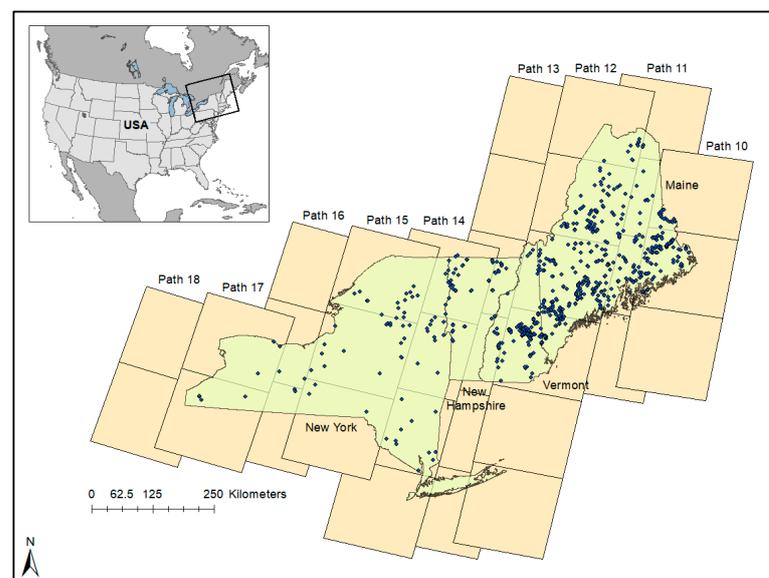

Figure 1. Landsat paths (tan rectangles) and Secchi measurement locations (blue diamonds) across the study region [60].

2.2. Field Observations of Secchi Depth

In this study, the in situ Secchi depth measurements originated from lake monitoring programs that rely on volunteers, lake organizations, and researchers to collect measurements (Supplementary Table S1). We compiled data from four organizations (Maine Department of Environmental Protection, New Hampshire Department of Environmental Services, Vermont Department of Environmental Conservation, and New York Department of Environmental Conservation) into a unified database containing 34 years of measurements across the study region [61–66]. Selecting pixels from near the central basin of a lake is the standard approach [67]. Here, Secchi measurements were limited to those that were taken more than 125 meters from shore to avoid mixed remote sensing pixels, varying lake depths due to changing lake level, and overhanging vegetation. Because many of the contributing data sources had duplicate entries or multiple entries for a lake on a given day, a ranking system was created to choose a single Secchi measurement for each lake on each day. Rank was weighted by distance from shore, where distances greater than 500 meters from the shore were favored, as well as time from remote sensing image, where Secchi measurements taken closer in time to the matched image were favored.

2.3. Data Extraction

We used Google Earth Engine [68], a cloud-based geospatial analysis tool, to extract spectral data from Landsat-4, -5, -7, and -8 surface reflectance images [69,70] (Supplementary Figure S1) for a buffer around site coordinates that have measurements between May and October from 1984–2017. Specifically, we utilized the Landsat Surface Reflectance tier-1 products from Landsat 4–5 TM, 7 ETM+, and 8 OLI [69,70] which have been atmospherically corrected and contain quality assurance bands. With the transition from ETM+ to OLI, both the radiometric resolution and atmospheric correction process were improved. The standard Landsat surface reflectance product is based on two different processors (LEDAPS for Landsat TM and ETM+, and LaSRC for Landsat OLI; [71,72]). In both cases, the algorithm was optimized for land rather than aquatic systems, but the standard LEDAPS/LaSRC products are nonetheless frequently used for lake monitoring applications. While other, potentially more effective atmospheric correction algorithms exist (e.g., ACOLITE [73] Polymer [74], and C2RCC [75]), they were not considered here due to constraints in implementation of the Earth Engine code, lack of applicability to TM/ETM+, or other factors. The results will therefore be conservative such that future work with better atmospheric correction algorithms will yield more transferable results. Supplementary Figure S2 compares Landsat-7 ETM+ and Landsat-8 OLI spectra for six randomly selected lakes with Landsat-7 (blue) and Landsat-8 (orange) images one day apart, for top of atmosphere (TOA) reflectance, ρ_λ (before atmospheric correction), and R_{rs} (after atmospheric correction, in units of sr^{-1}) to demonstrate the efficacy of the correction process.

We extracted spectral data only for those images that were captured within ±5 days of the in situ collection date, following Bohn et al. [36] and Boucher et al. [76]. We extracted pixel data within a buffer zone of 1.8 times the Landsat pixel size (30 meters) surrounding each Secchi observation location, based on comparing results from buffer zone sizes from 0.1–5 times the pixel size to the same buffer zone in a single Landsat scene downloaded from the USGS server. For Landsat-7 ETM+ SLC-off images, we did not include any pixels that fell inside a data gap [77]. The image data for each extracted buffer zone were filtered using the bit quality assessment (BQA) band to remove clouds and cloud shadows. The remaining data were filtered to remove any null values. We extracted the mean values of each buffer region surrounding a sample in bands 1–4 (blue, green, red, near-infrared) and 7 (short-wave infrared) for Landsat 4–5 TM and 7 ETM+ and bands 2–5 (blue, green, red, near-infrared) and 7 (short-wave infrared) for Landsat-8 OLI images. When pixels were removed due to quality considerations of the BQA bands, these data were not incorporated into the mean for the buffer region.

The extracted mean reflectance data for each buffer region surrounding a sample were then processed in the R Programming language [78]. To filter haze- and glare-affected

imagery, we removed all data associated with scaled reflectance values greater than 250 ($\rho_\lambda > 0.025$ or $R_{rs} > 0.00796$) in band 7, the second shortwave-infrared band between roughly 2.064 μm and 2.294 μm [79]. We also removed obvious outliers, defined as those outside 1.5 times the interquartile range beyond the first and third quartiles. There was a single image-measurement pair that was excluded from all analyses after these QA/QC steps were completed due to a very low red band value (scaled reflectance < 3; $\rho_\lambda < 0.0003$ or $R_{rs} < 9.56 \times 10^{-5}$) that caused issues with the application of historical algorithms that included the red band. To confirm that the Google Earth Engine code was handling the data extraction process correctly, spectral band values were extracted for the same sample locations in one Landsat-8 image downloaded directly from the United States Geological Survey (USGS) server and compared using ArcGIS Desktop (Supplementary Figure S3).

2.4. Implementation of Algorithms

We identified 13 unique algorithms previously published in the scientific literature (Table 1, Supplementary Table S2) for comparison to our machine learning methodology. When algorithms appeared multiple times in the literature, we cited the first (or a prominent early) appearance of the algorithm. In addition to these published algorithms, we constructed a new 4-band multivariate linear algorithm using the blue, green, red, and near-infrared bands to examine whether the contribution of all bands with no machine learning assistance would improve model fit and if the result of the machine learning model was a meaningful improvement. Our machine learning algorithm was a random forest regression model and the predictor variables were the same four bands used for previously published models and the 4-band linear model described (Table 1), allowing us to isolate the effectiveness of machine learning relative to traditional algorithm construction techniques.

An RF model uses a random sampling of training data to build independent decision trees, resulting in trees with high variance and low bias [80]. The final classification is the average of the probabilities from each tree [49,80]. RF modeling has been shown to work well with satellite imagery in prior contexts [81–84]. In practice, the RF selects a training sample with replacement from a dataset and uses repeated regression trees to decide the final class membership based on given predictor variables [80]. The samples that were not selected for training were used to evaluate model performance [49,80]. We used the randomForest package [85] within the R programming environment [78]. This package is often used in remote sensing of lake water quality across large datasets [11,86–89]. Though the default number of regression trees in R is 500, following Belgiu and Drăguţ [80], we used 128 regression trees because model performance past 128 trees requires significant computational costs for marginal improvements in accuracy [90].

We compared 13 existing algorithms, the 4-band linear models, and the RF model. In all cases, we used the same variables (e.g., individual spectral bands or band ratios) as in the originally published version, but the coefficients were re-calculated empirically for our dataset in R [78]. This recalculation was necessary because the coefficients were meant to be re-estimated for all iterations of the model in new contexts and, additionally, some of the original versions were based TOA reflectance rather than surface reflectance (SR) [69,70]. The sensor calibration coefficients and SR algorithms have both changed over time such that originally published coefficients would not be directly applicable to the current versions of image data [91]. All algorithms were linear models except Domínguez Gómez et al. [92]. For that model, we used a nonlinear least-squares method [93] to re-calculate the coefficients and exponents.

Table 1. Description of published algorithms for predicting water clarity, as reported by original sources. Ln indicates natural logarithm. R^2 is reported by the original authors on their predicted variable, including any transformations. Any operation (e.g., ln (Red)) is applied to the pixels in the buffer zone surrounding each sampling location.

Name	Source	Predicted Variable (m)	Formula	Samples	Images	R^2
Allee and Johnson	[34]	Secchi Depth	Red − mean (Red)	30	10	0.74
Baban	[94]	Secchi Depth	Blue	14	1	0.68
Chipman et al.	[95]	ln (Secchi Depth)	Blue/Red	15,615	17	0.85
Dekker and Peters 1	[26,96]	ln (Secchi Depth)	ln (Red)	15	1	0.86
Dekker and Peters 2	[26,97]	Secchi Depth	Red	15	1	0.81
Dominguez Gomez et al.	[92]	Secchi Depth	(Green)^x	16	5	0.9
Giardino et al.	[35]	Secchi Depth	Blue/Green	4	1	0.85
Kloiber et al.	[30,31,40, 98,99]	ln (Secchi Depth)	Blue/Red + Blue	374	13	0.93
Lathrop and Lillesand	[38,56]	ln (Secchi Depth)	Green	9	1	0.98
Lavery et al.	[100]	Secchi Depth	Red + Blue/Red	18–25	4	0.81
Mancino et al.	[101]	Secchi Depth	Red/Green + Blue/Green + Blue	60	1	0.82
Wu et al.	[41,102]	ln (Secchi Depth)	Blue + Red	25	5	0.83
Yip et al.	[14]	Secchi Depth	Infrared + Green + Blue	120	136	0.6

2.5. Algorithm Assessment

For each of the algorithms (all 13 published algorithms, the 4-band linear algorithm, and the RF), two types of models were constructed: one using the entire dataset ("overall model") and one using only single dates of imagery ("single-date model"). The overall model for a given algorithm was based on all in situ measurements and Landsat images across the entire study region for the 34-year time period. A total of 52 "single-date" models were developed for every date of Landsat imagery across the region that included at least 75 in situ Secchi measurements within five days of image acquisition [76]. Due to the nature of the Landsat orbital cycle, all the Landsat images from a given date in the study area lie along a single orbit path (Figure 1).

Two datasets were used to assess model performance: a "training" dataset and an out-of-bag "test" dataset. The test dataset was a stratified random selection of lakes totaling 10 percent of the total number of lakes in the complete dataset. To account for widely varying numbers of samples at individual lakes over the 34-year period, we tallied the

total number of Secchi-Landsat pairs within each lake and constructed deciles of lakes based on their sample count. We randomly chose 10 percent of lakes from each decile to create a testing dataset of 3937 samples. The goal of this test set was to isolate the influence of specific lake properties and be able to assess model performance on lakes that did not contribute to model training. In the text we refer to these two datasets as the training (n = 32,683 in the overall model, 90% of total) and testing (n = 3937 in the overall model, 10% of total) datasets. The data in these two datasets were similar, with the training dataset having a lower median Secchi depth (5.6 m; Supplementary Table S3) than the test dataset (6.5 m, Supplementary Table S3, Supplementary Figure S5). The single-date models used this same 90/10 split of data, so the total number of training data points and test data points varied by image date (Supplementary Table S4).

We assessed model performance for both the training and test datasets by calculating the root mean square error (RMSE), mean absolute error (MAE), bias, and a pseudo-coefficient of determination (pseudo-R^2). Bias is the mean of the residuals, and because the residuals are both positive and negative (and roughly equal in magnitude), it is not necessarily a valid measure of model performance without some constraint on variance, especially since most of our models are regression-based and the residuals will be zero by definition. While we have included bias, the pseudo-R^2 and RMSE values give a more complete picture of model performance. Pseudo-R^2, also known as the Nash-Sutcliffe efficiency coefficient [103], has been used in prior machine learning efforts [104,105].

In addition to these standard measures of algorithm performance, we also calculated the slope of the line describing the relationship between predicted Secchi depth and observed Secchi depth. Since RF models ostensibly do not require cross-validation from a separate test due to the method by which regression trees are bootstrapped during model construction, these error statistics have been asserted by other authors to provide a sensible test of the model [49]. In order to facilitate comparison among natural log-transformed Secchi (e.g., Chipman et al., Kloiber et al., etc.), untransformed Secchi (e.g., Allee and Johnson, Baban, etc.) and nonlinear (e.g., Dominguez-Gomez et al.) models, we report a pseudo-R^2 [55,106,107] for all algorithms using the formula (1—(residual sum of squares)/(total sum of squares)) [108].

3. Results

For the complete dataset, the overall random forest model had a lower MAE (Table 2, Figure 2) and RMSE (Table 2, Figure 3), as well as a higher pseudo-R^2 (Table 2, Figure 4) than the other algorithms. Of all the models, the random forest was the only model that, when constructed from the complete data set, explained more than 25% of the variability in observed Secchi depth (Figure 4). However, most algorithms (except Baban, Dominguez Gomez, and the untransformed Dekker and Peters) produced at least one single-date model with a pseudo-R^2 value over 0.6 (Figure 4).

In Figures 2–4, each model result is color coded based on the relationship between predicted and observed. For cases with very low slope values (shown in white) the model is essentially unresponsive such that the models are predicting consistent Secchi depths regardless of band characteristics and are therefore not useful, regardless of low MAE or RMSE values. The RF approach yielded a lower error (MAE: 0.47 m–0.86 m, RMSE: 0.66m–1.13 m) on individual scene training than nearly all other algorithms (MAE: 0.90 m–2.54 m, RMSE: 1.11 m–3.04 m).
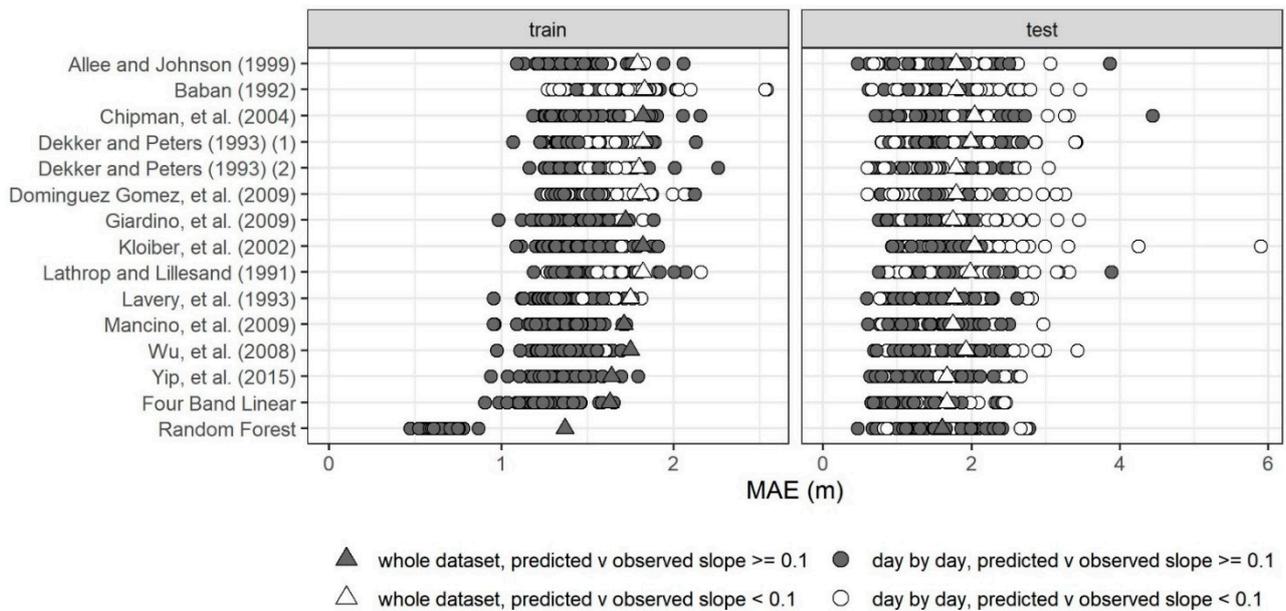
Figure 2. MAE for the training (**left**) and test (**right**) subsets, for the whole dataset (triangle) as compared to the day-by-day analyses (circles). Icons for the datasets are colored by the slope of the relationship between predicted and observed. Unfilled circles indicate a static predicted Secchi depth (predicted-observed slope < 0.1), gray fill indicates some response in the predicted Secchi depth such that there is a positive slope (>=0.1) in the relationship between predicted and observed Secchi depth. Note the difference in the x-axis scale between the two panels.
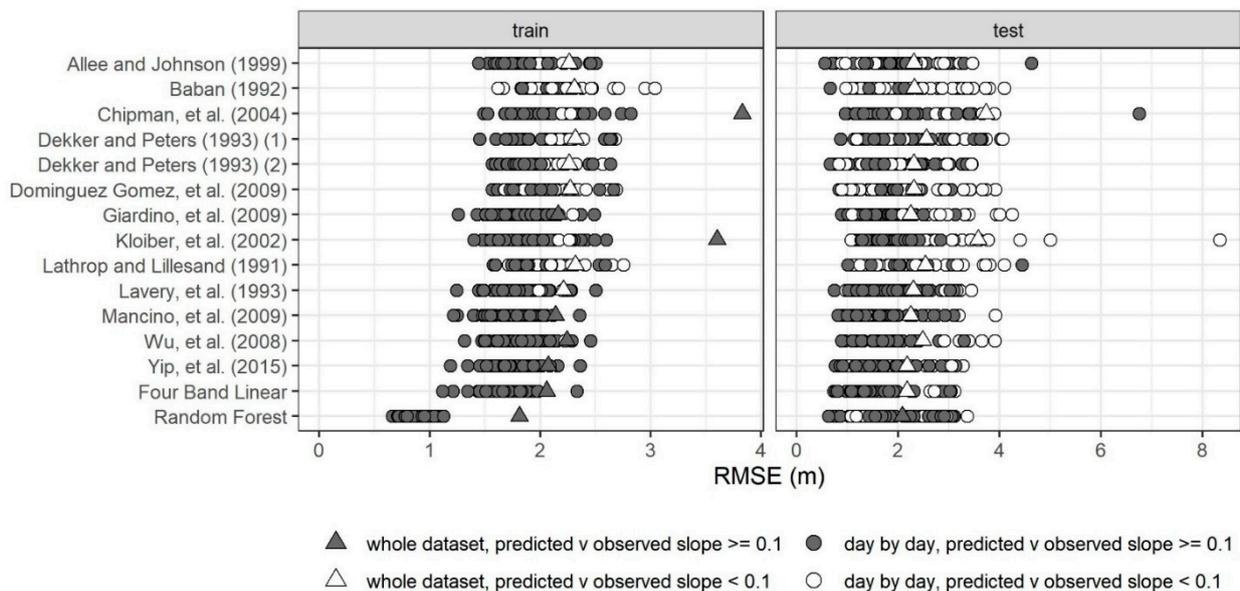

Figure 3. RMSE for the training (**left**) and test (**right**) subsets for the whole dataset (triangle) as compared to the day-by-day analyses (circles). Icons for the datasets are colored by the slope of the relationship between predicted and observed. Unfilled circles indicate a static predicted Secchi depth (predicted-observed slope < 0.1), gray fill indicates some response in the predicted Secchi depth such that there is a positive slope (>=0.1) in the relationship between predicted and observed Secchi depth. Note the difference in the x-axis scale between the two panels.

Figure 4. Pseudo-R^2 for the training (**left**) and test (**right**) subsets for the whole dataset (triangle) as compared to the day-by-day analyses (circles). Icons for the datasets are colored by the slope of the relationship between predicted and observed. Unfilled circles indicate a static predicted Secchi depth (predicted-observed slope < 0.1), gray fill indicates some response in the predicted Secchi depth such that there is a positive slope (>=0.1) in the relationship between predicted and observed Secchi depth. Note the difference in the x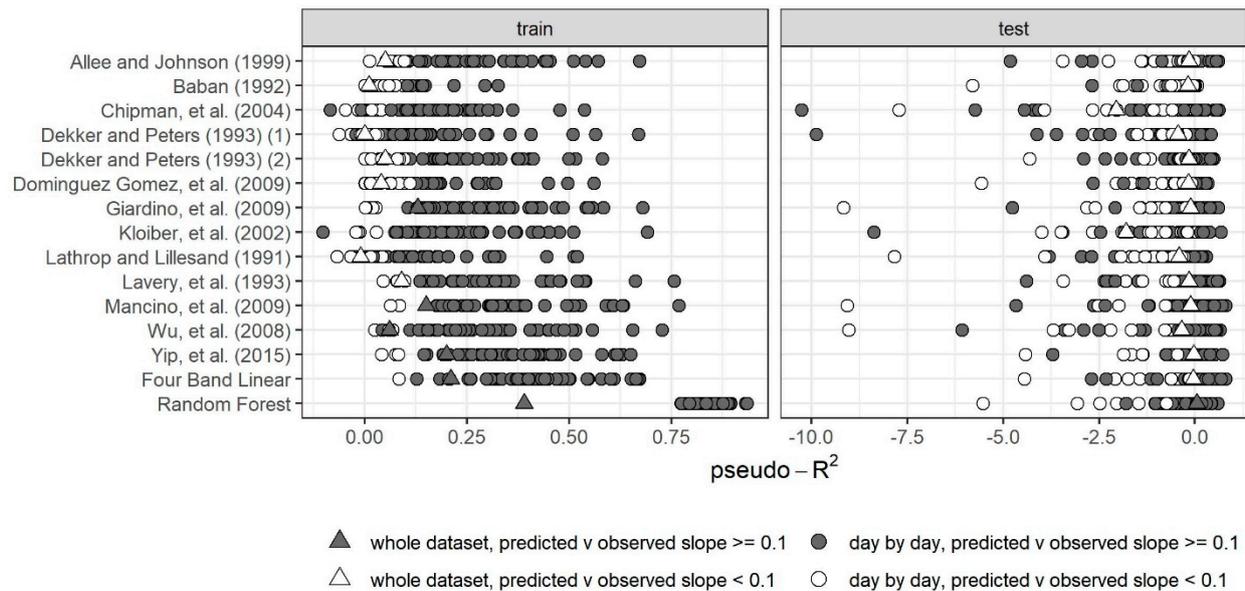-axis scale between the two panels. This figure has been truncated at −0.25 in the training set and −11 in the test set, and in doing this, 3 points were excluded: the whole dataset training result for Chipman et al. (pseudo-R^2: −1.73) and Kloiber et al. (pseudo-R^2: −1.41) and a single-date model testing result for Kloiber et al. (pseudo-R^2: −44.71). A full-scale figure is included as Supplementary Figure S6.

Table 2. Summary statistics for all algorithms created from the complete dataset of lake Secchi depths in ME, NH, VT, and NY, for both the training (n = 32,683) and testing (n = 3937) subsets. Pseudo-R^2 was calculated as 1-RSS/TSS.

	Training Data			Test Data			
Model Name	**MAE (m)**	**RMSE (m)**	**Pseudo-R^2**	**MAE (m)**	**RMSE (m)**	**Pseudo-R^2**	**Bias (m)**
Allee and Johnson	1.79	2.26	0.05	1.79	2.31	−0.16	0.85
Baban	1.83	2.31	0.01	1.80	2.32	−0.18	0.9
Chipman et al.	1.82	3.83	0.09	2.04	3.73	−0.33	1.28
Dekker and Peters	1.82	2.32	0.08	1.99	2.56	−0.33	1.35
Dekker and Peters 2	1.80	2.26	0.05	1.79	2.31	−0.16	0.85
Dominguez Gomez et al.	1.81	2.27	0.04	1.79	2.31	−0.17	0.86
Giardino et al.	1.72	2.16	0.13	1.75	2.25	−0.11	0.75
Kloiber et al.	1.82	3.60	0.10	2.04	3.58	−0.34	1.27
Lathrop and Lillesand	1.82	2.32	0.09	1.98	2.54	−0.31	1.32
Lavery et al.	1.75	2.21	0.09	1.77	2.30	−0.16	0.83
Mancino et al.	1.71	2.14	0.15	1.75	2.25	−0.11	0.74
Wu et al.	1.75	2.24	0.16	1.92	2.48	−0.27	1.25
Yip et al.	1.64	2.07	0.20	1.67	2.17	−0.03	0.65
4-Band	1.63	2.06	0.21	1.67	2.17	−0.04	0.65
Random Forest	1.37	1.81	0.39	1.60	2.08	0.05	0.61

The RF approach consistently explained more variability (had higher pseudo-R^2 values) in the training data than the historical algorithms (Figure 4). Performance for

all models was much weaker for the test data not used to fit the model, with negative pseudo-R^2 values indicating that the residual sum of squares was much higher than the total sum of squares. In other words, a simple mean of the data would have outperformed the statistical models. None of the algorithms differentiate between the deepest Secchi depths particularly well, even with the training data (Figure 5).
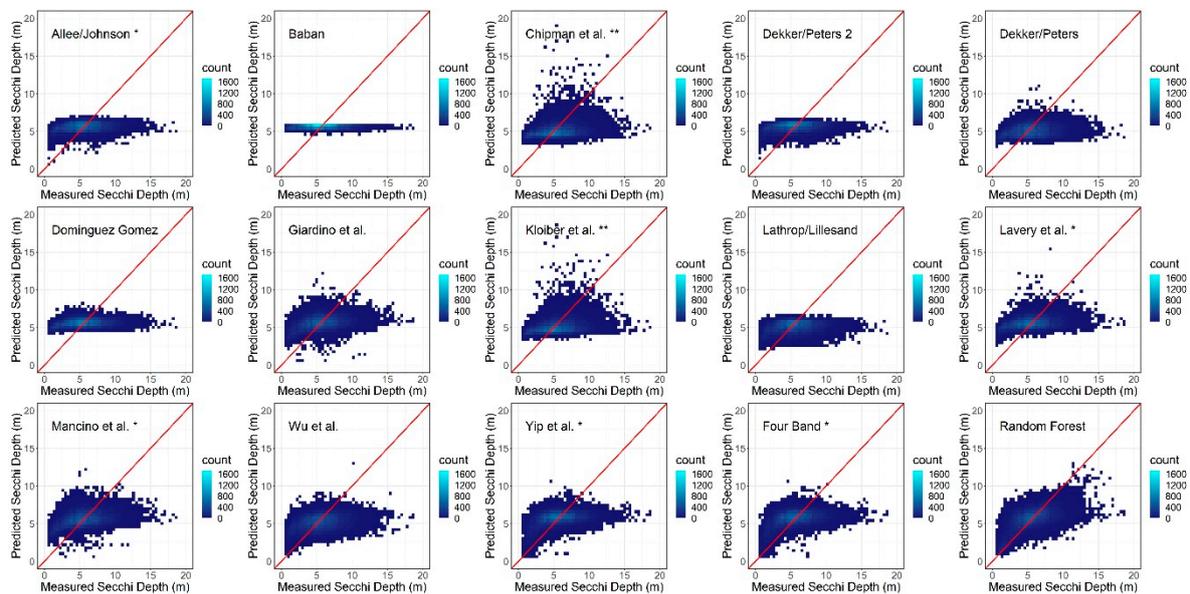

Figure 5. Training dataset model output for the overall dataset. Panels with a single asterisk ("*") after the model indicate that there are values that are not displayed because there are negatively predicted Secchi. Panels with two asterisks ("**") indicate that there are values that are not displayed because they are outside of the bounds of the limits displayed here (maximum Secchi depth displayed is 20 m).

4. Discussion

Algorithm Comparison

We found that, when compared across the complete dataset, the random forest approach predicts water clarity with lower error rates and higher pseudo-R^2 values than any of the 13 published algorithms evaluated here, even after the existing algorithms were re-fit to this dataset. Additionally, single-date models tended to outperform the overall models. The previously published algorithms, except for Yip et al., do not have meaningful model fits (where meaningful fit is defined as pseudo-$R^2 > 0.2$) in the overall dataset analysis (Table 2). Though some of the error metrics are relatively small, it is not clear how to interpret them given the poor model fits. This result suggests that these algorithms may not be widely applicable outside of the lakes and time frame for which they were developed and should be very cautiously utilized [42] for monitoring Secchi depth with remotely sensed imagery.

In terms of performance for the overall model versus single-date models, the single-date versions of all 15 algorithms generally provided more precise estimates of water clarity and higher pseudo-R^2 values, but only when applied within the same dates. This approach is similar to the approach taken by many early studies [31,56,94,95,97] of using a single date of imagery from one Landsat path to develop a model, and this finding suggests there are limits to the transferability of these algorithms across time. The most 'fair' comparison between the RF models fit here and the results reported from prior publications would be based on the single-date models with the highest pseudo-R^2 values. However, even the best single-date models in this analysis tend to perform less well than in prior studies elsewhere. This is likely due to the constraint of which dates we considered, which was driven by the availability of at least 75 in situ samples (to avoid overfitting by the RF algorithm), rather than image quality, as was the criterion in the past. Relaxing this minimum sample size

for the regression algorithms would allow more single-date models to be included, some of which may have pseudo-R^2 values falling in the range of those reported from previous studies ($R^2 > 0.8$).

While single-date models can perform well under optimal conditions, these models cannot necessarily be transferred to other locations and times because they are greatly influenced by lake-specific factors and the atmospheric conditions of that day [109], such as haze. Insofar as much of the problem is due to imperfect atmospheric correction algorithms and the low radiometric resolution of Landsat TM/ETM+, we can expect improvement in the future as better surface reflectance products are developed and the sensors' radiometric resolution improves. As shown in the right-hand column of Supplementary Figure S2b,d,f,h,j,l, the Landsat-7 atmospheric correction algorithm (LEDAPS) very closely matches the shape of the corresponding Landsat-8 spectrum. In some instances, such as the elevated near-infrared reflectance for Wilson Pond in Landsat-7 (Supplementary Figure S2b), differences in R_{rs} between the two sensors are also found in the TOA spectra (Supplementary Figure S2a), suggesting that changes occurred during the intervening day between the Landsat-7 and Landsat-8 images. More concerning are the spectra for Junior Lake (e,f) and China Lake (k,l), where the TOA spectra for Landsat-7 and -8 match quite closely in TOA reflectance but show slightly elevated values of R_{rs} for Landsat-7 in the short-wavelength bands, suggesting poorer performance by the LEDAPS algorithm in these cases. While this is just a small (but random) sample from the thousands in this study, it is reassuring to see the relative consistency between LEDAPS and LaSRC in most cases, but cautionary to see exceptions where the LEDAPS output may not be as reliable as that of LaSRC.

Given that both the radiometric resolution and atmospheric correction process changed from Landsat-4/5/7 to Landsat-8, we examined the RF model predictions for the full data set with and without Landsat-8. There was no large, consistent difference in the RF model predictions for the two cases (Supplementary Figure S7). While the radiometric properties and surface reflectance processor for Landsat-8 are superior to its predecessors, the combination of Landsat-8 with Landsat-4/5/7 in the RF analysis does not substantially change the outcome. Ultimately, it would be beneficial to not need expensive field data collection campaigns for every date of imagery, so some transferable overall model would be helpful. Even when field data are available, a transferable model is invaluable for variance reduction and anomaly detection [110].

The RF approach has a meaningful model fit and yields lower error across the overall model (Figures 2 and 3). However, the predicted values still exhibit error, especially for Secchi depths > 5 m—i.e., clearer-water lakes (Supplementary Figure S8). This indicates that for our dataset, a widely applicable RF model created from the same band data as the historical algorithms does not appear to be attainable. Even the best-performing algorithms (including the RF) do not differentiate Secchi depths of >5 m vs. >10 m well, especially in the complete dataset (Figure 5, Supplementary Figure S8). Though the RF's performance on the independent test data is very poor, there is no way to know if this is due to overfitting the training data or to small sample sizes (n < 200). Based on the difference in median Secchi depth between the training and testing datasets (5.5 m and 6.4 m, respectively, Supplementary Table S4, Supplementary Figure S5), it may be possible that the slight skew towards greater Secchi depths in the testing dataset greatly affects the performance in all test cases. In the day-by-day models, the number of samples included in the test set range from 3 to 20 (Supplementary Table S5) and summarizing fit metrics from that few samples could also contribute to the prevalence of negative pseudo-R^2 values [111].

The best performing algorithms tended to be those that included the near-infrared band, such as Yip et al., the 4-band multivariate algorithm, and the RF algorithm. These algorithms tended to predict Secchi depth more accurately in the overall model when the water clarity was relatively poor. A possible rationale for this performance is that the inclusion of a near-infrared band allows for greater ability to predict Secchi depths across a wide range of water optical properties and turbid conditions [112]. Given that

the RF algorithm produced nearly equal Gini-based importance values [113] for the four bands (Supplementary Table S5), it is possible that all are necessary to make progress in creating better algorithms to predict Secchi depth, although further examination is needed to test this hypothesis. Some of the historical algorithms tended to produce stronger fits for the single-date method later in the season, when lakes are more likely to have lower in situ Secchi depths (less clear water) and when the atmosphere is more stable [41]. The lack of sensitivity at the higher end of the Secchi scale is widely known from prior work [96,114,115], and the reason the RF, Yip et al., and 4-band algorithms perform better than others is that they are better able to differentiate among lakes with lower levels of water clarity.

The machine learning model does better estimate Secchi depth than the traditional algorithm development approach, when just comparing the 4-band linear model with the RF. Though these two algorithms use the same band data, the RF approach yields a lower MAE (1.37 m) and RMSE (1.81 m) as well as higher pseudo-R^2 (0.39) than the linear model approach (MAE: 1.63 m, RMSE: 2.06 m, pseudo-R^2: 0.21) in the training dataset.

Although our results indicate that published algorithms did not accurately predict Secchi depth for this dataset as a whole, this does not altogether invalidate those prior publications' results; the 13 algorithms work extremely well for the study areas and time periods for which they were originally developed (Table 1), especially when applied to clear-sky imagery and validated using data from the same location and time period. The high R^2 values of some previous studies may also have resulted from being developed with relatively turbid, productive lakes [20,24,30,56]. Instead, we believe the reduced applicability to a completely new study area (in most cases) and imagery with varying atmospheric conditions emphasizes the importance of lake-specific factors and atmospheric correction, which at the present time only imperfectly compensates for variations in scattering and absorption by the atmosphere.

Future work should include an investigation of the timing of precipitation events relative to the difference between measurement and image. We used a 5-day timeframe, but it is entirely possible that a precipitation event could occur within that window and introduce noise into the analysis. To do a full analysis, we would need consistent precipitation data for every location for every measurement date over the 34 years. Instead, we plot the residuals from the RF algorithm versus the time difference (either positive or negative) from the Secchi observation and demonstrate that there is no structure in the residuals for this dataset (Supplementary Figure S4).

Findings from this research also engender questions related to underlying distributions of lake bio-optical water properties and how these differences may require appropriate consideration in remote sensing-based algorithms. Optical water types are now frequently utilized when examining remotely sensed data from lakes [109,116,117], and these categories may provide additional information for machine learning approaches. Previously published empirical algorithms based on regression models with single or multiple spectral bands do not appear to adequately characterize water clarity in our large dataset across a range of lake types and Secchi depths. Various methods for optimizing the selection of band ratios or spectral features, such as optimal band ratio analysis (OBRA, [118]) and related methods [119] or a transformed feature space approach [120], could help maximize the effectiveness of these empirical spectral models. Additionally, our work focuses predominantly on Landsat-specific algorithm testing and does not address methods using other satellite platforms (e.g., MODIS, Sentinel-2); additional information captured in data containing a wider range of bands may yield more accurate models and should be explored. Biophysical or quasi-analytical modeling approaches, such as those that model and reconstruct the scattering of light within the water column, may also provide insight into atmospheric controls on algorithm performance [50,118–124], although we did not consider these methods in the research described here.

Altogether, we view our findings as a progressive step forward in identifying optimal methods for monitoring water quality changes over time using the Landsat archive.

5. Conclusions

Satellite remote sensing provides an attractive method for monitoring changes in lake water quality at regional scales. The analyses reported here use in situ measurements from 397 lakes to investigate the potential for random forest modeling with regression to improve estimates of lake water clarity from satellite imagery and evaluate the transferability of both RF models and traditional regression-based models across space and time.

Previously published historical algorithms were mostly developed using small sample sizes (<20 lakes) over short time periods (<3 years) and were therefore limited by the temporal and spatial specificity of the in situ data used to test these algorithms as well as the number of in situ samples (Table 1). These algorithms perform well on the images and timeframe for which they were developed but generally cannot be applied across the entire Landsat image archive as they often produce non-meaningful fits. While the RF approach and many of the historical algorithms yielded meaningful model fits for individual Landsat scenes, the RF approach resulted in models that had lower overall error than the historical algorithms. An RF approach seems promising in a regional, multi-Landsat-mission analysis, but it is still difficult to predict lake clarity with greater Secchi depths. We find that Secchi depth can be generally predicted with the RF algorithm, but there is not enough precision for explicit lake-specific depth predictions. For example, one could estimate whether a lake has low (Secchi depth < 3 m) or high (Secchi depth > 5 m) transparency reasonably well from the RF algorithm, whereas with previously published algorithms, the error is generally too large for such distinction.

The single-date models have better fits than the overall model for the non-machine learning algorithms; however, the improvement from machine learning is clouded by the possibility of overfitting. The RF also does very well in the overall model comparison. This suggests that the approach taken by most prior researchers of relying on single dates of imagery is effective, as demonstrated by their reported high R^2 values. However, with the increased emphasis on automating and operationalizing the process of extracting environmental information from remote sensing imagery, methods that are the most transferable over space and time will be best equipped to deal with spatially and temporally extensive data.

Our work provides an important contribution to the use of long-term satellite monitoring of water quality in freshwater lakes. Advancements in cloud processing and gathering of data from the USGS [91] provide the aquatic monitoring and remote sensing community the opportunity to investigate changes across space and time in new ways. However, as our results indicate, methodological designs and algorithms that were constructed during a time of limited data availability are not transferrable across very wide regions, particularly when lakes may vary considerably with respect to their bio-optical properties. Overall, this work furthers the research community's ability to remotely monitor water clarity as a complement to expensive and time-consuming in situ measurements and provides a framework for evaluating broad spatiotemporal trends in water clarity across the Northeast.

Supplementary Materials: All code is available from GitHub at https://github.com/steeleb/Rubin_etal_repository, Figure S1: Distribution of Landsat imagery used in this study by satellite and decade, Figure S2: Comparison of TOA reflectance (ρ_λ; a,c,e,g,i,k) and R_{rs} (b,d,f,h,j,l) for six randomly selected lakes with Landsat-7 (blue) and Landsat-8 (orange) images one day apart. Chart titles include lake name, latitude/longitude coordinates, and dates of the two images. Note that Landsat-8 includes an extra short-wavelength band, Figure S3: Landsat-8 image (ultrablue band). Green points are pixels properly included in the dataset when filtering with the BQA band and blue points are pixels properly excluded, Figure S4: The residuals from the RF algorithm versus the difference in days from the Secchi measurement. There is no clear pattern in the residuals, Figure S5: Histograms of the (top) test and (bottom) training set showing similar Scheme 6.4 m) was slightly higher than that of the "training" set (5.5 m), Figure S6: Full-range of pseudo-R^2 for 15 tested algorithms for predicting Secchi depth from Landsat imagery, Figure S7: The difference between predictions from the overall models with Landsat-8 and the overall model predictions without Landsat-8, Figure S8: Testing dataset model output for the overall dataset. Panels with a single asterisk ("*") after the

model indicate that there are values that are not displayed because there are negative predicted Secchi. Panels with two asterisks ("**") indicate that there are values that are not displayed because they are outside of the bounds of the limits displayed here (maximum Secchi depth displayed is 20 m), Table S1: The sources for a four-state (Maine, New Hampshire, Vermont, and New York) in-lake Secchi database consisting of six data providers, Table S2: Further description of published algorithms for predicting Secchi depth from Landsat imagery, as reported by original sources, Table S3: Summary statistics for the overall model training and testing data, in meters, Table S4: Summary statistics for the single-date model training and testing data, in meters, Table S5: This table reports the Gini-based importance values [113] for the four variables used in the random forest algorithm. Since Gini importance values are relative to one another, this indicates that the four bands used in this algorithm are all fairly balanced in importance in the building of the algorithm.

References

1. Doherty, E.; Murphy, G.; Hynes, S.; Buckley, C. Valuing Ecosystem Services across Water Bodies: Results from a Discrete Choice Experiment. *Ecosyst. Serv.* **2014**, *7*, 89–97. [CrossRef]
2. Mueller, H.; Hamilton, D.P.; Doole, G.J. Evaluating Services and Damage Costs of Degradation of a Major Lake Ecosystem. *Ecosyst. Serv.* **2016**, *22*, 370–380. [CrossRef]
3. Rose, K.C.; Winslow, L.A.; Read, J.S.; Hansen, G.J.A. Climate-Induced Warming of Lakes Can Be Either Amplified or Suppressed by Trends in Water Clarity. *Limnol. Oceanogr. Lett.* **2016**, *1*, 44–53. [CrossRef]
4. Boyle, K.M.J.; Poor, P.J.; Taylor, L.O. Estimating the Demand for Protecting Freshwater Lakes from Eutrophication. *Am. J. Agric. Econ.* **1999**, *81*, 1118–1122. [CrossRef]
5. Gibbs, J.P.; Halstead, J.M.; Boyle, K.M.J.; Huang, J.-C. An Hedonic Analysis of the Effects of Lake Water Clarity on New Hampshire Lakefront Properties. *Agric. Resour. Econ. Rev.* **2002**, *31*, 39. [CrossRef]
6. Poor, P.J.; Pessagno, K.L.; Paul, R.W. Exploring the Hedonic Value of Ambient Water Quality: A Local Watershed-Based Study. *Ecol. Econ.* **2007**, *60*, 797–806. [CrossRef]
7. Walsh, P.J.; Milon, J.W.; Scrogin, D.O. The Spatial Extent of Water Quality Benefits in Urban Housing Markets. *Land Econ.* **2011**, *87*, 628–644. [CrossRef]
8. Millennium Ecosystem Assessment. Available online: https://www.millenniumassessment.org/en/index.html (accessed on 15 January 2021).
9. Bruhn, L.C.; Soranno, P.A. Long Term (1974–2001) Volunteer Monitoring of Water Clarity Trends in Michigan Lakes and Their Relation to Ecoregion and Land Use/Cover. *Lake Reserv. Manag.* **2005**, *21*, 10–23.
10. Gunn, J.M.; Snucins, E.; Yan, N.D.; Arts, M.T. Use of Water Clarity to Monitor the Effects of Climate Change and Other Stressors on Oligotrophic Lakes. *Environ. Monit. Assess.* **2001**, *67*, 69–88. [CrossRef]
11. Read, E.K.; Patil, V.P.; Oliver, S.K.; Hetherington, A.L.; Brentrup, J.A.; Zwart, J.A.; Winters, K.M.; Corman, J.R.; Nodine, E.R.; Woolway, R.I.; et al. The Importance of Lake-Specific Characteristics for Water Quality across the Continental United States. *Ecol. Appl.* **2015**, *25*, 943–955. [CrossRef] [PubMed]
12. Soranno, P.A.; Bacon, L.C.; Beauchene, M.; Bednar, K.E.; Bissell, E.G.; Boudreau, C.K.; Boyer, M.G.; Bremigan, M.T.; Carpenter, S.R.; Carr, J.W.; et al. LAGOS-NE: A Multi-Scaled Geospatial and Temporal Database of Lake Ecological Context and Water Quality for Thousands of US Lakes. *GigaScience* **2017**, *6*. [CrossRef] [PubMed]
13. Stephens, D.L.B.; Carlson, R.E.; Horsburgh, C.A.; Hoyer, M.V.; Bachmann, R.W.; Canfield, D.E., Jr. Regional Distribution of Secchi Disk Transparency in Waters of the United States. *Lake Reserv. Manag.* **2015**, *31*, 55–63. [CrossRef]

14. Yip, H.D.; Johansson, J.; Hudson, J.J. A 29-Year Assessment of the Water Clarity and Chlorophyll-a Concentration of a Large Reservoir: Investigating Spatial and Temporal Changes Using Landsat Imagery. *J. Great Lakes Res.* **2015**, *41*, 34–44. [CrossRef]

15. Read, E.K.; Carr, L.; Cicco, L.D.; Dugan, H.A.; Hanson, P.C.; Hart, J.A.; Kreft, J.; Read, J.S.; Winslow, L.A. Water Quality Data for National-Scale Aquatic Research: The Water Quality Portal. *Water Resour. Res.* **2017**, *53*, 1735–1745. [CrossRef]

16. Lottig, N.R.; Wagner, T.; Norton Henry, E.; Spence Cheruvelil, K.; Webster, K.E.; Downing, J.A.; Stow, C.A.; Heffernan, J.; Soranno, P.; Angilletta, M.; et al. Long-Term Citizen-Collected Data Reveal Geographical Patterns and Temporal Trends in Lake Water Clarity. *PLoS ONE* **2014**, *9*, e95769. [CrossRef]

17. Preisendorfer Secchi Disk Science: Visual Optics of Natural Waters. *Limnol. Oceanogr.* **1986**, *31*, 909–926. [CrossRef]

18. Salvato, L.; Coordinator, S.D.-I.P. The 2015 Secchi Dip-in Report. 22. Available online: https://z0ku333mvy924cayk1kta4r1 -wpengine.netdna-ssl.com/wp-content/uploads/2015/02/Final-2015-Secchi-Dip-In-Report.pdf (accessed on 15 September 2020).

19. Williamson, C.E.; Neale, P.J. Ultraviolet Light. In *Encyclopedia of Inland Waters*; Likens, G.E., Ed.; Academic Press: Oxford, UK, 2009; pp. 705–714. ISBN 978-0-12-370626-3.

20. Brezonik, P.; Menken, K.D.; Bauer, M. Landsat-Based Remote Sensing of Lake Water Quality Characteristics, Including Chlorophyll and Colored Dissolved Organic Matter (CDOM). *Lake Reserv. Manag.* **2005**, *21*, 373–382. [CrossRef]

21. Lambou, V.W.; Taylor, W.D.; Hern, S.C.; Williams, L.R. Comparisons of Trophic State Measurements. *Water Res.* **1983**, *17*, 1619–1626. [CrossRef]

22. Lind, O.T. The Effect of Non-Algal Turbidity on the Relationship of Secchi Depth to Chlorophyll *a*. *Hydrobiologia* **1986**, *140*, 27–35. [CrossRef]

23. Sommer, U.; Adrian, R.; De Senerpont Domis, L.; Elser, J.J.; Gaedke, U.; Ibelings, B.; Jeppesen, E.; Lürling, M.; Molinero, J.C.; Mooij, W.M.; et al. Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. *Annu. Rev. Ecol. Evol. Syst.* **2012**, *43*, 429–448. [CrossRef]

24. Lathrop, R.G. Landsat Thematic Mapper Monitoring of Turbid Inland Water Quality. *Photogramm. Eng.* **1992**, *58*, 465–470.

25. Brown, D.; Skaggs, R.; Warwick, R. Reconnaissance Analysis of Lake Condition In east-Central Minnesota. Available online: https://conservancy.umn.edu/bitstream/handle/11299/205799/L1036.pdf?sequence=1 (accessed on 15 September 2020).

26. Dekker, A.G.; Peters, S.W.M. The Use of the Thematic Mapper for the Analysis of Eutrophic Lakes: A Case Study in the Netherlands. *Int. J. Remote Sens.* **1993**, *14*, 799–821. [CrossRef]

27. Lillesand, T.M. Use of Landsat Data to Predict the Trophic State of Minnesota Lakes. *Photogramm. Eng.* **1983**, *49*, 219–229.

28. Ritchie, J.C.; Cooper, C.M.; Schiebe, F.R. The Relationship of MSS and TM Digital Data with Suspended Sediments, Chlorophyll, and Temperature in Moon Lake, Mississippi. *Remote. Sens. Environ.* **1990**, *33*, 137–148. [CrossRef]

29. Cox, R.M.; Forsythe, R.D.; Vaughan, G.E.; Olmsted, L.L. Assessing Water Quality in Catawba River Reservoirs Using Landsat Thematic Mapper Satellite Data. *Lake Reserv. Manag.* **1998**, *14*, 405–416. [CrossRef]

30. Kloiber, S.M.; Brezonik, P.L.; Bauer, M.E. Application of Landsat Imagery to Regional-Scale Assessments of Lake Clarity. *Water Res.* **2002**, *36*, 4330–4340. [CrossRef]

31. Kloiber, S.M.; Brezonik, P.L.; Olmanson, L.G.; Bauer, M.E. A Procedure for Regional Lake Water Clarity Assessment Using Landsat Multispectral Data. *Remote Sens. Environ.* **2002**, *82*, 38–47. [CrossRef]

32. Lathrop, R.G.; Lillesand, T.M.; Yandell, B. Testing the Utility of Simple Multi-Date Thematic Mapper Calibration Algorithms for Monitoring Turbid Inland Waters. *Remote Sens.* **1991**, *12*, 2045–2063. [CrossRef]

33. Bukata, R.P.; Jerome, J.H.; Kondratyev, K.Y.; Pozdnyakov, D.V. *Optical Properties and Remote Sensing of Inland and Coastal Waters*, 1st ed.; CRC Press: Boca Raton, FL, USA, 1995; ISBN 978-0-203-74495-6.

34. Allee, R.J.; Johnson, J.E. Use of Satellite Imagery to Estimate Surface Chlorophyll a and Secchi Disc Depth of Bull Shoals Reservoir, Arkansas, USA. *Int. J. Remote Sens.* **1999**, *20*, 1057–1072. [CrossRef]

35. Giardino, C.; Pepe, M.; Brivio, P.A.; Ghezzi, P.; Zilioli, E. Detecting Chlorophyll, Secchi Disk Depth and Surface Temperature in a Sub-Alpine Lake Using Landsat Imagery. *Sci. Total Environ.* **2001**, *268*, 19–29. [CrossRef]

36. Bohn, V.Y.; Carmona, F.; Rivas, R.; Lagomarsino, L.; Diovisalvi, N.; Zagarese, H.E. Development of an Empirical Model for Chlorophyll-a and Secchi Disk Depth Estimation for a Pampean Shallow Lake (Argentina). *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 183–191. [CrossRef]

37. Ritchie, J.C.; Schiebe, F.R. Monitoring Suspended Sediments with Remote Sensing Techniques. *Int. Assoc. Hydrol. Sci. Hydrol. Appl. Space Technol.* **1986**, *160*, 233–243.

38. Harrington, J.A.; Schiebe, F.R.; Nix, J.F. Remote Sensing of Lake Chicot, Arkansas: Monitoring Suspended Sediments, Turbidity, and Secchi Depth with Landsat MSS Data. *Remote Sens. Environ.* **1992**, *39*, 15–27. [CrossRef]

39. Schiebe, F.R.; Ritchie, J.C. Suspended Sediment Monitored by Satellite. In Proceedings of the Fourth Federal Interagency Sedimentation Conference, Las Vegas, NV, USA, 24–27 March 1986.

40. Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-Year Landsat Water Clarity Census of Minnesota's 10,000 Lakes. *Remote Sens. Environ.* **2008**, *112*, 4086–4097. [CrossRef]

41. McCullough, I.M.; Loftin, C.S.; Sader, S.A. Combining Lake and Watershed Characteristics with Landsat TM Data for Remote Estimation of Regional Lake Clarity. *Remote Sens. Environ.* **2012**, *123*, 109–115. [CrossRef]

42. Deutsch, E.S.; Cardille, J.A.; Koll-Egyed, T.; Fortin, M.-J. Landsat 8 Lake Water Clarity Empirical Algorithms: Large-Scale Calibration and Validation Using Government and Citizen Science Data from across Canada. *Remote Sens.* **2021**, *13*, 1257. [CrossRef]

43. Michie, D.; Spiegelhalter, D.J.; Taylor, C. Others Machine Learning. *Neural Stat. Classif.* **1994**, *13*, 1–298.

44. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless Retrievals of Chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in Inland and Coastal Waters: A Machine-Learning Approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [CrossRef]

45. Pahlevan, N.; Sarkar, S.; Franz, B.A.; Balasubramanian, S.V.; He, J. Sentinel-2 MultiSpectral Instrument (MSI) Data Processing for Aquatic Science Applications: Demonstrations and Validations. *Remote Sens. Environ.* **2017**, *201*, 47–56. [CrossRef]

46. Najah Ahmed, A.; Binti Othman, F.; Abdulmohsin Afan, H.; Khaleel Ibrahim, R.; Ming Fai, C.; Shabbir Hossain, M.; Ehteram, M.; Elshafie, A. Machine Learning Methods for Better Water Quality Prediction. *J. Hydrol.* **2019**, *578*, 124084. [CrossRef]

47. Huo, S.; He, Z.; Su, J.; Xi, B.; Zhu, C. Using Artificial Neural Network Models for Eutrophication Prediction. *Procedia Environ. Sci.* **2013**, *18*, 310–316. [CrossRef]

48. Kuhn, C.; de Matos Valerio, A.; Ward, N.; Loken, L.; Sawakuchi, H.O.; Kampel, M.; Richey, J.; Stadler, P.; Crawford, J.; Striegl, R.; et al. 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. [CrossRef]

49. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

50. Gege, P. WASI-2D: A Software Tool for Regionally Optimized Analysis of Imaging Spectrometer Data from Deep and Shallow Waters. *Comput. Geosci.* **2014**, *62*, 208–215. [CrossRef]

51. Dekker, A.G.; Malthus, T.J.; Seyhan, E. Quantitative Modeling of Inland Water Quality for High-Resolution MSS Systems. *IEEE Trans. Geosci. Remote Sens.* **1991**, *29*, 89–95. [CrossRef]

52. Lee, Z.; Carder, K.L.; Arnone, R.A. Deriving Inherent Optical Properties from Water Color: A Multiband Quasi-Analytical Algorithm for Optically Deep Waters. *Appl. Opt.* **2002**, *41*, 5755–5772. [CrossRef] [PubMed]

53. Lee, Z.; Shang, S.; Qi, L.; Yan, J.; Lin, G. A Semi-Analytical Scheme to Estimate Secchi-Disk Depth from Landsat-8 Measurements. *Remote Sens. Environ.* **2016**, *177*, 101–106. [CrossRef]

54. Rodrigues, T.; Alcântara, E.; Watanabe, F.; Imai, N. Retrieval of Secchi Disk Depth from a Reservoir Using a Semi-Analytical Scheme. *Remote Sens. Environ.* **2017**, *198*, 213–228. [CrossRef]

55. Kvålseth, T.O. Cautionary Note about R 2. *Am. Stat.* **1985**, *39*, 279–285. [CrossRef]

56. Lathrop, R.G.; Lillesand, T.M. Use of Thematic Mapper Data to Assess Water Quality in Green Bay and Central Lake Michigan. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 671–680.

57. Peckham, S.D.; Lillesand, T.M. Detection of Spatial and Temporal Trends in Wisconsin Lake Water Clarity Using Landsat-Derived Estimates of Secchi Depth. *Lake Reserv. Manag.* **2006**, *22*, 331–341. [CrossRef]

58. Nelson, S.; Soranno, P.; Cheruvelil, K.S.; Batzli, S.A. Regional Assessment of Lake Water Clarity Using Satellite Remote Sensing. *J. Limnol.* **2003**, *62*, 27–32. [CrossRef]

59. Needelman, M.S.; Kealy, M.J. Recreational Swimming Benefits Of New Hampshire Lake Water Quality Policies: An Application of a Repeated Discrete Choice Model. *Agric. Resour. Econ. Rev.* **1995**, *24*, 1–10. [CrossRef]

60. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the Conterminous United States—Representing a Decade of Land Cover Change Information. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 345–354. [CrossRef]

61. New Hampshire Department of Environmental Services Volunteer Lake Assessment Program. Available online: https://www.des.nh.gov/organization/divisions/water/wmb/vlap/ (accessed on 15 November 2019).

62. National Water Quality Monitoring Council. Available online: https://www.waterqualitydata.us/ (accessed on 24 September 2020).

63. Maine Department of Environmental Protection. Available online: https://www.maine.gov/dep/ (accessed on 24 September 2020).

64. Lake Champlain Basin Program. Available online: https://www.lcbp.org/water-environment/ (accessed on 24 September 2020).

65. NYS Deptartment of Environmental Conservation. Available online: https://www.dec.ny.gov/25.html (accessed on 24 September 2020).

66. Vermont Department of Environmental Conservation. Available online: https://dec.vermont.gov/ (accessed on 24 September 2020).

67. Ross, M.R.V.; Topp, S.N.; Appling, A.P.; Yang, X.; Kuhn, C.; Butman, D.; Simard, M.; Pavelsky, T.M. AquaSat: A Data Set to Enable Remote Sensing of Water Quality for Inland Waters. *Water Resour. Res.* **2019**, *55*, 10012–10025. [CrossRef]

68. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18. [CrossRef]

69. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary Analysis of the Performance of the Landsat 8/OLI Land Surface Reflectance Product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [CrossRef]

70. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.-K. A Landsat Surface Reflectance Dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [CrossRef]

71. US Geological Survey. *Landsat 4-7 Collection 1 (C1) Surface Reflectance (LEDAPS) Product Guide*; EROS Data Center: Sioux Falls, SD, USA, 2020; p. 39.

72. US Geological Survey. *Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) Product Guide*; EROS Data Center: Sioux Falls, SD, USA, 2020; p. 38.

73. Vanhellemont, Q.; Ruddick, K. Turbid Wakes Associated with Offshore Wind Turbines Observed with Landsat 8. *Remote Sens. Environ.* **2014**, *145*, 105–115. [CrossRef]

74. Steinmetz, F.; Deschamps, P.-Y.; Ramon, D. Atmospheric Correction in Presence of Sun Glint: Application to MERIS. *Opt. Express* **2011**, *19*, 9783–9800. [CrossRef]

75. Doerffer, R.; Schiller, H. The MERIS Case 2 Water Algorithm. *Int. J. Remote Sens.* **2007**, *28*, 517–535. [CrossRef]

76. Boucher, J.; Weathers, K.C.; Norouzi, H.; Steele, B. Assessing the Effectiveness of Landsat 8 Chlorophyll a Retrieval Algorithms for Regional Freshwater Monitoring. *Ecol. Appl.* **2018**, *28*, 1044–1054. [CrossRef]

77. Preliminary Assessment of the Value of Landsat 7 ETM+ Data Following Scan Line Corrector Malfunction 2003. Available online: https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/SLC_off_Scientific_Usability.pdf (accessed on 12 March 2021).

78. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

79. Zhang, Y.; Pulliainen, J.T.; Koponen, S.S.; Hallikainen, M.T. Water Quality Retrievals from Combined Landsat TM Data and ERS-2 SAR Data in the Gulf of Finland. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 622–629. [CrossRef]

80. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

81. Belgiu, M.; Drăguţ, L. Comparing Supervised and Unsupervised Multiresolution Segmentation Approaches for Extracting Buildings from Very High Resolution Imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 67–75. [CrossRef] [PubMed]

82. Frazier, R.J.; Coops, N.C.; Wulder, M.A.; Kennedy, R. Characterization of Aboveground Biomass in an Unmanaged Boreal Forest Using Landsat Temporal Segmentation Metrics. *ISPRS J. Photogramm. Remote Sens.* **2014**. [CrossRef]

83. Karlson, M.; Ostwald, M.; Reese, H.; Sanou, J.; Tankoano, B.; Mattsson, E. Mapping Tree Canopy Cover and Aboveground Biomass in Sudano-Sahelian Woodlands Using Landsat 8 and Random Forest. *Remote Sens.* **2015**, *7*, 10017–10041. [CrossRef]

84. Tsutsumida, N.; Comber, A.J. Measures of Spatio-Temporal Accuracy for Time Series Land Cover Data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**. [CrossRef]

85. Liaw, A.; Wiener, M. Classification and Regression with Random Forest. *R News* **2002**, *2*, 18–22.

86. Hafeez, S.; Wong, M.; Ho, H.; Nazeer, M.; Nichol, J.; Abbas, S.; Tang, D.; Lee, K.; Pun, L. Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong. *Remote Sens.* **2019**, *11*, 617. [CrossRef]

87. Peterson, K.T.; Sagan, V.; Sidike, P.; Hasenmueller, E.A.; Sloan, J.J.; Knouft, J.H. Machine Learning-Based Ensemble Prediction of Water-Quality Variables Using Feature-Level and Decision-Level Fusion with Proximal Remote Sensing. *Photogramm. Eng. Remote Sens.* **2019**, *85*, 269–280. [CrossRef]

88. Wang, C.; Jia, M.; Chen, N.; Wang, W. Long-Term Surface Water Dynamics Analysis Based on Landsat Imagery and the Google Earth Engine Platform: A Case Study in the Middle Yangtze River Basin. *Remote Sens.* **2018**, *10*, 1635. [CrossRef]

89. Lin, S.; Novitski, L.N.; Qi, J.; Stevenson, R.J. Landsat TM/ETM+ and Machine-Learning Algorithms for Limnological Studies and Algal Bloom Management of Inland Lakes. *J. Appl. Remote Sens.* **2018**, *12*, 1. [CrossRef]

90. Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How Many Trees in a Random Forest? In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7376, pp. 154–168.

91. Chander, G.; Markham, B.L.; Helder, D.L. Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors. *Remote Sens. Environ.* **2009**, *113*, 893–903. [CrossRef]

92. Domínguez Gómez, J.A.; Chuvieco Salinero, E.; Sastre Merlín, A. Monitoring Transparency in Inland Water Bodies Using Multispectral Images. *Int. J. Remote Sens.* **2009**, *30*, 1567–1586. [CrossRef]

93. Bates, D.M.; Watts, D.G. *Nonlinear Regression Analysis and Its Applications*; Wiley: New York, NY, USA, 1988; Volume 2.

94. Baban, S.M.J. Detecting Water Quality Parameters in the Norfolk Broads, U.K., Using Landsat Imagery. *Int. J. Remote Sens.* **1993**, *14*, 1247–1267. [CrossRef]

95. Chipman, J.W.; Lillesand, T.M.; Schmaltz, J.E.; Leale, J.E.; Nordheim, M.J. Mapping Lake Water Clarity with Landsat Images in Wisconsin, U.S.A. *Can. J. Remote Sens.* **2004**, *30*, 1–7. [CrossRef]

96. Odermatt, D.; Gitelson, A.; Brando, V.E.; Schaepman, M. Review of Constituent Retrieval in Optically Deep and Complex Waters from Satellite Imagery. *Remote Sens. Environ.* **2012**, 116–126. [CrossRef]

97. Duane Nellis, M.; Harrington, J.A.; Wu, J. Remote Sensing of Temporal and Spatial Variations in Pool Size, Suspended Sediment, Turbidity, and Secchi Depth in Tuttle Creek Reservoir, Kansas: 1993. *Geomorphology* **1998**, *21*, 281–293. [CrossRef]

98. Olmanson, L.G.; Brezonik, P.L.; Bauer, M.E. Evaluation of Medium to Low Resolution Satellite Imagery for Regional Lake Water Quality Assessments. *Water Resour. Res.* **2011**, *47*. [CrossRef]

99. Hellweger, F.L.; Schlosser, P.; Lall, U.; Weissel, J.K. Use of Satellite Imagery for Water Quality Studies in New York Harbor. *Estuar. Coast. Shelf Sci.* **2004**, *61*, 437–448. [CrossRef]

100. Lavery, P.; Pattiaratchi, C.; Wyllie, A.; Hick, P. Water Quality Monitoring in Estuarine Waters Using the Landsat Thematic Mapper. *Remote Sens. Environ.* **1993**, *46*, 268–280. [CrossRef]

101. Mancino, G.; Nolè, A.; Urbano, V.; Amato, M.; Ferrara, A. Assessing Water Quality by Remote Sensing in Small Lakes: The Case Study of Monticchio Lakes in Southern Italy. *IForest Biogeosci. For.* **2009**, *2*, 154. [CrossRef]

102. Wu, M.; Zhang, W.; Wang, X.; Luo, D. Application of MODIS Satellite Data in Monitoring Water Quality Parameters of Chaohu Lake in China. *Environ. Monit. Assess.* **2009**, *148*, 255–264. [CrossRef] [PubMed]

103. Nash, J.E.; Sutcliffe, J.V. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]

104. Palace, M.; Herrick, C.; DelGreco, J.; Finnell, D.; Garnello, A.J.; McCalley, C.; McArthur, K.; Sullivan, F.; Varner, R.K. Determining Subarctic Peatland Vegetation Using an Unmanned Aerial System (UAS). *Remote Sens.* **2018**, *10*, 1498. [CrossRef]

105. Paliwal, M.; Kumar, U.A. Neural Networks and Statistical Techniques: A Review of Applications. *Expert Syst. Appl.* **2009**, *36*, 2–17. [CrossRef]

106. Cragg, J.G.; Uhler, R.S. The Demand for Automobiles. *Can. J. Econ.* **1970**, *3*, 386–406. [CrossRef]

107. Nagelkerke, N. A Note on a General Definition of the Coefficient of Determination. *Biometrika* **1991**, *78*, 691–692. [CrossRef]

108. Nakagawa, S.; Schielzeth, H. A General and Simple Method for Obtaining R2 from Generalized Linear Mixed-Effects Models. *Methods Ecol. Evol.* **2013**, *4*, 133–142. [CrossRef]

109. Moore, T.S.; Dowell, M.D.; Bradt, S.; Verdu, A.R. An Optical Water Type Framework for Selecting and Blending Retrievals from Bio-Optical Algorithms in Lakes and Coastal Waters. *Remote Sens. Environ.* **2014**, *143*, 97–111. [CrossRef]

110. Thompson, S.K. *Sampling*, 3rd ed.; John Wiley & Sons: New York, NY, USA, 2012.

111. Cornell, J.A. Factors That Influence the Value of the Coefficient of Determination in Simple Linear and Nonlinear Regression Models. *Phytopathology* **1987**, *77*, 63. [CrossRef]

112. 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*, 6855–6899. [CrossRef]

113. Buchan, I. Calculating the Gini Coefficient of Inequality. *Northwest. Inst. BioHealth Inform.* **2002**. Available online: https://www.nibhi.org.uk/Training/Forms/AllItems.aspx (accessed on 6 April 2021).

114. 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. [CrossRef]

115. Gons, H.J.; Auer, M.; Effler, S.W. MERIS Satellite Chlorophyll Mapping of Oligotrophic and Eutrophic Waters in the Laurentian Great Lakes. *Remote Sens. Environ.* **2008**, 4098–4106. [CrossRef]

116. Soomets, T.; Uudeberg, K.; Jakovels, D.; Zagars, M.; Reinart, A.; Brauns, A.; Kutser, T. Comparison of Lake Optical Water Types Derived from Sentinel-2 and Sentinel-3. *Remote Sens.* **2019**, *11*, 2883. [CrossRef]

117. Soomets, T.; Uudeberg, K.; Jakovels, D.; Brauns, A.; Zagars, M.; Kutser, T. Validation and Comparison of Water Quality Products in Baltic Lakes Using Sentinel-2 MSI and Sentinel-3 OLCI Data. *Sensors* **2020**, *20*, 742. [CrossRef]

118. Legleiter, C.J.; Roberts, D.A.; Lawrence, R.L. Spectrally Based Remote Sensing of River Bathymetry. *Earth Surf. Process. Landf.* **2009**, *34*, 1039–1059. [CrossRef]

119. Niroumand-Jadidi, M.; Vitti, A.; Lyzenga, D.R. Multiple Optimal Depth Predictors Analysis (MODPA) for River Bathymetry: Findings from Spectroradiometry, Simulations, and Satellite Imagery. *Remote Sens. Environ.* **2018**, *218*, 132–147. [CrossRef]

120. Niroumand-Jadidi, M.; Bovolo, F.; Bruzzone, L. Novel Spectra-Derived Features for Empirical Retrieval of Water Quality Parameters: Demonstrations for OLI, MSI, and OLCI Sensors. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 10285–10300. [CrossRef]

121. Sagan, V.; Peterson, K.T.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, B.A.; Maalouf, S.; Adams, C. Monitoring Inland Water Quality Using Remote Sensing: Potential and Limitations of Spectral Indices, Bio-Optical Simulations, Machine Learning, and Cloud Computing. *Earth Sci. Rev.* **2020**, *205*, 103187. [CrossRef]

122. Schowengerdt, R.A. *Techniques for Image Processing and Classifications in Remote Sensing*; Academic Press: Cambridge, MA, USA, 2012; ISBN 978-0-323-13855-0.

123. 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. [CrossRef]

124. Chavez, P.S. An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data. *Remote Sens. Environ.* **1988**, *24*, 459–479. [CrossRef]